Pathogen Origins and Evolution in the New World:

A Molecular and Bioarchaeological Approach to Tuberculosis and Leishmaniasis

by

Kelly M. Harkins

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2014 by the
Graduate Supervisory Committee:

Jane Buikstra, Co-Chair
Anne Stone, Co-Chair
Kelly Knudson
Johannes Krause
Sudhir Kumar

ARIZONA STATE UNIVERSITY

August 2014

ABSTRACT

Studies of ancient pathogens are moving beyond simple confirmatory analysis of diseased bone; researchers are now asking nuanced questions and utilizing novel methods capable of confronting the debates surrounding pathogen origins and evolution, and the relationships between humans and disease in the past. This dissertation examines two ancient human diseases through molecular and bioarchaeological lines of evidence, relying on techniques in paleogenetics and phylogenetics to detect, isolate, sequence and analyze ancient and modern pathogen DNA within an evolutionary framework. Specifically this research addresses outstanding issues regarding a) the evolution, origin and phylogenetic placement of the pathogen causing skeletal tuberculosis in New World prior to European contact, and b) the phylogeny and origins of the parasite causing the human leishmaniasis disease complex. A fourth manuscript (Chapter 5) presents a review of the major technological and theoretical advances in ancient pathogen genomics to frame the contributions of this work within a rapidly developing field. This overview emphasizes that understanding the evolution of human disease is critical to contextualizing relationships between humans and pathogens, and the epidemiological shifts observed both in the past and in the present era of (re)emerging infectious diseases. These questions continue to be at the forefront of not only pathogen research, but also bioarchaeological and paleopathological scholarship.

DEDICATION

To my parents, partner, labmates, friends and climbing buddies, without whom the last seven years would have been a. impossible, b. perilous, c. way less fun, or d. all of the above.

ACKNOWLEDGMENTS

This dissertation would not have been possible without a number of people and institutions providing training, funding, access to samples, and mentorship. I am grateful for the funding sources that enabled my graduate education and research: the Jacob K. Javits Graduate Fellowship, the National Science Foundation (NSF) Graduate Research Program Fellowship and the NSF Doctoral Dissertation Improvement Grant, as well as the Philanthropic Educational Organization Scholar Award, from the local Tempe chapter.

I would like to thank: Dr. McMahon-Pratt at Yale School of Public Health for hosting me in her laboratory while giving me access not only to *Leishmania* parasites, but also lab materials; Dr. Lucile Floeter-Winter at the University of Sao Paolo for additional DNA samples; Dr. Rachel Schwartz for her dedication to developing a bioinformatics pipeline suitable for my dataset, as well as her endless patience accommodating my research questions; Drs. Johannes Krause and Kirsten Bos and the members of Paleogenetics Lab at the Institute for Archaeological Science at the University of Tuebingen for all of their formal and informal traning and guidance throughout the past 2 years, including Alissa Mittnik who was always willing to pass on her expertise; Dr. Christina Torres-Rouff and the RP Gustavo Le Paige Archaeological Museum for hosting me and facilitating sampling of their collections; Drs. Sonia Guillen and Jane Buikstra and the Centro Mallqui for their permission, help and hard work acquiring skeletal samples; the women of the Stone Molecular Anthropology Lab for their perpetual assistance, support and comradery, in and outside of the classroom and lab; and my committee members for sharing their expertise with me, especially Dr. Sudhir Kumar for continually offering eye-opening insight. A special thanks to my co-chairs, Drs. Jane Buikstra and Anne Stone, first for giving me the opportunity five years ago to be a part of a project investigating ancient tuberculosis, and for all of their support and encouragement since, both personally and academically.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

For more than three decades, bioarchaeological and molecular approaches have offered independent evidence to address questions about health, disease, and adaptation in the recent and deep past. Rapidly advancing “next generation” sequencing (NGS) technologies adapted for paleogenetics (e.g. Burbano et al., 2010; Hodges et al., 2009; Kircher et al., 2011; Maricic et al., 2010; Meyer and Kircher, 2010) have enhanced the integration of these two approaches, reshaping our understanding of pathogen origins and evolution, as well as the historical and cultural processes that are central to contextualizing disease in past human groups. Across three manuscripts (Chapters 2-4), this dissertation examines two ancient human diseases through molecular and bioarchaeological lines of evidence, relying on techniques in paleogenetics and phylogenetics to detect, isolate, sequence and analyze ancient and modern pathogen DNA within an evolutionary framework. Specifically this research addresses outstanding issues regarding a) the evolution, origin and phylogenetic placement of the pathogen causing skeletal tuberculosis in New World prior to European contact, and b) the phylogeny and origins of the parasite causing the human leishmaniasis disease complex. A fourth manuscript (Chapter 5) presents a review of the major technological and theoretical advances in ancient pathogen genomics to frame the contributions of this work within a rapidly developing field. This overview emphasizes that understanding the evolution of human disease is critical to contextualizing relationships between humans and pathogens, and the epidemiological shifts observed both in the past and in the present era of (re)emerging infectious diseases (Armelagos et al., 2005; Barrett et al., 1998; Roberts and Manchester, 2005).

Outstanding issues in ancient tuberculosis

The New World prior and subsequent to the Age of Exploration provides a unique context to explore the relationship between humans and infectious disease, as both autochthonous pathogens and as a function of anthropogenic introduction. Diseases present in the New World

before European contact are inferred from the paleopathological record, namely skeletal changes consistent with current clinical manifestations. Research on skeletal tuberculosis from bioarchaeological contexts has a long tradition of scholarship in anthropology, likely owing to its conspicuous skeletal signature. Twentieth century physicians did not accept the extra-pulmonary spinal lesions as evidence of TB (Pott's disease) in the ancient New World (Cockburn, 1963; Hrdlička, 1909a; Morse, 1961) and thus proposed that TB was only present in the Old World prior to the Age of Exploration. The high morbidity and mortality from TB in Native American populations following European contact also bolstered the post-contact Old World origin scenario, with New World groups considered "virgin soil" for the organism. Because Native American groups reacted as if they were epidemiologically "naíve" to the pathogen, for decades it was believed that *M. tuberculosis* could not have existed in the ancient New World (Cockburn, 1963; Hrdlička, 1909b; Morse, 1961; Stead et al., 1995; Stead, 1997).

The modern phylogeography of the *Mycobacterium tuberculosis* complex (MTBC) demonstrates that human strains of *M. tuberculosis* are most diverse in Africa, while current strains in the Americas most closely resemble European strains (Gagneux and Small, 2007), leading to the hypotheses that any strains present in the New World before European exploration were either replaced or did not exist (Cockburn, 1963; Pepperell et al., 2013; Stone et al., 2009). Although decades of paleopathological evidence for pre-Columbian tuberculosis are now well accepted (see Roberts and Buikstra 2003), molecular investigations using ancient pre-Columbian material, which have identified short conserved regions of mobile elements considered to be diagnostic for tuberculosis, offer no information about phylogenetic placement, and are thus difficult to authenticate as ancient (e.g. Arriaza et al., 1995; Klaus et al., 2010; Salo et al., 1994; Shapiro and Gilbert, 2006; Spigelman and Lemma, 1993). The question thus remains: what precisely *is* precontact New World tuberculosis?

A main objective of this dissertation is to discern the exact causative agent of tuberculosis-like lesions in New World skeletal remains, thus addressing one of the most enduring debates in

pathogen evolution: whether *M. tuberculosis* in the New World was the result of a zoonotic jump within the last few thousand years or an heirloom pathogen, emerging in Africa and dispersing across the globe with human migration. Genomic reconstructions of ancient pathogens produce robust evidence of DNA authenticity and can be integrated into the well-resolved modern phylogenomic reconstructions of the MTBC, providing crucial calibration points for dating estimates (e.g. Bos et al., 2011; Shapiro et al., 2011). Therefore, the specific goals of Chapters 3 and 4 are to:

1. Determine the presence and evaluate the preservation of MTBC DNA in a global sample of archaeological human remains with indicators of skeletal tuberculosis;
2. Determine the genome sequence of the causative pathogen in New World samples and test the hypothesis that the phylogenetic placement of the ancient New World TB genomes falls within the MTBC and will be most closely related to strains in Asia.
3. Use radiocarbon dates as tip calibration in Bayesian dating analyses to calculate substitution rates and estimate divergence dates of MTBC lineages, evaluating the suitability of the model whereby TB attained its current distribution following human dispersals out of Africa during the Pleistocene.

Establishing an approach to studying ancient leishmaniasis

Leishmaniasis is an emerging human disease that results in high morbidity and mortality in contemporary societies worldwide (WHO 2014). Like tuberculosis, skeletal, ethnohistoric, iconographic, and genetic evidence indicate that the disease has persisted in human groups for millennia (Comas et al., 2013; Costa Junqueira et al., 2009; Gade, 1979; Morse, 1961; Roberts and Buikstra, 2003; Urteaga-Ballon, 1991) The process of identifying paleopathological cases, detecting the causative pathogen DNA, capturing the genome by targeting the expected diversity, and integrating the sequence data into the overall evolutionary history is demonstrated as a successful approach to tuberculosis, as for a select few other pathogens, e.g. leprosy, the Black

Death (Chan et al., 2013; Chen and Holmes, 2006; Schuenemann et al., 2011; Schuenemann et al., 2013; Wagner et al., 2014). But what about applying this process to a neglected, understudied pathogen like *Leishmania*?

Unlike research on TB, leishmaniasis does not benefit from a century of bioarchaeological investigation, an abundance of paleopathological evidence and decades of ancient DNA studies, nor are the extant *Leishmania* parasites characterized by a well-understood modern phylogeny. This lack of foundation is due to a number of intrinsic and extrinsic factors that influence the focus of current scholarship, the availability of modern sequence data, the difficulty of genomic analyses (i.e. complex eukaryotic genome 10x the size of TB), and the ability to recognize a mainly soft-tissue disease in human remains. Without a strong foundation, the application of the above methodological procedure that integrates the bioarchaeological record with molecular data cannot be the first objective. Instead, it must follow an investigation of the unresolved evolutionary history and phylogeny of the genus *Leishmania* with the intention of contributing to a more comprehensive understanding of the evolution and origins of *Leishmania* and its relationship with humans throughout history.

Outstanding issues in the evolutionary history of Leishmania

Over thirty *Leishmania* species infect humans as well as reptilian and mammalian reservoirs via a sandfly vector, but given recent leishmaniasis epidemics and the global increase in disease incidence, molecular research naturally is biased towards human-adapted strains (Schönian et al., 2013). Humans display a range of three main clinical manifestations: visceral, cutaneous, and mucocutaneous. The severity and susceptibility to *Leishmania* infection relies not only on the infecting species, but also on complex interactions between host-pathogen genetics and immune response (Lipoldova and Demant, 2006). There remains no predictable association, however, between the species genotypes and their role in disease development (Lukes et al., 2007), nor is there a strict correlation with geographic location. Consequently, molecular-based research has

long focused on identifying loci that enable rapid discrimination among human pathogenic species.

Such molecular data, in the absence of ancient DNA or fossil data, permit reconstructions of evolutionary relationships by applying theories of molecular evolution, coalescent models, and statistics to phylogenetic methods (Marjoram and Tavare, 2006; Nei and Kumar, 2000). However, a focus on the human disease at the expense of reptile- or mammalian-restricted taxa has limited comprehensive examinations on the genus level. Even when attempting to sample broadly, relatively few genetic loci have been used to create *Leishmania* phylogenies to date (e.g. Asato et al., 2009; Croan et al., 1997; Fraga et al., 2010; Zelazny, 2005). These gene trees may not accurately reflect the true history of the genus (Castresana, 2007). Three origins hypotheses have emerged to explain the current biogeographical patterns of the parasite (Kerr, 2000; Lukes et al., 2007; Momen and Cupolillo, 2000; Noyes et al., 2000), but phylogenetic analyses have been unsuccessful in reaching a consensus. This is in part due to the incongruence of molecular and non-molecular evidence (Kerr, 2000; Lysenko, 1971), or the lack of sequence data for unclassified or newly described *Leishmania* species, despite the fact that these species are likely informative for resolving branching order and interspecific relationships (e.g. Dougall, 2011; Reuss et al. 2012).

With the unresolved evolutionary history of Leishmania in mind, the goals of the current *Leishmania* component are to:

1. Obtain whole genome data via NGS from pathogenic and non-pathogenic *Leishmania* isolates for which few molecular data are available;
2. Build phylogenomic reconstructions to support/reject proposed origins hypotheses.

Although whole genome data from modern organisms should theoretically be able to resolve every part of the tree, from deep to shallow divergences (Delsuc et al., 2005; Gee, 2003), they

pose their own problems for phylogenetic inference, especially for rapid speciation events deep in the tree (Kumar et al., 2012; Maddison and Knowles, 2006; Nielsen, 2005). First are the challenges in acquiring a dataset, which vary depending on the size of the genome and whether there are references available. The difficulties of obtaining phylogenetically informative markers are exacerbated in whole genome data, especially among widely divergence taxa that each evolve uniquely, as shown in *Leishmania* (Roure and Philippe, 2011). Large datasets can negatively affect both the accurate construction of phylogenies and the statistical methods used to evaluate their support. This is observed when an increase of data leads to an increase of noise (Philippe et al., 2011), and perhaps more alarmingly, the use of large datasets arbitrarily decreases variance, producing high statistical support without increasing accuracy (Kumar et al., 2012). The methodological challenges of extracting phylogenetically informative loci within large genome datasets in particular are confronted in Chapter 4.

Broadly, this leishmaniasis study represents the beginning of a long path to acquire a similar level of resolution as found in present studies of modern tuberculosis; additional lines of evidence from the bioarchaeological record, i.e. pre-Columbian skeletal remains identified with leishmaniasis, are under current investigation and discussed further in the conclusion. Although the dissertation must approach the evolution of tuberculosis and leishmaniasis from different stages within their research trajectories, the inferences drawn from the results run parallel.

Contents of the dissertation

The dissertation consists of four manuscripts that address the goals discussed above.

Paper 1:

Title: Screening for ancient tuberculosis with qPCR: Challenges and Opportunities

Author list: Kelly M. Harkins, Jane E. Buikstra, Tessa C. Campbell, Kirsten I. Bos, Eric D. Johnson, Johannes Krause, Anne C. Stone

Publication status: Philosophical Transactions of the Royal Society B [accepted]

Paper 2: Pre-Columbian Mycobacterial Genomes Reveal Seals as a Source of Ancient Tuberculosis

Author List: Kirsten I. Bos, Kelly M. Harkins, Alexander Herbig, Mireia Coscolla, Nico Weber, Iñaki Comas, Stephen A. Forrest, Josephine M. Bryant, Simon R. Harris, Verena J. Schuenemann, Tessa J. Campbell, Kerrtu Majander, Alicia K. Wilbur, Ricardo A. Guichon[1], Dawnie L. Wolfe Steadman, Della Collins Cook, Stefan Niemann, Marcel A. Behr, Martin Zumarraga, Ricardo Bastida, Daniel Huson, Kay Nieselt, Douglas Young, Julian Parkhill, Jane E. Buikstra, Sebastien Gagneux, Anne C. Stone, and Johannes Krause.

Publication status: Nature [accepted]

Paper 3:

Title: Phylogenomic Reconstruction Supports Supercontinent Origin of *Leishmania*

Publication list: Kelly M. Harkins, Rachel S. Schwartz, Reed Cartwright, Anne C. Stone, other authors pending;

Publication status: manuscript

Paper 4:

Title: Ancient pathogen genomics: insights into timing and adaptation

Author list: Kelly M. Harkins, Anne C. Stone

Publication status: Journal of Human Evolution, Special Issue: Ancient DNA and Human Evolution [in review]

References

Armelagos, G.J., Brown, P.J., Turner, B., 2005. Evolutionary, historical and political economic perspectives on health and disease. Soc Sci Med 61, 755-765.

Arriaza, B.T., Salo, W., Aufderheide, A.C., Holcomb, T.A., 1995. Pre-Columbian tuberculosis in northern Chile: molecular and skeletal evidence. American Journal of Physical Anthropology 98, 37-45.

Asato, Y., Oshiro, M., Myint, C.K., Yamamoto, Y.-i., Kato, H., Marco, J.D., Mimori, T., Gomez, E.a.L., Hashiguchi, Y., Uezato, H., 2009. Phylogenic analysis of the genus Leishmania by cytochrome b gene sequencing. Experimental parasitology 121, 352-361.

Barrett, R., Kuzawa, C.W., McDade, T., Armelagos, G.J., 1998. Emerging and re-emerging infectious diseases: the third epidemiologic transition. Annual Review of Anthropology 27, 247-271.

Bos, K.I., Schuenemann, V.J., Golding, G.B., Burbano, H.A., Waglechner, N., Coombes, B.K., McPhee, J.B., Dewitte, S.N., Meyer, M., Schmedes, S., Wood, J., Earn, D.J., Herring, D.A., Bauer, P., Poinar, H.N., Krause, J., 2011. A draft genome of Yersinia pestis from victims of the Black Death. Nature 478, 506-510.

Burbano, H.a., Hodges, E., Green, R.E., Briggs, a.W., Krause, J., Meyer, M., Good, J.M., Maricic, T., Johnson, P.L.F., Xuan, Z., Rooks, M., Bhattacharjee, a., Brizuela, L., Albert, F.W., de la Rasilla, M., Fortea, J., Rosas, a., Lachmann, M., Hannon, G.J., Paabo, S., 2010. Targeted Investigation of the Neandertal Genome by Array-Based Sequence Capture. Science 328, 723-725.

Castresana, J., 2007. Topological variation in single-gene phylogenetic trees. Genome Biol 8, 216.

Chan, J.Z., Sergeant, M.J., Lee, O.Y., Minnikin, D.E., Besra, G.S., Pap, I., Spigelman, M., Donoghue, H.D., Pallen, M.J., 2013. Metagenomic analysis of tuberculosis in a mummy. New England Journal of Medicine 369, 289-290.

Chen, R., Holmes, E.C., 2006. Avian influenza virus exhibits rapid evolutionary dynamics. Molecular biology and evolution 23, 2336-2341.

Cockburn, A., 1963. The evolution and eradication of Infectious diseases. Cambridge Univ Press, Baltimore.

Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K.E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., Bothamley, G., Mei, J., Wei, L., Bentley, S., Harris, S.R., Niemann, S., Diel, R., Aseffa, A., Gao, Q., Young, D., Gagneux, S., 2013. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nature genetics 45, 1176-1182.

Costa Junqueira, M.A., Matheson, C., Iachetta, L., Llagostera, A., Appenzeller, O., 2009. Ancient Leishmaniasis in a highland desert of Northern Chile. PLoS ONE 4, e6983.

Croan, D.G., Morrison, D.a., Ellis, J.T., 1997. Evolution of the genus Leishmania revealed by comparison of DNA and RNA polymerase gene sequences. Molecular and biochemical parasitology 89, 149-159.

Delsuc, F., Brinkmann, H., Philippe, H., 2005. Phylogenomics and the reconstruction of the tree of life. Nat Rev Genet 6, 361-375.

Fraga, J., Montalvo, A.M., De Doncker, S., Dujardin, J.-C., Van der Auwera, G., 2010. Phylogeny of Leishmania species based on the heat-shock protein 70 gene. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases 10, 238-245.

Gade, D.W., 1979. Inca and colonial settlement, coca cultivation and endemic disease in the tropical forest. Journal of Historical Geography 5, 263-279.

Gagneux, S., Small, P.M., 2007. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. Lancet Infect Dis 7, 328-337.

Gee, H., 2003. Evolution: ending incongruence. Nature 425, 782.

Hodges, E., Rooks, M., Xuan, Z., Bhattacharjee, A., Benjamin Gordon, D., Brizuela, L., Richard McCombie, W., Hannon, G.J., 2009. Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. Nat Protoc 4, 960-974.

Hrdlička, A., 1909. Tuberculosis among certain Indian tribes of the United States: Bureau of American ethnology. Smithsonian Institute, Washington, DC.

Kerr, S.F., 2000. Palaearctic origin of Leishmania. Memórias do Instituto Oswaldo Cruz 95, 75-80.

Kircher, M., Sawyer, S., Meyer, M., 2011. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Research 40, e3.

Klaus, H.D., Wilber, A.K., Temple, D.H., Buikstra, J.E., Stone, A.C., Fernandez, M., Wester, C., Tam, M.E., 2010. Tuberculosis on the north coast of Peru: skeletal and molecular paleopathology of late pre-Hispanic and postcontact mycobacterial disease. Journal of Archaeological Science 37, 2587-2597.

Kumar, S., Filipski, A.J., Battistuzzi, F.U., Kosakovsky Pond, S.L., Tamura, K., 2012. Statistics and truth in phylogenomics. Mol Biol Evol 29, 457-472.

Lipoldova, M., Demant, P., 2006. Genetic susceptibility to infectious disease: lessons from mouse models of leishmaniasis. Nat Rev Genet 7, 294-305.

Lukes, J., Mauricio, I.L., Schönian, G., Dujardin, J.-C., Soteriadou, K., Dedet, J.-P., Kuhls, K., Tintaya, K.W.Q., Jirků, M., Chocholová, E., Haralambous, C., Pratlong, F., Oborniik, M., Horák, A., Ayala, F.J., Miles, M.A., 2007. Evolutionary and geographical history of the Leishmania donovani complex with a revision of current taxonomy. Proceedings of the National Academy of Sciences USA 104, 9375-9380.

Lysenko, A., 1971. Distribution of leishmaniasis in the Old World. Bulletin of the World Health Organization 44, 515-520.

Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. Syst Biol 55, 21-30.

Maricic, T., Whitten, M., Paabo, S., 2010. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. PLoS ONE 5, e14004.

Marjoram, P., Tavare, S., 2006. Modern computational approaches for analysing molecular genetic variation data. Nat Rev Genet 7, 759-770.

Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc 2010, pdb prot5448.

Momen, H., Cupolillo, E., 2000. Speculations on the origin and evolution of the genus Leishmania. Memórias do Instituto Oswaldo Cruz 95, 583-588.

Morse, D., 1961. Prehistoric tuberculosis in America. Am Rev Respir Dis 83, 489.

Nei, M., Kumar, S., 2000. Molecular evolution and phylogenetics. Oxford University Press, Oxford; New York.

Nielsen, R., 2005. Statistical methods in molecular evolution. Springer, New York.

Noyes, H.A., Morrison, D.A., Chance, M.L., Ellis, J.T., 2000. Evidence for a neotropical origin of Leishmania. Memórias do Instituto Oswaldo Cruz 95, 575-578.

Pepperell, C.S., Casto, A.M., Kitchen, A., Granka, J.M., Cornejo, O.E., Holmes, E.C., Birren, B., Galagan, J., Feldman, M.W., 2013. The role of selection in shaping diversity of natural M. tuberculosis populations. PLoS Pathogens 9, e1003543.

Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T., Manuel, M., Worheide, G., Baurain, D., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol 9, e1000602.

Roberts, C.A., Buikstra, J.E., 2003. The Bioarchaeology of Tuberculosis: A Global View on a Reemerging Disease. University Press of Florida, Gainesville, FL.

Roberts, C.A., Manchester, K., 2005. The Archaeology of Disease, 3rd edition ed. Cornell University PRess, Ithaca.

Roure, B., Philippe, H., 2011. Site-specific time heterogeneity of the substitution process and its impact on phylogenetic inference. BMC Evol Biol 11, 17.

Salo, W.L., Aufderheide, A.C., Buikstra, J.E., Holcomb, T.A., 1994. Identification of Mycobacterium tuberculosis DNA in a Pre-Columbian Peruvian Mummy. PNAS 91, 2091-2094.

Schönian, G., Cupolillo, E., Mauricio, I., 2013. Molecular Evolution and Phylogeny of *Leishmania*, in: Ponte-Sucre, A.e.a. (Ed.), Drug Resistance in Leishmania Parasites. Springer-Verlag, Wien.

Schuenemann, V.J., Bos, K., Dewitte, S., Schmedes, S., Jamieson, J., Mittnik, A., Forrest, S., Coombes, B.K., Wood, J.W., Earn, D.J., White, W., Krause, J., Poinar, H.N., 2011. From the Cover: Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of Yersinia pestis from victims of the Black Death. Proceedings of the National Academy of Sciences USA 108, E746-752.

Schuenemann, V.J., Singh, P., Mendum, T.A., Krause-Kyora, B., Jager, G., Bos, K.I., Herbig, A., Economou, C., Benjak, A., Busso, P., Nebel, A., Boldsen, J.L., Kjellstrom, A., Wu, H., Stewart, G.R., Taylor, G.M., Bauer, P., Lee, O.Y., Wu, H.H., Minnikin, D.E., Besra, G.S., Tucker, K., Roffey, S., Sow, S.O., Cole, S.T., Nieselt, K., Krause, J., 2013. Genome-wide comparison of medieval and modern Mycobacterium leprae. Science 341, 179-183.

Shapiro, B., Gilbert, M.P.T., 2006. No proof that typhoid fever caused the Plague of Athens (a reply to Papagrigorakis et al.). Int J Infect Dis 10, 334-340.

Shapiro, B., Ho, S.Y.W., Drummond, A.J., Suchard, M.A., Pybus, O.G., Rambaut, A., 2011. A Bayesian phylogenetic method to estimate unknown sequence ages. Molecular biology and evolution 28, 879-887.

Spigelman, M., Lemma, E., 1993. The use of the polymerase chain reaction (PCR) to detect Mycobacterium tuberculosis in ancient skeletons. International Journal of Osteoarchaeology 3, 137-143.

Stead, W., Aufderheide, A., Buikstra, J., Holcomb, T., 1995. When did Mycobacterium tuberculosis infection first occur in the New World. American Journal of Respiratory and Critical Care Medicine 151, 1267-1268.

Stead, W.W., 1997. The origin and erratic global spread of tuberculosis: how the past explains the present and is the key to the future. Clin Chest Med 18, 65-77.
Stone, A.C., Wilbur, A.K., Buikstra, J.E., Roberts, C.A., 2009. Tuberculosis and leprosy in perspective. American Journal of Physical Anthropology 140 Suppl 49, 66-94.

Urteaga-Ballon, O., 1991. Medical ceramic representation of nasal leishmaniasis and surgical amputation in ancient Peruvian civilization., Human Paleopathology: Current Synthesis and Future Options. Smithsonian Institution Press, Washington, DC, pp. 95-101.

Wagner, D.M., Klunk, J., Harbeck, M., Devault, A., Waglechner, N., Sahl, J.W., Enk, J., Birdsell, D.N., Kuch, M., Lumibao, C., Poinar, D., Pearson, T., Fourment, M., Golding, B., Riehm, J.M., Earn, D.J.D., DeWitte, S., Rouillard, J.-M., Grupe, G., Wiechmann, I., Bliska, J.B., Keim, P.S., Scholz, H.C., Holmes, E.C., Poinar, H., 2014. Yersinia pestis and the Plague of Justinian 541-543 AD: a genomic analysis. Lancet Infect Dis.

Zelazny, A.M., 2005. Evaluation of 7SL RNA gene sequences for the identification of Leishmania spp,. Am. J. Trop. Med. Hyg. 72, 415-420.

CHAPTER 2

SCREENING ANCIENT TUBERCULOSIS WITH QPCR: CHALLENGES AND OPPORTUNITIES

Kelly M. Harkins[1], Jane E. Buikstra[1], Tessa J. Campbell[2], Kirsten I. Bos[3], Eric D. Johnson[4] and Johannes Krause[2], and Anne C. Stone[1]

[1]School of Human Evolution and Social Change, Arizona State University
[2]Department of Archaeology, University of Cape Town
[3]Department of Archeological Sciences, University of Tübingen
[4]School of Life Sciences, Arizona State University

Author contributions: ACS and JEB conceived of the investigation; KMH performed data analysis; labwork was performed by KMH, KIB and TC; qPCR database and queries written by EDJ and designed by EDJ, KMH (and GH in acknowledgements). KMH wrote the manuscript.

Abstract

The field of ancient DNA (aDNA) has rapidly accelerated in recent years as a result of new methods in next-generation sequencing, library preparation and targeted enrichment. Such research is restricted, however, by the highly variable DNA preservation within different tissues, especially when isolating ancient pathogens from human remains. Identifying positive candidate samples via quantitative PCR (qPCR) for downstream procedures can reduce reagents costs, increase capture efficiency and maximize the number of sequencing reads of the target. This study uses four qPCR assays designed to target regions within the *Mycobacterium tuberculosis complex* to examine 133 human skeletal samples from a wide geographical and temporal range, identified by the presence of skeletal lesions typical of chronic disseminated tuberculosis. Given the inherent challenges working with ancient mycobacteria, strict criteria must be used and primer/probe design continually re-evaluated as new data from bacteria become available. Seven samples tested positive for multiple MTBC loci, supporting them as strong candidates for downstream analyses. Using strict and conservative criteria, qPCR remains a fast and effective screening tool when compared to screening by more expensive sequencing and enrichment technologies.

Introduction

Thirty years have passed since the first ancient DNA (aDNA) analyses, largely made possible by the advent of the polymerase chain reaction (PCR) in 1985. As early as 1993, researchers began to target ancient pathogens such as *Mycobacterium tuberculosis, M. leprae, Treponema pallidum, Yersinia pestis*, and the influenza virus (e.g. Arriaza et al., 1995; Basler et al., 2001; Drancourt et al., 1998; Haas et al., 2000; Kolman et al., 1999; Konomi et al., 2002; Reid et al., 2000; Salo et al., 1994; Spigelman and Lemma, 1993; Taylor et al., 1999). These pioneering studies were largely limited to identifying the presence of the pathogen in archaeological bone or tissue, often by focusing on repetitive or high copy number sequences or by targeting a few phylogenetically informative sites, and were often controversial (Barnes and Thomas, 2006; Gilbert et al., 2005; Gilbert et al., 2004; Stone et al., 2009). The noted exceptions were the studies of the 1918 flu virus and early HIV samples where small complete genomes were accessible via PCR and Sanger sequencing alone (Tumpey et al., 2005; Worobey et al., 2008).

Over the past decade, however, technical advances in library construction, enrichment, and sequencing, specifically for ancient DNA (Gansauge and Meyer, 2013; Hodges et al., 2009; Kircher et al., 2011; Maricic et al., 2010; Meyer and Kircher, 2010), have revolutionized pathogen aDNA studies. These “next-generation” methods enable the study of complete genomes and of samples with little remaining endogenous DNA (Dabney et al., 2013a; Green et al., 2010; Meyer et al., 2014; Meyer et al., 2012), thereby allowing direct access to the evolutionary timescale inherent in heterochronously sampled individuals (Drummond et al., 2004; Rambaut, 2000; Shapiro et al., 2011). Additionally, bioinformatics tools that estimate postmortem damage characteristic for aDNA fragments offer greater confidence in establishing authenticity (Ginolhac et al., 2011; Skoglund et al., 2014). Current protocols can capture complete pathogen genomes that comprise a very small portion of the endogenous DNA in a sample (Bos et al., 2011; Schuenemann et al., 2013) and thus shed light on questions about the causative agents of disease and epidemics, the evolutionary history of pathogens, and co-infection with multiple strains or species (e.g. Chan et al., 2013).

In the current era of "omics", aDNA research remains restricted by the highly variable biomolecular preservation of each bone, tooth, or tissue fragment. To maximize the application of next generation methods, it is beneficial first to identify positive candidate samples for downstream procedures, ultimately reducing reagents costs, increasing capture efficiency and maximizing the number of sequencing reads of the target. Quantitative PCR (qPCR), which is widely used in modern diagnostics, offers a rapid, sensitive, and economic method that assesses the presence and preservation of nucleic acids, and is known to be a reliable screening tool for identifying samples that have sufficient endogenous DNA preservation to warrant deeper sequencing (Enk et al., 2013; Schuenemann et al., 2011). This study reports the results of four qPCR assays designed specifically for detection of Mycobacterium tuberculosis complex (MTBC) DNA in over 130 archaeological human remains from 29 sites in the New and Old World. We find that qPCR continues to be a rapid and dependable screening tool, but care must be taken in primer and probe design to avoid false positives and to properly accommodate pathogen diversity.

Previous tuberculosis aDNA research

Mycobacterium tuberculosis complex (MTBC) includes a group of pathogenic mycobacteria that cause disease in both animals and humans. *M. tuberculosis* is the species most commonly associated with clinical tuberculosis (TB) in humans, although members of the complex that are adapted to other mammalian hosts such as bovids, rodent, and pinnipeds, as well as the ancestral *"M. canetti*" (Brosch et al., 2002) also cause disease in humans.

Human-adapted TB strains are a major cause of mortality and morbidity worldwide, exacerbated by a recent resurgence driven through HIV co-infection and the development of antibiotic resistant forms. The global prevalence and virulence of TB today is the culmination of host-pathogen interactions that extend across millennia, although when these interactions began is controversial (Chapter 3; Comas et al., 2013; Gutierrez et al., 2005; Namouchi et al., 2012). The antiquity of the disease is well documented by historians, archaeologists and paleopathologists,

relying primarily upon skeletal evidence for disease consistent with modern, clinical bone lesions in extrapulmonary TB patients (see Aufderheide et al., 1998; Roberts and Buikstra, 2003). Advancements in aDNA research can in turn be applied to molecular investigations of affected skeletal remains to address the evolutionary history of this remarkably enduring pathogen.

Many of the earliest pathogen aDNA studies focused on TB, with the presence of ancient MTBC DNA first reported in the early 1990s (Salo et al., 1994; Spigelman and Lemma, 1993). These studies used PCR solely to identify the presence or absence of pathogenic bacilli in bone or mummified tissue. Isolating a small fraction of pathogen DNA within an archaeological sample poses a technical challenge, and is further complicated by the close genetic relationship of most pathogenic bacteria to those of environmental origin, which are easily introduced from the burial environment (Campana et al., 2014). Direct sequencing of ancient products, or clones thereof, is required to authenticate ancient PCR results (Cooper and Poinar, 2000). However, genetic regions commonly used for PCR-based detection of MTBC DNA are homologous in some related soil mycobacteria, such as the IS6110 mobile element (Kent et al.; McHugh et al.) and the rpoB gene (discussed below). Additional regions used in past investigations have related homologs in non-tuberculous mycobacteria (NTM): IS1395 in *M. xenopi*, for example, shares 86% sequence identity with the IS1081 element of the MTBC (Picardeau et al., 1996). If primers are located in conserved regions, homology can cause false-positive PCR results in NTM, as shown by (Savelkoul et al., 2006) with the IS6110 element. Ideally, primers should target regions with enough genetic diversity to avoid false-positives from related bacteria, and should be evaluated manually through multiple sequence alignments or by utilizing search features of online databases such as NCBI's Basic Local Sequence Alignment (BLAST) tool. Unfortunately, our understanding of environmental mycobacterial sequence diversity is limited by available data; in other words, sequence alignments may not represent the true amount of diversity amplifiable with a given primer set. Thus performing periodic cross-checks of primer pairs as new data become accessible helps ensure the PCR target regions remain specific.

With the availability of more extensive sequence information and improved methods of detection, ancient TB studies are moving from mere identification to more complex and more intriguing questions about the history and evolution of the MTBC across host species and through time (e.g. Bouwman et al., 2012; Chan et al., 2013). The rapid identification of well-preserved DNA remains the first step in the costly current state of ancient TB research.

Detecting ancient DNA via Quantitative Real-Time PCR

Quantitative PCR is capable of detecting short and low-quantity starting templates and thus has many applications for ancient genetics, as reviewed by Bunce et al. (2012). Taqman qPCR technology provides additional specificity over SYBR chemistries by including a non-extendable DNA probe, which hybridizes to a region within the amplicon, reducing the amplification of non-specific molecules (Heid et al., 1996). The assumption when targeting ancient DNA is that the fragment of interest is conserved enough to be amplified by the primers/probe set. If the assumption is incorrect, or if miscoding lesions and other damage common to aDNA occur at crucial binding sites, qPCR efficiency is reduced mimicking the amplification curves of a sample co-extracted with inhibitory substances (King et al., 2009).

Aside from mismatches, a number of other factors affect the success and efficiency of an amplification reaction, including production of secondary structures, repetitive sequences, and high GC/AT content (Frey et al., 2008; Gallup, 2011; Roux, 1995). Such factors are further compounded when templates are degraded, damaged, and low in quantity (e. g. Binladen et al., 2006; Bunce et al., 2012; Dabney et al., 2013b; King et al., 2009; Mitchell et al., 2005). While we are often concerned with false-positives in ancient DNA, inhibition can cause false-negatives—an equally undesirable result. Methods are being developed continuously to improve retention of endogenous molecules and to reduce the quantity of inhibitors co-extracted during purification (i.e. Boom et al., 1990; Hänni et al., 1995; Rohland and Hofreiter, 2007). In addition, creating a dilution series with an ancient extract, while reducing the number of endogenous templates, can reveal the presence of inhibitory substances. A 10-fold dilution theoretically amplifies 3.32 PCR

cycles later than an undiluted sample. Deviations from this pattern due to delays in amplification of undiluted samples provide an indication of PCR inhibition. Similarly, levels of inhibition are gauged by examining variation in the slopes of exponential amplification plots, spiking internal controls with ancient extracts (King et al., 2009) or by simultaneously running UNG-treated extracts in parallel (Pruvost et al., 2005). Unfortunately, the exact chemistry of each qPCR reaction can differ; according to (Huggett et al., 2008): "It is possible for one PCR reaction to be unaffected by a potential inhibitor whilst another is completely suppressed." Late occurring or non-exponential amplification often is seen with both inhibited and low-copy number samples, as well as with template mismatches (Smith et al., 2002; Süss et al., 2009); consequently the cycle numbers are often increased compared to conventional PCR with modern template.
Interpretations of qPCR data rely on user-based threshold and baseline determinations.

Therefore, quantification estimates in qPCR software output must be interpreted with care. Given the obstacles of quantifying aDNA and the potential inaccuracies of copy number estimates, a simple presence/absence determination for a target sequence may be the most appropriate level of resolution, rather than relying on quantity estimates when qPCR is used as a screening tool for downstream analyses.

Material and methods

The objective of this research is to screen archaeological human remains for preserved MTBC DNA and to identify candidates for targeted enrichment and genome sequencing on NGS platforms. We examined 133 human skeletal samples from a wide geographical and temporal range identified by the presence of skeletal lesions typical of skeletal TB, listed in Table 1 and Table S1. Despite methodological advances, sample preparation and contamination remain critical considerations. Contaminant DNA, even at low quantities, can be preferentially amplified (Krause et al.). Ancient genetic research benefits from stringent, well-defined and controlled procedures for nucleic acid isolation, which are requisite for any aDNA facility (Hofreiter et al., 2001; Paabo et al., 2004). All work performed in this investigation was carried out in the Ancient

DNA Laboratory at Arizona State University. This facility is separated from post-PCR labs and contains a positive pressure class 10,000 HEPA filtered clean room, UV irradiated every evening and after every use, as is standard practice (Cooper and Poinar, 2000). All personnel and reagents maintain a unidirectional workflow between ancient and modern DNA facilities to eliminate post-PCR contamination. Disposable facemask, double gloves and hooded coveralls are worn at all times; gloves are changed between samples. After each use, surfaces, bench top equipment and pipettes are washed with bleach, 70% ethanol and sterile water.

Bone fragments and teeth were briefly rinsed with 3% bleach and sterile water to remove surface contaminants. Between 50-250 mg of bone powder were used in each extraction. Four silica and spin-column based extraction protocols were used over the course of the project (e. g. (Dabney et al., 2013a; Höss and Paabo, 1993; Rohland and Hofreiter, 2007; Yang et al., 1998). Contamination during DNA extraction was evaluated with extraction blanks; 5 µl of each extract and blank were used in conventional PCR to test for presence of an 80 bp region of human mtDNA (Wrischnik et al., 1987). Extraction blanks showing any potential human contamination were re-extracted. PCR blanks also confirmed uncontaminated reactions.

qPCR assays

The presence and quantity of MTBC DNA in each extract were examined with four separate TaqMan® qPCR assays, two targeting the single-copy rpoB gene, referred to as rpoB1 and rpoB2, and two targeting multi-copy insertion elements, IS6110 and IS1081, thought to be specific to the MTBC complex (Collins and Stephens, 1991; Eisenach et al., 1990; Thierry et al., 1990; van Soolingen et al., 1993). The rpoB2 and IS1081 primer/probe sequences were designed for this investigation (Table 1). The IS6110 primer/probe set targets a 63bp fragment that avoids a region of homology in other mycobacterial species as identified by McHugh et al. (1997) (see Klaus et al. 2010). Although the rpoB1 assay reportedly contains 15 differences between the MTBC ecotypes and other mycobacterial species (Klaus et al.), a multiple alignment of newly available mycobacterial sequences showed that the rpoB1 probe/primers could amplify species outside of the complex. In testing this observation, the rpoB1 qPCR assay detected the presence

of species outside the MTB complex, including *M. marinum, M. avium, M. leprae, M. kansasii,* and *"M. lufu".* As discussed above, PCR efficiency is negatively impacted if mismatches or inhibition are present. Interestingly, we observe a number of non-exponential or late amplification signals that mimic results caused by inhibition and/or mismatches. A representative multiple sequence alignment is crucial for MTBC-specific primer design because mismatches between the reference and closely related bacteria can be placed at the 3' or 5' priming sites, known to affect binding efficiency most heavily (Hummel, 2003). To reduce the risk of mispriming, our redesigned rpoB2 assay amplifies a 70 bp region where nucleotide differences are most heavily concentrated at the 3' and 5' ends of the primers and probe, respectively, thus increasing specificity of the intended product within the MTBC.

To evaluate whether non-specific products, such as primer-dimers, are recorded as false-positives, 51 sterile water samples were amplified with our rpoB1 and IS6110 assays for 60 cycles. No non-specific products were recorded with IS6110. For rpoB1, seven wells displayed a late amplifying curve, crossing the threshold after 50 cycles (n=6), and at 46 cycles (n=1). These false-positives underscore the caution taken during data analysis, particularly with the rpoB1 locus.

All ancient extracts and a subset of extract dilutions (1:10, 1:100) were tested for each of these regions by comparison to a standard curve of serial dilutions created from genomic DNA of *M. tuberculosis* H37rv (ATCC 25618 or ATCC 27294) in a 10-fold serial dilution of 1 to 100,000 copy numbers per microliter, optimized for each assay. It is important to note that once the ancient samples are loaded into the qPCR plate in the cleanroom, they are sealed and not reopened in any post-PCR lab; the standards are then loaded into one corner of the plate that remains unsealed. The primers used for the standard and unknown ancient samples are identical and thus produce the same amplicon; fluorescence is recorded only when the probe hybridizes to the target sequence. In each run, two or three replicates, depending on available extract, of each ancient sample are included, and non-template controls are included in triplicate. Reactions are in

a 20-µL total volume; 10 µL of TaqMan® 2x Universal Master Mix, primers, probe (see Table 2 for primer/probe sequence and concentration), 10mg/ml RSA, and 2 µL of sample. Thermocycling conditions for the ABI 7900 consisted of an initial enzyme activation at 95˚C for 10 min, followed by 45-50 cycles at 95˚C for 15s and 60˚C for 1 min.

Results and Discussion

Our qPCR screening of archaeological skeletal remains identified 28 samples with evidence of mycobacterial DNA, but nine with evidence for MTBC-specific DNA (i.e. excluding rpoB1), with seven of those positive for multiple loci. To avoid misinterpreting false positives identified by the software's output, each sample was evaluated individually. With low copy number templates, as expected with ancient material, fluorescence often occurred relatively late in the reaction and had to be evaluated against any low-level background signal. Multi-component plots were used in late amplifications to evaluate each replicate's fluorescence signal crossing over the background signal or the passive reference before confirming the well as positive. A sample extract is considered positive for an assay if more than 50% of all replicates cross the automatic or user-defined threshold. With these criteria, 28 of 133 samples showed a positive signal in at least one of the four qPCR assays. These data are listed in Table 3 and SI Tables.

To evaluate inhibition, extract dilutions of a subsample (N=32) were also tested, revealing at least seven inhibited samples. Inhibition can produce false-negative results but also prohibit reliable estimation of starting copy numbers when compared to a standard of known quantity, in some cases producing theoretically impossible values that are well below one copy number (see amplification plot in Figure 2). In eight extracts, amplification signals were recorded only in the dilutions but not in the undiluted extract, only some of which met the 50% positive replicate criterion. Five samples showed lower or similar Ct values in the dilutions compared to the raw extracts, rather than approximately +3.32 cycle difference expected. The five potentially inhibited samples categorized as negative with our criteria could be considered again if further measures are taken to address inhibition. Nevertheless, it is useful information for downstream steps.

As mentioned earlier, the rpoB1 assay amplified a conserved gene region in species outside the complex at low levels. By factoring in additional sequence information and mismatch placement, our redesigned assay, rpoB2, achieved increased strain specificity. We therefore omit rpoB1 in our discussion of positive MTBC results, although all assays are reported in the Table S3. With this additional criterion, nine samples are positive for at least one of the remaining three assays: IS1081 (N=7), IS6110 (N=7), rpoB2 (N=7). These samples are located throughout North and South America (Figure 1). In most cases, the same and/or a new extraction of each sample were repeated. Extraction blanks associated with these samples were negative in all assays for MTBC DNA, as were non-template controls. Most convincing as candidate samples were those positive for multiple (2+) MTBC loci. These include samples from seven sites: one historic (Highland Park) and one proto-historic site (Arikara) in North America and five pre-contact South American sites, Chiribaya Alta (Sample 54), El Algodonal (Sample 58), El Yaral (Sample 64), Las Delícias y Candeleria (Sample 281) and Estuquiña (Sample 82). All efforts to clone and/or direct sequence the qPCR product via Sanger sequencing were unsuccessful, the latter likely due to short amplicon length and/or low DNA concentration. We acknowledge that it is impossible to authenticate ancient fragments with PCR data alone, and that the qPCR results were used primarily for identifying candidate samples for targeted enrichment. Four of the seven positive samples listed above from Chiribaya Alta, El Algodonal, El Yaral and Estuquiña were independently replicated by in-solution capture designed to target genes within the MTBC (Chapter 3). Sample 128 (Highland Park) did not meet the cutoff for full genome capture but nonetheless showed some reads mapping to MTBC (See Chapter 3). Three of these four positive samples (54, 58, 64) were enriched further, generating three complete genome sequences (Chapter 3). The genome enrichment for the fourth sample (82) is underway. The two remaining samples testing positive for multiple assays in our qPCR screening (15, 281) only did so after extraction with a newer protocol (Dabney et al.), which could account for the discrepancy between qPCR and the negative in-solution enrichment results. Reanalysis of earlier qPCR data as well as newly tested extracts from Lambayeque ILL-22, reported positive in (Klaus et al.), reveal the absence of IS6110 and other MTBC loci, according to our stricter criteria. Likewise,

bone from Chiribaya Alta T-30, originally detected from mummified lung tissues reported in (Salo et al.), was not found to be positive in the current study. This negative result may highlight a fundamental difference in recovering pathogen DNA from different tissue types.

Interestingly, four of the seven most well-preserved MTBC-positive samples are from four sites located in the Osmore River valley on the coast of southern Peru (Figure 1). This general region has been the focus of other bioarchaeological, paleopathological and molecular investigations of ancient TB (e.g. Arriaza et al., 1995; Buikstra and Williams, 1991; Burgess, 1999). Identification of acid-fast bacilli within a paravertebral abscess in a subadult near Nasca, Peru (AD 200-800) (Allison et al., 1973), and diagnostic lung lesions from mummified remains at Chiribaya Alta were among the first reports of ancient TB DNA in the area and in the world (Salo et al. 1994), though our analysis of bone from the latter individual was negative for qPCR. Most recently, as Chapter 3 reports, the individuals from three sites along the Osmore River drainage share a unique MTBC strain most closely related to *M. pinnipedii*, a pathogen of marine mammals. The arid desert of Peru's southern coast likely contributes to the surprisingly high recovery of positive MTBC cases in our study. An increasing number of skeletal cases in the region approaching AD 1000 suggest this area could have been a hotspot of ancient "tuberculosis" (Roberts and Buikstra, 2003). This is not surprising considering TB-like diseases are density dependent. The expanding evidence that humans were infected with pathogenic MTBC DNA in pre-contact South America provide a new starting point to examine the larger, evolutionary questions regarding the origin, dispersal and adaptation of this disease-causing clade of mycobacteria; thus the rapid identification of pathogen DNA in ancient samples is an important first step for further analyses targeting pathogen genomes. This study also indicates the presence of MTBC DNA in an individual from the South Dakota Arikara site of Mobridge, occupied around the time of European contact (AD 1500-1750). Several disseminated skeletal TB cases have been identified from a number of Arikara sites (Palkovich, 1981). A transmission scenario very different from that of the southern continent may emerge from this period in North America, e.g. evidence for infection from

European strains or of an MTBC strain that crossed the Bering Strait, if planned enrichment and sequencing efforts are successful.

Conclusion

Moving beyond simple confirmatory analysis of diseased bone, researchers are now asking nuanced research questions capable of confronting the debate regarding TB's origins and evolution. These questions continue to be at the forefront of not only ancient pathogen research, but also (bio)archaeological and paleopathological scholarship. Here we report the results of screening for presence of mycobacterial DNA archaeological skeletal remains. The speed of these analyses is unrivaled compared to screening with other methods, e.g. in-solution capture; in a 384-well plate, results for multiple samples are available in less than 3 hours and require no high-throughput computation. Using multiple qPCR loci currently thought to be specific for the TB complex, we report nine positive samples from skeletons ranging in age from AD 770 to as late as AD 1850 from the Americas. It is important to note that mycobacterial species other than *M. tuberculosis sensu stricto* can cause TB-like disease in humans, and thus focused MTBC assays are best designed around conserved regions within the complex that differ from species outside the complex at crucial binding sites. Although the rpoB1 qPCR assay alone could not be used to determine the presence of MTBC DNA, there remains a possibility that samples identified solely with rpoB1 reflect zoonotic events by other related mycobacteria, e. g. *M. avium*; such events could be the topic of further investigation.

In summary, there are several ways to overcome the obstacles of using damaged, degraded DNA in qPCR analyses: 1. Stringency in sample preparation to reduce co-extraction of inhibitors; 2. Taqman probe-based assays to increase sensitivity of detection; 3. Regular in silico re-examination of newly published sequence data and when appropriate, primer/probe re-design, 4. Serial dilutions of extractions to evaluate inhibition; 5. Use of multiple replicates in evaluating a positive signal. If no enrichment/sequencing projects are planned with the candidate samples, direct sequencing can strengthen claims of authenticity. Reporting the presence or absence of

pathogen aDNA is no longer an acceptable final research objective, and as (Wilbur et al., 2009) expressed, its simple detection tells us little of value about the existence of the disease in the past. NGS genome data are the new objective for ancient pathogen research. Improving molecular techniques, specifically whole genome sequencing of modern MTBC strains, continues to strengthen investigations of the microevolutionary processes that contribute to virulence and the host-pathogen co-evolution of the TB disease complex evolution of the TB disease complex (Comas et al., 2013; Gagneux et al., 2006). While in its nascent stages, this level of ancient data acquisition portends investigations on the same scale as modern genomics.

Ultimately, there are known limits to qPCR, and it is important to avoid over-interpreting data and to be conservative in each evaluation to avoid false-positives. In ancient DNA, the adage that the absence of evidence is not evidence of absence certainly plays a role in sample choice and experimental design. It is expectable and demonstrated here that samples qPCR positive for multiple MTBC loci were the best indicators of aDNA preservation for full genome reconstruction. As a proxy for downstream analyses, probe-based qPCR technology remains a fast and more cost-effective assessment of aDNA preservation when compared to screening by more expensive enrichment and sequencing technologies.

Acknowledgements:

Funding: We would like to acknowledge the following sources of funding: The National Science Foundation (ACS and JEB) for NSF BCS-1063939, NSF-REU BCS-0612222 and NSF BCS-0612222, the George E. Burch Fellow in Theoretic Medicine and Affiliated Sciences at the Smithsonian Institution (2003-2007, to JEB), the NSF Graduate Research Fellowship DGE-1311230 and Jacob K. Javits Fellowship to KMH. Social Sciences and Humanities Research Council of Canada postdoctoral fellowship grant 756-2011-501(to KIB). The Wenner-Gren Wadsworth Fellowship (to TJC).

People and Institutions: We would like to thank the following people and institutions for assistance and/or permission for sampling: Museo Contisuyu, Centro Mallqui, Sonia Guillen,

Instituto Nacional de Cultura, Peru, Guillermo Cock, Cathy Gaither, Melissa Murphy, Maria Cecilia Lozada, Shelley Burgess, Deborah Blom, Bruce Owen, Lic. Antonio Oquiche Hernani, Lic. Patricia Palacios Filinich, Sloan Williams, Bertha Vargas, Don Rice, Ricardo Guichón and Haagen Klaus. Susan Pfieffer, the University of Toronto, the Upper Mississippi Valley Archaeological Research Foundation, Larry Conrad, Indiana University, Della Cook, the Illinois State Museum, George Milner, the Pennsylvania State University, the American Museum of Natural History, Ann Stodder, Dawnie Steadman, Bob Brier, Ian Tattersall, Ken Mowbray, the National Museum of Natural History (Smithsonian Institution), Bill Billeck, Noreen Tuross, Rochester Museum & Science Center, Lorraine P. Saunders, the Museum of London, Rebecca Redfern, Takao Suzuki and the Tokyo Metropolitan Institute of Gerontology, Soprintendenza ai Beni Archeologici della Liguria, Vincenzo Formicola, Silvia Smith, Rafael Goñi and Solana Garcia Guraieb, Alexandra Buzhilova/Russian Academy of Sciences, members of the Stone lab, Tanvi Honap, Maria Nieves-Colon, Charlotte Till, and also Genevieve Housman for her extensive contribution to the qPCR database. The Huron Wendat Nation is aware of the sampling of Uxbridge bone and is a recipient of information from this study.

References

Allison, M.J., Mendoza, D., Pezzia, A., 1973. Documentation of a case of tuberculosis in Pre-Columbian America. Am Rev Respir Dis 107, 985-991.

Arriaza, B.T., Salo, W., Aufderheide, A.C., Holcomb, T.A., 1995. Pre-Columbian tuberculosis in northern Chile: molecular and skeletal evidence. American Journal of Physical Anthropology 98, 37-45.

Aufderheide, A.C., Conrado, R., Langsjoen, O., 1998. The Cambridge Encyclopedia of Human Paleopathology. Cambridge University Press.

Barnes, I., Thomas, M.G., 2006. Evaluating bacterial pathogen DNA preservation in museum osteological collections. Proceedings of the Royal Society - biological sciences 273, 645-653.

Basler, C.F., Reid, A.H., Dybing, J.K., Janczewski, T.A., Fanning, T.G., Zheng, H., Salvatore, M., Perdue, M.L., Swayne, D.E., Garcia-Sastre, A., Palese, P., Taubenberger, J.K., 2001. Sequence of the 1918 pandemic influenza virus nonstructural gene (NS) segment and characterization of recombinant viruses bearing the 1918 NS genes. Proceedings of the National Academy of Sciences USA 98, 2746-2751.

Binladen, J., Wiuf, C., Gilbert, M.T.P., Bunce, M., Barnett, R., Larson, G., Greenwood, A.D., Haile, J., Ho, S.Y., Hansen, A.J., 2006. Assessing the fidelity of ancient DNA sequences amplified from nuclear genes. Genetics 172, 733-741.

Boom, R., Sol, C.J., Salimans, M.M., Jansen, C.L., Wertheim-van Dillen, P.M., van der Noordaa, J., 1990. Rapid and simple method for purification of nucleic acids. J Clin Microbiol 28, 495-503.

Bos, K.I., Harkins, K.M., Herbig, A., Coscolla, M., Weber, N., Comas, I., Forrest, S.A., Bryant, J.M., Harris, S.R., Schuenemann, V.J., Campbell, T.J., Majander, K., Wilbur, A.K., Guichon, R.A., Wolfe Steadman, D.L., Cook, D.C., Niemann, S., Behr, M.A., Zumarraga, M., Huson, D., Knieselt, K., Young, D., Parkhill, J., Buikstra, J.E., Gagneux, S., Stone, A.C., Krause, J., in press. Pre-Columbian Mycobacterial Genomes Reveal Seals as a Source of New World Human Tuberculosis. Nature.

Bos, K.I., Schuenemann, V.J., Golding, G.B., Burbano, H.A., Waglechner, N., Coombes, B.K., McPhee, J.B., Dewitte, S.N., Meyer, M., Schmedes, S., Wood, J., Earn, D.J., Herring, D.A., Bauer, P., Poinar, H.N., Krause, J., 2011. A draft genome of Yersinia pestis from victims of the Black Death. Nature 478, 506-510.

Bouwman, A.S., Kennedy, S.L., Muller, R., Stephens, R.H., Holst, M., Caffell, A.C., Roberts, C.A., Brown, T.A., 2012. Genotype of a historic strain of Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences USA 109, 18511-18516.

Brosch, R., Gordon, S.V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L.M., Pym, A.S., Samper, S., van Soolingen, D., Cole, S.T., 2002. A new evolutionary scenario for the Mycobacterium tuberculosis complex. Proceedings of the National Academy of Sciences USA 99, 3684-3689.

Buikstra, J.E., Williams, S.R., 1991. Tuberculosis in the Americas: Current perspectives, in: Ortner, D.,

Aufderheide, A.C. (Eds.), Human paleopathology: Current syntheses and future options. . Smithsonian Institution Press, Washington, D.C., pp. 161–172.

Bunce, M., Oskam, C.L., Allentoft, M.E., 2012. Quantitative Real-Time PCR in aDNA Research, in: Shapiro, B., Hofreiter, M. (Eds.), Ancient DNA: Methods and Protocols. Humana Press, California, pp. 121-132.

Burgess, S.D., 1999. Chiribaya Skeletal Pathology on the South Coast of Peru: Patterns of Production and Consumption, Department of Anthropology. University of Chicago.

Campana, M.G., Robles Garcia, N., Ruhli, F.J., Tuross, N., 2014. False positives complicate ancient pathogen identifications using high-throughput shotgun sequencing. BMC Res Notes 7, 111.

Chan, J.Z., Sergeant, M.J., Lee, O.Y., Minnikin, D.E., Besra, G.S., Pap, I., Spigelman, M., Donoghue, H.D., Pallen, M.J., 2013. Metagenomic analysis of tuberculosis in a mummy. New England Journal of Medicine 369, 289-290.

Collins, D.M., Stephens, D.M., 1991. Identification of an insertion sequence, IS1081, in Mycobacterium bovis. FEMS Microbiol Lett 67, 11-15.

Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K.E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., Bothamley, G., Mei, J., Wei, L., Bentley, S., Harris, S.R., Niemann, S., Diel, R., Aseffa, A., Gao, Q., Young, D., Gagneux, S., 2013. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nature genetics 45, 1176-1182.

Cooper, A., Poinar, H.N., 2000. Ancient DNA: do it right or not at all. Science 289, 1139.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.T., Weihmann, A., Nickel, B., Valdiosera, C., Garcia, N., Paabo, S., Arsuaga, J.L., Meyer, M., 2013a. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences USA 110, 15758-15763.

Dabney, J., Meyer, M., Paabo, S., 2013b. Ancient DNA damage. Cold Spring Harb Perspect Biol 5.

Drancourt, M., Aboudharam, G., Signoli, M., Dutour, O., Raoult, D., 1998. Detection of 400-year-old Yersinia pestis DNA in human dental pulp: an approach to the diagnosis of ancient septicemia. Proceedings of the National Academy of Sciences USA 95, 12637-12640.

Drummond, A., Nicholls, G.K., Rodrigo, A.G., Solomon, W., 2004. Genealogies from time-stamped sequence data, in: Buck, C.E., Millard, A. (Eds.), Tools for Constructing Chronologies: Crossing Disciplinary Boundaries. Springer, London, pp. 149-171.

Eisenach, K.D., Cave, M.D., Bates, J.H., Crawford, J.T., 1990. Polymerase chain reaction amplification of a repetitive DNA sequence specific for Mycobacterium tuberculosis. The Journal of infectious diseases 161, 977-981.

Enk, J., Rouillard, J.-M., Poinar, H., 2013. Quantitative PCR as a predictor of aligned ancient DNA read counts following targeted enrichment. BioTechniques 55, 300-309.

Frey, U.H., Bachmann, H.S., Peters, J., Siffert, W., 2008. PCR-amplification of GC-rich regions: 'slowdown PCR'. Nat Protoc 3, 1312-1317.

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B.C., Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M.C., Hilty, M., Hopewell, P.C., Small, P.M., 2006. Variable

host-pathogen compatibility in Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences USA 103, 2869-2873.

Gallup, J.M., 2011. qPCR Inhibition and Amplification of Difficult Templates, in: Kennedy, S., Oswald, N. (Eds.), PCR Troubleshooting and Optimization: The Essential Guide. Caister Academic Press, pp. 23-66.

Gansauge, M.-T., Meyer, M., 2013. Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nat Protoc 8, 737-748.

Gilbert, M.T., Barnes, I., Collins, M.J., Smith, C., Eklund, J., Goudsmit, J., Poinar, H., Cooper, A., 2005. Long-term survival of ancient DNA in Egypt: response to Zink and Nerlich (2003). American Journal of Physical Anthropology 128, 110-114; discussion 115-118.

Gilbert, M.T., Cuccui, J., White, W., Lynnerup, N., Titball, R.W., Cooper, A., Prentice, M.B., 2004. Absence of Yersinia pestis-specific DNA in human teeth from five European excavations of putative plague victims. Microbiology 150, 341-354.

Ginolhac, A., Rasmussen, M., Gilbert, M.T., Willerslev, E., Orlando, L., 2011. mapDamage: testing for damage patterns in ancient DNA sequences. Bioinformatics 27, 2153-2155.

Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H., Hansen, N.F., Durand, E.Y., Malaspinas, A.S., Jensen, J.D., Marques-Bonet, T., Alkan, C., Prufer, K., Meyer, M., Burbano, H.A., Good, J.M., Schultz, R., Aximu-Petri, A., Butthof, A., Hober, B., Hoffner, B., Siegemund, M., Weihmann, A., Nusbaum, C., Lander, E.S., Russ, C., Novod, N., Affourtit, J., Egholm, M., Verna, C., Rudan, P., Brajkovic, D., Kucan, Z., Gusic, I., Doronichev, V.B., Golovanova, L.V., Lalueza-Fox, C., de la Rasilla, M., Fortea, J., Rosas, A., Schmitz, R.W., Johnson, P.L., Eichler, E.E., Falush, D., Birney, E., Mullikin, J.C., Slatkin, M., Nielsen, R., Kelso, J., Lachmann, M., Reich, D., Paabo, S., 2010. A draft sequence of the Neandertal genome. Science 328, 710-722.

Gutierrez, M.C., Brisse, S., Brosch, R., Fabre, M., Omaïs, B., Marmiesse, M., Supply, P., Vincent, V., 2005. Ancient origin and gene mosaicism of the progenitor of Mycobacterium tuberculosis. PLoS Pathogens 1, e5.

Haas, C.J., Zink, A., Palfi, G., Szeimies, U., Nerlich, A.G., 2000. Detection of leprosy in ancient human skeletal remains by molecular identification of Mycobacterium leprae. Am J Clin Pathol 114, 428-436.

Hänni, C., Brousseau, T., Laudet, V., Stehelin, D., 1995. Isopropanol precipitation removes PCR inhibitors from ancient bone extracts. Nucleic Acids Research 23, 881.
Heid, C.A., Stevens, J., Livak, K.J., Williams, P.M., 1996. Real time quantitative PCR. Genome Res 6, 986-994.

Hodges, E., Rooks, M., Xuan, Z., Bhattacharjee, A., Benjamin Gordon, D., Brizuela, L., Richard McCombie, W., Hannon, G.J., 2009. Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. Nat Protoc 4, 960-974.

Hofreiter, M., Serre, D., Poinar, H.N., Kuch, M., Paabo, S., 2001. Ancient DNA. Nature Reviews Genetics 2, 353-359.

Höss, M., Paabo, S., 1993. DNA extraction from Pleistocene bones by a silica-based purification method. Nucleic Acids Research 21, 3913-3914.

Huggett, J.F., Novak, T., Garson, J.A., Green, C., Morris-Jones, S.D., Miller, R.F., Zumla, A., 2008. Differential susceptibility of PCR reactions to inhibitors: an important and unrecognised phenomenon. BMC Res Notes 1, 70.

Hummel, S., 2003. Ancient DNA typing: methods, strategies and applications. Springer.

Kent, L., McHugh, T.D., Billington, O., Dale, J.W., H., G.S., 1995. Demonstration of homology between IS6110 of Mycobacterium tuberculosis and DNAs of other Mycobacterium spp.? J Clin Microbiol 33, 2290-2293.

King, C., Debruyne, R., Kuch, M., Schwarz, C., Poinar, H., 2009. A quantitative approach to detect and overcome PCR inhibition in ancient DNA extracts. BioTechniques 47, 941-949.

Kircher, M., Sawyer, S., Meyer, M., 2011. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Research 40, e3.

Klaus, H.D., Wilber, A.K., Temple, D.H., Buikstra, J.E., Stone, A.C., Fernandez, M., Wester, C., Tam, M.E., 2010. Tuberculosis on the north coast of Peru: skeletal and molecular paleopathology of late pre-Hispanic and postcontact mycobacterial disease. Journal of Archaeological Science 37, 2587-2597.

Kolman, C.J., Centurion-Lara, A., Lukehart, S.A., Owsley, D.W., Tuross, N., 1999. Identification of Treponema pallidum subspecies pallidum in a 200-year-old skeletal specimen. Journal of Infectious Diseases 180, 2060-2063.

Konomi, N., Lebwohl, E., Mowbray, K., Tattersall, I., Zhang, D., 2002. Detection of mycobacterial DNA in Andean mummies. J Clin Microbiol 40, 4738-4740.

Krause, J., Briggs, A.W., Kircher, M., Maricic, T., Zwyns, N., Derevianko, A., Paabo, S., 2010. A complete mtDNA genome of an early modern human from Kostenki, Russia. Curr Biol 20, 231-236.

Maricic, T., Whitten, M., Paabo, S., 2010. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. PLoS ONE 5, e14004.

McHugh, T.D., Newport, L.E., Gillespie, S.H., 1997. IS6110 homologs are present in multiple copies in mycobacteria other than tuberculosis-causing mycobacteria. J Clin Microbiol 35, 1769-1771.

Meyer, M., Fu, Q., Aximu-Petri, A., Glocke, I., Nickel, B., Arsuaga, J.L., Martinez, I., Gracia, A., de Castro, J.M., Carbonell, E., Paabo, S., 2014. A mitochondrial genome sequence of a hominin from Sima de los Huesos. Nature 505, 403-406.

Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc 2010, pdb prot5448.

Meyer, M., Kircher, M., Gansauge, M.T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prufer, K., de Filippo, C., Sudmant, P.H., Alkan, C., Fu, Q., Do, R., Rohland, N., Tandon, A., Siebauer, M., Green, R.E., Bryc, K., Briggs, A.W., Stenzel, U., Dabney, J., Shendure, J., Kitzman, J., Hammer, M.F., Shunkov,
M.V., Derevianko, A.P., Patterson, N., Andres, A.M., Eichler, E.E., Slatkin, M., Reich, D., Kelso, J., Paabo, S., 2012. A high-coverage genome sequence from an archaic Denisovan individual. Science 338, 222-226.

Mitchell, D., Willerslev, E., Hansen, A., 2005. Damage and repair of ancient DNA. Mutat Res 571, 265-
276.

Namouchi, A., Didelot, X., Schock, U., Gicquel, B., Rocha, E.P., 2012. After the bottleneck: Genome-wide diversification of the Mycobacterium tuberculosis complex by mutation, recombination, and natural selection. Genome Res 22, 721-734.

Paabo, S., Poinar, H., Serre, D., Jaenicke-Despres, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., Hofreiter, M., 2004. Genetic analyses from ancient DNA. Annu Rev Genet 38, 645-679.

Palkovich, A.M., 1981. Tuberculosis epidemiology in two Arikara skeletal samples: A study of disease impact, in: Buikstra, J.E. (Ed.), Prehistoric tuberculosis in the Americas. Northwestern University Archeological Program, Evanston, Ill., pp. 161–175.

Picardeau, M., Varnerot, A., Rauzier, J., Gicquel, B., Vincent, V., 1996. Mycobacterium xenopi IS1395, a novel insertion sequence expanding the IS256 family. Microbiology 142 (Pt 9), 2453-2461.

Pruvost, M., Grange, T., Geigl, E.M., 2005. Minimizing DNA contamination by using UNG-coupled quantitative real-time PCR on degraded DNA samples: application to ancient DNA studies. BioTechniques 38, 569-575.

Rambaut, A., 2000. Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. Bioinformatics 16, 395-399.

Reid, A.H., Fanning, T.G., Janczewski, T.A., Taubenberger, J.K., 2000. Characterization of the 1918 "Spanish" influenza virus neuraminidase gene. Proceedings of the National Academy of Sciences USA 97, 6785-6790.

Roberts, C.A., Buikstra, J.E., 2003. The Bioarchaeology of Tuberculosis: A Global View on a Reemerging Disease. University Press of Florida, Gainesville, FL.

Rohland, N., Hofreiter, M., 2007. Comparison and optimization of ancient DNA extraction. BioTechniques 42, 343-352.

Roux, K.H., 1995. Optimization and troubleshooting in PCR. Genome Research 4, S185-S194.

Salo, W.L., Aufderheide, A.C., Buikstra, J.E., Holcomb, T.A., 1994. Identification of Mycobacterium tuberculosis DNA in a Pre-Columbian Peruvian Mummy. PNAS 91, 2091-2094.

Savelkoul, P.H., Catsburg, A., Mulder, S., Oostendorp, L., Schirm, J., Wilke, H., van der Zanden, A.G., Noordhoek, G.T., 2006. Detection of Mycobacterium tuberculosis complex with Real Time PCR: comparison of different primer-probe sets based on the IS6110 element. J Microbiol Methods 66, 177-180.

Schuenemann, V.J., Bos, K., Dewitte, S., Schmedes, S., Jamieson, J., Mittnik, A., Forrest, S., Coombes, B.K., Wood, J.W., Earn, D.J., White, W., Krause, J., Poinar, H.N., 2011. From the Cover: Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of Yersinia pestis from victims of the Black Death. Proceedings of the National Academy of Sciences USA 108, E746-752.

Schuenemann, V.J., Singh, P., Mendum, T.A., Krause-Kyora, B., Jager, G., Bos, K.I., Herbig, A., Economou, C., Benjak, A., Busso, P., Nebel, A., Boldsen, J.L., Kjellstrom, A., Wu, H., Stewart, G.R.,

Taylor, G.M., Bauer, P., Lee, O.Y., Wu, H.H., Minnikin, D.E., Besra, G.S., Tucker, K., Roffey, S., Sow, S.O., Cole, S.T., Nieselt, K., Krause, J., 2013. Genome-wide comparison of medieval and modern Mycobacterium leprae. Science 341, 179-183.
Shapiro, B., Ho, S.Y.W., Drummond, A.J., Suchard, M.A., Pybus, O.G., Rambaut, A., 2011. A Bayesian phylogenetic method to estimate unknown sequence ages. Molecular biology and evolution 28, 879-887.

Skoglund, P., Northoff, B.H., Shunkov, M.V., Derevianko, A.P., Paabo, S., Krause, J., Jakobsson, M., 2014. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. Proc Natl Acad Sci U S A 111, 2229-2234.
Smith, S., Vigilant, L., Morin, P.A., 2002. The effects of sequence length and oligonucleotide mismatches on 5' exonuclease assay efficiency. Nucleic Acids Research 30, e111.
Spigelman, M., Lemma, E., 1993. The use of the polymerase chain reaction (PCR) to detect Mycobacterium tuberculosis in ancient skeletons. International Journal of Osteoarchaeology 3, 137-143.

Stone, A.C., Wilbur, A.K., Buikstra, J.E., Roberts, C.A., 2009. Tuberculosis and leprosy in perspective. American Journal of Physical Anthropology 140 Suppl 49, 66-94.
Süss, B., Flekna, G., Wagner, M., Hein, I., 2009. Studying the effect of single mismatches in primer and probe binding regions on amplification curves and quantification in real-time PCR. J Microbiol Methods 76, 316-319.

Taylor, G.M., Goyal, M., Legge, A.J., Shaw, R.J., Young, D., 1999. Genotypic analysis of Mycobacterium tuberculosis from medieval human remains. Microbiology 145 (Pt 4), 899-904.
Thierry, D., Cave, M.D., Eisenach, K.D., Crawford, J.T., Bates, J.H., Gicquel, B., Guesdon, J.L., 1990. IS6110, an IS-like element of Mycobacterium tuberculosis complex. Nucleic Acids Research 18, 188.

Tumpey, T.M., Basler, C.F., Aguilar, P.V., Zeng, H., Solorzano, A., Swayne, D.E., Cox, N.J., Katz, J.M.,

Taubenberger, J.K., Palese, P., Garcia-Sastre, A., 2005. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. Science 310, 77-80.
van Soolingen, D., de Haas, P.E., Hermans, P.W., Groenen, P.M., van Embden, J.D., 1993. Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of Mycobacterium tuberculosis. J Clin Microbiol 31, 1987-1995.

Wilbur, A.K., Bouwman, A.S., Stone, A.C., Roberts, C.A., Pfister, L.-A., Buikstra, J.E., Brown, T.A., 2009. Deficiencies and challenges in the study of ancient tuberculosis DNA. Journal of Archaeological Science 36, 1990-1997.

Worobey, M., Gemmel, M., Teuwen, D.E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J.J., Kabongo, J.M., Kalengayi, R.M., Van Marck, E., Gilbert, M.T., Wolinsky, S.M., 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature 455, 661-664.

Wrischnik, L.A., Higuchi, R.G., Stoneking, M., Erlich, H.A., Arnheim, N., Wilson, A.C., 1987. Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. Nucleic Acids Research 15, 529-542.

Yang, D.Y., Eng, B., Waye, J.S., Dudar, J.C., Saunders, S.R., 1998. Technical note: improved DNA extraction from ancient bones using silica-based spin columns. American Journal of Physical Anthropology 105, 539-543.

Figure 1: Map of samples positive for MTBC loci in the qPCR screening

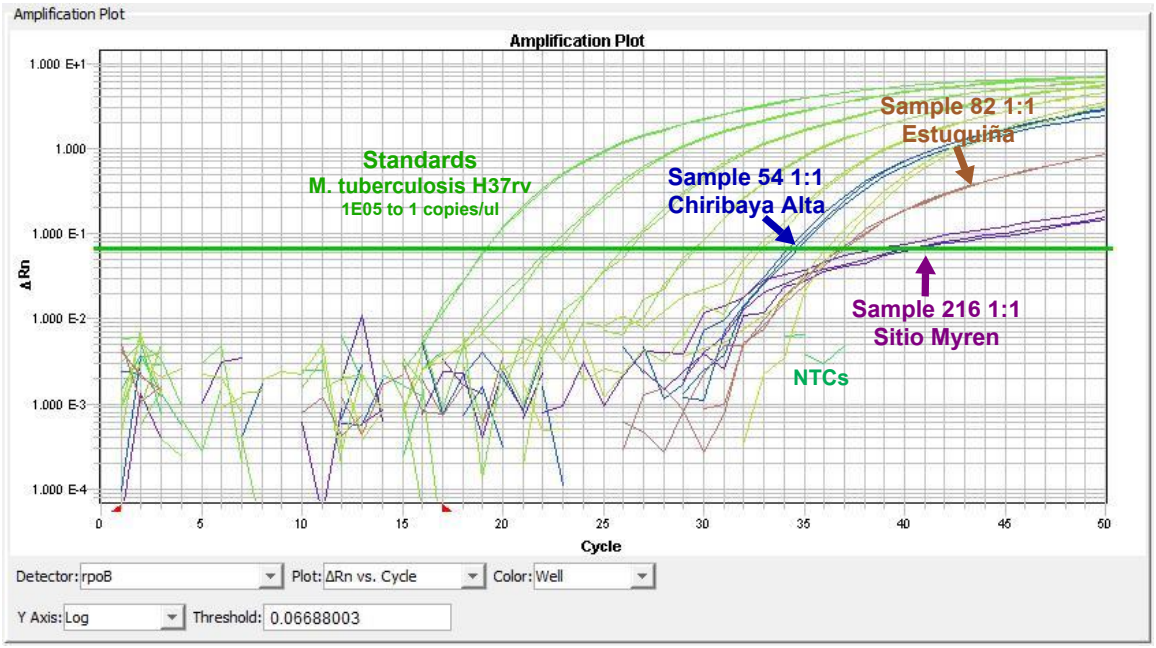


Figure 2: Amplification plot of standards and positive samples. (Note that Sample 54 was successful for whole genome capture, see Chapter 3).

Primer/Probe	*Final concentration*	*Sequence (5'-3')*	*Amplicon size (bp)*
IS1081_F	900 nM	*GCACTCCATCTACGACCAG*	
IS1081_Probe	300 nM	*6FAM-ATTGGGCAACAACTGATTCGGCGTCG-TAMRA*	*84*
IS1081_R	900 nM	*GGGAGTTTGTCGGTCAGAG*	
rpoB2_F	300 nM	*CAACGTCGAGGTGCTATCG*	
rpob2_Probe	300 nM	*FAM-TCGCCGCACCGTCACT-NFQMGB*	*70*
rpoB2_R	300 nM	*CTCCAGGTCCTCGTCCTCA*	

Table 1: primers and Taqman® probes used in the analysis

Site	*N*	*Dates*	*Location*
OLD WORLD			
Arena Candida	1	3000-4000 BC	Italy
Arma dell Aquila	1	3850+-90 BC	Italy
St Mary Spital	16	AD 1200-1539	England
Ryukyu (Okinawa)	1	AD 1600-1800	Japan
Russian Academy of Science	1	Early Medieval	Russia
NEW WORLD (SOUTH AMERICA)			
Moquegua (Chen Chen)	10	AD 500-700	Moquegua, Peru
Las Delicias y Candeleria	1	AD 770-990	Colombia
Lambayeque	2	AD 900-1750	Peru
Moquegua (Estuquina)	14	AD 1000-1476	Moquegua, Peru
Yaral	2	AD 900 - 1200	Osmore Valley, Peru
Chiribaya Alta	4	AD 772-1350	Ilo, Peru
Algodonal	1	AD 1031-1179	Ilo, Peru
Santa Cruz	1	AD 1200	Patagonia, Argentina
Punta Arenas, Sitio Myren-1	1	AD 1310+-20	Argentina
Purucucho	3	AD 1470-1540	Peru
Misión Salesiana	15	AD 1870-1930	Tierra del Fuego
NEW WORLD (NORTH AMERICA)			
Yokem	1	AD 1039-1401	Illinois, USA
Schild	1	AD 1040-1407	Illinois, USA
Norris Farms	5	AD 1100-1250	Illinois, USA
Orendorf	2	AD 1150-1300	Illinois, USA
Uxbridge	19	AD 1410-1570	Ontario, Canada
Arikara, Mobridge	2	AD 1500-1750	South Dakota, USA
Arikara, Sully School Village	7	AD 1550-1700	South Dakota, USA
Arikara, Leavenworth	1	AD 1725-1775	South Dakota, USA
Arikara, Leavitt	1	AD 1725-1775	South Dakota, USA
Arikara, Cheyenne River village	6	AD 1750-1775	South Dakota, USA
Chirikov	2	AD 1750-1899	Alaska, USA
San Cristobal	2	AD 1450-1680 (contact AD 1598)	New Mexico, USA
Highland Park	4	AD 1847-1850	New York, USA

Table 2: Site information for samples used in this analysis

Archaeological Site Name	*Individual*	*Lab Sample Name*	*No. loci pos*	*rpoB1*	*rpoB 2*	*IS6110*	*IS1081*	*Is Inhibited*
Samples positive for multiple loci								
Arikara, Mobridge	382961	AD15*	4/4	+	+	+	+	No
Chiribaya Alta, Cemetery 1	2069	AD54*	4/4	+	+	+	+	
El Algodonol (AL)	386-1	AD58*	4/4	+	+	+	+	No
Las Delicias	LD-90-1X-	AD281	3/4	-	+	+	+	
Estuquina	4165	AD82*	4/4	+	+	+	+	Yes
Highland Park	141a	AD128	4/4	+	?	+	+	
Yaral-2 (Cemetery	10291	AD64	3/3	+	nt	+	+	
Samples positive for one MTBC-specific locus								
Highland Park	138a	AD127	2/4	+	+	-	-	
Norris Farms	228	AD160	1/4	-	+	-	-	
Samples positive only for rpoB1								
Punta Arenas, Sitio Myren-7	TDF 30477	AD216	1/4	+	?	-	-	Yes
Okinawa Island,		AD224	1/4	+	-	-	-	
Arikara, Cheyenne River village (39STI)	382713	AD11*	1/4	+	-	-	-	
Arikara, Cheyenne River village (39STI)	382669	AD6*	1/4	+	-	-	-	
Arikara,	325341	AD1*	1/4	+	-	-		
Arikara, Sully School Village (39SL4)	388120	AD19*	1/3	+	nt	-	-	Yes
Misión Salesiana	D-E 14	AD256	1/4	+	-	-	-	
Misión Salesiana	E 15-16	AD257	1/4	+	-	-	-	Maybe
Misión Salesiana	C 15-16	AD263	1/4	+	-	-	-	
Misión Salesiana	D 14	AD264	1/4	+	-	-		
Misión Salesiana	C 15	AD265	1/3	+	nt	-	-	
Estuquina	2279b	AD76*	1/4	+	-	-	-	
Estuquina	3215	AD79	1/3	+	nt	-	-	
Estuquina	5859	AD87	1/3	+	nt	-	-	
Estuquina	99407a	AD89	1/4	+	-	-	-	
Norris Farms	21	AD145	1/4	+	-	-	-	
Lambayeque, Peru	ILL-22	AD227	1/4	+	-	-	-	
Highland Park	225a	AD134	1/4	+	-	-	-	
Highland Park	16a	AD114	1/4	+	-	-	-	

Table 3: Results of qPCR assays indicating number of independently positive extractions per sample. ?=inconclusive, nt= not tested. *Independent replication: Separate positive extraction repeated by different personnel, or confirmed by enrichment for MTBC DNA (Chapter 3). Samples not listed were negative for all assays (Tables S2-5).

Table S1 – Appendix A
Table S2 – Appendix A
Table S3 – Appendix A
Table S4 – Appendix A
Table S5 – Appendix A

CHAPTER 3

PRE-COLUMBIAN MYCOBACTERIAL GENOMES REVEAL SEALS AS A SOURCE OF NEW WORLD HUMAN TUBERCULOSIS

Kirsten I. Bos[1]*ł, **Kelly M. Harkins[2]***, Alexander Herbig[1,3]*, Mireia Coscolla[4,5]*, Nico Weber[3], Iñaki Comas[6,7], Stephen A. Forrest[1], Josephine M. Bryant[8], Simon R. Harris[8], Verena J. Schuenemann[1], Tessa J. Campbell[9], Kerrtu Majander[1], Alicia K. Wilbur[2], Ricardo A. Guichon[10], Dawnie L. Wolfe Steadman[11], Della Collins Cook[12], Stefan Niemann[13,14], Marcel A. Behr[15], Martin Zumarraga[16], Ricardo Bastida[17], Daniel Huson[3], Kay Nieselt[3], Douglas Young[18,19], Julian Parkhill[8], Jane E. Buikstra[2], Sebastien Gagneux[4,5], Anne C. Stone[2]ł, and Johannes Krause[1,20]ł.

[1]Department of Archeological Sciences, University of Tübingen
[2]School of Human Evolution and Social Change, Arizona State University
[3]Center for Bioinformatics, University of Tübingen, Tübingen, Germany
[4]Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute
[5]University of Basel, Basel Switzerland
[6]Genomics and Health Unit, FISABIO-Public Health, Valencia, Spain
[7]CIBER (Centros de Investigación Biomédica en Red) in Epidemiology and Public Health, Spain
[8]Pathogen Genomics, The Wellcome Trust Sanger Institute
[9]Department of Archaeology, University of Cape Town
[10]ONICET, Laboratorio de Ecología Evolutiva Humana (FACSO, UNCPBA); Dpto. Biología (FCEyN,UNMDP)
[11]Department of Anthropology, University of Tennessee
[12]Department of Anthropology, Indiana University
[13]Molecular Mycobacteriology, Forschungszentrum Borstel, Borstel, Germany
[14]German Center for Infection Research, Borstel Site, Borstel, Germany
[15]McGill International TB Centre, McGill University
[16]Biotechnology Institute, CICVyA-INTA, Castelar, Buenos Aires, Argentina
[17]Instituto de Investigaciones Marinas y Costeras (CONICET-UNMdP) and Facultad de Ciencias Exactas y Naturales; Universidad Nacional de Mar del Plata, Argentina
[18]Imperial College London
[19]MRC National Institute for Medical Research, London, UK
[20]Senckenberg Centre for Human Evolution and Palaeoenvironment, University of Tübingen, Tübingen, Germany

*** these authors contributed equally to this work**
ł to whom correspondence should be addressed: Kirsten I. Bos (kirsten.bos@ifu.uni-tuebingen.de), Anne C. Stone (acstone@asu.edu) or Johannes Krause (johannes.krause@uni-tuebingen.de)

Author contributions:

ACS, JEB, JK, KIB, and KMH conceived of the investigation. JK, KIB, ACS, SAF, NW, and AKW designed experiments. JP, RAG, DLWS, DCC, SN, MAB, MZ, and RB provided samples for analysis. KIB, KMH, VJS, TJC, and AKW performed laboratory work. AH, JK, SG, MC, NW, KIB, IC, DY, JP, JMB, SRH, DH, KN, ACS, KMH, JEB, TJC, DCC, and DLWS performed analyses. KIB wrote the manuscript with contributions from all coauthors.

Summary paragraph

Modern strains of *Mycobacterium tuberculosis* from the Americas are most closely related to those common in Europe, a result that supports the assumption that human tuberculosis was introduced post-contact[1]. This notion, however, is incompatible with the evidence of pre-contact tuberculosis in the New World identified by diagnostic morphological lesions[2]. Comparative genomics of modern isolates suggests that *M. tuberculosis* attained its worldwide distribution following human dispersals out of Africa during the Pleistocene[3], though ancient genomes would be highly informative for evaluating the suitability of this model. Here we present three approximately 1000 year old ancient mycobacterial genomes from human remains found in Peru, providing unequivocal evidence that a member of the *M. tuberculosis* complex caused human disease in the pre-contact New World. The ancient strains are distinct from any known human-adapted *M. tuberculosis* form and are most closely related to strains adapted to seals and sea lions. Two independent dating approaches suggest a most recent common ancestor for the *M. tuberculosis* complex less than 6000 years ago, which supports a dispersal of the disease during the Holocene. Our results implicate sea mammals as having played a role in transmitting the disease to human populations across the ocean.

Mycobacterium tuberculosis has had a long history with humans, though consensus has not been reached on when this interaction began[1,3,4]. Previous models held that the human-adapted pathogen evolved from a zoonotic transfer of *M. bovis* following animal domestication during the Neolithic[5]. Comparative genomic analyses, however, suggest that the bovine form and those adapted to other animal hosts are in fact derived from human strains[3,6]. This supports a rather different disease history where humans may have been the most susceptible host species for early progenitors of strains currently circulating. Today the majority of *M. tuberculosis* diversity exists in Africa[7], implying that the pathogen likely originated from a monoclonal expansion therein and achieved its worldwide distribution via human movements[3,4,9]. The observations that *M. tuberculosis* strains tend to be associated with human populations[9] and that selection in the bacterium exists at loci associated with host immune responses[10] indicate that host and pathogen have had sufficient time to co-evolve. Dating approaches that use human demographic events for calibration generate coalescence estimates for all *M. tuberculosis* lineages, collectively referred to as the *M. tuberculosis* complex (MTBC), that differ by over an order of magnitude based on mutation rates that are derived from different models.

Given the pathogen's phylogeography, current models are unable to explain the abundant archaeological evidence for the presence of tuberculosis in the Americas prior to European contact. Strains currently circulating in the Americas group with those of European origin, and this has been used to support a European dissemination from either early settlement or trade associations[1,8]. This model, however, is incompatible with bioarchaeological data indicating the presence of tuberculosis in the pre-contact New World[2] (see Supplementary Information). Molecular investigations using ancient pre-Columbian material have identified short conserved regions of mobile elements considered to be diagnostic for tuberculosis, though these markers offer no information about phylogenetic placement, and are thus difficult to authenticate as ancient[11]. While a Pleistocene dispersal following human movements out of Africa could explain its presence in the pre-contact New World[3,4], the dominance of European-derived lineages in the

Americas today makes this difficult to reconcile without data to support a complete strain replacement within the last 500 years.

Genomic reconstructions of ancient pathogens provide robust evidence of DNA authenticity and permit genome-level comparisons[12]. The success of DNA capture[13] and genomic assembly of an historical MTBC via metagenomic sequencing[14] implies that DNA preservation of this pathogen may be adequate to address outstanding evolutionary questions requiring use of archaeological material. Here we apply these techniques to demonstrate that a previously uncharacterised member of the MTBC caused human infection in the Americas prior to European contact.

We screened 68 skeletal samples representing New World pre- and post-contact sites (Table S1). All individuals showed skeletal indicators associated with tuberculosis infections. Samples were processed via established protocols and were screened for *M. tuberculosis* DNA by an in-solution capture assay designed for the *rpoB, gyrA, gyrB, katG*, and *mpt40* genes (Table S2). Capture products for samples and negative controls were sequenced on an Illumina MiSeq and mapped to the corresponding regions in the *M. tuberculosis* H37Rv reference genome (NC_000962.2). No tuberculosis fragments were found in our negative controls. Only three of the 68 samples, referred to here as samples 54, 58, and 64, showed convincing preservation of tuberculosis DNA (see Supplementary Information, Extended Data Figure 1, and Table S1): all three samples were recovered from excavations in Peru and derive from Chiribaya cultures associated with the Middle Horizon/Late Intermediate Period AD 750 – AD 1350) (Figure 1), and radiocarbon dates ranging from AD 1028 – AD 1280 (at ≥98.5% probability) (Table 1, Table S3) confirm that they predate European contact. DNA damage spectra displayed a pattern expected of ancient molecules[15]. For comparison, non-enriched libraries were sequenced on an Illumina MiSeq producing 32,687 to 182,908 reads for each sample, of which 4.6% to 1.6% mapped to the human genome (hg19). In contrast, a maximum of only 1.8% of the reads mapped to the *M. tuberculosis* reference genomes (Table S1), indicating that DNA capture would be necessary for genome retrieval.

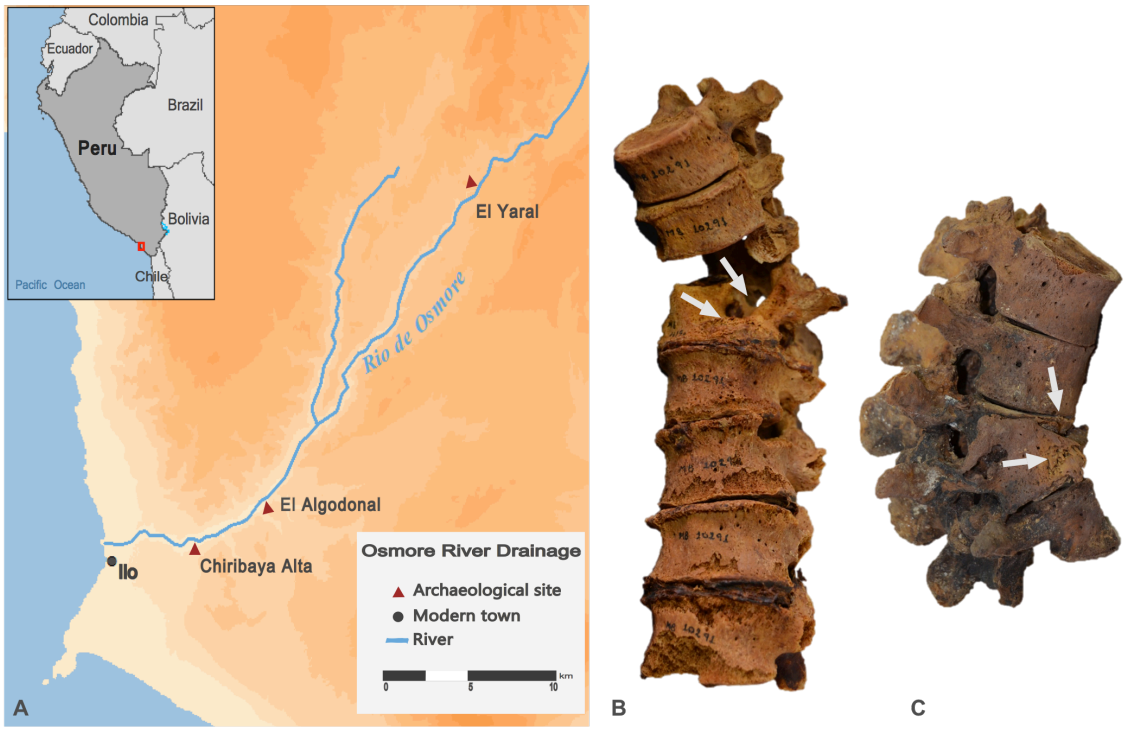


Figure 1 – A) Map of Peru showing locations of the three archaeological sites, Peru Elevation (n.d.) CGIAR SRTM 90 M Digital Elevation Database v4.1 (*http://srtm.csi.cgiar.org*). B and C) Skeletal lesions from two of the individuals positive for ancient *M. tuberculosis* DNA, showing evidence suggestive of active tuberculosis infections (B = individual 58, C = individual 64). Arrows point to vertebral lesions, collapse, fusion, and kyphosis.

DNA libraries treated with uracil DNA glycosylase were generated to remove and repair damaged nucleotides, and were subsequently used for full genome hybridization capture (Agilent). Array probes were designed to accommodate genetic diversity in the MTBC, as well as portions of the *M. avium* and *M. kansasii* genomes (Table S4). Enriched products were sequenced on one lane of an Illumina HiSeq 2000. For comparison against a larger dataset of 259 modern MTBC genomes including the outgroup *M. canettii*[3], all ancient reads were mapped against a computationally constructed ancestor for the MTBC[10]. The recently published genome from an 18th century Hungarian mummy[14], as well as 14 animal strains from the *M. caprae*, *M. microti*, and *M. pinnipedii* lineages were added, along with a strain recently isolated from a wild chimpanzee[16]. Standard mapping resulted in heterozygous positions for the mummy, Peruvian

samples 54 and 64, and all modern samples (Extended Data Figure 2). Increased mapping stringency removed many heterozygous positions for the Peruvian samples, suggesting they derived from non-tuberculosis reads; however, the Hungarian mummy and 8 modern samples still displayed heterozygosity consistent with mixed strains[14] (Extended Data Figure 3). Our more stringent mapping reduced overall genomic coverage for all samples. The final dataset thus consisted of 262 genomes with a minimum of 75% coverage, four of which were ancient (Table S5). A minimum of 20-fold average coverage was obtained for each of the Peruvian genomes (Table 1, Figure 2, Table S6), implying a 40- to 120-fold enrichment (Table S6).

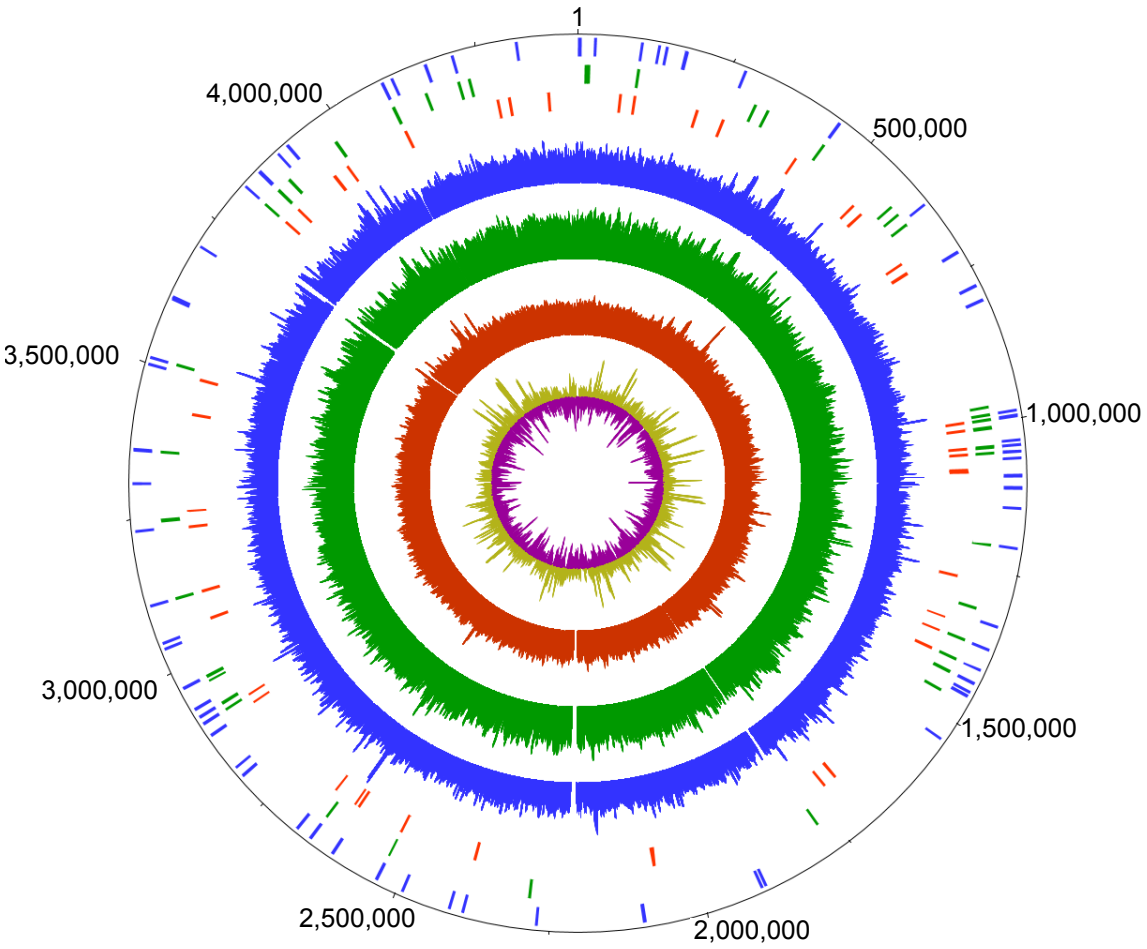


Figure 2 – Coverage plots for three ancient genomes. GC content is shown in the inner ring, with purple representing AT content and gold GC content. Coverage rings for samples 64, 58, and 54 are shown in red, green, and blue, respectively. Locations of SNPs that differ from the hypothetical ancestral sequence and that are specific for the ancient strains are shown by vertical lines in the corresponding colours. SNPs shown here were identified before exclusion of positions with missing data from SNP calls from the 262 full genome dataset, hence not all of them are considered in phylogenetic reconstruction.

Single nucleotide polymorphism (SNP) analyses were performed by comparing all genomes against the constructed ancestor. This identified 53,177 SNPs for the entire dataset, which ranged between 489 and 1415 per genome (Table S8). As input for phylogenetic assessments we used an alignment of 22,480 variable positions after removing all positions with missing data. Tree reconstructions revealed that our Peruvian genomes do not cluster with other human strains, but rather are more closely related to the animal lineage (Extended Data Figures 4-6), sharing 76 SNPs with modern *M. pinnipedii* strains (Figure 3). Genomic architecture revealed a region of difference (RD) deletion pattern common to all animal lineages (Table S7), as well as absence of the microti-specific *RDmic* and presence of the seal-specific *RDseal*. To our knowledge, *M. pinnipedii* strains have been isolated only from seal species restricted to the southern hemisphere[17]. Here they were harvested from both captive and wild animals from South America and Australia. The three ancient strains share 5 SNPs, all of which are non-synonymous (Table S9): this indicates that these strains derive from a common progenitor, with subsequent accumulation of 10 to 23 substitutions along the three strain-specific branches. To investigate possible signals of adaptation, we screened these 5 shared SNPs for putative functional effects.

Our computational analysis predicted a functional impact of the *P44L* mutation in *Rv2258c*, encoding a methyltransferase involved in ubiquinone metabolism (Table S9). The SNP in the *ctpA* gene at codon 62 (D62N) was not predicted to have a functional impact; however, we identified two other non-synonymous SNPs (D62G and D62E), also not predicted to have functional impacts, in the same codon of *ctpA* at different positions, each in a lineage 4 modern strain. The occurrence of homoplasies is uncommon in the MTBC, and therefore potentially indicates positive selective pressure[10]. A site-wise analysis of positive selection on codon 62 of *ctpA* confirmed that all three SNPs may be under diversifying selection (Table S10). The *ctp* genes encode efflux ion pumps that are thought to prevent metal accumulation in the bacterium[18], hence adaptation may relate to host metal ion availability. This notion is supported by the existence of homoplasies in other genes of the efflux pump family in modern MTBC strains (Table S9).

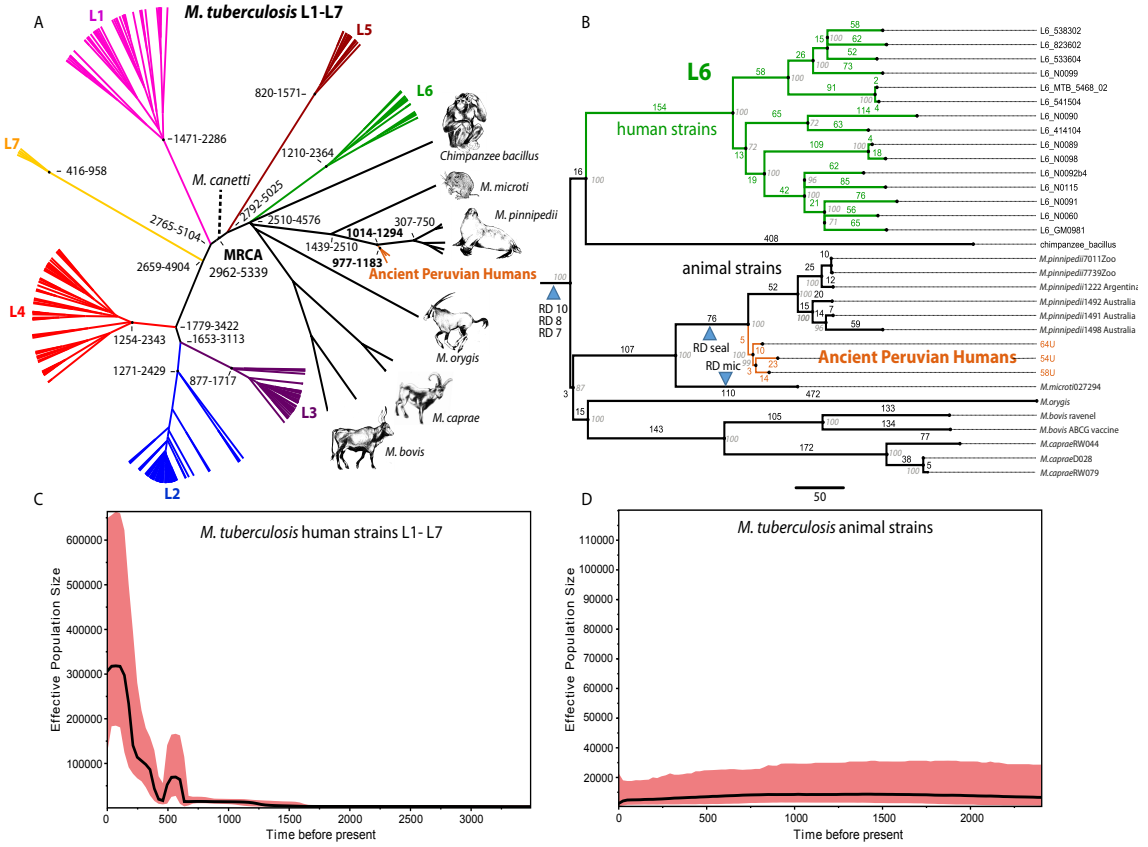


Figure 3 – A) Bayesian maximum clade credibility tree of 261 MTBC genomes (which excludes the Hungarian mummy sample) showing relative placement of modern and ancient strains, with estimated divergence dates shown in YBP using a model of population expansion. B) Maximum parsimony tree for the lineage 6 and animal-adapted MTBC genomes, and SNPs that define all branches. Bootstrap values are shown in grey italics. Deletions specific to the animal lineages are shown as triangles. C) Bayesian skyline plot showing estimated effective population sizes for the human lineages. D) Bayesian skyline plot showing estimated effective population sizes for the animal lineages.

Bayesian dating analysis was performed using radiocarbon dates as tip calibration (Table 1). The Hungarian mummy sample was excluded due to the presence of multiple strains. A clock test rejected the molecular clock for all 258 modern genomes (p = 5E-147) (Extended Data Figure 7). Dating analysis using a relaxed clock model and a constant population size generated a mutation rate of 4.6 $x10^{-8}$ substitutions per site per year (3 $x10^{-8}$ – 6.2 $x10^{-8}$ 95% HPD). Bayesian skyline plots revealed constant population sizes for the animal strains and clear indications of expansions in the human adapted lineages (Figure 3). Using an expansion model had a negligible influence on the mutation rate, generating 4.9 $x10^{-8}$ substitutions/site/year (3.4 $x10^{-8}$ – 6.4 $x10^{-8}$ 95% HPD),

which corresponds to 0.20 and 0.21 substitutions/genome/year for a constant and expanding population model, respectively. This rate agrees well with estimates for MTBC evolution in modern epidemiological contexts[20], and is more than ten-fold faster than those using human out of Africa dispersals to calibrate the molecular clock[3]. Our mutation rates date the most recent common ancestor (MRCA) for the MTBC (excluding *M.canettii*) at 4,449 YBP (2,990 – 6,062 YBP 95% HPD) and 4,064 YBP (2,951 – 5,339 YBP 95% HPD) for constant size and expansion models, respectively (Extended Data Figure 8). This dating was corroborated via an independent analysis using the sequences from the Hungarian mummy[14] sample as single calibration point. We separated the individual variants of the two mummy strains by reconstructing them onto the MTBC Lineage 4 phylogeny (Extended Data Figure 9). Lengths from the terminal branches were estimated by using the number of heterozygous variants not present in the modern strains, under the assumption that both isolates are equidistant from the root of the tree. The year of death 1797 and estimated ages for the penultimate nodes were used as priors for Bayesian phylogenetic reconstruction. Using only synonymous variants, a relaxed clock, and constant population size, we estimate the age of the MRCA of the MTBC (excluding *M. canettii*) to be 5268.5 YBP (2689.6 - 8417.7 95% HPD) with a synonymous substitution rate of 7.07 $x10^{-8}$ ($3.70x10^{-8}$ – 1.12 $x10^{-7}$ 95% HPD) per site (Extended Data Figure 10, Table S11). This higher substitution rate may be due to lower selective pressure on synonymous sites.

Our results provide unequivocal evidence of human infection caused by members of the MTBC in pre-Columbian South America. Our MRCA, which is at least an order of magnitude younger than previous estimates[3,4], presented us with a challenge to explain how a mammalian pathogen could have reached human populations in the Americas ca. 9000 years after inundation of the Bering Land Bridge[21]. The fact that our ancient genomes share a common ancestor with strains that are restricted to seals and sea lions[17] provides a plausible, if unexpected, route of entry into the New World: within the last 2500 years pinnipeds likely contracted the disease from an African host species, carried the disease across ocean waters, and exploitation of marine mammals amongst coastal peoples of South America facilitated a zoonotic transfer of the bacterium within the first

millennium AD. This parallels similar zoonoses of marine parasites acquired from seal consumption amongst archaeological coastal populations[22] (Supplementary Information).

Due to the abundance of publications reporting morphological evidence of pre-Columbian tuberculosis in the region, the coasts of Peru and northern Chile have long been recognised in the archaeological literature as locations where tuberculosis first came into view in the New World[2]. Some have even suggested marine mammals as a potential source of the infection[23]. The three individuals considered here show pathological changes consistent with either pulmonary or disseminated tuberculosis, so a non-contagious infection acquired from consumption of contaminated animal products in each case cannot be ruled out. In the absence of these data, however, the five unique derived positions shared by the ancient Peruvian genomes may provide preliminary evidence of host specificity. All three genomes share a common ancestor that pre-dates the radiocarbon age of our skeletal material by more than 100 years, and two SNPs show potential signals of adaptation. These observations could support a single zoonotic transfer from pinnipeds to humans between AD700 and AD1000 (Figure 3). Subsequent host adaptation and dissemination is a compelling prospect for future work. If confirmed, this would constitute the first example of a zoonotic transfer followed by re-adaptation to the human host in the MTBC.

Such a model could explain the abundance of tuberculosis-like lesions in the region that accumulate beginning approximately AD700[24,25]. The later appearance of similar skeletal lesions in North America that first appear ca. AD900 is consistent with either a transcontinental spread of the pathogen via established trade routes[26] or a later independent introduction of tuberculosis from a different source. The lack of representation of this or any other American-specific strain in modern groups supports replacement by a European strain after contact[1] that quickly moved through indigenous populations on account of additional adverse factors such as social marginalization, food insecurity, and potentially facilitative co-circulating infections that reached epidemic levels, such as those recorded in northern North America during the decline of the fur trade[27]. Our data also indicate a subsequent introduction of *M. pinnipedii* to Australian seal

colonies within the last 700 years (Figure 3); the potential for similar zoonotic transfers, therefore, exist in Oceanian populations, though lesions suggestive of tuberculosis have not been identified in relevant skeletal material[2].

M. pinnipedii has caused infection in several mammalian host species, including humans, in the context of zoo outbreaks[28]. Further sampling of animal-adapted MTBC from both modern and ancient contexts will be of great value to determine its range of potential host species and clarify directions of transmission. While a human transfer of the bacterium to marine mammals cannot be ruled out from our data, we consider this extremely unlikely: humans did not herd or farm seals, and close regular contacts would be required for anthroponotic transmission, as is observed in domestic cattle[29].

The above assertion of an introduction of MTBC via pinnipeds followed by human adaptation and subsequent transmission throughout the Americas can only be confirmed via comparison with additional North and South American pre-Columbian MTBC genomes from non-coastal groups, which remain elusive despite our inclusion of suitable material in our screening (Table S1). In addition, our dating analyses are based on two independent approaches, though each relies on (effectively) a single calibration point. Mutation rate heterogeneity is documented in other clonal pathogens[30], and the rejection of our molecular clock indicates that MTBC evolution is not constant amongst lineages. Additional calibration points from ancient MTBC lineages around the world will be essential to evaluate the legitimacy of our proposed models. Such caveats are of paramount importance considering the many investigations that report on members of the MTBC identified in skeletal samples that predate our inferred MRCA, or American material from periods that predate our proposed time of MTBC entry. Such claims could only be reconciled with what we propose here if 1) rate heterogeneity or horizontal gene transfer is obscuring our dating analysis, perhaps as a result of human population expansions which increase the availability of susceptible hosts and allow selection to operate more quickly, 2) the pathogens identified in the earlier archaeological material are in fact not members of the MTBC, but rather are ancestral

forms that have since undergone replacements, or 3) certain techniques for MTBC identification in archeological material lack specificity.

Acknowledgements

We would like to thank the following people and institutions for assistance and/or permission for sampling: Museo Contisuyu, Centro Mallqui, Sonia Guillen, Instituto Nacional de Cultura, Peru, Guillermo Cock, Cathy Gaither, Melissa Murphy, Maria Cecilia Lozada, Shelley Burgess, Deborah Blom, Bruce Owen, Lic. Antonio Oquiche Hernani, Lic. Patricia Palacios Filinich, Sloan Williams, Bertha Vargas, Don Rice, and Haagen Klaus, Susan Pfieffer, the University of Toronto, the Upper Mississippi Valley Archaeological Research Foundation (UMVARF), Larry Conrad, Indiana University, George Milner, the Pennsylvania State University, the American Museum of Natural History, Ann Stodder, Bob Brier, Ian Tattersall, Ken Mowbray, the National Museum of Natural History (Smithsonian), Bill Billeck, the Smithsonian Institution, Noreen Tuross, Rochester Museum & Science Center, Lorraine P. Saunders, Christopher Grivas, Genevieve Housman, and Maria Nieves-Colon. We thank Hendrik Poinar for helpful discussions regarding capture regions for *M. tuberculosis* screening. We thank Brian Coombes and Ben Krause-Kyora for providing modern tuberculosis for bait manufacture. The Huron Wendat Nation is aware of the sampling of Uxbridge bone and is a recipient of information from this study. We would like to acknowledge the following sources of funding: European Research Council (ERC) starting grant APGREID (to JK), The National Science Foundation (ACS and JEB) for NSF BCS-1063939, NSF-REU BCS-0612222 and NSF BCS-0612222, the George E. Burch Fellow in Theoretic Medicine and Affiliated Sciences at the Smithsonian Institution (2003-2007, to JEB), Social Sciences and Humanities Research Council of Canada postdoctoral fellowship grant 756-2011-501 (to KIB), the NSF Graduate Research Fellowship DGE-1311230 and Jacob K. Javits Fellowship to KMH, Ramón y Cajal spanish research grant RYC-2012-10627 (to IC), Swiss National Science Foundation (PP0033_119205 to SG), the National Institutes of Health (AI090928 to SG), the European Research Council (309540 to SG), PICT0575 Argentina to RAG, the Wadsworth

Fellowship from the Wenner-Gren Foundation (to TJC), the Wellcome Trust (098051 to JP) and funding from MRC (to JMB).

Raw sequencing data have been deposited in the NCBI Sequence Read Archive (SRA) under numbers SRP041177 for the ancient Peruvian samples and SRP041181 for the *M. pinnipedii* strains.

Competing financial interests:

The authors declare no competing financial interests.

Site	Sample number	sample type	Dates cal AD (probability)	Number of reads from whole *M. tuberculosis* genome capture	Percent of reads mapping to *M. tuberculosis* hypothetical ancestor	Number of mapped reads post duplicate removal	Average fold coverage
Chiribaya Alta	54	Bone (L4)	AD 1028-1224 (99.7%)	30,080,214	37.02%	2,980,807	31.35
El Yaral	64	Bone (T12)	AD 1141-1280 (98.5%)	5,999,216	29.16%	1,944,666	20.45
El Algodonal	58	Bone (L5)	AD 1023-1211 (99.7%)	8,683,999	17.20%	2,479,743	22.73

Table 1 – Radiocarbon dates and mapping statistics for samples used in whole genome capture.

References Cited

1. Pepperell, C.S., *et al*. The Role of Selection in Shaping Diversity of Natural *M. tuberculosis* Populations. *PLOS Path* **9**(8), e1003543 (2013).

2. Roberts, C., and Buikstra, J.E. The Bioarchaeology of Tuberculosis. A Global View on a Reemerging Disease. Gainesville: University Press of Florida (2003).

3. Comas I., *et al*. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat Gen* **45**, 1176-1182 (2013).

4. Wirth, T., *et al*. Origin, Spread, and Demography of the *Mycobacterium tuberculosis* Complex. *PLOS Path* **4**(9), e1000160 (2008).

5. Cockburn A. The Evolution and Eradication of Infectious Diseases. Baltimore: Johns Hopkins Press (1963).

6. Brosch, R., *et al*. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Nat Acad Sci* **99**(6), 3684-3689 (2002).

7. Gagneux, S. and Small, P.M. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infect Dis* **7**, 328-337 (2007).

8. Hershberg, R., *et al*. High Functional Diversity in *Mycobacterium tuberculosis* Driven by Genetic Drift and Human Demography. *PLOS Biol* **6**(2), e311 (2008).

9. Gagneux, S., *et al*. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *Proc Nat Acad Sci* **103**(8), 2869-2873 (2006).

10. Comas, I., *et al*. Human T-cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat Genet* **42**(6), 498-503 (2010).

11. Shapiro, B. and Gilbert, M.P.T. No proof that typhoid fever caused the Plague of Athens (a reply to Papagrigorakis et al.). *Int J Infect Dis* **10**, 334-340 (2006).

12. Bos, K.I., *et al*. A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature* **478**, 506-510 (2011).

13. Bouwman, A., *et al*. Genotype of an historic strain of *Mycobacterium tuberculosis*. *Proc Nat Acad Sci* **109**(45), 8511-18516 (2012).

14. Chan, J.Z.-M., *et al.* Metagenomic Analysis of Tuberculosis in a Mummy. *New Engl J Med* **369**, 3 (2013).

15. Briggs, A. W. et al. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Nat Acad Sci* **104**(37), 14616-14621 (2007).

16. Coscolla, M., *et al*. Novel *Mycobacterium tuberculosis* Complex from a Wild Chimpanzee. *Emerg Infect Dis* **19**(6), 969-976 (2013).

17. Bastida, R., *et al*. Tuberculosis in a Wild Subantarctic Fur Seal from Argentina. *J Wildlife Dis* **35**(4), 796-798 (1999).

18. Botella, H., *et al*. Metallobiology of host-pathogen interactions: an intoxicating new insight. *Trends Micorbiol*, **20**(3), 106–12. doi:10.1016/j.tim.2012.01.005 (2012).

19. Comas, I., *et al*. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet* **18**(1), 106-10 (2012).

20. Bryant, J. M., *et al*. Inferring patient to patient transmission of *Mycobacterium tuberculosis* from whole genome sequencing data. *BMC infectious diseases* **13**(1), 110 (2013).

21. Pitblado, B.L. A Tale of Two Migrations: Reconciling Recent Biological and Archaeological Evidence for the Pleistocene Peopling of the Americas. *J Archaeol Res* doi:10.1007/s10814-011-9049-y (2011).

22. Patrucco, R., *et al.* Parasitological Studies of Coprolites of Pre-Hispanic Peruvian Populations. *Curr Anthropol* **24**(3), 393-394 (1983).

23. Bastida, R., *et al*. La tuberculosis en grupos de cazadores recolectores de Patagonia y Tierra del Fuego: nuevas alternativas de contagio a través de la fauna silvestre. Revista Argentina de Antropología Biológica **13**(1), 83-95 (2011).

24. Salo, W., *et al*. Identification of *Mycobacterium tuberculosis* DNA in a pre-Columbian Peruvian mummy. *Proc Nat Acad Sci* **91**, 2091-2094 (1994).

25. Arriaza, R.T., *et al*. Pre-Columbian tuberculosis in Northern Chile: Molecular and Skeletal Evidence. *Am J Phys Anthropol* **98**, 37-45 (1995).

26. Anawalt, P.R. Cultural Contacts between Ecuador, West Mexico, and the American Southwest: Clothing Similarities. *Lat Am Antiq* **3**(2), 114-129 (1992).

27. Herring, D.A., and Sattenspiel, L. Social Contexts, Syndemics, and Infectious Disease in Northern Aboriginal Populations. *Am J Hum Biol* **19**, 190-202 (2007).

28. Jurczynski, K., *et al*. Pinniped Tuberculosis in Malayan Tarpirs (*Tapirus indicus*) and its Transmission to Other Terrestrial Mammals. *J Zoo Wildlife Med* **42**(2), 222-227 (2011).

29. Berg, S. *et al*. The Burden of Mycobacterial Disease in Ethiopian Cattle: Implications for Public Health. *PLOS ONE* DOI: 10.1371/journal.pone.0006068 (2009).

30. Cui, Y., *et al*. Historical variations in mutation rate in an epidemic pathogen, *Yersinia pestis*. *Proc Nat Acad Sci* **110**(2), 577-582 (2012).

Supplementary information for "Pre-Columbian Mycobacterial Genomes Reveal Seals as a Source of New World Human Tuberculosis"

DNA Extraction

Kirsten I. Bos, Kelly Harkins*, Jane E. Buikstra, and Anne C. Stone*

* kelly.harkins@asu.edu, acstone@asu.edu

Extractions were performed for 68 archaeological bone samples from individuals that showed osteological lesions classically associated with *M. tuberculosis* infections, including vertebral cavitation, angular kyphosis/Pott's disease and/or periosteal remodeling on the internal aspect of the ribs[1]. All samples were of New World origin and represented time periods ranging from ca. AD 500 to ca. AD 1900 based on archaeological dating (Table S1). All extractions were performed in facilities specifically dedicated to the processing of ancient DNA, either in the Arizona State University's ancient DNA facility, or the Palaeogenetics laboratories at the University of Tuebingen. Extractions were performed following established protocols for archaeological bone[2], using approximately 100mg of pulverised material, and between 16 and 24hr rotation at 37°C for the lysis step. A minimum of one negative control was included for every ten samples.

Molecular Screening for Presence of *M. tuberculosis* DNA via In-solution Capture

Kirsten I. Bos*, Kelly Harkins, Johannes Krause, and Verena J. Schuenemann

*kirsten.bos@ifu.uni-tuebingen.de

A 5µl aliquot of each extract was converted into an Illumina DNA library following established protocols[3], with replacement of the final MinElute purification by an 80°C incubation for 20 minutes to denature the *Bst* polymerase. A minimum of one negative library control was included for every ten samples. Libraries were quantified using an IS7/IS8 primer quantification assay (Lightcycler 480 Roche)[3], using chemistry from the DyNAmo SYBR Green qPCR Kit (Termo

Scientific). A unique P5 and P7 index combination was subsequently added to all amplifiable molecules of each library in four parallel 100µl reactions following established protocols for double indexing amplification, using 0.074 units per micoliter of Ampli *Taq* Gold DNA polymerase (Applied Biosystems)[4]. Indexed products for each library were then pooled and purified over MinElute columns, eluted in 30µl TET (10mM Tris, 1mM EDTA, and 0.05% Tween), and quantified via qPCR using IS5 and IS6 primers[3], again following the DyNAmo SYBR Green qPCR Kit chemistry. Five microliters of MinElute purified product were subsequently amplified in 100µl reactions using *AccuPfx* (Invitrogen) following the manufacturer's protocol with 0.3µM each IS5 and IS6 primer to generate the high concentration of molecules needed for the bead capture approach used here[5]. Cycle numbers were strategically minimized to suspend the reaction during the amplification phase (extrapolated as the number of cycles to yield 10^{13} final copies), thus discouraging the formation of chimeric sequences which can form toward the plateau phase inlibraries of low complexity[4]. Libraries were then quantified in serial dilutions to 1:100,000 using the aforementioned IS5/IS6 qPCR assay, and subsequently combined in equimolar concentrations into pools of maximum 9.2×10^{12} copies. Pools were quantified using a DNA 1000 chip (Agilent) with an Agilent 2100 Bioanalyzer, and volumes corresponding to 2µg of DNA were determined for each pool. In addition, a 10nM pool consisting of all libraries was made for downstream shotgun sequencing.

Biotinylated baits specific to *M. tuberculosis* DNA were generated from long range PCR products via methods previously described[5], following the amplification chemistry for 100µl reactions described in the Roche Long Range PCR Kit (Roche), and using a modern *M. tuberculosis* extract as template. Capture baits were generated for five regions, and PCR primers specific to these regions generated via Primer 3[6] are shown in Table S2. Our 3538bp stretch of the beta subunit for RNA polymerase (*rpoB*) and the 4468bp region comprising two subunits of DNA gyrase (*gyrA and gyrB*) are shared with other bacteria for some of their length, and hence were used to investigate the amount of non-specific capture that might be expected from an eventual whole genome capture approach. Regions for the catalase peroxidase enzyme (*katG*, 2203bp)

and the phospholipase C gene (*mpt40*, 361bp) are both exclusive to members of the *M. tuberculosis* complex to the best of our knowledge (though *mpt40* is missing in *M. bovis, M. caprae*, and perhaps in some additional tuberculosis strains), and hence were used as specific markers for presence of tuberculosis DNA. PCRs were carried out using the Roche long-range PCR kit in 100 µl reactions (1 µl DNA template, 1 unit 10´ PCR buffer 2, 0.4 mg/mL BSA, 3µl DMSO, 3125 µM each dNTP, 7 U polymerase and 0.3 mM each primer) with the following thermal profile: an initial denaturation at 92°C for 2 min, 10 cycles consisting of a denaturation at 92°C for 10 sec, an annealing at 58°C for 30 sec and a 6-min elongation at 68°C, followed by 30 cycles using the same thermal profile with an additional increase of the elongation time by 20 sec each cycle and a final elongation at 68°C for 7 min. The long range *katG* amplification had inconsistent success, hence this region was excluded for 28 of the 96 libraries captured. Purified long range PCR products were sheared via sonication to lengths of ca. 300bp using a Bioruptor, Diogenode, Liege, Belgium. The shearing process was unnecessary for the shorter 361bp *mpt40* product. Products were visualised on a 2% agarose gel, ligated to biotinylated adapters[5], and immobilised on streptavidin-coated magnetic beads. DNA capture took place following methods described elsewhere[5], with all library pools distributed between two capture reactions: one that contained the *gyrA*/*gyrB* baits, and another that contained the *rpoB*, *katG* (if present), and *mpt40*. Capture products were eluted in 15µl TET, qPCR quantified using the IS5/IS6 assay[3] (Lightcycler 480 Roche), and reamplified to reach the amplification phase only using *AccuPfx* (Invitrogen). Products were again quantified on the Agilent 2100 Bioanalyzer with a DNA 1000 chip, and dilution pools to 10nM were made for downstream high-throughput sequencing.

Sequencing and Analysis of Screening Results

Nico Weber, Kirsten I. Bos*, Daniel Huson, and Johannes Krause

*kirsten.bos@ifu.uni-tuebingen.de

Sequencing for both shotgun and captured DNA was performed on the Illumina MiSeq platform with 2x150+8 +8 cycles using the MiSeq reagent kit v2 and the manufacturers' protocol for multiplex sequencing. A training set for the base caller Ibis 1.1.1[7] was developed by aligning all reads to the PhiX reference, and all raw reads were subsequently called. Raw reads were filtered according to their individual indices, and adapter and index sequences were subsequently removed. Paired end reads overlapping by at least 6 nucleotides and differing not more than 8% were merged, retaining the base with the higher quality score for each position.

Shotgun analysis. Samples were mapped against 2 *M. avium* strains, 14 MTBC strains and one *M. kansasii* strain (those used for array design discussed below) using Bowtie2 in local alignment mode. Samples were also mapped against the human genome (hg19, GCA_000001405.1), and summary statistics are reported in Table S1. Duplication removal was performed using *rmdup* in samtools[8] and damage pattern calculated using the mapDamage tool[9]. Average and maximum coverage for each reference was calculated. The total number of merged reads ranged from 32,452 and 181,211 for the samples, and 917 and 73,309 for the negative controls. The controls contained a maximum of 227 reads mapping to the human genome.

Screening Capture Analysis. Merged reads generated from DNA capture were mapped against corresponding regions in the H37Rv MTBC reference (accession number NC_000962.2[10]) using BWA with default parameters. Additional mapping was done using extensions of the capture regions, which included 150 bp upstream and downstream to correctly map fragments that annealed to the 5' and 3' ends of our capture probes. Environmental samples are known to contain high amounts of mycobacterial DNA[11], so the percentage of the capture regions covered at least one-fold was considered as opposed to average coverage to avoid the bias of high

coverage due to mismapping in conserved regions. No MTBC reads were found in our negative controls. We further investigated all samples that had reads mapping to a minimum of 60% of the capture regions for *gyrA/B*, *rpoβ*, and *katG*, and a minimum of 50% for *mpt40*. Although *mpt40* does not exist in some animal *M. tuberculosis* lineages, there were no samples that fulfilled criteria for *gyrA/B*, *rpoβ*, and *katG*, and not for *mpt40*. Only three samples satisfied our filtering criteria (samples 54, 58, and 64), and they did so prominently. For these samples, damage plots were generated to characterise the pattern of nucleotide misincorporations in the MTBC-specific reads using MapDamage[9]. Damage plots and the associated coverages are shown for all capture regions in Extended Data Figure 1. The damage plots for sample 54 show a pattern that is expected of ancient DNA with damage concentrated at the ends of the molecule[12,13] owing to faster deamination of cytosine in the 5' single strand overhangs common in degraded DNA[14]. Similar damage plots were demonstrated for the remaining two samples, though with some unanticipated misincorporation patterns such as the *katG* for sample 64 and the *mtp40* for both samples. These patterns which show misincorporations in the middle of the molecule potentially indicate the presence of low amounts of non-specific capture products. This suggestion is supported by the lower coverage for these regions as compared to sample 54.

Archaeological Provenience for Samples 54, 58, and 64

Kelly Harkins*, Jane E. Buikstra*, and Anne C. Stone

*kelly.harkins@asu.edu, buikstra@asu.edu

All three samples that fulfilled our criteria for whole genome capture represent burials from a narrow geographical region along the Lower Osmore drainage in southern Peru, near the modern city of Ilo (Figure 1, main manuscript). The coastal site Chiribaya Alta (Sample 54) and the mid-valley site El Yaral (Sample 64) are both attributed to the Chiribaya polity (AD 772-1350), while a contemporary Ilo-Tumilaca/Cabuza group occupied the coastal site of El Algodonal (Sample 58)[15-18]. The extreme aridity along the Western slopes of the South Central Andes may have contributed to both archaeological and biomolecular preservation. These three sites have been

the focus of numerous bioarchaeological and archaeological studies, providing us with a broad contextualization of political, socio-economic, and ritual organization of pre-contact Andean communities, including regional studies on the pathoecology[19,20]. Overall, the late Middle Horizon (AD 500-1000)/Late Intermediate (AD 1000-1450) period of Andean prehistory in the Osmore drainage reflects changes following the weakening and eventual collapse of the expansive Tiwanaku state, which either resulted in a diversification and strengthening of local cultural/stylistic traditions present in the coastal valley, e.g. Chiribaya, Ilo-Tumilaca/Cabuza[21-23] or an influx of groups repopulating areas of the lower Osmore drainage from the middle valley[24-28].

At least two distinct ethnic identities are discernible in material culture and mortuary traditions within the Chiribaya domestic and mortuary sites, highlighting different ethnic and economic strategies, namely the *laboradores* and *pescadores*. Inland *laboradores*, associated with agricultural specialization and terrestrial subsistence strategies, are distinct from their coastal neighbours, the *pescadores,* characterized by marine specialization. Despite conflicting evidence concerning the biological and cultural origins of groups within the lower Osmore drainage during the late Middle Horizon/Late Intermediate, skeletal non-metric biodistance suggest genetic continuity amongst groups[21]. Internal cultural differentiation was, nevertheless, pronounced within the Chiribaya polity.

The first excavation of Chiribaya Alta began in 1989 by Maria Lozada Cerna and Barra O'Donnabhain. *Proyecto Chiribaya*, a part of *el Programa Contisuyo*, continued in 1990-1991 under the supervision of Jane Buikstra, Sloan Williams, and Maria Lozada Cerna. The site of Chiribaya Alta, located on a coastal promontory only seven kilometers from the sea, is the most extensive and complex Chiribaya site. With nine cemeteries, multiple domestic units and plazas as well as a defensive wall structure, Chiribaya Alta is considered a centre of political power and prestige for the Chiribaya polity[15,16,18]. Chiribaya Alta is also the only site with unequivocal evidence for cohabitation of multiple ethnic identities[29], with numerous cemeteries demonstrating affiliations with distinct communities. Although evidence of the *pescadores* identity is more prevalent at the coastal Chiribaya sites, Chiribaya Alta Burial 103 (referred to here as Sample 54)

was within a cemetery noted for its similarity to inland agricultural Chiribaya sites like El Yaral, both in material culture and dietary patterns[21].

The mid-valley Chiribaya site of El Yaral, from where our Sample 64 originated, was also excavated as a part of Proyecto Chiribaya, with excavation of the two cemeteries overseen by Jane Buikstra and Maria Lozada Cerna. The mid-valley, known for its farmable land, is isolated from the coastal valley by 25km of uninhabitable land[18]. The site of El Yaral is located 50 km from the sea at 1000 meters above sea level and is surrounded by Tiwanaku-affiliated sites. El Yaral is considered an agro-pastoral settlement, consisting of more than 300 residential terraces with a large mortuary component. Evidence for a highly ritualized tradition of animal sacrifice and extensive camelid herding[30-33] are also found at El Yaral. Strong reliance on agricultural and terrestrial resources characterizes the archaeological assemblage. Contact with the coast, however, is inferred from the presence of some marine resources in domestic contexts. There is evidence for a strong *laboradores* ethnic identity reflected in mortuary patterns and confirmed by Tomzcak's[34] dietary examination.

The mortuary and domestic sites at El Algodonal, from where our Sample 58 originated, were excavated as a part of *El Proyecto Colonias Costeras de Tiwanaku*[18]. In contrast to Chiribaya Alta and El Yaral, El Algodonal was occupied by an Ilo-Tumilaca/Cabuza cultural group, located less than 10 km inland from Chiribaya Alta[18]. Owen[18] associates this Tumilaca phase in the lower Osmore with the aftermath of Tiwanaku collapse, supported by dental and non-metric data[26]; however, despite clear evidence for cultural differentiation, other biodistances derived from metric and non-metric skeletal traits reflect biological continuity between the Chiribaya sites and El Algodonal[21,22,35], rather than with descendants of Tiwanaku colonies located along the upper Osmore. Whether an early Chiribaya polity was developing in-situ in the lower Osmore during the late Middle Horizon or its emergence was a result of repopulation following abandonment of Tiwanaku-influence, it is clear the Chiribaya dominated the lower river valley by the later Late Intermediate period[15,18], persisting long after the decline (or absorption) of the Ilo-

Tumilaca/Cabuza population. While archaeological assemblages at the domestic site of El Algodonal demonstrate multiple cultural phases, Owen[18] attributes the burials in the cemetery with confidence to the Ilo-Tumilaca/Cabuza tradition.

For discussion of the paleopathological evidence of tuberculosis at sites within the Osmore River drainage, see below.

Radiocarbon Dating

Kelly Harkins*, Anne C. Stone, and Jane E. Buikstra

*kelly.harkins@asu.edu

Vertebral body fragments of each sample were sent to Beta Analytic, Inc., for radiocarbon dating and nitrogen stable isotope analyses. The calibrated date estimates, reported in Table S3, were produced using CALIB 4.0[36] with the Southern Hemisphere correction (ShCal13)[37]. Radiocarbon dates derived from human bone are interpreted with certain caveats. The marine reservoir effect, known as the "old carbon problem", occurs when an upwelling of ^{14}C-depleted deep water can cause marine organisms to have ^{14}C ages centuries older than the apparent ages of terrestrial material. Although a regional correction (**Δ**R) can be estimated, the size of the marine reservoir effect is extremely variable, both geographically and temporally, and it poses problems for accurate radiocarbon dating of material from coastal sites[38,39]. In archaeological contexts, the concern is heightened when human samples are from littoral environments and/or when the diet is expected to be composed mainly of marine resources. Systematic studies comparing terrestrial and marine mammals from archaeological sites, as well as humans remains from individuals with a heavy reliance on seafood, have shown that the marine reservoir effect influences date estimates[40]. Owen[41] directly addressed the old carbon problem over time within the Chiribaya archaeological context in the Osmore Valley and concluded that the **Δ**R with a 40-y Southern Hemisphere correction is 382+-102 years – larger and more variable than other nearby regions. Dietary composition can be estimated with 15N/14N and 13C/12C ratios, but there is no

straightforward method for determining whether these values necessitate the application of a delta R correction for the “old carbon” effect.

At the cemetery site of Algodonal, radiocarbon dates of raw wool from multiple burials were first reported by Owen[18]. Despite Owen’s[41] observed influence of the marine reservoir, our ^{14}C date estimates from human bone at this site closely match those of non-collagen material from the same cemetery[18]. According to carbon and nitrogen isotope value ranges of marine and terrestrial resources determined specifically for the Osmore Valley[42,34], isotope ratios for Sample 58 do suggest a marine dietary component. Nevertheless, the concordance of dates between collagen-derived and non-collage material suggests that at Algodonal the marine reservoir effect was not sufficiently large to influence sample age estimates. Similarly, the calibrated radiocarbon dates for the two individuals sampled from Chiribaya Alta and El Yaral closely match previous dates from skeletal material[30,35], and fall well within the expected archaeological chronology for the Chiribaya polity, thus providing further evidence that the marine reservoir effect is negligible. We acknowledge that marine component in the diet can vary over time periods and short geographical distances, and can even vary between two individuals from the same culture or archaeological site (Bruce Owen, personal communication, October, 2013). Recent re-sampling in October 2013 of textiles from the two Chiribaya and El Yaral burials will potentially clarify any apparent discrepancies between estimated ages derived from human samples and terrestrial material.

In regions where plants are fertilized with seabird guano or camelid dung, human dietary signatures could reflect in part the leaching process of the manure rather than the diet itself[43], further complicating estimations of the marine proportion in an individual’s diet, and subsequently the potential marine reservoir effect. Seabird guano-fertilized maize plants can have δ15N values similar to or greater than those of marine mammals, fish, and seabirds themselves [44,45], which has implications for reconstructing the diets of pre-contact coastal and valley groups in Southern Peru and northern Chile, who practiced seabird guano fertilization[46]. Furthermore, camelids who

also consume these nitrogen-rich foods would have elevated levels of δ15N, and would thus contribute to higher nitrogen levels in the humans who consumed plants fertilized with their dung[47], or in wool textiles manufactured from their coats. Unfortunately bone collagen does not provide the resolution required to disentangle the relative impacts of the consumption of high trophic level marine resources and guano-fertilized maize[45].

Array design

Kirsten I. Bos*, Nico Weber, and Stephen A. Forrest

*kirsten.bos@ifu.uni-tuebingen.de

Probe design was accomplished through use of an in-house modification of available probe design software[48], such that degenerate templates could be accommodated. A degenerated consensus sequence was created from a MAUVE[49] multiple genome alignment of 21 genomes that form part of the *M. tuberculosis* complex. Genomes used in this generation along with their accession numbers are shown in Table S4. Regions with greater than 10 degenerate positions per 60bp region were interpreted as alignment artifacts as opposed to true sequence polymorphism, and were replaced with the corresponding sequence of the H37rV reference genome (accession number NC_000962.2[10]). Indels were treated as length polymorphisms, and probes were generated to span all possible indel combinations. Probes were designed with 4bp tiling density across the template. For regions with fewer than 10 degenerate bases per 60bp stretch, probes representing all nucleotide variants were made for the first degenerate base in the probe, and nucleotides were assigned evenly to represent possible diversity from the remaining degenerate positions. Efforts were made to identify and mask repetitive elements based on 15-mer frequency counts of 100 or more. Probes identical to those already generated were removed. This resulted in 1,493,747 probes for the *M. tuberculosis* complex. Probes were distributed between two Agilent 1-million feature arrays, and the remaining probes were distributed evenly between the genetically divergent *M. avium* and *M. kansasii*. With 5bp tiling, retention of only

unique probes, and repeat masking, 860,009bp and 925,601bp could be accommodated for *M. kansaii* and *M. avium*, respectively. The end result was two arrays, each with 974,016 probes.

Array capture

Kirsten I. Bos, Kelly Harkins, and Verena J. Schuenemann

*kirsten.bos@ifu.uni-tuebingen.de

Highly concentrated libraries were generated from 80µl of DNA extract from samples 54, 58, and 64, and their associated negative controls following the methods described above. Pretreatment with a uracil DNA glycolsylase (UDG) enzyme followed by endonuclease VIII was done to remove deaminated cytosines and replace them with an intact nucleotide, thus preventing erroneous substitution assignments from the most common form of DNA damage[50]. Library molecules then received unique index combinations in an IS5/IS6 amplification reaction[4], were purified over MinElute columns (Qiagen), and reamplified with AccuPrime Pfx (Invitrogen) to attain two 20µg pools of purified product for capture on each of our two arrays, where 80% consisted of the three UDG treated libraries, and the remaining was distributed evenly between negative controls and the non-UDG libraries prepared for bead capture described above. Array hybridization was performed following methods described elsewhere[51], with two nights of incubation at 65°C. Post-harvest, the 490µl eluate was reamplified to again generate two 20µg pools of purified product for serial capture on a set of arrays identical to those used in the first round of capture. Products were eluted following established protocols[51], reamplified in IS5/IS6 reactions described above, purified over MinElute comlumns, and diluted to produce a 10nM pool for high-throughput sequencing. Sequencing was performed on an Illumina HiSeq 2000 platform by 2*101 + 7 cycles using the manufacturer's protocols for multiplex sequencing and a TruSeq PE Cluster Kit v3-cBot-HS cluster generation kit and a TruSeq SBS kit v3 chemistry.

Modern MTBC Genome Preparation and Sequencing

Mireia Coscolla and Sebastien Gagneux*

*sebastien.gagneux@unibas.ch

Genomic DNA was extracted using a standard CTAB method[52], and purified using Genomic DNA Clean & Concentrator (Zymo Research). Sequencing libraries were constructed using Nextera® XT-DNA from Illumina according to the manufacturer's instructions. Libraries for up to 12 strains were pooled and run in a single MiSeq flow cell 2x251 cycles in the quantitative genomics facility from DBSSE (ETHZ) in Basel.

Read Processing and Mapping

Alexander Herbig*, Mireia Coscolla, Johannes Krause, Sebastien Gagneux, and Kirsten I. Bos

*alexander.herbig@uni-tuebingen.de

High-throughput sequencing on the HiSeq 1500 yielded between 81,257,798 and 50,483,372 raw reads per sample library, and between 409,545 and 110,088 reads for the controls. Adapters were subsequently clipped from all reads and overlapping read pairs were merged as described elsewhere[55]. Reads that could not be merged were quality-trimmed for a minimal Phred score of 20. All reads with a length below 30 nt (nucleotides) after clipping and trimming were discarded.

Array captured reads from our Peruvian samples were analysed alongside short reads of 275 *M. tuberculosis* libraries for comparison. This set was comprised of 269 genomic samples that have been described elsewhere, 259 of which were used in a recent comprehensive study of *M. tuberculosis* evolution[54]. The remaining genomes consisted of one ancient sample that was recently obtained from a Hungarian mummy[55], four *M. caprae* samples[56], four *M. microti* samples (EBI ID: ERR027294, ERR027295, ERR027297, ERR027298) and one sample that was isolated from a wild chimpanzee[57]. An additional six *M. pinnipedii* strains (referred to as Pinnipedi7011,

Pinnipedi_G01222, Pinnipedi7739, Pinnipedi_G01491, Pinnipedi_G01498, Pinnipedi_G01492) were also included[58,59]. Genomes used for analyses are listed in Table S5.

All reads were mapped via BWA to a computationally extrapolated MTBC ancestral genome[60] using either the *sampe* or *samse* algorithms for single read or paired-end data[61]. Using BWA standard parameters a high number of "heterozygous" positions were observed in the mappings of Peruvian samples 54 and 64. This phenomenon was also observed in the Hungarian mummy and for some of the modern MTBC isolates. Asymmetries in read coverage in the original analysis were previously interpreted as an indication that the Hungarian mummy was infected with two different MTBC strains, which could partially explain the high number of heterozygous positions in that individual. To evaluate this possibility, SNP allele frequency distributions of several samples showing high heterozygosity were further investigated. Extended Data Figure 2 shows histograms of the SNP allele frequency distributions of the three Peruvian samples and the Hungarian mummy. For the Hungarian mummy, a normal distribution of SNP allele frequencies with a mean of about 50% is observed, which suggests an infection with two strains in equal abundance. For samples 54 and 64, a large number of heterozygous positions with an allele frequency of around 20% are observed. If this were due to a mixed infection, another set of heterozygotes with SNP allele frequencies of around 80% would be expected; however, in both samples this is clearly not the case, hence the perceived heterozygosity is likely the result of mismapping of non-tuberculosis reads.

To avoid cross-mapping of reads from other organisms, the mapping stringency of BWA was increased to 0.1 (-n parameter) and all mapped reads were filtered for a minimal mapping quality of 37. Histograms of SNP allele frequency distributions under these stricter parameters are shown in Extended Data Figure 3. The distribution in the Hungarian mummy sample did not change significantly, providing further support of a mixed infection. The number of heterozygous positions of the two Peruvian samples 54 and 64, however, decreased significantly, showing a

distribution which is very similar to those of most modern isolates such as sample L1_N0141b4, which is included in Extended Data Figure 3 for comparison. This persistent low level of heterozygosity is likely the result of non-specific mappings to regions that are conserved amongst bacteria, such as the rRNA and tRNA regions. Exclusion of these regions in SNP calling largely circumvented this mapping artifact, and confirmed that only single strains were represented in each of the Peruvian samples. Several modern isolates had a high proportion of heterozygous positions. One example is sample L6_N0090 (Extended Data Figure 3) which had bimodal SNP allele frequencies with maxima around 0.15 and 0.85, even under our stricter mapping, the result of either a mixed infection or minor contamination in growth media.

A small number of reads were found in our negative controls, which presumably arose from chimeras that formed during the pooled reamplification after either of the serial array captures[4]. The number of mapped reads post duplicate removal ranged from 215 to 881 (Table S6).

Our stricter mapping reduced the five-fold genomic coverage to below 75% for 16 of our modern genomes, hence these were removed from subsequent analyses. For the three Peruvian samples 54, 58 and 64, the percentage of mapping reads with the stricter mapping criteria were 37.0, 29.2 and 17.2, respectively, yielding a genomic coverage of about 31 , 20 and 23-fold. The cluster factors of the three libraries ranged between 3 and 10, thus indicating that they were sequenced to a great depth. A summary of mapping statistics for the three samples is provided in Table S6.

Mapping analysis also confirmed presence of the IS6110 and IS1081 mobile elements that are frequently used for PCR-based detection of MTBC DNA. Asymmetric coverage across the genome, however, prevented us from reliably estimating copy number for these elements.

Regions of Difference

Mireia Coscolla*, Alexander Herbig, and Kirsten I. Bos

*mireia.coscolla@unibas.ch

In order to investigate if regions of difference (RD) reported deleted in the animal strains of the MTBC (RD7, RD8, RD9, RD10)[62] are also found in our Peruvian strains, we evaluated the number of reads mapping to the respective regions as performed elsewhere[57]. We observe that no reads could be mapped to regions RD7, RD8, RD9, and RD10. Single reads, likely stemming from mapping artifacts, existed for Sample 54 in RD7 and Sample 64 in RD9.

As confirmation of the detected RDs we performed an additional mapping of the sequencing reads using BWA-MEM[63], which is able to perform split alignments, meaning the algorithm can align reads that span a deletion in the reference genome by splitting it into two segments, which are aligned separately. This allows for a precise determination of the coordinates of a deleted region in the reference genome. We applied BWA-MEM to the merged reads of the three samples using standard parameters. As the reference we used the constructed common ancestor as described above.

By this we could confirm the deletion of RD7, RD8, RD9, and RD10 in all three Peruvian samples with numbers of reads spanning the deletion ranging from 3 (RD7 in Sample 54) to 25 (RD8 in Sample 54).

The determined coordinates are in concordance with the genomic intervals between external and internal primers that have been used to detect RDs in members of the MTBC by PCR[62,64].
In addition, we detected the seal-specific RDseal deletion[65] in the three Peruvian strains. The *M. microtii*-specific deletion RDmic is not present in any of the Peruvian strains.

Precise genomic coordinates as determined from our split read mapping and the number of spanning reads for each deletion in each sample are provided in Table S7.

SNP Calling and Phylogenetic Analysis

Alexander Herbig*, Mireia Coscolla, Johannes Krause, Sebastien Gagneux, Iñaki Comas, and Kirsten I. Bos

*alexander.herbig@uni-tuebingen.de

Comparative SNP typing was performed on mapped data for all 262 samples using the Unified Genotyper of the Genome Analysis Toolkit (GATK)[66]. A *vcf* file (variant call format) for each sample was produced, and the 'EMIT_ALL_SITES' option was set to generate calls for both variant and non-variant sites.

A custom Java program (MultiVCFanalyzer) was applied to all generated *vcf* files in parallel to comparatively analyse and filter the detected SNPs and to produce a multiple sequence alignment of all positions, for which a SNP was called in at least one of the strains in the complete dataset. We called a SNP if GATK called a homozygous SNP, if the position was covered by at least 5 reads and the quality of the GATK call was at least 30. A reference base was called if GATK called the reference base, the position was covered by at least 5 reads, and the quality was at least 30. If GATK called a heterozygous SNP, we called the SNP allele if it was supported by at least 5 reads with a quality of at least 30 and if the fraction of reads containing the SNP was at least 90%. In this case we still considered the SNP to be homozygous. If a SNP call was not possible, the reference base was called instead if the respective thresholds were reached. If neither a reference call nor a SNP call was possible based on these thresholds, the 'N' character was inserted in the SNP alignment. All regions that contain a high degree of variability or that were suspected to be affected by cross-mapping of reads from other organisms were excluded. These included all repeat regions, insertion and mobile elements, phage-related genes, *PE*, *PPE*

and *PGRS* genes, *muturase* and *resolvase* genes, *REP* family genes, as well as *tRNA*s and *rRNA*s.

Using this approach a total of 42,526 variant positions were called in the complete data set. Inclusion of the *M. canettii* outgroup raised this total to 53,177. The numbers of SNP calls for the three Peruvian samples 54, 58, and 64 were 756, 697, and 740, respectively. In the complete dataset the number of variant positions ranged from 489 in strain L4_erdman to 1415 in the *M. orygis* strain. A summary of the number of variant positions detected in each strain can be found in Table S8.

For phylogenetic analyses, the SNP alignment generated from the *vcf* files was used as input for the phylogenetic reconstruction software MEGA5[67]. All alignment columns that contained missing values (i.e., 'N') were excluded from the analysis (complete deletion) in order to avoid any bias resulting from different genomic coverages between samples. Post filtering, 22,480 positions remained for the phylogenetic reconstruction.

A Maximum Parsimony tree of all 262 samples is shown in Extended Data Figure 4. A close-up view of the Maximum Parsimony subtree of all lineage 6 and animal strains including branch lengths and bootstrap statistics is shown in Extended Data Figure 5. Maximum Likelihood and Neighbor Joining trees of the complete dataset and respective subtrees focusing on lineage 6 and animal strains are shown in Extended Data Figures 6 – 9. All trees show a high bootstrap support for separation of *M. microti*, *M. pinnipedii*, and the ancient Peruvian strains.

SNP Effect Analysis

Inaki Comas*, Douglas Young, Alexander Herbig, Sebastien Gagneux, and Mireia Coscolla

*inaki.coma@uv.es

All homozygous SNPs that were detected in at least one strain of the dataset excluding the outgroup were investigated for their effect on protein-coding genes using *SnpEff* [60]. For this a custom annotation database was generated for all protein-coding and non-protein-coding genes annotated in the genome of *M. tuberculosis* H37Rv, on which the computational construction of the hypothetical MTBC ancestral genome was based.

SnpEff was applied to the complete set of SNPs using standard parameters except that the size of the upstream and downstream region of genes in which SNPs are classified respectively was set to 100 nt. A total of 38,304 SNPs were located in protein-coding regions. Of these 24,176 were classified as non-synonymous: 23,432 cause a non-synonymous amino acid change, 55 lead to the loss of a start-codon, 56 lead to the loss of a stop codon, and 633 represent a stop mutation. Altogether 66 SNPs were located in non-coding RNA. A list of all detected SNPs and their effect on annotated genes for the animal cluster, which includes the ancient Peruvian strains, is provided in Table S9.

The ancient human strains belong to a phylogenetic cluster that includes human Lineage 6 and *M. bovis* and has the broadest host range among lineages of the MTBC. The whole cluster is defined by deletion of the *Mce3* locus and 79 non-synonymous SNPs enriched amongst proteins associated with the cell envelope (Tuberculist functional class II.C[10]), including mutations predicted to affect the function of mycolic acid modification enzymes MmaA2 and MmaA4. There are two nsSNPs in the MmpL3 transporter that is the target of multiple mutations conferring resistance to a series of novel drug candidates[69]. While MmpL3 resistance mutations are generally located in transmembrane domains and are predicted to impair function, the Lineage 6/*M. bovis* nsSNPs are in less restrictive cytoplasmic and extracellular domains. Several nsSNPs

target central carbon metabolism, including a mutation predicted to affect the function of RamB (*Rv0465c*), a regulator of isocitrate lyase expression in the presence of glucose[70]. Similar RamB nsSNPs occur in human Lineages 1 and 5.

Subsequent niche adaptation of the *M. pinnipedii*/*M. microti* branch is suggested by acquisition of stop codons that reduce the potential for transcriptional adaptation by truncating alternative sigma factor SigF at position 122 to leave a protein that contains only the core sigma-70 region 2, and excising the kinase domain from the predicted Rv0845 two-component regulator. nsSNPs common to the *M. pinnipedii*/*M. microti* branch are enriched for transport and binding proteins (Tuberculist functional class III.A) including multiple metal ion transporters.

The ancient human isolates are characterised by a further nsSNP (D62N) in Rv0092, encoding the CtpA copper efflux protein. Aspartate 62 is conserved in *M. canettii* and the related animal-adapted pathogens *M. ulcerans* and *M. marinum*, but has been substituted by alternative amino acids (D62G and D62E) in two independent *M. tuberculosis* isolates. The occurrence of homoplasies and positions with more than two different amino acid residues is very uncommon in the MTBC and a site-wise analysis of codon 62 confirms that this position may be under diversifying selection (see below). Asp62 is located in a heavy metal binding domain that is a characteristic feature of copper-transporting P-type ATPases and is proposed to play an allosteric role in regulating transfer of Cu^{+} from a cytoplasmic chaperone[71]. Additional nsSNPs targeting genes involved in copper ion homeostasis in the *M. pinnipedii*/*M. microti* branch include *I79F* in transcriptional regulator *CsoR* and *I4M* in metallothioinine *MymT* (Table S9). Taken together, these nsSNPs suggest the potential adaptation to a niche with altered metal ion availability. Starvation by metal ion restriction and intoxication by metal ion overload have both been implicated as mechanisms for control of bacterial infection[72]; resistance to copper ion overload is required for survival of *M. tuberculosis* H37Rv in a murine infection model[73].

Four other nsSNPs are specific to the ancient human strains (Table S9), including a predicted functional SNP in Rv2258c, encoding a methyltransferase linked to ubiquinone metabolism, an uncharacterised transcriptional regulator, and two further hypothetical proteins.

Positive selection analysis

Iñaki Comas*

*inaki.coma@uv.es

To determine if codon 62 was likely under diversifying selection in the MTBC we carried out a site-by-site positive selection analysis of the gene *Rv0092*. For the analysis we used the SLAC method as implemented in the datamonkey server[74]. The SLAC method is suited for analyses with large numbers of sequences in the alignment and is based on reconstructing ancestral codon estates and quantifying the extent to which the number of synonymous and non-synonymous substitutions per codon deviates from the expectation for a neutrally evolving codon. As an alternative we carried out the site-wise position selection analysis by implementing the in codeml as part of the PAML package. We compared a model that assumes that no site can have a dN/dS > 1 (M8a) with a model where sites are allowed to have dN/dS >1 (M8 model). As the two models are nested the likelihood associated to them can be compared by a likelihood ratio test (LRT). The LRT was highly significant (LRT = 240.61, P-value < 0.001 according to X2 df =1). To identify the codons likely under positive selection a Bayes Empirical Bayes analysis (BEB) was carried out based on *a priori* distribution of dN/dS in eleven categories between 0 < dN/dS > 1 (Table S10). Both the SLAC and codeml analyses identified codon 62 to be under diversifying selection (codeml: dN/dS 10.092, posterior probability of dN/dS > 1 = 0.999; SLAC probability that dN>dS 0.998).

Estimation of substitution rates and divergence times of Peruvian strains

Johannes Krause*

*johannes.krause@uni-tuebingen.de

To estimate divergence times of modern and ancient *M. tuberculosis* strains and lineages 1-7 from the MTB complex, substitution rates were calculated in a Bayesian framework using the software package BEAST 1.7.5[75]. After variable positions between the reconstructed common ancestor[60] and all modern and ancient *M. tuberculosis* strains falling into the complex were identified, a multiple sequence alignment was created that included all variable positions of the 258 modern *M. tuberculosis,* genomes including *M. canettii* as outgroup as well as the three pre-Columbian Peruvian genomes to the reference. The ancient Hungarian *M. tuberculosis* genome was excluded from this dating analysis as it displayed high levels of heterozygosity, likely as a result of a multiple strain infection as discussed above. Thus the total number of genomes used here is 261. All positions containing gaps and missing data were eliminated, leaving a total of 22,544 variable positions in the dataset. In order to test if the strains evolve in a clock-like manner, a likelihood ratio test was performed in MEGA5[67] by comparing the ML value for the given topology using only the modern strains (excluding *M. canettii*) with and without the molecular clock constraints. The null hypothesis of equal evolutionary rate throughout the tree was rejected at a 5% significance level (P = 5.16E-147).

As a result, an uncorrelated lognormal-distributed relaxed clock model of rate variation among branches was chosen. The calibrated radiocarbon dates (oxcal 4.2) for all three ancient Peruvian samples were used as priors (Table S3). As tree priors, both the constant size and expansion growth coalescence models were considered. For each model, three MCMC runs were carried out with 50,000,000 iterations each, sampling every 10,000 steps. The first 5,000,000 iterations were discarded as burn-in. For each model all three independent runs were combined resulting in 135,000,000 iterations. We evaluated chain convergence to stationarity for all model parameters

using Tracer v1.5[76]. Maximum Clade Credibility (MCC) trees were summarized using TreeAnnotator v1.7x (available as part of the BEAST package) and visualized and manipulated in FigTree v1.4 (http://tree.bio.ed.ac.uk/software/figtree/) including a maximum clade credibility tree showing the amount of rate heterogeneity throughout the MTBC phylogeny (Fig .S10).

Using the uncorrelated lognormal distributed relaxed clock and constant size model we estimate a substitution rate of 4.63 $x10^{-8}$ substitutions per site in the genome per year (3.03 $x10^{-8}$ – 6.21 $x10^{-8}$ 95% HPD) and with the expansion growth prior we estimate a substitution rate of 4.9 $x10^{-8}$ substitutions per site in the genome per year (3.39 $x10^{-8}$ – 6.36 $x10^{-8}$ 95% HPD). This corresponds to 0.22 substitutions per genome per year for the constant size and 0.23 for the expansion size model, respectively. The resulting divergence times for the TMRCA for all *M. tuberculosis* strains is 4,449 YBP (2,990 – 6,062 YBP 95% HPD) using the constant size tree prior and 4,046 YBP (2,951 – 5,339 YBP 95% HPD) using the coalescence expansion tree prior, respectively (Extended Data Figure 11).

For the purposes of exploring the demographic histories of animal and human tuberculosis strains, we used the Bayesian skyline model implemented in BEAST 1.7.5[75]. As tree priors the expansion growth coalescence model and the uncorrelated lognormal-distributed relaxed clock were used. Two MCMC runs were carried out with 50,000,000 iterations each sampling every 10,000 steps, with 100 population size intervals. The first 5,000,000 iterations were discarded as burn-in. Both independent runs were combined resulting in 90,000,000 iterations. Skyline plots were generated with Tracer v1.5[76].

Dating analysis for the mummified Hungarian individual

Julian Parkhill, Josephine Bryant*, and Simon R. Harris

*jb31@sanger.ac.uk

Raw sequencing reads were obtained for 141 modern MTBC isolates described elsewhere[54,57], the Hungarian ancient TB sample[55] and one modern isolate from Malawi, sequenced at the Sanger Institute as part of another project, and available from ENA accession number ERR245743, which was found to be the closest match to one of the Hungarian strains. Reads were mapped to a corrected version of the H37Rv reference[77] using the program SMALT[78], and a combination of *samtools* and *bcftools*[61] were used to call bases as part of an in-house pipeline. Appropriate filters were used to keep the number of false positive variant calls to a minimum as previously described[79]. Excluding the mixed Hungarian TB sample we built a maximum likelihood phylogeny using RAxML[80] based on variants called in the remaining strains and mapped the Hungarian sequence variants back onto the resultant phylogeny using ACCTRAN parsimony algorithms[81].

To deconvolute the Hungarian TB sample we extracted mapped bases which had been identified as mixed by samtools (n=1174), confirming that the sample comprised a 50/50 mixture of two TB strains, as described previously[55]. The majority of these mixed sites (n=661) occurred at positions where variants occurred on the internal branches of the MTBC phylogeny constructed above. For each of these sites we identified the splits on the tree defined by the two variants, which all mapped to two paths through the tree (as shown in Extended Data Figure 12 – red branches), thus allowing us to separate the non-unique variants of the sample onto these two paths. Divergence points of the isolates on these branches were determined by identifying the proportion of variants that occurred on that branch that were found mixed in the sample (see Extended Data Figure 12). The remaining mixed sites were not shared with any modern isolate, and could either represent mapping error or variants unique to the ancient strains, which would form their terminal branches. We applied two filtering techniques to these variants. For the first

"relaxed" method we used variants that had a base and mapping quality of 50 and 30 respectively (n=220). For the "stringent" method we applied additional filters, including strand bias (P>0.05)[61], coverage within 50% of the mean, only two possible alleles and a distance of at least 200bp from another variant; resulting in 105 variants. Assuming the two mixed samples should have approximately equal distances from the root of the tree, this enabled us to determine an upper (stringent) and lower (relaxed) bound of 4.5 and 60.5 SNPs on the terminal branch of Hungarian 1 and 91.5 and 147.5 SNPs on the terminal branch of Hungarian 2. Using two published substitution rates for TB, of 0.3[79] and 0.5[82] SNPs per genome per year, we estimated the bounds of the probable age of the penultimate node (marked by a star Extended Data Figure 12) for these isolates as shown in Table S11. We used these values to construct lognormal prior distributions for the ages of the nodes in BEAUti[75]. For Hungarian 1 we used a mean of 97.5, offset of 201.9 and a standard deviation of 1. For Hungarian 2 we used a mean of 406 and offset of 201.9 and a standard deviation of 0.8. Although we use a hard cutoff minimum bound of 201.9 years (which we know it cannot be older than) the distributions are generous (213 – 508 and 281-1301 95% CI), encompassing the bounds calculated in Table S11. The ages for all other isolates were set to the present day (year 2000). We built an alignment based on the synonymous variants identified in the modern isolates and the separated out the Hungarian samples, excluding any variants called in the PE and PPE genes which are known to cause alignment issues. We ran three independent MCMC chains of 100,000,000 states in BEAST v1.7.5[75] using a lognormal relaxed clock model and constant population size model under a GTR model of evolution. Tracer (v1.5) was used to assess convergence (after an initial burn-in period of 10,000,000) and one run with all ESS values > 200 was used to produce a maximum clade credibility tree in TreeAnnotator v1.7.1[75] shown in Extended Data Figure 13.

Pinniped Exploitation in Pre-contact South America

Kelly Harkins*

*kelly.harkins@asu.edu

Evidence that human groups exploited pinnipeds for subsistence, tool, and ritual use is well-documented along the Peruvian and Chilean coasts, including Tierra del Fuego, as early as the Pleistocene period. In Peru specifically, Dillehay et al[83] report stone tool processing of *Otaria* on the northern coast ca. 14,000 BP, and DeFrance and Alvarez[84] see evidence of active sea lion hunting ca. 11,000 BP at the late Pleistocene site at Quebrada Tacahauy on the southern coast. By 6000 BP, specialized gatherers and fishers from the southern tip of South America were hunting pinnipeds as their main dietary staple[85-87]. Pinnipeds in this area are available year-round, either in permanent or seasonal colonies or as isolated individuals scattered along the coast[88]. Reliance on pinniped meat and blubber may have provided coastal groups with adequate nutrition caloric intake to withstand the cold conditions of the southern continent.

Over time, pinnipeds had an increasing representation in Andean coastal archaeological sites, as their remains were associated with tool use, implementation in domestic structures, ritual representation, and use in the process of mummification or body preparation. The use of seals or sea lions in a mortuary context is also evident for the archaic Chinchorro group of northern Chilean coast, as pinniped coats were often used in place of skin of the deceased[89,90]. Pinniped remains were also found in Early Intermediate period sites (200 BC- AD 800)[91] and during the late Moche Period (AD 550-800) along the north coast of Peru[92]. Here sea lion remains were often represented in domestic and funerary offerings and depicted in Moche iconography, often on ceramics with or without hunting scenes[93,94].

Paleoparasitological and isotopic evidence also support the consumption of pinnipeds at coastal South American archaeological sites. Ferreira et al.[95] report the presence of *Diphyllobothrium pacificium* eggs, a tapeworm found in South American sea lions and other fish-eating mammals, in human coprolites of the early Chinchorro inhabitants of northern Chile. Human infection results

from ingestion of raw or undercooked, infected meat. Interestingly, tapeworm infection was not found in coprolites in the Alta Ramirez (1000 BC-AD 500) cultural phase directly following the Chinchorro; the authors thus suggest that marine resources, shown to comprise a third of their diet, might have been cooked[96]. Fluctuating rates of infection, however, could have been caused by El Niño disturbances affecting marine mammal and related *D. pacificum* distribution[97]. Callen and Cameron[98] also reported possible *Diphyllobotrium* eggs from coprolites at Huaca Prieta mound on the coast of northern Peru, dating to 3000 – 1200 B.C. Subsequent work has further implicated *Diphyllobothrium* spp. parasites as a common infection in Peruvian and Chilean coastal groups, with overall prevalence in some areas, including at Chiribaya sites in the Osmore River valley during the Late Intermediate Period, as high as 56.6%[99,-102]. In general, the authors relate these finding to the foraging strategies and hunting of marine resources.

Specialized sea lion hunting persisted until European contact, as evident in chronicles[103], highlighted by Santoro et al.[104]: "Those who kill sea lions do not kill other fish… In this way, each category of fisherman kills the type of fish of his specialty". As also suggested by Bastida et al.[105], this long history of pinniped exploitation by indigenous South American communities suggests a plausible scenario for the initial transmission of *M. pinnipedii* to humans.

Paleopathological Evidence of Tuberculosis in the New World

Kelly Harkins*, Jane Buikstra, Anne Stone

kelly.harkins@asu.edu

For over a century, (bio)archaeologists have studied the relationship between morphological indicators of clinical tuberculosis and changes observed in archaeological human remains that are considered consistent for the disease. Most notably these changes include kyphosis of the vertebral column following collapse and fusion of thoracic and lumbar vertebral bodies, referred to as "Pott's Disease", and/or the presence of a Ghon complex or cavitary lung lesions in soft tissue remains[106], although no changes alone are fully pathognomonic. A pattern of skeletal involvement may include lesions of the spine, as well as any joint, especially those that bear loads, and the

skull[107]. Resorptive lesions or proliferation on the rib surfaces are also evaluated within the context of tuberculosis diagnosis, but alone are considered non-specific, as many pulmonary and other infections can cause these changes[108-111]. Thus, there are a number of caveats when associating paleopathological signatures with contemporary disease, and extensive scholarship is devoted to this topic (see review[112]). Much of the impetus for developing accurate differential paleopathological diagnoses was fuelled by the ardent debate regarding the arrival of tuberculosis in the New World[113-116], further evidenced by numerous PCR-based identifications targeting small fragments of MTBC DNA, e.g. [117-120], although the exact identity and phylogenetic placement of these remained elusive.

Paleopathological and/or molecular analyses have reported hundreds of individuals from more than 200 sites worldwide with evidence of tuberculosis infection; these cases are synthesized in a number of recent reviews [1,112,121-123]. In the New World, the density of reported cases increases in both North and South America through the first millennium AD and until European contact, an observation that is made more profound by the extremely low percentage of skeletal involvement in clinically documented tuberculosis examples [124,125]. A small number of pre-Columbian examples have been identified in Colombia, Venezuela and western Mexico, but the majority of evidence for tuberculosis is found along the coast of modern-day Peru/Chile and in eastern and southwestern North America (see reviews cited above). The relative abundance of evidence in these areas could result from biases relating to burial/DNA preservation conditions (e.g. dry coastal deserts vs. tropics, differential burial practices), the presence of a well-established tradition of paleopathological work in those regions, or differential host responses or pathogen virulence affecting skeletal involvement [1,126]. As stated above, the three individuals for whom full mycobacterial genomes were reconstructed in this study were found at three archaeological sites along the extent of the Osmore River Valley. Of the Chiribaya-related sites (Chiribaya Alta, Chiribaya Baja, El Yaral, and San Geronimo), 19 individuals display evidence of tuberculosis infection, nearly 14% of all excavated human remains[1]. For this study, 5 of these 19 individuals were included in the screening (see Table S1), recovering two full tuberculosis genomes. At El

Algodonal, nine individuals are reported with tuberculosis; here, we screened one, and recovered one full genome. Further inland and occupying the late Intermediate/Terminal period are Estuquiña sites, distinct from the Chiribaya culture tradition. Most notably, at the type site of Estuquiña (~AD 1350), 37 cases of skeletal tuberculosis have been recorded, representing 37.5% of all males and 14.3% of all females found at the site[127]. We screened seven of these individuals; one sample was promising but did not meet our criteria for full genome capture.

Most morphological identifications consistent with tuberculosis in the New World are compatible with a scenario where an infectious agent acquired in the early first millennium AD on the central/southern coast of South America spread inland and northward, disseminating to North America by AD 900, possibly via sea trade routes. As Roberts and Buikstra[1] (p194) propose, human disease could have accompanied trade wares carried by seafarers from the northern coast of South America (modern-day Ecuador) along the western coast of the Mexico to North America, a route supported by archaeological evidence [128-130]. This seafaring model is consistent with the dearth of tuberculosis evidence in central Mesoamerica, and simultaneously accounts for the multiple pre-Columbian TB cases reported along the coast of Western Mexico.

Several of these New World cases, however, may predate our estimated date of *M. pinnipedii* transmission to human groups along the coast of Peru and Chile. For example, the earliest evidence for pre-Columbian tuberculosis in the South American paleopathological record describes mummified individuals from the Atacama desert of northern coastal Chile, c. AD 290, displaying cavitary pulmonary disease, a healed Ghon complex, and/or positive acid-fast bacilli tests, as well as one case from Paracas, Peru, AD 160 [131]. Additionally, cases in northern Chile occur c. AD 500 [117]. Isolated zoonotic events from pinnipeds along the South American coast that predate c. AD 700 would not be necessarily surprising given the region's longstanding reliance on pinnipeds for subsistence, and use of their remains in household or ritual contexts, discussed above. One outlier in skeletal material with signatures of tuberculosis in North America is an unpublished report from an Ipiutak site in Beringia (modern-day Point Hope, Alaska), c. 100 BCE-

AD 500[132]. The causative agent of the infection has not been investigated, but as stated in the main manuscript (lines 193-195), further analysis would be necessary to evaluate whether cases such as these represent an American-specific strain or an earlier introduction from a different source. Ultimately, successful isolation and sequencing of pathogen genomes from inland South American sites with reported cases of tuberculosis, e.g. the Amazonian site of Chachapoyas, AD 500-1000[133], or from the many cases in North America e.g. [118,134-137], are necessary to test the hypothesis that human adaptation of the Peruvian tuberculosis strain or a close relative resulted in the dissemination of *M. pinnipedii* throughout the New World via human movements.

References Cited

1. Roberts, C.A., Buikstra, J.E. The Bioarchaeology of Tuberculosis: A Global View on a Reemerging Disease, University Press of Florida, Gainesville, FL (2003).

2. Rohland, N., and Hofreiter M. Ancient DNA extraction from Bones and Teeth. *Nat Protoc* **2**, 1756 – 1762 (2007).

3. Meyer, M. and Kircher M. Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. *Cold Spring Harb Protoc* **6**; doi:10.1101/pdb.prot5448 (2010).

4. Kircher M., Sawyer S., and Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* **40**(1); doi:10.1093/nar/gkr771 (2012).

5. Maricic T., Whitten M., Pääbo S. Multiplexed DNA Sequence Capture of Mitochondrial Genomes Using PCR Products. *PLoS ONE* **5**(11):e14004, doi:10.1371/joutnal.pone.0014004 (2010).
6. Untergasser A., *et al*. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res* **35**(2),W71-W74 (2007).

7. Kircher, M., Stenzel U, and Kelso J. Improved base calling for the Illumina Genome Analyzer using machine learning strategies. *Genome Biol* **10**, R83 (2009).

8. Li, H. and Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics* **25**,1754-60 (2009).

9. Jónsson, H. *et al*. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics*, doi: 10.1093/bioinformatics/btt193 (2013).

10. Cole, S. T. et al. Deciphering the biololgy of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393**, 537-544 (1998).

11. Covert T.C., et al. Occurrence of Non-tuberculosis Mycobacteria in Environmental Samples. *Appl Environ Microbiol* **65**(6), 2492-2496 (1999).

12. Krause, J. *et al*. A Complete mtDNA Genome of an Early Modern Human from Kostenki, Russia. *Current Biology* **20**,231-236 (2010).

13. Sawyer, S. *et al*. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* **7**, e34131 (2012).

14. Lindahl, T. and Nyberg. B. Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry* **13**(16), 3405-3410 (1974).

15. Buikstra, J.E. Tombs for the living. . .or. . .for the dead: the Osmore ancestors. In: Dillehay, T.D. (Ed.), Tombs for the Living: Andean Mortuary Practices. Dumbarton Oaks, Washington, DC, 229–280 (1995).

16. Buikstra, J.E., *et al*. Chiribaya Political Economy: A Bioarchaeological Perspective, in: Rakita, G.F., Buikstra, J.E., Beck, L.A., Williams, S.R. (Eds.), Interacting with the dead: Perspectives on mortuary archaeology for the new millennium. University Press of Florida, Gainesville, 66-80 (1995).

17. Ghersi, H. Informe sobre las excavaciones en Chiribaya. *Revista del Museo Nactional* **25**, 89-119 (1956).

18. Owen, B.D. A model of multiethnicity: state collapse, competition, and social complexity from Tiwanaku to Chiribaya in the Osmore Valley, Peru. University of California Los Angeles (1993).

19. Martinson, E. et al. Pathoecology of Chiribaya parasitism. *Mem Inst Oswaldo Cruz* **98**(1),195-205 (2003).

20. Reinhard, K.J., and Buikstra, J.E. Louse infestation of the Chiribaya culture, southern Peru: variation in prevalence by age and sex. *Mem Inst Oswaldo Cruz* **98**(1),173-179 (2003).

21. Lozada, M.C., and Buikstra, J.E.. Pescadores and labradores among the señorío of Chiribaya in southern Peru. Us and Them: Archaeology and Ethnicity in the Andes, Cotsen Institute of Archaeology, University of California, Los Angeles, 206-225 (2005).

22. Lozada, M.C. The Señorío of Chiribaya: A Bio-archaeological Study in the Osmore Drainage of Southern Perú. University of Chicago, Department of Anthropology (1998).

23. Tomczak, P.D. Prehistoric Socio-Economic Relations and Population Organization in the Smore Valley of Southern Peru. University of New Mexico, Albuquerque (2001).

24. Sutter, R.C. Prehistoric Genetic and Culture Change: A Bioarcheological Search for Pre-Inka Altiplano Colonies in the Coastal Valleys of Moquegua, Peru, and Azapa, Chile. *Latin American Antiquity* **11**, 43-70 (2000).

25. Sutter, R. Post-Tiwanaku Ethnogenesis in the Coastal Osmore Valley: A Test of Competing Models for the Origins of the Moquegua Chiribaya Using Matrix Correlations, in: Knudson, K.J., Stojanowski, C.M. (Eds.), Bioarchaeology and Identity in the Americas, University Press of Florida (2009).

26. Sutter, R.C., and Sharratt, N. Continuity and Transformation During the Terminal Middle Horizon (AD 950-1150): A Bioarchaeological Assessment of Tumilaca Origins within the Middle Moquegua Valley, Peru. *Latin American Antiquity* **21**, 67-86 (2010).

27. Owen, B.D. Distant colonies and explosive collapse: The two stages of the Tiwanaku diaspora in the Osmore drainage*. Latin American Antiquity*, 45-80 (2005).

28. Goldstein, P. Andean Diaspora: The Tiwanaku Colonies and the Origins of Andean Empire. University Press of Florida, Gainesville (2005).

29. Minkes, W. Wrap the Dead: The Funerary Textile Tradition from the Osmore Valley, South Peru, and its Social-Political Implications, Archaeological Studies. Leiden University (2005).

30. Lozada, M.C., *et al*. Camelid herders: the forgotten specialists in the coastal senorio of Chiribaya, southern Peru. Andean Civilization: A Tribute to Michael E. Moseley Marcus J, Williams PR (eds). Cotsen Institute of Archaeology: Los Angeles, 351-364 (2009).

31. Rofes, J. Sacrificio de cuyes en El Yaral, comunidad prehispánica del extremo sur Peruano. *Bulletin Institue Francais d'Êtudes Andines* **29**, 1-12 (2000).

32. Thornton, E., *et al*. Isotopic evidence for Middle Horizon to 16th century camelid herding in the Osmore Valley, Peru. *Int J of Osteoarch* **21**, 544-567 (2011).

33. Wheeler, J.C., Russel, A., and Redden, H. Llamas and alpacas: pre-conquest breeds and post-conquest hybrids. *J Arch Sci* **22**, 833-840 (1995).

34. Tomczak, P. Prehistoric diet and socioeconomic relationships within the Osmore Valley of southern Peru. *J Anthropol Arch* **22**, 262-278 (2003).

35. Nystrom, K.C., and Malcom, C.M. Sex-Specific Phenotypic Variability and Social Organization in the Chiribaya of Southern Peru. *Latin American Antiquity* **21**, 375-397 (2010).
36. Stuiver, M., Reimer, P.J. Extended 14C database and revised CALIB radiocarbon calibration program. *Radiocarbon* **35**, 215-230 (1993).

37. Hogg, A.G, *et al*. SHCal13 Southern Hemisphere Calibration, 0–50,000 Years cal BP (2013).

38. Stuiver, M., Pearson, G.W., and Braziunas, T.F. Radiocarbon age calibration of marine samples back to 9000 cal yr BP. *Radiocarbon* **28**, 980-1021 (1986).

39. Stuiver, M., and Braziunas, T.F. Modeling atmospheric 14C influences and 14C ages of marine samples to 10,000 BC. *Radiocarbon* **36**, 137-189 (1993).

40. Yoneda, M., *et al*. Radiocarbon marine reservoir effect in human remains from the Kitakogane site, Hokkaido, Japan*, J Archaeol Sci* **29**, 529-536 (2002).

41. Owen, B.D. Marine carbon reservoir age estimates for the far south coast of Peru. *Radiocarbon* **44**, 701-708 (2002).

42. Sandness, K.L. Temporal and spatial dietary variability in the prehistoric lower and middle Osmore drainage: the carbon and nitrogen isotope evidence. University of Nebraska—Lincoln (1992).

43. Choi, W.-J., Ro, H.-M., and Hobbie, E.A. Patterns of natural 15N in soils and plants from chemically and organically fertilized uplands. *Soil Biology and Biochemistry* **35**, 1493-1500 (2003).

44. Poulson, S.R., *et al*. Paleodiet in northern Chile through the Holocene: extremely heavy δ15N values in dental calculus suggest a guano-derived signature? *J Archaeol Science* **40**(12), 4576-4585 (2013).

45. Szpak, P. *et a*l. Influence of Seabird guano and Camelid dung fertilization on the nitrogen isotopic composition of field-grown maize (*Zea mays*). J Archaeol Sci **39**(12), 3721-3740 (2012).

46. Julien, C. Guano and resource control in sixteenth-century Arequipa, in: Masuda, S., Shimada, I., Morris, C.P. (Eds.), Andean ecology and civilization, University of Tokyo Press, Tokyo, 185-231 (1985).

47. Thornton, E., *et al*. Isotopic evidence for Middle Horizon to 16th century camelid herding in the Osmore Valley, Peru. *International Journal of Osteoarchaeology* **21**, 544-567 (2011).

48. Bos, KI. Genetic investigations into the Black Death. PhD thesis: McMaster University (2011).

49. Darling A.C., *et al*. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* **14**, 1394-1403 (2004).

50. Briggs, A.W., *et al*. Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. *Nuc. Acids Res.***38**, e87 (2009).

51. Hodges, E., *et al*. Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. *Nature Protoc* **4**, 960 – 974 (2009).

52. Somerville, W., *et al*. Extraction of Mycobacterium tuberculosis DNA: a Question of Containment. *J. Clin. Microbiol.* **43**(6)**,** 2996-2997 (2005).

53. Schuenemann, V. J., *et al*. Genome-wide comparison of medieval and modern *Mycobacterium leprae*. Science **341**(6132),179-183.

54. Comas, I., M. Coscolla, *et al.* Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat Gen* **45**(10), 1176-1182 (2013).

55. Chan, J. Z., *et al*. Metagenomic analysis of tuberculosis in a mummy. *New England J Med* **369**(3), 289-290 (2013).

56. Domogalla, J., *et al*. Region of difference 4 in alpine Mycobacterium caprae isolates indicates three variants. *J Clin Microbiol* **51**(5), 1381-1388 (2013).

57. Coscolla, M., *et al*. Novel *Mycobacterium tuberculosis* Complex from a Wild Chimpanzee. *Emerg Infect Dis* **19**(6), 969-976 (2013).

58. Zumárraga M.J., *et al*. Molecular characterization of mycobacteria isolated from seals. *Microbiology* **145**, 2519-26 (1999).

59. Bernardelli, A., *et al*. Tuberculosis in sea lions and fur seals from the south-western Atlantic coast. *Rev Sci Tec Off Int Epiz* **15**(3), 985-1005 (1996).

60. Comas, I., *et al*. Human T-cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat Genet* **42**(6), 498-503 (2010).

61. Li, H, and Durbin, R. Fast and Accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**(14), 1754-1760 (2009).

62. Brosch, R. et al. A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proc Nat Acad Sci* **99**(6), 3684-3689.

63. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 (2013).

64. Tsolaki A. G. et al. Functional and evolutionary genomics of Mycobacterium tuberculosis: Insights from genomic deletions in 100 strains. *Proc Nat Acad Sci* **101**(14), 2865-4870 (2004).

65. Marmiesse, M., et al. Macro-array and bioniformatic analyses reveal mycobacterial 'core' genes, variation in the ESAT-6 gene family and new phylogenetic markers for the Mycobacterium tuberculosis complex. *Microbiology* **150**, 483-496 (2004).

66. DePristo, M., *et al*. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* **43**, 491-498 (2011).

67. Tamura, K., et al. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* **28**, 2731-2739 (2011).

68. Cingolani P, *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2;iso-3. Fly (Austin) **6**, 80–92 (2012).

69. Goldman, R.C. Why are membrane targets discovered by phenotypic screens and genome sequencing in *Mycobacterium tuberculosis*? *Tuberculosis (Edinb)* **93**,569-88 (2013).

70. Micklinghoff, J.C., *et al*. Role of the transcriptional regulator RamB (Rv0465c) in the control of the glyoxylate cycle in *Mycobacterium tuberculosis. J Bacteriol*. **191**:7260-9 (2009).

71. Mattle, D., *et al*. On allosteric modulation of P-type Cu^{+} -ATPases*. J Mol Biol* **425**, 2299-2308 (2013).

72. Porcheron, G., *et al*. Iron, copper, zinc, and manganese transport and regulation in pathogenic Enterobacteria: correlations between strains, site of infection and the relative importance of the different metal transport systems for virulence. *Front Cell Infect Microbiol* **3**, 90 (2013).

73. Botella, H., *et al*. Metallobiology of host-pathogen interactions: an intoxicating new insight. *Trends Microbiol* **20**, 106–12 (2012).

74. Pond, S.L. and Frost, S.D. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* **21**(10), 2531-3 (2005).

75. Drummond, A. J. and A. Rambaut. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC evolutionary biology* **7**, 214 (2007).

76. Rambaut, A. Drummond, A. J. Tracer, version 1.4. *Computer program and documentation distributed by the author, http://beast.bio.ed.ac.uk/Tracer* (2007).

77. Casali, N., V., *et al*. Microevolution of extensively drug-resistant tuberculosis in Russia. *Genome Res* **22**(4), 735-745 (2012).

78. Ponstingl, H. "SMALT v0.5.8." Retrieved 5/1, 2012, from http://www.sanger.ac.uk/resources/software/smalt/ (2011).

79. Bryant, J. M., A. C. Schurch, *et al*. Inferring patient to patient transmission of *Mycobacterium tuberculosis* from whole genome sequencing data. *BMC infectious diseases* **13**(1), 110 (2013).

80. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**(21), 2688-2690 (2006).

81. Farris. Methods for computing Wagner Trees. *Syst Zool* **19**, 83–92 (1970).

82. Walker, T. M., *et al*. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. *Lancet infect Dis* **13**(2), 137-146 (2013).

83. Dillehay, T.D., *et al*. A late Pleistocene human presence at Huaca Prieta, Peru, and early Pacific Coastal adaptations. *Quaternary Research* **77**(3), 418-423 (2012).

84. DeFrance, S.D., and Alvarez, A.U. Quebrada Tacahuay: un sitio marítimo del Pleistoceno Tardío en la costa sur del Perú, Chungara. *Revista de Antropología Chilena* **36**, 257-278 (2004).

85. Orquera, L.A. Mid-Holocene littoral adaptation at the southern end of South America. *Quaternary International* **132**, 107-115 (2005).

86. Orquera, L.A., Legoupil, D., and Piana, E.L. Littoral adaptation at the southern end of South America. *Quaternary International* **239**, 61-69 (2011).

87. Schiavini, A. Los lobos marinos como recurso para cazadores-recolectores marinos: el caso de Tierra del Fuego. *Latin American Antiquity*, 346-366 (1993).

88. Muñoz, A.S. Taphonomic properties of pinniped bone assemblages from the Atlantic coast of Isla Grande de Tierra del Fuego, Argentina. *Geobios* **41**, 113-122 (2008).

89. Arriaza, B. Preparation of the dead in coastal Andean preceramic populations, Human Mummies. Springer, 131-140 (1996).

90. Arriaza, B.T., and Standen, V.G. Differential Mortuary Treatment among the Andean Chinchorro Fishers: Social Inequalities or In Situ Regional Cultural Evolution? *Curr Anthropol* **46**, 662-671 (2005).

91. Masur LJ. Peanuts and prestige on the Peruvian North Coast: the archaeology of peanuts at Huaca Gallinazo (V-59) and Huaca Santa Clara (V-67). MA thesis: University of British Columbia (2008).

92. Swenson, E.R. Competitive Feasting, Religious Pluralism, and Decentralized Power in the Late Moche Period, in: Isbell, W.H., Silverman, H. (Eds.), Andean Archaeology III: North and South, Springer/Plenum Press, New York, 112-142 (2006).

93. Donnan, C.B. Moche Art of Peru: Pre-Columbian Symbolic Communication. Museum of Cultural History, University of California (1978).

94. Benson, E.P. The Worlds of the Moche on the North Coast of Peru. University of Texas Press (2012).

95. Ferreira L.F., *et al.* L. The finding of eggs of Diphyllobothrium in human coprolites (4,100-1,950 B.C.) from northern Chile. Mem Inst Oswaldo Cruz 79, 175-180 (1984).

96. Aufderheide, A.C., *et al*.. Pulmonary disease in a sample of mummies from the az-75 cemetery in northern Chile's Azapa Valley. Chungara Revista de Antropología Chilena **34**, 253-263 (2002).

97. Arriaza, B.T. *et al.* Possible influence of the ENSO phenomenon on the pathoecology of diphyllobothriasis and anisakiasis in ancient Chinchorro populations. Mem Inst Oswaldo Cruz **105**, 66-72 (2010).

98. Callen, E.O. & Cameron, T.W.M. A prehistoric diet revealed in coprolites. New Scientist 8, 35-**40** (1960).

99. Patrucco, R., *et al.* Parasitological Studies of Coprolites of Pre-Hispanic Peruvian Populations. Current Anthropology **24**(3), 393-394 (1983).

100. Reinhard, K. & Urban, O. Diagnosing ancient diphyllobothriasis from Chinchorro mummies. Mem Inst Oswaldo Cruz 98 Suppl 1, 191-193 (2003).

101. Martinson, E., *et al*. Pathoecology of Chiribaya parasitism. Mem Inst Oswaldo Cruz **98** Suppl 1, 195-205 (2003).

102. Holiday, D.M., *et al*. Diphyllobothriasis of the Chiribaya Culture (700–1476 AD) of Southern Peru. Comparative Parasitology 70, 167-171 (2003).

103. Bibar, J. Fondo Histórico y Bibliográfico Jos Toribio Medina, Santiago, 1565 (1966).

104. Santoro, C.M., *et al*. People of the Coastal Atacama Desert: Living Between Sand Dunes and Waves of the Pacific Ocean, in Veth, P., Smith, M. Hiscock, P. (Eds.), Desert Peoples: Archaeological Perspectives, Blackwell Publishing, 243-260 (2005).

105. Bastida, R., *et al*. La Tuberculosis en grupos de Cazadores recolectores de Patagonia y Tierra del Fuego: nuevas alternativas de contagio a traves de la fauna silvestre. Revista Argentina de Antropología Biológica. 13 (1): 83-95 (2011)

106. Aufderheide, A. C. & Rodríguez-Martín, C. The Cambridge encyclopedia of human paleopathology. (Cambridge University Press, 1998).

107. Ortner, D. Identification of pathological conditions in human skeletal remains. (Academic Press, 2003).

108. Kelley, M. A. & Micozzi, M. S. Rib lesions in chronic pulmonary tuberculosis. Am J Phys Anthropol 65, 381-386, doi:10.1002/ajpa.1330650407 (1984).

109. Roberts, C., Lucy, D. & Manchester, K. Inflammatory lesions of ribs: an analysis of the Terry Collection. American Journal of Physical Anthropology 95, 169-182 (1994).

110. Mays, S., Fysh, E. & Taylor, G. M. Investigation of the link between visceral surface rib lesions and tuberculosis in a medieval skeletal series from England using ancient DNA. American Journal of Physical Anthropology 119, 27-36, doi:Doi 10.1002/Ajpa.10099 (2002).

111. Matos, V. & Santos, A. L. On the trail of pulmonary tuberculosis based on rib lesions: results from the Human Identified Skeletal Collection from the Museu Bocage (Lisbon, Portugal). Am J Phys Anthropol 130, 190-200, doi:10.1002/ajpa.20309 (2006).

112. Grauer, A. L. A Companion to Paleopathology. (John Wiley & Sons, Malden, MA, 2011).

113. Hrdlička, A. Tuberculosis among certain Indian tribes of the United States. (Smithsonian Institution, Bureau of American Ethnology, 1909).

114. Morse, D. Prehistoric tuberculosis in America. The American review of respiratory disease 83, 489 (1961).

115. Buikstra, J. E. Differential Diagnosis: An Epidemiological Model. *Yearbook of Physical Anthropology [1976]*, 316-328 (1977).

116. Cockburn, A. The evolution and eradication of Infectious diseases. (Cambridge Univ Press, 1963).

117. Arriaza, B. T., Salo, W., Aufderheide, A. C. & Holcomb, T. A. Pre-Columbian tuberculosis in northern Chile: molecular and skeletal evidence. American Journal of Physical Anthropology 98, 37-45, doi:10.1002/ajpa.1330980104 (1995).

118. Braun, M., Cook, D. C. & Pfeiffer, S. DNA from Mycobacterium tuberculosis complex identified in North American, Pre-Columbian human skeletal remains. Journal of Archaeological Science 25 (1998).

119. Raff, J., Cook, D. C. & Kaestle, F. Tuberculosis in the New World: a study of ribs from the Schild Mississippian population, West-Central Illinois. Mem Inst Oswaldo Cruz 101 Suppl 2, 25-27 (2006).

120. Salo, W. L., Aufderheide, A. C., Buikstra, J. E. & Holcomb, T. A. Identification of Mycobacterium tuberculosis DNA in a Pre-Columbian Peruvian Mummy. PNAS 91, 2091-2094 (1994).

121. Galagan, J. E. Genomic insights into tuberculosis. Nature Reviews Genetics Early View, doi:http://dx.doi.org/10.1038/nrg3664 (2014).

122. Holloway, K. L., Henneberg, R. J., de Barros Lopes, M. & Henneberg, M. Evolution of human tuberculosis: a systematic review and meta-analysis of paleopathological evidence. Homo : internationale Zeitschrift fur die vergleichende Forschung am Menschen 62, 402-458, doi:10.1016/j.jchb.2011.10.001 (2011).

123. Stone, A. C., Wilbur, A. K., Buikstra, J. E. & Roberts, C. A. Tuberculosis and leprosy in perspective. American Journal of Physical Anthropology 140 Suppl 49, 66-94, doi:10.1002/ajpa.21185 (2009).

124. Wilbur, A. K. & Buikstra, J. E. Patterns of tuberculosis in the Americas: how can modern biomedicine inform the ancient past? Memórias do Instituto Oswaldo Cruz 101 Suppl 59-66 (2006).

125. Lombardi, G. P. & Cáceres, U. G. Multisystemic Tuberculosis in a Pre-Columbian Peruvian Mummy: Four Dagnostic levels, and a paleoepidemiological hypothesis. Chungará (Arica) 32, 55-60 (2000).

126. Wilbur, a. K., Farnbach, a. W., Knudson, K. J. & Buikstra, J. E. Diet, Tuberculosis, and the Paleopathological Record. Current Anthropology 49, 963-991 (2008).

127. Buikstra, J. E. & Williams, S. R. in *Human paleopathology: Current syntheses and future options.* (eds Donald Ortner & A. C. Aufderheide) 161–172 (Smithsonian Institution Press, Washington, D.C., 1991).

128. Hosler, D. Ancient West Mexican Metallurgy: South and Central American Origins and West Mexican Transformations. American Anthropologist 90, doi:doi: 10.1525/aa.1988.90.4.02a00040 (1988).

129. Paulsen, A. C. in The Sea in the Pre-Columbian World (ed E. P. Benson) (Dumbarton Oaks, 1977).

130. Anawalt, P. Ancient cultural contacts between Ecuador, West Mexico, and the American southwest. Latin American Antiquity 3, 114–129. (1992).

131. Allison, M., Gerzten, E., Munizaga, J., Santoro, C. & Mendoza, D. in Prehistoric tuberculosis in the Americas (ed Jane E. Buikstra) 49-62 (Northwestern University Archeological Program, 1981).

132. Dabbs, G. R. Resuscitating the epidemiological model of differential diagnosis: tuberculosis at prehistoric Point Hope, Alaska. *Paleopathology Association Newsletter* **148**, 11-24 (2009).

133. Friedrich, K. M. et al. The story of 12 Chachapoyan mummies through multidetector computed tomography. European journal of radiology 76, 143-150, doi:10.1016/j.ejrad.2009.07.009 (2010).

134. Buikstra, J. E. & Cook, D. C. in Prehistoric tuberculosis in the Americas (ed Jane E. Buikstra) 115-139 (Northwestern University Archaeological Program, 1981).

135. Pfeiffer, S. Paleopathology in an Iroquoian ossuary, with special reference to tuberculosis. American Journal of Physical Anthropology 65, 181-189 (1984).

136. El-Najjar, M. Y. Human treponematosis and tuberculosis: evidence from the New World. American Journal of Physical Anthropology 51, 599-618, doi:10.1002/ajpa.1330510412 (1979).

137. Milner, G. R. & Smith, V. G. Oneota human skeletal remains. Archaeological Investigations at the Morton Village and Norris Farms 36, 111-148 (1990).

All Extended Data Figures 1-13 in Appendix B

Table S1 – Appendix B

Table S2 – Long range PCR primers used for *M. tuberculosis* bait generation.

Primer	Sequence	Tm (°C)	Amplicon Length (bp)
rpoB_forward	GAGCAAAACAGCCGCTAGTC	60.16	3537
rpoB_reverse	ACGCAAGATCCTCGACACTT	59.87	
katG_forward	GAGCAACACCCACCCATTAC	60.24	2223
katG_reverse	ACCTGTCGAGGTTCATCACC	59.97	
mpt40_forward	ATAGGGAATGCTCGGCAAC	60.06	426
mpt40_reverse	CAACATCGACGCAGTACCC	60.13	
gyrA_forward	TTGACATCGAGCAGGAGATG	59.94	2368
gyrA_reverse	CCCAGATTCATCAACCGAAC	60.32	
gyrB_forward	AAAGAAGGCCCAAGACGAAT	60.07	2145
gyrB_reverse	ACATCCAGGAACCGAACATC	59.97	

Table S3 – radiocarbon dates for the samples used in *M. tuberculosis* genome capture.

Site	Sample	Burial	Tissue type	15N/ 14N	13C/ 12C	Dates uncal BP	Dates cal AD	Cal program	Southern Hemisphere cal program
Chiribaya Alta	54	103	bone	-13.5	17.9	940 +-30	AD 1105-1214 (.7 prob dist), AD 1045-1100 (.3 prob dist)	CALIB	ShCal13
							AD 1028-1224 (99.7%)	OxCal4.2	ShCal13
El Yaral	64	205	bone	-14.6	14.1	880 +- 30	AD 1158-1266 (prob dist=1)	CALIB	ShCal13
							AD 1141-1280 (98.5% prob)	OxCal4.2	ShCal13
El Algodonal	58	386	bone	-13.4	13.1	970 +-30	AD 1023- 1211 (99.7% prob)	OxCal 4.2	ShCal13
							AD 1031-1179 (prob dist=1)	CALIB	ShCal13

Table S4 – Publicly available genomes used for array probe design.

Accession Number	Sequence Name
NC_015758.1	*Mycobacterium africanum* GM041182 chromosome
NC_002945.3	*Mycobacterium bovis* AF2122/97 chromosome
NC_008769.1	*Mycobacterium bovis BCG* str. Pasteur 1173P2 chromosome
NC_012207.1	*Mycobacterium bovis BCG* str. Tokyo 172
NC_016804.1	*Mycobacterium bovis BCG* str. Mexico chromosome
NC_015848.1	*Mycobacterium canettii* CIPT 140010059
NC_016768.1	*Mycobacterium tuberculosis* KZN 4207 chromosome
NC_012943.1	*Mycobacterium tuberculosis* KZN 1435
NC_009565.1	*Mycobacterium tuberculosis* F11
NZ_CM000789.2	*Mycobacterium tuberculosis* KZN R506 chromosome
NZ_CM000788.2	*Mycobacterium tuberculosis* KZN V2475 chromosome
NC_017026.1	*Mycobacterium tuberculosis* RGTB327 chromosome
NC_017528.1	*Mycobacterium tuberculosis* RGTB423 chromosome
NC_017523.1	*Mycobacterium tuberculosis* CCDC5079 chromosome
NC_017522.1	*Mycobacterium tuberculosis* CCDC5180 chromosome
NC_002755.2	*Mycobacterium tuberculosis* CDC1551 chromosome
NC_017524.1	*Mycobacterium tuberculosis* CTRI-2 chromosome
NC_009525.1	*Mycobacterium tuberculosis* H37Ra
NC_000962.2	*Mycobacterium tuberculosis* H37Rv chromosome
NZ_CM000787.2	*Mycobacterium tuberculosis* KZN 4207 chromosome

Table S5 – Appendix B

Table S6 – Appendix B

Table S7 – Coordinates and numbers of spanning reads for all Regions of Difference detected in the three Peruvian samples.

			# of spanning reads		
	start	**end**	**Sample 54**	**Sample 58**	**Sample 64**
RD7	2208006	2220724	3	4	4
RD8	4056841	4062734	25	14	10
RD9	2330074	2332101	8	6	7
RD10	264755	266656	24	18	11
RDseal	2220948	2222888	4	3	7

Table S8 – Appendix B
Table S9 – Appendix B

Table S10 - Bayesian identification of sites under positive selection according to model M8 from the *codeml* package.

Codon position with potential dN/dS > 1	BEB Pr of dN/dS > 1	dN/dS +- SE
3	0.516	5.357 +- 4.917
62	0.999***	10.092 +- 0.761
144	0.759	7.757 +- 4.226
156	0.609	6.280 +- 4.807
254	0.676	6.938 +- 4.616
382	0.788	8.041 +- 4.043
413	0.506	5.264 +- 4.919

Table S11 - Estimated ages for Hungarian ancient TB samples. Present day assumes year 2000 as this is the average isolation date for the "modern" strains

Sample bound (No. variants)	Years before present (0.3 SNPs/genome/year)	Years before present (0.5 SNPs/genome/year)
Hungarian 1 Lower (4.5)	225	211
Hungarian 1 Upper (60.5)	403.66	324
Hungarian 2 Lower (91.5)	709	385
Hungarian 2 Upper (147.5)	895.6	497

CHAPTER 4

PHYLOGENOMIC RECONSTRUCTION SUPPORTS SUPERCONTINENT ORIGINS FOR *LEISHMANIA*

Kelly M. Harkins[1], Rachel S. Schwartz[2], Reed Cartwright[2], and Anne C. Stone[1]

[1]School of Human Evolution and Social Change, Arizona State University
[2]Center for Evolutionary Medicine. Biodesign, Arizona State University

Among infectious diseases, human leishmaniasis is estimated to cause the ninth largest disease burden. Endemic to poverty-stricken countries and rising in incidence by nearly 2 million cases annually, it lacks effective treatment or vaccine (Alvar et al., 2012). Over twenty *Leishmania* species cause disease in humans with three main clinical manifestations: visceral, cutaneous and mucocutaneous. The severity and susceptibility to *Leishmania* infection relies not only on the infecting species, but also on complex interactions between host-pathogen genetics and immune response (Lipoldova and Demant, 2006). There remains no predictable association, however, between species genotypes and their role in disease development (Lukes et al., 2007), nor is there a strict correlation with geographic location. Consequently, molecular-based research on the parasite has long focused on identifying loci that enable rapid discrimination among human pathogenic species (e.g. Boite et al., 2012; Marfurt et al., 2003; Weirather et al., 2011), determining function of genes under selective pressures (e.g. Peacock et al., 2007) and examining population structure (e.g. Schönian et al., 2011) to better anticipate clinical outcome and target drug/vaccine candidates. Such molecular data are also critical for phylogenetic reconstructions that characterize intra- and interspecific diversity and are key to understanding the evolution of *Leishmania*, namely to inform geographic origins (Croan et al., 1997; Fraga et al., 2010; Momen and Cupolillo, 2000; Noyes et al., 2000), estimate timing of divergence events (Lukes et al., 2007; Stevens and Rambaut, 2001), and reconsider the taxonomy of the parasite (Figure 1), which has been in continuous revision since its first description in 1903 (Cupolillo et al., 2000; Lainson, 2010; Noyes et al., 2002; Schönian et al., 2010; Van der Auwera et al., 2011).

Phylogenetic analyses have been unsuccessful, however, in reaching a consensus on the overall evolutionary history of Leishmania. This lack of consensus is due to the incongruence of molecular and non-molecular data (Kerr, 2000, 2006; Lysenko, 1971). Despite the increasing availability of sequence data, including recent draft genomes (Ivens et al., 2005; Peacock et al., 2007; Raymond et al., 2011; Real et al., 2013), the majority of molecular phylogenetic studies rely on one or few genetic loci to infer evolutionary relationships (see Schönian 2012). These gene trees may not accurately reflect the true history of the genus (Castresana, 2007). Unclassified *Leishmania* species with uncertain phylogenetic or taxonomic position are often represented by the fewest data, but are likely informative for resolving branching order and species relationships. For example, several of these unclassified *Leishmania* species (e.g. *L. hertigi, L. hererri, L. deanei, L. enrietti, L.adleri, L. equatorensis, L. columbiensis*) are not considered human pathogens, having shared millions of years of co-evolution with distinct reptilian and mammalian reservoirs (e.g. sloths, porcupines, guinea pigs). However, the emergence of HIV co-infection in the 1980s and its associated increase in mortality (Cruz et al., 2006) has highlighted a major global health problem: opportunistic infections with “non-pathogenic”, distantly-related *Leishmania* parasites. For example, HIV-*Leishmania* co-infection with species closely related to the guinea pig form, *L. enriettii*, have been reported from the Martinique Islands (Noyes et al. 2002), and more recently in Southeast Asia (Bualert et al., 2012). Other obscure and/or recently described species, e.g. *Leishmania sp.* AM-2004 in Australia (Dougall et al. 2011), and *Leishmania siemanisis* in the US, Europe and Asia (Reuss et al., 2012), are ambiguously classified and confound our understanding of *Leishmania* biogeographic distribution, raising questions about the emergence of mechanisms enabling host jumps causing human infection, and prompting controversy regarding the origins and dispersal of the parasite, already a longstanding source of debate.

Recent improvements in sequencing allow us to obtain unprecented, large numbers of molecular markers (O'Neill et al., 2013) for phylogenetic analyses. Phylogenetic methods, however, were not initially developed to accommodate such extensive datasets, leading to substantial challenges

in the interpretation of statistical support for tree reconstructions (Kumar et al., 2012). Divergent taxa cannot be easily aligned to reference genomes. The uncertainties and computational cost of de-novo assembly required us to find an alternative method to obtain phylogenetically informative sites from whole genome data. Thus we must build upon traditional techniques of multi-gene phylogenies together with novel bioinformatics methods to access informative sites in whole genome data. This study represents the first phylogenomic analysis of *Leishmania*, employing nearly 400,000 variable sites and 43 genes from across the genome of over 20 species and related taxa. We discover that the divergence of the *Leishmania* lineage in Australia is not a recent or anthropogenic introduction but likely an ancient event. This observation supports at minimum a southern supercontinent origin of *Leishmania* before the separation of Australia and Antarctica 40-50 million years ago (Flynn and Wyss, 1998). Our future analyses aim to identify the effects among-lineage rate variation and to find constant rate sites that will faciliate use of molecular clock methods (e.g. *BEAST (Heled and Drummond, 2010) and RelTime (Tamura et al., 2012)) to empirically relate the number of accumulated nucleotide substitutions to divergence times.

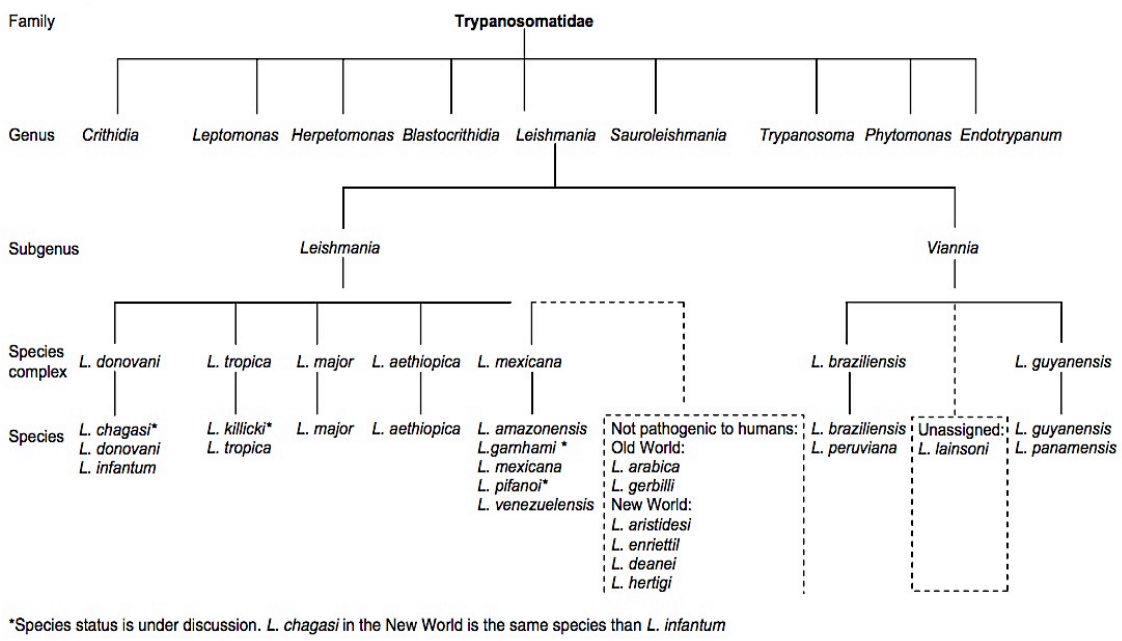


Figure 1: Leishmania taxonomy, including species status. World Health Organization 2010.

Background

Taxonomy and biogeography

The early taxonomy of *Leishmania* was inferred from morphological and life cycle data, namely the location of development in the sand fly vector (Lainson and Shaw, 1972). Based on the distinction between mid-foregut and hindgut development, respectively, Lainson and Shaw (1987) suggested a division into two subgenera: *L. (Leishmania),* consisting of all Old World (OW) species and at least one species complex found in the New World (NW)[1]; and *L. (Viannia)*, comprising exclusively NW species. Reptile-infecting *Leishmania* sp., formerly assigned to a separate genus (Lainson and Shaw), are now often referred to as a third subgenus, "Sauroleishmania", as proposed by (Saf'janova, 1982). Cupolillo et al. (2000) subsequently proposed that these three subgenera be called Euleishmania, and all other unclassified species that are restricted to New World mammals be called Paraleishmania. In this schema, *Endotrypanum* would be considered a member of a polyphyletic Paraleishmania. Marcili et al. (2014) recently suggest that the unclassified Paraleishmania species instead be considered part of the *Endotrypanum* genus, but this proposal has not been debated in depth and may not be supported with molecular data. The enigmatic *L. enrietti*, once assumed to be a member of the NW Leishmania subgenus, is positioned at the root of the Euleishmania section in sequence-based trees (Dougall et al. 2011; Noyes et al. 1997). The recent discovery that a *Leishmania* species most closely related to *L. enrietti* is found in naturally-infected Australian mammals (Dougall et al.; Rose et al.) will play a key role in our understanding of *Leishmania* biogeography and the parasite's ancient global dispersal.

[1] Phylogenetic investigations of *Leishmania infantum* genetic diversity recently enabled scholars to confirm

Proposed origins hypotheses

The origin of *Leishmania* is disputed: Neotropic, Palaearctic, and independent origins of *(L.)Sauro/Leishmania* sp. and Paraleishmania*/(L.)Viannia* sp. following the split of Africa and South America have all been proposed (see reviews Schönian et al 2013.; Tuon et al. 2008). The Palaearctic hypothesis assumes an origin in Cretaceous reptiles with later migrations to the Nearctic and Neotropics; this hypothesis is only supported by non-molecular data, namely host-phylogenies, biogeography, and evidence of an ancestral "*Paleoleishmania*" parasite found fossilized in Cretaceous amber in Burma (Kerr, 2000; Kerr, 2006; Kerr et al., 2000; Lysenko, 1971; Poinar Jr, 2007).

Rooted sequence-based phylogenies support a Neotropical origin of *Leishmania* (Croan et al., 1997; Cupolillo et al., 2000; Fraga et al., 2010; Noyes et al., 2000; Stevens et al., 2001) (Figure 2). The position of the subgenus *Sauroleishmania*, which only infects Old World reptiles, in the middle rather than at the root of molecular-based trees revealed that reptilian forms of *Leishmania* were surprisingly not ancestral but derived from mammalian forms of the parasite. However, this result requires either two separate intercontinental migrations to the Old World, both either at a) timescales that also coincide with appropriate land bridges and reservoir radiation or, b) the emergence of *Sauroleishmania* in the New World before migration to Asia followed by extinction thereafter (Figure 3).

Finally, the multiple origins hypothesis (Momen and Cupolillo, 2000) implies cladogenesis of the Paraleishmania and Euleishmania on Gondwana, where continental split c. 100 mya separated the MRCA of *Leishmania* and *Viannia* subgenera in Africa and South America, respectively. It is not explicitly stated that *Sauroleishmania*, as a subgenus, would also have needed to separate in Africa. The amount of diversity observed between the *Leishmania* and Viannia subgenera has been cited as evidence for vicariance due to the separation of Africa from the Neotropics (Fernandes et al., 1993). Momen and Cupollilo (2000) propose three ancient introductions to the Neotropics that must accommodate movement of New World *L. (Leishmania)* from Africa through

Eurasia to North America. Due to the sandfly's short lifecycle and weak flying ability, dispersion of the parasite must be related to conditions conducive to vector survival, and the migration of hosts and reservoirs, thus constraining *Leishmania* movement throughout history (Tuon et al., 2008). The most recent date of an introduction of *Leishmania* subgenus into the Neartic via Beringia would have been the mid-Miocene (26-13 mya [sic]) when temperatures were warm enough for sandfly survival (Stevens e al. 2001).

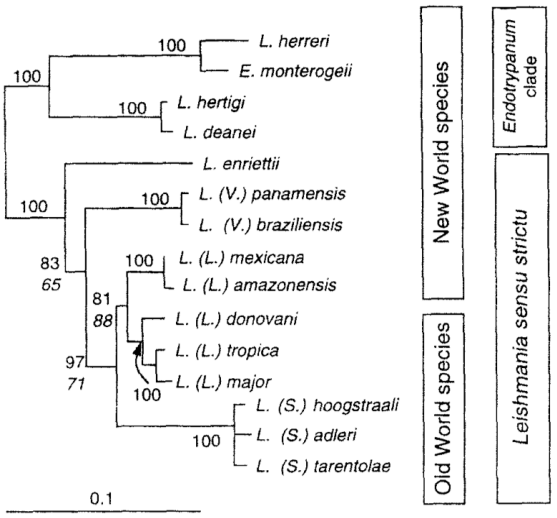


Figure 2: This published maximum likelihood tree constructed with RNA and DNA polymerase genes is used to support Neotropical origins hypothesis (Figure 7, (Stevens et al., 2001)).

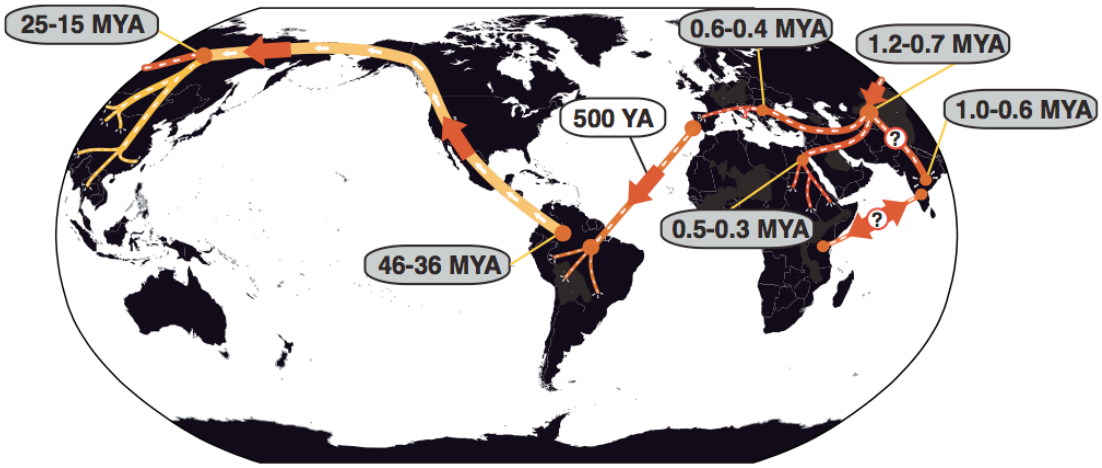


Figure 3: Divergence times associated with proposed Neotropical origins scenario (Lukes et al. 2007).

Slight derivations in timelines and terminology within different presentations of the same origins hypotheses produce ambiguities for testing the proposed sequence of events, i.e. Neotropical genus origins in the Paleocene or in the Eocene (Lukes et al. 2007; Noyes 2000; Noyes et al. 1997). For this reason, the scenario depicted in Figure 3 will represent the Neotropical origins hypothesis. It is possible that one origins scenario is supported by multiple topologies, or vice versa. It is implicit however that the hypotheses rely on the timescale at which these events occurred, thus corroboration of biogeographical events and the plausibility of relative branch lengths and divergence times at nodes throughout the entire tree are crucial to evaluate these proposed hypotheses or pose alternative scenarios. A reworking of the origins hypothesis may be in order since the discovery and description of *Leishmania* in Australia, calling into question the antiquity of that lineage and its associated global dispersal.

Materials and Methods

Raw data

We sequenced isolates of twelve *Leishmania* and one *Endotrypanum*. The genomic DNA from the majority of isolates was extracted in 2011 by KH, acquired from the laboratory of Dr. McMahon-Pratt (Yale School of Public Health); three strains were received from the lab of Dr. Lucile Floeter-Winter (University Sao Paolo). See Table 1.

Leishmania strains were grown as promastigote-stage organisms (Wirth and Pratt). Briefly, promastigotes were grown at 22°C in Schneider's Medium supplemented with 20% heat inactivated FCS and 17.5 mg/mL gentamycin. The cells were pelleted and washed twice in phosphate-buffered saline (PBS) and gDNA extracted according to established methods. The cells were incubated in lysis buffer (100mM NaCL, 10mM TrisCl pH 8.0, 0.5mM EDTA, 0.5% SDS, 0.1mg/ml fresh Proteinase K) for 12-18hrs at 50C, and purified via phenol/chloroform/isopropyl alcohol. Extract was incubated in RNase A (10mg/ml) for 2.5hrs at 37C and dialyzed overnight at 4C with 3 changes of PBS buffer. DNA concentration was

evaluated with spectrophotometry (Beckman Coulter DU730) and re-extracted if contaminated with phenol. Paired-end reads (100 bp) were sequenced on an Illumina HiSeq2000 by the University of Arizona Genome Core. The number of pairs of reads for each species ranged from 2.1-14.3 million, or 10 - 67x average coverage for the ~34 million bp genome.

Additional shotgun genomic data were downloaded from the European Nucleotide Archive (ENA; Table S1). These data represent all publicly available genome data per *Leishmania* taxon.

Phylogenetically informative data

The caveats of generating species trees from multiple genes are better understood than for variable site data. Although gene trees are not expected always to reflect the true species tree, the parameters guiding gene evolution are also more readily modeled. If rates are unequal, adding more data rather than more taxa is shown to help resolve inconsistencies (Poe and Swofford, 1999). Our goal is to reconstruct species trees with whole genome data and a comparison to more traditional gene data. Thus we obtained two datasets – one from across the genome and another of known genes (Table 2).

Known genes

Additionally, 42 genes with putative or known function were extracted from the shotgun sequencing reads using the following pipeline: (1) Reads from each species were aligned to the reference genes using Bowtie2 (Langmead and Salzberg). (2) The base for each site was identified based on whether at least 80% of the reads at that site contained a single base. (3) Alignments were adjusted using MAFFT with default settings. The reference for each gene sequence was downloaded from the *L.V. braziliensis* MHOM/BR/75/M2904 or M2903 reference genome.

We also obtained the RNA polymerase II (RPolII) gene sequence for 38 taxa from NCBI and ENA, including *Leishmania* strains AM-2004, and MHOM/MQ/92/MAR1. As RPolII is often the

only representative of many obscure or recently described species/isolates, it provides the highest taxon density for any single gene phylogeny in our study (29 species). Additionally, the hsp70 gene sequence was downloaded from NCBI for the recently described *L. siamensis* isolate CU1 (Kanjanopas et al., 2013).

Variable sites from across the genome

To extract phylogenetically informative sites from the available shotgun sequences, we used a bioinformatics pipeline, Site Identification from Short Read Sequences (SISRS) (Schwartz et al., in review). This method identifies sites that are fixed for each species and variable across species to construct a multiple species alignment. This approach is advantageous for new *Leishmania* genome data, as it requires neither *a priori* knowledge of homologous regions nor a reference genome. We produced a dataset with up to 6 taxa missing per site (missing data were allowed to ensure sufficient data to determine the phylogeny). To ensure that linked sites would not bias our results, we subsampled this dataset by sampling only one site per linkage region. We produced an additional alignment with no missing data allowed, but with a lower calling threshold of 80%. We called these datasets VS_m6, VS_m6s, and VS_t80.

Phylogenetic analysis

Concatenated gene analysis

Gene sequences were first analyzed as a concatenated alignment with a total of 73819 sites. The best way to partition the data was determined by the program *partitionfinder* (Lanfear et al., 2012), using the RAxML option to estimate likelihoods, which finds the best-fit partitioning scheme for the nucleotide alignment and the best-fit model of rate evolution for each partition. For coding genes, partitioning by codon position is found to provide a better fit than by gene alone (Mueller et al., 2004) Input partitions were separate codon positions for each gene. We employed the Bayesian Information Criterion for model selection to avoid overparameterisation. The resulting twenty-one partitions were used to estimate the phylogeny in a ML framework; this

analysis was implemented in RAxML 8.0.20 (Stamatakis et al. 2014) using the GTRCAT approximation, which allows rate heterogeneity across sites, for 1000 bootstrap replicates. This analysis was repeated five times with random starting seeds to identify the ML tree and avoid local optima.

First and second codon positions, or sites that are conserved, can be highly variable, as opposed to the third position where rates are more homogenous (Yang, 1996). However, saturation may underestimate the substitutions between highly divergent species for third positions (Nei and Kumar, 2000). We constructed a separate ML phylogeny in RAxML, with the parameters listed above, from the first, second and third codon positions, to evaluate differences in phylogenetic signal from variable evolutionary rate across the positions. Branch lengths were estimated for the concatenated dataset with RAxML on each phylogeny.

Individual gene analysis

ML trees for each gene were constructed identically in RAxML for 1000 bootstrap replicates. Although concatenated multi-gene datasets are found to be robust in phylogenetic inference, even without explicitly accounting for large variations in rates, lengths and GC content, systematic biases in the dataset can lead to high support of the wrong tree (Gadagkar et al., 2005). We follow the authors' recommendation to report the gene support frequency for each given partition. A majority rule (MR) consensus tree was generated in RAxML from the best scoring ML trees of 42 individual gene trees for which we had all ingroup taxa. When necessary the outgroup *Crithidia* was pruned using Newick Utilites (Junier and Zdobnov, 2010).

The RNA polymerase II (RPolII) gene and heat shock protein 70 (hsp-70) genes allow us to analyze more species than from our genome data alone. The purpose of these gene trees is to evaluate whether or not the distribution of species closely related to the New World guinea-pig form, *L. enrietti*, likely represent an anthropogenic introduction or ancient migratory events. RPOlII is the only locus where inclusion of the Australian *Leishmania sp.* AM-2004 was possible.

Of the handful of genes available for *L. siamensis*, we chose to consider hsp-70, which we extracted from raw data for all other taxa. An individual ML tree for both hsp70 and RPolII with 100 bootstrap replicates were constructed in RAxML. Branch lengths were estimated for each dataset using RAxML on the best ML tree.

Concatenated variable site analysis (3 datasets)

The practice of concatenating thousands of likely unlinked sites into a single alignment (Philippe et al., 2011) can elide the variable genealogical histories of different chromosomal regions. We thus treated the SISRS variable site data as a single concatenated locus (Yoder et al., 2013). For each dataset, we constructed a phylogeny using maximum likelihood (ML) in RAxML 8.0.20 (Stamatakis) with a General Time Reversible (GTR) model and substitution rates following a discrete gamma distribution with four categories for 1000 bootstrap replicates and allowing for acquisition bias as recommended in the manual (Lewis, 2001).

Estimating divergence time

Because no secure fossil calibration dates within the genus exist, we evaluate the plausibility of Neotropical and Multiple Origins hypotheses given the current temporal understanding of ancient mammal migration patterns or biogeographical events required to explain our phylogeny. Due to the importance of *Leishmania* sp. AM-2004 in re-evaluating the biogeography of the parasite, a phylogeny was constructed for the RNA polymerase II large subunit gene in RAxML, as above, and in *BEAST given a GTR substitution model with rates estimated from a gamma distribution with 4 rate categories and a strict molecular clock for 10,000,000 generations. The molecular clock may not be an appropriate model for the genus, as high amounts of rate heterogeneity within the gene or among lineages may be present; the major goal however is to examine the deep or shallow divergence events of *L. enriettii* isolates within its clade. Branch length estimates from the trees constructed with the strict clock model are used with a calibration date of 40 mya for the split between *L. enriettii* and the Australian isolate with the assumption that the results of the previous phylogenetic trees firmly support an ancient dispersal event. This calibration, like

other fossil dates, provides only a lower bound on the true time since two species diverged. However, an examination of the resulting dates on other nodes within the tree will be compared with previously published dates.

To corroborate the dates estimated with the single RPolII gene tree and a single 40 mya calibration point, the concatenated gene dataset including two outgroup species, *T. cruzi* and *T. brucei*, are used also to estimate branch lengths on a maximum likelihood tree constructed in RAxML, with the above parameters. These branch lengths are converted to time with RelTime (Tamura et al. 2012), given a date of 100 mya for the *T. cruzi* and *T. brucei* split, as published by Stevens et al. 1999 and used by Lukes et al. (2007). The dates for the split of *L. donovani* correspond to those published from the same RPolII gene in a previous study (Lukes et al 2007); therefore we consider the dates of deeper nodes to evaluate the origins hypotheses.

Results

Phylogenetic analyses

Concatenated genes

A total of 73,819 sites in 42 genes were concatenated in the alignment; ~20% of all sites were removed from the analysis due to missing data. The resulting ML tree is similar to that for the variable sites data, but shows an unresolved placement of *L. enrietti* (Fig. 5), here as basal to the *Sauroleishmania*/*Leishmania* subgenera. The branch length at this node is short and the less than half of the individual genes support this partition. Individual gene trees rarely agreed on the species that formed the most ancestral branch of the Euleishmania section and bootstrap consensus trees frequently contained polytomies. The ML trees constructed from first and second codon position, however, fully resolves the phylogeny, supported 100% at each node (with the exception of one crown node), and is identical to that constructed for the variable site data (Fig. 7). The third codon position ML tree share the same topology but are not as well resolved. The

conflict in the topology may be a result of short branch lengths (0.01 and lower) associated with the split between *L. enrietti* and the clades within Euleishmania.

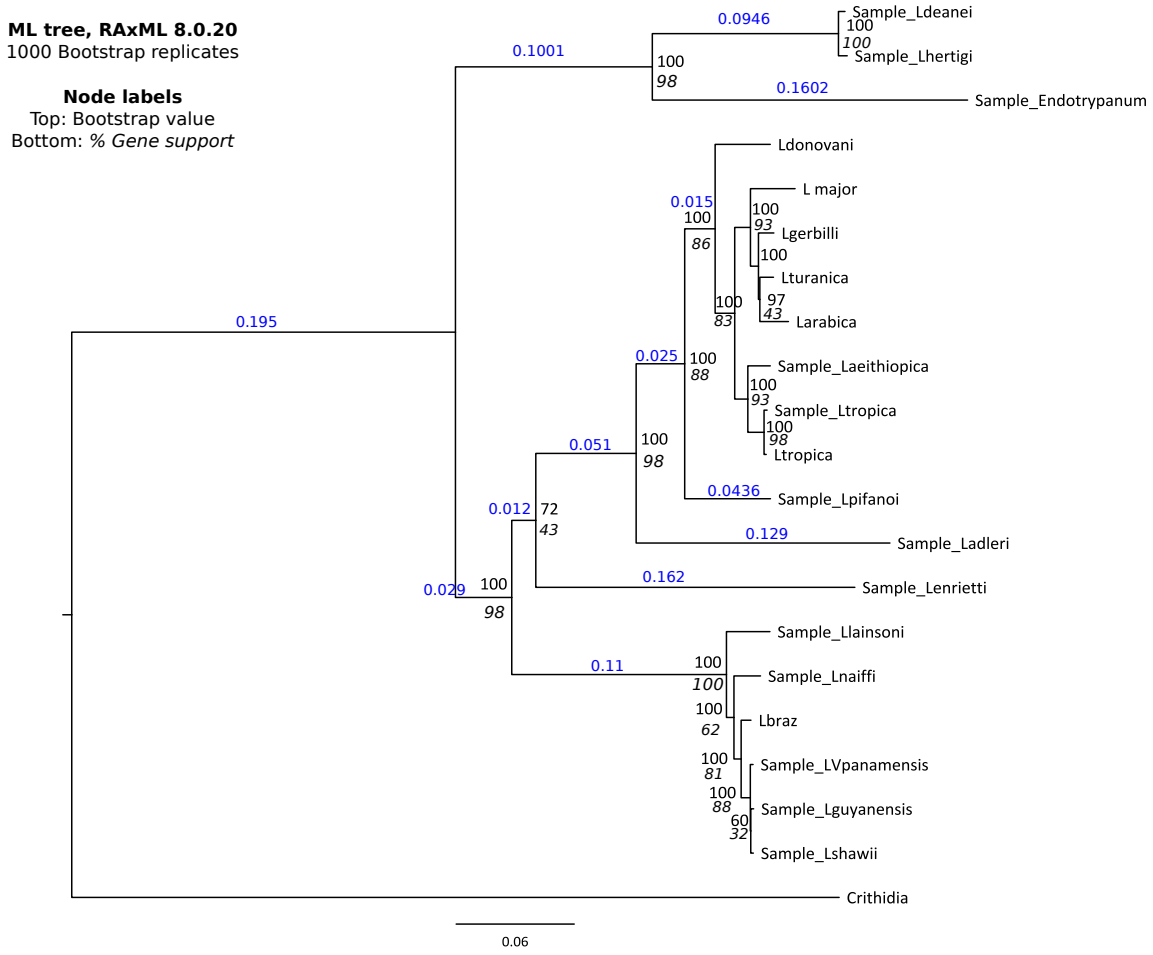


Figure 4: ML tree of 42 concatenated genes.

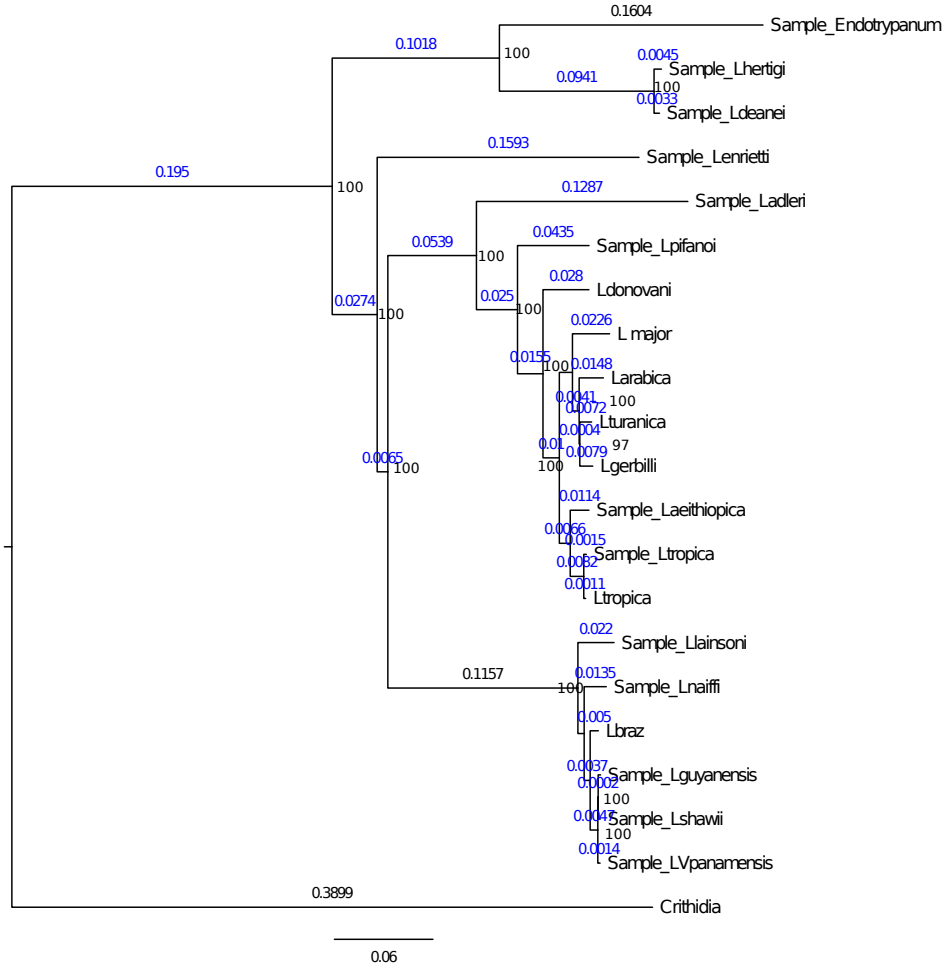


Figure 5: ML tree of concatenated genes, first position only. Branch lengths in blue, node values indicate bootstrap support.

Single gene analyses

As anticipated, Leishmania sp. MAR1 and AM-2004, as well as *L. siamensis* are most closely related to the guinea pig for of Leishmania, *L. enrietti*, and form a monophyletic group. The phylogeny for the RPolII and hsp-70 genes, 1201 bp and 1852 bp, respectively, reflect the same overall topology as the variable site data with regards to the ancestral placement of *L. enrietti*. As with the concatenated dataset, however, support values for the branches leading to *L. enrietti* and Euleishmania less than 80% supported (Figure 6 and 7).

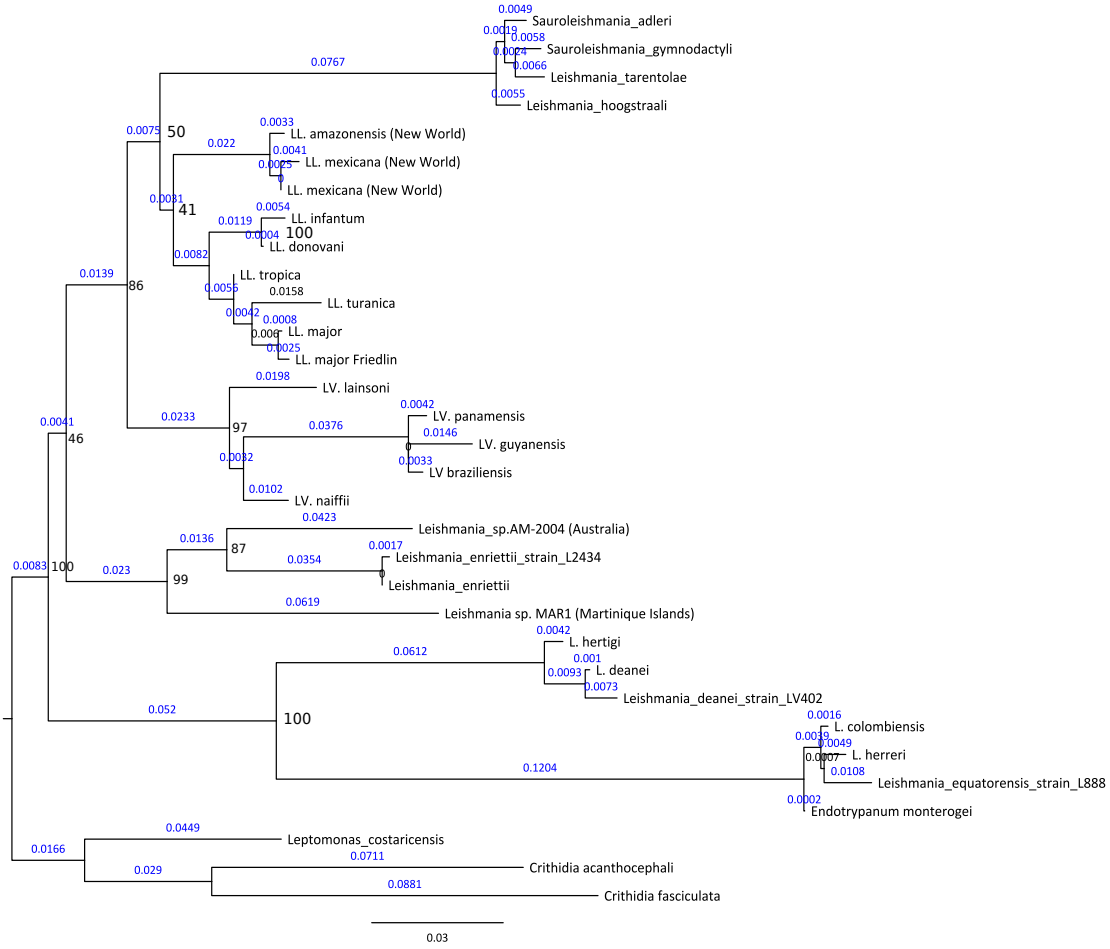


Figure 6: ML tree of RPolII gene. Branch lengths in blue; node values indicate bootstrap support. L.enriettii complex is circled in red.

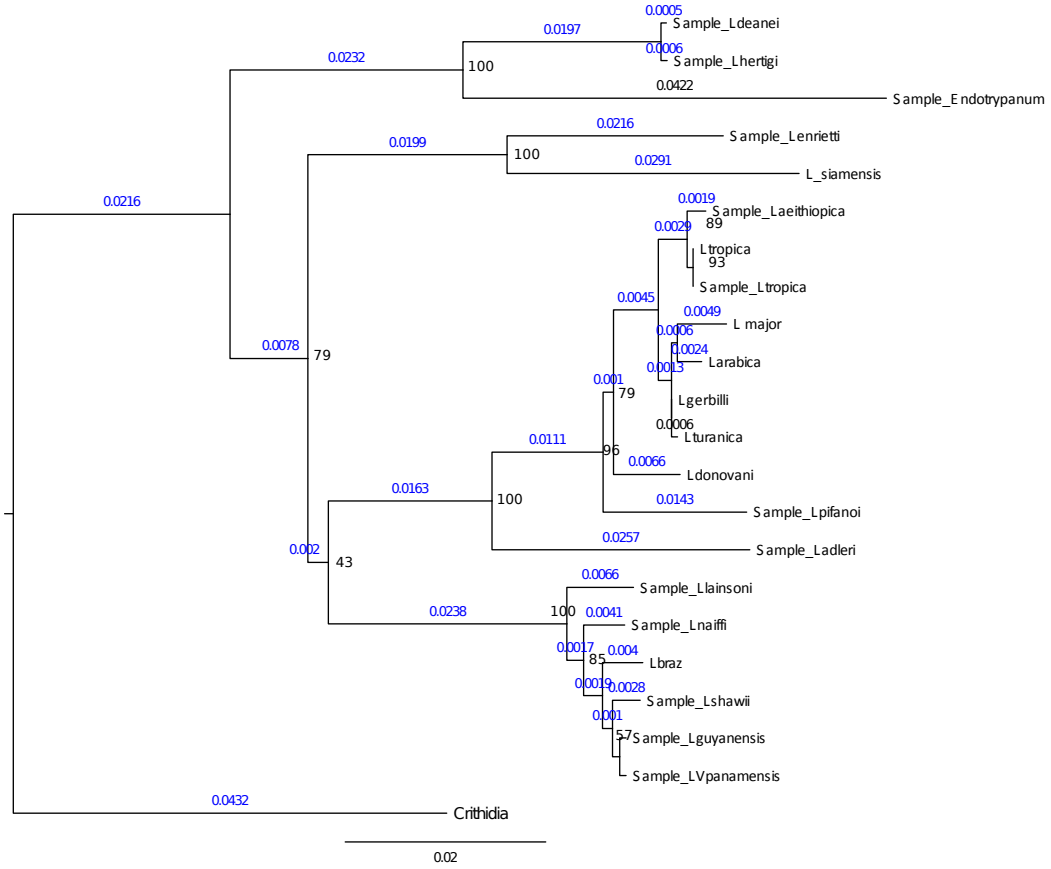


Figure 7: ML tree of hsp70 gene. Branch lengths are in blue; node values indicate bootstrap support; L. enriettii complex is circled.

Variable sites

The number of variable sites identified for each dataset were 388111 (VS_m6), 18312 (VS_m6s), and 2790 (VS_t80). Based on an alignment with the reference genome for *L. tarentolae*, we determined these sites are distributed across the *Leishmania* genome. GC content was over 70%, which is higher than the GC content estimated previously of 50-60% (Peacock et al. 2007). The ML phylogeny for all of these datasets supports the *L.Leishmania* and *L.Viannia* subgenera as monophyletic clades; Sauroleishmania is also monophyletic, although only two taxa were available for phylogenetic analysis (Figure 4). The branch leading to *L. enrietti* is likewise fully

supported at the root of the Euleishmania section. As shown in other sequenced-based trees (e.g. (Croan et al., 1997), the Paraleishmania section, comprised here by *L. hertigi*, *L. deanei*, and *Endotrypanum*, are ancestral to Euleishmania, i.e. all other *Leishmania* lineages. The OW *Leishmania* subgenus lies at the crown.

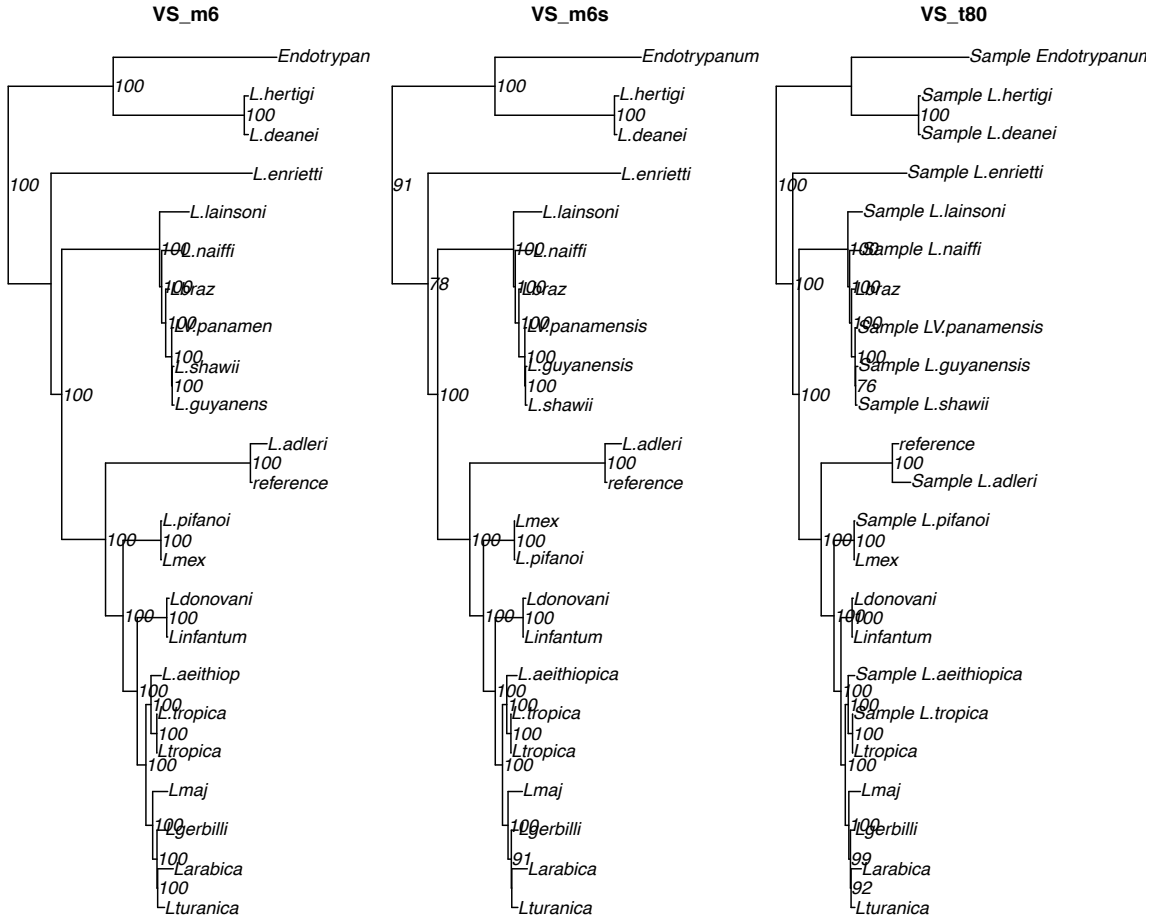


Figure 8: Maximum likelihood trees based on three datasets of variable sites. Node values indicate bootstrap support.

Divergence time estimates

A visual examination of branch lengths in previous reported phylogenies (Rose et al. 2011, not shown) and those generated in this study demonstrate a deep divergence within the *L.enrietti* complex. This divergence is most clearly seen in the phylogenies of hsp-70 and RPolII where multiple isolates within the complex are available for analysis. The depth of the split in the RpolII phylogeny leading to the Australian isolate rends the argument for a recent or anthropogenic introduction to that continent implausible. The concept that *Leishmania* must have reached Australia before the isolation of the landmass is thus justifiable. Likewise, the deep split at *L. siamensis* and *L. enrettii* also suggests an ancient divergence event. Thus *L. enriettii* complex may have reached other parts of present-day Asia and Europe in addition to Australia, but more data are necessary to consider this distribution. Interestingly, the inclusion of four Sauroleishmania species in the RPolII phylogeny illustrates a relatively recent jump into to OW reptiles, relative to the length of the entire branch.

The *BEAST analysis of the RPolII gene shows high ESS values and convergence, as observed in Tracer 1.5. The results of the Bayesian phylogeny implemented with a strict molecular clock indicate that by manually calculating a calibration date of 40 mya for the split the Australian isolate, and working backwards from the tips of only those branches sharing the same total node height, the previously estimated dates of the split of *L. donovani-L. major* (24.6-14.7 mya) (Lukes et al. 2007) correspond well to our estimate of 17.7 mya. TimeTree (Hedges et al. 2006) estimates the median split of *L. donovani* and *L. major* at 5.5 mya, where we estimate 3.3 mya. The estimates at basal nodes, shown in Figure 9, implicate an origin of the genus >90 million years at the youngest, if we believe this constant rate to be true. A relaxed clock in BEAST (not shown) lengthens the branches slightly.

An Reltime tree estimated from the concatenated gene dataset with a 100 mya calibration point at the split between two outgroup species, *T. cruzi* and *T. brucei*, produces extremely large confidence intervals that create ambiguities for interpreting the dates, likely due to an extremely long branch length leading to *T. brucei*. With this caveat in mind, the median date for the *L.*

donovani split is not far from the expectation given previously reports, at 25.2 mya (5.1-45.3 95% CI). The dates estimated on the multi-gene tree for the origin of the entire Endotrypanum/Leishmania genus is ~112 mya, although with large CI intervals (tree not shown).

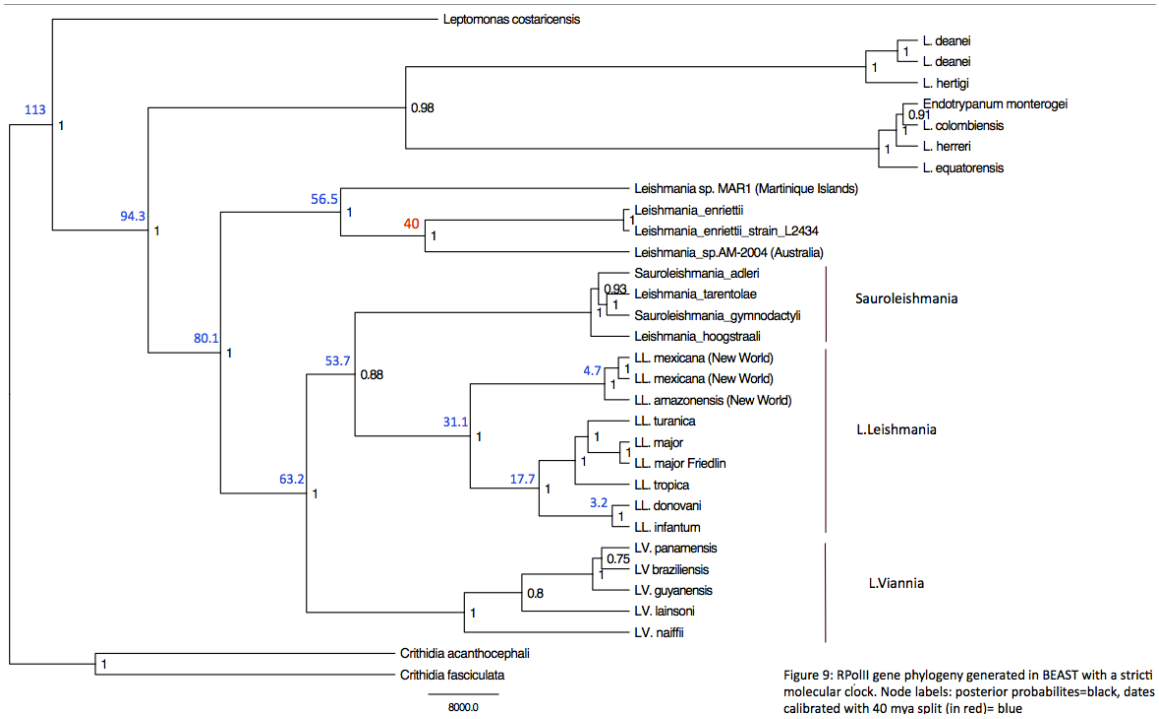


Figure 9: Species tree for RNA polymerase II gene, constructed in *BEAST (Heled and Drummond 2010) with strict molecular clock. Node values indicate estimated dates (blue) and posterior probability (black).

Discussion

Phylogeny

As expected, our results place species of Old and New World subgenus *Leishmania* in a monophyletic clade separate from species of the exclusively New World subgenus *Viannia*. These results are consistent with other molecular-based trees, including those that firmly place Sauroleishmania at the root of the *Leishmania* subgenus (e.g. Fraga et al. 2010, Rose et al. 2011). Interestingly, our genome wide variable site results place *L. enrietti*, a parasite of the New

World guinea pig, basal to all Euleishmania. This result differs from early hypotheses that suggest *L. enriettii* is a member of the New World *L. Leishmania* subgenus (Lainson, 1997; Lainson and Shaw, 1987) or an outgroup to the entire genus (Kerr, 2006; Yurchenko et al., 2006), rather than ancestral to Euleishmania, which has been corroborated by molecular trees (Croan et al., 1997; Noyes et al., 2002; Stevens et al., 2001).

When all codon positions are considered, the concatenated gene dataset produces a conflicting topology whereby *L. enrietti* is not basal to all Euleishmania but to the *Sauroleishmania/Leihsmania* subgenera. This is not necessarily surprising given the extremely short branch lengths leading to the splits of *L. enrietti* and the *Euleishmania* clades, indicating the potentially distorting effects of long branch attraction (Felsenstein, 1988), or incomplete lineage sorting. The latter scenario, specifically where an old but rapid set of divergence events yield short branches deep in the tree, has been shown to pose difficulty in recovering the species tree, although adding more loci to increase resolution is advocated in this case (Maddison and Knowles, 2006). It is interesting to note that phylogenetic resolution and 100% support is obtained using only the first and second codon position, decreasing in support with the third position. All codon phylogenies support the variable site results. As with the variable site phylogeny, however, we observe a short branch at the splits of the subgenera. Substitutions are expected to be most rapid at the third position (Kimura, 1968). At these time scales, saturation in the third position may negatively affect the phylogenetic signal, reflected in the decrease of support.

As with nearly all other analyses to date, the molecular evidence does not support the taxonomic classification of *Endotrypanum* as a separate genus. Relative to molecular markers, there are few morphological characters used to classify unicellular organisms, calling their utility into question, when compared to molecular data (Perkins et al., 2011). Sequence data supports their inclusion in the *Leishmania* genus, unless all other Paraleishmania species are misclassified.

Leishmania Origins

The multi-gene and variable site *Leishmania* phylogenies do not support Kerr's (2000, 2006) Palaearctic origin of *Leishmania*. This result is not only due to the position of *Sauroleishmania* in the middle rather than at the root of the tree, but also given our relative branch lengths, a Pliocene introduction of the MRCA of *(L.) Viannia* to the Neotropics at ~5 mya after the reformation of the Panama isthmus is implausible. If Kerr is correct in asserting that reptiles were the first hosts of *Leishmania* in the Cretaceous, extinction events associated with the K-T boundary could have erased evidence of those lineages; this speculation does not affect our interpretation.

In the Neotropical Origins hypothesis, the current global distribution of *Leishmania* is a purely a result of dispersal through vector/reservoir migration and not vicariance. The ancient, global distribution of the *L. enrietti* complex, however cast doubts on this scenario. At the latest, this distribution must have occurred before the isolation of Australia from the southern supercontinent at 40-50 mya (Flynn and Wyss 1998). Luke et al.'s (2007) version of the hypothesis proposes the ancestor of Paraleishmania evolved in the Neotropics between 46-36 mya and dispersed north. Similarly, Noyes et al. (1997) and Noyes (1998) propose that the Leishmania/Endotrypanum clade evolved in the Neotropics between 65-40 mya and dispersed to the Nearctic and to the Palearctic through land bridges; the Neotropics however were isolated at this time. A minimum age of 40 mya for the split of the Australian *Leishmania* species would require a much older genus than proposed by this scenario; therefore we can also reject this hypothesis.

The less cited Multiple Origins scenario refers to the separation and subsequent independent evolution of the *Viannia* and *Leishmania* subgenera before 90-100 mya. This is the only formalized hypothesis that includes vicariance as a mechanism for the evolution of the major *Leishmania* clades. In proposing this hypothesis, however, the authors make small, erroneous suppositions. For example, the emergence of Paraleishmania with the introduction of

hystricomorph rodents in the “early Cenozoic” is unlikely since they do not appear in the fossil record until 23 mya (Paleobiology database, fossilworks.org).

We instead propose a Supercontinent hypothesis for the origin of the *Leishmania* genus. In this scenario, the MRCA of Leishmania/Endotrypanum emerged together from monoxneous parasites, i.e. restricted to one host (Yurchenko et al. 2006) on Gondwana, which can accommodate biogeographic, fossil and molecular evidence.

Biogeographic and fossil evidence

The correlation between extant distributions and the phylogeny can be distorted by the uncertainties of the paleobiogeographical record and further confounded by the digenetic life cycle of the *Leishmania* parasite, i.e. unknown host/reservoir ranges, extinct hosts/vectors, extinct *Leishmania* lineages. The Supercontinent proposal is in agreement with Shaw (1997) who suggested that an adaptation to mammals occurred around 90 mya when mammals began to radiate and Africa became fully isolated. The worldwide distribution of the sandfly genera that serve as vectors for the parasite likely results from the breakup of Pangaea and subsequent continental splintering (Filho and Brazil, 2003; Galati, 1995). There were no significant changes in insect taxa across the K-T boundary that would have changed the vector (Labandeira and Sepkoski, 1993) but it is interesting to consider the potential change in host species that occurred alongside the K-T extinction event if reptiles were also hosts during the diversification of *Leishmania*.

The Early Cretaceous fossils of *Paleoleismania proterus* are found in sandflies trapped within Burmese amber ~100 mya and are reportedly evidence of first digenetic trypanosomatids, associated with reptile hosts (Poinar Jr, 2007). There is no way to confirm the genus but the organisms are morphologically similar to *Leptomonas*, the sister of *Leishmania*. The fossil taxon *Paleoleishmania neotropicum* is found in the Dominican Republic between 15-45 mya and is hypothesized to be the progenitor of Neotropical clades, although the dating of Dominican amber

is controversial (Poinar, 2008). These fossils cannot be used refute the proposed hypotheses but introduce the possibility of an MRCA as old as the Early Cretaceous.

Molecular evidence

The levels of genetic diversity between the *Viannia* subgenera and *Sauro/Leishmania* subgenera have been cited previously as a reflection of vicariance after the separation of South America and Africa. Additional support for this hypothesis is inferred from the absence of the surface glycoprotein gene family (GP36/M-2) in Paraleishmania, *L. enriettii* and the *Viannia* subgenus (Cupolillo et al., 2000; McMahon-Pratt et al., 1992). Without vicariance, this absence is difficult to explain and is in conflict with the Neotropical origins, which would require the unlikely gain of this gene family as the subgenera evolved. Instead, the loss of the GP46-M-2 could have occurred shortly after geographic isolation of the MRCA while conserved in Sauro/Leishmania.

The concept of a molecular clock can be extremely useful in correlating substitution rates to time, but it is easily violated by loci or lineages that do no evolve at a constant rate. Future analyses that can accurately account for gene and among-lineage rate heterogeneity are underway and will test the assumptions of the Supercontinent hypothesis, namely that the MRCA of Paraleishmania, i.e. *Leishmania/Endotrypanum*, emerged in Gondwana before c. 90-100 mya, and. *L. enrietti* dispersed before the separation of Australia from Antarctica c. 40-50 mya. If the short branch lengths/rapid split leading to *Viannia* and *Sauro/Leishmania* subgenera occurred c. 65 mya (Figure 9), we could be observing highly punctuated diversification via the explosion of mammalian reservoirs closely following the mass reptilian extinction of the K-T boundary, i.e. the “Explosive Model” (Jablonski, 2001; Springer et al., 2003). This hypothesis of course requires testing with appropriate data and clock methods.

Conclusion

Our results support the origins of the genus *Leishmania* at a time when continents were united prior to 90-100 mya and a global dispersal prior to separation of the southern supercontinents.

There have been multiple independent transitions to human infection; this is consistent with contemporary epidemiology; the increase of immunocompromised patient populations illustrates how capable the parasites are at switching between mammalian, but not reptilian hosts/reservoirs. Leishmania may have diversified in mammals as mammals themselves radiated after K-T.

Advantages of genome data and a comparison with traditional phylogenetic data

This study is the first phylogenomic analysis of the disease-causing parasite *Leishmania*. We use whole genome data from thirteen new species to characterize genus-level diversity. Although not every dataset was able to fully resolve the important relationship at the split between the Leishmania *subgenera*, this likely reflects a reality of the genealogical history or issues inherent issues in using only extant taxa to infer deep phylogenetic relationships (Donoghue et al., 1989).

New genome wide data provide markers for future analyses

Why are (*L.)Leishmania* and *(L.)Viannia* capable of infecting humans and a wide range of mammals, while (*L.)Sauroleishmania* and the Paraleishmania species are primarily host-specific? Answers to this and other questions are accessible through molecular sequence analysis. The current study contributes thousands of loci across pathogenic and non-pathogenic lineages aligned to annotated reference genomes, made available to the broad community of scholars invested in illuminating *Leishmania* molecular evolution for a wide range of clinical and intellectual goals.

Uncertainty in *Leishmania* evolutionary history impacts how we interpret the timing and origin of the parasite and its global dispersal. Overall, the bias in current taxon and genetic loci availability limits broadly representative phylogenetic reconstructions. Furthermore, an insufficient understanding of *Leishmania* genomics has continued to impede efforts to create affordable

vaccines and effective treatment, in part because there is no predictable association between the species genotype and the clinical outcome.

Future Directions

There are several ongoing analyses that are confronting problems associated with heterogeneous rates of evolution within particular genes (e.g. strong selective pressures) and along specific lineages in respect to accurately estimating divergence dates.

These analyses will:

1. Test whether all genes in analysis reject clock-like behavior (Molecular Clock test implemented in MEGA 6.0 (Tamura et al. 2013) and classify genes under strong purifying or positive selection (codon-based Z-test or HyPhy) for separate analyses;
2. Extract loci around each SISRS variable site to find regions evolving under constant rates and the inclusion of invariable sites;
3. Examine the loci determined to evolve at constant rates with RelTime (Tamura et al. 2012) or other molecular dating methods.

Funding:

This work was supported by ASU startup funds to RAC and a National Science Foundation Doctoral Dissertation Improvement Grant [grant number BCS-1232582 to KMH, JEB and ACS].

Acknowledgements:

Genomic DNA from *Leishmania* isolates were obtained from the lab of Diane McMahon-Pratt at the Yale School of Public Health and from Lucille Floeter-Winter at the University of Sao Paolo.

Species name	Isolate ID	Region	Location	Host	WHO Number	Subgenus	Lab PI
L. adleri	LRC-L123	OW	Kenya	Lizard/*Latastia longicaudalia*	RLAT/KE/57/SKINK-7	"Sauroleishmania"	DMcp
L. L. aethiopica	LRC-L147	OW	Ethiopia	man	MHOM/ET/71/L100	Leishmania	DMcp
L.mexicana pifanoi	Ltrod	NW	Venezuela	man	MHOM/VE/00/Ltrod	Leishmania	DMcp
Endotrypanum schaudini	M6159	NW	Sloth	Sloth	MCHO/BR/80/M6159	n/a	DMcp
L. hertigi hertigi	M4051	NW	Panama	porcupine, *Coendou rothschildi*	MCOE/PA/00/M4051, also referred to as LV42 and C-8	n/a	DMcp
L. V. panamensis	WR120	NW	Panama	man	MHOM/PA/74/WR120	Viannia	DMcp
L. deanei	M5088	NW	Brazil	*Coendou*	MCOE/BR/00/M5088	n/a	DMcp
L.enriettii	L88	NW	Brazil	Cavia	MCAV/BR/45/L88	n/a	DMcp
L. V. guyanensis	M4147	NW	Brazil	man	MHOM/BR/75/M4147	Viannia	DMcp
L. L. tropica	LRC-L39	OW	USSR	man	MHOM/SU/60/LRC-L39	Leishmania	DMcp
L.V. shawii		NW				Viannia	LFW
L.V. naiffi		NW		armadllo		Viannia	LFW
L.V. lainsoni		NW				Viannia	LFW

Table 1. DMcP= Diane McMahon-Pratt, LFW=Lucile Floeter-Winter

Table 1: Leishmania strains sequenced in this analysis

Species name	Isolate ID	Region	Location	Host	WHO Number	Subgenus	Lab PI
L. adleri	LRC-L123	OW	Kenya	Lizard/*Latastia longicaudalia*	RLAT/KE/57/SKINK-7	"Sauroleishmania"	DMcp
L. L. aethiopica	LRC-L147	OW	Ethiopia	man	MHOM/ET/71/L100	Leishmania	DMcp
L.mexicana pifanoi	Ltrod	NW	Venezuela	man	MHOM/VE/00/Ltrod	Leishmania	DMcp
Endotrypanum schaudini	M6159	NW	Sloth	Sloth	MCHO/BR/80/M6159	n/a	DMcp
L. hertigi hertigi	M4051	NW	Panama	porcupine, *Coendou rothschildi*	MCOE/PA/00/M4051, also referred to as LV42 and C-8	n/a	DMcp
L. V. panamensis	WR120	NW	Panama	man	MHOM/PA/74/WR120	Viannia	DMcp
L. deanei	M5088	NW	Brazil	*Coendou*	MCOE/BR/00/M5088	n/a	DMcp
L.enriettii	L88	NW	Brazil	Cavia	MCAV/BR/45/L88	n/a	DMcp
L. V. guyanensis	M4147	NW	Brazil	man	MHOM/BR/75/M4147	Viannia	DMcp
L. L. tropica	LRC-L39	OW	USSR	man	MHOM/SU/60/LRC-L39	Leishmania	DMcp
L.V. shawii		NW				Viannia	LFW
L.V. naiffi		NW		armadllo		Viannia	LFW
L.V. lainsoni		NW				Viannia	LFW

Table 1. DMcP= Diane McMahon-Pratt, LFW=Lucile Floeter-Winter

Table 1: Leishmania strains sequenced in this analysis

GeneInfo Identifier	ReferenceID	L. braziliensis strain ID	Gene
>gi 154345881	XM_001568828.1	MHOM/BR/75/M2904	tyrosine aminotransferase (TAT)
>gi 154337345	XM_001564856.1	MHOM/BR/75/M2904	RNA polymerase II (LBRM_21_2050)
>gi 389602506		MHOM/BR/75/M2904	prostaglandin f2-alpha synthase/D-arabinose dehydrogenase (pgfs)
>gi 389603341	XM_001569011.2	MHOM/BR/75/M2904	putative oxidoreductase (LBRM_35_4420)
>gi 112383574	gbDQ836162.1	HOM/BR/75/M2903	N-acetylglucosamine-1-phosphate transferase gene partial cds (nagt)
>gi 154346249	XM_001569012.1	MHOM/BR/75/M2904	putative UDP-N-acetylglucosamine-dolichyl-phosphate N-acetylglucosaminephosphotransferase (LBRM_35_4430)
>gi 154336055	XM_001564214.1	MHOM/BR/75/M2904	mitogen-activated protein kinase (MPK4)
>gi 155675719	gbEU053119.1	MHOM/PE/88/BAA2079	mannose phosphate isomerase (MPI) gene
>gi 155675703	gbEU053111.1	MHOM/PE/88/BAA2079	malate dehydrogenase (MDH) gene
>gi 389602969		MHOM/BR/75/M2904	HSP83 heat shock protein 83-1
>gi 9864198	gbAF291716.1		heat shock protein 70 (hsp70) gene
>gi 322505745		MHOM/BR/75/M2904	chromosome 36 GPI alpha-mannosyltransferase III
>gi 154341936	XM_001566870.1	MHOM/BR/75/M2904	glyceraldehyde 3-phosphate dehydrogenase, glycosomal (LBRM_30_2950)
>gi 154332907	XM_001562666.1	MHOM/BR/75/M2904	elongation factor-1 gamma (LBRM_09_1020)
>gi 154335073	XM_001563727.1	MHOM/BR/75/M2904	elongation factor 1-alpha (LBRM_17_0090)
>gi 2581879	gbAF009138.1		DNA polymerase alpha gene
>ENA CAM37041.1		MHOM/BR/75/M2904	dihydrofolate reductase-thymidylate synthase (dhfr-ts)
>ENA AB434681.1		MHOM/BR/00/LTB300	kinetoplast pre-edited Cytb gene for cytochrome b
>gi 154345431	XM_001568603.1	MHOM/BR/75/M2904	elongation factor 2 (LBRM_35_0270)
>gi 154342453	XM_001567125.1	MHOM/BR/75/M2904	putative cytochrome c oxidase VIII (COX VIII) (LBRM_31_1780)
>gi 154343626	XM_001567709.1	MHOM/BR/75/M2904	myosin XXI (LBRM_32_4110)
>gi 154340941	XM_001566374.1	MHOM/BR/75/M2904	putative heat shock protein 90 (LBRM_29_0780)
>gi 389604019	XM_003723107.1	MHOM/BR/75/M2904	arginine N-methyltransferase-like protein (LBRM_20_6060)
>gi 389604005	XM_003723100.1	MHOM/BR/75/M2904	serine peptidase (LBRM_20_6000)
>gi 389603991	XM_003723093.1	MHOM/BR/75/M2904	putative N-acyl-L-amino acid amidohydrolase (LBRM_20_5930)
>gi 389603983	XM_003723089.1	MHOM/BR/75/M2904	mitochondrial RNA ligase 2 (LBRM_20_5890)
>gi 389603973	XM_003723084.1	MHOM/BR/75/M2904	putative separin (LBRM_20_5840)
>gi 389603961	XM_003723078.1	MHOM/BR/75/M2904	cell division cycle protein-like protein (LBRM_20_5780)
>gi 389603935	XM_003723065.1	MHOM/BR/75/M2904	putative axoneme central apparatus protein (LBRM_20_5640)
>gi 389603929	XM_003723062.1	MHOM/BR/75/M2904	putative RNA-binding regulatory protein (LBRM_20_5610)
>gi 389603919	XM_003723057.1	MHOM/BR/75/M2904	putative coatomer beta subunit (LBRM_20_5560)
>gi 389603917	XM_003723056.1	MHOM/BR/75/M2904	kinase-like protein (LBRM_20_5550)
>gi 389603921	XM_003723058.1	MHOM/BR/75/M2904	putative exosome associated protein 1 (Rrp42 homologue) (LBRM_20_5570)
>gi 389603879	XM_003723037.1	MHOM/BR/75/M2904	triosephosphate isomerase (LBRM_20_5360)
>gi 389603873	XM_003723034.1	MHOM/BR/75/M2904	endo-1, 4-beta-xylanase z precursor-like protein (LBRM_20_5330)
>gi 389603843	XM_003723019.1	MHOM/BR/75/M2904	putative glutaredoxin (LBRM_20_5180)
>gi 389603811	XM_003723003.1	MHOM/BR/75/M2904	cytochrome c oxidase assembly factor-like protein (LBRM_20_5020)
>gi 389603809	XM_003723002.1	MHOM/BR/75/M2904	phosphopantetheinyl transferase-like protein (LBRM_20_5010)
>gi 389603777	XM_003722986.1	MHOM/BR/75/M2904	conserved hypothetical protein (LBRM_20_4850)
>gi 389603767	XM_003722981.1	MHOM/BR/75/M2904	chaperone protein DNAJ-like protein (LBRM_20_4790)
>gi 389603765	XM_003722980.1	MHOM/BR/75/M2904	conserved hypothetical protein (LBRM_20_4780)
>gi 389603747	XM_003722971.1	MHOM/BR/75/M2904	glycerol-3-phosphate dehydrogenase-like protein (LBRM_20_4680)

Table 2: Genes used in analysis according to the reference sequence.

References Cited

Alvar J, Velez ID, Bern C, Herrero M, Desjeux P, Cano J, Jannin J, den Boer M, and Team WHOLC. 2012. Leishmaniasis worldwide and global estimates of its incidence. PLoS One 7(5):e35671.

Boite MC, Mauricio IL, Miles MA, and Cupolillo E. 2012. New insights on taxonomy, phylogeny and population genetics of Leishmania (Viannia) parasites based on multilocus sequence analysis. PLoS Neglected Tropical Diseases 6(11):e1888.

Bualert L, Charungkiattikul W, Thongsuksai P, Mungthin M, Siripattanapipong S, Khositnithikul R, Naaglor T, Ravel C, El Baidouri F, and Leelayoova S. 2012. Autochthonous disseminated dermal and visceral leishmaniasis in an AIDS patient, southern Thailand, caused by Leishmania siamensis. Am J Trop Med Hyg 86(5):821-824.

Castresana J. 2007. Topological variation in single-gene phylogenetic trees. Genome biology 8(6):216.

Croan DG, Morrison Da, and Ellis JT. 1997. Evolution of the genus Leishmania revealed by comparison of DNA and RNA polymerase gene sequences. Molecular and biochemical parasitology 89(2):149-159.

Cruz I, Nieto J, Moreno J, C1 nnavate C, Desjeux P, and Alvar J. 2006. Leishmania/HIV co-infections in the second decade. Indian J Med Res 123:357-388.

Cupolillo E, Medina-Acosta E, Noyes H, Momen H, and Grimaldi G, Jr. 2000. A revised classification for Leishmania and Endotrypanum. Parasitol Today 16(4):142-144.

Donoghue MJ, Doyle JA, Gauthier J, Kluge AG, and Rowe T. 1989. The importance of fossils in phylogeny reconstruction. Annual Review of Ecology and Systematics 20(1):431-460.

Dougall AM, Alexander B, Holt DC, Harris T, Sultan AH, Bates PA, Rose K, and Walton SF. 2011. Evidence incriminating midges (Diptera: Ceratopogonidae) as potential vectors of Leishmania in Australia. International Journal for Parasitology 41(5):571-579.

Felsenstein J. 1988. Phylogenies from molecular sequences: inference and reliability. Annual review of genetics 22:521-565.

Fernandes AP, Nelson K, and Beverley SM. 1993. Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism. Proc Natl Acad Sci U S A 90(24):11608-11612.

Filho JD, and Brazil RP. 2003. Relationships of new world phlebotomine sand flies (Diptera: Psychodidae) based on fossil evidence. Mem Inst Oswaldo Cruz 98 Suppl 1:145-149.

Flynn JJ, and Wyss AR. 1998. Recent advances in South American mammalian paleontology. Trends Ecol Evol 13(11):449-454.

Fraga J, Montalvo AM, De Doncker S, Dujardin J-C, and Van der Auwera G. 2010. Phylogeny of Leishmania species based on the heat-shock protein 70 gene. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases 10(2):238-245.

Gadagkar SR, Rosenberg MS, and Kumar S. 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. Journal of experimental zoology Part B, Molecular and developmental evolution 304(1):64-74.

Galati E. 1995. Phylogenetic systematics of Phlebotominae (Diptera, Psychodidae) with emphasis on American groups. Bol Direc Malariol y San Amb 35(1):133-142.

Heled J, and Drummond AJ. 2010. Bayesian inference of species trees from multilocus data. Mol Biol Evol 27(3):570-580.

Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E, Aert R et al. . 2005. The genome of the kinetoplastid parasite, Leishmania major. Science 309(5733):436-442.

Jablonski D. 2001. Lessons from the past: evolutionary impacts of mass extinctions. Proc Natl Acad Sci U S A 98(10):5393-5398.

Junier T, and Zdobnov EM. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics 26(13):1669-1670.

Kanjanopas K, Siripattanapipong S, Ninsaeng U, Hitakarun A, Jitkaew S, Kaewtaphaya P, Tan-ariya P, Mungthin M, Charoenwong C, and Leelayoova S. 2013. Sergentomyia (Neophlebotomus) gemmea, a potential vector of Leishmania siamensis in southern Thailand. BMC infectious diseases 13:333.

Kerr SF. 2000. Palaearctic origin of Leishmania. Memórias do Instituto Oswaldo Cruz 95:75-80.

Kerr SF. 2006. Molecular trees of trypanosomes incongruent with fossil records of hosts. Memórias do Instituto Oswaldo Cruz 101(1):25-30.

Kerr SF, Merkelz R, and Mackinnon C. 2000. Further support for a Palaearctic origin of Leishmania. Memórias do Instituto Oswaldo Cruz 95(4):579-581.
Kimura M. 1968. Evolutionary rate at the molecular level. Nature 217(5129):624-626.

Kumar S, Filipski AJ, Battistuzzi FU, Kosakovsky Pond SL, and Tamura K. 2012. Statistics and truth in phylogenomics. Mol Biol Evol 29(2):457-472.

Labandeira CC, and Sepkoski JJ, Jr. 1993. Insect diversity in the fossil record. Science 261(5119):310-315.

Lainson R. 1997. On Leishmania enriettii and other enigmatic Leishmania species of the Neotropics. Mem Inst Oswaldo Cruz 92(3):377-387.

Lainson R. 2010. The Neotropical Leishmania species: a brief historical review of their discovery, ecology and taxonomy. Rev Pan-Amaz Saude 1(2):13-32.

Lainson R, and Shaw JJ. 1972. Leishmaniasis of the New World: taxonomic problems. British medical bulletin 28(1):44-48.

Lainson R, and Shaw JJ. 1987. Evolution, classification, and geographical distribution. In: Peters W, and
Killick-Kendrick R, editors. The leishmaniases in biology and medicine. London: Academic Press. p 1-120.

Lanfear R, Calcott B, Ho SY, and Guindon S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol 29(6):1695-1701.

Langmead B, and Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357-359.

Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst Biol 50(6):913-925.

Lipoldova M, and Demant P. 2006. Genetic susceptibility to infectious disease: lessons from mouse models of leishmaniasis. Nat Rev Genet 7(4):294-305.

Lukes J, Mauricio IL, Schönian G, Dujardin J-C, Soteriadou K, Dedet J-P, Kuhls K, Tintaya KWQ, Jirků M, Chocholová E et al. . 2007. Evolutionary and geographical history of the Leishmania donovani complex with a revision of current taxonomy. Proceedings of the National Academy of Sciences USA 104(22):9375-9380.

Lysenko A. 1971. Distribution of leishmaniasis in the Old World. Bulletin of the World Health Organization 44:515-520.

Maddison WP, and Knowles LL. 2006. Inferring phylogeny despite incomplete lineage sorting. Syst Biol 55(1):21-30.

Marcili A, Speranca MA, da Costa AP, Madeira MD, Soares HS, Sanches CD, Acosta ID, Girotto A, Minervino AH, Horta MC et al. . 2014. Phylogenetic relationships of Leishmania species based on trypanosomatid barcode (SSU rDNA) and gGAPDH genes: Taxonomic revision of Leishmania (L.) infantum chagasi in South America. Infect Genet Evol.

Marfurt J, Niederwieser I, Makia ND, Beck H-P, and Felger I. 2003. Diagnostic genotyping of Old and New World Leishmania species by PCR-RFLP. Diagnostic Microbiology and Infectious Disease 46(2):115-124.

McMahon-Pratt D, Traub-Cseko Y, Lohman KL, Rogers DD, and Beverley SM. 1992. Loss of the GP46/M-2 surface membrane glycoprotein gene family in the Leishmania braziliensis complex. Mol Biochem Parasitol 50(1):151-160.

Momen H, and Cupolillo E. 2000. Speculations on the origin and evolution of the genus Leishmania. Memórias do Instituto Oswaldo Cruz 95(4):583-588.

Mueller RL, Macey JR, Jaekel M, Wake DB, and Boore JL. 2004. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders inferred from complete mitochondrial genomes. Proc Natl Acad Sci U S A 101(38):13820-13825.

Nei M, and Kumar S. 2000. Molecular evolution and phylogenetics. Oxford ; New York: Oxford University Press. xiv, 333 p. p.

Noyes H, Pratlong F, Chance M, Ellis J, Lanotte G, and Dedet JP. 2002. A previously unclassified trypanosomatid responsible for human cutaneous lesions in Martinique (French West Indies) is the most divergent member of the genus Leishmania ss. Parasitology 124(Pt 1):17-24.

Noyes HA. 1998. Can Trypanosoma Trees be Trusted? Parasitology 14(2):49-50.

Noyes HA, Morrison DA, Chance ML, and Ellis JT. 2000. Evidence for a neotropical origin of Leishmania. Memórias do Instituto Oswaldo Cruz 95(4):575-578.

O'Neill EM, Schwartz R, Bullock CT, Williams JS, Shaffer HB, Aguilar-Miguel X, Parra-Olea G, and Weisrock DW. 2013. Parallel tagged amplicon sequencing reveals major lineages and

phylogenetic structure in the North American tiger salamander (Ambystoma tigrinum) species complex. Molecular ecology 22(1):111-129.

Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA, Peters N, Adlem E, Tivey A, Aslett M et al. 2007. Comparative genomic analysis of three Leishmania species that cause diverse human disease. Nature Genetics 39(7):839-847.

Perkins SL, Martinsen ES, and Falk BG. 2011. Do molecules matter more than morphology? Promises and pitfalls in parasites. Parasitology 138(13):1664-1674.

Philippe H, Brinkmann H, Lavrov DV, Littlewood DT, Manuel M, Worheide G, and Baurain D. 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol 9(3):e1000602.

Poe S, and Swofford DL. 1999. Taxon sampling revisited. Nature 398(6725):299-300.

Poinar G, Jr. 2008. Lutzomyia adiketis sp. n. (Diptera: Phlebotomidae), a vector of Paleoleishmania neotropicum sp. n. (Kinetoplastida: Trypanosomatidae) in Dominican amber. Parasit Vectors 1(1):22.

Poinar Jr G. 2007. Early Cretaceous trypanosomatids associated with fossil sand fly larvae in Burmese amber. Memórias do Instituto Oswaldo Cruz 102(5):635-637.

Poux C, Chevret P, Huchon D, de Jong WW, and Douzery EJ. 2006. Arrival and diversification of caviomorph rodents and platyrrhine primates in South America. Syst Biol 55(2):228-244.

Raymond F, Boisvert S, Roy G, Ritt J-F, Légaré D, Isnard A, Stanke M, Olivier M, Tremblay MJ, Papadopoulou B et al. . 2011. Genome sequencing of the lizard parasite Leishmania tarentolae reveals loss of genes associated to the intracellular stage of human pathogenic species. Nucleic Acids Research.

Real F, Vidal RO, Carazzolle MF, Mondego JM, Costa GG, Herai RH, Wurtele M, de Carvalho LM, Carmona e Ferreira R, Mortara RA et al. . 2013. The genome sequence of Leishmania (Leishmania) amazonensis: functional annotation and extended analysis of gene models. DNA research : an international journal for rapid publication of reports on genes and genomes 20(6):567-581.

Reuss SM, Dunbar MD, Calderwood Mays MB, Owen JL, Mallicote MF, Archer LL, and Wellehan JF, Jr. 2012. Autochthonous Leishmania siamensis in horse, Florida, USA. Emerg Infect Dis 18(9):1545-1547.

Rose K, Curtis J, Baldwin T, Mathis A, Kumar B, Sakthianandeswaren A, Spurck T, Low Choy J, and Handman E. 2004. Cutaneous leishmaniasis in red kangaroos: isolation and characterisation of the causative organisms. International Journal for Parasitology. 2004/04/28 ed. p 655-664.

Rull V. 2008. Speciation timing and neotropical biodiversity: the Tertiary-Quaternary debate in the light of molecular phylogenetic evidence. Molecular ecology 17(11):2722-2729.

Saf'janova V. 1982. The problem of taxonomy with Leishmania. Ser Protozool Sov Acad Sci Leningr 7:5-109.

Schönian G, Cupolillo E, and Mauricio I. 2013. Molecular Evolution and Phylogeny of *Leishmania*. In: Ponte-Sucre Aea, editor. Drug Resistance in Leishmania Parasites. Wien: Springer-Verlag.

Schönian G, Kuhls K, and Mauricio IL. 2011. Molecular approaches for a better understanding of the epidemiology and population genetics of Leishmania. Parasitology 138(4):405-425.

Schönian G, Mauricio I, and Cupolillo E. 2010. Is it time to revise the nomenclature of Leishmania? Trends in parasitology 26(10):466-469.

Shaw J. 1997. Ecological and evolutionary pressures on leishmanial parasites. Brazilian Journal of Genetics 20.

Springer MS, Murphy WJ, Eizirik E, and O'Brien SJ. 2003. Placental mammal diversification and the Cretaceous-Tertiary boundary. Proc Natl Acad Sci U S A 100(3):1056-1061.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312-1313.

Stevens J, and Rambaut A. 2001. Evolutionary rate differences in trypanosomes. Infect Genet Evol 1(2):143-150.

Stevens JR, Noyes HA, Schofield CJ, and Gibson W. 2001. The molecular evolution of Trypanosomatidae. Advances in parasitology 48:1-56.

Tamura K, Battistuzzi FU, Billing-Ross P, Murillo O, Filipski A, and Kumar S. 2012. Estimating divergence times in large molecular phylogenies. Proceedings of the National Academy of Sciences USA 109(47):19333-19338.

Tuon FF, Amato Neto V, and Sabbaga Amato V. 2008. Leishmania : origin, evolution and future since the Precambrian. FEMS Immunology & Medical Microbiology 54(2):158-166.

Van der Auwera G, Fraga J, Montalvo AM, and Dujardin J-C. 2011. Leishmania taxonomy up for promotion? Trends in parasitology 27(2):49-50.

Weirather JL, Jeronimo SM, Gautam S, Sundar S, Kang M, Kurtz MA, Haque R, Schriefer A, Talhari S, Carvalho EM et al. . 2011. Serial quantitative PCR assay for detection, species discrimination, and quantification of Leishmania spp. in human samples. Journal of clinical microbiology 49(11):3892-3904.

Wirth DF, and Pratt DM. 1982. Rapid identification of Leishmania species by specific hybridization of kinetoplast DNA in cutaneous lesions. Proc Natl Acad Sci U S A 79(22):6999-7003.

Wyss AR, Flynn JJ, Norell MA, Swisher CC, Charrier R, Novacek MJ, and McKenna MC. 1993. South America's earliest rodent and recognition of a new interval of mammalian evolution. Nature 365(6445):434-437.

Yang Z. 1996. Maximum-Likelihood Models for Combined Analyses of Multiple Sequence Data. Journal of molecular evolution 42(5):587-596.

Yoder JB, Briskine R, Mudge J, Farmer A, Paape T, Steele K, Weiblen GD, Bharti AK, Zhou P, May GD et al. 2013. Phylogenetic signal variation in the genomes of *Medicago* (Fabaceae). Sys Bio:Early View.

Yurchenko VY, Lukes J, Jirku M, Zeledon R, and Maslov DA. 2006. Leptomonas costaricensis sp. n. (Kinetoplastea: Trypanosomatidae), a member of the novel phylogenetic group of insect trypanosomatids closely related to the genus Leishmania. Parasitology 133(Pt 5):537-546.

CHAPTER 5
ANCIENT PATHOGEN GENOMICS: INSIGHTS INTO TIMING AND ADAPTATION

Special Issue: Ancient DNA & Human Evolution

Kelly M. Harkins[1*] and Anne C. Stone[1]

1 School of Human Evolution and Social Change, Arizona State University,

* Corresponding author

Ancient pathogen genomics: insights into timing and adaptation

Keywords: ancient DNA, pathogen evolution, human disease

Abstract

Disease is a major selective force affecting human evolution, whether a pathogen causes persistent morbidity and mortality or a sudden pandemic. Recent contributions in the field of ancient pathogen genomics have advanced our understanding of the antiquity and nature of human-pathogen interactions throughout history. Technical advancements have facilitated the recovery, enrichment and high-throughput sequencing of pathogen and parasite DNA from archived and archaeological human remains. These time-stamped genomes are crucial for calibrating molecular clocks to infer the timing of evolutionary events, while providing finer-grain resolution to phylogenetic reconstructions and complex biogeographical patterns. Additionally, genome scale data allow better identification of substitutions linked to co-evolutionary adaptations of the pathogen to their human hosts. As methodology continues to improve, ancient genomes of humans and their diverse microbiomes from a range of eras and archaeological contexts will enable population-level ancient analyses in the near future.

Introduction

The driving force behind the field of ancient pathogen genomics is the simple observation that disease is a major selective force affecting human evolution. Today, populations and individuals differ in susceptibility to disease as a direct result of our evolutionary history with pathogens. Likewise, pathogens evolve as a direct response to human biological change, as well as our

sociocultural and technological developments. One of the most influential developments in human history—the intensification of agriculture and domestication— also carried with it enduring consequences for human health. This shift, facilitated by an increase in population size, density, fertility, and eventually, expansion, is referred to as the Neolithic Demographic Transition (NDT). The NDT greatly contrasted the longstanding population structure and demographic patterns of hominin foragers, thus representing one of the most fundamental structural processes in our history (Bocquet-Appel and Bar-Yosef, 2008; Bocquet-Appel, 2002, 2011). Our understanding of pre-Neolithic population structure forms a 'baseline' from which we evaluate health and disease and develop hypotheses about the nature of human-pathogen interaction in pre- and post-Neolithic society (Armelagos et al., 2005). A demographic shift as immense as the NDT has many unintended "side effects" for human health and consequently corresponds with the 'first epidemiological transition' (Armelagos et al., 1996; Dobson and Carper, 1996; Moodley et al., 2012). As scholars of the 20th-21st centuries contemplated the origins of infectious disease, many focused upon this demographic transition to examine the persistence of old or the emergence of new pathogens.

Modern and ancient genome data are suitable for addressing key assumptions and unanswered questions associated with the first epidemiological transition. These have included, namely, whether "heirloom" diseases, present long before the NDT, co-evolved with our hominid ancestors and whether density-dependent diseases entered human populations only after the intensification of agriculture. Additionally, did the "souvenir" pathogens that we acquired after the NDT "jump" primarily from domesticated animals? Is the long-held "conventional wisdom" that avirulence is a consequence of a long co-evolutionary history in a host species true? Do pandemics signify human exposure to a "new" pathogen or genetic adaptation on the part of the pathogen? Recent research has disputed many of these assumptions and raised new questions: Which pathogens are, in fact, old? What diseases did humans encounter in the New World, and which did they carry with them? How does our identification of disease in the skeletal record affect what pathogens we target? How (and why) has virulence changed over time? Why do

pandemics occur and why do some pathogens have repeated pandemics? Both timing and genome scale data are crucial to examining these outstanding questions.

Ancient pathogen research is quickly contributing to our knowledge of infectious disease evolution by providing time-stamped sequence data to integrate into phylogenetic reconstructions and to use as crucial calibration points to estimate the timing of divergence events, thus testing long held hypotheses regarding our co-evolutionary history with pathogenic organisms. In addition to providing finer-grain resolution to robust phylogenetic inferences and complex biogeographical patterns, genome scale data are enabling us to identify substitutions linked to co-evolutionary adaptations of the pathogen to their human hosts. Over a decade of technical improvements to ancient human genomics research has provided the foundation to investigate the less explored realm of microbes. Thus, the nascent stages of ancient pathogen genomics have begun, often with surprising results, most notably this early research has focused on the 1918 “Spanish” influenza virus (Taubenberger et al., 2005), HIV-1 virus (Worobey et al. 2008), *Yersinia pestis* (Achtman, 2012; Bos et al., 2011; Schuenemann et al., 2011), leprosy (Schuenemann et al., 2013), tuberculosis (Bouwman et al., 2012; Chan et al., 2013), and most recently, cholera (Devault et al., 2014). This review examines how recent work in ancient pathogen genomics has contributed to our knowledge of infectious disease in the context of human evolution.

Fitting human-pathogen coevolution to a temporal framework

Timing is critical to understanding the emergence of a human disease and is essential for testing hypotheses about human interaction with pathogens in the course of our evolution. Ancient DNA sequences with time estimates from archaeological/historical contexts or radiocarbon dates can provide calibrations for estimating substitution rates and divergence times at important phylogenetic nodes (Drummond et al., 2002; Rambaut, 2000). Ancient and archived samples also can resolve inconsistencies when timing of older divergence events is underestimated (Wertheim

and Pond, 2011). Although calibrating evolutionary clocks is notoriously difficult, current methods account for temporally sampled sequences and the uncertainty in sample ages, which pose problems for phylogenetic and dating reconstructions (Debruyne and Poinar, 2009; Ho and Phillips, 2009; Ho et al., 2005; Ho et al., 2007; Ho et al., 2011; Ho and Larson, 2006). Sample-dating errors, however, appear to inflict only minor effects on substitution rate and dating estimates, a conclusion that Molak et al. (2013) further extrapolate to include errors inherent in calibrating radiocarbon dates to calendar ages. Having a restricted number of ancient sequences in their dataset, however, negatively impacted root age estimates. This observation is the most problematic for current ancient genetics research, considering the difficulties of obtaining multiple ancient sequences with sufficient quality and variation to incorporate into phylogenetic analyses. Population-level ancient analyses are not currently the standard, but given the current trajectory of the field, the number of ancient sequences will no longer be a limiting factor in the near future.

RNA or ssDNA viruses are ideal for studying rates of evolutionary change because of their fast mutation rate (Duffy et al. 2008), and because they change on a scale such that they are “measurably evolving” (Drummond et al., 2003). These features allow researchers to use samples collected over relatively short time periods, for example tissue samples from infected patients, to calibrate molecular clocks and test models of evolution (Holmes et al. 2003, Shapiro et al. 2010). Research on HIV evolution demonstrates the importance of historically-sampled sequences for rate calibration to help define the timescale of evolution, and in turn, the timing of divergence events in phylogenies. Methodology specifically designed for recovering degraded and damaged nucleic acid (e. g. Dabney et al., 2013a; Rohland and Hofreiter, 2007) and phylogenetics programs designed to integrate heterochronous data (Anderson et al., 2005; Drummond and Rambaut, 2007; Yang et al., 2007) have benefitted research in viral evolution. Before the AIDS epidemic, many hospitals in west-central Africa collected Bouin’s-fixed paraffin-embedded samples from patients. Ancient DNA extraction protocols facilitated the isolation of genetic information from a 1960 archived sample, DRC60. The resulting sequence information coupled with a previously successful isolation from a 1959 human blood sample, ZR.59 (Korber

et al., 2000) allowed Worobey et al. (2008) to make a number of significant conclusions, namely that HIV-1 evolves in a clock-like fashion, and that high sequence diversity within the region, in contrast to founder effects seen outside, is consistent with the hypothesis that central Africa is the site of origin.

In addition to using dates from time-stamped samples, using dates inferred from biogeographical patterns is another tool to calibrate phylogenetic analyses and convert evolutionary distance into time. Worobey et al. (2010) employ this strategy to address similar questions of rate evolution and issues of dating deeper time scales in simian immunodeficiency virus (SIV). SIV is widespread throughout African non-human primates, suggesting an ancient origin for the virus. This is in contrast to molecular clocks calibrated with the HIV sequences sampled from the 1950s and 60s that indicate an origin of SIV that is just centuries old (Wertheim and Worobey, 2009). To address this discrepancy, the authors sampled non-human primates on Bioko island, a landmass known to have separated from mainland Africa ca. 10-12 kya. Worobey et al.'s (2010) results identify the misleading effect of using recent tip dates to calibrate deeper nodes by showing that the previous rate estimates are ~125x faster than those calibrated with Bioko landmass separation. Furthermore, their results demonstrate that the SIV phylogeny is on the order of thousands, not millions (or hundreds) of years old and adjust our estimates of when SIV "jumped" into humans and evolved into HIV.

Generally, sampling of non-human primates, as with human remains, is often difficult, but in appropriate cases, it is worthwhile since non-human primates are known to harbor the closest-known relatives of many human infectious diseases (Wolfe et al., 1998). To date, there is no systematic effort to monitor pathogens emerging from animals (Wolfe et al., 2007). For this reason, very little is known about wild or natural reservoirs, or the effect of viruses and other pathogens on fitness costs across non-human primate taxa. Although Wertheim and Worobey (2009) state that SIV is generally benign in its natural host, recent research demonstrates fitness costs in chimpanzees (Keele et al., 2009). Such observations are important because symptoms

of disease may not mimic those of humans, in the case of HIV/AIDS, to have a health burden with evolutionary implications. Much remains unknown regarding the fitness costs of pathogens in natural hosts or reservoirs, and considerable scholarship and debate is devoted to understanding the relationship of pathogen virulence to long term host specificity or coevolution (see Alizon et al., 2009; Lenski and May, 1994; Read, 1994). This research underscores the importance of systematically sampling natural hosts (both modern and ancient where available) for future research on the evolution of infectious disease.

Timing is not only important for understanding zoonotic events in the past, but it also informs us about significant human transmission events, such as the emergence of pandemics. For example, hepatitis C virus (HCV) in the US currently infects between 2-3.9 million people; it is estimated that 80,000 will die annually from the infection, an average of 15 years before the general population (Mahajan et al., 2013), surpassing HIV as a cause of death for Americans (Ly et al., 2012). Many studies have attempted to estimate the origins of HCV as a human pathogen, but those dates are uncertain, ranging between 200 years to 150,000 years ago (Gonzalez-Perez et al., 1997; Simmonds, 2001, 2004; Smith et al., 1997). Although the HCV virus was only first described in the 1980s, Gray et al. (2013) were able to isolate HCV gene sequences from archived samples dating to 1953, identified in a previous study (Seeff et al., 2000) to date the origin of the US epidemic to 1901 (95% CIs: 1874-1926). Ancient genetics and phylogenetic methods have informed our understanding of other historic outbreaks, as well, including the confirmation of variola virus from mummified remains dating to a suspected smallpox pandemic in 17^{th}-18^{th} century Siberia (Biagini et al. 2012).

Pathogens other than fast evolving viruses can also be integrated into phylogenetic analyses to estimate divergence times. A disease of interest to the public because of its persistence, its clinical phenotype, and unique history of stigmatization is leprosy, also known as Hansen's disease. *Mycobacterium leprae,* the unculturable bacteria that causes the disease, was among the earliest pathogens targeted in skeletal remains because of its paleopathological signature in

bone, with first reports of *M. leprae* PCR product published in the mid-1990s (Rafi et al., 1994). Since then many publications have targeted loci within *M. leprae* to identify the pathogen in archaeological remains (Donoghue et al., 2005; Haas et al., 2000; Montiel et al., 2003; Taylor et al., 2013). The advancements used to isolate, enrich, and sequence the Neanderthal genome (Burbano et al., 2010) and an ancient *Yersinia pestis* genome (Bos et al., 2011, Schuenemann et al., 2011) were quickly applied to samples with signatures of leprosy. Considering that *M. leprae* has never been sequenced directly from a human infection, the ancient pathogen research was surprisingly successful. Schuenemann et al. (2013) used high-throughput sequencing, both with and without array-capture, to reconstruct five ancient *M. leprae* genomes from medieval leprosy cases in Europe. Due to a historically recognized decline of the disease in the 16th century, scholars had questioned whether a change in virulence could explain the decrease. Their discovery of the close similarity between these ancient and modern strains precludes a purely genetic cause for the changes in disease rates or virulence, emphasizing the importance of external factors in explaining such observations.

Comparative genomic analyses between the ancient and modern strains of *M. leprae* reveal a remarkably reduced and conserved genome, with roughly 1300 pseudogenes (Monot et al., 2009; Schuenemann et al., 2011). Using modern and ancient sequence data, tip sampling dates from 11 modern strains, and radiocarbon dates from archaeological samples, the authors calibrated substitution rates to infer that the all human leprosy strains diverged at most ~5000 years ago, consistent with paleopathological evidence but much earlier than previous estimations (Monot et al., 2005). Still unanswered is the puzzling link between the clinical disease leprosy and another infectious agent, *M. lepromatosis*, the closest known related species to *M. leprae*. *M. lepromatosis* causes leprosy-like clinical manifestations but appears to have diverged millions of years ago from M. leprae (Han et al., 2009). While modern and ancient data both support leprosy as a relatively young disease in human evolution, very few complete genomes are available to date, and there is little observed genetic diversity between available strains despite a wide geographic representation (Monot et al., 2005, Monot et al., 2009). If the conserved genetic

diversity in *M. leprae* reflects a recent bottleneck, from where and what did *M. leprae* originate or jump? How does *M. lepromatosis* fit in with the evolutionary history of the clinical disease, especially as it is recognized in the paleopathological record? Discoveries about these questions, in addition to information about currently unknown environmental reservoirs and/or vectors will further elucidate leprosy's history.

The recent discovery that *M. lepromatosis* causes the clinical disease leprosy (Han et al., 2009) raises an important point. Ancient DNA research targets pathogens based on disease recognition in skeletal remains. Until 2009, *M. leprae* was the only pathogen known to cause the skeletal indicators of leprosy. If related species can share the identical paleopathological signature, we must be cautious with experimental design, also recognizing that the absence of a pathogen may be the result of targeting the wrong causative agent. Given the millions of years of divergence between *M. leprae* and *M. lepromatosis*, researchers may benefit from more inclusive screening processes that focuses on conserved core genomes of closely, as well as more distantly related species, further emphasizing the need for broad, representative sampling, both geographically and temporally. This strategy also may impact the field of paleopathology, contributing new criteria for identifying evidence of pathogens in archaeological remains, prompting important theoretical discussion about the limits of defining health, disease, and illness in the past.

Re-classifying the "heirloom" and "souvenir" pathogens

In recent years, genetic evidence and phylogenetic reconstructions have been essential in testing hypotheses regarding the timing of pathogen emergence in human populations (see Pearce-Duvet, 2006). Pathogens thought to share a long coevolutionary history with humans or our hominid ancestors are referred to as "heirloom" species (Sprent, 1962). Alternatively, "souvenir" pathogens and parasites ostensibly were acquired after the NDT through exposure to new vectors and reservoirs as populations expanded, migrated, and changed their landscape to accommodate agricultural and pastoral practices. Before the availability of molecular data,

scholars formulated hypotheses about the origins, dispersal, causes, and attributes of diseases or pandemics based on biogeographical patterns inferred from clinical data and/or evidence of disease in the paleontological, archaeological, ethnohistoric, or historic record (e. g. Brothwell and Sandison, 1967; Cockburn, 1971; Hare, 1955; Merbs, 1992; Roberts and Manchester, 2005). It is thus thought that farming/domestication and the concomitant increase in human population size was accompanied by the emergence of several density-dependent human pathogens (Diamond, 1999, 2002; McNeill, 1976; Pearce-Duvet, 2006). The link between agriculture and the emergence of deadly pathogens prompted the traditional view that domestic animals transmitted diseases to their human caretakers (Weiss, 2001), who now had a population large enough to sustain them.

A number of important and unexpected phylogenetic analyses have rejected the traditional "animal-to-human" hypothesis. The earliest and most notable case involves the discovery that human tuberculosis is ancestral to bovine tuberculosis, refuting the popular belief that bovids transmitted the disease to humans (Brosch et al., 2002). Likewise, phylogenies of taenid tapeworms show that humans were infected long before cattle and pig domesticates, probably as a result of scavenging during the Paleolithic, supporting a non-domestic origin of the parasite (De Queiroz and Alkire, 1998; Hoberg, 2006; Hoberg et al., 2001; Hoberg et al., 2000). Taenid tapeworms are one pathogen, like lice and *H. pylori* (Falush et al., 2003; Linz et al., 2007; Moodley et al., 2012), which have shared millennia of co-evolutionary relationship with humans.

Did agriculture stimulate the emergence of some of the deadliest human diseases in history? To answer this important question we must understand disease ecologies that may influence the probability of a pathogen's jump from another host to humans. Prehistoric communities did not occupy "pristine landscapes" (Denevan, 1992). For example, if 10 million people occupied the Andes before European contact, as Denevan (1976, 1992) suggests, human impact on the environment in some areas, namely deforestation, land development, agricultural and herding practices, terracing, and urbanization would have been extensive. These very modifications to the landscape are the same forces that today affect vector survival, epidemiology, and increase

incidence for vector-borne disease and risks for infectivity (González et al., 2010; Shaw, 2007). This observation highlights the persistence of anthropogenic factors, including migration and agricultural intensification, that affect pathogen distribution, density, and interaction with humans. Each pathogen, however, shares a highly variable relationship with its host and thus benefits from individual examination. Pearce-Duvet (2006) reviewed several molecular phylogenetic analyses of pathogens for several human diseases associated with the rise of agriculture, including measles, pertussis, and falciparum malaria, and was unable to support or reject the domestic-origins hypothesis fully. This review highlights both the obstacles that define many of the research problems in pathogen evolution, mainly ambiguities caused by a limited number and diversity of available genetic data, the extinction of specific lineages, and biased signals due to demographic events (i.e. sweeps, bottlenecks), as well as the reality that each pathogen shares a unique history with humans. Technological improvements that make the field of modern and ancient genomics possible overcome such obstacles by increasing the amount of genetic data used in analyses and creating the potential to sample from extinct lineages and date the divergence of animal and human forms of disease. These improvements may enable us to look at how these pathogens acted as a selective force on humans at different time points in the past.

Heirloom treponemes

As it stands, genetic data, including complete genomes, are crucial to phylogenetic reconstructions, but they do not always resolve questions about the timing of a jump to humans or the timing of changes in virulence or patterns of transmission in an organism's evolutionary history. The treponemal diseases, for example, remain enigmatic, as genetic data have not resolved the origins scenario for syphilis (see Harper and Armelagos, 2013; Mulligan et al., 2008). *Treponema pallidum,* is the causative agent of four clinical diseases, one of which, venereal syphilis (subspecies *pallidum*), is sexually transmitted, whereas the others are not: yaws (subspecies *pertenue*), bejel (subspecies *endemicum*), and pinta (subspecies *carateum)*. The subspecies are morphologically identical and for this reason, scientists have offered alternative

biological, cultural, and climatic explanations for virulence (Hackett, 1963; Hudson, 1965). Yaws has long been thought to be an heirloom pathogen, infecting our ancestors as early as the Pleistocene (Rothschild et al., 1995) and also present in African apes (Harper et al., 2008). It remains unclear, however, whether syphilis arose in the Old World or in the New World and then was transmitted by returning explorers after contact (Hackett, 1963). This decades-long debate, reviewed in Baker and Armelagos (1988) and de Melo et al. (2010), has been revisited in recent years due to the availability of *Treponema* sequence data and the growing number of purported pre-Columbian paleopathological cases of skeletal treponematosis in the Old World (e. g. Cole and Waldron, 2011; Mays et al., 2012; Schwarz et al., 2013), although many of these cases have been refuted (Harper et al., 2011). *Treponoma*, however, are uncultivable, which continues to pose problems for obtaining taxonomically and geographically broad based sequence data for phylogenetic analyses. Numerous other factors complicate evolutionary analyses, including a lack of genetic diversity between subspecies, the similarity in skeletal lesions among the clinical diseases, the dearth of comparative sequences, the eradication of extant species and strain diversity, and the difficulties isolating ancient DNA (Barnes and Thomas, 2006; Bouwman and Brown, 2005; von Hunnius et al., 2007).

The timing of syphilis' emergence is also central to this discussion, but no studies have had the opportunity to use time-stamped sequence data or other methods of molecular calibration to test alternative hypotheses directly. Unfortunately, spirochetes are notoriously fragile bacteria, with unique structures that differ from their more robust GC-rich, gram-negative counterparts, e.g. *M. tuberculosis,* rendering them susceptible to post-mortem degradation (Cox et al., 1992; Radolf and Lukehart, 2006). Furthermore, pathogen load is low during the disease stage where skeletal changes are observable, decreasing the likelihood of isolating *Treponema* DNA from archaeological remains. While ancient DNA sequence data theoretically would contribute to this debate, the unique traits of this bacterium generate pessimism. Bouwman and Brown (2005) boldly but perhaps prematurely claim, "Ancient DNA cannot be used to study venereal syphilis" after finding no treponemal DNA in clear skeletal cases. Equally pessimistic are von Hunnius et

al. (2007:2098), who concur with both Bouwman and Brown (2005) and Barnes and Thomas (2006) to send a warning: "…the discussion presented here and elsewhere, should serve as a warning for paleopathologists who hope that the new and exciting world of ancient DNA can help them with their observations of diseases in the past". Often, the allure of the new methods, such as those in ancient genetics, is associated with the tendency for scholars to forego grounded, problem-oriented research. In the case of *Treponema*, we should be hopeful that the methods and analytic tools will catch up to the questions the research community has sought to answer for decades. For example, Montiel et al. (2012) report *T.pallidum pallidum* in neonates dating from post-European exploration Spain, suggesting that active infection with congenital syphillis, especially in neonates, results in higher bacterial load and thus, in better spirochete DNA preservation and recovery. If this is true, new sampling strategies combined with enrichment techniques may yet be useful. Alternative methods of accessing preserved pathogen DNA, for example, from dental calculus, may succeed as well.

Historically, we know that non-venereal treponematosis was found endemically worldwide. If treponemal disease diverged in Eurasia and traveled to the New World with human migrants, then pre-Columbian Old World *Treponema* genomes should harbor greater diversity, whereas New World strains would have experienced a bottleneck, as we see in others that evolved with a migrating host (e. g. Fisher et al., 2001). These are testable hypotheses if ancient *Treponema* DNA is successfully isolated and sequenced; the debate is thus ready for successes in ancient genomics.

Cholera, a free living souvenir?

Cholera, transmitted by contaminated water, is problematic for urban environments where adequate sanitation and clean water are necessary for prevention of waterborne infection, and thus the disease is often associated with the increasing population densities after the NDT. There have been seven *V. cholerae* pandemics recorded since 1817; the most recent outbreak began in

early 1960s and continues to infect humans today, having been most recently brought to world attention in the 2010 Haiti epidemic (Colwell, 1996; Grad and Waldor, 2013). Although some scholars suggested the pathogen adapted suddenly to humans before its first detection in 1817 (McNicol and Doetsch, 1983), translated historical documents and accounts of travelers, reviewed in depth by Macpherson (1884), suggest the presence of a cholera-like disease predating the first documented human pandemic by hundreds or perhaps thousands of years (e. g. Bhishagaratna, 1963; Peters, 1885 cited in Colwell, 1996). Whether these early accounts describe *V. cholerae* or a different pathogen remains unknown. However, two genetically distinct 01 biotypes are thought to be responsible for the seven main human pandemics: classical and "El Tor". El Tor emerged during the 6th pandemic in the 1960s, and it was thus assumed the classical biotype was responsible for the pandemics prior to the 1961. To test this hypothesis, Devault et al. (2014) reconstruct a *V. cholerae* genome from an ancient tissue sample archived during the second pandemic (1849) with targeted enrichment techniques and NGS, demonstrating that the second cholera pandemic was caused by a strain most-closely related to the classical *V. cholera* biotype, as opposed to the El Tor type circulating today. Due to limitations of estimating divergence time using dates from only the terminal branches of the tree, the authors use a large El Tor dataset (Mutreja et al., 2011), to estimate substitution rates. According to a strict molecular clock, the emergence of pathogenic *V. cholerae* occurred ~430 - 440 years ago; the authors, however, highlight the misleading effects of genetic recombination and saturation on dating estimates, concluding their data could also be consistent with a longer evolutionary history of human cholera, potentially coinciding with the first epidemiological transition during the NDT (Devault et al. 2014:6). Studies that integrate ancient samples into a large phylogenetic dataset inform our understanding of both past and recent epidemics. Interestingly, cholera is a free-living environmental organism and has also been used as an example of how human disease patterns respond to climate change (Colwell, 1996; Rodo et al., 2002). In the future, ancient samples may help elucidate how climactic events, like El Niño, played a role in ancient disease transmission to create awareness for prevention of potential outbreaks.

Phylogenies are not only strengthened by taxonomically-, geographically- and temporally-representative data, but data representative of the entire genome. Combining genome scale data with the archaeological or historical contexts from which they are isolated enables examinations of how pathogens evolve or how conditions change to produce a fatal human outbreak.

Understanding the evolution of pathogens

What causes a pandemic? The timing of pathogen divergence events in phylogenetic reconstructions ultimately sheds lights on the emergence of human disease throughout our history, but additional data are needed in order to infer the causative conditions, biological or environmental, that set the stage for human pandemics. Biopsy or autopsy tissue specimens from patients during influenza outbreaks and skeletal remains from associated plague contexts provide an excellent source of data from which to reconstruct the evolutionary history of pandemics. The first ancient pathogen genome was recovered from the notorious 1918 "Spanish" influenza pandemic (Taubenberger et al., 1997). Accomplished before the extensive use of NGS in aDNA research, this work represented a significant achievement and exemplifies the advantages of genome data for examining evolutionary history and origins of disease. Outbreaks of influenza A RNA viruses (genus O*rthomyxoviridae)* occur seasonally each year, but only occasionally result in extremely virulent pandemics. Most notable is the 1918 "Spanish" influenza, which caused nearly 50 million deaths worldwide (Johnson and Mueller, 2002). In the late 1990s, Taubenberger et al. (1997) first isolated fragments of the RNA virus from formalin-fixed paraffin-embedded (FFPE) tissues taken from U.S. serviceman killed in the 1918 pandemic. More recent studies sequenced fragments of the virus from the frozen lung tissue of an Inuit woman buried in Alaskan permafrost since her death in 1918 (Reid and Taubenberger, 1999) and two additional fixed samples from 1918 influenza victims at the Royal London Hospital (Reid et al., 2003b). The natural reservoir of influenza A viruses is thought to be aquatic birds (Webster et al., 1992), and thus the phylogenetic relationships among avian, mammalian, and the 1918 human strains are critical to understanding the pandemic's origins.

There are eight RNA segments in the genome, encoding at least 10 proteins. The virus phylogenies vary depending on the segments available for analysis, which increased as research progressed, from small gene fragments (Taubenberger et al. 1997), to full gene sequences (Basler et al., 2001; Reid et al., 1999; Reid et al., 2003a), and eventually all eight RNA segments—the entire genome (Taubenberger et al 2005). For example, phylogenetic analyses of the haemagglutinin (HA) gene, a crucial element involved in host interaction, always place the 1918 strain with the mammalian rather than the avian clades, despite some similarities to HA in avian viruses (Reid et al., 1999). By contrast, nonsynonymous or amino acid changes observed in the 1918 neuraminidase segment result in a phylogenetic placement with the avian viruses (Reid et al., 2000). It is important to understand the different pathways from which a lethal pandemic can emerge. In this case, the evolutionary scenario neither resembles direct reassortment nor adaptation from an intermediate host (Reid and Taubenberger, 2003), but the sequences isolated from archived samples ultimately reflect an avian influenza-like genome, suggesting an avian common ancestor, with or without an intermediate host (Taubenberger et al., 2012).

The phylogenetic ambiguity observed with the Influenza A strains reflects broader issues crucial for the evolutionary analysis of pathogens: biases caused by each site/gene's variable evolutionary history, insufficient taxonomic coverage and/or genetic diversity to resolve phylogenetic relationships, and an overall limited understanding and sampling of potential wild reservoirs/hosts. Ancient pathogen genome research must confront such challenges as it matches the speed of technological development. In 2013, a study applied NGS technology without enrichment to a previously reported archived tissue sample and produced in a single sequencing run a 3000x coverage genome of the 1918 influenza A virus. Their analyses of this genome in comparison with that of the 2009 H1N1 pandemic virus found that the 1918 virus demonstrated significantly increased inflammatory and cell death responses (Xiao et al., 2013). This technical advance further demonstrates the promise of newly developing methods with

"ancient" medical specimens and highlights the excitement of obtaining DNA/RNA from temporally distinct outbreaks. It would be of great interest to apply these new methods to different samples to examine how the virus changed over the course of this pandemic.

Perhaps one of the most famous pandemics in history is the Black Death, one of three historical plague pandemics. The Black Death decimated the European population in the 14th century, and the twin papers reporting on a *Yersinia pestis* genome from 14th century plague victims are the first and clearest successes of targeted genome enrichment in ancient pathogen DNA (Bos et al., 2011; Schuenemann et al., 2011). Although many previous studies had reported the presence of *Y. pestis* in plague contexts (e. g. Haensch et al., 2010; Raoult et al., 2000; Wiechmann et al., 2010), Bos et al. (2011) and Scheunemann et al. (2011) ultimately had two clear advantages: a historically-documented Black Death cemetery with a tight temporal constraint (AD 1348-1350) and genome-scale sequence information. They confirmed *Y. pestis* as the causative agent of the Black Death, as opposed to a separate biotype as previously hypothesized (Devignat, 1951) and concluded that its phylogenetic position placed the ancient genome as ancestral to all extant human strains (i.e. branch 1 and 2 isolates). These cemetery dates were used to calibrate the molecular clock, and their analyses traced the divergence time of all currently circulating isolates of branch 1 and 2 strains to the thirteenth century, proposed to correspond to an origin in or near China followed by plague transmission along the Silk Road (Bos et al., 2011; Morelli et al., 2010).

With a later publication, Cui et al. (2013) greatly expand the number of *Y. pestis* whole genome sequences, which impacted the estimates for clock rate variation and thus challenged Bos et al.'s (2011) divergence dates. The authors hypothesize that transmission history affects substitution rates such that epidemic periods result in more SNPs than enzootic phases, and that all *Y. pestis* strains are capable of causing disease in humans, not just branches 1 and 2. Furthermore, they emphasize that an analysis of fewer genomes or less geographically representative genomes would not have provided enough statistical power or resolution to demonstrate their observations from the full dataset. There were only 17 reference sequences available before Morelli et al.'s (2010) and Cui et al.'s (2013) collective contribution of over 300 new genomes. It is thus clear

that including the full range of species diversity is important to phylogenetic resolution. Likewise, a recent paper examines two samples from the first “Justinian” plague (AD 541-543), integrating these genomes with 131 *Y. pestis* strains from the following two pandemics to demonstrate that the earliest historic pandemic reflects an independent, and possibly extinct, outbreak compared to the Black Death 800 years later (Wagner et al., 2014). As with all paleogenetics studies, the ancient *Yersinia* publications were strengthened by the modern genome work and extensive sampling, and they highlight a weakness for studies of some current pathogens that are either unculturable (e. g. leprosy, syphilis), extinct, rare (e. g. those causing bejel, yaws, and smallpox), or neglected (e. g. those causing leishmaniasis, dengue, and schistosomiasis), in that often few comparative data are available.

The Black Death is also an important case for understanding the biological and demographic consequences of pandemics. DeWitte and Wood (2008) suggest, for example, that despite perceived virulence of the Black Death in London, the plague did not kill indiscriminately but reflected selective mortality with respect to frailty. In fact, no unique derived positions were found in the ancient genome to suggest a difference in virulence from currently circulating strains that would explain the massive death toll seen in the Black Death (Bos et al., 2011). As observed in work on HIV evolution (Worobey et al., 2010), recent research contradicts the outdated assumption that high virulence must be associated with a short-term relationship with a host (Read, 1994, reviewed in Alizon et al., 2009). If a genetic explanation is not plausible, perhaps evidence for an environmental one, e.g. adaptation to the vector, a louse in the case of the Black Death (Ayyadurai et al., 2010), is yet to be uncovered. Virulence cannot always be explained clearly with genetics, and alternative factors, such as changes in vectors, variation in immune responses among individuals, as well as social conditions that affect response to disease and transmission (Armelagos et al., 2005) can also alter the impact of the same pathogen within and among populations.

The impact of ancient migration on disease transmission

Migration is one of most influential processes impacting human-pathogen interactions; it facilitates the movement of existing human pathogens to new environments and/or communities, as well as the expansion of humans into new ecological niches, resulting in interactions with new vectors and disease reservoirs. Sequence data representative of the entire genome better facilitate inferences concerning biogeographical patterning of pathogens over time as they disperse or are carried across the landscape. In the increasingly globalized world, the sudden introduction of pathogens and parasites from a single traveler to naïve populations can have devastating effects, as witnessed in the recent 2010 Haiti cholera epidemic, which likely originated from Nepalese UN Security Forces during relief work (Chin et al., 2011; Grad and Waldor, 2013; Hendriksen et al., 2011). While globalization plays a role in current disease transmission, similar factors were important in the past, albeit at a slower pace.

As a result of European exploration and settlement, indigenous populations of the New World suffered massive demographic collapse, largely from the diseases transmitted during the "Columbian exchange" (Cook, 1998; Crosby, 1972; Thornton, 1987; Verano and Ubelaker, 1992). Any diseases present in the New World before European contact are inferred from skeletal changes consistent with current clinical manifestations. Unfortunately, the exact causative agents are often unknown, although the recent ancient genome research is contributing to debates on New World disease origins. Tuberculosis is one such disease that has been the focus of archaeological inquiry.

Although research has demonstrated that livestock did not transmit cattle-adapted tuberculosis to humans during domestication (Gordon et al. 1999; Brosch et al. 2002, Wirth et al. 2008, Comas et al. 2013), there are still outstanding questions regarding the disease's origins. Whether *M. tuberculosis* was the result of a zoonotic jump within the last few thousand years or an heirloom pathogen, emerging in Africa and dispersing across the globe with human migration is still one of

the most enduring debates in pathogen evolution. Today, human strains of *M. tuberculosis* are most diverse in Africa, while current strains in the Americas most closely resemble European strains (Gagneux and Small, 2007), leading to the hypotheses that any strains present in the New World before European exploration were either replaced or did not exist (Cockburn, 1963; Pepperell et al., 2013; Stone et al., 2009). Twentieth century medical doctors such as Hrdlička (1909), Morse (1961), and Cockburn (1963) did not accept the extra-pulmonary spinal lesion as evidence of TB (Pott's disease) in the ancient New World, and thus proposed that TB was only present in the Old World prior to the Age of Exploration. This supported the view that tuberculosis was acquired from cattle during domestication in the Old World (Cockburn, 1963; Rich, 1944). The high morbidity and mortality from TB in Native American populations following European contact also bolstered the post-contact Old World origin scenario, with New World groups considered "virgin soil" for the organism. Because Native American groups reacted as if they were epidemiologically "naíve" to the pathogen, for decades it was believed that *M. tuberculosis* could not have existed in the ancient New World (Cockburn, 1963; Hrdlička, 1909; Morse, 1961; Stead et al., 1995; Stead, 1997). However, decades of paleopathological evidence for pre-Columbian tuberculosis is now well accepted and many studies have reported the presence of *M. tuberculosis* in ancient skeletal or mummified remains (Bouwman et al., 2012; Chan et al., 2013; Donoghue et al., 2005; Müller et al., 2013; reviewed in Roberts and Buikstra, 2003).

The most recent modern genomic analysis supports an 'heirloom' status for tuberculosis by using models of demographic history to generate coalescence dates; these dates corroborate a biogeographical model of a dispersal out-of-Africa with human migration at the end of the Pleistocene (Comas et al., 2013). Nevertheless, the true identity of the infectious agent causing lesions consistent with chronic military tuberculosis in the New World and how the pathogen arrived in North and South America remains unknown. Ancient genomic methods applied to New World skeletal samples could address directly the hypotheses concerning the dispersal of the pathogen; however, genomic scale studies have only been successful in the Old World, namely with historic skeletal material from England and Hungary. The genotype of the historic strain of

tuberculosis found in remains buried at St. George's Crypt in Yorkshire, England demonstrated a close genetic relationship of the 19th century strain with the modern reference strain H37rv, not unexpected for the time and region (Bouwman et al., 2012). Given the ubiquity of environmental mycobacteria, this study also discusses important caveats associated with working with ancient tuberculosis and alleviates some obstacles by employing enrichment via hybridization capture. However, mycobacterial DNA tends to preserve well, corroborated by earlier work without enrichment (Schuenemann et al. 2013), in addition to a recent analysis of remains from a late 18th century Hungarian crypt (Chan et al. 2013). In the Hungarian remains, Chan et al. (2013) recover more mycobacterial than human DNA, and 8% of the sequencing reads likewise mapped to H37rv. Interestingly, the authors identify a mixed-strain infection in the individual, linking the strains to different 20th century European outbreaks, and highlighting the importance of recognizing signatures of co-infection in NGS data. Historic *M. tuberculosis* infection in Europe is becoming better characterized through ancient DNA; unfortunately, the origins story of how tuberculosis came to infect humans in the New World remains elusive and currently awaits successes from the field of ancient genomics.

Paleogenomics has its limitations.

Genomics is not a panacea, of course. Genomes are not always capable of resolving evolutionary relationships. Researchers also must contend with massive computational loads and face the immense task of sequence assembly, among other challenges. Ancient genomics faces the same obstacles as modern genomics but is worsened by the inherent problems of isolating degraded and damaged DNA. Aside from the profound methodological and computational challenges, there are limits to ancient DNA analyses, especially when working with skeletal remains. While such diseases as tuberculosis, leprosy, and syphilis have paleopathological signatures that develop over time, many diseases do not manifest skeletally, are fatal before they can, or are not the cause of death. Additionally, experimental design is constrained by the expectation of what pathogen to target; however, not all pathogens present in the past exist

today, and the species that most commonly cause disease in humans today may not be exclusively responsible for disease in the past. Much remains unknown about the pathogenicity of many microbes, thus the presence of a microorganism or parasite may not tell us much about disease load.

By solely focusing on skeletal remains to identify disease, we are limited to a small portion of the disease load in any population (Roberts and Manchester, 2005). This is partially due to a phenomenon Wood et al. (1992) described as the osteological paradox; in the burial record, those who appear healthy in the burial record were the sickest. Conversely, individuals with chronic signs of pathology were "healthier" in life—immunologically capable of fighting infection long enough for the formation of skeletal lesions. Reluctance to use destructive analyses on skeletal remains without pathological lesions is justifiable, but does restrict sampling to those with visible lesions, or those interred in cemeteries historically associated with plagues or epidemics (e. g. Baziotopoulou-Valavani, 2002; Harbeck et al., 2013; Schuenemann et al., 2011). A potential solution, however, is new methodology that requires very little bone or tooth powder (< 50 mg) that may mitigate some concerns regarding sample destruction. Other caveats inherent to studying disease in archaeological remains, e.g. poor burial preservation, uncertain differential diagnoses, and differential burial treatment of the sick, further limit what samples are available for ancient DNA analyses. Finally, there is the reality that some bacterial DNA and most virus RNA/DNA are not stable over time and may never be isolated from ancient remains (Bouwman and Brown, Barnes and Thomas 2006), regardless of the technological advancements. As mentioned here and examined thoroughly elsewhere (e. g. Briggs et al., 2007; Dabney et al., 2013b; Gilbert et al., 2007; Mitchell et al., 2005), ancient DNA is subject to degradation, chemical alteration, and contamination, which can complicate, or sometimes prevent extraction, sequencing, and downstream analysis. In the face of these obstacles, information from other biomolecules, e.g. mycolic acids and proteins, can complement ancient genomic studies. Tran et al. (2011) review the alternative methodologies and the resulting advances seen in ancient pathogen research.

Dental calculus, on the other hand, has unique properties that may propel the field forward by enabling the isolation of an individual's entire microbiome (as well as his or her genome) in cases where bone and tissue fail to preserve biomolecules (Preus et al., 2011). The unique human microbiomes reflect coevolutionary pressures that act both on the level of the host and on microbiota, whether mutualistic or pathogenic. Interesting comparisons could be made; for example, in pre-Neolithic foraging communities, were men and women differentially exposed to parasites and pathogens? Did this affect the composition of the microbiome? How does the pre- and post-Neolithic microbiome differ? Some of this exciting work already has begun (Adler et al., 2013).

Future Directions

Ancient pathogen research has focused heavily on methodological challenges: isolating degraded molecules that are present in miniscule proportions, less than 0.1% of a DNA extract. As our technological sophistication increases, however, other unresolved issues become apparent. For example, to interpret the evolutionary history of a pathogen, we must explicitly consider the evolution of the pathogen in terms of co-evolution with the human host. Additionally, modern and ancient human evolutionary genomics now enable the detection of changes unique to human lineages by comparing human genomes to the available Neanderthal and Denisovan genomes (see review, O'Bleness et al., 2012). Infectious disease, especially at pandemic levels, has the power to alter population demographics and affect cultural systems and modes of behavior. As discovered with the transition to farming, these changes often are accompanied by complex consequences for human health, many of which are currently being examined.

Ancient genomics is mutually beneficial to research in modern genetics, as both fields strive to identify polymorphisms that may have been important for human adaptation to pathogens or vice-versa over time. We also foresee the potential for population-level surveys of ancient human exomes and genomes in the very near future, especially with respect to identifying the selective

effects of infectious disease over time and particularly in pre/post-pandemic eras, or as a result of the transition to agriculture. On a small scale, the nascent stages of this research began a number of years ago, most notably with the investigation of the CCR5 delta-32 allele frequencies in ancient European skeletal collections (e. g. Wang et al., 2012). The allele, which confers HIV-1 resistance, is thought to have arisen during a single event between 700 and 3500 years ago, possibly coinciding with selection pressures following the Black Death in Europe or smallpox (Galvani and Slatkin, 2003; Libert et al., 1998; Martinson et al., 1997). This allele was then detected in comparable frequencies in Bronze Age skeletal remains and in post-plague remains in Central Europe, demonstrating that the Black Death was not a selective force influencing the increase in CCR5 delta-32 frequencies observed in northeastern European populations (Hummel et al., 2005) . More recently, Olalde et al. 2014 take a more sophisticated approach to detecting this allele and other polymorphisms that may have influenced susceptibility to infection in European populations. The authors report on a newly sequenced Mesolithic genome with which they examine the genotypes at a number of loci implicated in immune response and host resistance. The date of this early human genome allows the authors to examine directly hypotheses about the remodeling of the immune system purportedly due to post-Neolithic events in Europe; they find a mixed picture, however, with some variants, thought to be derived, as ancestral, while others were confirmed absent in the Mesolithic individual (Olalde et al., 2014). Recent paleogenomic research not only has questioned the antiquity of human adaptation to disease, but also has challenged the antiquity of traits thought to be adaptations on the part of the pathogen. For example, D'Costa et al. (2011) identify loci implicated in antibiotic resistance in the metagenome of 30,000-year-old Beringian permafrost sediments, refuting the notion that antibiotic resistance is an exclusively modern phenomenon. Future ancient genome and metagenome reconstructions from various eras and archaeological contexts are imminent; eventually we will see a more fine-grained picture of circulating genetic diversity at numerous points in history, and conceivably the variability of individual responses to infection within a population. These population-level inferences will be made possible by the increasing standardization of NGS methods in ancient DNA laboratories worldwide.

One of the main benefits of ancient genomics adopting NGS methods is the potential to immortalize precious material. These ancient DNA sequencing libraries are becoming standard, such that over time, the thousands of skeletal and mummified remains processed by labs worldwide will generate a new fount of samples to target the host or any number of organisms within; a single aDNA library theoretically could contribute to a multitude of interesting research questions. As costs decrease and techniques improve, one can imagine an archive of ancient DNA libraries, removing the need for future destructive analyses, and forging the way for future collaborations and research in ancient population genomics.

Conclusion

It has been asked, "does genomic evidence raise more questions than it answers?" (Harper and Armelagos, 2013). The answer is obviously, yes. Each infectious disease has a unique history. Migration, agriculture and animal domestication, anthropogenic change to the landscape, changes to cultural or institutional practices, behavior, and most recently, increasing globalization and wide-reaching human interaction have all impacted our co-evolutionary history with pathogens. Although recent genomics research has called into question exactly which diseases were introduced **before or after** the development of agriculture and domestication (see review, Harper and Armelagos, 2013), these practices indisputably had an epidemiological impact on human societies and on human biology. Archaeologists and biologists continue to debate whether the epidemiological impact resulted in increased mortality (Cohen and Armelagos, 1984; Steckel and Rose, 2002) or overall healthier populations (Pinhasi and Stock 2011, Gage and DeWitte 2009). Olalde et al. (2014) review genome-wide association studies that address the susceptibility of human populations to disease, revealing the complex relationship of background genetics on disease expression that is often population specific. The microbiome, representing the particular suite of organisms existing and co-evolving in a single human host, helps shape the individual genetic background that influences a host's immune response. Soon the field will combine

information about pathogen evolution on a particular host background and on multiple scales, i.e. community, individual, and population, to draw inferences about human health in the past, and in the future.

Methodological improvements for the isolation and sequencing of ancient genomes will continue to result in a number of important contributions to the field. Genomic-scale sequence information from historical, archaeological, or paleoanthropological timescales can better address both macro and micro-evolutionary questions, whether sourced from humans or microbes. Nevertheless, taxonomic or biogeographic coverage is as crucial as genomic coverage in accurate phylogenetic inferences, urging more comprehensive sampling strategies. Phylogenetic inference, dating calibration, disease origins, host-pathogen interactions and co-evolution, biogeographical patterns, including local extinctions, ancient population structure, and paleopathological diagnoses are a few of the research areas or methods that will benefit from continuing work in both ancient and modern pathogen genomics. Understanding the evolution of human infectious disease has practical importance as it does scholarly interest, especially in fields of medicine as we enter the era of personalized genomics. Until the next big revolution in the field of ancient DNA, enrichment, NGS and phylogenomics are an exciting way forward in ancient pathogen research.

Acknowledgements

We would like to acknowledge the following sources of funding: The National Science Foundation for NSF BCS-1063939 and BCS-0612222 (PIs: ACS and Jane E. Buikstra) and NSF Doctoral Dissertation Improvement Grant BCS-1232582 (PIs: KMH, JEB, ACS).

References cited

Achtman, M., 2012. Insights from genomic comparisons of genetically monomorphic bacterial pathogens. Philosophical Transactions of the Royal Society B: Biological Sciences 367, 860-867.

Adler, C.J., Dobney, K., Weyrich, L.S., Kaidonis, J., Walker, A.W., Haak, W., Bradshaw, C.J., Townsend, G., Soltysiak, A., Alt, K.W., Parkhill, J., Cooper, A., 2013. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nature genetics 45, 450-455, 455e451.

Alizon, S., Hurford, A., Mideo, N., Van Baalen, M., 2009. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. J Evol Biol 22, 245-259.
Anderson, C.N., Ramakrishnan, U., Chan, Y.L., Hadly, E.A., 2005. Serial SimCoal: a population genetics model for data from multiple populations and points in time. Bioinformatics 21, 1733-1734.

Armelagos, G.J., Barnes, K.C., Lin, J., 1996. Disease in human evolution: The re-emergence of infectious disease in the third epidemiological transition. AnthroNotes 18, 1-7.

Armelagos, G.J., Brown, P.J., Turner, B., 2005. Evolutionary, historical and political economic perspectives on health and disease. Soc Sci Med 61, 755-765.

Ayyadurai, S., Sebbane, F., Raoult, D., Drancourt, M., 2010. Body lice, yersinia pestis orientalis, and black death. Emerging Infectious Diseases 16, 892-893.

Baker, B.J., Armelagos, G.J., 1988. The origin and antiquity of syphilis: paleopathological diagnosis and interpretation. Current Anthropology 29, 703-738.

Barnes, I., Thomas, M.G., 2006. Evaluating bacterial pathogen DNA preservation in museum osteological collections. Proceedings of the Royal Society - biological sciences 273, 645-653.

Basler, C.F., Reid, A.H., Dybing, J.K., Janczewski, T.A., Fanning, T.G., Zheng, H., Salvatore, M., Perdue, M.L., Swayne, D.E., Garcia-Sastre, A., Palese, P., Taubenberger, J.K., 2001. Sequence of the 1918 pandemic influenza virus nonstructural gene (NS) segment and characterization of recombinant viruses bearing the 1918 NS genes. Proceedings of the National Academy of Sciences USA 98, 2746-2751.

Baziotopoulou-Valavani, E., 2002. A mass burial from the cemetery of Kerameikos. Excavating classical culture. Recent archaeological discoveries in Greece. Studies in classical archaeology I. BAR International Series 1031, 187-201.

Bhishagaratna, K.K., 1963. An English translation of Sushruta Samhita. Chaukhamba Sanskrit series Office, Varanasi.

Bocquet-Appel, J.-P., Bar-Yosef, O., 2008. The neolithic demographic transition and its consequences, 1st ed. Springer, New York.

Bocquet-Appel, J.P., 2002. Paleoanthropological traces of a neolithic demographic transition. Current Anthropology 43, 637-650.

Bocquet-Appel, J.P., 2011. When the world's population took off: the springboard of the Neolithic Demographic Transition. Science 333, 560-561.

Bos, K.I., Schuenemann, V.J., Golding, G.B., Burbano, H.A., Waglechner, N., Coombes, B.K., McPhee, J.B., Dewitte, S.N., Meyer, M., Schmedes, S., Wood, J., Earn, D.J., Herring, D.A.,

Bauer, P., Poinar, H.N., Krause, J., 2011. A draft genome of Yersinia pestis from victims of the Black Death. Nature.

Bouwman, A.S., Brown, T.A., 2005. The limits of biomolecular palaeopathology: ancient DNA cannot be used to study venereal syphilis. Journal of Archaeological Science 32, 703-713.

Bouwman, A.S., Kennedy, S.L., Muller, R., Stephens, R.H., Holst, M., Caffell, A.C., Roberts, C.A., Brown, T.A., 2012. Genotype of a historic strain of Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences USA 109, 18511-18516.

Briggs, A.W., Stenzel, U., Johnson, P.L., Green, R.E., Kelso, J., Prufer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., Paabo, S., 2007. Patterns of damage in genomic DNA sequences from a Neandertal. Proceedings of the National Academy of Sciences USA 104, 14616-14621.

Brosch, R., Gordon, S.V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L.M., Pym, A.S., Samper, S., van Soolingen, D., Cole, S.T., 2002. A new evolutionary scenario for the Mycobacterium tuberculosis complex. Proceedings of the National Academy of Sciences USA 99, 3684-3689.

Brothwell, D.R., Sandison, A.T., 1967. Disease in Antiquity: A Survey of the Diseases, Injuries and Surgery of Early Populations. Charles C. Thomas, Springfield, IL.

Burbano, H.A., Hodges, E., Green, R.E., Briggs, A.W., Krause, J., Meyer, M., Good, J.M., Maricic, T., Johnson, P.L., Xuan, Z., Rooks, M., Bhattacharjee, A., Brizuela, L., Albert, F.W., de la Rasilla, M., Fortea, J., Rosas, A., Lachmann, M., Hannon, G.J., Paabo, S., 2010. Targeted investigation of the Neandertal genome by array-based sequence capture. Science 328, 723-725.

Chan, J.Z., Sergeant, M.J., Lee, O.Y., Minnikin, D.E., Besra, G.S., Pap, I., Spigelman, M., Donoghue, H.D., Pallen, M.J., 2013. Metagenomic analysis of tuberculosis in a mummy. New England Journal of Medicine 369, 289-290.

Chin, C.S., Sorenson, J., Harris, J.B., Robins, W.P., Charles, R.C., Jean-Charles, R.R., Bullard, J., Webster, D.R., Kasarskis, A., Peluso, P., Paxinos, E.E., Yamaichi, Y., Calderwood, S.B., Mekalanos, J.J., Schadt, E.E., Waldor, M.K., 2011. The origin of the Haitian cholera outbreak strain. New England Journal of Medicine 364, 33-42.

Cockburn, A., 1963. The evolution and eradication of Infectious diseases. Cambridge Univ Press, Baltimore.

Cockburn, T.A., 1971. Infectious Diseases in Ancient Populations. Current Anthropology 12, 45-62.

Cohen, M.N., Armelagos, G.J., 1984. Paleopathology at the origins of agriculture. Academic Press, Orlando.

Cole, G., Waldron, T., 2011. Apple Down 152: a putative case of syphilis from sixth century AD Anglo-Saxon England. American Journal of Physical Anthropology 144, 72-79.

Colwell, R.R., 1996. Global climate and infectious disease: the cholera paradigm. Science 274, 2025-2031.

Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K.E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., Bothamley, G., Mei, J., Wei, L., Bentley, S., Harris, S.R., Niemann, S., Diel, R., Aseffa, A., Gao, Q., Young, D., Gagneux, S., 2013. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nature genetics 45, 1176-1182.

Cook, N.D., 1998. Born to Die: Disease and New World Conquest, 1492-1650. Cambridge University Press.

Cox, D.L., Chang, P., McDowall, A.W., Radolf, J.D., 1992. The outer membrane, not a coat of host proteins, limits antigenicity of virulent Treponema pallidum. Infect Immun 60, 1076-1083.

Crosby, A.W.J., 1972. The Columbian Exchange: Biological and Cultural Consequences of 1492. Greenwood Press, Westport, CT.

Cui, Y., Yu, C., Yan, Y., Li, D., Li, Y., Jombart, T., Weinert, L.A., Wang, Z., Guo, Z., Xu, L., Zhang, Y., Zheng, H., Qin, N., Xiao, X., Wu, M., Wang, X., Zhou, D., Qi, Z., Du, Z., Wu, H., Yang, X., Cao, H., Wang, H., Wang, J., Yao, S., Rakin, A., Li, Y., Falush, D., Balloux, F., Achtman, M., Song, Y., Wang, J., Yang, R., 2013. Historical variations in mutation rate in an epidemic pathogen, Yersinia pestis. Proceedings of the National Academy of Sciences USA 110, 577-582.
D'Costa, V.M., King, C.E., Kalan, L., Morar, M., Sung, W.W., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., Golding, G.B., Poinar, H.N., Wright, G.D., 2011. Antibiotic resistance is ancient. Nature 477, 457-461.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.T., Weihmann, A., Nickel, B., Valdiosera, C., Garcia, N., Paabo, S., Arsuaga, J.L., Meyer, M., 2013a. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences USA 110, 15758-15763.

Dabney, J., Meyer, M., Paabo, S., 2013b. Ancient DNA damage. Cold Spring Harb Perspect Biol 5.

de Melo, F.L., de Mello, J.C.M., Fraga, A.M., Nunes, K., Eggers, S., 2010. Syphilis at the Crossroad of Phylogenetics and Paleopathology. PLoS neglected tropical diseases 4.

De Queiroz, A., Alkire, N.L., 1998. The phylogenetic placement of Taenia cestodes that parasitize humans. The Journal of Parasitology 84, 379.

Debruyne, R., Poinar, H.N., 2009. Time dependency of molecular rates in ancient DNA data sets, a sampling artifact? Systematic biology 58, 348-360.

Denevan, W.M., 1976. The Native population of the Americas in 1492. The Board of Regents of the Unviersity of Wisconsin System, Madison, Wisconsin.

Denevan, W.M., 1992. The pristine myth: The landscape of the Americas in 1492. Annals of the Association of American Geographers 82, 369-385.

Devault, A.M., Golding, G.B., Waglechner, N., Enk, J.M., Kuch, M., Tien, J.H., Shi, M., Fisman, D.N., Dhody, A.N., Forrest, S., Bos, K.I., Earn, D.J., Holmes, E.C., Poinar, H.N., 2014. Second-Pandemic Strain of Vibrio cholerae from the Philadelphia Cholera Outbreak of 1849. New England Journal of Medicine.

Devignat, R., 1951. Varieties of Pasteurella pestis; new hypothesis. Bulletin of the World Health Organization 4, 247-263.

DeWitte, S.N., Wood, J.W., 2008. Selectivity of black death mortality with respect to preexisting health. Proceedings of the National Academy of Sciences USA 105, 1436-1441.

Diamond, J., 1999. Guns, Germs, and Steel: The Fates of Human Societies. W. W. Norton.

Diamond, J., 2002. Evolution, consequences and future of plant and animal domestication. Nature 418, 700-707.

Dobson, A.P., Carper, E.R., 1996. Infectious diseases and human population history. BioScience 46, 115-126.

Donoghue, H.D., Marcsik, A., Matheson, C., Vernon, K., Nuorala, E., Molto, J.E., Greenblatt, C.L., Spigelman, M., 2005. Co-infection of Mycobacterium tuberculosis and Mycobacterium leprae in human archaeological samples: a possible explanation for the historical decline of leprosy. Proceedings of the Royal Society - biological sciences 272, 389-394.

Drummond, A.J., Nicholls, G.K., Rodrigo, A.G., Solomon, W., 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161, 1307-1320.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7, 214.

Falush, D., Wirth, T., Linz, B., Pritchard, J.K., Stephens, M., Kidd, M., Blaser, M.J., Graham, D.Y., Vacher, S., Perez-Perez, G.I., Yamaoka, Y., Megraud, F., Otto, K., Reichard, U., Katzowitsch, E., Wang, X., Achtman, M., Suerbaum, S., 2003. Traces of human migrations in Helicobacter pylori populations. Science 299, 1582-1585.

Fisher, M.C., Koenig, G.L., White, T.J., San-Blas, G., Negroni, R., Alvarez, I.G., Wanke, B., Taylor, J.W., 2001. Biogeographic range expansion into South America by Coccidioides immitis mirrors New World patterns of human migration. Proceedings of the National Academy of Sciences USA 98, 4558-4562.

Gagneux, S., Small, P.M., 2007. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. Lancet Infect Dis 7, 328-337.
Galvani, A.P., Slatkin, M., 2003. Evaluating plague and smallpox as historical selective pressures for the CCR5-Delta 32 HIV-resistance allele. Proceedings of the National Academy of Sciences USA 100, 15276-15279.

Gilbert, M.T., Binladen, J., Miller, W., Wiuf, C., Willerslev, E., Poinar, H., Carlson, J.E., Leebens-Mack, J.H., Schuster, S.C., 2007. Recharacterization of ancient DNA miscoding lesions: insights in the era of sequencing-by-synthesis. Nucleic Acids Research 35, 1-10.

González, C., Wang, O., Strutz, S.E., González-Salazar, C., Sánchez-Cordero, V., Sarkar, S., 2010. Climate change and risk of leishmaniasis in north america: predictions from ecological niche models of vector and reservoir species. PLoS neglected tropical diseases 4, e585.

Gonzalez-Perez, M.A., Norder, H., Bergstrom, A., Lopez, E., Visona, K.A., Magnius, L.O., 1997. High prevalence of GB virus C strains genetically related to strains with Asian origin in Nicaraguan hemophiliacs. J Med Virol 52, 149-155.

Grad, Y.H., Waldor, M.K., 2013. Deciphering the origins and tracking the evolution of cholera epidemics with whole-genome-based molecular epidemiology. MBio 4, e00670-00613.

Gray, R.R., Tanaka, Y., Takebe, Y., Magiorkinis, G., Buskell, Z., Seeff, L., Alter, H.J., Pybus, O.G., 2013. Evolutionary analysis of hepatitis C virus gene sequences from 1953. Philosophical Transactions of the Royal Society B: Biological Sciences 368, 20130168.

Haas, C.J., Zink, A., Palfi, G., Szeimies, U., Nerlich, A.G., 2000. Detection of leprosy in ancient human skeletal remains by molecular identification of Mycobacterium leprae. Am J Clin Pathol 114, 428-436.

Hackett, C.J., 1963. On the Origin of the Human Treponematoses (Pinta, Yaws, Endemic Syphilis and Venereal Syphilis). Bull World Health Organ 29, 7-41.

Haensch, S., Bianucci, R., Signoli, M., Rajerison, M., Schultz, M., Kacki, S., Vermunt, M., Weston, D.A., Hurst, D., Achtman, M., Carniel, E., Bramanti, B., 2010. Distinct Clones of Yersinia pestis Caused the Black Death. PLoS Pathogens 6.

Han, X.Y., Sizer, K.C., Thompson, E.J., Kabanja, J., Li, J., Hu, P., Gomez-Valero, L., Silva, F.J., 2009. Comparative sequence analysis of Mycobacterium leprae and the new leprosy-causing Mycobacterium lepromatosis. J Bacteriol 191, 6067-6074.

Harbeck, M., Seifert, L., Hansch, S., Wagner, D.M., Birdsell, D., Parise, K.L., Wiechmann, I., Grupe, G., Thomas, A., Keim, P., Zoller, L., Bramanti, B., Riehm, J.M., Scholz, H.C., 2013. Yersinia pestis DNA from skeletal remains from the 6(th) century AD reveals insights into Justinianic Plague. PLoS Pathogens 9, e1003349.

Hare, R., 1955. Pomp and pestilence: infectious disease, its origins and conquest. Philosophical Library.

Harper, K.N., Armelagos, G.J., 2013. Genomics, the origins of agriculture, and our changing microbe-scape: Time to revisit some old tales and tell some new ones. American Journal of Physical Anthropology 152, 135-152.

Harper, K.N., Ocampo, P.S., Steiner, B.M., George, R.W., Silverman, M.S., Bolotin, S., Pillay, A., Saunders, N.J., Armelagos, G.J., 2008. On the origin of the treponematoses: a phylogenetic approach. PLoS neglected tropical diseases 2, e148.

Harper, K.N., Zuckerman, M.K., Harper, M.L., Kingston, J.D., Armelagos, G.J., 2011. The origin and antiquity of syphilis revisited: an appraisal of Old World pre-Columbian evidence for treponemal infection. American Journal of Physical Anthropology 146 Suppl 53, 99-133.

Hendriksen, R.S., Price, L.B., Schupp, J.M., Gillece, J.D., Kaas, R.S., Engelthaler, D.M., Bortolaia, V., Pearson, T., Waters, A.E., Upadhyay, B.P., Shrestha, S.D., Adhikari, S., Shakya, G., Keim, P.S., Aarestrup, F.M., 2011. Population genetics of Vibrio cholerae from Nepal in 2010: evidence on the origin of the Haitian outbreak. MBio 2, e00157-00111.

Ho, S.Y., Phillips, M.J., 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst Biol 58, 367-380.

Ho, S.Y., Phillips, M.J., Cooper, A., Drummond, A.J., 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Molecular biology and evolution 22, 1561-1568.

Ho, S.Y., Shapiro, B., Phillips, M.J., Cooper, A., Drummond, A.J., 2007. Evidence for time dependency of molecular rate estimates. Systematic biology 56, 515-522.

Ho, S.Y.W., Lanfear, R., Phillips, M.J., Barnes, I., Thomas, J.A., Kolokotronis, S.-O., Shapiro, B., 2011. Bayesian Estimation of Substituion Rates from Ancient DNA Sequences with Low Information Content. Systematic biology In Press.

Ho, S.Y.W., Larson, G., 2006. Molecular clocks: when times are a-changin'. Trends in Genetics 22, 79-83.

Hoberg, E.P., 2006. Phylogeny of Taenia: Species definitions and origins of human parasites. Parasitol Int 55 Suppl, S23-30.

Hoberg, E.P., Alkire, N.L., de Queiroz, A., Jones, A., 2001. Out of Africa: Origins of the Taenia tapeworms in humans. Proceedings of the Royal Society, Biological Sciences 268, 781-787.
Hoberg, E.P., Jones, A., Rausch, R.L., Eom, K.S., Gardner, S.L., 2000. A phylogenetic hypothesis for species of the genus Taenia (Eucestoda : Taeniidae). Journal of Parasitology 86, 89-98.

Hrdlička, A., 1909. Tuberculosis among certain Indian tribes of the United States: Bureau of American ethnology. Smithsonian Institute, Washington, DC.
Hudson, E.H., 1965. Treponematosis and man's social evolution. American Anthropologist 67, 885-901.

Hummel, S., Schmidt, D., Kremeyer, B., Herrmann, B., Oppermann, M., 2005. Detection of the CCR5-Delta32 HIV resistance gene in Bronze Age skeletons. Genes Immun 6, 371-374.
Johnson, N.P., Mueller, J., 2002. Updating the accounts: global mortality of the 1918-1920" Spanish" influenza pandemic. Bulletin of the History of Medicine 76, 105-115.

Keele, B.F., Jones, J.H., Terio, K.A., Estes, J.D., Rudicell, R.S., Wilson, M.L., Li, Y., Learn, G.H., Beasley, T.M., Schumacher-Stankey, J., Wroblewski, E., Mosser, A., Raphael, J., Kamenya, S., Lonsdorf, E.V., Travis, D.A., Mlengeya, T., Kinsel, M.J., Else, J.G., Silvestri, G., Goodall, J., Sharp, P.M., Shaw, G.M., Pusey, A.E., Hahn, B.H., 2009. Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. Nature 460, 515-519.
Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S., Bhattacharya, T., 2000. Timing the ancestor of the HIV-1 pandemic strains. Science 288, 1789-1796.

Lenski, R.E., May, R.M., 1994. The evolution of virulence in parasites and pathogens: reconciliation between two competing hypotheses. J Theor Biol 169, 253-265.
Libert, F., Cochaux, P., Beckman, G., Samson, M., Aksenova, M., Cao, A., Czeizel, A., Claustres, M., De La Rúa, C., Ferrari, M., 1998. The ΔCCR5 mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe. Human Molecular Genetics 7, 399-406.

Linz, B., Balloux, F., Moodley, Y., Manica, A., Liu, H., Roumagnac, P., Falush, D., Stamer, C., Prugnolle, F., van der Merwe, S.W., Yamaoka, Y., Graham, D.Y., Perez-Trallero, E., Wadstrom, T., Suerbaum, S., Achtman, M., 2007. An African origin for the intimate association between humans and Helicobacter pylori. Nature 445, 915-918.

Ly, K.N., Xing, J., Klevens, R.M., Jiles, R.B., Ward, J.W., Holmberg, S.D., 2012. The increasing burden of mortality from viral hepatitis in the United States between 1999 and 2007. Ann Intern Med 156, 271-278.

Macpherson, J., 1884. Annals of Cholera: From the Earliest Periods to the Year 1817. HK Lewis.
Mahajan, R., Xing, J., Liu, S., Ly, K., Moorman, A.C., Rupp, L., Xu, F., Holmberg, S., 2013. Rates and causes of mortality among people in care with hepatitis C virus infection--Chronic Hepatitis Cohort Study (CHeCS), 2006-2010, IDWeek 2013, San Francisco.

Martinson, J.J., Chapman, N.H., Rees, D.C., Liu, Y.T., Clegg, J.B., 1997. Global distribution of the CCR5 gene 32-basepair deletion. Nature genetics 16, 100-103.

Mays, S., Vincent, S., Meadows, J., 2012. A possible case of treponemal disease from England dating to the 11th–12th century AD. International Journal of Osteoarchaeology 22, 366-372.

McNeill, W., 1976. Plagues and Peoples. Knopf Doubleday Publishing Group.
McNicol, L., Doetsch, R., 1983. A hypothesis accounting for the origin of pandemic cholera: a retrograde analysis. Perspectives in biology and medicine 26, 547.
Merbs, C.F., 1992. A new world of infectious disease. Yearbook of Physical Anthropology 35, 3-42.

Mitchell, D., Willerslev, E., Hansen, A., 2005. Damage and repair of ancient DNA. Mutat Res 571, 265-276.

Molak, M., Lorenzen, E.D., Shapiro, B., Ho, S.Y., 2013. Phylogenetic estimation of timescales using ancient DNA: The effects of temporal sampling scheme and uncertainty in sample ages. Molecular biology and evolution 30, 253-262.

Monot, M., Honore, N., Garnier, T., Araoz, R., Coppee, J.Y., Lacroix, C., Sow, S., Spencer, J.S., Truman, R.W., Williams, D.L., Gelber, R., Virmond, M., Flageul, B., Cho, S.N., Ji, B., Paniz-Mondolfi, A., Convit, J., Young, S., Fine, P.E., Rasolofo, V., Brennan, P.J., Cole, S.T., 2005. On the origin of leprosy. Science 308, 1040-1042.

Monot, M., Honore, N., Garnier, T., Zidane, N., Sherafi, D., Paniz-Mondolfi, A., Matsuoka, M., Taylor, G.M., Donoghue, H.D., Bouwman, A., Mays, S., Watson, C., Lockwood, D., Khamesipour, A., Dowlati, Y., Jianping, S., Rea, T.H., Vera-Cabrera, L., Stefani, M.M., Banu, S., Macdonald, M., Sapkota, B.R., Spencer, J.S., Thomas, J., Harshman, K., Singh, P., Busso, P., Gattiker, A., Rougemont, J., Brennan, P.J., Cole, S.T., 2009. Comparative genomic and phylogeographic analysis of Mycobacterium leprae. Nature genetics 41, 1282-1289.
Montiel, R., Garcia, C., Canadas, M.P., Isidro, A., Guijo, J.M., Malgosa, A., 2003. DNA sequences of Mycobacterium leprae recovered from ancient bones. FEMS Microbiol Lett 226, 413-414.

Montiel, R., Solorzano, E., Diaz, N., Alvarez-Sandoval, B.A., Gonzalez-Ruiz, M., Canadas, M.P., Simoes, N., Isidro, A., Malgosa, A., 2012. Neonate human remains: a window of opportunity to the molecular study of ancient syphilis. PLoS ONE 7, e36371.

Moodley, Y., Linz, B., Bond, R.P., Nieuwoudt, M., Soodyall, H., Schlebusch, C.M., Bernhoft, S., Hale, J., Suerbaum, S., Mugisha, L., van der Merwe, S.W., Achtman, M., 2012. Age of the Association between Helicobacter pylori and Man. PLoS Pathogens 8.

Morelli, G., Song, Y.J., Mazzoni, C.J., Eppinger, M., Roumagnac, P., Wagner, D.M., Feldkamp, M., Kusecek, B., Vogler, A.J., Li, Y.J., Cui, Y.J., Thomson, N.R., Jombart, T., Leblois, R., Lichtner, P., Rahalison, L., Petersen, J.M., Balloux, F., Keim, P., Wirth, T., Ravel, J., Yang, R.F., Carniel, E., Achtman, M., 2010. Yersinia pestis genome sequencing identifies patterns of global phylogenetic diversity. Nature genetics 42, 1140-+.

Morse, D., 1961. Prehistoric tuberculosis in America. Am Rev Respir Dis 83, 489.

Müller, R., Roberts, C.A., Brown, T.A., 2013. Biomolecular identification of ancient Mycobacterium tuberculosis complex DNA in human remains from Britain and continental Europe. American Journal of Physical Anthropology.

Mulligan, C.J., Norris, S.J., Lukehart, S.A., 2008. Molecular studies in Treponema pallidum evolution: toward clarity? PLoS neglected tropical diseases 2, e184.

Mutreja, A., Kim, D.W., Thomson, N.R., Connor, T.R., Lee, J.H., Kariuki, S., Croucher, N.J., Choi, S.Y., Harris, S.R., Lebens, M., Niyogi, S.K., Kim, E.J., Ramamurthy, T., Chun, J., Wood, J.L., Clemens, J.D., Czerkinsky, C., Nair, G.B., Holmgren, J., Parkhill, J., Dougan, G., 2011. Evidence for several waves of global transmission in the seventh cholera pandemic. Nature 477, 462-465.

O'Bleness, M., Searles, V.B., Varki, A., Gagneux, P., Sikela, J.M., 2012. Evolution of genetic and genomic features unique to the human lineage. Nature Reviews Genetics 13, 853-866.

Olalde, I., Allentoft, M.E., Sanchez-Quinto, F., Santpere, G., Chiang, C.W., Degiorgio, M., Prado-Martinez, J., Rodriguez, J.A., Rasmussen, S., Quilez, J., Ramirez, O., Marigorta, U.M., Fernandez-Callejo, M., Prada, M.E., Encinas, J.M., Nielsen, R., Netea, M.G., Novembre, J., Sturm, R.A., Sabeti, P., Marques-Bonet, T., Navarro, A., Willerslev, E., Lalueza-Fox, C., 2014. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. Nature.

Pearce-Duvet, J.M., 2006. The origin of human pathogens: evaluating the role of agriculture and domestic animals in the evolution of human disease. Biol Rev Camb Philos Soc 81, 369-382.
Pepperell, C.S., Casto, A.M., Kitchen, A., Granka, J.M., Cornejo, O.E., Holmes, E.C., Birren, B., Galagan, J., Feldman, M.W., 2013. The role of selection in shaping diversity of natural M. tuberculosis populations. PLoS Pathogens 9, e1003543.

Peters, J., 1885. Early history of Asiatic cholera, in India as known to Europeans AD 1503-1800. William Wood, New York.

Preus, H.R., Marvik, O.J., Selvig, K.A., Bennike, P., 2011. Ancient bacterial DNA (aDNA) in dental calculus from archaeological human remains. Journal of Archaeological Science 38, 1827-1831.

Radolf, J.D., Lukehart, S.A., 2006. Pathogenic Treponema: Molecular and Cellular Biology. Caister Academic.

Rafi, A., Spigelman, M., Stanford, J., Lemma, E., Donoghue, H., Zias, J., 1994. Mycobacterium leprae DNA from ancient bone detected by PCR. Lancet 343, 1360-1361.

Rambaut, A., 2000. Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. Bioinformatics 16, 395-399.

Raoult, D., Aboudharam, G., Crubezy, E., Larrouy, G., Ludes, B., Drancourt, M., 2000. Molecular identification by "suicide PCR" of Yersinia pestis as the agent of Medieval Black Death. Proceedings of the National Academy of Sciences USA 97, 12800-12803.

Read, A.F., 1994. The evolution of virulence. Trends Microbiol 2, 73-76.

Reid, A.H., Fanning, T.G., Hultin, J.V., Taubenberger, J.K., 1999. Origin and evolution of the 1918 "Spanish" influenza virus hemagglutinin gene. Proceedings of the National Academy of Sciences USA 96, 1651-1656.

Reid, A.H., Fanning, T.G., Janczewski, T.A., Taubenberger, J.K., 2000. Characterization of the 1918 "Spanish" influenza virus neuraminidase gene. Proceedings of the National Academy of Sciences USA 97, 6785-6790.

Reid, A.H., Fanning, T.G., Slemons, R.D., Janczewski, T.A., Dean, J., Taubenberger, J.K., 2003a. Relationship of pre-1918 avian influenza HA and NP sequences to subsequent avian influenza strains. Avian Dis 47, 921-925.

Reid, A.H., Janczewski, T.A., Lourens, R.M., Elliot, A.J., Daniels, R.S., Berry, C.L., Oxford, J.S., Taubenberger, J.K., 2003b. 1918 influenza pandemic caused by highly conserved viruses with two receptor-binding variants. Emerging Infectious Diseases 9, 1249-1253.

Reid, A.H., Taubenberger, J.K., 1999. The 1918 flu and other influenza pandemics: "over there" and back again. Laboratory Investigation 79, 95-101.

Reid, A.H., Taubenberger, J.K., 2003. The origin of the 1918 pandemic influenza virus: a continuing enigma. Journal of General Virology 84, 2285-2292.

Rich, A.R., 1944. The Pathogenesis of Tuberculosis. Charles C. Thomas., Springfield, IL.

Roberts, C.A., Buikstra, J.E., 2003. The Bioarchaeology of Tuberculosis: A Global View on a Reemerging Disease. University Press of Florida, Gainesville, FL.

Roberts, C.A., Manchester, K., 2005. The Archaeology of Disease, 3rd edition ed. Sutton Publishing, Stroud.

Rodo, X., Pascual, M., Fuchs, G., Faruque, A.S., 2002. ENSO and cholera: a nonstationary link related to climate change? Proceedings of the National Academy of Sciences USA 99, 12901-12906.

Rohland, N., Hofreiter, M., 2007. Comparison and optimization of ancient DNA extraction. BioTechniques 42, 343-352.

Rothschild, B.M., Hershkovitz, I., Rothschild, C., 1995. Origin of yaws in the Pleistocene. Nature 378, 343-344.

Schuenemann, V.J., Bos, K., Dewitte, S., Schmedes, S., Jamieson, J., Mittnik, A., Forrest, S., Coombes, B.K., Wood, J.W., Earn, D.J., White, W., Krause, J., Poinar, H.N., 2011. From the Cover: Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of Yersinia pestis from victims of the Black Death. Proceedings of the National Academy of Sciences USA 108, E746-752.

Schuenemann, V.J., Singh, P., Mendum, T.A., Krause-Kyora, B., Jager, G., Bos, K.I., Herbig, A., Economou, C., Benjak, A., Busso, P., Nebel, A., Boldsen, J.L., Kjellstrom, A., Wu, H., Stewart, G.R., Taylor, G.M., Bauer, P., Lee, O.Y., Wu, H.H., Minnikin, D.E., Besra, G.S., Tucker, K., Roffey, S., Sow, S.O., Cole, S.T., Nieselt, K., Krause, J., 2013. Genome-wide comparison of medieval and modern Mycobacterium leprae. Science 341, 179-183.

Schwarz, S., Skytte, L., Rasmussen, K.L., 2013. Pre-Columbian treponemal infection in Denmark?-a paleopathological and archaeometric approach. Heritage Science 1, 19.

Seeff, L.B., Miller, R.N., Rabkin, C.S., Buskell-Bales, Z., Straley-Eason, K.D., Smoak, B.L., Johnson, L.D., Lee, S.R., Kaplan, E.L., 2000. 45-year follow-up of hepatitis C virus infection in healthy young adults. Ann Intern Med 132, 105-111.

Shaw, J.J., 2007. The leishmaniases- survival and expansion in a changing world. A mini-review. Memórias do Instituto Oswaldo Cruz 102, 541-547.

Simmonds, P., 2001. The origin and evolution of hepatitis viruses in humans. Journal of General Virology 82, 693-712.

Simmonds, P., 2004. Genetic diversity and evolution of hepatitis C virus--15 years on. Journal of General Virology 85, 3173-3188.

Smith, D.B., Pathirana, S., Davidson, F., Lawlor, E., Power, J., Yap, P.L., Simmonds, P., 1997. The origin of hepatitis C virus genotypes. Journal of General Virology 78 (Pt 2), 321-328.
Sprent, J.F.A., 1962. Parasitism, immunity and evolution, in: Leeper, G.S. (Ed.), The evolution of living organisms. Melbourne University Press, Melbourne, pp. 149-165.

Stead, W., Aufderheide, A., Buikstra, J., Holcomb, T., 1995. When did Mycobacterium tuberculosis infection first occur in the New World. American Journal of Respiratory and Critical Care Medicine 151, 1267-1268.

Stead, W.W., 1997. The origin and erratic global spread of tuberculosis: how the past explains the present and is the key to the future. Clin Chest Med 18, 65-77.
Steckel, R.H., Rose, J.C., 2002. The Backbone of History: Health and Nutrition in the Western Hemisphere. Cambridge University Press.

Stone, A.C., Wilbur, A.K., Buikstra, J.E., Roberts, C.A., 2009. Tuberculosis and leprosy in perspective. American Journal of Physical Anthropology 140 Suppl 49, 66-94.
Taubenberger, J.K., Baltimore, D., Doherty, P.C., Markel, H., Morens, D.M., Webster, R.G., Wilson, I.A., 2012. Reconstruction of the 1918 influenza virus: unexpected rewards from the past. MBio 3.

Taubenberger, J.K., Reid, A.H., Krafft, A.E., Bijwaard, K.E., Fanning, T.G., 1997. Initial genetic characterization of the 1918 "Spanish" influenza virus. Science 275, 1793-1796.
Taubenberger, J.K., Reid, A.H., Lourens, R.M., Wang, R., Jin, G., Fanning, T.G., 2005. Characterization of the 1918 influenza virus polymerase genes. Nature 437, 889-893.

Taylor, G.M., Tucker, K., Butler, R., Pike, A.W., Lewis, J., Roffey, S., Marter, P., Lee, O.Y., Wu, H.H., Minnikin, D.E., Besra, G.S., Singh, P., Cole, S.T., Stewart, G.R., 2013. Detection and strain typing of ancient Mycobacterium leprae from a medieval leprosy hospital. PLoS ONE 8, e62406.

Thornton, R., 1987. American Indian Holocaust and Survival: A Population History: Since 1492. University of Oklahoma Press.

Tran, T.N., Aboudharam, G., Raoult, D., Drancourt, M., 2011. Beyond ancient microbial DNA: nonnucleotidic biomolecules for paleomicrobiology. BioTechniques 50, 370-380.

Verano, J.W., Ubelaker, D.H., 1992. Disease and Demography in the Americas. Smithsonian Institution Press, Washington, D.C.
von Hunnius, T.E., Yang, D., Eng, B., Waye, J.S., Saunders, S.R., 2007. Digging deeper into the limits of ancient DNA research on syphilis. Journal of Archaeological Science 34, 2091-2100.

Wagner, D.M., Klunk, J., Harbeck, M., Devault, A., Waglechner, N., Sahl, J.W., Enk, J., Birdsell, D.N., Kuch, M., Lumibao, C., Poinar, D., Pearson, T., Fourment, M., Golding, B., Riehm, J.M., Earn, D.J.D., DeWitte, S., Rouillard, J.-M., Grupe, G., Wiechmann, I., Bliska, J.B., Keim, P.S., Scholz, H.C., Holmes, E.C., Poinar, H., 2014. Yersinia pestis and the Plague of Justinian 541-543 AD: a genomic analysis. Lancet Infect Dis.

Wang, X., Mitra, N., Secundino, I., Banda, K., Cruz, P., Padler-Karavani, V., Verhagen, A., Reid, C., Lari, M., Rizzi, E., Balsamo, C., Corti, G., De Bellis, G., Longo, L., Program, N.C.S., Beggs, W., Caramelli, D., Tishkoff, S.A., Hayakawa, T., Green, E.D., Mullikin, J.C., Nizet, V., Bui, J.,
Varki, A., 2012. Specific inactivation of two immunomodulatory SIGLEC genes during human evolution. Proceedings of the National Academy of Sciences USA 109, 9935-9940.

Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M., Kawaoka, Y., 1992. Evolution and ecology of influenza A viruses. Microbiol Rev 56, 152-179.

Weiss, R.A., 2001. Animal origins of human infectious disease. Philosophical Transactions of the Royal Society B: Biological Sciences 356, 957-977.

Wertheim, J.O., Pond, S.L.K., 2011. Purifying Selection Can Obscure the Ancient Age of Viral Lineages. Molecular biology and evolution 28, 3355-3365.

Wertheim, J.O., Worobey, M., 2009. Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2. PLoS Comput Biol 5, e1000377.
Wiechmann, I., Harbeck, M., Grupe, G., 2010. Yersinia pestis DNA Sequences in Late Medieval Skeletal Finds, Bavaria. Emerging Infectious Diseases 16, 1806-1807.

Wolfe, N.D., Dunavan, C.P., Diamond, J., 2007. Origins of major human infectious diseases. Nature 447, 279-283.

Wolfe, N.D., Escalante, A.A., Karesh, W.B., Kilbourn, A., Spielman, A., Lal, A.A., 1998. Wild primate populations in emerging infectious disease research: the missing link? Emerging Infectious Diseases 4, 149-158.

Wood, J.W., Milner, G., Harpending, J., Weiss, K., 1992. The osteological paradox. Current Anthropology 33, 343-370.

Worobey, M., Gemmel, M., Teuwen, D.E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J.J., Kabongo, J.M., Kalengayi, R.M., Van Marck, E., Gilbert, M.T., Wolinsky, S.M., 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature 455, 661-664.

Worobey, M., Telfer, P., Souquiere, S., Hunter, M., Coleman, C.A., Metzger, M.J., Reed, P., Makuwa, M., Hearn, G., Honarvar, S., Roques, P., Apetrei, C., Kazanji, M., Marx, P.A., 2010. Island biogeography reveals the deep history of SIV. Science 329, 1487.

Xiao, Y.L., Kash, J.C., Beres, S.B., Sheng, Z.M., Musser, J.M., Taubenberger, J.K., 2013. High-throughput RNA sequencing of a formalin-fixed, paraffin-embedded autopsy lung tissue sample from the 1918 influenza pandemic. Journal of Pathology 229, 535-545.

Yang, Z., O'Brien, J.D., Zheng, X., Zhu, H.Q., She, Z.S., 2007. Tree and rate estimation by local evaluation of heterochronous nucleotide data. Bioinformatics 23, 169-176.

CHAPTER 6

CONCLUSIONS

As is common in scientific scholarship, the current research resolves certain issues and also generates many new and exciting questions. Studies of ancient pathogens are moving beyond simple confirmatory analysis of diseased bone; researchers are now asking nuanced questions and utilizing novel methods capable of confronting the debates surrounding pathogen origins and evolution, and the relationships between humans and disease in the past. These questions continue to be at the forefront of not only pathogen research, but also bioarchaeological and paleopathological scholarship.

This dissertation demonstrates a bioarchaeological and molecular approach to evaluating ancient pathogens, and when possible, a marriage of the two approaches. The second and third chapters discuss the process and utility of molecular screening for adequately preserved samples via qPCR, and the subsequent approach of targeting complete MTBC genomes from ancient skeletal material for subsequent integration into phylogenetic reconstructions. The fourth chapter confronts the challenges of whole genome data for reconstructing the evolutionary history of *Leishmania* and addressing unresolved origins hypotheses in the absence of ancient DNA and/or internal calibration points for confident dating. The fifth chapter places these studies in the context of a field rapidly adapting to new technologies and considers future directions.

Issues addressed in ancient tuberculosis research

At the outset, the first two chapters highlight how critical informed paleopathological diagnoses are to downstream analyses. Although the exact causative agent is often unknown, uniformitarian principles guide our understanding of how pathogens operate and manifest skeletally in living (pre-antibiotic) patients. Likewise, these principles guide our expectations of the causative agent, and thus the experimental design of all subsequent DNA analyses, from primer design to genome

enrichment. The studies of ancient tuberculosis presented here relied first on a well-developed differential diagnosis for TB (Aufderheide and Rodríguez-Martín, 1998; Buikstra, 1976), but the design of an array method capable of capturing what was ultimately an unknown pathogen was informed by knowledge of the modern genetic diversity within 21 MTCB genomes and related mycobacteria (*M. avium* and *M. kansasii*). Ultimately, the phylogenomic reconstructions were comprised of over 270 MTBC genomes, including those infecting chimpanzee, goats, rodents, pinnipeds, and bovids (see Table S5 in Chapter 3).

One of the remarkable results of the ancient genome capture is the discovery that a member of the MTBC most closely related to M*. pinnipedii* caused the tuberculosis-like disease observed in three pre-Columbian humans associated with the Chiribaya cultural tradition in the Osmore River Valley, Peru. This result rejects the hypothesis that all precontact New World TB strains are most closely related strains in Asia. The archaeological evidence for persistent and heavy reliance on pinnipeds along the coast for thousands of years provides a plausible explanation for the contraction of *M. pinnipedii* in humans. The lack of any ancient New World genomes from outside of this valley in Peru prevents, however, a confirmation of whether this strain was transmissible among humans or whether or not it is responsible for the abundance of skeletal cases that culminate towards the end of the 1st millennium AD in both South and North America (Roberts and Buikstra, 2003). It is interesting to note that nearly 14% of all human remains recovered at four Chiribaya-related sites display evidence of tuberculosis-like disease. This percentage increases two centuries later to more than 37% of males and 14% of females at sites in the nearby Moquegua valley. These numbers prompt paleoepidemological-scale questions, i.e. whether agriculture and animal domestication, population density or other factors (or a combination of these) had an impact on Pre-Columbian groups in southern Peru that affected their interactions with the MTBC. The most pressing hypothesis to test, however, must be whether or not the causative agent of North American Pre-Columbian tuberculosis is closely related to the strains identified in Chapter 3. Future analyses following the approaches

demonstrated here, namely with samples identified during the MTBC qPCR screening process (Chapter 2) are underway.

The results of the Bayesian dating analysis were similarly surprising, supporting a divergence time for the MTBC clade no earlier than 6000 years ago. This date estimate rejects the hypothesis that *M. tuberculosis* is an heirloom pathogen that achieved its global distribution through Pleistocene migration events. The estimated substitution rate and corresponding dates are independently corroborated with an additional sample from 18th century Hungary, reported in Chan et al. (2013). This recent date reshapes our understanding of the emergence of "tuberculosis" as a human pathogen of the Holocene era and underscores the potential for other zoonotic mycobacterial transfers to infect humans and/or become endemic.

Implications for paleopathology and paleoepidemiology

The collective successes of these methods illustrate a) the utility of ancient sequences for calibrating molecular clocks and dating MTBC divergence, b) the importance of characterizing modern genetic diversity for experimental design, as well as c) the potential to reinvigorate the field of paleopathology. To expound upon the latter point, a bioarchaeological approach to disease seeks to understand population-level health status in the past. But the limits of paleopathology currently confine the number of infectious diseases that can be studied within an evolutionary context (Dustugue, 1980). Lesions associated with disseminated tuberculosis, for example, are well understood (see Roberts and Buikstra, 2003), and in ancient remains, the disease can often be diagnosed with confidence. Even so, only a minority of cases yields bony expression. This observation is explained by the principle of hidden heterogeneity, wherein individuals vary in their susceptibility to disease (Wood et al., 1992). Recognizing the majority of conditions, however, depends on the severity of infection, whether the lesions were active or healed at the time of death, the individual immune response, and coexisting pathologies (Ortner, 1998). Likewise, other conditions, i.e. those afflicting soft tissue like leishmaniasis, are obscured

in the skeletal record. By solely focusing on skeletal remains to identify disease we are limited to a small portion of the disease load in any population (Roberts and Manchester, 2005). These considerations, including the observation that selective effects of mortality influence the composition of a cemetery sample, are aspects of osteological paradox that prevent a straightforward examination of population-level health (Boldsen and Milner, 2011) and hinder the development of differential diagnoses.

In chapters 2-3, candidate ancient samples are identified for whole genome capture based first on the recognition of pathology, without which it is difficult to justify destructive analyses. The question remains, how do these few represent the whole? Epidemiological/population-level approaches require representative, large collections that are unavailable for most ancient communities, and the non-comparability of data from different sources prevent a simple compilation (Stodder, 2011). Additional sources of data that are not necessarily destructive and that extend beyond bones and mummified tissue, e.g. from dental calclus, mycolic acids, and museum samples from historic disease epidemics, to name a few (Devault et al., 2014a; Taylor et al., 2013; Warinner et al., 2014; Worobey et al., 2008), also contribute a new perspective to these paleoepidemiological questions. As “immortal” ancient DNA sequencing libraries become standard practice in paleogenetics, samples prepared for other purposes may be reclaimed to address new questions, removing the need for further sample destruction. This reallocation of aDNA libraries may have a profound effect on paleoepidemiology and paleopathology in the near future, confronting aspects of the osteological paradox directly by applying capture techniques equally across large numbers of sample libraries, irrespective of visible pathology, from one site or comparatively across many sites of interest.

In fact, newer enrichment methods can now search for hundreds or thousands of pathogenic organisms in each library (e.g. Devault et al., 2014b). The adoption of NGS techniques in paleogenetics combined with novel enrichment methods will permit examinations of disease, pathogen load, and human-pathogen co-evolution on the population-level, increasing accuracy of

prevalence estimates. These data interpreted within the bioarchaeological context will approach the paleopathology envisioned by Roberts and Manchester (1999): empirical, multidisciplinary, and holistic.

Characterizing modern pathogen evolution to inform ancient enrichment strategies

Also with the advent of NGS and the surge of genome sequencing projects comes the desire for methods of inferring evolutionary relationships from genome-scale data (e.g. Burleigh et al., 2011; Cohen and Chor, 2012). Relative to the computational challenges of a small, prokaryotic MTBC genome (~4 Mb), analysis of NGS genome data from many larger, complex eukaryotic *Leishmania* genomes is a more formidable task. Thirteen genomes were sequenced on 3 lanes of the Illumina HiSeq2000 for purposes of this project. Additional sequence reads for several species were downloaded from publicly available databases. Chapter 4 compares traditional and novel methods (i.e. SISRS pipeline (Schwartz et al., 2013)) of extracting phylogenetic informative sites from whole genome data to address outstanding issues regarding the evolutionary history of *Leishmania*. In reference to the unresolved *Leishmania* phylogeny, Shaw (1994:476) stated: "One might suggest that the ideal tree should be based on the DNA sequence of the whole genome. At the present time and possibly for a long time to come, however, this is clearly an impossible goal". Twenty years later, we've arrived in the era of the $1000 genome. Although the entire 34 Mb genome is not represented in full, 42 known or putative genes and over 300,000 genome-wide loci are used to reconstruct phylogenies of the genus, representing the largest, most inclusive phylogenetic study to date.

The placement and apparent deep divergence of a recently described species from Australia requires that this lineage must predate the split of Australia from the southern supercontinent, and further supports an origin of the genus that predates the split of Gondwana and the southern supercontinent. However, a number of limitations afflict this study, namely, that the Australian isolate is only represented by one gene, and the dataset otherwise lacks an internal calibration

point to date divergence events confidently along the tree. Loci and lineages that are used in this study are found to experience rate heterogeneity, reflected in variable and, at times, long branch lengths that cannot easily be used to infer mean substitution rates over the length of the tree. Short branches that may indicate rapid radiation events hinder phylogenetic resolution. To confront these challenges as a part of an ongoing analysis, loci that demonstrate constant rates of evolution will be extracted from the whole genome data, essentially recreating the ideal dataset to best utilize methods that implement the molecular clock (Sudhir Kumar, personal communication; Tamura et al., 2012).

The ongoing ancient leishmaniasis project, not reported in the dissertation, aims to follow the same trajectory as the ancient tuberculosis project, but alternatively focusing upon this neglected, relatively unknown parasite. It builds upon current theories and methods used to detect, isolate, and sequence ancient pathogen DNA with the ultimate goal of integrating ancient sequence data into this newly established phylogenomic framework. I have sampled eight individuals whose lesions had been previously identified as leishmaniasis from four archaeological sites AD 500-1400) near San Pedro de Atacama, with the help of Dr. Christina Torres-Rouff and two from the southern coast of Peru (AD 800-1200) with the help of Dr. Jane Buikstra. MycroArray MyBaits were designed to cover 42 genes used in the previous analysis, every 27bp on average with 3x tiling (n=2182), as well as over 219,000 variable sites in 13,000 loci, also with 3x tiling (n=34598). The design is thus intended to capture any species within the genus, much the way the TB genome array probes were designed. These probes will be used for in-solution capture of the ancient libraries generated from the skeletal remains. If successful, ancient sequence data can be aligned to loci from the previous study using the same pipeline, with the goal of determining its phylogenetic placement.

Given the restriction of mucocutaneous leishmaniasis in the New World today, I would expect ancient strains to be most closely related to species in the *L. (Viannia)* subgenus. As we witnessed with the first New World ancient MTBC genome, however, the causative agent might

represent a lineage that made a host switch from an unsuspecting animal reservoir. Recent discoveries of endemic leishmaniasis in horses and cows in the US and central Europe, for example, reveal new lineages that fall on branches formerly thought to be host-restricted to caviormorphs (Lobsiger et al., 2010; Muller et al., 2009; Reuss et al., 2012); the placement of these species sheds light on crucial parts of the evolutionary tree (e.g. *L. siamensis*). To date however, only a handful of genetic loci and no publicly accessible read data are available, preventing their inclusion in the full phylogenomic analysis.

The overrepresentation of human adapted lineages in population genetics and phylogenetics causes a taxon bias shifted away from autochthonous animal and reptile hosts. Schönian et al. (2013) call for the inclusion of more species collected from vertebrate hosts or insect vectors, or even asymptomatic human hosts. This call to action should be applied to both tuberculosis and leishmaniasis, as it is evident that there is a need for sampling strategies of zoonotic reservoirs, including primates, in endemic areas (Wolfe et al. 2007). Furthermore, the emergence of HIV co-infection in the 1980s and its associated increase in mortality (Cruz et al., 2006) has highlighted a major global health problem: opportunistic infections with species of mycobacteria (i.e. *M. avium, M. kansasii, M. ulcerans*) or *Leishmania* that are distantly-related or generally considered "non-pathogenic" (Bualert et al., 2012; Porter, 1996). This ever-increasing public health issue further emphasizes the importance of a non-biased approach to taxon sampling.

Final thoughts

The inaugural studies identifying ancient pathogen DNA from archaeological remains suffered from the limitations common to pioneering efforts. However, the period of intense criticism directed at the oversights of early paleogenetic research resulted in sweeping improvements in the field by way of stringent laboratory practices and methods for authentication (Cooper and Poinar, 2000). For years, ancient DNA studies were met with heavy scrutiny, until NGS technology revolutionized the field of molecular bioarchaeology/paleogenetics, from a

methodological, theoretical, and computational (i.e. bioinformatic) standpoint. Genomic-scale sequence data from historical, archaeological, or paleoanthropological timescales not only can be authenticated more confidently, they can also better address macro and micro-evolutionary questions, whether sourced from humans or microbes. Understanding the evolution of human infectious disease has practical importance as it does scholarly interest, especially in fields of medicine as we enter the era of personalized genomics. Soon the field will combine information about pathogen evolution amidst a unique human genetic background and on multiple scales, i.e. community, individual, and population, to draw inferences about human health in the past, and in the future.

In summary, the last chapter contextualizes the studies presented in this dissertation as it reviews the trajectory of "ancient" pathogen genomics research, considers the limitations of both human remains and ancient DNA, and contemplates current and future research directions. When available, DNA from historical or archaeological contexts enables a finer resolution to evolutionary reconstructions, offering calibration information for molecular clocks and providing insight to pathogen/host biogeography and transmission at timescales that are otherwise unobservable (Bos et al., 2011; Devault et al., 2014a). With or without ancient DNA, taxonomic or biogeographic coverage is shown to be as crucial as genomic coverage in accurate phylogenetic inferences. This observation has prompted a call for more comprehensive sampling strategies, as urged by other scholars. In the absence of molecular data, inferences about disease, pathogen load, immune responses, frailty/susceptibility, and more nuanced questions, such as the effect of sex, gender, and status on health or cultural responses to illness, are in part accessible from the bioarchaeological record (Dewitte and Hughes-Morey, 2012; Marsteller et al., 2011; Pinhasi and Turner, 2008; Roberts and Buikstra, 2003), although some contexts are more suitable than others, i.e. the particular skeletal assemblage, preservation conditions, and historical/cultural setting (Boldsen and Milner, 2011; Wood et al., 1992). While modern DNA, ancient DNA and bioarchaeological data each contribute their own temporal frame of reference, together their integration strengthens our understanding of pathogen evolution, origins and their relationship

between humans and disease in recent and in deep time. Driven by bioarchaeological inquiry, and moving at the current speed of (paleo)genomic progress, I look forward to witnessing many exciting discoveries before the next big ancient DNA revolution.

References Cited

Aufderheide, A.C., Rodríguez-Martín, C., 1998. The Cambridge encyclopedia of human paleopathology. Cambridge University Press.

Boldsen, J.L., Milner, G.R., 2011. An Epidemiological Approach to Paleopathology, in: Grauer, A.L. (Ed.), A Companion to Paleopathology. John Wiley & Sons, Malden, MA, pp. 114-132.

Bos, K.I., Schuenemann, V.J., Golding, G.B., Burbano, H.A., Waglechner, N., Coombes, B.K., McPhee, J.B., Dewitte, S.N., Meyer, M., Schmedes, S., Wood, J., Earn, D.J., Herring, D.A., Bauer, P., Poinar, H.N., Krause, J., 2011. A draft genome of Yersinia pestis from victims of the Black Death. Nature 478, 506-510.

Bualert, L., Charungkiattikul, W., Thongsuksai, P., Mungthin, M., Siripattanapipong, S., Khositnithikul, R., Naaglor, T., Ravel, C., El Baidouri, F., Leelayoova, S., 2012. Autochthonous disseminated dermal and visceral leishmaniasis in an AIDS patient, southern Thailand, caused by Leishmania siamensis. Am J Trop Med Hyg 86, 821-824.

Buikstra, J.E., 1976. Hopewell in the lower Illinois valley : a regional approach to the study of human biological variability and prehistoric behavior. Northwestern Archeological Program, Evanston, Ill.

Burleigh, J.G., Bansal, M.S., Eulenstein, O., Hartmann, S., Wehe, A., Vision, T.J., 2011. Genome-scale phylogenetics: inferring the plant tree of life from 18,896 gene trees. Syst Biol 60, 117-125.

Cohen, E., Chor, B., 2012. Detecting phylogenetic signals in eukaryotic whole genome sequences. J Comput Biol 19, 945-956.

Cooper, A., Poinar, H.N., 2000. Ancient DNA: do it right or not at all. Science 289, 1139.

Cruz, I., Nieto, J., Moreno, J., C1 nnavate, C., Desjeux, P., Alvar, J., 2006. Leishmania/HIV co-infections in the second decade. Indian J. Med. Res. 123, 357-388.

Devault, A.M., Golding, G.B., Waglechner, N., Enk, J.M., Kuch, M., Tien, J.H., Shi, M., Fisman, D.N., Dhody, A.N., Forrest, S., Bos, K.I., Earn, D.J., Holmes, E.C., Poinar, H.N., 2014a. Second-Pandemic Strain of Vibrio cholerae from the Philadelphia Cholera Outbreak of 1849. New England Journal of Medicine.

Devault, A.M., McLoughlin, K., Jaing, C., Gardner, S., Porter, T.M., Enk, J.M., Thissen, J., Allen, J., Borucki, M., DeWitte, S.N., Dhody, A.N., Poinar, H.N., 2014b. Ancient pathogen DNA in archaeological samples detected with a Microbial Detection Array. Sci Rep 4, 4245.

Dewitte, S.N., Hughes-Morey, G., 2012. Stature and frailty during the Black Death: the effect of stature on risks of epidemic mortality in London, A.D. 1348-1350. J Archaeol Sci 39.

Dustugue, J., 1980. Possibiities, Limits, and Prospects in Paleopathology of the Human Skeleton. Journal of Human Evolution 9, 3-8.

Grauer, A.L., Stuart-Macadam, P., 1998. Sex and Gender in Paleopathological Perspective. Cambridge University Press.

Lobsiger, L., Muller, N., Schweizer, T., Frey, C.F., Wiederkehr, D., Zumkehr, B., Gottstein, B., 2010. An autochthonous case of cutaneous bovine leishmaniasis in Switzerland. Vet Parasitol 169, 408-414.

Marsteller, S.J., Torres-Rouff, C., Knudson, K.J., 2011. Pre-Columbian Andean sickness ideology and the social experience of leishmaniasis: A contextualized analysis of bioarchaeological and paleopathological data from San Pedro de Atacama, Chile. International Journal of Paleopathology 1, 24-34.

Muller, N., Welle, M., Lobsiger, L., Stoffel, M.H., Boghenbor, K.K., Hilbe, M., Gottstein, B., Frey, C.F., Geyer, C., von Bomhard, W., 2009. Occurrence of Leishmania sp. in cutaneous lesions of horses in Central Europe. Vet Parasitol 166, 346-351.

Ortner, D., 1998. Male-female immune reactivity and its implications for interpreting evidence in human skeletal paleopathology, in: Grauer, A.L., Stuart-Macadam, P. (Eds.), Sex and Gender in Paleopathological Perspective. Cambridge University Press, Cambridge, pp. 79-92.

Pinhasi, R., Turner, K., 2008. Epidemiological approaches in palaeopathology, in: Pinhasi, R., Mays, S. (Eds.), Advances in Human Palaeopathology. Wiley Blackwell, pp. 45-56.

Porter, J.D., 1996. Mycobacteriosis and HIV infection: the new public health challenge. J Antimicrob Chemother 37 Suppl B, 113-120.
Reuss, S.M., Dunbar, M.D., Calderwood Mays, M.B., Owen, J.L., Mallicote, M.F., Archer, L.L., Wellehan, J.F., Jr., 2012. Autochthonous Leishmania siamensis in horse, Florida, USA. Emerg Infect Dis 18, 1545-1547.

Roberts, C.A., Buikstra, J.E., 2003. The Bioarchaeology of Tuberculosis: A Global View on a Reemerging Disease. University Press of Florida, Gainesville, FL.

Roberts, C.A., Manchester, K., 2005. The Archaeology of Disease, 3rd edition ed. Cornell University PRess, Ithaca.

Schönian, G., Cupolillo, E., Mauricio, I., 2013. Molecular Evolution and Phylogeny of *Leishmania*, in: Ponte-Sucre, A.e.a. (Ed.), Drug Resistance in Leishmania Parasites. Springer-Verlag, Wien.

Schwartz, R.S., Harkins, K., Stone, A.C., Cartwright, R.A., 2013. SISRS: SNP Identification from Short Read Sequences. arXiv preprint arXiv:1305.3665.

Shaw, J.J., 1994. Taxonomy of the genus Leishmania: present and future trends and their implications. Memórias do Instituto Oswaldo Cruz 89, 471-478.

Stodder, A.L.W., 2011. Data and Data Analysis Issues in Paleopathology, in: Grauer, A.L. (Ed.), A Companion to Paleopathology. John Wiley & Sons, Malden, MA, pp. 339-256.

Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A., Kumar, S., 2012. Estimating divergence times in large molecular phylogenies. Proceedings of the National Academy of Sciences USA 109, 19333-19338.

Taylor, G.M., Tucker, K., Butler, R., Pike, A.W., Lewis, J., Roffey, S., Marter, P., Lee, O.Y., Wu, H.H., Minnikin, D.E., Besra, G.S., Singh, P., Cole, S.T., Stewart, G.R., 2013. Detection and strain typing of ancient Mycobacterium leprae from a medieval leprosy hospital. PLoS ONE 8, e62406.

Warinner, C., Rodrigues, J.F., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R.Y., Fiddyment, S., Speller, C., Hendy, J., Charlton, S., Luder, H.U., Salazar-Garcia, D.C., Eppler, E., Seiler, R., Hansen, L.H., Castruita, J.A., Barkow-Oesterreicher, S.,
Teoh, K.Y., Kelstrup, C.D., Olsen, J.V., Nanni, P., Kawai, T., Willerslev, E., von Mering, C., Lewis, C.M., Jr., Collins, M.J., Gilbert, M.T., Ruhli, F., Cappellini, E., 2014. Pathogens and host immunity in the ancient human oral cavity. Nat Genet 46, 336-344.

Wood, J.W., Milner, G., Harpending, J., Weiss, K., 1992. The osteological paradox. Current Anthropology 33, 343-370.

Worobey, M., Gemmel, M., Teuwen, D.E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J.J., Kabongo, J.M., Kalengayi, R.M., Van Marck, E., Gilbert, M.T., Wolinsky, S.M., 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature 455, 661-664.

Complete References Cited

Achtman, M., 2008. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. Annual review of microbiology 62, 53-70.

Adler, C.J., Dobney, K., Weyrich, L.S., Kaidonis, J., Walker, A.W., Haak, W., Bradshaw, C.J., Townsend, G., Soltysiak, A., Alt, K.W., Parkhill, J., Cooper, A., 2013. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. Nature genetics 45, 450-455, 455e451.

Alizon, S., Hurford, A., Mideo, N., Van Baalen, M., 2009. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. J Evol Biol 22, 245-259.

Allison, M., Gerzten, E., Munizaga, J., Santoro, C., Mendoza, D., 1981. Tuberculosis in pre-Columbian Andean populations. , in: Buikstra, J.E. (Ed.), Prehistoric tuberculosis in the Americas. Northwestern University Archeological Program, Evanston, pp. 49-62.

Alvar, J., Velez, I.D., Bern, C., Herrero, M., Desjeux, P., Cano, J., Jannin, J., den Boer, M., 2012. Leishmaniasis worldwide and global estimates of its incidence. PLoS ONE 7.

Anawalt, P., 1992. Ancient cultural contacts between Ecuador, West Mexico, and the American southwest. Latin American Antiquity 3, 114–129.

Anderson, C.N., Ramakrishnan, U., Chan, Y.L., Hadly, E.A., 2005. Serial SimCoal: a population genetics model for data from multiple populations and points in time. Bioinformatics 21, 1733-1734.

Armelagos, G.J., Barnes, K.C., Lin, J., 1996. Disease in human evolution: The re-emergence of infectious disease in the third epidemiological transition. AnthroNotes 18, 1-7.

Armelagos, G.J., Brown, P.J., Turner, B., 2005. Evolutionary, historical and political economic perspectives on health and disease. Soc Sci Med 61, 755-765.

Arriaza, B., 1996. Preparation of the dead in coastal Andean preceramic populations, Human Mummies. Springer, pp. 131-140.

Arriaza, B.T., Reinhard, K.J., Araujo, A.G., Orellana, N.C., Standen, V.G., 2010. Possible influence of the ENSO phenomenon on the pathoecology of diphyllobothriasis and anisakiasis in ancient Chinchorro populations. Mem Inst Oswaldo Cruz 105, 66-72.

Arriaza, B.T., Salo, W., Aufderheide, A.C., Holcomb, T.A., 1995. Pre-Columbian tuberculosis in northern Chile: molecular and skeletal evidence. American Journal of Physical Anthropology 98, 37-45.

Arriaza, B.T., Standen, V.G., 2005. Differential Mortuary Treatment among the Andean Chinchorro Fishers: Social Inequalities or In Situ Regional Cultural Evolution? Current Anthropology 46, 662-671.

Aufderheide, A.C., Rodríguez-Martín, C., 1998. The Cambridge encyclopedia of human paleopathology. Cambridge University Press.

Aufderheide, A.C., Aturaliya, S., Focacci, G., 2002. Pulmonary disease in a sample of mummies from the az-75 cemetery in northern Chile s Azapa Valley. Chungara Revista de Antropología Chilena 34, 253-263

Ayyadurai, S., Sebbane, F., Raoult, D., Drancourt, M., 2010. Body lice, yersinia pestis orientalis, and black death. Emerging Infectious Diseases 16, 892-893.

Baker, B.J., Armelagos, G.J., 1988. The origin and antiquity of syphilis: paleopathological diagnosis and interpretation. Current Anthropology 29, 703-738.
Barnes, I., Thomas, M.G., 2006. Evaluating bacterial pathogen DNA preservation in museum osteological collections. Proceedings of the Royal Society - biological sciences 273, 645-653.

Basler, C.F., Reid, A.H., Dybing, J.K., Janczewski, T.A., Fanning, T.G., Zheng, H., Salvatore, M., Perdue, M.L., Swayne, D.E., Garcia-Sastre, A., Palese, P., Taubenberger, J.K., 2001. Sequence of the 1918 pandemic influenza virus nonstructural gene (NS) segment and characterization of recombinant viruses bearing the 1918 NS genes. Proceedings of the National Academy of Sciences USA 98, 2746-2751.

Bastida, R., Loureiro, J., Quse, V., Bernardelli, A., Rodriguez, D., Costa, E., 1999. Tuberculosis in a wild subantarctic fur seal from Argentina. J Wildl Dis 35, 796-798.

Bastida, R., Quse, V., Guichón, R., 2011. La tuberculosis en grupos de cazadores recolectores de Patagonia y Tierra del Fuego: nuevas alternativas de contagio a través de la fauna silvestre/Tuberculosis in hunter-gatherer groups of Patagonia and Tierra del Fuego: new alternatives of disease. Revista Argentina de Antropología Biológica 13, 83-95.

Baziotopoulou-Valavani, E., 2002. A mass burial from the cemetery of Kerameikos. Excavating classical culture. Recent archaeological discoveries in Greece. Studies in classical archaeology I. BAR International Series 1031, 187-201.

Benson, E.P., 2012. The Worlds of the Moche on the North Coast of Peru. University of Texas Press.
Berg, S., Firdessa, R., Habtamu, M., Gadisa, E., Mengistu, A., Yamuah, L., Ameni, G., Vordermeier, M., Robertson, B.D., Smith, N.H., Engers, H., Young, D., Hewinson, R.G., Aseffa, A., Gordon, S.V., 2009. The burden of mycobacterial disease in ethiopian cattle: implications for public health. PLoS ONE 4, e5068.

Bernardelli, A., Bastida, R., Loureiro, J., Michelis, H., Romano, M.I., Cataldi, A., Costa, E., 1996. Tuberculosis in sea lions and fur seals from the south-western Atlantic coast. Rev Sci Tech 15, 985-1005.

Bhishagaratna, K.K., 1963. An English translation of Sushruta Samhita. Chaukhamba Sanskrit series Office, Varanasi.

Bibar, J., 1966 [1565]. Crónica y Relación Copiosa y Verdadera de los Reinos de Chile. Fondo Histórico y Bibliográfico Jos Toribio Medina, Santiago.

Bocquet-Appel, J.-P., Bar-Yosef, O., 2008. The neolithic demographic transition and its consequences, 1st ed. Springer, New York.

Bocquet-Appel, J.P., 2002. Paleoanthropological traces of a neolithic demographic transition. Current Anthropology 43, 637-650.

Bocquet-Appel, J.P., 2011. When the world's population took off: the springboard of the Neolithic Demographic Transition. Science 333, 560-561.

Boite, M.C., Mauricio, I.L., Miles, M.A., Cupolillo, E., 2012. New insights on taxonomy, phylogeny and population genetics of Leishmania (Viannia) parasites based on multilocus sequence analysis. PLoS Negl Trop Dis 6, e1888.

Boldsen, J.L., Milner, G.R., 2011. An Epidemiological Approach to Paleopathology, in: Grauer, A.L. (Ed.), A Companion to Paleopathology. John Wiley & Sons, Malden, MA, pp. 114-132.

Bos, K.I., Schuenemann, V.J., Golding, G.B., Burbano, H.A., Waglechner, N., Coombes, B.K., McPhee, J.B., Dewitte, S.N., Meyer, M., Schmedes, S., Wood, J., Earn, D.J., Herring, D.A., Bauer, P., Poinar, H.N., Krause, J., 2011. A draft genome of Yersinia pestis from victims of the Black Death. Nature 478, 506-510.

Botella, H., Stadthagen, G., Lugo-Villarino, G., de Chastellier, C., Neyrolles, O., 2012. Metallobiology of host-pathogen interactions: an intoxicating new insight. Trends Microbiol 20, 106-112.

Bouwman, A.S., Brown, T.A., 2005. The limits of biomolecular palaeopathology: ancient DNA cannot be used to study venereal syphilis. Journal of Archaeological Science 32, 703-713.

Bouwman, A.S., Kennedy, S.L., Muller, R., Stephens, R.H., Holst, M., Caffell, A.C., Roberts, C.A., Brown, T.A., 2012. Genotype of a historic strain of Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences USA 109, 18511-18516.

Braun, M., Cook, D.C., Pfeiffer, S., 1998. DNA from Mycobacterium tuberculosis complex identified in North American, Pre-Columbian human skeletal remains. Journal of Archaeological Science 25.

Briggs, A.W., Stenzel, U., Johnson, P.L., Green, R.E., Kelso, J., Prufer, K., Meyer, M., Krause, J., Ronan, M.T., Lachmann, M., Paabo, S., 2007. Patterns of damage in genomic DNA sequences from a Neandertal. Proceedings of the National Academy of Sciences USA 104, 14616-14621.

Briggs, A.W., Stenzel, U., Meyer, M., Krause, J., Kircher, M., Paabo, S., 2010. Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. Nucleic Acids Res 38, e87.

Brosch, R., Gordon, S.V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L.M., Pym, A.S., Samper, S., van Soolingen, D., Cole, S.T., 2002. A new evolutionary scenario for the Mycobacterium tuberculosis complex. Proceedings of the National Academy of Sciences USA 99, 3684-3689.

Brothwell, D.R., Sandison, A.T., 1967. Disease in Antiquity: A Survey of the Diseases, Injuries and Surgery of Early Populations. Charles C. Thomas, Springfield, IL.

Bryant, J.M., Schurch, A.C., van Deutekom, H., Harris, S.R., de Beer, J.L., de Jager, V., Kremer, K., van Hijum, S.A., Siezen, R.J., Borgdorff, M., Bentley, S.D., Parkhill, J., van Soolingen, D., 2013. Inferring patient to patient transmission of Mycobacterium tuberculosis from whole genome sequencing data. BMC Infect Dis 13, 110.

Bualert, L., Charungkiattikul, W., Thongsuksai, P., Mungthin, M., Siripattanapipong, S., Khositnithikul, R., Naaglor, T., Ravel, C., El Baidouri, F., Leelayoova, S., 2012. Autochthonous disseminated dermal and visceral leishmaniasis in an AIDS patient, southern Thailand, caused by Leishmania siamensis. Am J Trop Med Hyg 86, 821-824.

Buikstra, J.E., 1976. Hopewell in the lower Illinois valley : a regional approach to the study of human biological variability and prehistoric behavior. Northwestern Archeological Program, Evanston, Ill.

Buikstra, J.E., 1977. Differential Diagnosis: An Epidemiological Model. Yearbook of Physical Anthropology [1976], 316-328.

Buikstra, J.E., 1981. Prehistoric tuberculosis in the Americas. Northwestern University Archaeological Program, Evanston, IL.

Buikstra, J.E., 1995. Tombs for the Living... or ...for the Dead: The Osmore Ancestors, in: Dillehay, T.D. (Ed.), Tombs for the Living: Andean Mortuary Practices. Dumbarton Oaks, Washington, D.C., pp. 229-280.

Buikstra, J.E., Tomczak, P.D., Lozada Cerna, M., Rakita, G.F., 2005. Chiribaya Political Economy: A Bioarchaeological Perspective, in: Rakita, G.F., Buikstra, J.E., Beck, L.A., Williams, S.R. (Eds.), Interacting with the dead: Perspectives on mortuary archaeology for the new millennium. University Press of Florida, Gainesville, pp. 66-80.

Buikstra, J.E., Williams, S.R., 1991. Tuberculosis in the Americas: Current perspectives, in: Ortner, D., Aufderheide, A.C. (Eds.), Human paleopathology: Current syntheses and future options. . Smithsonian Institution Press, Washington, D.C., pp. 161–172.

Burbano, H.A., Hodges, E., Green, R.E., Briggs, A.W., Krause, J., Meyer, M., Good, J.M., Maricic, T., Johnson, P.L., Xuan, Z., Rooks, M., Bhattacharjee, A., Brizuela, L., Albert, F.W., de la Rasilla, M., Fortea, J., Rosas, A., Lachmann, M., Hannon, G.J., Paabo, S., 2010. Targeted investigation of the Neandertal genome by array-based sequence capture. Science 328, 723-725.

Burleigh, J.G., Bansal, M.S., Eulenstein, O., Hartmann, S., Wehe, A., Vision, T.J., 2011. Genome-scale phylogenetics: inferring the plant tree of life from 18,896 gene trees. Syst Biol 60, 117-125.

Callen, E.O., Cameron, T.W.M., 1960. A prehistoric diet revealed in coprolites. New Scientist 8, 35-40.

Casali, N., Nikolayevskyy, V., Balabanova, Y., Ignatyeva, O., Kontsevaya, I., Harris, S.R., Bentley, S.D., Parkhill, J., Nejentsev, S., Hoffner, S.E., Horstmann, R.D., Brown, T., Drobniewski, F., 2012. Microevolution of extensively drug-resistant tuberculosis in Russia. Genome Res 22, 735-745.

Castresana, J., 2007. Topological variation in single-gene phylogenetic trees. Genome Biol 8, 216.

Chan, J.Z., Sergeant, M.J., Lee, O.Y., Minnikin, D.E., Besra, G.S., Pap, I., Spigelman, M., Donoghue, H.D., Pallen, M.J., 2013. Metagenomic analysis of tuberculosis in a mummy. New England Journal of Medicine 369, 289-290.

Choi, W.-J., Ro, H.-M., Hobbie, E.A., 2003. Patterns of natural 15N in soils and plants from chemically and organically fertilized uplands. Soil Biology and Biochemistry 35, 1493-1500.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ruden, D.M., 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6, 80-92.

Cockburn, A., 1963. The evolution and eradication of Infectious diseases. Cambridge Univ Press, Baltimore.

Cockburn, T.A., 1971. Infectious Diseases in Ancient Populations. Current Anthropology 12, 45-62.

Cohen, E., Chor, B., 2012. Detecting phylogenetic signals in eukaryotic whole genome sequences. J Comput Biol 19, 945-956.

Cohen, M.N., Armelagos, G.J., 1984. Paleopathology at the origins of agriculture. Academic Press, Orlando.

Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S., Barrell, B.G., 1998. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 393, 537-544.

Colwell, R.R., 1996. Global climate and infectious disease: the cholera paradigm. Science 274, 2025-2031.

Comas, I., Borrell, S., Roetzer, A., Rose, G., Malla, B., Kato-Maeda, M., Galagan, J., Niemann, S., Gagneux, S., 2012. Whole-genome sequencing of rifampicin-resistant Mycobacterium tuberculosis strains identifies compensatory mutations in RNA polymerase genes. Nat Genet 44, 106-110.

Comas, I., Chakravartti, J., Small, P.M., Galagan, J., Niemann, S., Kremer, K., Ernst, J.D., Gagneux, S., 2010. Human T cell epitopes of Mycobacterium tuberculosis are evolutionarily hyperconserved. Nat Genet 42, 498-503.

Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K.E., Kato-Maeda, M., Parkhill, J., Malla, B., Berg, S., Thwaites, G., Yeboah-Manu, D., Bothamley, G., Mei, J., Wei, L., Bentley, S., Harris, S.R., Niemann, S., Diel, R., Aseffa, A., Gao, Q., Young, D., Gagneux, S., 2013. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nature genetics 45, 1176-1182.

Cook, N.D., 1998. Born to Die: Disease and New World Conquest, 1492-1650. Cambridge University Press.

Cooper, A., Poinar, H.N., 2000. Ancient DNA: do it right or not at all. Science 289, 1139.

Coscolla, M., Lewin, A., Metzger, S., Maetz-Rennsing, K., Calvignac-Spencer, S., Nitsche, A., Dabrowski, P.W., Radonic, A., Niemann, S., Parkhill, J., Couacy-Hymann, E., Feldman, J., Comas, I., Boesch, C., Gagneux, S., Leendertz, F.H., 2013. Novel Mycobacterium tuberculosis complex isolate from a wild chimpanzee. Emerg Infect Dis 19, 969-976.

Covert, T.C., Rodgers, M.R., Reyes, A.L., Stelma, G.N., Jr., 1999. Occurrence of nontuberculous mycobacteria in environmental samples. Appl Environ Microbiol 65, 2492-2496.

Cox, D.L., Chang, P., McDowall, A.W., Radolf, J.D., 1992. The outer membrane, not a coat of host proteins, limits antigenicity of virulent Treponema pallidum. Infect Immun 60, 1076-1083.

Croan, D.G., Morrison, D., Ellis, J.T., 1997. Evolution of the genus Leishmania revealed by comparison of DNA and RNA polymerase gene sequences. Molecular and biochemical parasitology 89, 149-159.

Crosby, A.W.J., 1972. The Columbian Exchange: Biological and Cultural Consequences of 1492. Greenwood Press, Westport, CT.

Cruz, I., Nieto, J., Moreno, J., C1 nnavate, C., Desjeux, P., Alvar, J., 2006. Leishmania/HIV co-infections in the second decade. Indian J. Med. Res. 123, 357-388.

Cui, Y., Yu, C., Yan, Y., Li, D., Li, Y., Jombart, T., Weinert, L.A., Wang, Z., Guo, Z., Xu, L., Zhang, Y., Zheng, H., Qin, N., Xiao, X., Wu, M., Wang, X., Zhou, D., Qi, Z., Du, Z., Wu, H., Yang, X., Cao, H., Wang, H., Wang, J., Yao, S., Rakin, A., Li, Y., Falush, D., Balloux, F., Achtman, M., Song, Y., Wang, J., Yang, R., 2013. Historical variations in mutation rate in an epidemic pathogen, Yersinia pestis. Proceedings of the National Academy of Sciences USA 110, 577-582.

Cupolillo, E., Medina-Acosta, E., Noyes, H., Momen, H., Grimaldi, G., Jr., 2000. A revised classification for Leishmania and Endotrypanum. Parasitol Today 16, 142-144.

D'Costa, V.M., King, C.E., Kalan, L., Morar, M., Sung, W.W., Schwarz, C., Froese, D., Zazula, G., Calmels, F., Debruyne, R., Golding, G.B., Poinar, H.N., Wright, G.D., 2011. Antibiotic resistance is ancient. Nature 477, 457-461.

Dabbs, G.R., 2009. Resuscitating the epidemiological model of differential diagnosis: tuberculosis at prehistoric Point Hope, Alaska. Paleopathology Association Newsletter 148, 11-24.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.T., Weihmann, A., Nickel, B., Valdiosera, C., Garcia, N., Paabo, S., Arsuaga, J.L., Meyer, M., 2013a. Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences USA 110, 15758-15763.

Dabney, J., Meyer, M., Paabo, S., 2013b. Ancient DNA damage. Cold Spring Harb Perspect Biol 5.

Darling, A.C., Mau, B., Blattner, F.R., Perna, N.T., 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14, 1394-1403.

de Melo, F.L., de Mello, J.C.M., Fraga, A.M., Nunes, K., Eggers, S., 2010. Syphilis at the Crossroad of Phylogenetics and Paleopathology. PLoS neglected tropical diseases 4.

De Queiroz, A., Alkire, N.L., 1998. The phylogenetic placement of Taenia cestodes that parasitize humans. The Journal of Parasitology 84, 379.

Debruyne, R., Poinar, H.N., 2009. Time dependency of molecular rates in ancient DNA data sets, a sampling artifact? Systematic biology 58, 348-360.

DeFrance, S.D., Alvarez, A.U., 2004. Quebrada Tacahuay: un sitio marítimo del Pleistoceno Tardío en la costa sur del Perú. Chungara: Revista de Antropología Chilena 36, 257-278.

Denevan, W.M., 1976. The Native population of the Americas in 1492. The Board of Regents of the Unviersity of Wisconsin System, Madison, Wisconsin.

Denevan, W.M., 1992. The pristine myth: The landscape of the Americas in 1492. Annals of the Association of American Geographers 82, 369-385.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M., Sivachenko, A.Y.,
Cibulskis, K., Gabriel, S.B., Altshuler, D., Daly, M.J., 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43, 491-498.

Devault, A.M., Golding, G.B., Waglechner, N., Enk, J.M., Kuch, M., Tien, J.H., Shi, M., Fisman, D.N., Dhody, A.N., Forrest, S., Bos, K.I., Earn, D.J., Holmes, E.C., Poinar, H.N., 2014a. Second-Pandemic Strain of Vibrio cholerae from the Philadelphia Cholera Outbreak of 1849. New England Journal of Medicine.

Devault, A.M., McLoughlin, K., Jaing, C., Gardner, S., Porter, T.M., Enk, J.M., Thissen, J., Allen, J.,
Borucki, M., DeWitte, S.N., Dhody, A.N., Poinar, H.N., 2014b. Ancient pathogen DNA in archaeological samples detected with a Microbial Detection Array. Sci Rep 4, 4245.

Devignat, R., 1951. Varieties of Pasteurella pestis; new hypothesis. Bulletin of the World Health Organization 4, 247-263.

DeWitte, S.N., Wood, J.W., 2008. Selectivity of black death mortality with respect to preexisting health. Proceedings of the National Academy of Sciences USA 105, 1436-1441.

Diamond, J., 1999. Guns, Germs, and Steel: The Fates of Human Societies. W. W. Norton.

Diamond, J., 2002. Evolution, consequences and future of plant and animal domestication. Nature 418, 700-707.

Dillehay, T.D., 1999. The late Pleistocene cultures of South America. Evolutionary Anthropology: Issues, News, and Reviews 7, 206-216.

Dobson, A.P., Carper, E.R., 1996. Infectious diseases and human population history. BioScience 46, 115-126.

Domogalla, J., Prodinger, W.M., Blum, H., Krebs, S., Gellert, S., Muller, M., Neuendorf, E., Sedlmaier, F., Buttner, M., 2013. Region of difference 4 in alpine Mycobacterium caprae isolates indicates three variants. J Clin Microbiol 51, 1381-1388.

Donnan, C., 1995. Moche funerary practice, Tombs for the Living; Andean Mortuary practices, pp. 111-159.

Donoghue, H.D., Marcsik, A., Matheson, C., Vernon, K., Nuorala, E., Molto, J.E., Greenblatt, C.L., Spigelman, M., 2005. Co-infection of Mycobacterium tuberculosis and Mycobacterium leprae in human archaeological samples: a possible explanation for the historical decline of leprosy. Proceedings of the Royal Society - biological sciences 272, 389-394.

Donoghue, M.J., Doyle, J.A., Gauthier, J., Kluge, A.G., Rowe, T., 1989. The importance of fossils in phylogeny reconstruction. Annual Review of Ecology and Systematics 20, 431-460.

Dougall, A.M., Alexander, B., Holt, D.C., Harris, T., Sultan, A.H., Bates, P.A., Rose, K., Walton, S.F., 2011. Evidence incriminating midges (Diptera: Ceratopogonidae) as potential vectors of Leishmania in Australia. International journal for parasitology 41, 571-579.

Drummond, A.J., Nicholls, G.K., Rodrigo, A.G., Solomon, W., 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. Genetics 161, 1307-1320.

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular biology and evolution 29, 1969-1973.

Dustugue, J., 1980. Possibiities, Limits, and Prospects in Paleopathology of the Human Skeleton. Journal of Human Evolution 9, 3-8.

El-Najjar, M.Y., 1979. Human treponematosis and tuberculosis: evidence from the New World. Am J Phys Anthropol 51, 599-618.

Falush, D., Wirth, T., Linz, B., Pritchard, J.K., Stephens, M., Kidd, M., Blaser, M.J., Graham, D.Y., Vacher, S., Perez-Perez, G.I., Yamaoka, Y., Megraud, F., Otto, K., Reichard, U., Katzowitsch, E., Wang, X., Achtman, M., Suerbaum, S., 2003. Traces of human migrations in Helicobacter pylori populations. Science 299, 1582-1585.

Farris, J.S., 1970. Methods for computing Wagner trees. Systematic biology 19, 83-92.
Felsenstein, J., 1988. Phylogenies from molecular sequences: inference and reliability. Annu Rev Genet 22, 521-565.

Fernandes, A.P., Nelson, K., Beverley, S.M., 1993. Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism. Proc Natl Acad Sci U S A 90, 11608-11612.

Ferreira, L.F., de Araujo, A.J., Confalonieri, U.E., Nunez, L., 1984. The finding of eggs of Diphyllobothrium in human coprolites (4,100-1,950 B.C.) from northern Chile. Mem Inst Oswaldo Cruz 79, 175-180.

Filho, J.D., Brazil, R.P., 2003. Relationships of new world phlebotomine sand flies (Diptera: Psychodidae) based on fossil evidence. Mem Inst Oswaldo Cruz 98 Suppl 1, 145-149.

Fisher, M.C., Koenig, G.L., White, T.J., San-Blas, G., Negroni, R., Alvarez, I.G., Wanke, B., Taylor, J.W., 2001. Biogeographic range expansion into South America by Coccidioides immitis mirrors New World patterns of human migration. Proceedings of the National Academy of Sciences USA 98, 4558-4562.

Flynn, J.J., Wyss, A.R., 1998. Recent advances in South American mammalian paleontology. Trends Ecol Evol 13, 449-454.

Fraga, J., Montalvo, A.M., De Doncker, S., Dujardin, J.-C., Van der Auwera, G., 2010. Phylogeny of Leishmania species based on the heat-shock protein 70 gene. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases 10, 238-245.

Friedrich, K.M., Nemec, S., Czerny, C., Fischer, H., Plischke, S., Gahleitner, A., Viola, T.B., Imhof, H., Seidler, H., Guillen, S., 2010. The story of 12 Chachapoyan mummies through multidetector computed tomography. Eur J Radiol 76, 143-150.

Gadagkar, S.R., Rosenberg, M.S., Kumar, S., 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. J Exp Zool B Mol Dev Evol 304, 64-74.

Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B.C., Narayanan, S., Nicol, M., Niemann, S., Kremer, K., Gutierrez, M.C., Hilty, M., Hopewell, P.C., Small, P.M., 2006. Variable host-pathogen compatibility in Mycobacterium tuberculosis. Proceedings of the National Academy of Sciences USA 103, 2869-2873.

Gagneux, S., Small, P.M., 2007. Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development. Lancet Infect Dis 7, 328-337.
Galagan, J.E., 2014. Genomic insights into tuberculosis. Nature Reviews Genetics Early View.

Galati, E., 1995. Phylogenetic systematics of Phlebotominae (Diptera, Psychodidae) with emphasis on American groups. Bol Direc Malariol y San Amb 35, 133-142.

Galvani, A.P., Slatkin, M., 2003. Evaluating plague and smallpox as historical selective pressures for the CCR5-Delta 32 HIV-resistance allele. Proceedings of the National Academy of Sciences USA 100, 15276-15279.

Ghersi, H., 1956. Informe sobre las excavaciones en Chiribaya. Revista del Museo Nactional 25, 89-119.

Gilbert, M.T., Binladen, J., Miller, W., Wiuf, C., Willerslev, E., Poinar, H., Carlson, J.E., Leebens-Mack, J.H., Schuster, S.C., 2007. Recharacterization of ancient DNA miscoding lesions: insights in the era of sequencing-by-synthesis. Nucleic Acids Research 35, 1-10.

Goldman, R.C., 2013. Why are membrane targets discovered by phenotypic screens and genome sequencing in Mycobacterium tuberculosis? Tuberculosis (Edinb) 93, 569-588.

Goldstein, P., 2005. Andean Diaspora: The Tiwanaku Colonies and the Origins of Andean Empire. University Press of Florida, Gainesville.

González, C., Wang, O., Strutz, S.E., González-Salazar, C., Sánchez-Cordero, V., Sarkar, S., 2010. Climate change and risk of leishmaniasis in north america: predictions from ecological niche models of vector and reservoir species. PLoS neglected tropical diseases 4, e585.

Gonzalez-Perez, M.A., Norder, H., Bergstrom, A., Lopez, E., Visona, K.A., Magnius, L.O., 1997. High prevalence of GB virus C strains genetically related to strains with Asian origin in Nicaraguan hemophiliacs. J Med Virol 52, 149-155.

Grad, Y.H., Waldor, M.K., 2013. Deciphering the origins and tracking the evolution of cholera epidemics with whole-genome-based molecular epidemiology. MBio 4, e00670-00613.

Grauer, A.L., 2011. A Companion to Paleopathology. John Wiley & Sons, Malden, MA.

Grauer, A.L., Stuart-Macadam, P., 1998. Sex and Gender in Paleopathological Perspective. Cambridge University Press.

Gray, R.R., Tanaka, Y., Takebe, Y., Magiorkinis, G., Buskell, Z., Seeff, L., Alter, H.J., Pybus, O.G., 2013. Evolutionary analysis of hepatitis C virus gene sequences from 1953. Philosophical Transactions of the Royal Society B: Biological Sciences 368, 20130168.

Haas, C.J., Zink, A., Palfi, G., Szeimies, U., Nerlich, A.G., 2000. Detection of leprosy in ancient human skeletal remains by molecular identification of Mycobacterium leprae. Am J Clin Pathol 114, 428-436.

Hackett, C.J., 1963. On the Origin of the Human Treponematoses (Pinta, Yaws, Endemic Syphilis and Venereal Syphilis). Bull World Health Organ 29, 7-41.

Haensch, S., Bianucci, R., Signoli, M., Rajerison, M., Schultz, M., Kacki, S., Vermunt, M., Weston, D.A., Hurst, D., Achtman, M., Carniel, E., Bramanti, B., 2010. Distinct Clones of Yersinia pestis Caused the Black Death. PLoS Pathogens 6.

Han, X.Y., Sizer, K.C., Thompson, E.J., Kabanja, J., Li, J., Hu, P., Gomez-Valero, L., Silva, F.J., 2009. Comparative sequence analysis of Mycobacterium leprae and the new leprosy-causing Mycobacterium lepromatosis. J Bacteriol 191, 6067-6074.

Harbeck, M., Seifert, L., Hansch, S., Wagner, D.M., Birdsell, D., Parise, K.L., Wiechmann, I., Grupe, G., Thomas, A., Keim, P., Zoller, L., Bramanti, B., Riehm, J.M., Scholz, H.C., 2013. Yersinia pestis DNA from skeletal remains from the 6(th) century AD reveals insights into Justinianic Plague. PLoS Pathogens 9, e1003349.

Hare, R., 1955. Pomp and pestilence: infectious disease, its origins and conquest. Philosophical Library.

Harper, K.N., Armelagos, G.J., 2013. Genomics, the origins of agriculture, and our changing microbe-scape: Time to revisit some old tales and tell some new ones. American Journal of Physical Anthropology 152, 135-152.

Harper, K.N., Ocampo, P.S., Steiner, B.M., George, R.W., Silverman, M.S., Bolotin, S., Pillay, A., Saunders, N.J., Armelagos, G.J., 2008. On the origin of the treponematoses: a phylogenetic approach. PLoS neglected tropical diseases 2, e148.

Harper, K.N., Zuckerman, M.K., Harper, M.L., Kingston, J.D., Armelagos, G.J., 2011. The origin and antiquity of syphilis revisited: an appraisal of Old World pre-Columbian evidence for treponemal infection. American Journal of Physical Anthropology 146 Suppl 53, 99-133.

Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. Mol Biol Evol 27, 570-580.

Hendriksen, R.S., Price, L.B., Schupp, J.M., Gillece, J.D., Kaas, R.S., Engelthaler, D.M., Bortolaia, V., Pearson, T., Waters, A.E., Upadhyay, B.P., Shrestha, S.D., Adhikari, S., Shakya, G., Keim, P.S., Aarestrup, F.M., 2011. Population genetics of Vibrio cholerae from Nepal in 2010: evidence on the origin of the Haitian outbreak. MBio 2, e00157-00111.

Herring, D.A., Sattenspiel, L., 2007. Social contexts, syndemics, and infectious disease in northern Aboriginal populations. Am J Hum Biol 19, 190-202.

Hershberg, R., Lipatov, M., Small, P.M., Sheffer, H., Niemann, S., Homolka, S., Roach, J.C., Kremer, K., Petrov, D.a., Feldman, M.W., Gagneux, S., 2008. High functional diversity in Mycobacterium tuberculosis driven by genetic drift and human demography. PLoS biology 6, e311.

Ho, S.Y., Phillips, M.J., 2009. Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst Biol 58, 367-380.

Ho, S.Y., Phillips, M.J., Cooper, A., Drummond, A.J., 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. Molecular biology and evolution 22, 1561-1568.

Ho, S.Y., Shapiro, B., Phillips, M.J., Cooper, A., Drummond, A.J., 2007. Evidence for time dependency of molecular rate estimates. Systematic biology 56, 515-522.

Ho, S.Y.W., Lanfear, R., Phillips, M.J., Barnes, I., Thomas, J.A., Kolokotronis, S.-O., Shapiro, B., 2011. Bayesian Estimation of Substituion Rates from Ancient DNA Sequences with Low Information Content. Systematic biology In Press.

Ho, S.Y.W., Larson, G., 2006. Molecular clocks: when times are a-changin'. Trends in Genetics 22, 79-83.

Hoberg, E.P., 2006. Phylogeny of Taenia: Species definitions and origins of human parasites. Parasitol Int 55 Suppl, S23-30.

Hoberg, E.P., Alkire, N.L., de Queiroz, A., Jones, A., 2001. Out of Africa: Origins of the Taenia tapeworms in humans. Proceedings of the Royal Society, Biological Sciences 268, 781-787.

Hoberg, E.P., Jones, A., Rausch, R.L., Eom, K.S., Gardner, S.L., 2000. A phylogenetic hypothesis for species of the genus Taenia (Eucestoda : Taeniidae). Journal of Parasitology 86, 89-98.

Hodges, E., Rooks, M., Xuan, Z., Bhattacharjee, A., Benjamin Gordon, D., Brizuela, L., Richard McCombie, W., Hannon, G.J., 2009. Hybrid selection of discrete genomic intervals on custom-designed microarrays for massively parallel sequencing. Nat Protoc 4, 960-974.

Hogg, A.G., Hua, Q., Blackwell, P.G., Niu, M., Buck, C.E., Guilderson, T.P., Heaton, T.J., Palmer, J.G., Reimer, P.J., Reimer, R.W., Turney, C.S.M., Zimmerman, S.R.H., 2013. SHCal13 Southern Hemisphere Calibration, 0–50,000 Years cal BP.

Holiday, D.M., Guillen, S., Richardson, D.J., 2003. Diphyllobothriasis of the Chiribaya Culture (700–1476 AD) of Southern Peru. Comparative Parasitology 70, 167-171.

Holloway, K.L., Henneberg, R.J., de Barros Lopes, M., Henneberg, M., 2011. Evolution of human tuberculosis: a systematic review and meta-analysis of paleopathological evidence. Homo 62, 402-458.

Hosler, D., 1988. Ancient West Mexican Metallurgy: South and Central American Origins and West Mexican Transformations. American Anthropologist 90.

Hrdlička, A., 1909. Tuberculosis among certain Indian tribes of the United States: Bureau of American ethnology. Smithsonian Institute, Washington, DC.

Hudson, E.H., 1965. Treponematosis and man's social evolution. American Anthropologist 67, 885-901.

Hummel, S., Schmidt, D., Kremeyer, B., Herrmann, B., Oppermann, M., 2005. Detection of the CCR5-Delta32 HIV resistance gene in Bronze Age skeletons. Genes Immun 6, 371-374.

Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S.M., Bianchettin, G., Borzym, K., Bothe, G., Bruschi, C.V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R.M., Cronin, A., Cruz, A.K., Davies, R.M., De Gaudenzi, J., Dobson, D.E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A.C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J.C., Muller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'Neil, S., Pentony, M., Pohl, T.M., Price, C., Purnelle, B., Quail, M.A., Rabbinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J.C., Rutter, S., Saunders, D., Schafer, M., Schein, J., Schwartz, D.C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou, S., Zimmermann, W., Smith, D.F., Blackwell, J.M., Stuart, K.D., Barrell, B., Myler, P.J., 2005. The genome of the kinetoplastid parasite, Leishmania major. Science 309, 436-442.

Jablonski, D., 2001. Lessons from the past: evolutionary impacts of mass extinctions. Proc Natl Acad Sci 98, 5393-5398.

Johnson, N.P., Mueller, J., 2002. Updating the accounts: global mortality of the 1918-1920" Spanish" influenza pandemic. Bulletin of the History of Medicine 76, 105-115.

Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L., Orlando, L., 2013. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29, 1682-1684.

Julien, C., 1985. Guano and resource control in sixteenth-century Arequipa, in: Masuda, S., Shimada, I., Morris, C.P. (Eds.), Andean ecology and civilization. University of Tokyo Press, Tokyo, pp. 185-231.

Junier, T., Zdobnov, E.M., 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics 26, 1669-1670.

Jurczynski, K., Lyashchenko, K.P., Gomis, D., Moser, I., Greenwald, R., Moisson, P., 2011. Pinniped tuberculosis in Malayan tapirs (Tapirus indicus) and its transmission to other terrestrial mammals. J Zoo Wildl Med 42, 222-227.

Kanjanopas, K., Siripattanapipong, S., Ninsaeng, U., Hitakarun, A., Jitkaew, S., Kaewtaphaya, P., Tan-ariya, P., Mungthin, M., Charoenwong, C., Leelayoova, S., 2013. Sergentomyia (Neophlebotomus) gemmea, a potential vector of Leishmania siamensis in southern Thailand. BMC Infect Dis 13, 333.

Keele, B.F., Jones, J.H., Terio, K.A., Estes, J.D., Rudicell, R.S., Wilson, M.L., Li, Y., Learn, G.H., Beasley, T.M., Schumacher-Stankey, J., Wroblewski, E., Mosser, A., Raphael, J., Kamenya, S., Lonsdorf, E.V., Travis, D.A., Mlengeya, T., Kinsel, M.J., Else, J.G., Silvestri, G., Goodall, J., Sharp, P.M., Shaw, G.M., Pusey, A.E., Hahn, B.H., 2009. Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. Nature 460, 515-519.

Kelley, M.A., Micozzi, M.S., 1984. Rib lesions in chronic pulmonary tuberculosis. Am J Phys Anthropol 65, 381-386.

Kerr, S.F., 2000. Palaearctic origin of Leishmania. Memórias do Instituto Oswaldo Cruz 95, 75-80.

Kerr, S.F., 2006. Molecular trees of trypanosomes incongruent with fossil records of hosts. Memórias do Instituto Oswaldo Cruz 101, 25-30.

Kerr, S.F., Merkelz, R., Mackinnon, C., 2000. Further support for a Palaearctic origin of Leishmania. Memórias do Instituto Oswaldo Cruz 95, 579-581.

Kimura, M., 1968. Evolutionary rate at the molecular level. Nature 217, 624-626.

Kircher, M., Sawyer, S., Meyer, M., 2011. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Research 40, e3.

Kircher, M., Stenzel, U., Kelso, J., 2009. Improved base calling for the Illumina Genome Analyzer using machine learning strategies. Genome Biol 10, R83.

Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S., Bhattacharya, T., 2000. Timing the ancestor of the HIV-1 pandemic strains. Science 288, 1789-1796.

Krause, J., Briggs, A.W., Kircher, M., Maricic, T., Zwyns, N., Derevianko, A., Paabo, S., 2010. A complete mtDNA genome of an early modern human from Kostenki, Russia. Curr Biol 20, 231-236.

Kumar, S., Filipski, A.J., Battistuzzi, F.U., Kosakovsky Pond, S.L., Tamura, K., 2012. Statistics and truth
in phylogenomics. Mol Biol Evol 29, 457-472.
Labandeira, C.C., Sepkoski, J.J., Jr., 1993. Insect diversity in the fossil record. Science 261, 310-315.

Lainson, R., 1997. On Leishmania enriettii and other enigmatic Leishmania species of the Neotropics. Mem Inst Oswaldo Cruz 92, 377-387.

Lainson, R., 2010. The Neotropical Leishmania species: a brief historical review of their discovery, ecology and taxonomy. Rev Pan-Amaz Saude 1, 13-32.

Lainson, R., Shaw, J.J., 1972. Leishmaniasis of the New World: taxonomic problems. Br Med Bull 28, 44-48.

Lainson, R., Shaw, J.J., 1987. Evolution, classification and geographical distribution, The Leishmaniasis in Biology and Epidemiology. Academic Press, London, pp. 1-120.

Lanfear, R., Calcott, B., Ho, S.Y., Guindon, S., 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol Biol Evol 29, 1695-1701.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357-359.

Lenski, R.E., May, R.M., 1994. The evolution of virulence in parasites and pathogens: reconciliation between two competing hypotheses. J Theor Biol 169, 253-265.

Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst Biol 50, 913-925.

Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754-1760.

Libert, F., Cochaux, P., Beckman, G., Samson, M., Aksenova, M., Cao, A., Czeizel, A., Claustres, M., De La Rúa, C., Ferrari, M., 1998. The ΔCCR5 mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe. Human Molecular Genetics 7, 399-406.

Lindahl, T., Nyberg, B., 1974. Heat-induced deamination of cytosine residues in deoxyribonucleic acid. Biochemistry 13, 3405-3410.

Linz, B., Balloux, F., Moodley, Y., Manica, A., Liu, H., Roumagnac, P., Falush, D., Stamer, C., Prugnolle, F., van der Merwe, S.W., Yamaoka, Y., Graham, D.Y., Perez-Trallero, E., Wadstrom, T., Suerbaum, S., Achtman, M., 2007. An African origin for the intimate association between humans and Helicobacter pylori. Nature 445, 915-918.

Lipoldova, M., Demant, P., 2006. Genetic susceptibility to infectious disease: lessons from mouse models of leishmaniasis. Nat Rev Genet 7, 294-305.

Lobsiger, L., Muller, N., Schweizer, T., Frey, C.F., Wiederkehr, D., Zumkehr, B., Gottstein, B., 2010. An autochthonous case of cutaneous bovine leishmaniasis in Switzerland. Vet Parasitol 169, 408-414.

Lombardi, G.P., Cáceres, U.G., 2000. Multisystemic Tuberculosis in a Pre-Columbian Peruvian Mummy: Four Dagnostic levels, and a paleoepidemiological hypothesis. Chungará (Arica) 32, 55-60.

Lozada, M.C., 1998. The Señorío of Chiribaya: A Bio-archaeological Study in the Osmore Drainage of Southern Perú. University of Chicago, Department of Anthropology.

Lozada, M.C., Buikstra, J.E., 2005. Pescadores y Labradores among the Chiribaya of Southern Perú. Intentional Cranial Deformation and Ethnicity., in: Reycraft, R.M. (Ed.), Us and Them: The Archaeology of Ethnicity in the Andes. University of California Archaeological Research Facility, Los Angeles, CA.

Lozada, M.C., Buikstra, J.E., Rakita, G., Wheeler, J.C., 2009. Camelid herders: the forgotten specialists in the coastal senorio of Chiribaya, southern Peru. Andean Civilization: A Tribute to Michael E. Moseley Marcus J, Williams PR (eds). Cotsen Institute of Archaeology: Los Angeles, 351-364.

Lukes, J., Mauricio, I.L., Schönian, G., Dujardin, J.-C., Soteriadou, K., Dedet, J.-P., Kuhls, K., Tintaya, K.W.Q., Jirků, M., Chocholová, E., Haralambous, C., Pratlong, F., Oborniik, M., Horák, A., Ayala, F.J., Miles, M.A., 2007. Evolutionary and geographical history of the Leishmania donovani complex with a revision of current taxonomy. Proceedings of the National Academy of Sciences USA 104, 9375-9380.

Ly, K.N., Xing, J., Klevens, R.M., Jiles, R.B., Ward, J.W., Holmberg, S.D., 2012. The increasing burden of mortality from viral hepatitis in the United States between 1999 and 2007. Ann Intern Med 156, 271-278.

Lysenko, A., 1971. Distribution of leishmaniasis in the Old World. Bulletin of the World Health Organization 44, 515-520.

Macpherson, J., 1884. Annals of Cholera: From the Earliest Periods to the Year 1817. HK Lewis.
Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. Syst Biol 55, 21-30.

Mahajan, R., Xing, J., Liu, S., Ly, K., Moorman, A.C., Rupp, L., Xu, F., Holmberg, S., 2013. Rates and causes of mortality among people in care with hepatitis C virus infection--Chronic Hepatitis Cohort Study (CHeCS), 2006-2010, IDWeek 2013, San Francisco.

Marcili, A., Speranca, M.A., da Costa, A.P., Madeira, M.D., Soares, H.S., Sanches, C.D., Acosta, I.D., Girotto, A., Minervino, A.H., Horta, M.C., Shaw, J.J., Gennari, S.M., 2014. Phylogenetic relationships of Leishmania species based on trypanosomatid barcode (SSU rDNA) and gGAPDH genes: Taxonomic revision of Leishmania (L.) infantum chagasi in South America. Infect Genet Evol.

Marfurt, J., Niederwieser, I., Makia, N.D., Beck, H.-P., Felger, I., 2003. Diagnostic genotyping of Old and New World Leishmania species by PCR-RFLP. Diagnostic Microbiology and Infectious Disease 46, 115-124.

Maricic, T., Whitten, M., Paabo, S., 2010. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. PLoS ONE 5, e14004.

Marmiesse, M., Brodin, P., Buchrieser, C., Gutierrez, C., Simoes, N., Vincent, V., Glaser, P., Cole, S.T., Brosch, R., 2004. Macro-array and bioinformatic analyses reveal mycobacterial 'core' genes, variation in the ESAT-6 gene family and new phylogenetic markers for the Mycobacterium tuberculosis complex. Microbiology 150, 483-496.

Marsteller, S.J., Torres-Rouff, C., Knudson, K.J., 2011. Pre-Columbian Andean sickness ideology and the social experience of leishmaniasis: A contextualized analysis of bioarchaeological and paleopathological data from San Pedro de Atacama, Chile. International Journal of Paleopathology 1, 24-34.

Martinson, E., Reinhard, K.J., Buikstra, J.E., Dittmar de la Cruz, K., 2003. Pathoecology of Chiribaya parasitism. Mem Inst Oswaldo Cruz 98 Suppl 1, 195-205.

Martinson, J.J., Chapman, N.H., Rees, D.C., Liu, Y.T., Clegg, J.B., 1997. Global distribution of the CCR5 gene 32-basepair deletion. Nature genetics 16, 100-103.

Masur, L.J., 2008. Peanuts and prestige on the Peruvian North Coast: the archaeology of peanuts at Huaca Gallinazo (V-59) and Huaca Santa Clara (V-67), Anthropology. University of British Columbia.

Matos, V., Santos, A.L., 2006. On the trail of pulmonary tuberculosis based on rib lesions: results from the Human Identified Skeletal Collection from the Museu Bocage (Lisbon, Portugal). Am J Phys Anthropol 130, 190-200.

Mattle, D., Sitsel, O., Autzen, H.E., Meloni, G., Gourdon, P., Nissen, P., 2013. On allosteric modulation of P-type Cu(+)-ATPases. J Mol Biol 425, 2299-2308.

Mays, S., Fysh, E., Taylor, G.M., 2002. Investigation of the link between visceral surface rib lesions and tuberculosis in a medieval skeletal series from England using ancient DNA. American Journal of Physical Anthropology 119, 27-36.

Mays, S., Vincent, S., Meadows, J., 2012. A possible case of treponemal disease from England dating to the 11th–12th century AD. International Journal of Osteoarchaeology 22, 366-372.

McMahon-Pratt, D., Traub-Cseko, Y., Lohman, K.L., Rogers, D.D., Beverley, S.M., 1992. Loss of the GP46/M-2 surface membrane glycoprotein gene family in the Leishmania braziliensis complex. Mol Biochem Parasitol 50, 151-160.

McNeill, W., 1976. Plagues and Peoples. Knopf Doubleday Publishing Group.

McNicol, L., Doetsch, R., 1983. A hypothesis accounting for the origin of pandemic cholera: a retrograde analysis. Perspectives in biology and medicine 26, 547.

Merbs, C.F., 1992. A new world of infectious disease. Yearbook of Physical Anthropology 35, 3-42.

Meyer, M., Kircher, M., 2010. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc 2010, pdb prot5448.

Micklinghoff, J.C., Breitinger, K.J., Schmidt, M., Geffers, R., Eikmanns, B.J., Bange, F.C., 2009. Role of the transcriptional regulator RamB (Rv0465c) in the control of the glyoxylate cycle in Mycobacterium tuberculosis. J Bacteriol 191, 7260-7269.

Milner, G.R., Smith, V.G., 1990. Oneota human skeletal remains. Archaeological Investigations at the Morton Village and Norris Farms 36, 111-148.

Minkes, W., 2005. Wrap the Dead: The Funerary Textile Tradition from the Osmore Valley, South Peru, and its Social-Political Implications, Archaeological Studies. Leiden University.

Mitchell, D., Willerslev, E., Hansen, A., 2005. Damage and repair of ancient DNA. Mutat Res 571, 265-276.

Molak, M., Lorenzen, E.D., Shapiro, B., Ho, S.Y., 2013. Phylogenetic estimation of timescales using ancient DNA: The effects of temporal sampling scheme and uncertainty in sample ages. Molecular biology and evolution 30, 253-262.

Momen, H., Cupolillo, E., 2000. Speculations on the origin and evolution of the genus Leishmania. Memórias do Instituto Oswaldo Cruz 95, 583-588.

Monot, M., Honore, N., Garnier, T., Araoz, R., Coppee, J.Y., Lacroix, C., Sow, S., Spencer, J.S., Truman, R.W., Williams, D.L., Gelber, R., Virmond, M., Flageul, B., Cho, S.N., Ji, B., Paniz-Mondolfi, A., Convit, J., Young, S., Fine, P.E., Rasolofo, V., Brennan, P.J., Cole, S.T., 2005. On the origin of leprosy. Science 308, 1040-1042.

Monot, M., Honore, N., Garnier, T., Zidane, N., Sherafi, D., Paniz-Mondolfi, A., Matsuoka, M., Taylor, G.M., Donoghue, H.D., Bouwman, A., Mays, S., Watson, C., Lockwood, D., Khamesipour, A., Dowlati, Y., Jianping, S., Rea, T.H., Vera-Cabrera, L., Stefani, M.M., Banu, S., Macdonald, M., Sapkota, B.R., Spencer, J.S., Thomas, J., Harshman, K., Singh, P., Busso, P., Gattiker, A., Rougemont, J., Brennan, P.J., Cole, S.T., 2009. Comparative genomic and phylogeographic analysis of Mycobacterium leprae. Nature genetics 41, 1282-1289.

Montiel, R., Garcia, C., Canadas, M.P., Isidro, A., Guijo, J.M., Malgosa, A., 2003. DNA sequences of Mycobacterium leprae recovered from ancient bones. FEMS Microbiol Lett 226, 413-414.

Montiel, R., Solorzano, E., Diaz, N., Alvarez-Sandoval, B.A., Gonzalez-Ruiz, M., Canadas, M.P., Simoes, N., Isidro, A., Malgosa, A., 2012. Neonate human remains: a window of opportunity to the molecular study of ancient syphilis. PLoS ONE 7, e36371.

Moodley, Y., Linz, B., Bond, R.P., Nieuwoudt, M., Soodyall, H., Schlebusch, C.M., Bernhoft, S., Hale, J., Suerbaum, S., Mugisha, L., van der Merwe, S.W., Achtman, M., 2012. Age of the Association between Helicobacter pylori and Man. PLoS Pathogens 8.

Morelli, G., Song, Y.J., Mazzoni, C.J., Eppinger, M., Roumagnac, P., Wagner, D.M., Feldkamp, M., Kusecek, B., Vogler, A.J., Li, Y.J., Cui, Y.J., Thomson, N.R., Jombart, T., Leblois, R., Lichtner, P., Rahalison, L., Petersen, J.M., Balloux, F., Keim, P., Wirth, T., Ravel, J., Yang, R.F., Carniel, E., Achtman, M., 2010. Yersinia pestis genome sequencing identifies patterns of global phylogenetic diversity. Nature genetics 42, 1140-+.

Morse, D., 1961. Prehistoric tuberculosis in America. Am Rev Respir Dis 83, 489.

Mueller, R.L., Macey, J.R., Jaekel, M., Wake, D.B., Boore, J.L., 2004. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders inferred from complete mitochondrial genomes. Proc Natl Acad Sci U S A 101, 13820-13825.

Muller, N., Welle, M., Lobsiger, L., Stoffel, M.H., Boghenbor, K.K., Hilbe, M., Gottstein, B., Frey, C.F., Geyer, C., von Bomhard, W., 2009. Occurrence of Leishmania sp. in cutaneous lesions of horses in Central Europe. Vet Parasitol 166, 346-351.

Müller, R., Roberts, C.A., Brown, T.A., 2013. Biomolecular identification of ancient Mycobacterium tuberculosis complex DNA in human remains from Britain and continental Europe. American Journal of Physical Anthropology.

Mulligan, C.J., Norris, S.J., Lukehart, S.A., 2008. Molecular studies in Treponema pallidum evolution: toward clarity? PLoS neglected tropical diseases 2, e184.

Muñoz, A.S., 2008. Taphonomic properties of pinniped bone assemblages from the Atlantic coast of Isla Grande de Tierra del Fuego, Argentina. GEOBIOS 41, 113-122.

Mutreja, A., Kim, D.W., Thomson, N.R., Connor, T.R., Lee, J.H., Kariuki, S., Croucher, N.J., Choi, S.Y., Harris, S.R., Lebens, M., Niyogi, S.K., Kim, E.J., Ramamurthy, T., Chun, J., Wood, J.L., Clemens, J.D., Czerkinsky, C., Nair, G.B., Holmgren, J., Parkhill, J., Dougan, G., 2011. Evidence for several waves of global transmission in the seventh cholera pandemic. Nature 477, 462-465.

Nei, M., Kumar, S., 2000. Molecular evolution and phylogenetics. Oxford University Press, Oxford ; New York.

Noyes, H., Pratlong, F., Chance, M., Ellis, J., Lanotte, G., Dedet, J.P., 2002. A previously unclassified trypanosomatid responsible for human cutaneous lesions in Martinique (French West Indies) is the most divergent member of the genus Leishmania ss. Parasitology 124, 17-24.

Noyes, H.A., 1998. Can Trypanosoma Trees be Trusted? Parasitology 14, 49-50.

Noyes, H.A., Morrison, D.A., Chance, M.L., Ellis, J.T., 2000. Evidence for a neotropical origin of Leishmania. Memórias do Instituto Oswaldo Cruz 95, 575-578.

Nystrom, K.C., Malcom, C.M., 2010. Sex-Specific Phenotypic Variability and Social Organization in the Chiribaya of Southern Peru. Latin American Antiquity 21, 375-397.

O'Bleness, M., Searles, V.B., Varki, A., Gagneux, P., Sikela, J.M., 2012. Evolution of genetic and genomic features unique to the human lineage. Nature Reviews Genetics 13, 853-866.

O'Neill, E.M., Schwartz, R., Bullock, C.T., Williams, J.S., Shaffer, H.B., Aguilar-Miguel, X., Parra-Olea, G., Weisrock, D.W., 2013. Parallel tagged amplicon sequencing reveals major lineages and phylogenetic structure in the North American tiger salamander (Ambystoma tigrinum) species complex. Mol Ecol 22, 111-129.

Olalde, I., Allentoft, M.E., Sanchez-Quinto, F., Santpere, G., Chiang, C.W., Degiorgio, M., Prado-Martinez, J., Rodriguez, J.A., Rasmussen, S., Quilez, J., Ramirez, O., Marigorta, U.M., Fernandez-Callejo, M., Prada, M.E., Encinas, J.M., Nielsen, R., Netea, M.G., Novembre, J., Sturm, R.A., Sabeti, P., Marques-Bonet, T., Navarro, A., Willerslev, E., Lalueza-Fox, C., 2014. Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European. Nature.

Orquera, L.A., Legoupil, D., Piana, E.L., 2011. Littoral adaptation at the southern end of South America. Quaternary International 239, 61-69.

Ortner, D., 1998. Male-female immune reactivity and its implications for interpreting evidence in human skeletal paleopathology, in: Grauer, A.L., Stuart-Macadam, P. (Eds.), Sex and Gender in Paleopathological Perspective. Cambridge University Press, Cambridge, pp. 79-92.

Ortner, D., 2003. Identification of pathological conditions in human skeletal remains. Academic Press, New York.

Owen, B.D., 1993. A model of multiethnicity: state collapse, competition, and social complexity from Tiwanaku to Chiribaya in the Osmore Valley, Peru. University of California Los Angeles.

Owen, B.D., 2002. Marine carbon reservoir age estimates for the far south coast of Peru. Radiocarbon 44, 701-708.

Owen, B.D., 2005. Distant colonies and explosive collapse: The two stages of the Tiwanaku diaspora in the Osmore drainage. Latin American Antiquity, 45-80.

Patrucco, R., Tello, R., Bonavia, D., 1983. Parasitological studies of coprolites of pre-Hispanic Peruvian populations. Current Anthropology, 393-394.

Paulsen, A.C., 1977. Patterns of maritime trade between south coastal Ecuador and western Mesoamerica, in: Benson, E.P. (Ed.), The Sea in the Pre-Columbian World. Dumbarton Oaks, Washington, D.C.

Peacock, C.S., Seeger, K., Harris, D., Murphy, L., Ruiz, J.C., Quail, M.A., Peters, N., Adlem, E., Tivey, A., Aslett, M., Kerhornou, A., Ivens, A., Fraser, A., Rajandream, M.A., Carver, T., Norbertczak, H., Chillingworth, T., Hance, Z., Jagels, K., Moule, S., Ormond, D., Rutter, S., Squares, R., Whitehead, S., Rabbinowitsch, E., Arrowsmith, C., White, B., Thurston, S., Bringaud, F., Baldauf, S.L., Faulconbridge, A., Jeffares, D., Depledge, D.P., Oyola, S.O., Hilley, J.D., Brito, L.O., Tosi, L.R., Barrell, B., Cruz, A.K., Mottram, J.C., Smith, D.F., Berriman, M., 2007. Comparative genomic analysis of three Leishmania species that cause diverse human disease. Nature genetics 39, 839-847.

Pearce-Duvet, J.M., 2006. The origin of human pathogens: evaluating the role of agriculture and domestic animals in the evolution of human disease. Biol Rev Camb Philos Soc 81, 369-382.

Pepperell, C.S., Casto, A.M., Kitchen, A., Granka, J.M., Cornejo, O.E., Holmes, E.C., Birren, B., Galagan, J., Feldman, M.W., 2013. The role of selection in shaping diversity of natural M. tuberculosis populations. PLoS Pathogens 9, e1003543.

Perkins, S.L., Martinsen, E.S., Falk, B.G., 2011. Do molecules matter more than morphology? Promises and pitfalls in parasites. Parasitology 138, 1664-1674.

Peters, J., 1885. Early history of Asiatic cholera, in India as known to Europeans AD 1503-1800. William Wood, New York.

Pfeiffer, S., 1984. Paleopathology in an Iroquoian ossuary, with special reference to tuberculosis. American Journal of Physical Anthropology 65, 181-189.

Philippe, H., Brinkmann, H., Lavrov, D.V., Littlewood, D.T., Manuel, M., Worheide, G., Baurain, D., 2011. Resolving difficult phylogenetic questions: why more sequences are not enough. PLoS Biol 9, e1000602.

Pinhasi, R., Turner, K., 2008. Epidemiological approaches in palaeopathology, in: Pinhasi, R., Mays, S. (Eds.), Advances in Human Palaeopathology. Wiley Blackwell, pp. 45-56.

Pitblado, B.L., 2011. A tale of two migrations: reconciling recent biological and archaeological evidence for the Pleistocene peopling of the Americas. Journal of Archaeological Research 19, 327-375.

Poe, S., Swofford, D.L., 1999. Taxon sampling revisited. Nature 398, 299-300.

Poinar, G., Jr., 2008. Lutzomyia adiketis sp. n. (Diptera: Phlebotomidae), a vector of Paleoleishmania neotropicum sp. n. (Kinetoplastida: Trypanosomatidae) in Dominican amber. Parasit Vectors 1, 22.

Poinar Jr, G., 2007. Early Cretaceous trypanosomatids associated with fossil sand fly larvae in Burmese amber. Memórias do Instituto Oswaldo Cruz 102, 635-637.

Pond, S.L., Frost, S.D., 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21, 2531-2533.

Ponstingl, H., 2011. SMALT v0.5.8. Retrieved 5/1, 2012, from http://www.sanger.ac.uk/resources/software/smalt/.

Porcheron, G., Garenaux, A., Proulx, J., Sabri, M., Dozois, C.M., 2013. Iron, copper, zinc, and manganese transport and regulation in pathogenic Enterobacteria: correlations between strains, site of infection and the relative importance of the different metal transport systems for virulence. Front Cell Infect Microbiol 3, 90.

Porter, J.D., 1996. Mycobacteriosis and HIV infection: the new public health challenge. J Antimicrob Chemother 37 Suppl B, 113-120.

Poulson, S.R., Kuzminsky, S.C., Scott, G.R., Standen, V.G., Arriaza, B., Muñoz, I., Dorio, L., 2013. Paleodiet in northern Chile through the Holocene: extremely heavy δ15N values in dental calculus suggest a guano-derived signature? Journal of Archaeological Science.

Poux, C., Chevret, P., Huchon, D., de Jong, W.W., Douzery, E.J., 2006. Arrival and diversification of caviomorph rodents and platyrrhine primates in South America. Syst Biol 55, 228-244.

Preus, H.R., Marvik, O.J., Selvig, K.A., Bennike, P., 2011. Ancient bacterial DNA (aDNA) in dental calculus from archaeological human remains. Journal of Archaeological Science 38, 1827-1831.

Radolf, J.D., Lukehart, S.A., 2006. Pathogenic Treponema: Molecular and Cellular Biology. Caister Academic.

Raff, J., Cook, D.C., Kaestle, F., 2006. Tuberculosis in the New World: a study of ribs from the Schild Mississippian population, West-Central Illinois. Mem Inst Oswaldo Cruz 101 Suppl 2, 25-27.

Rafi, A., Spigelman, M., Stanford, J., Lemma, E., Donoghue, H., Zias, J., 1994. Mycobacterium leprae DNA from ancient bone detected by PCR. Lancet 343, 1360-1361.

Rambaut, A., 2000. Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. Bioinformatics 16, 395-399.

Rambaut, A., Drummond, A.J., 2007. Tracer, version 1.4. . Computer program and documentation distributed by the author, http://beast.bio.ed.ac.uk/Tracer.

Raoult, D., Aboudharam, G., Crubezy, E., Larrouy, G., Ludes, B., Drancourt, M., 2000. Molecular identification by "suicide PCR" of Yersinia pestis as the agent of Medieval Black Death. Proceedings of the National Academy of Sciences USA 97, 12800-12803.

Raymond, F., Boisvert, S., Roy, G., Ritt, J.-F., Légaré, D., Isnard, A., Stanke, M., Olivier, M., Tremblay, M.J., Papadopoulou, B., Ouellette, M., Corbeil, J., 2011. Genome sequencing of the lizard parasite Leishmania tarentolae reveals loss of genes associated to the intracellular stage of human pathogenic species. Nucleic Acids Research.

Read, A.F., 1994. The evolution of virulence. Trends Microbiol 2, 73-76.

Real, F., Vidal, R.O., Carazzolle, M.F., Mondego, J.M., Costa, G.G., Herai, R.H., Wurtele, M., de Carvalho, L.M., Carmona e Ferreira, R., Mortara, R.A., Barbieri, C.L., Mieczkowski, P., da Silveira, J.F., Briones, M.R., Pereira, G.A., Bahia, D., 2013. The genome sequence of Leishmania (Leishmania) amazonensis: functional annotation and extended analysis of gene models. DNA Res 20, 567-581.

Reid, A.H., Fanning, T.G., Hultin, J.V., Taubenberger, J.K., 1999. Origin and evolution of the 1918 "Spanish" influenza virus hemagglutinin gene. Proceedings of the National Academy of Sciences USA 96, 1651-1656.

Reid, A.H., Fanning, T.G., Janczewski, T.A., Taubenberger, J.K., 2000. Characterization of the 1918 "Spanish" influenza virus neuraminidase gene. Proceedings of the National Academy of Sciences USA 97, 6785-6790.

Reid, A.H., Fanning, T.G., Slemons, R.D., Janczewski, T.A., Dean, J., Taubenberger, J.K., 2003a. Relationship of pre-1918 avian influenza HA and NP sequences to subsequent avian influenza strains. Avian Dis 47, 921-925.

Reid, A.H., Janczewski, T.A., Lourens, R.M., Elliot, A.J., Daniels, R.S., Berry, C.L., Oxford, J.S., Taubenberger, J.K., 2003b. 1918 influenza pandemic caused by highly conserved viruses with two receptor-binding variants. Emerging Infectious Diseases 9, 1249-1253.

Reid, A.H., Taubenberger, J.K., 1999. The 1918 flu and other influenza pandemics: "over there" and back again. Laboratory Investigation 79, 95-101.

Reid, A.H., Taubenberger, J.K., 2003. The origin of the 1918 pandemic influenza virus: a continuing enigma. Journal of General Virology 84, 2285-2292.

Reinhard, K., Urban, O., 2003. Diagnosing ancient diphyllobothriasis from Chinchorro mummies. Mem Inst Oswaldo Cruz 98 Suppl 1, 191-193.

Reinhard, K.J., Buikstra, J., 2003. Louse infestation of the Chiribaya culture, southern Peru: variation in prevalence by age and sex. Memórias do Instituto Oswaldo Cruz 98 Suppl 1, 173-179.

Reuss, S.M., Dunbar, M.D., Calderwood Mays, M.B., Owen, J.L., Mallicote, M.F., Archer, L.L., Wellehan, J.F., Jr., 2012. Autochthonous Leishmania siamensis in horse, Florida, USA. Emerg Infect Dis 18, 1545-1547.

Rich, A.R., 1944. The Pathogenesis of Tuberculosis. Charles C. Thomas., Springfield, IL.

Roberts, C.A., Buikstra, J.E., 2003. The Bioarchaeology of Tuberculosis: A Gobal View on a Reemerging Disease. University Press of Florida, Gainesville, FL.

Roberts, C.A., Manchester, K., 2005. The Archaeology of Disease, 3rd edition ed. Cornell University PRess, Ithaca.

Rodo, X., Pascual, M., Fuchs, G., Faruque, A.S., 2002. ENSO and cholera: a nonstationary link related to climate change? Proceedings of the National Academy of Sciences USA 99, 12901-12906.

Rofes, J., 2000. Sacrificio de cuyes en El Yaral, comunidad prehispánica del extremo sur Peruano. Bulletin Institue Francais d'Êtudes Andines 29, 1-12.

Rohland, N., Hofreiter, M., 2007a. Ancient DNA extraction from bones and teeth. Nat Protoc 2, 1756-1762.

Rohland, N., Hofreiter, M., 2007b. Comparison and optimization of ancient DNA extraction. BioTechniques 42, 343-352.

Rose, K., Curtis, J., Baldwin, T., Mathis, A., Kumar, B., Sakthianandeswaren, A., Spurck, T., Low Choy, J., Handman, E., 2004. Cutaneous leishmaniasis in red kangaroos: isolation and characterisation of the causative organisms, International journal for parasitology, 2004/04/28 ed, pp. 655-664.

Rothschild, B.M., Hershkovitz, I., Rothschild, C., 1995. Origin of yaws in the Pleistocene. Nature 378, 343-344.

Rull, V., 2008. Speciation timing and neotropical biodiversity: the Tertiary-Quaternary debate in the light of molecular phylogenetic evidence. Mol Ecol 17, 2722-2729.

Saf'janova, V., 1982. The problem of taxonomy with Leishmania. Ser Protozool Sov Acad Sci Leningr 7, 5-109.

Salo, W.L., Aufderheide, A.C., Buikstra, J.E., Holcomb, T.A., 1994. Identification of Mycobacterium tuberculosis DNA in a Pre-Columbian Peruvian Mummy. PNAS 91, 2091-2094.

Sandness, K.L., 1992. Temporal and spatial dietary variability in the prehistoric lower and middle Osmore drainage: the carbon and nitrogen isotope evidence. University of Nebraska--Lincoln.

Santoro, C.M., Arriaza, B.T., Standen, V.G., Marquet, P.A., 2005. People of the Coastal Atacama Desert: Living Between Sand Dunes and Waves of the Pacific Ocean, in: Veth, P., Smith, M., Hiscock, P. (Eds.), Desert Peoples: Archaeological Perspectives. Blackwell Publishing, Malden, MA, pp. 243-260.

Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., Paabo, S., 2012. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. PLoS ONE 7, e34131.

Schiavini, A., 1993. Los lobos marinos como recurso para cazadores-recolectores marinos: el caso de Tierra del Fuego. Latin American Antiquity, 346-366.

Schönian, G., Cupolillo, E., Mauricio, I., 2013. Molecular Evolution and Phylogeny of Leishmania, in: Ponte-Sucre, A.e.a. (Ed.), Drug Resistance in Leishmania Parasites. Springer-Verlag, Wien.

Schönian, G., Kuhls, K., Mauricio, I.L., 2011. Molecular approaches for a better understanding of the epidemiology and population genetics of Leishmania. Parasitology 138, 405-425.

Schönian, G., Mauricio, I., Cupolillo, E., 2010. Is it time to revise the nomenclature of Leishmania? Trends Parasitol 26, 466-469.

Schuenemann, V.J., Bos, K., Dewitte, S., Schmedes, S., Jamieson, J., Mittnik, A., Forrest, S., Coombes, B.K., Wood, J.W., Earn, D.J., White, W., Krause, J., Poinar, H.N., 2011. Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of Yersinia pestis from victims of the Black Death. Proceedings of the National Academy of Sciences USA 108, E746-752.

Schuenemann, V.J., Singh, P., Mendum, T.A., Krause-Kyora, B., Jager, G., Bos, K.I., Herbig, A., Economou, C., Benjak, A., Busso, P., Nebel, A., Boldsen, J.L., Kjellstrom, A., Wu, H., Stewart, G.R., Taylor, G.M., Bauer, P., Lee, O.Y., Wu, H.H., Minnikin, D.E., Besra, G.S., Tucker, K.,

Roffey, S., Sow, S.O., Cole, S.T., Nieselt, K., Krause, J., 2013. Genome-wide comparison of medieval and modern Mycobacterium leprae. Science 341, 179-183.

Schwartz, R.S., Harkins, K., Stone, A.C., Cartwright, R.A., 2013. SISRS: SNP Identification from Short Read Sequences. arXiv preprint arXiv:1305.3665.

Schwarz, S., Skytte, L., Rasmussen, K.L., 2013. Pre-Columbian treponemal infection in Denmark?-a paleopathological and archaeometric approach. Heritage Science 1, 19.

Seeff, L.B., Miller, R.N., Rabkin, C.S., Buskell-Bales, Z., Straley-Eason, K.D., Smoak, B.L., Johnson, L.D., Lee, S.R., Kaplan, E.L., 2000. 45-year follow-up of hepatitis C virus infection in healthy young adults. Ann Intern Med 132, 105-111.

Shapiro, B., Gilbert, M.P.T., 2006. No proof that typhoid fever caused the Plague of Athens (a reply to Papagrigorakis et al.). Int J Infect Dis 10, 334-340.

Shaw, J., 1997. Ecological and evolutionary pressures on leishmanial parasites. Brazilian Journal of Genetics 20.

Shaw, J.J., 1994. Taxonomy of the genus Leishmania: present and future trends and their implications. Memórias do Instituto Oswaldo Cruz 89, 471-478.

Shaw, J.J., 2007. The leishmaniases- survival and expansion in a changing world. A mini-review. Memórias do Instituto Oswaldo Cruz 102, 541-547.

Simmonds, P., 2001. The origin and evolution of hepatitis viruses in humans. Journal of General Virology 82, 693-712.

Simmonds, P., 2004. Genetic diversity and evolution of hepatitis C virus--15 years on. Journal of General Virology 85, 3173-3188.

Smith, D.B., Pathirana, S., Davidson, F., Lawlor, E., Power, J., Yap, P.L., Simmonds, P., 1997. The origin of hepatitis C virus genotypes. Journal of General Virology 78 (Pt 2), 321-328.

Somerville, W., Thibert, L., Schwartzman, K., Behr, M.A., 2005. Extraction of Mycobacterium tuberculosis DNA: a question of containment. J Clin Microbiol 43, 2996-2997.

Sprent, J.F.A., 1962. Parasitism, immunity and evolution, in: Leeper, G.S. (Ed.), The evolution of living organisms. Melbourne University Press, Melbourne, pp. 149-165.

Springer, M.S., Murphy, W.J., Eizirik, E., O'Brien, S.J., 2003. Placental mammal diversification and the Cretaceous-Tertiary boundary. Proc Natl Acad Sci U S A 100, 1056-1061.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688-2690.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312-1313.

Stead, W., Aufderheide, A., Buikstra, J., Holcomb, T., 1995. When did Mycobacterium tuberculosis infection first occur in the New World. American Journal of Respiratory and Critical Care Medicine 151, 1267-1268.

Stead, W.W., 1997. The origin and erratic global spread of tuberculosis: how the past explains the present and is the key to the future. Clin Chest Med 18, 65-77.

Steckel, R.H., Rose, J.C., 2002. The Backbone of History: Health and Nutrition in the Western Hemisphere. Cambridge University Press.

Stevens, J., Rambaut, A., 2001. Evolutionary rate differences in trypanosomes. Infect Genet Evol 1, 143-150.

Stevens, J.R., Noyes, H.A., Schofield, C.J., Gibson, W., 2001. The molecular evolution of Trypanosomatidae. Advances in parasitology 48, 1-56.

Stodder, A.L.W., 2011. Data and Data Analysis Issues in Paleopathology, in: Grauer, A.L. (Ed.), A Companion to Paleopathology. John Wiley & Sons, Malden, MA, pp. 339-256.

Stone, A.C., Wilbur, A.K., Buikstra, J.E., Roberts, C.A., 2009. Tuberculosis and leprosy in perspective. American Journal of Physical Anthropology 140 Suppl 49, 66-94.

Stuiver, M., Braziunas, T.F., 1993. Modeling atmospheric 14C influences and 14C ages of marine samples to 10,000 BC. Radiocarbon 36, 137-189.

Stuiver, M., Pearson, G.W., Braziunas, T.F., 1986. Radiocarbon age calibration of marine samples back to 9000 cal yr BP. Radiocarbon 28, 980-1021.

Stuiver, M., Reimer, P.J., 1993. Extended 14C database and revised CALIB radiocarbon calibration program. Radiocarbon 35, 215-230.

Sutter, R., 2009. Post-Tiwanaku Ethnogenesis in the Coastal Osmore Valley: A Test of Competing Models for the Origins of the Moquegua Chiribaya Using Matrix Correlations, in: Knudson, K.J., Stojanowski, C.M. (Eds.), Bioarchaeology and Identity in the Americas, University Press of Florida.

Sutter, R.C., 2000. Prehistoric Genetic and Culture Change: A Bioarcheological Search for Pre-Inka Altiplano Colonies in the Coastal Valleys of Moquegua, Peru, and Azapa, Chile. Latin American Antiquity 11, 43-70.

Swenson, E.R., 2006 Competitive Feasting, Religious Pluralism, and Decentralized Power in the Late Moche Period, in: Isbell, W.H., Silverman, H. (Eds.), Andean Archaeology III: North and South. Springer/Plenum Press, New York, pp. 112-142.

Szpak, P., Millaire, J.-F., White, C.D., Longstaffe, F.J., 2012. Influence of Seabird guano and Camelid dung fertilization on the nitrogen isotopic composition of field-grown maize (< i> Zea mays</i>). Journal of Archaeological Science.

Tamura, K., Battistuzzi, F.U., Billing-Ross, P., Murillo, O., Filipski, A., Kumar, S., 2012. Estimating divergence times in large molecular phylogenies. Proceedings of the National Academy of Sciences USA 109, 19333-19338.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular biology and evolution 28, 2731-2739.

Taubenberger, J.K., Baltimore, D., Doherty, P.C., Markel, H., Morens, D.M., Webster, R.G., Wilson, I.A., 2012. Reconstruction of the 1918 influenza virus: unexpected rewards from the past. MBio 3.

Taubenberger, J.K., Reid, A.H., Krafft, A.E., Bijwaard, K.E., Fanning, T.G., 1997. Initial genetic characterization of the 1918 "Spanish" influenza virus. Science 275, 1793-1796.

Taubenberger, J.K., Reid, A.H., Lourens, R.M., Wang, R., Jin, G., Fanning, T.G., 2005. Characterization of the 1918 influenza virus polymerase genes. Nature 437, 889-893.

Taylor, G.M., Tucker, K., Butler, R., Pike, A.W., Lewis, J., Roffey, S., Marter, P., Lee, O.Y., Wu, H.H., Minnikin, D.E., Besra, G.S., Singh, P., Cole, S.T., Stewart, G.R., 2013. Detection and strain typing of ancient Mycobacterium leprae from a medieval leprosy hospital. PLoS ONE 8, e62406.

Thornton, E., Defrance, S., Krigbaum, J., Williams, P., 2011. Isotopic evidence for Middle Horizon to 16th century camelid herding in the Osmore Valley, Peru. International Journal of Osteoarchaeology 21, 544-567.

Thornton, R., 1987. American Indian Holocaust and Survival: A Population History: Since 1492. University of Oklahoma Press.

Tomczak, P., 2003. Prehistoric diet and socioeconomic relationships within the Osmore Valley of southern Peru. Journal of Anthropological Archaeology 22, 262-278.

Tomczak, P.D., 2001. Prehistoric Socio-Economic Relations and Population Organization in the Smore Valley of Southern Peru. University of New Mexico, Albuquerque.

Tran, T.N., Aboudharam, G., Raoult, D., Drancourt, M., 2011. Beyond ancient microbial DNA: nonnucleotidic biomolecules for paleomicrobiology. BioTechniques 50, 370-380.

Tsolaki, A.G., Hirsh, A.E., DeRiemer, K., Enciso, J.A., Wong, M.Z., Hannan, M., Goguet de la Salmoniere, Y.O., Aman, K., Kato-Maeda, M., Small, P.M., 2004. Functional and evolutionary genomics of Mycobacterium tuberculosis: insights from genomic deletions in 100 strains. Proc Natl Acad Sci U S A 101, 4865-4870.

Tuon, F.F., Amato Neto, V., Sabbaga Amato, V., 2008. Leishmania : origin, evolution and future since the Precambrian. FEMS Immunology & Medical Microbiology 54, 158-166.
Van der Auwera, G., Fraga, J., Montalvo, A.M., Dujardin, J.-C., 2011. Leishmania taxonomy up for promotion? Trends Parasitol 27, 49-50.

Verano, J.W., Ubelaker, D.H., 1992. Disease and Demography in the Americas. Smithsonian Institution Press, Washington, D.C.
von Hunnius, T.E., Yang, D., Eng, B., Waye, J.S., Saunders, S.R., 2007. Digging deeper into the limits of ancient DNA research on syphilis. Journal of Archaeological Science 34, 2091-2100.

Wagner, D.M., Klunk, J., Harbeck, M., Devault, A., Waglechner, N., Sahl, J.W., Enk, J., Birdsell, D.N., Kuch, M., Lumibao, C., Poinar, D., Pearson, T., Fourment, M., Golding, B., Riehm, J.M., Earn, D.J.D., DeWitte, S., Rouillard, J.-M., Grupe, G., Wiechmann, I., Bliska, J.B., Keim, P.S., Scholz, H.C., Holmes, E.C., Poinar, H., 2014. Yersinia pestis and the Plague of Justinian 541-543 AD: a genomic analysis. Lancet Infect Dis.

Walker, T.M., Ip, C.L., Harrell, R.H., Evans, J.T., Kapatai, G., Dedicoat, M.J., Eyre, D.W., Wilson, D.J., Hawkey, P.M., Crook, D.W., 2013. Whole-genome sequencing to delineate< i> Mycobacterium tuberculosis</i> outbreaks: a retrospective observational study. Lancet Infect Dis 13, 137-146.
Wang, R., Taubenberger, J.K., 2013. Characterization of the non-coding regions of the 1918 influenza A H1N1 virus. J Virol.

Warinner, C., Rodrigues, J.F., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y., Tito, R.Y., Fiddyment, S., Speller, C., Hendy, J., Charlton, S., Luder, H.U., Salazar-Garcia, D.C., Eppler, E., Seiler, R., Hansen, L.H., Castruita, J.A., Barkow-Oesterreicher, S., Teoh, K.Y., Kelstrup, C.D., Olsen, J.V., Nanni, P., Kawai, T., Willerslev, E., von Mering, C., Lewis, C.M., Jr., Collins, M.J., Gilbert, M.T.,
Ruhli, F., Cappellini, E., 2014. Pathogens and host immunity in the ancient human oral cavity. Nat Genet 46, 336-344.

Webster, R.G., Bean, W.J., Gorman, O.T., Chambers, T.M., Kawaoka, Y., 1992. Evolution and ecology of influenza A viruses. Microbiol Rev 56, 152-179.

Weirather, J.L., Jeronimo, S.M., Gautam, S., Sundar, S., Kang, M., Kurtz, M.A., Haque, R., Schriefer, A., Talhari, S., Carvalho, E.M., Donelson, J.E., Wilson, M.E., 2011. Serial quantitative PCR assay for detection, species discrimination, and quantification of Leishmania spp. in human samples. J Clin Microbiol 49, 3892-3904.

Weiss, R.A., 2001. Animal origins of human infectious disease. Philosophical Transactions of the Royal Society B: Biological Sciences 356, 957-977.

Wertheim, J.O., Pond, S.L.K., 2011. Purifying Selection Can Obscure the Ancient Age of Viral Lineages. Molecular biology and evolution 28, 3355-3365.

Wertheim, J.O., Worobey, M., 2009. Dating the age of the SIV lineages that gave rise to HIV-1 and HIV-2. PLoS Comput Biol 5, e1000377.

Wheeler, J.C., Russel, A., Redden, H., 1995. Llamas and alpacas: pre-conquest breeds and post-conquest hybrids. Journal of Archaeological Science 22, 833-840.

Wiechmann, I., Harbeck, M., Grupe, G., 2010. Yersinia pestis DNA Sequences in Late Medieval Skeletal Finds, Bavaria. Emerging Infectious Diseases 16, 1806-1807.

Wilbur, A.K., Buikstra, J.E., 2006. Patterns of tuberculosis in the Americas: how can modern biomedicine inform the ancient past? Memórias do Instituto Oswaldo Cruz 101 Suppl 59-66.
Wilbur, a.K., Farnbach, a.W., Knudson, K.J., Buikstra, J.E., 2008. Diet, Tuberculosis, and the Paleopathological Record. Current Anthropology 49, 963-991.

Wirth, D.F., Pratt, D.M., 1982. Rapid identification of Leishmania species by specific hybridization of kinetoplast DNA in cutaneous lesions. Proc Natl Acad Sci U S A 79, 6999-7003.

Wirth, T., Hildebrand, F., Allix-Beguec, C., Wolbeling, F., Kubica, T., Kremer, K., van Soolingen, D., Rusch-Gerdes, S., Locht, C., Brisse, S., Meyer, A., Supply, P., Niemann, S., 2008. Origin, spread and demography of the Mycobacterium tuberculosis complex. PLoS Pathogens 4, e1000160.

Wolfe, N.D., Dunavan, C.P., Diamond, J., 2007. Origins of major human infectious diseases. Nature 447, 279-283.

Wolfe, N.D., Escalante, A.A., Karesh, W.B., Kilbourn, A., Spielman, A., Lal, A.A., 1998. Wild primate populations in emerging infectious disease research: the missing link? Emerging Infectious Diseases 4, 149-158.

Wood, J.W., Milner, G., Harpending, J., Weiss, K., 1992. The osteological paradox. Current Anthropology 33, 343-370.

Worobey, M., Gemmel, M., Teuwen, D.E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J.J., Kabongo, J.M., Kalengayi, R.M., Van Marck, E., Gilbert, M.T., Wolinsky, S.M., 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature 455, 661-664.

Worobey, M., Telfer, P., Souquiere, S., Hunter, M., Coleman, C.A., Metzger, M.J., Reed, P., Makuwa, M., Hearn, G., Honarvar, S., Roques, P., Apetrei, C., Kazanji, M., Marx, P.A., 2010. Island biogeography reveals the deep history of SIV. Science 329, 1487.

Wyss, A.R., Flynn, J.J., Norell, M.A., Swisher, C.C., Charrier, R., Novacek, M.J., McKenna, M.C., 1993. South America's earliest rodent and recognition of a new interval of mammalian evolution. Nature 365, 434-437.

Xiao, Y.L., Kash, J.C., Beres, S.B., Sheng, Z.M., Musser, J.M., Taubenberger, J.K., 2013. High-throughput RNA sequencing of a formalin-fixed, paraffin-embedded autopsy lung tissue sample from the 1918 influenza pandemic. Journal of Pathology 229, 535-545.

Yang, Z., 1996. Maximum-Likelihood Models for Combined Analyses of Multiple Sequence Data. J Mol Evol 42, 587-596.

Yang, Z., O'Brien, J.D., Zheng, X., Zhu, H.Q., She, Z.S., 2007. Tree and rate estimation by local evaluation of heterochronous nucleotide data. Bioinformatics 23, 169-176.

Yoder, J.B., Briskine, R., Mudge, J., Farmer, A., Paape, T., Steele, K., Weiblen, G.D., Bharti, A.K., Zhou, P., May, G.D., Young, N.D., Tiffin, P., 2013. Phylogenetic signal variation in the genomes of Medicago (Fabaceae). Syst Biol 62, 424-438.

Yoneda, M., Tanaka, A., Shibata, Y., Morita, M., Uzawa, K., Hirota, M., Uchida, M., 2002. Radiocarbon marine reservoir effect in human remains from the Kitakogane site, Hokkaido, Japan. Journal of Archaeological Science 29, 529-536.
Yurchenko, V.Y., Lukes, J., Jirku, M., Zeledon, R., Maslov, D.A., 2006. Leptomonas costaricensis sp. n. (Kinetoplastea: Trypanosomatidae), a member of the novel phylogenetic group of insect trypanosomatids closely related to the genus Leishmania. Parasitology 133, 537-546.

Zumrrága, M.J., Bernardelli, A., Bastida, R., Quse, V., Loureiro, J., Cataldi, A., Bigi, F., Alito, A., Castro Ramos, M., Samper, S., Otal, I., Martin, C., Romano, M.I., 1999. Molecular characterization of mycobacteria isolated from seals. Microbiology 145 (Pt 9), 2519-2526.

APPENDIX A

SUPPLEMENTAL DATA FOR “SCREENING ANCIENT TUBERCULOSIS WITH qPCR: CHALLENGES & OPPORTUNITIES”

Appendix A
Harkins KM, Buikstra JE, Campbell TJ, Bos KI, Johnson ED, Krause J, Stone AC. Screening Ancient Tuberculosis with qPCR: Challenges and Opportunities.
Table S1: Site information for samples used in this analysis including permission to access and references.

Site	N	Dates	Location	Contact and/or curational facility	Reference
OLD WORLD					
Arena Candida	1	3000-4000 BC	Italy	Silvia Smith	[1]
Arma dell Aquila	1	3850+-90 BC	Italy	Silvia Smith	[2]
St Mary Spital	16	AD 1200-1539	England	Museum of London	[3]
Ryukyu (Okinawa)	1	AD 1600-1800	Japan	Takao Suzuki	Takao Suzuki, pers. comm.
Russian Academy of Science	1	Early Medieval	Russia	Alexandra Buzhilova	Buzhilova, pers. comm.
NEW WORLD (SOUTH AMERICA)					
Moquegua (Chen Chen)	10	AD 500-700	Moquegua, Peru	Bertha Vargas, INC	121[4]
Las Delicias y Candeleria	1	AD 770-990	Colombia		[122, 123] [5, 6]
Lambayeque	2	AD 900-1750	Peru	Haagen Klaus	[7]
Moquegua (Estuquina)	14	AD 1000-1476	Moquegua, Peru	Jane Buikstra, INC	[8]
Yaral	2	AD 900 - 1200	Osmore Valley, Peru	Jane Buikstra	[9-11]
Chiribaya Alta	4	AD 772-1350	Ilo, Peru	Jane Buikstra	[9, 10, 12]
Algodonal	1	AD 1031-1179	Ilo, Peru	Bruce Owen	[4, 11]
Santa Cruz	1	AD 1200	Patagonia, Argentina	Solana Garcia Guraieb, Rafael Goñi	[13]
Punta Arenas, Sitio Myren-1	1	AD 1310+-20	Argentina	Ricardo Guichón	Guichón, pers. comm.
Purucucho	3	AD 1470-1540	Peru	Guillermo Cock Carrasco, INC	[14, 15]
Misión Salesiana	15	AD 1870-1930	Tierra del Fuego	Ricardo Guichon	[16]
NEW WORLD (NORTH AMERICA)					
Yokem	1	AD 1039-1401	Illinois, USA	Della Cook, U. Indiana	[17, 18]
Schild	1	AD 1040-1407	Illinois, USA	Della Cook, U. Indiana	[19, 20]
Norris Farms	5	AD 1100-1250	Illinois, USA	Illinois State Museum	[21]
Orendorf	2	AD 1150-1300	Illinois, USA	Ken Mowbray, AMNH	[22]
Uxbridge	19	AD 1410-1570	Ontario, Canada	Dr. Susan Pfeiffer, U. Toronto	[23]
Arikara, Mobridge	2	AD 1500-1750	South Dakota, USA	NMNH	[24]
Arikara, Sully School Village	7	AD 1550-1700	South Dakota, USA	NMNH	[24]
Arikara, Leavenworth	1	AD 1725-1775	South Dakota, USA	NMNH	[25]
Arikara, Leavitt	1	AD 1725-1775	South Dakota, USA	NMNH	[24]
Arikara, Cheyenne River Village	6	AD 1750-1775	South Dakota, USA	NMNH	[24]
Chirikov	2	AD 1750-1899	Alaska, USA	Della Cook, U. Indiana	[26]
San Cristobal	2	AD 1450-1680 (contact AD 1598)	New Mexico, USA	Ken Mowbray, AMNH	[27]
Highland Park	4	AD 1847-1850	New York, USA	Rochester Museum & Science Center	[28]

[1] Formicola, V., Milanesi, Q. & Scarsini, C. 1987 Evidence of spinal tuberculosis at the beginning of the fourth millennium BC from Arene Candide cave (Liguria, Italy). *American Journal of Physical Anthropology* **72**, 1-6.
[2] Canci, A., Minozzi, S. & Tarli, S.M.B. 1996 New evidence of tuberculous spondylitis from Neolithic Liguria (Italy). *International Journal of Osteoarchaeology* **6**, 497-501.
[3] Connell, B., Jones, A., Redfern, R. & Walker, D. 2012 A Bioarchaeological Study of Medieval Burials on the Site of St Mary Spital. *Excavations at Spitalfields Market, London E* **1**, 1991-2007.
[4] Owen, B.D. 2005 Distant colonies and explosive collapse: The two stages of the Tiwanaku diaspora in the Osmore drainage. *Latin American Antiquity*, 45-80.
[5] Cárdenas Arroyo, F. 1993 Paleodieta y paleodemografía en poblaciones arqueológicas muiscas (sitios Las Delicias y Candelaria). *Revista Colombiana de Antropología* **30**, 129-148.
[6] Enciso, B. 1991 Arqueología de rescate en el barrio Las Delicias (Bogotá). *Revista Colombiana de Antropología* **28**, 155-160.
[7] Klaus, H.D., Wilber, A.K., Temple, D.H., Buikstra, J.E., Stone, A.C., Fernandez, M., Wester, C. & Tam, M.E. 2010 Tuberculosis on the north coast of Peru: skeletal and molecular paleopathology of late pre-Hispanic and postcontact mycobacterial disease. *Journal of Archaeological Science* **37**, 2587-2597.
[8] Williams, S.R. 1990 The skeletal biology of Estuquiña: a late intermediate period site in southern Peru, Northwestern University.
[9] Buikstra, J.E., Tomczak, P.D., Lozada Cerna, M. & Rakita, G.F. 2005 Chiribaya Political Economy: A Bioarchaeological Perspective. In *Interacting with the dead: Perspectives on mortuary archaeology for the new millennium* (eds. G.F. Rakita, J.E. Buikstra, L.A. Beck & S.R. Williams), pp. 66-80. Gainesville, University Press of Florida.
[10] Lozada, M.C., Buikstra, J.E., Rakita, G. & Wheeler, J.C. 2009 Camelid herders: the forgotten specialists in the coastal senorio of Chiribaya, southern Peru. *Andean Civilization: A Tribute to Michael E. Moseley Marcus J, Williams PR (eds). Cotsen Institute of Archaeology: Los Angeles*, 351-364.
[11] Owen, B.D. 1993 A model of multiethnicity: state collapse, competition, and social complexity from Tiwanaku to Chiribaya in the Osmore Valley, Peru, University of California Los Angeles.
[12] Lozada, M.C. & Buikstra, J.E. 2005 Pescadores and labradores among the señorío of Chiribaya in southern Peru. *Us and Them: Archaeology and Ethnicity in the Andes, Cotsen Institute of Archaeology, University of California, Los Angeles*, 206-225.
[13] Goñi, R., Bosio, L. & García Guraieb, S. 2003-2005 Un caso de enfermedad infecciosa en cazadores-recolectores prehispánicos de Patagonia. *Cuadernos del Instituto Nacional de Antropología y Pensamiento Latinoamericano* **20**, 399-404.
[14] Cock, G.A. 2006 Proyecto de Recuperaciòn de Contextos Funerarios en el Cementerio 57AS03-Zona Arqueológica Purchuco-Huaquerones. (Report on file Instituto Nacional de Cultura, Lima.
[15] Cock, G. & Goycochea Dıaz, C. 2004 Puruchuco y el cementerio Inca de la quebrada de Huaquerones. In *Puruchuco y la sociedad de Lima: un homenaje a Arturo Jiménez Borja* (eds. L.F. Villacorta, L. Vetter Parodi & C. Ausejo Castilla), pp. 179-197. Lima, Concytec.
[16] Guichón, R.A., Suby, J.A., Casali, R. & Fugassa, M.H. 2006 Health at the time of Native-European contact in Southern Patagonia: First steps, results, and prospects. *Memórias do Instituto Oswaldo Cruz* **101 (Suppl. 2)**, 97-105.
[17] Crane, H. & Griffin, J.B. 1972 University of Michigan radiocarbon dates XV. *Radiocarbon* **14**, 195-222.
[18] Perino, G. 1971 The Yokem site, Pike County, Illinois. *Illinois Archaeological Survey Bulletin* **8**, 149-186.
[19] Perino, G.H. 1973 The Late Woodland component at the Schild sites, Greene County, Illinois. *Late Woodland Site Archaeology in Illinois* **1**, 90-137.
[20] Crane, H.R. & Griffin, J.B. 1966 University of Michigan Radiocarbon Dates XI. . *Radiocarbon* **8**, 256-285.
[21] Milner, G.R. & Smith, V.G. 1990 Oneota human skeletal remains. *Archaeological Investigations at the Morton Village and Norris Farms* **36**, 111-148.
[22] Hanson, A.J. 2000 Skeletal evidence of tuberculosis and treponematosis in a prehistoric population from west-central Illinois, Iowa State University.
[23] Pfeiffer, S. 1984 Paleopathology in an Iroquoian ossuary, with special reference to tuberculosis. *American Journal of Physical Anthropology* **65**, 181-189.
[24] Owsley, D.W. & Bruwelheide, K.L. 1997 Bioarchaeological research in northeastern Colorado, northern Kansas, Nebraska and South Dakota. In *Bioarcheology of the North Central United States* (eds. D.W. Owsley & J.C. Rose), pp. 7-56. Fayetteville, Arkansas Archaeological Survey Research Series.
[25] Krause, R.A. 1972 *The Leavenworth site: archaeology of an historic Arikara community*, University of Kansas.
[26] Cook, D.C. & Dougherty, S.P. 2001 Row, row, row your boat: activity patterns and skeletal robusticity in a series from Chirikof island, Alaska. *American Journal of Physical Anthropology* **53 (Suppl. 32)**, (abstract).
[27] Stodder, A.L.W. 1990 *Paleoepidemiology of eastern and western Pueblo communities in protohistoric New Mexico*, University of Colorado at Boulder.
[28] Lanphear, K.M. 1988 Health and mortality in a nineteenth century poorhouse skeletal sample, University at Albany, State University of New York, Department of Anthropology.

Appendix A
Harkins KM, Buikstra JE, Campbell TJ, Bos KI, Johnson ED, Krause J, Stone AC. Screening Ancient Tuberculosis with qPCR: Challenges and Opportunities.

Table S2 Summary of qPCR results by sample

World Region	Archaeological Site Name	Burial/ Indiv Info	Lab Sample Name	No. loci positive	Target Name	Total Percent Positive Wells	Is Positive Assay	Is Inhibited
Positive for Multiple Loci								
New World	Arikara, Mobridge (39WW1)	382961	AD15	4/4	rpoB1	100.0%	Yes	
					IS6110	100.0%	Yes	No
					IS1081	83.3%	Yes	
					rpoB2	50.0%	Yes	No
New World	Chiribaya Alta, Cemetery 1	2069	AD54	4/4	IS1081	100.0%	Yes	
					rpoB2	100.0%	Yes	
					IS6110	100.0%	Yes	
					rpoB1	100.0%	Yes	
New World	El Algodonol (AL)	386-1	AD58	4/4	IS6110	66.7%	Yes	No
					rpoB1	83.3%	Yes	
					IS1081	50.0%	Yes	Yes
					rpoB2	50.0%	Yes	No
New World	Las Delicias	LD-90-1X-11	AD281	3/4	IS6110	100.0%	Yes	
					IS1081	67.7%	Yes	
					rpoB1	0.0%	No	
					rpoB2	66.7%	Yes	
New World	Moquegua, M6: Estuquina	4165	AD82	4/4	rpoB2	75.0%	Yes	Yes
					rpoB1	96.2%	Yes	No
					IS6110	100.0%	Yes	
					IS1081	100.0%	Yes	
New World	Rochester, Highland Park	138a	AD127	2/4	rpoB2	100.0%	Yes	
					rpoB1	100.0%	Yes	
New World	Rochester, Highland Park	141a	AD128	4/4	rpoB2	50.0%	Inconclusive	
					IS1081	100.0%	Yes	
					IS6110	100.0%	Yes	
					rpoB1	100.0%	Yes	
New World	Yaral-2 (Cemetery 2)	10291	AD64	3/3	IS1081	66.7%	Yes	
					IS6110	100.0%	Yes	
					rpoB1	83.3%	Yes	
Positive for one loci (excluding rpoB1)								
New World	Norris Farms	228	AD160	1/4	rpoB2	66.7%	Yes	
Positive for only rpoB1								

New World	Punta Arenas, Sitio Myren-7	T.D.F. 30477	AD216	1/4	rpoB1	100.0%	Yes	
					rpoB2	37.5%	No	Yes
Old World	Okinawa Island, Japan		AD224	1/4	rpoB1	66.7%	Yes	
New World	Arikara, Cheyenne River village (39STI)	382713	AD11	1/4	rpoB1	100.0%	Yes	
New World	Arikara, Cheyenne River village (39STI)	382669	AD6	1/4	rpoB1	50.0%	Yes	
New World	Arikara, Leavenworth (39CO9)	325341	AD1	1/4	rpoB1	100.0%	Yes	
New World	Arikara, Sully School Village (39SL4)	388120	AD19	1/3	rpoB1	100.0%	Yes	Yes
New World	Misión Salesiana	D-E 14	AD256	1/4	rpoB1	75.0%	Yes	
New World	Misión Salesiana	E 15-16 (1)	AD257	1/4	rpoB1	64.7%	Yes	Maybe
New World	Misión Salesiana	C 15-16	AD263	1/4	rpoB1	100.0%	Yes	
New World	Misión Salesiana	D 14	AD264	1/4	rpoB1	100.0%	Yes	
New World	Misión Salesiana	C 15	AD265	1/3	rpoB1	100.0%	Yes	
New World	Moquegua, M6: Estuquina	2279b	AD76	1/4	rpoB1	100.0%	Yes	
New World	Moquegua, M6: Estuquina	3215	AD79	1/3	rpoB1	83.3%	Yes	
New World	Moquegua, M6: Estuquina	5859	AD87	1/3	rpoB1	50.0%	Yes	
New World	Moquegua, M6: Estuquina	99407a	AD89	1/4	rpoB1	83.3%	Yes	
New World	Norris Farms	21	AD145	1/4	rpoB1	60.0%	Yes	
New World	Lambayeque, Peru	ILL-22	AD227	1/4	rpoB1	85.0%	Yes	
New World	Rochester, Highland Park	225a	AD134	1/4	rpoB1	100.0%	Yes	
New World	Rochester, Highland Park	16a	AD114	1/4	rpoB1	100.0%	Yes	
Negative samples (inhibition detected)								
New World	San Cristobal	8634	AD165	0/4	rpoB1	44.4%	No	Yes
New World	Moquegua, M6: Estuquina	99407b	AD90	0/4	rpoB1	33.3%	No	Yes
New World	Schild	SA-81	AD93	0/2	rpoB1	40.0%	No	Yes
New World	Misión Salesiana	E 14-15 (1)	AD250	0/3	rpoB1	12.5%	No	Maybe
New World	Misión Salesiana	D 16 (Bis)	AD258	0/3	rpoB1	12.5%	No	Maybe
New World	Moquegua, M6: Estuquina	3644a	AD81	0/4	rpoB1	16.7%	No	Maybe
Negative samples (inhibition negative or not tested)								
Old World	Arena Candide, Italy		AD282	0/4	rpoB1	0.0%	No	
					IS6110	16.7%	No	
					rpoB2	0.0%	No	
					IS1081	0.0%	No	
New World	Arikara, Cheyenne River village (39STI)	382713	AD11	0/4	IS1081	0.0%	No	
					IS6110	33.3%	No	
					rpoB2	0.0%	No	
New World	Arikara, Cheyenne River village (39STI)	382669	AD6	0/4	IS1081	0.0%	No	
					rpoB2	0.0%	No	
					IS6110	0.0%	No	
New World	Arikara, Leavenworth (39CO9)	325341	AD1	0/4	IS1081	0.0%	No	
					rpoB2	0.0%	No	

					IS6110	0.0%	No	
New World	Arikara, Sully School Village (39SL4)	388120	AD19	0/3	IS1081	0.0%	No	
					IS6110	0.0%	No	
New World	Arikara, Sully School Village (39SL4)	381390	AD3	0/2	rpoB2	0.0%	No	
					IS1081	0.0%	No	
New World	Chiribaya Alta, Cemetery 1, CHA-1	196	AD51	0/3	IS6110	0.0%	No	
					rpoB1	0.0%	No	
					rpoB2	0.0%	No	
New World	Chiribaya Alta, Cemetery 1, CHA-1	316	AD52	0/4	IS1081	0.0%	No	
					rpoB2	0.0%	No	
					rpoB1	0.0%	No	
					IS6110	0.0%	No	
New World	Chirikof	62CF/BSc-18	AD105	0/4	IS1081	0.0%	No	
					rpoB1	40.0%	No	
					IS6110	0.0%	No	
					rpoB2	50.0%	Inconclusive	
New World	Chirikof	62CF/AE3-14	AD106	0/4	rpoB1	0.0%	No	
					IS6110	0.0%	No	
					rpoB2	0.0%	No	
					IS1081	0.0%	No	
New World	Gooden Mds	F85-48	AD112	0/3	rpoB2	0.0%	No	
					rpoB1	0.0%	No	
					IS1081	0.0%	No	
New World	Lambayeque, Peru	ILL-22	AD227	0/3	IS6110	0.0%	No	
					rpoB2	12.5%	No	
					IS1081	0.0%	No	
New World	Lambayeque, Peru	CSJ-21	AD228	0/4	IS1081	0.0%	No	
					IS6110	0.0%	No	
					rpoB1	0.0%	No	
New World	Misión Salesiana	E 15-16 (2 bis)	AD247	0/3	rpoB1	0.0%	No	
					IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Misión Salesiana	E 15-16 (3)	AD248	0/4	rpoB2	0.0%	No	
					rpoB1	21.4%	No	No
					IS1081	0.0%	No	
					IS6110	16.7%	No	
New World	Misión Salesiana	E 14-15 (2)	AD249	0/3	rpoB2	0.0%	No	
					rpoB1	18.2%	No	No
					IS6110	0.0%	No	
New World	Misión Salesiana	E 14-15 (1)	AD250	0/3	rpoB2	0.0%	No	

New World	Misión Salesiana	E 14-15 (1)	AD250	0/3	IS1081	10.0%	No	No
New World	Misión Salesiana	C 14 (1)	AD251	0/4	IS1081	0.0%	No	
					IS6110	0.0%	No	
					rpoB2	0.0%	No	
					rpoB1	21.4%	No	
New World	Misión Salesiana	C 7-8	AD252	0/4	IS6110	33.3%	No	
					rpoB1	9.1%	No	No
					rpoB2	0.0%	No	
					IS1081	33.3%	No	
New World	Misión Salesiana	C 13	AD253	0/2	IS6110	0.0%	No	
					rpoB1	0.0%	No	
New World	Misión Salesiana	E 12-13	AD255	0/2	rpoB1	12.5%	No	
					rpoB2	0.0%	No	
New World	Misión Salesiana	D-E 14	AD256	0/4	IS1081	0.0%	No	
					IS6110	0.0%	No	
					rpoB2	0.0%	No	
New World	Misión Salesiana	E 15-16 (1)	AD257	0/4	IS1081	0.0%	No	
					IS6110	0.0%	No	
					rpoB2	0.0%	No	
New World	Misión Salesiana	D 16 (Bis)	AD258	0/3	rpoB2	0.0%	No	
					IS1081	0.0%	No	
New World	Misión Salesiana	C 14 (2)	AD259	0/4	rpoB2	0.0%	No	
					rpoB1	0.0%	No	
					IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Misión Salesiana	D 15-16	AD262	0/3	IS1081	0.0%	No	
					rpoB1	0.0%	No	
					rpoB2	0.0%	No	
New World	Misión Salesiana	C 15-16	AD263	0/4	IS1081	0.0%	No	
					IS6110	0.0%	No	
					rpoB2	0.0%	No	
New World	Misión Salesiana	D 14	AD264	0/4	IS1081	0.0%	No	
					rpoB2	0.0%	No	
					IS6110	0.0%	No	
New World	Misión Salesiana	C 15	AD265	0/3	IS1081	0.0%	No	
					IS6110	0.0%	No	
New World	Moquegua, Chen Chen	M1-2705	AD191	0/3	rpoB1	0.0%	No	
					IS1081	0.0%	No	
					IS6110	0.0%	No	
					IS1081	0.0%	No	

New World	Moquegua, M6: Estuquina	1002	AD69	0/4	IS6110	0.0%	No	
					rpoB1	0.0%	No	
					rpoB2	0.0%	No	
New World	Moquegua, M6: Estuquina	2279b	AD76	0/4	rpoB2	0.0%	No	
					IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Moquegua, M6: Estuquina	3215	AD79	0/3	rpoB2	0.0%	No	
					IS6110	0.0%	No	
New World	Moquegua, M6: Estuquina	3644a	AD81	0/4	rpoB2	0.0%	No	
					IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Moquegua, M6: Estuquina	4256	AD85	0/3	rpoB2	0.0%	No	
					rpoB1	16.7%	No	No
					IS6110	0.0%	No	
New World	Moquegua, M6: Estuquina	5859	AD87	0/3	IS6110	0.0%	No	
					rpoB2	0.0%	No	
New World	Moquegua, M6: Estuquina	99407a	AD89	0/4	rpoB2	0.0%	No	
					IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Moquegua, M6: Estuquina	99407b	AD90	0/4	rpoB2	0.0%	No	
					IS6110	16.7%	No	
					IS1081	0.0%	No	
New World	Norris Farms	21	AD145	0/4	IS1081	0.0%	No	
					rpoB2	0.0%	No	
					IS6110	0.0%	No	
New World	Norris Farms	38	AD146	0/4	rpoB2	0.0%	No	
					rpoB1	0.0%	No	
					IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Norris Farms	95	AD151	0/4	IS6110	0.0%	No	
					rpoB2	0.0%	No	
					rpoB1	0.0%	No	
					IS1081	0.0%	No	
New World	Norris Farms	160	AD156	0/3	rpoB1	0.0%	No	
					IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Norris Farms	224	AD159	0/3	rpoB1	0.0%	No	
					IS1081	0.0%	No	
					IS6110	0.0%	No	
New World	Norris Farms	228	AD160	0/3	rpoB1	0.0%	No	

New World	Norris Farms	228	AD100	0/3	IS1081	0.0%	No	
Old World	Okinawa Island, Japan		AD224	0/3	IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Orendorf	Burial 116	AD171	0/3	rpoB2	0.0%	No	
					rpoB1	0.0%	No	
					IS6110	0.0%	No	
New World	Orendorf	Burial 153	AD172	0/3	rpoB2	0.0%	No	
					IS6110	0.0%	No	
					rpoB1	0.0%	No	
New World	Punta Arenas, Sitio Myren-7	T.D.F. 30477	AD216	0/4	IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Puruchuco	Tomb 5, Indv. 2	AD209	0/3	rpoB2	0.0%	No	
					rpoB1	0.0%	No	
					IS6110	0.0%	No	
New World	Puruchuco	Tomb 49	AD211	0/3	IS6110	0.0%	No	
					rpoB1	0.0%	No	
					rpoB2	0.0%	No	
New World	Rochester, Highland Park	16a	AD114	0/4	rpoB2	14.3%	No	
New World					IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Rochester, Highland Park	138a	AD127	0/4	IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Rochester, Highland Park	225a	AD134	0/4	rpoB2	0.0%	No	
					IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	San Cristobal	8708	AD162	0/4	IS1081	0.0%	No	
					rpoB2	20.0%	No	No
					rpoB1	0.0%	No	
					IS6110	0.0%	No	
New World	San Cristobal	8634	AD165	0/4	rpoB2	0.0%	No	
					IS6110	0.0%	No	
					IS1081	0.0%	No	
New World	Santa Cruz, Patagonia (Argentina)	SAC4-1-1	AD225	0/3	IS1081	0.0%	No	
					IS6110	0.0%	No	
					rpoB1	0.0%	No	
New World	Schild	SB-262A	AD102	0/2	IS6110	0.0%	No	
					rpoB2	0.0%	No	
New World	Schild	SA-81	AD93	0/2	rpoB2	0.0%	No	
Old World	St. Mary Spital	2783	AD229	0/4	IS1081	0.0%	No	
					rpoB1	0.0%	No	

Old World	St. Mary Spital	3549	AD230	0/2	rpoB1	0.0%	No	
					IS1081	0.0%	No	
Old World	St. Mary Spital	5426	AD231	0/2	IS1081	0.0%	No	
					rpoB1	0.0%	No	
Old World	St. Mary Spital	8490	AD232	0/2	IS1081	0.0%	No	
					rpoB1	0.0%	No	
Old World	St. Mary Spital	8852	AD233	0/2	IS1081	0.0%	No	
					rpoB1	0.0%	No	
Old World	St. Mary Spital	10686	AD234	0/3	rpoB1	0.0%	No	
					IS1081	0.0%	No	
					IS6110	33.3%	No	
Old World	St. Mary Spital	12188	AD235	0/3	rpoB2	20.0%	No	No
					rpoB1	0.0%	No	
					IS1081	0.0%	No	
Old World	St. Mary Spital	13759	AD236	0/3	rpoB1	0.0%	No	
					rpoB2	20.0%	No	No
					IS1081	0.0%	No	
Old World	St. Mary Spital	17769	AD237	0/2	IS1081	0.0%	No	
					rpoB1	0.0%	No	
Old World	St. Mary Spital	19599	AD238	0/2	rpoB1	0.0%	No	
					IS1081	0.0%	No	
Old World	St. Mary Spital	22806	AD240	0/3	rpoB2	0.0%	No	
					rpoB1	0.0%	No	
					IS1081	16.7%	No	
Old World	St. Mary Spital	22829	AD241	0/2	rpoB1	0.0%	No	
					IS1081	0.0%	No	
Old World	St. Mary Spital	23630	AD242	0/3	rpoB1	33.3%	No	
					IS6110	0.0%	No	
					IS1081	0.0%	No	
Old World	St. Mary Spital	25714	AD243	0/2	rpoB1	0.0%	No	
					IS1081	0.0%	No	
Old World	St. Mary Spital	27372	AD244	0/2	rpoB1	0.0%	No	
					IS1081	0.0%	No	
Old World	St. Mary Spital	29490	AD245	0/2	rpoB1	0.0%	No	
					IS1081	0.0%	No	
Old World	St. Mary Spital	30702	AD246	0/2	rpoB1	0.0%	No	
					IS1081	0.0%	No	
New World	Uxbridge (Bb Gs-3)	P16b	AD28	0/3	IS6110	0.0%	No	
					rpoB2	0.0%	No	
					rpoB1	0.0%	No	

New World	Uxbridge (Bb Gs-3)	P47	AD31	0/3	rpoB2	0.0%	No	
					rpoB1	16.7%	No	No
					IS6110	0.0%	No	
New World	Uxbridge (Bb Gs-3)	P55	AD33	0/4	rpoB1	0.0%	No	
					IS1081	0.0%	No	
					IS6110	0.0%	No	
					rpoB2	0.0%	No	
New World	Uxbridge (Bb Gs-3)	P219	AD41	0/3	IS1081	0.0%	No	
					IS6110	0.0%	No	
					rpoB1	0.0%	No	
New World	Uxbridge (Bb Gs-3)	P280	AD44	0/3	rpoB2	0.0%	No	
					rpoB1	0.0%	No	
					IS6110	0.0%	No	
New World	Uxbridge (Bb Gs-3)	P1		0/2	rpoB1	0.0%	No	
					IS6110	0.0%	No	
New World	Uxbridge (Bb Gs-3)	P5		0/2	rpoB1	0.0%	No	
					IS6110	0.0%	No	
New World	Uxbridge (Bb Gs-3)	P6		0/2	rpoB1	0.0%	No	
					IS6110	0.0%	No	
New World	Uxbridge (Bb Gs-3)	P7		0/2	rpoB1	0.0%	No	
					IS6110	0.0%	No	
New World	Uxbridge (Bb Gs-3)	P45		0/2	rpoB1	0.0%	No	
					rpoB1	0.0%	No	
New World	Uxbridge (Bb Gs-3)	P57		0/2	rpoB1	0.0%	No	
					IS6110	0.0%	No	
New World	Yaral-1 (Cemetery 1)	59-1-E	AD60	0/3	rpoB2	0.0%	No	
					rpoB1	16.7%	No	
					IS6110	0.0%	No	
New World	Yaral-2 (Cemetery 2)	10335	AD65	0/3	IS6110	0.0%	No	
					rpoB1	0.0%	No	
					rpoB2	0.0%	No	
New World	Yokem	Md 2-5	AD109	0/3	IS6110	0.0%	No	
					rpoB1	0.0%	No	
					rpoB2	0.0%	No	

Appendix A

Harkins KM, Buikstra JE, Campbell TJ, Bos KI, Johnson ED, Krause J, Stone AC. Screening Ancient Tuberculosis with qPCR: Challenges and Opportunities.

Table S3: Summary of qPCR results of all qPCR assays (2008-2014) for each extract and number of wells recorded per target

World Region	Archaeological Region	Archaeological Site	Burial/ Indiv Info	Extract Name	Target Name	Total Percent Positive Wells	Is Positive Assay	Dilution 1:1 Positive Wells	Dilution 1:10 Positive Wells	Dilution 1:100 Positive Wells	Dilution 1:1000 Positive Wells	Extraction Protocol
New World	South Dakota	Arikara, Cheyenne River village	382669	AD6.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Cheyenne River village	382669	AD6.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Cheyenne River village	382669	AD6.1	rpoB	100.0%	Yes	3/3				modified Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Cheyenne River village	382669	AD6Tu	rpoB2	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Cheyenne River village	382713	AD11.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Cheyenne River village	382713	AD11.1	rpoB	100.0%	Yes	3/3	2/2			modified Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Cheyenne River village	382713	AD11.1	rpoB2	0.0%	No		0/2			modified Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Cheyenne River village	382713	AD11.1	IS6110	33.3%	No	1/3				modified Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Leavenworth	325341	AD1.2	IS6110	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Leavenworth	325341	AD1.2	IS1081	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Leavenworth	325341	AD1.3	rpoB	100.0%	Yes	3/3				Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Leavenworth	325341	AD1.3	IS6110	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Leavenworth	325341	AD1.2	rpoB	100.0%	Yes	3/3	2/2			Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Leavenworth	325341	AD1.2	rpoB2	0.0%	No		0/2			Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Leavenworth	325341	AD1.3	IS1081	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Leavenworth	325341	AD1.1	rpoB	100.0%	Yes	3/3	3/3			Yang [2]
New World	South Dakota	Arikara, Leavenworth	325341	AD1.1	rpoB2	0.0%	No		0/3			Yang [2]
New World	South Dakota	Arikara, Leavenworth	325341	AD1.1	IS6110	0.0%	No	0/3				Yang [2]
New World	South Dakota	Arikara, Leavenworth	325341	AD1.1	IS1081	0.0%	No	0/3				Yang [2]
New World	South Dakota	Arikara, Mobridge	382961	AD15.1	IS1081	83.3%	Yes	3/3	2/3			Dabney et al. [3]
New World	South Dakota	Arikara, Mobridge	382961	AD15.1	IS6110	100.0%	Yes	3/3	3/3			Dabney et al. [3]
New World	South Dakota	Arikara, Mobridge	382961	AD15.1	rpoB	100.0%	Yes	3/3	3/3			Dabney et al. [3]
New World	South Dakota	Arikara, Mobridge	382961	AD15.1	rpoB2	50.0%	Yes	3/3	0/3			Dabney et al. [3]
New World	South Dakota	Arikara, Sully School Village	381390	AD3Tu	IS1081	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Sully School Village	381390	AD3Tu	rpoB2	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Sully School Village	388120	AD19.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	South Dakota	Arikara, Sully School Village	388120	AD19.1	IS1081	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]
New World	Moquegua, Peru	Chen Chen	M1-2705	AD191.1	rpoB	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Moquegua, Peru	Chen Chen	M1-2705	AD191.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]

New World	Moquegua, Peru	Chen Chen	M1-2705	AD191.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	196	AD51.1	rpoB	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	196	AD51.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	196	AD51.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	316	AD52.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	316	AD52.1	rpoB2	0.0%	No		0/2			modified Rohland & Hofreiter [1]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	316	AD52.1	rpoB	0.0%	No	0/3	0/2			modified Rohland & Hofreiter [1]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	316	AD52.1	IS1081	0.0%	No	0/6				modified Rohland & Hofreiter [1]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	2069	AD54.1	rpoB	100.0%	Yes	3/3	2/2			modified Rohland & Hofreiter [1]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	2069	AD54.1	rpoB2	100.0%	Yes		2/2			modified Rohland & Hofreiter [1]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	2069	AD54.1	IS6110	100.0%	Yes	3/3				modified Rohland & Hofreiter [1]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	2069	AD54.1	IS1081	100.0%	Yes	3/3				modified Rohland & Hofreiter [1]
New World	Chirikof Island, Alaska	Chirikof	62CF/AE3-14	AD106.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Chirikof Island, Alaska	Chirikof	62CF/AE3-14	AD106.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Chirikof Island, Alaska	Chirikof	62CF/AE3-14	AD106.1	rpoB2	0.0%	No		0/2			modified Rohland & Hofreiter [1]
New World	Chirikof Island, Alaska	Chirikof	62CF/AE3-14	AD106.1	rpoB	0.0%	No	0/3	0/2			modified Rohland & Hofreiter [1]
New World	Chirikof Island, Alaska	Chirikof	62CF/BSc-18	AD105.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Chirikof Island, Alaska	Chirikof	62CF/BSc-18	AD105.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Chirikof Island, Alaska	Chirikof	62CF/BSc-18	AD105.1	rpoB2	50.0%	Yes		1/2			modified Rohland & Hofreiter [1]
New World	Chirikof Island, Alaska	Chirikof	62CF/BSc-18	AD105.1	rpoB	40.0%	No	0/3	2/2			modified Rohland & Hofreiter [1]
New World	Osmore Valley, Peru	El Algodonal	386-1	AD58.1	IS1081	50.0%	Yes	2/3	1/3			Dabney et al. [3]
New World	Osmore Valley, Peru	El Algodonol (AL)	386-1	AD58.1	IS6110	66.7%	Yes	3/3	1/3			Dabney et al. [3]
New World	Osmore Valley, Peru	El Algodonol (AL)	386-1	AD58.1	rpoB	83.3%	Yes	3/3	2/3			Dabney et al. [3]
New World	Osmore Valley, Peru	El Algodonol (AL)	386-1	AD58.1	rpoB2	50.0%	Yes	3/3	0/3			Dabney et al. [3]
New World	Illinois	Gooden Mds	F85-48	AD112Tu	rpoB2	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Illinois	Gooden Mds	F85-48	AD112Tu	rpoB	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Illinois	Gooden Mds	F85-48	AD112Tu	IS1081	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	138a	AD127.1	rpoB2	100.0%	Yes		2/2			modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	138a	AD127.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	138a	AD127.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	138a	AD127.1	rpoB	100.0%	Yes	3/3	2/2			modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	141a	AD128.1	rpoB	100.0%	Yes	3/3	2/2			modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	141a	AD128.1	IS6110	100.0%	Yes	3/3				modified Rohland & Hofreiter [1]

New World	Rochester, NY	Highland Park	141a	AD128.1	rpoB2	50.0%	Yes		1/2			modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	141a	AD128.1	IS1081	100.0%	Yes	3/3				modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	16a	AD114.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	16a	AD114.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	16a	AD114.1	rpoB2	0.0%	No		0/2			modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	16a	AD114.1	rpoB	100.0%	Yes	3/3	2/2			modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	16a	AD114Tu	rpoB2	20.0%	No	1/5				modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	225a	AD134.1	rpoB	100.0%	Yes	3/3	2/2			modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	225a	AD134.1	rpoB2	0.0%	No		0/2			modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	225a	AD134.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Rochester, NY	Highland Park	225a	AD134.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Lambayeque, Peru	Lambayeque, Peru	CSJ-21	AD228.1	rpoB	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Lambayeque, Peru	Lambayeque, Peru	CSJ-21	AD228.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.1	IS1081	0.0%	No	0/3				Yang [2]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.2	rpoB	100.0%	Yes	3/3	3/3			Yang [2]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.3B	IS1081	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.3	rpoB	100.0%	Yes	3/3	2/2			modified Rohland & Hofreiter [1]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.3	rpoB2	0.0%	No		0/2			modified Rohland & Hofreiter [1]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.3	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.3	IS1081	0.0%	No	0/6				modified Rohland & Hofreiter [1]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.3B	rpoB	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.1	IS6110	0.0%	No	0/3				Yang [2]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.2	rpoB2	33.3%	No		1/3			Yang [2]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.2	IS6110	0.0%	No	0/3				Yang [2]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.2	IS1081	0.0%	No	0/3				Yang [2]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.1	rpoB2	0.0%	No		0/3			Yang [2]
New World	Lambayeque, Peru	Lambayeque, Peru	ILL-22	AD227.1	rpoB	100.0%	Yes	3/3	3/3			Yang [2]
New World	Colombia	Las Delicias	LD-90-1 X-11	AD281.1	rpoB2	66.7%	Yes	3/3	1/3			Dabney et al. [3]
New World	Colombia	Las Delicias	LD-90-1 X-11	AD281.1	IS6110	100.0%	Yes	3/3	3/3			Dabney et al. [3]
New World	Colombia	Las Delicias	LD-90-1 X-11	AD281.1	rpoB	0.0%	No		0/3			Dabney et al. [3]
New World	Colombia	Las Delicias	LD-90-1 X-11	AD281.1	IS081	66.7%	Yes	3/3	1/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	1002	AD69.2	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	1002	AD69.2	rpoB2	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	1002	AD69.2	rpoB	0.0%	No	0/6	0/3			modified Rohland & Hofreiter [1]

New World	Moquegua, Peru	M6: Estuquiña	1002	AD69.2	IS1081	0.0%	No	0/6				modified Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	3215	AD79.1	rpoB	83.3%	Yes	2/3	3/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	3215	AD79.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	3215	AD79.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.7	IS1081	100.0%	Yes	3/3				Yang [2]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.8	IS1081	100.0%	Yes	3/3				Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.3	rpoB	100.0%	Yes		2/2			Yang [2]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.3	rpoB2	0.0%	No	0/2				Yang [2]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.4	rpoB	100.0%	Yes	3/3				Yang [2]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.4	IS6110	100.0%	Yes	3/3				Yang [2]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.4	IS1081	100.0%	Yes	3/3				Yang [2]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.5	rpoB	91.7%	Yes	3/3	3/3	3/3	2/3	Yang [2]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.5	rpoB2	85.7%	Yes		8/8	3/3	1/3	Yang [2]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.7	IS6110	100.0%	Yes	3/3				Yang [2]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.9	rpoB	100.0%	Yes	3/3				Yang [2]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.7	rpoB	100.0%	Yes	3/3				Yang [2]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.8	IS6110	100.0%	Yes	3/3				Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	4165	AD82.8	rpoB	100.0%	Yes	3/3				Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	4256	AD85.1	rpoB	16.7%	No	1/3	0/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	4256	AD85.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	4256	AD85.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	5859	AD87.1	rpoB	50.0%	Yes	0/3	3/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	5859	AD87.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	5859	AD87.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	2279b	AD76.2	IS6110	0.0%	No	0/6				Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	2279b	AD76Tu	rpoB2	0.0%	No	0/5				modified Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	2279b	AD76.2	rpoB	100.0%	Yes	6/6				Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	2279b	AD76.2	IS1081	0.0%	No	0/6				Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	3644a	AD81Tu	rpoB2	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	3644a	AD81.1	rpoB	16.7%	No	0/3	1/3			modified Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	3644a	AD81.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	3644a	AD81.1	IS1081	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.1	rpoB2	0.0%	No		0/3			Höss & Pääbo [5]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.1	IS6110	0.0%	No	0/3				Höss & Pääbo [5]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.1	IS1081	0.0%	No	0/3				Höss & Pääbo [5]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.2	rpoB	0.0%	No	0/3				Höss & Pääbo [5]

New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.2	IS6110	0.0%	No	0/3				Höss & Pääbo [5]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.2	IS1081	0.0%	No	0/3				Höss & Pääbo [5]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.3	rpoB	100.0%	Yes	3/3				Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.1	rpoB	100.0%	Yes	3/3	3/3			Höss & Pääbo [5]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.3	IS6110	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.3	IS1081	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.4	IS6110	0.0%	No	0/6				Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.4	rpoB	100.0%	Yes	6/6				Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	99407a	AD89.4	IS1081	0.0%	No	0/6				Rohland & Hofreiter [1]
New World	Moquegua, Peru	M6: Estuquiña	99407b	AD90.1	rpoB	50.0%	Yes	0/3	3/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	99407b	AD90.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	99407b	AD90.1	IS6110	16.7%	No	1/3	0/3			Dabney et al. [3]
New World	Moquegua, Peru	M6: Estuquiña	99407b	#221	IS1081	0.0%	No	0/3				Höss & Pääbo [5]
New World	Moquegua, Peru	M6: Estuquiña	99407b	#221	rpoB	0.0%	No	0/3				Höss & Pääbo [5]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 13	AD253.1	IS6110	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 13	AD253.1	rpoB	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 14 (1)	AD251.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 14 (1)	AD251.2	rpoB	37.5%	No	0/2	2/3	1/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 14 (1)	AD251.1	rpoB	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 14 (1)	AD251.1	rpoB2	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 14 (1)	AD251.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 14 (1)	AD251.2	rpoB2	0.0%	No	0/2	0/3	0/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 14 (2)	AD259.1	rpoB	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 14 (2)	AD259.1	rpoB2	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 14 (2)	AD259.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 14 (2)	AD259.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 15	AD265.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]

New World	Tierra del Fuego, Argentina	Misión Salesiana	C 15	AD265.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 15	AD265.1	rpoB	100.0%	Yes	3/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 15-16	AD263.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 15-16	AD263.1	rpoB	100.0%	Yes	3/3	2/2			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 15-16	AD263.1	rpoB2	0.0%	No		0/2			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 7-8	AD252.1	rpoB	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 7-8	AD252.1	IS6110	33.3%	No	1/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 7-8	AD252.1	IS1081	33.3%	No	1/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 7-8	AD252.3	rpoB2	0.0%	No	0/2	0/3	0/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	C 7-8	AD252.3	rpoB	12.5%	No	0/2	1/3	0/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D 14	AD264.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D 14	AD264.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D 14	AD264.1	rpoB2	0.0%	No		0/2			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D 14	AD264.1	rpoB	100.0%	Yes	3/3	2/2			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D 14	AD264Tu	rpoB2	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D 15-16	AD262Tu	rpoB	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D 15-16	AD262Tu	rpoB2	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D 15-16	AD262Tu	IS1081	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D 15-16	AD262TuB	rpoB	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D 16 (Bis)	AD258Tu	IS1081	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D 16 (Bis)	AD258.1	rpoB	12.5%	No	0/2	0/3	1/3		modified Rohland & Hofreiter [1]

New World	Tierra del Fuego, Argentina	Misión Salesiana	D 16 (Bis)	AD258.1	rpoB2	0.0%	No	0/2	0/3	0/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256.3	IS6110	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256.3	rpoB	100.0%	Yes	3/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256.2	IS1081	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256.2	IS6110	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256.3	IS1081	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256.2	rpoB	100.0%	Yes	3/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256.1	IS6110	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256.1	rpoB	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256Tu	rpoB2	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256.4	rpoB	100.0%	Yes	3/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256.4	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	D-E 14	AD256.4	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 12-13	AD255.1	rpoB2	0.0%	No	0/2	0/3	0/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 12-13	AD255.1	rpoB	12.5%	No	0/2	1/3	0/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 14-15 (1)	AD250.2	rpoB	12.5%	No	0/2	0/3	1/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 14-15 (1)	AD250Tu	IS1081	16.7%	No	1/3	0/3			modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 14-15 (1)	AD250.2	rpoB2	0.0%	No	0/2	0/3	0/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 14-15 (2)	AD249.1	rpoB	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 14-15 (2)	AD249.1	IS6110	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 14-15 (2)	AD249.2	rpoB	25.0%	No	0/2	1/3	1/3		modified Rohland & Hofreiter [1]

New World	Tierra del Fuego, Argentina	Misión Salesiana	E 14-15 (2)	AD249.2	rpoB2	0.0%	No	0/2	0/3	0/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (1)	AD257.1	rpoB	50.0%	Yes	3/6				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (1)	AD257.1	IS6110	0.0%	No	0/6				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (1)	AD257.1	IS1081	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (1)	AD257.3	rpoB	100.0%	Yes	3/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (1)	AD257.3	IS6110	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (1)	AD257.3	IS1081	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (1)	AD257.4	rpoB2	0.0%	No	0/2	0/3	0/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (1)	AD257.4	rpoB	62.5%	Yes	2/2	2/3	1/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (2 bis)	AD247.1	rpoB	0.0%	No	0/6				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (2 bis)	AD247.1	IS6110	0.0%	No	0/6				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (2 bis)	AD247.1	IS1081	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (3)	AD248.1	rpoB	0.0%	No	0/6				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (3)	AD248.1	IS6110	16.7%	No	1/6				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (3)	AD248.1	IS1081	0.0%	No	0/3				Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (3)	AD248.3	rpoB	37.5%	No	1/2	1/3	1/3		modified Rohland & Hofreiter [1]
New World	Tierra del Fuego, Argentina	Misión Salesiana	E 15-16 (3)	AD248.3	rpoB2	0.0%	No	0/2	0/3	0/3		modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	21	AD145.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	21	AD145.1	rpoB	60.0%	Yes	1/3	2/2			modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	21	AD145.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	21	AD145.1	rpoB2	0.0%	No		0/2			modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	38	AD146.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	38	AD146.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	38	AD146.1	rpoB	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]

New World	Illinois	Norris Farms	38	AD146.1	rpoB2	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	95	AD151.1	rpoB	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	95	AD151.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	95	AD151.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	95	AD151Tu	rpoB2	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	160	AD156.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	160	AD156.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	160	AD156.1	rpoB	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	224	AD159.1	rpoB	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	224	AD159.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	224	AD159.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	228	AD160Tu	rpoB2	66.7%	Yes		2/3			modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	228	AD160Tu	rpoB	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	Illinois	Norris Farms	228	AD160Tu	IS1081	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
New World	Illinois	Orendorf	Burial 116	AD171.1	rpoB	0.0%	No		0/3			Dabney et al. [3]
New World	Illinois	Orendorf	Burial 116	AD171.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Illinois	Orendorf	Burial 116	AD171.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Illinois	Orendorf	Burial 153	AD172.1	rpoB	0.0%	No		0/3			Dabney et al. [3]
New World	Illinois	Orendorf	Burial 153	AD172.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Illinois	Orendorf	Burial 153	AD172.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Peru	Puruchuco	Tomb 49	AD211.1	rpoB	0.0%	No		0/3			Dabney et al. [3]
New World	Peru	Puruchuco	Tomb 49	AD211.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Peru	Puruchuco	Tomb 49	AD211.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Peru	Puruchuco	Tomb 5, Indv. 2	AD209.1	rpoB	0.0%	No		0/3			Dabney et al. [3]
New World	Peru	Puruchuco	Tomb 5, Indv. 2	AD209.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Peru	Puruchuco	Tomb 5, Indv. 2	AD209.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	New Mexico	San Cristobal	8634	AD165.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	New Mexico	San Cristobal	8634	AD165.1	rpoB2	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	New Mexico	San Cristobal	8634	AD165.1	rpoB	44.4%	No	1/3	3/6			modified Rohland & Hofreiter [1]
New World	New Mexico	San Cristobal	8634	AD165.1	IS1081	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]
New World	New Mexico	San Cristobal	8708	AD162Tu	rpoB2	20.0%	No	1/2	0/3			modified Rohland & Hofreiter [1]
New World	New Mexico	San Cristobal	8708	AD162.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	New Mexico	San Cristobal	8708	AD162.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	New Mexico	San Cristobal	8708	AD162.1	rpoB	0.0%	No	0/3				modified Rohland & Hofreiter [1]

New World	Patagonia (Argentina)	Santa Cruz, Patagonia	SAC4-1-1	AD225.1	IS1081	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]
New World	Patagonia (Argentina)	Santa Cruz, Patagonia	SAC4-1-1	AD225.1B	IS1081	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Patagonia (Argentina)	Santa Cruz, Patagonia	SAC4-1-1	AD225.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Patagonia (Argentina)	Santa Cruz, Patagonia	SAC4-1-1	AD225.1B	rpoB	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Patagonia (Argentina)	Santa Cruz, Patagonia	SAC4-1-1	AD225.1	rpoB	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]
New World	Illinois	Schild	SA-81	AD93.1	rpoB	40.0%	No	0/2	2/3			Dabney et al. [3]
New World	Illinois	Schild	SA-81	AD93.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Illinois	Schild	SB-262A	AD102.1	IS6110	0.0%	No	0/3				Dabney et al. [3]
New World	Illinois	Schild	SB-262A	AD102.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Punta Arenas	Sitio Myren-7	T.D.F. 30477	AD216.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Punta Arenas	Sitio Myren-7	T.D.F. 30477	AD216.1	IS1081	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]
New World	Punta Arenas	Sitio Myren-7	T.D.F. 30477	AD216.2	rpoB	100.0%	Yes	2/2	3/3	3/3		modified Rohland & Hofreiter [1]
New World	Punta Arenas	Sitio Myren-7	T.D.F. 30477	AD216.1	rpoB	100.0%	Yes	3/3	3/3			modified Rohland & Hofreiter [1]
New World	Punta Arenas	Sitio Myren-7	T.D.F. 30477	AD216.2	rpoB2	37.5%	No	2/2	1/3	0/3		modified Rohland & Hofreiter [1]
New World	Ontario, Canada	Uxbridge	47	AD31.1	rpoB	16.7%	No	1/3	0/3			Dabney et al. [3]
New World	Ontario, Canada	Uxbridge	47	AD31.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Ontario, Canada	Uxbridge	47	AD31.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Ontario, Canada	Uxbridge	55	AD33.1	IS6110	0.0%	No	0/3				Yang [2]
New World	Ontario, Canada	Uxbridge	55	AD33.2	rpoB2	0.0%	No		0/3			modified Rohland & Hofreiter [1]
New World	Ontario, Canada	Uxbridge	55	AD33.2	rpoB	0.0%	No	0/3	0/3			modified Rohland & Hofreiter [1]
New World	Ontario, Canada	Uxbridge	55	AD33.2	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Ontario, Canada	Uxbridge	55	AD33.2	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
New World	Ontario, Canada	Uxbridge	55	AD33.1	rpoB	0.0%	No	0/3				Yang [2]
New World	Ontario, Canada	Uxbridge	55	AD33.1	IS1081	0.0%	No	0/3				Yang [2]
New World	Ontario, Canada	Uxbridge	219	AD41.1	IS1081	0.0%	No	0/3				Yang [2]
New World	Ontario, Canada	Uxbridge	219	AD41.1	IS6110	0.0%	No	0/3				Yang [2]
New World	Ontario, Canada	Uxbridge	219	AD41.1	rpoB	0.0%	No	0/3				Yang [2]
New World	Ontario, Canada	Uxbridge	280	AD44.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Ontario, Canada	Uxbridge	280	AD44.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Ontario, Canada	Uxbridge	280	AD44.1	rpoB	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Ontario, Canada	Uxbridge	16b	AD28.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Ontario, Canada	Uxbridge	16b	AD28.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Ontario, Canada	Uxbridge	16b	AD28.1	rpoB	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Osmore Valley, Peru	Yaral-1 (Cemetery 1)	59-1-E	AD60.1	rpoB	16.7%	No	0/3	1/3			Dabney et al. [3]
New World	Osmore Valley, Peru	Yaral-1 (Cemetery 1)	59-1-E	AD60.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Osmore Valley, Peru	Yaral-1 (Cemetery 1)	59-1-E	AD60.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Osmore Valley, Peru	Yaral-2 (Cemetery 2)	10291	AD64.1	IS1081	66.7%	Yes	3/3	1/3			modified Rohland & Hofreiter [1]
New World	Osmore Valley, Peru	Yaral-2 (Cemetery 2)	10291	AD64.1	IS6110	100.0%	Yes	3/3				modified Rohland & Hofreiter [1]
New World	Osmore Valley, Peru	Yaral-2 (Cemetery 2)	10291	AD64.1	rpoB	83.3%	Yes	3/3	2/3			modified Rohland & Hofreiter [1]

New World	Osmore Valley, Peru	Yaral-2 (Cemetery 2)	10335	AD65.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Osmore Valley, Peru	Yaral-2 (Cemetery 2)	10335	AD65.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Osmore Valley, Peru	Yaral-2 (Cemetery 2)	10335	AD65.1	rpoB	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Illinois	Yokem	Md 2-5	AD109.1	rpoB2	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Illinois	Yokem	Md 2-5	AD109.1	IS6110	0.0%	No	0/3	0/3			Dabney et al. [3]
New World	Illinois	Yokem	Md 2-5	AD109.1	rpoB	0.0%	No		0/3			Dabney et al. [3]
Old World	Italy	Arena Candide	n/a	AD282.1	rpoB	0.0%	No	0/3				modified Rohland & Hofreiter [1]
Old World	Italy	Arena Candide	n/a	AD282Tu	rpoB2	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
Old World	Italy	Arena Candide	n/a	AD282.1	IS6110	0.0%	No	0/3				modified Rohland & Hofreiter [1]
Old World	Italy	Arena Candide	n/a	AD282.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
Old World	Japan	Okinawa Island	n/a	AD224.2	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
Old World	Japan	Okinawa Island	n/a	AD224.2	rpoB	33.3%	No	1/3				modified Rohland & Hofreiter [1]
Old World	Japan	Okinawa Island	n/a	AD224.1	IS1081	0.0%	No	0/3				modified Rohland & Hofreiter [1]
Old World	Japan	Okinawa Island	n/a	AD224.1	rpoB	100.0%	Yes	3/3				modified Rohland & Hofreiter [1]
Old World	England	St. Mary Spital	2783	sp2783	rpoB	0.0%	No	0/6				Yang [2]
Old World	England	St. Mary Spital	2783	sp2783	IS1081	0.0%	No	0/6				Yang [2]
Old World	England	St. Mary Spital	3549	sp3549	IS1081	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	3549	sp3549	rpoB	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	5426	#224	IS1081	0.0%	No	0/3				Höss & Pääbo [5]
Old World	England	St. Mary Spital	5426	#224	rpoB	0.0%	No	0/3				Höss & Pääbo [5]
Old World	England	St. Mary Spital	8490	sp8490	IS1081	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	8490	sp8490	rpoB	0.0%	No	0/6				Yang [2]
Old World	England	St. Mary Spital	8852	sp8852	rpoB	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	8852	sp8852	IS1081	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	10686	sp10686	IS6110	33.3%	No	1/3				Yang [2]
Old World	England	St. Mary Spital	10686	#223	rpoB	0.0%	No	0/3				Höss & Pääbo [5]
Old World	England	St. Mary Spital	10686	sp10686	IS1081	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	10686	#223	IS1081	0.0%	No	0/3				Höss & Pääbo [5]
Old World	England	St. Mary Spital	10686	sp10686	rpoB	0.0%	No	0/6				Yang [2]
Old World	England	St. Mary Spital	12188	AD235Tu	rpoB2	20.0%	No	1/2	0/3			modified Rohland & Hofreiter [1]
Old World	England	St. Mary Spital	12188	sp12188	rpoB	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	12188	sp12188	IS1081	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	13759	AD236Tu	rpoB2	20.0%	No	1/2	0/3			modified Rohland & Hofreiter [1]
Old World	England	St. Mary Spital	13759	sp13759	rpoB	0.0%	No	0/6				Yang [2]
Old World	England	St. Mary Spital	13759	sp13759	IS1081	0.0%	No	0/6				Yang [2]
Old World	England	St. Mary Spital	17769	sp17769	IS1081	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	17769	sp17769	rpoB	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	19599	sp19599	IS1081	0.0%	No	0/6				Yang [2]
Old World	England	St. Mary Spital	19599	sp19599	rpoB	0.0%	No	0/6				Yang [2]
Old World	England	St. Mary Spital	22806	sp22806	IS1081	16.7%	No	1/6				Yang [2]
Old World	England	St. Mary Spital	22806	sp22806	rpoB	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	22806	AD240Tu	rpoB2	0.0%	No	0/2	0/3			modified Rohland & Hofreiter [1]
Old World	England	St. Mary Spital	22829	sp22829	rpoB	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	22829	sp22829	IS1081	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	23630	sp23630	rpoB	33.3%	No	2/6				Yang [2]
Old World	England	St. Mary Spital	23630	sp23630	IS6110	0.0%	No	0/3				Yang [2]

Old World	England	St. Mary Spital	23630	sp23630	IS1081	0.0%	No	0/6				Yang [2]
Old World	England	St. Mary Spital	25714	sp25714	rpoB	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	25714	sp25714	IS1081	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	27372	sp27372	rpoB	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	27372	sp27372	IS1081	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	29490	sp29490	rpoB	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	29490	sp29490	IS1081	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	30702	sp30702	rpoB	0.0%	No	0/3				Yang [2]
Old World	England	St. Mary Spital	30702	sp30702	IS1081	0.0%	No	0/3				Yang [2]

Samples Tested with qPCR 2007-2008 (see Table S4)

World Region	Archaeological Region	Archaeological Site	Burial/ Indiv Info	Extract Name	Target Name	Total Percent Positive	Is Positive Assay	Dilution 1:1 Positive	Date			Extraction protocol
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	196	#163	rpoB	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	196	#163	IS6110	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
Old World	Italy	Arena Candide	n/a	#212	rpoB	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
Old World	Italy	Arena Candide	n/a	#212	IS6110	33.3%	No	1/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
Old World	Italy	Arma dell'Aquila	n/a	#211	rpoB	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
Old World	Italy	Arma dell'Aquila	n/a	#211	IS6110	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 6	1942-51	#204	rpoB	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 6	1942-51	#204	IS6110	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Peru	Lambayeque	CSJ-21	#215	rpoB	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Peru	Lambayeque	CSJ-21	#215	IS6110	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Peru	Lambayeque	ILL-22	#214	rpoB	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Peru	Lambayeque	ILL-22	#214	IS6110	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
Old World	Japan	Okinawa Island	Okinawa	#210	rpoB	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
Old World	Japan	Okinawa Island	Okinawa	#210	IS6110	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
Old World	Russia	Russian Academy of Science	TM-1999 p-3, n.5	#218	rpoB	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
Old World	Russia	Russian Academy of Science	TM-1999 p-3, n.5	#218	IS6110	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Tierra del Fuego, Argentina	Sitio Myren-7	T.D.F. 30477	#216	rpoB	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Tierra del Fuego, Argentina	Sitio Myren-7	T.D.F. 30477	#216	IS6110	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Arikara	Sully School Village	388283	#217	rpoB	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Arikara	Sully School Village	388283	#217	IS6110	0.0%	No	0/3	2/21/2008			Höss & Pääbo [5] or Boom[4]
New World	Ontario, Canada	Uxbridge	P1	UbEx11	IS6110	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P5	UbEx14.2	IS6110	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P6	UbEx17	IS6110	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P7	UbEx12	IS6110	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P57	UbEx13	IS6110	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P132	UbEx9.2	IS6110	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P159	UbEx15.2	IS6110	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P165	UbEx10.2	IS6110	0.0%	No	0/3	8/8/2008			Boom [4]

New World	Ontario, Canada	Uxbridge	P253	UbEx16	IS6110	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P1	UbEx11	rpoB	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P5	UbEx14.2	rpoB	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P6	UbEx17	rpoB	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P7	UbEx12	rpoB	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P57	UbEx13	rpoB	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P132	UbEx9.2	rpoB	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P159	UbEx15.2	rpoB	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P165	UbEx10.2	rpoB	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Ontario, Canada	Uxbridge	P253	UbEx16	rpoB	0.0%	No	0/3	8/8/2008			Boom [4]
New World	Moquegua, Peru	M6: Estuquina	4213a	#143	rpoB	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	M6-6464	#82	rpoB	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	M6-6464	#82	IS6110	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	2256a	#167	rpoB	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	2256a	#167	IS6110	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	2279b	#169	IS6110	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	4165	#171	rpoB	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	4165	#171	IS6110	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	4213a	#143	IS6110	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	5859	#155	rpoB	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	5859	#155	IS6110	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Puruchuco	Tomb 5, Indv. 2	#147	rpoB	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Peru	Puruchuco	Tomb 5, Indv. 2	#147	IS6110	0.0%	No	0/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Peru	Uxbridge	P1	UbEx11	rpoB	0.0%	No	0/3	8/27/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	P45	UbEx5	rpoB	0.0%	No	0/3	8/27/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	P48	UbEx3	rpoB	0.0%	No	0/3	8/27/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	P55	UbEx6	rpoB	0.0%	No	0/3	8/27/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	P132	UbEx9	rpoB	0.0%	No	0/3	8/27/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	P135	UbEx8	rpoB	0.0%	No	0/3	8/27/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	P165	UbEx10	rpoB	0.0%	No	0/3	8/27/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	P219	UbEx4	rpoB	0.0%	No	0/3	8/27/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	P16a	UbEx2	rpoB	0.0%	No	0/3	8/27/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	U3 L7 S8	UbEx7	rpoB	0.0%	No	0/3	8/27/2007			Boom [4]
New World	Ontario, Canada	Arikara, Cheyenne River village	382674	382674	rpoB	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382674	382674	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382699	82699 (femu	rpoB	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382699	82699 (femu	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382699	382699 (rib	rpoB	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382699	382699 (rib	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382706	382706	rpoB	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382706	382706	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382713	382713	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Sully School Village	388106	388106	rpoB	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Sully School Village	388106	388106	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]

New World	South Dakota	Chen Chen	M1-2289	#184	rpoB	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Chen Chen	M1-2289	#184	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-2479	#186	rpoB	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-2479	#186	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-3353-1	#192	rpoB	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-3353-1	#192	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chiribaya Alta, Cemetery 1	2069	#205	rpoB	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chiribaya Alta, Cemetery 1	2069	#205	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 1	2069	#207	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 6	1942-51	#204	rpoB	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Osmore Valley, Peru	Chiribaya Alta, Cemetery 6	1942-51	#204	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Osmore Valley, Peru	M6: Estuquina	99407a	#176	rpoB	33.3%	No	1/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Osmore Valley, Peru	M6: Estuquina	99407a	#176	IS6110	0.0%	No	0/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	Osmore Valley, Peru	Uxbridge	P55	UbEx6	rpoB	0.0%	No	0/3	9/11/2007			Boom [4]
New World	Moquegua, Peru	Uxbridge	P55	UbEx6	IS6110	0.0%	No	0/3	9/11/2007			Boom [4]
New World	Moquegua, Peru	Uxbridge	P135	Ubx8	rpoB	0.0%	No	0/3	9/11/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	P135	Ubx8	IS6110	0.0%	No	0/3	9/11/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	U3 L7 S8	UbEx7	rpoB	0.0%	No	0/3	9/11/2007			Boom [4]
New World	Ontario, Canada	Uxbridge	U3 L7 S8	UbEx7	IS6110	0.0%	No	0/3	9/11/2007			Boom [4]
New World	Ontario, Canada	Arikara, Leavenworth	325341	325341	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Ontario, Canada	Puruchuco	Tomb 69	#149	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Ontario, Canada	Puruchuco	Tomb 69	#149	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	M6: Estuquiña	99407b	#221	IS1081	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Peru	M6: Estuquiña	99407b	#221	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Peru	M6: Estuquina	4176	#174	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	4176	#174	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	4256	#175	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	4256	#175	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	M6-5390	#78	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	M6-5390	#78	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-0243	#178	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-0243	#178	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-0646	#177	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-0646	#177	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-1340	#180	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-1340	#180	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-1573	#181	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-1573	#181	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-1658	#182	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-1658	#182	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-2289	#187	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-2289	#187	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-2705	#188	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Chen Chen	M1-2705	#188	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	M6: Estuquina	2279b	#169	rpoB	100.0%	Yes	3/3	8/27/2007			Höss & Pääbo [5] or Boom[4]
New World	Moquegua, Peru	Arikara, Cheyenne River village	382713	382713	rpoB	66.7%	Yes	2/3	9/11/2007			Höss & Pääbo [5] or Boom[4]

New World	Moquegua, Peru	Chiribaya Alta, Cemetery 1	2069	#207	rpoB	66.7%	Yes	2/3	9/11/2007			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Sully School Village	381390	381390	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Sully School Village	381390	381390	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Sully School Village	381407	381407	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Sully School Village	381407	381407	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Leavenworth	325341	325341	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382669	382669	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382669	382669	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382710	382710	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Cheyenne River village	382710	382710	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Leavitt	382742	382742	rpoB	33.3%	No	1/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Leavitt	382742	382742	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Mobridge	382990	382990	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Mobridge	382990	382990	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Sully School Village	388120	388120	rpoB	66.7%	Yes	2/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Sully School Village	388120	388120	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Sully School Village	388243	388243	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Arikara, Sully School Village	388243	388243	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Sully School Village	388403	388403	rpoB	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]
New World	South Dakota	Sully School Village	388403	388403	IS6110	0.0%	No	0/3	2/8/2008			Höss & Pääbo [5] or Boom[4]

[1] Rohland, N. & Hofreiter, M. 2007 Comparison and optimization of ancient DNA extraction. BioTechniques 42, 343-352.

[2] Yang, D.Y., Eng, B., Waye, J.S., Dudar, J.C. & Saunders, S.R. 1998 Technical note: improved DNA extraction from ancient bones using silica-based spin columns. AJPA 105, 539-543.

[3] Dabney, J., Knapp, M., Glocke, I., Gansauge, M.T., Weihmann, A., Nickel, B., Valdiosera, C., Garcia, N., Paabo, S., Arsuaga, J.L., et al. 2013 Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proceedings of the National Academy of Sciences USA 110, 15758-15763.

[4] Boom, R., Sol, C.J., Salimans, M.M., Jansen, C.L., Wertheim-van Dillen, P.M. & van der Noordaa, J. 1990 Rapid and simple method for purification of nucleic acids. J Clin Microbiol 28, 495-503.

[5] Höss, M. & Paabo, S. 1993 DNA extraction from Pleistocene bones by a silica-based purification method. Nucleic Acids Research 21, 3913-3914

Table S4: Assay summary information

PlateID	TargetID	Y-Intercept	Slope	R-Squared	Efficiency	Threshold Value	Threshold Type	Baseline Start	Baseline Finish	Baseline Type	Is Excluded	Exclusion Reason
aTBassay1	rpoB	35.6856	-3.268946	0.9981854	2.022595327	0.050758332	Manual	2	15	Manual	FALSE	
aTBassay2	IS6110	37.368637	-3.322249	0.9979378	1.999866099	0.093504556	Manual	1	16	Manual	FALSE	
aTBassay2	IS1081	40.918583	-3.2590034	0.99713284	2.026946424	0.5085541	Automatic			Automatic	FALSE	
aTBassay3	IS6110	36.60743	-3.3223803	0.9985782	1.999811322	0.055	Manual	2	15	Manual	FALSE	
aTBassay4	IS6110	38.948536	-3.516551	0.99934256	1.924728791	0.25362164	Automatic			Automatic	FALSE	
aTBassay9	IS1081	33.171753	-3.2485185	0.9968982	2.031573934	0.017285265	Manual			Automatic	FALSE	
aTBassay11	rpoB	34.860554	-3.3333712	0.9988623	1.995246658	0.036885902	Manual	3	18	Manual	FALSE	
aTBassay14	rpoB	34.8234	-3.2310762	0.9869933	2.039362378	0.07140249	Manual	1	13	Manual	FALSE	
aTBassay14	IS1081	37.333687	-3.3044276	0.9971145	2.007355411	0.39103064	Manual	2	14	Manual	FALSE	
aTBassay17	rpoB	34.38798	-3.4420607	0.99532735	1.952196991	0.07	Manual	1	17	Manual	FALSE	
aTBassay22	rpoB2	40.613144	-4.4181743	0.9980346	1.683983507	0.15	Manual			Automatic	FALSE	
aTBassay23	IS6110	42.10982	-4.519468	0.9971744	1.664427876	0.092528135	Manual	3	11	Manual	FALSE	
aTBassay29	rpoB	39.714043	-3.6846309	0.98837173	1.868089178	0.20689133	Manual	1	20	Manual	FALSE	
aTBassay5	IS6110	36.160122	-3.3673632	0.9934009	1.981382226	0.03402724	Manual	2	14	Manual	FALSE	
aTBassay5	IS1081	37.84281	-3.3502429	0.9977471	1.988317891	0.19034876	Manual	2	16	Manual	FALSE	
aTBassay33	IS1081	40.815266	-4.106871	0.98376596	1.751839757	0.19337916	Manual			Automatic	FALSE	
aTBassay28.2	rpoB2	40.727398	-4.709445	0.97380966	1.630569242	0.024408793	Manual			Automatic	FALSE	
aTBassay6	IS1081	35.48128	-3.3196585	0.9981917	2.000948011	0.08299999	Manual	1	15	Manual	FALSE	
aTBassay2	rpoB	36.71058	-3.4018943	0.992633	1.967677282	0.03834868	Manual	3	19	Manual	FALSE	
aTBassay3	rpoB	37.887726	-3.562026	0.9980446	1.908706302	0.055	Manual	1	20	Manual	FALSE	
aTBassay4	rpoB	39.25199	-3.6599162	0.99841076	1.875989052	0.18291628	Automatic			Automatic	FALSE	
aTBassay5	rpoB	39.150284	-3.4055078	0.997586	1.966264618	0.02484407	Manual	1	19	Manual	FALSE	
aTBassay17	rpoB2	34.616123	-3.379525	0.99949175	1.976512529	0.113281846	Manual	1	8	Manual	FALSE	
aTBassay20	IS1081	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	FALSE	Stanards inconsistent, only rec positive/negative
TB1	rpoB	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	TRUE	Stanards inconsistent, only rec positive/negative
TB1	IS6110	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	TRUE	Stanards inconsistent, only rec positive/negative

TB3	rpoB	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	TRUE	Stanards inconsistent, only rec positive/negative
TB3	IS6110	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	TRUE	Stanards inconsistent, only rec positive/negative
TB4	rpoB	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	TRUE	Stanards inconsistent, only positive/negative
TB4	IS6110	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	TRUE	Stanards inconsistent, only rec positive/negative

[redacted]

Appendix A

Harkins KM, Buikstra JE, Campbell TJ, Bos KI, Johnson ED, Krause J, Stone AC. Screening Ancient Tuberculosis with qPCR: Challenges and Opportunities.

Table S5: SDS 2.4 software output for all extracts tested 2012-2014

Plate Date	Assay Name	Lab Sample Name	Extract Name	Dilution Name	Ct Value	Ct Mean	Ct Std Dev	Quantity Copies	Quantity Copies Mean	Quantity Copies Std Dev	Mean Copies X Dilution	Is Excluded	Exclusion Reason
25-Jan-12	IS1081 on aTBassay1	AD1	AD1.3	AD1.3 1:1	Undetermined			0			0		
25-Jan-12	IS1081 on aTBassay1	AD1	AD1.3	AD1.3 1:1	Undetermined			0			0		
25-Jan-12	IS1081 on aTBassay1	AD1	AD1.3	AD1.3 1:1	Undetermined			0			0		
25-Jan-12	IS1081 on aTBassay1	AD247	AD247.1	AD247.1 1:1	Undetermined			0			0		
25-Jan-12	IS1081 on aTBassay1	AD248	AD248.1	AD248.1 1:1	Undetermined			0			0		
25-Jan-12	IS1081 on aTBassay1	AD76	AD76.2	AD76.2 1:1	Undetermined			0			0		
25-Jan-12	IS1081 on aTBassay1	AD89	AD89.3	AD89.3 1:1	Undetermined			0			0		
25-Jan-12	IS6110 on aTBassay1	AD1	AD1.3	AD1.3 1:1	Undetermined			0			0		
25-Jan-12	IS6110 on aTBassay1	AD247	AD247.1	AD247.1 1:1	Undetermined			0			0		
25-Jan-12	IS6110 on aTBassay1	AD248	AD248.1	AD248.1 1:1	Undetermined			0			0		
25-Jan-12	IS6110 on aTBassay1	AD248	AD248.1	AD248.1 1:1	32.162243			46.652596			46.652596		
25-Jan-12	IS6110 on aTBassay1	AD82	AD82.7	AD82.7 1:1	33.85747	33.40942	0.4103887	15.744047	21.45291	5.362256	21.45291		
25-Jan-12	IS6110 on aTBassay1	AD89	AD89.3	AD89.3 1:1	Undetermined			0			0		
25-Jan-12	IS6110 on aTBassay1	AD89	AD89.3	AD89.3 1:1	Undetermined			0			0		
25-Jan-12	IS6110 on aTBassay1	AD89	AD89.3	AD89.3 1:1	Undetermined			0			0		
25-Jan-12	rpoB on aTBassay1	AD1	AD1.3	AD1.3 1:1	27.684273	27.985117	0.3313375	280.33548	230.86176	52.028664	230.86176		
25-Jan-12	rpoB on aTBassay1	AD1	AD1.3	AD1.3 1:1	28.340244	27.985117	0.3313375	176.60803	230.86176	52.028664	230.86176		
25-Jan-12	rpoB on aTBassay1	AD1	AD1.3	AD1.3 1:1	27.930836	27.985117	0.3313375	235.64174	230.86176	52.028664	230.86176		
25-Jan-12	rpoB on aTBassay1	AD247	AD247.1	AD247.1 1:1	Undetermined			0			0		
25-Jan-12	rpoB on aTBassay1	AD247	AD247.1	AD247.1 1:1	Undetermined			0			0		
25-Jan-12	rpoB on aTBassay1	AD247	AD247.1	AD247.1 1:1	Undetermined			0			0		
25-Jan-12	rpoB on aTBassay1	AD248	AD248.1	AD248.1 1:1	Undetermined			0			0		
25-Jan-12	rpoB on aTBassay1	AD248	AD248.1	AD248.1 1:1	Undetermined			0			0		
25-Jan-12	rpoB on aTBassay1	AD248	AD248.1	AD248.1 1:1	Undetermined			0			0		
25-Jan-12	rpoB on aTBassay1	AD76	AD76.2	AD76.2 1:1	25.236998	25.187243	0.0814652	1571.5266	1629.3844	94.98563	1629.3844		
25-Jan-12	rpoB on aTBassay1	AD76	AD76.2	AD76.2 1:1	25.231504	25.187243	0.0814652	1577.6189	1629.3844	94.98563	1629.3844		
25-Jan-12	rpoB on aTBassay1	AD76	AD76.2	AD76.2 1:1	25.09323	25.187243	0.0814652	1739.008	1629.3844	94.98563	1629.3844		
25-Jan-12	rpoB on aTBassay1	AD82	AD82.7	AD82.7 1:1	34.002556	34.29558	0.4732429	3.27234	2.7552462	0.8179836	2.7552462		
25-Jan-12	rpoB on aTBassay1	AD82	AD82.7	AD82.7 1:1	34.84155	34.29558	0.4732429	1.8121866	2.7552462	0.8179836	2.7552462		
25-Jan-12	rpoB on aTBassay1	AD82	AD82.7	AD82.7 1:1	34.042652	34.29558	0.4732429	3.1812122	2.7552462	0.8179836	2.7552462		
25-Jan-12	rpoB on aTBassay1	AD89	AD89.3	AD89.3 1:1	26.257193	26.126778	0.3360174	766.0108	856.0798	212.44626	856.0798		
25-Jan-12	rpoB on aTBassay1	AD89	AD89.3	AD89.3 1:1	26.378036	26.126778	0.3360174	703.50574	856.0798	212.44626	856.0798		
30-Jan-12	IS1081 on aTBassay2	AD1	AD1.2	AD1.2 1:1	Undetermined			0			0		
30-Jan-12	IS1081 on aTBassay2	AD1	AD1.2	AD1.2 1:1	Undetermined			0			0		
30-Jan-12	IS1081 on aTBassay2	AD76	AD76.2	AD76.2 1:1	Undetermined			0			0		
30-Jan-12	IS1081 on aTBassay2	AD76	AD76.2	AD76.2 1:1	Undetermined			0			0		
30-Jan-12	IS1081 on aTBassay2	AD89	AD89.2	AD89.2 1:1	Undetermined			0			0		
30-Jan-12	IS6110 on aTBassay2	AD1	AD1.2	AD1.2 1:1	Undetermined			0			0		
30-Jan-12	IS6110 on aTBassay2	AD76	AD76.2	AD76.2 1:1	Undetermined			0			0		

30-Jan-12	IS6110 on aTBassay2	AD89	AD89.2	AD89.2 1:1	Undetermined			0			0
30-Jan-12	IS6110 on aTBassay2	AD89	AD89.4	AD89.4 1:1	Undetermined			0			0
30-Jan-12	rpoB on aTBassay2	AD76	AD76.2	AD76.2 1:1	33.707455	33.1731	0.7216854	7.6344895	11.923077	6.2599106	11.923077
30-Jan-12	rpoB on aTBassay2	AD89	AD89.4	AD89.4 1:1	27.550304	28.995832	1.2113615	492.83124	240.78381	175.7623	240.78381
30-Jan-12	rpoB on aTBassay2	AD89	AD89.4	AD89.4 1:1	29.749743	28.995832	1.2113615	111.215355	240.78381	175.7623	240.78381
30-Jan-12	rpoB on aTBassay2	AD89	AD89.1	AD89.1 1:1	30.68103	30.795412	0.1039272	59.21193	54.891876	3.9058845	54.891876
31-Jan-12	IS1081 on aTBassay3	AD1	AD1.1	AD1.1 1:1	Undetermined			0			0
31-Jan-12	IS1081 on aTBassay3	AD227	AD227.1	AD227.1 1:1	Undetermined			0			0
31-Jan-12	IS1081 on aTBassay3	AD33	AD33.1	AD33.1 1:1	Undetermined			0			0
31-Jan-12	IS1081 on aTBassay3	AD33	AD33.1	AD33.1 1:1	Undetermined			0			0
31-Jan-12	IS6110 on aTBassay3	AD227	AD227.2	AD227.2 1:1	Undetermined			0			0
31-Jan-12	IS6110 on aTBassay3	AD227	AD227.2	AD227.2 1:1	Undetermined			0			0
31-Jan-12	rpoB on aTBassay3	AD1	AD1.1	AD1.1 1:1	32.759636	32.345814	0.3961684	27.520443	36.741787	9.164526	36.741787
31-Jan-12	rpoB on aTBassay3	AD227	AD227.1	AD227.1 1:1	31.996664	31.883635	0.1286957	45.066246	48.595142	4.0894246	48.595142
31-Jan-12	rpoB on aTBassay3	AD82	AD82.4	AD82.4 1:1	33.976902	33.634228	0.3705407	12.529137	15.9412155	3.8796096	15.9412155
30-Mar-12	IS6110 on aTBassay4	AD256	AD256.1	AD256.1 1:1	Undetermined			0			0
30-Mar-12	IS6110 on aTBassay4	AD257	AD257.1	AD257.1 1:1	Undetermined			0			0
30-Mar-12	rpoB on aTBassay4	AD248	AD248.1	AD248.1 1:1	Undetermined			0			0
30-Mar-12	rpoB on aTBassay4	AD249	AD249.1	AD249.1 1:1	Undetermined			0			0
30-Mar-12	rpoB on aTBassay4	AD253	AD253.1	AD253.1 1:1	Undetermined			0			0
30-Mar-12	rpoB on aTBassay4	AD256	AD256.1	AD256.1 1:1	Undetermined			0			0
6-Apr-12	IS1081 on aTBassay5	AD256	AD256.3	AD256.3 1:1	Undetermined			0			0
6-Apr-12	IS1081 on aTBassay5	AD256	AD256.3	AD256.3 1:1	Undetermined			0			0
6-Apr-12	IS6110 on aTBassay5	AD256	AD256.2	AD256.2 1:1	Undetermined			0			0
6-Apr-12	IS6110 on aTBassay5	AD257	AD257.3	AD257.3 1:1	Undetermined			0			0
6-Apr-12	rpoB on aTBassay5	AD256	AD256.3	AD256.3 1:1	36.39482	36.536434	0.2875801	6.4434648	5.9264603	1.083536	5.9264603
6-Apr-12	rpoB on aTBassay5	AD256	AD256.3	AD256.3 1:1	36.86736	36.536434	0.2875801	4.681257	5.9264603	1.083536	5.9264603
6-Apr-12	rpoB on aTBassay5	AD257	AD257.3	AD257.3 1:1	39.381855	40.779903	1.3469353	0.8550686	0.43424007	0.37419754	0.43424007
6-Apr-12	rpoB on aTBassay5	AD257	AD257.1	AD257.1 1:1	39.355225	38.494007	1.0800674	0.8706043	1.878828	1.4465832	1.878828
4-Jun-12	rpoB on aTBassay7	AD229	sp2783	sp2783 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD230	sp3549	sp3549 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD231	#224	#224 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD232	sp8490	sp8490 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD232	sp8490	sp8490 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD233	sp8852	sp8852 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD234	sp10686	sp10686 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD234	#223	#223 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD235	sp12188	sp12188 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD236	sp13759	sp13759 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD236	sp13759	sp13759 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD236	sp13759	sp13759 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD237	sp17769	sp17769 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD238	sp19599	sp19599 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD238	sp19599	sp19599 1:1	Undetermined			0			0
4-Jun-12	rpoB on aTBassay7	AD240	sp22806	sp22806 1:1	Undetermined			0			0

4-Jun-12	rpoB on aTBassay7	AD240	sp22806	sp22806 1:1	Undetermined			0			0		
4-Jun-12	rpoB on aTBassay7	AD241	sp22829	sp22829 1:1	Undetermined			0			0		
4-Jun-12	rpoB on aTBassay7	AD242	sp23630	sp23630 1:1	Undetermined			0			0		
4-Jun-12	rpoB on aTBassay7	AD243	sp25714	sp25714 1:1	Undetermined			0			0		
4-Jun-12	rpoB on aTBassay7	AD244	sp27372	sp27372 1:1	Undetermined			0			0		
4-Jun-12	rpoB on aTBassay7	AD244	sp27372	sp27372 1:1	Undetermined			0			0		
4-Jun-12	rpoB on aTBassay7	AD245	sp29490	sp29490 1:1	Undetermined			0			0		
4-Jun-12	rpoB on aTBassay7	AD90	#221	#221 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD229	sp2783	sp2783 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD229	sp2783	sp2783 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD229	sp2783	sp2783 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD229	sp2783	sp2783 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD229	sp2783	sp2783 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD229	sp2783	sp2783 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD230	sp3549	sp3549 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD230	sp3549	sp3549 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD230	sp3549	sp3549 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD231	#224	#224 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD231	#224	#224 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD231	#224	#224 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD232	sp8490	sp8490 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD232	sp8490	sp8490 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD232	sp8490	sp8490 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD233	sp8852	sp8852 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD233	sp8852	sp8852 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD233	sp8852	sp8852 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD234	sp10686	sp10686 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD234	sp10686	sp10686 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD234	sp10686	sp10686 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD234	#223	#223 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD234	#223	#223 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD234	#223	#223 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD235	sp12188	sp12188 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD235	sp12188	sp12188 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD235	sp12188	sp12188 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD236	sp13759	sp13759 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD236	sp13759	sp13759 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD236	sp13759	sp13759 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD236	sp13759	sp13759 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD236	sp13759	sp13759 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD236	sp13759	sp13759 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD237	sp17769	sp17769 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD237	sp17769	sp17769 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD237	sp17769	sp17769 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD238	sp19599	sp19599 1:1	49.79908	47.404118	3.7937336	4.8637E-05	0.003028042	0.005880296	0.003028042	Yes	false-pos

4-Jun-12	IS1081 on aTBassay6	AD238	sp19599	sp19599 1:1	49.96428	47.404118	3.7937336	4.3371E-05	0.003028042	0.005880296	0.003028042	Yes	false-pos
4-Jun-12	IS1081 on aTBassay6	AD238	sp19599	sp19599 1:1	Undetermined	47.404118	3.7937336	0	0.003028042	0.005880296	0.003028042	Yes	false-pos
4-Jun-12	IS1081 on aTBassay6	AD238	sp19599	sp19599 1:1	47.976994	47.404118	3.7937336	0.00017213	0.003028042	0.005880296	0.003028042	Yes	false-pos
4-Jun-12	IS1081 on aTBassay6	AD238	sp19599	sp19599 1:1	41.876118	47.404118	3.7937336	0.01184803	0.003028042	0.005880296	0.003028042	Yes	false-pos
4-Jun-12	IS1081 on aTBassay6	AD238	sp19599	sp19599 1:1	Undetermined	47.404118	3.7937336	0	0.003028042	0.005880296	0.003028042	Yes	false-pos
4-Jun-12	IS1081 on aTBassay6	AD240	sp22806	sp22806 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD240	sp22806	sp22806 1:1	37.63528			0.22445966			0.22445966		
4-Jun-12	IS1081 on aTBassay6	AD240	sp22806	sp22806 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD241	sp22829	sp22829 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD241	sp22829	sp22829 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD241	sp22829	sp22829 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD242	sp23630	sp23630 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD242	sp23630	sp23630 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD242	sp23630	sp23630 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD243	sp25714	sp25714 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD243	sp25714	sp25714 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD243	sp25714	sp25714 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD244	sp27372	sp27372 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD244	sp27372	sp27372 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD244	sp27372	sp27372 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD245	sp29490	sp29490 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD245	sp29490	sp29490 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD245	sp29490	sp29490 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD246	sp30702	sp30702 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD246	sp30702	sp30702 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD246	sp30702	sp30702 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD90	#221	#221 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD90	#221	#221 1:1	Undetermined			0			0		
4-Jun-12	IS1081 on aTBassay6	AD90	#221	#221 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD105	AD105.1	AD105.1 1:1	48.68984	48.82946	0.1974549	0.00026452	0.000241972	3.1883E-05	0.000241972	Yes	
15-Jun-12	rpoB on aTBassay8	AD105	AD105.1	AD105.1 1:1	48.96908	48.82946	0.1974549	0.00021943	0.000241972	3.1883E-05	0.000241972	Yes	
15-Jun-12	rpoB on aTBassay8	AD105	AD105.1	AD105.1 1:1	Undetermined	48.82946	0.1974549	0	0.000241972	3.1883E-05	0.000241972	Yes	
15-Jun-12	rpoB on aTBassay8	AD106	AD106.1	AD106.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD106	AD106.1	AD106.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD106	AD106.1	AD106.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD127	AD127.1	AD127.1 1:1	37.331852	37.926563	0.7525219	0.5292624	0.38388428	0.16678853	0.38388428		
15-Jun-12	rpoB on aTBassay8	AD127	AD127.1	AD127.1 1:1	38.77258	37.926563	0.7525219	0.20179999	0.38388428	0.16678853	0.38388428		
15-Jun-12	rpoB on aTBassay8	AD127	AD127.1	AD127.1 1:1	37.67526	37.926563	0.7525219	0.42059034	0.38388428	0.16678853	0.38388428		
15-Jun-12	rpoB on aTBassay8	AD134	AD134.1	AD134.1 1:1	37.85269	37.8782	0.1178484	0.3734978	0.36793056	0.028640829	0.36793056		
15-Jun-12	rpoB on aTBassay8	AD134	AD134.1	AD134.1 1:1	38.006718	37.8782	0.1178484	0.33691484	0.36793056	0.028640829	0.36793056		
15-Jun-12	rpoB on aTBassay8	AD134	AD134.1	AD134.1 1:1	37.7752	37.8782	0.1178484	0.39337903	0.36793056	0.028640829	0.36793056		
15-Jun-12	rpoB on aTBassay8	AD145	AD145.1	AD145.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD145	AD145.1	AD145.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD145	AD145.1	AD145.1 1:1	44.30404			0.00497958			0.00497958		
15-Jun-12	rpoB on aTBassay8	AD151	AD151.1	AD151.1 1:1	Undetermined			0			0		

15-Jun-12	rpoB on aTBassay8	AD151	AD151.1	AD151.1 1:1	45.732662			0.00191408			0.001914083	Yes	
15-Jun-12	rpoB on aTBassay8	AD151	AD151.1	AD151.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD156	AD156.1	AD156.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD156	AD156.1	AD156.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD156	AD156.1	AD156.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD162	AD162.1	AD162.1 1:1	46.108604	43.44858	2.538108	0.00148831	0.018628908	0.022315273	0.018628908	Yes	barely crosses MC
15-Jun-12	rpoB on aTBassay8	AD162	AD162.1	AD162.1 1:1	43.184025	43.44858	2.538108	0.01053724	0.018628908	0.022315273	0.018628908	Yes	barely crosses MC
15-Jun-12	rpoB on aTBassay8	AD162	AD162.1	AD162.1 1:1	41.053112	43.44858	2.538108	0.04386118	0.018628908	0.022315273	0.018628908	Yes	barely crosses MC
15-Jun-12	rpoB on aTBassay8	AD216	AD216.1	AD216.1 1:1	38.806175	39.88925	0.9600745	0.19731332	0.11053572	0.075706206	0.11053572		
15-Jun-12	rpoB on aTBassay8	AD216	AD216.1	AD216.1 1:1	40.63561	39.88925	0.9600745	0.05800005	0.11053572	0.075706206	0.11053572		
15-Jun-12	rpoB on aTBassay8	AD216	AD216.1	AD216.1 1:1	40.225975	39.88925	0.9600745	0.07629377	0.11053572	0.075706206	0.11053572		
15-Jun-12	rpoB on aTBassay8	AD225	AD225.1	AD225.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD225	AD225.1	AD225.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD225	AD225.1	AD225.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD228	AD228.1	AD228.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD228	AD228.1	AD228.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD228	AD228.1	AD228.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD234	sp10686	sp10686 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD234	sp10686	sp10686 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD234	sp10686	sp10686 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD242	sp23630	sp23630 1:1	40.156033	43.15644	4.243221	0.07994991	0.040695462	0.055514168	0.040695462		
15-Jun-12	rpoB on aTBassay8	AD242	sp23630	sp23630 1:1	Undetermined	43.15644	4.243221	0	0.040695462	0.055514168	0.040695462		
15-Jun-12	rpoB on aTBassay8	AD242	sp23630	sp23630 1:1	46.156853	43.15644	4.243221	0.00144102	0.040695462	0.055514168	0.040695462		
15-Jun-12	rpoB on aTBassay8	AD251	AD251.1	AD251.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD251	AD251.1	AD251.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD251	AD251.1	AD251.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD252	AD252.1	AD252.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD252	AD252.1	AD252.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD252	AD252.1	AD252.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD282	AD282.1	AD282.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD282	AD282.1	AD282.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD282	AD282.1	AD282.1 1:1	Undetermined			0			0		
15-Jun-12	rpoB on aTBassay8	AD54	AD54.1	AD54.1 1:1	34.49956	34.45901	0.2203478	3.5227458	3.6462898	0.5466218	3.6462898		
15-Jun-12	rpoB on aTBassay8	AD54	AD54.1	AD54.1 1:1	34.221203	34.45901	0.2203478	4.2441106	3.6462898	0.5466218	3.6462898		
15-Jun-12	rpoB on aTBassay8	AD54	AD54.1	AD54.1 1:1	34.656265	34.45901	0.2203478	3.1720135	3.6462898	0.5466218	3.6462898		
15-Jun-12	rpoB on aTBassay8	AD82	AD82.8	AD82.8 1:1	37.052067	36.876575	0.1543203	0.6382504	0.7202995	0.07235576	0.7202995		
15-Jun-12	rpoB on aTBassay8	AD82	AD82.8	AD82.8 1:1	36.815628	36.876575	0.1543203	0.7476743	0.7202995	0.07235576	0.7202995		
15-Jun-12	rpoB on aTBassay8	AD82	AD82.8	AD82.8 1:1	36.762043	36.876575	0.1543203	0.7749738	0.7202995	0.07235576	0.7202995		
19-Jun-12	IS1081 on aTBassay9	AD105	AD105.1	AD105.1 1:1	Undetermined			0			0		
19-Jun-12	IS1081 on aTBassay9	AD105	AD105.1	AD105.1 1:1	Undetermined			0			0		
19-Jun-12	IS1081 on aTBassay9	AD105	AD105.1	AD105.1 1:1	Undetermined			0			0		
19-Jun-12	IS1081 on aTBassay9	AD106	AD106.1	AD106.1 1:1	Undetermined			0			0		

19-Jun-12	IS1081 on aTBassay9	AD106	AD106.1	AD106.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD106	AD106.1	AD106.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD127	AD127.1	AD127.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD127	AD127.1	AD127.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD127	AD127.1	AD127.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD134	AD134.1	AD134.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD134	AD134.1	AD134.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD134	AD134.1	AD134.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD145	AD145.1	AD145.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD145	AD145.1	AD145.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD145	AD145.1	AD145.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD151	AD151.1	AD151.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD151	AD151.1	AD151.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD151	AD151.1	AD151.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD156	AD156.1	AD156.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD156	AD156.1	AD156.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD156	AD156.1	AD156.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD162	AD162.1	AD162.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD162	AD162.1	AD162.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD162	AD162.1	AD162.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD216	AD216.1	AD216.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD216	AD216.1	AD216.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD216	AD216.1	AD216.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD225	AD225.1	AD225.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD225	AD225.1	AD225.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD225	AD225.1	AD225.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD227	AD227.3	AD227.3 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD227	AD227.3	AD227.3 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD227	AD227.3	AD227.3 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD228	AD228.1	AD228.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD228	AD228.1	AD228.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD228	AD228.1	AD228.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD240	sp22806	sp22806 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD240	sp22806	sp22806 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD240	sp22806	sp22806 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD242	sp23630	sp23630 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD242	sp23630	sp23630 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD242	sp23630	sp23630 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD251	AD251.1	AD251.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD251	AD251.1	AD251.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD251	AD251.1	AD251.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD252	AD252.1	AD252.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD252	AD252.1	AD252.1 1:1	Undetermined	0	0
19-Jun-12	IS1081 on aTBassay9	AD252	AD252.1	AD252.1 1:1	34.992153	0.27518287	0.27518287
19-Jun-12	IS1081 on aTBassay9	AD282	AD282.1	AD282.1 1:1	Undetermined	0	0

19-Jun-12	IS1081 on aTBassay9	AD282	AD282.1	AD282.1 1:1	Undetermined			0			0
19-Jun-12	IS1081 on aTBassay9	AD282	AD282.1	AD282.1 1:1	Undetermined			0			0
19-Jun-12	IS1081 on aTBassay9	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0
19-Jun-12	IS1081 on aTBassay9	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0
19-Jun-12	IS1081 on aTBassay9	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0
19-Jun-12	IS1081 on aTBassay9	AD54	AD54.1	AD54.1 1:1	30.334835	30.141047	0.175067	7.4695754	8.612762	1.041772	8.612762
19-Jun-12	IS1081 on aTBassay9	AD54	AD54.1	AD54.1 1:1	29.994324	30.141047	0.175067	9.508596	8.612762	1.041772	8.612762
19-Jun-12	IS1081 on aTBassay9	AD54	AD54.1	AD54.1 1:1	30.093979	30.141047	0.175067	8.860114	8.612762	1.041772	8.612762
19-Jun-12	IS1081 on aTBassay9	AD82	AD82.8	AD82.8 1:1	30.442732	30.459463	0.1229783	6.919612	6.8552704	0.59177434	6.8552704
19-Jun-12	IS1081 on aTBassay9	AD82	AD82.8	AD82.8 1:1	30.345705	30.459463	0.1229783	7.412245	6.8552704	0.59177434	6.8552704
19-Jun-12	IS1081 on aTBassay9	AD82	AD82.8	AD82.8 1:1	30.589949	30.459463	0.1229783	6.2339544	6.8552704	0.59177434	6.8552704
20-Jun-12	IS6110 on aTBassay10	AD105	AD105.1	AD105.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD105	AD105.1	AD105.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD105	AD105.1	AD105.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD106	AD106.1	AD106.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD106	AD106.1	AD106.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD106	AD106.1	AD106.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD127	AD127.1	AD127.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD127	AD127.1	AD127.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD127	AD127.1	AD127.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD134	AD134.1	AD134.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD134	AD134.1	AD134.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD134	AD134.1	AD134.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD145	AD145.1	AD145.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD145	AD145.1	AD145.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD145	AD145.1	AD145.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD151	AD151.1	AD151.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD151	AD151.1	AD151.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD151	AD151.1	AD151.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD156	AD156.1	AD156.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD156	AD156.1	AD156.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD156	AD156.1	AD156.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD162	AD162.1	AD162.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD162	AD162.1	AD162.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD162	AD162.1	AD162.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD216	AD216.1	AD216.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD216	AD216.1	AD216.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD216	AD216.1	AD216.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD225	AD225.1	AD225.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD225	AD225.1	AD225.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD225	AD225.1	AD225.1 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD227	AD227.3	AD227.3 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD227	AD227.3	AD227.3 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD227	AD227.3	AD227.3 1:1	Undetermined			0			0
20-Jun-12	IS6110 on aTBassay10	AD228	AD228.1	AD228.1 1:1	Undetermined			0			0

20-Jun-12	IS6110 on aTBassay10	AD228	AD228.1	AD228.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD228	AD228.1	AD228.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD234	sp10686	sp10686 1:1	44.110493	36.750534	6.6836977	0.00107958	3.0313756	4.7914166	3.0313756	Yes
20-Jun-12	IS6110 on aTBassay10	AD234	sp10686	sp10686 1:1	35.08179	36.750534	6.6836977	0.5377019	3.0313756	4.7914166	3.0313756	
20-Jun-12	IS6110 on aTBassay10	AD234	sp10686	sp10686 1:1	31.059319	36.750534	6.6836977	8.555346	3.0313756	4.7914166	3.0313756	Yes
20-Jun-12	IS6110 on aTBassay10	AD242	sp23630	sp23630 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD242	sp23630	sp23630 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD242	sp23630	sp23630 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD251	AD251.1	AD251.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD251	AD251.1	AD251.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD251	AD251.1	AD251.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD252	AD252.1	AD252.1 1:1	42.043194			0.00447562			0.004475617	
20-Jun-12	IS6110 on aTBassay10	AD252	AD252.1	AD252.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD252	AD252.1	AD252.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD282	AD282.1	AD282.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD282	AD282.1	AD282.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD282	AD282.1	AD282.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0	
20-Jun-12	IS6110 on aTBassay10	AD54	AD54.1	AD54.1 1:1	37.872177	39.07609	1.2607741	0.07887256	0.043314185	0.03288927	0.043314185	
20-Jun-12	IS6110 on aTBassay10	AD54	AD54.1	AD54.1 1:1	38.969185	39.07609	1.2607741	0.03708512	0.043314185	0.03288927	0.043314185	
20-Jun-12	IS6110 on aTBassay10	AD54	AD54.1	AD54.1 1:1	40.386917	39.07609	1.2607741	0.01398487	0.043314185	0.03288927	0.043314185	
20-Jun-12	IS6110 on aTBassay10	AD82	AD82.8	AD82.8 1:1	35.383278	35.3202	0.1566023	0.43699217	0.458169	0.050605245	0.458169	
20-Jun-12	IS6110 on aTBassay10	AD82	AD82.8	AD82.8 1:1	35.1419	35.3202	0.1566023	0.51592255	0.458169	0.050605245	0.458169	
20-Jun-12	IS6110 on aTBassay10	AD82	AD82.8	AD82.8 1:1	35.435432	35.3202	0.1566023	0.42159235	0.458169	0.050605245	0.458169	
29-Jun-12	IS1081 on aTBassay12	AD114	AD114.1	AD114.1 1:1	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD114	AD114.1	AD114.1 1:1	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD114	AD114.1	AD114.1 1:1	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD128	AD128.1	AD128.1 1:1	39.39968	38.75894	1.1177733	0.4015378	0.787961	0.67155045	0.787961	
29-Jun-12	IS1081 on aTBassay12	AD128	AD128.1	AD128.1 1:1	37.468258	38.75894	1.1177733	1.5633992	0.787961	0.67155045	0.787961	
29-Jun-12	IS1081 on aTBassay12	AD128	AD128.1	AD128.1 1:1	39.408882	38.75894	1.1177733	0.39894602	0.787961	0.67155045	0.787961	
29-Jun-12	IS1081 on aTBassay12	AD159	AD159.1	AD159.1 1:1	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD159	AD159.1	AD159.1 1:1	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD159	AD159.1	AD159.1 1:1	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD165	AD165.1	AD165.1 1:1	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD165	AD165.1	AD165.1 1:10	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD165	AD165.1	AD165.1 1:10	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD165	AD165.1	AD165.1 1:1	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD165	AD165.1	AD165.1 1:10	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD165	AD165.1	AD165.1 1:1	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD19	AD19.1	AD19.1 1:10	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD19	AD19.1	AD19.1 1:10	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD19	AD19.1	AD19.1 1:10	Undetermined			0			0	
29-Jun-12	IS1081 on aTBassay12	AD19	AD19.1	AD19.1 1:1	Undetermined			0			0	

29-Jun-12	IS1081 on aTBassay12	AD19	AD19.1	AD19.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD19	AD19.1	AD19.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD216	AD216.1	AD216.1 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD216	AD216.1	AD216.1 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD216	AD216.1	AD216.1 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD225	AD225.1B	AD225.1B 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD225	AD225.1	AD225.1 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD225	AD225.1	AD225.1 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD225	AD225.1	AD225.1 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD225	AD225.1B	AD225.1B 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD225	AD225.1B	AD225.1B 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD227	AD227.3	AD227.3 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD227	AD227.3	AD227.3 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD227	AD227.3B	AD227.3B 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD227	AD227.3B	AD227.3B 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD227	AD227.3B	AD227.3B 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD227	AD227.3	AD227.3 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD265	AD265.1	AD265.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD265	AD265.1	AD265.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD265	AD265.1	AD265.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD6	AD6.1	AD6.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD6	AD6.1	AD6.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD6	AD6.1	AD6.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD64	AD64.1	AD64.1 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD64	AD64.1	AD64.1 1:1	41.765427	42.33548	1.2015291	0.07596813	0.06172937	0.03744739	0.06172937
29-Jun-12	IS1081 on aTBassay12	AD64	AD64.1	AD64.1 1:1	41.525085	42.33548	1.2015291	0.08996887	0.06172937	0.03744739	0.06172937
29-Jun-12	IS1081 on aTBassay12	AD64	AD64.1	AD64.1 1:1	43.71593	42.33548	1.2015291	0.01925112	0.06172937	0.03744739	0.06172937
29-Jun-12	IS1081 on aTBassay12	AD64	AD64.1	AD64.1 1:10	40.33744			0.20753837			2.0753837
29-Jun-12	IS1081 on aTBassay12	AD64	AD64.1	AD64.1 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD81	AD81.1	AD81.1 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD81	AD81.1	AD81.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD81	AD81.1	AD81.1 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD81	AD81.1	AD81.1 1:1	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD81	AD81.1	AD81.1 1:10	Undetermined			0			0
29-Jun-12	IS1081 on aTBassay12	AD81	AD81.1	AD81.1 1:1	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD114	AD114.1	AD114.1 1:1	41.324253	40.711346	0.6599832	0.01150568	0.01884501	0.008709816	0.01884501
29-Jun-12	rpoB on aTBassay11	AD114	AD114.1	AD114.1 1:1	40.79711	40.711346	0.6599832	0.01655972	0.01884501	0.008709816	0.01884501
29-Jun-12	rpoB on aTBassay11	AD114	AD114.1	AD114.1 1:1	40.012672	40.711346	0.6599832	0.02846964	0.01884501	0.008709816	0.01884501
29-Jun-12	rpoB on aTBassay11	AD128	AD128.1	AD128.1 1:1	39.669426	39.361572	0.513361	0.03608732	0.046658073	0.017809216	0.046658073

29-Jun-12	rpoB on aTBassay11	AD128	AD128.1	AD128.1 1:1	39.646347	39.361572	0.513361	0.03666724	0.046658073	0.017809216	0.046658073
29-Jun-12	rpoB on aTBassay11	AD128	AD128.1	AD128.1 1:1	38.768944	39.361572	0.513361	0.06721966	0.046658073	0.017809216	0.046658073
29-Jun-12	rpoB on aTBassay11	AD159	AD159.1	AD159.1 1:1	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD159	AD159.1	AD159.1 1:1	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD159	AD159.1	AD159.1 1:1	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD165	AD165.1	AD165.1 1:10	45.83507	45.29688	1.2648888	0.00051013	0.000970985	0.00089936	0.009709851
29-Jun-12	rpoB on aTBassay11	AD165	AD165.1	AD165.1 1:1	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD165	AD165.1	AD165.1 1:10	46.203674	45.29688	1.2648888	0.00039546	0.000970985	0.00089936	0.009709851
29-Jun-12	rpoB on aTBassay11	AD165	AD165.1	AD165.1 1:10	43.851902	45.29688	1.2648888	0.00200736	0.000970985	0.00089936	0.009709851
29-Jun-12	rpoB on aTBassay11	AD165	AD165.1	AD165.1 1:1	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD165	AD165.1	AD165.1 1:1	48.0254			0.00011235			0.000112355
29-Jun-12	rpoB on aTBassay11	AD19	AD19.1	AD19.1 1:10	31.48521	31.756353	0.274449	10.294173	8.638258	1.6248996	86.38258
29-Jun-12	rpoB on aTBassay11	AD19	AD19.1	AD19.1 1:1	31.774508	31.41205	0.4206088	8.42953	11.144485	3.351286	11.144485
29-Jun-12	rpoB on aTBassay11	AD19	AD19.1	AD19.1 1:1	31.510788	31.41205	0.4206088	10.113893	11.144485	3.351286	11.144485
29-Jun-12	rpoB on aTBassay11	AD19	AD19.1	AD19.1 1:1	30.950857	31.41205	0.4206088	14.890034	11.144485	3.351286	11.144485
29-Jun-12	rpoB on aTBassay11	AD19	AD19.1	AD19.1 1:10	32.033993	31.756353	0.274449	7.0462604	8.638258	1.6248996	86.38258
29-Jun-12	rpoB on aTBassay11	AD19	AD19.1	AD19.1 1:10	31.74985	31.756353	0.274449	8.57434	8.638258	1.6248996	86.38258
29-Jun-12	rpoB on aTBassay11	AD216	AD216.1	AD216.1 1:10	33.57707	34.25308	0.621249	2.4268408	1.6203026	0.71972317	16.203026
29-Jun-12	rpoB on aTBassay11	AD216	AD216.1	AD216.1 1:10	34.79895	34.25308	0.621249	1.0434722	1.6203026	0.71972317	16.203026
29-Jun-12	rpoB on aTBassay11	AD216	AD216.1	AD216.1 1:10	34.383213	34.25308	0.621249	1.3905947	1.6203026	0.71972317	16.203026
29-Jun-12	rpoB on aTBassay11	AD225	AD225.1	AD225.1 1:10	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD225	AD225.1	AD225.1 1:10	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD225	AD225.1	AD225.1 1:10	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD225	AD225.1B	AD225.1B 1:10	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD225	AD225.1B	AD225.1B 1:10	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD225	AD225.1B	AD225.1B 1:10	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD227	AD227.3B	AD227.3B 1:10	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD227	AD227.3B	AD227.3B 1:10	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD227	AD227.3B	AD227.3B 1:10	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD227	AD227.3	AD227.3 1:1	36.2203	35.819473	1.4985716	0.3909139	0.74275684	0.76559854	0.74275684
29-Jun-12	rpoB on aTBassay11	AD227	AD227.3	AD227.3 1:1	37.07687	35.819473	1.4985716	0.21632798	0.74275684	0.76559854	0.74275684
29-Jun-12	rpoB on aTBassay11	AD227	AD227.3	AD227.3 1:1	34.161243	35.819473	1.4985716	1.6210287	0.74275684	0.76559854	0.74275684
29-Jun-12	rpoB on aTBassay11	AD265	AD265.1	AD265.1 1:1	39.642628	39.425858	1.1682674	0.03676157	0.053190973	0.043073528	0.053190973
29-Jun-12	rpoB on aTBassay11	AD265	AD265.1	AD265.1 1:1	40.47056	39.425858	1.1682674	0.02074999	0.053190973	0.043073528	0.053190973
29-Jun-12	rpoB on aTBassay11	AD265	AD265.1	AD265.1 1:1	38.164387	39.425858	1.1682674	0.10206136	0.053190973	0.043073528	0.053190973
29-Jun-12	rpoB on aTBassay11	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD52	AD52.1	AD52.1 1:1	Undetermined			0			0
29-Jun-12	rpoB on aTBassay11	AD6	AD6.1	AD6.1 1:1	34.69739	33.805622	0.9058427	1.1193042	2.3544767	1.4231514	2.3544767
29-Jun-12	rpoB on aTBassay11	AD6	AD6.1	AD6.1 1:1	33.833145	33.805622	0.9058427	2.0333824	2.3544767	1.4231514	2.3544767
29-Jun-12	rpoB on aTBassay11	AD6	AD6.1	AD6.1 1:1	32.886333	33.805622	0.9058427	3.9107437	2.3544767	1.4231514	2.3544767
29-Jun-12	rpoB on aTBassay11	AD64	AD64.1	AD64.1 1:10	42.732346	40.263214	3.4918818	0.00435	0.06808062	0.0901287	0.6808062
29-Jun-12	rpoB on aTBassay11	AD64	AD64.1	AD64.1 1:10	37.79408	40.263214	3.4918818	0.13181123	0.06808062	0.0901287	0.6808062
29-Jun-12	rpoB on aTBassay11	AD64	AD64.1	AD64.1 1:10	Undetermined	40.263214	3.4918818	0	0.06808062	0.0901287	0.6808062
29-Jun-12	rpoB on aTBassay11	AD64	AD64.1	AD64.1 1:1	38.503975	38.639812	1.0897253	0.0807209	0.08736817	0.057800703	0.08736817

29-Jun-12	rpoB on aTBassay11	AD64	AD64.1	AD64.1 1:1	39.79109	38.639812	1.0897253	0.03317848	0.08736817	0.057800703	0.08736817		
29-Jun-12	rpoB on aTBassay11	AD64	AD64.1	AD64.1 1:1	37.624374	38.639812	1.0897253	0.14820512	0.08736817	0.057800703	0.08736817		
29-Jun-12	rpoB on aTBassay11	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0		
29-Jun-12	rpoB on aTBassay11	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0		
29-Jun-12	rpoB on aTBassay11	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0		
29-Jun-12	rpoB on aTBassay11	AD81	AD81.1	AD81.1 1:10	46.50326			0.00032153			0.003215342		
29-Jun-12	rpoB on aTBassay11	AD81	AD81.1	AD81.1 1:10	Undetermined			0			0		
29-Jun-12	rpoB on aTBassay11	AD81	AD81.1	AD81.1 1:1	Undetermined			0			0		
29-Jun-12	rpoB on aTBassay11	AD81	AD81.1	AD81.1 1:1	Undetermined			0			0		
29-Jun-12	rpoB on aTBassay11	AD81	AD81.1	AD81.1 1:1	Undetermined			0			0		
29-Jun-12	rpoB on aTBassay11	AD81	AD81.1	AD81.1 1:10	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD11	AD11.1	AD11.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD11	AD11.1	AD11.1 1:1	35.97456			0.00767661			0.007676611		
5-Jul-12	IS6110 on aTBassay13	AD11	AD11.1	AD11.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD114	AD114.1	AD114.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD114	AD114.1	AD114.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD114	AD114.1	AD114.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD128	AD128.1	AD128.1 1:1	31.448997	31.179731	0.2331932	0.20825444	0.25581315	0.041187372	0.25581315		
5-Jul-12	IS6110 on aTBassay13	AD128	AD128.1	AD128.1 1:1	31.045877	31.179731	0.2331932	0.27943346	0.25581315	0.041187372	0.25581315		
5-Jul-12	IS6110 on aTBassay13	AD128	AD128.1	AD128.1 1:1	31.044317	31.179731	0.2331932	0.2797516	0.25581315	0.041187372	0.25581315		
5-Jul-12	IS6110 on aTBassay13	AD146	AD146.1	AD146.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD146	AD146.1	AD146.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD146	AD146.1	AD146.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD159	AD159.1	AD159.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD159	AD159.1	AD159.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD159	AD159.1	AD159.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD165	AD165.1	AD165.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD165	AD165.1	AD165.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD165	AD165.1	AD165.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD19	AD19.1	AD19.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD19	AD19.1	AD19.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD19	AD19.1	AD19.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD191	AD191.1	AD191.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD191	AD191.1	AD191.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD191	AD191.1	AD191.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD224	AD224.1	AD224.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD224	AD224.1	AD224.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD224	AD224.1	AD224.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD224	AD224.2	AD224.2 1:1	16.691792	18.318901	2.301079	9835.262	5375.79	6306.645	5375.79	Yes	nonsense
5-Jul-12	IS6110 on aTBassay13	AD224	AD224.2	AD224.2 1:1	19.946009	18.318901	2.301079	916.31885	5375.79	6306.645	5375.79	Yes	nonsense
5-Jul-12	IS6110 on aTBassay13	AD224	AD224.2	AD224.2 1:1	Undetermined	18.318901	2.301079	0	5375.79	6306.645	5375.79	Yes	nonsense
5-Jul-12	IS6110 on aTBassay13	AD256	AD256.4	AD256.4 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD256	AD256.4	AD256.4 1:1	Undetermined			0			0	Yes	
5-Jul-12	IS6110 on aTBassay13	AD256	AD256.4	AD256.4 1:1	47.819286			1.3598E-06			1.3598E-06	Yes	
5-Jul-12	IS6110 on aTBassay13	AD259	AD259.1	AD259.1 1:1	Undetermined			0			0		

5-Jul-12	IS6110 on aTBassay13	AD259	AD259.1	AD259.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD259	AD259.1	AD259.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD263	AD263.1	AD263.1 1:1	40.65094			0.0002535			0.000253498	Yes	
5-Jul-12	IS6110 on aTBassay13	AD263	AD263.1	AD263.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD263	AD263.1	AD263.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD264	AD264.1	AD264.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD264	AD264.1	AD264.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD264	AD264.1	AD264.1 1:1	31.803402			0.16082011			0.16082011	Yes	
5-Jul-12	IS6110 on aTBassay13	AD265	AD265.1	AD265.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD265	AD265.1	AD265.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD265	AD265.1	AD265.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD33	AD33.2	AD33.2 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD33	AD33.2	AD33.2 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD33	AD33.2	AD33.2 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD6	AD6.1	AD6.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD6	AD6.1	AD6.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD6	AD6.1	AD6.1 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD64	AD64.1	AD64.1 1:1	37.38414	38.54736	1.0354004	0.00274599	0.001435965	0.001142512	0.001435965		
5-Jul-12	IS6110 on aTBassay13	AD64	AD64.1	AD64.1 1:1	39.368233	38.54736	1.0354004	0.00064604	0.001435965	0.001142512	0.001435965		
5-Jul-12	IS6110 on aTBassay13	AD64	AD64.1	AD64.1 1:1	38.889698	38.54736	1.0354004	0.00091586	0.001435965	0.001142512	0.001435965		
5-Jul-12	IS6110 on aTBassay13	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0		
5-Jul-12	IS6110 on aTBassay13	AD81	AD81.1	AD81.1 1:1	47.831867	45.008987	3.9921522	1.3474E-06	4.20446E-05	5.75546E-05	4.20446E-05	Yes	nonsense, nonsense, sporadic
5-Jul-12	IS6110 on aTBassay13	AD81	AD81.1	AD81.1 1:1	Undetermined	45.008987	3.9921522	0	4.20446E-05	5.75546E-05	4.20446E-05	Yes	nonsense, nonsense, sporadic
5-Jul-12	IS6110 on aTBassay13	AD81	AD81.1	AD81.1 1:1	42.18611	45.008987	3.9921522	8.2742E-05	4.20446E-05	5.75546E-05	4.20446E-05	Yes	positive
6-Jul-12	IS1081 on aTBassay14	AD11	AD11.1	AD11.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD11	AD11.1	AD11.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD11	AD11.1	AD11.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD146	AD146.1	AD146.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD146	AD146.1	AD146.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD146	AD146.1	AD146.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD191	AD191.1	AD191.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD191	AD191.1	AD191.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD191	AD191.1	AD191.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD224	AD224.2	AD224.2 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD224	AD224.2	AD224.2 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD224	AD224.1	AD224.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD224	AD224.1	AD224.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD224	AD224.1	AD224.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD224	AD224.2	AD224.2 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD256	AD256.4	AD256.4 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD256	AD256.4	AD256.4 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD256	AD256.4	AD256.4 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD259	AD259.1	AD259.1 1:1	Undetermined			0			0		

6-Jul-12	IS1081 on aTBassay14	AD259	AD259.1	AD259.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD259	AD259.1	AD259.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD263	AD263.1	AD263.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD263	AD263.1	AD263.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD263	AD263.1	AD263.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD264	AD264.1	AD264.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD264	AD264.1	AD264.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD264	AD264.1	AD264.1 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD33	AD33.2	AD33.2 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD33	AD33.2	AD33.2 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD33	AD33.2	AD33.2 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0		
6-Jul-12	IS1081 on aTBassay14	AD69	AD69.2	AD69.2 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD11	AD11.1	AD11.1 1:1	45.137203	45.63615	0.6185693	0.00064259	0.000478116	0.000186825	0.000478116		
6-Jul-12	rpoB on aTBassay14	AD11	AD11.1	AD11.1 1:1	46.328247	45.63615	0.6185693	0.00027499	0.000478116	0.000186825	0.000478116		
6-Jul-12	rpoB on aTBassay14	AD11	AD11.1	AD11.1 1:1	45.442993	45.63615	0.6185693	0.00051677	0.000478116	0.000186825	0.000478116		
6-Jul-12	rpoB on aTBassay14	AD146	AD146.1	AD146.1 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD146	AD146.1	AD146.1 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD146	AD146.1	AD146.1 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD191	AD191.1	AD191.1 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD191	AD191.1	AD191.1 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD191	AD191.1	AD191.1 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD224	AD224.1	AD224.1 1:1	41.73798	41.862606	0.8017903	0.00724396	0.007337404	0.003785617	0.007337404		
6-Jul-12	rpoB on aTBassay14	AD224	AD224.1	AD224.1 1:1	41.13043	41.862606	0.8017903	0.01116888	0.007337404	0.003785617	0.007337404		
6-Jul-12	rpoB on aTBassay14	AD224	AD224.2	AD224.2 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD224	AD224.2	AD224.2 1:1	43.445587			0.00214527			0.002145265		
6-Jul-12	rpoB on aTBassay14	AD224	AD224.2	AD224.2 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD224	AD224.1	AD224.1 1:1	42.719414	41.862606	0.8017903	0.00359938	0.007337404	0.003785617	0.007337404		
6-Jul-12	rpoB on aTBassay14	AD256	AD256.4	AD256.4 1:1	35.427807	35.47415	0.1017575	0.6500384	0.6300087	0.04475208	0.6300087		
6-Jul-12	rpoB on aTBassay14	AD256	AD256.4	AD256.4 1:1	35.59083	35.47415	0.1017575	0.5787403	0.6300087	0.04475208	0.6300087		
6-Jul-12	rpoB on aTBassay14	AD256	AD256.4	AD256.4 1:1	35.403816	35.47415	0.1017575	0.66124743	0.6300087	0.04475208	0.6300087		
6-Jul-12	rpoB on aTBassay14	AD259	AD259.1	AD259.1 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD259	AD259.1	AD259.1 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD259	AD259.1	AD259.1 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD263	AD263.1	AD263.1 1:1	34.831837	35.136578	0.2646983	0.9940047	0.8098219	0.15984727	0.8098219		
6-Jul-12	rpoB on aTBassay14	AD263	AD263.1	AD263.1 1:1	35.268562	35.136578	0.2646983	0.7281558	0.8098219	0.15984727	0.8098219		
6-Jul-12	rpoB on aTBassay14	AD263	AD263.1	AD263.1 1:1	35.30933	35.136578	0.2646983	0.7073054	0.8098219	0.15984727	0.8098219		
6-Jul-12	rpoB on aTBassay14	AD264	AD264.1	AD264.1 1:1	34.669647	35.45527	1.0368942	1.115797	0.7443754	0.42829928	0.7443754		
6-Jul-12	rpoB on aTBassay14	AD264	AD264.1	AD264.1 1:1	35.065617	35.45527	1.0368942	0.84146273	0.7443754	0.42829928	0.7443754		
6-Jul-12	rpoB on aTBassay14	AD264	AD264.1	AD264.1 1:1	36.630543	35.45527	1.0368942	0.2758665	0.7443754	0.42829928	0.7443754		
6-Jul-12	rpoB on aTBassay14	AD33	AD33.2	AD33.2 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD33	AD33.2	AD33.2 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD33	AD33.2	AD33.2 1:1	Undetermined			0			0		
6-Jul-12	rpoB on aTBassay14	AD69	AD69.2	AD69.2 1:1	26.544273	17.588026	12.666046	365.04205	63845524	90290688	63845524	Yes	

6-Jul-12	rpoB on aTBassay14	AD69	AD69.2	AD69.2 1:1	Undetermined	17.588026	12.666046	0	63845524	90290688	63845524	Yes
6-Jul-12	rpoB on aTBassay14	AD69	AD69.2	AD69.2 1:1	8.631779	17.588026	12.666046	127690680	63845524	90290688	63845524	Yes
30-May-13	rpoB on aTBassay16	AD1	AD1.1	AD1.1 1:10	35.38221	35.6022	0.2266186	0.6769932	0.58616275	0.09102348	5.8616275	
30-May-13	rpoB on aTBassay16	AD1	AD1.1	AD1.1 1:10	35.589485	35.6022	0.2266186	0.5865477	0.58616275	0.09102348	5.8616275	
30-May-13	rpoB on aTBassay16	AD1	AD1.1	AD1.1 1:10	35.83491	35.6022	0.2266186	0.49494743	0.58616275	0.09102348	5.8616275	
30-May-13	rpoB on aTBassay16	AD146	AD146.1	AD146.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD146	AD146.1	AD146.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD146	AD146.1	AD146.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD165	AD165.1	AD165.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD165	AD165.1	AD165.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD165	AD165.1	AD165.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD227	AD227.1	AD227.1 1:10	32.945393	32.980663	0.0913476	3.6541417	3.5707567	0.22202504	35.707567	
30-May-13	rpoB on aTBassay16	AD227	AD227.1	AD227.1 1:10	33.084385	32.980663	0.0913476	3.319111	3.5707567	0.22202504	35.707567	
30-May-13	rpoB on aTBassay16	AD227	AD227.2	AD227.2 1:10	33.650192	33.740498	0.1603685	2.2439396	2.1165006	0.22687653	21.165006	
30-May-13	rpoB on aTBassay16	AD227	AD227.2	AD227.2 1:10	33.64565	33.740498	0.1603685	2.2510042	2.1165006	0.22687653	21.165006	
30-May-13	rpoB on aTBassay16	AD227	AD227.2	AD227.2 1:10	33.92566	33.740498	0.1603685	1.8545579	2.1165006	0.22687653	21.165006	
30-May-13	rpoB on aTBassay16	AD227	AD227.1	AD227.1 1:10	32.912205	32.980663	0.0913476	3.7390175	3.5707567	0.22202504	35.707567	
30-May-13	rpoB on aTBassay16	AD251	AD251.1	AD251.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD251	AD251.1	AD251.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD251	AD251.1	AD251.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD259	AD259.1	AD259.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD259	AD259.1	AD259.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD259	AD259.1	AD259.1 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD33	AD33.2	AD33.2 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD33	AD33.2	AD33.2 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD33	AD33.2	AD33.2 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD69	AD69.2	AD69.2 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD69	AD69.2	AD69.2 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD69	AD69.2	AD69.2 1:10	Undetermined			0			0	
30-May-13	rpoB on aTBassay16	AD82	AD82.5	AD82.5 1:10	34.695602	34.966076	0.3444077	1.0886599	0.91941804	0.20572601	9.1941804	
30-May-13	rpoB on aTBassay16	AD82	AD82.5	AD82.5 1:10	35.353798	34.966076	0.3444077	0.69043267	0.91941804	0.20572601	9.1941804	
30-May-13	rpoB on aTBassay16	AD82	AD82.5	AD82.5 1:10	34.84882	34.966076	0.3444077	0.9791615	0.91941804	0.20572601	9.1941804	
30-May-13	rpoB on aTBassay16	AD89	AD89.1	AD89.1 1:10	33.445156	33.472523	0.1330446	2.585945	2.5445497	0.23078118	25.445497	
30-May-13	rpoB on aTBassay16	AD89	AD89.1	AD89.1 1:10	33.617123	33.472523	0.1330446	2.2958722	2.5445497	0.23078118	25.445497	
30-May-13	rpoB on aTBassay16	AD89	AD89.1	AD89.1 1:10	33.35529	33.472523	0.1330446	2.7518318	2.5445497	0.23078118	25.445497	
30-May-13	rpoB2 on aTBassay16	AD1	AD1.1	AD1.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD1	AD1.1	AD1.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD1	AD1.1	AD1.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD146	AD146.1	AD146.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD146	AD146.1	AD146.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD146	AD146.1	AD146.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD165	AD165.1	AD165.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD165	AD165.1	AD165.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD165	AD165.1	AD165.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD227	AD227.1	AD227.1 1:10	Undetermined			0			0	

30-May-13	rpoB2 on aTBassay16	AD227	AD227.1	AD227.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD227	AD227.2	AD227.2 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD227	AD227.2	AD227.2 1:10	41.95341			0.00690229			0.069022877	
30-May-13	rpoB2 on aTBassay16	AD227	AD227.2	AD227.2 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD227	AD227.1	AD227.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD251	AD251.1	AD251.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD251	AD251.1	AD251.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD251	AD251.1	AD251.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD259	AD259.1	AD259.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD259	AD259.1	AD259.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD259	AD259.1	AD259.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD33	AD33.2	AD33.2 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD33	AD33.2	AD33.2 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD33	AD33.2	AD33.2 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD69	AD69.2	AD69.2 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD69	AD69.2	AD69.2 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD69	AD69.2	AD69.2 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD82	AD82.5	AD82.5 1:10	32.32157	32.860394	0.4910615	6.101723	4.349337	1.5627154	43.49337	
30-May-13	rpoB2 on aTBassay16	AD82	AD82.5	AD82.5 1:10	33.282757	32.860394	0.4910615	3.1004014	4.349337	1.5627154	43.49337	
30-May-13	rpoB2 on aTBassay16	AD82	AD82.5	AD82.5 1:10	32.976852	32.860394	0.4910615	3.8458877	4.349337	1.5627154	43.49337	
30-May-13	rpoB2 on aTBassay16	AD89	AD89.1	AD89.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD89	AD89.1	AD89.1 1:10	Undetermined			0			0	
30-May-13	rpoB2 on aTBassay16	AD89	AD89.1	AD89.1 1:10	Undetermined			0			0	
4-Jun-13	rpoB on aTBassay17	AD1	AD1.2	AD1.2 1:10	30.21398	30.208202	0.0081704	16.317219	16.380526	0.08952935	163.80526	
4-Jun-13	rpoB on aTBassay17	AD1	AD1.2	AD1.2 1:10	30.202425	30.208202	0.0081704	16.443832	16.380526	0.08952935	163.80526	
4-Jun-13	rpoB on aTBassay17	AD105	AD105.1	AD105.1 1:10	40.01338	40.160263	0.207724	0.02321076	0.021140214	0.002928185	0.21140214	
4-Jun-13	rpoB on aTBassay17	AD105	AD105.1	AD105.1 1:10	40.307144	40.160263	0.207724	0.01906968	0.021140214	0.002928185	0.21140214	
4-Jun-13	rpoB on aTBassay17	AD106	AD106.1	AD106.1 1:10	43.28223			0.00260622			0.02606216	Yes
4-Jun-13	rpoB on aTBassay17	AD106	AD106.1	AD106.1 1:10	Undetermined			0			0	
4-Jun-13	rpoB on aTBassay17	AD11	AD11.1	AD11.1 1:10	35.18763	35.215584	0.0395304	0.5857108	0.57496107	0.015202494	5.7496107	
4-Jun-13	rpoB on aTBassay17	AD11	AD11.1	AD11.1 1:10	35.243534	35.215584	0.0395304	0.56421125	0.57496107	0.015202494	5.7496107	
4-Jun-13	rpoB on aTBassay17	AD114	AD114.1	AD114.1 1:10	36.763386	36.846596	0.1176767	0.20412157	0.19336903	0.015206404	1.9336903	
4-Jun-13	rpoB on aTBassay17	AD114	AD114.1	AD114.1 1:10	36.929806	36.846596	0.1176767	0.18261647	0.19336903	0.015206404	1.9336903	
4-Jun-13	rpoB on aTBassay17	AD127	AD127.1	AD127.1 1:10	31.281073	31.51407	0.3295088	7.9915395	6.921388	1.5134225	69.21388	
4-Jun-13	rpoB on aTBassay17	AD127	AD127.1	AD127.1 1:10	31.747068	31.51407	0.3295088	5.851237	6.921388	1.5134225	69.21388	
4-Jun-13	rpoB on aTBassay17	AD128	AD128.1	AD128.1 1:10	35.55193	35.370827	0.256118	0.45903474	0.52196276	0.088993706	5.2196276	
4-Jun-13	rpoB on aTBassay17	AD128	AD128.1	AD128.1 1:10	35.189724	35.370827	0.256118	0.58489084	0.52196276	0.088993706	5.2196276	
4-Jun-13	rpoB on aTBassay17	AD134	AD134.1	AD134.1 1:10	32.88526	33.00648	0.1714305	2.7325974	2.5280452	0.28928062	25.280452	
4-Jun-13	rpoB on aTBassay17	AD134	AD134.1	AD134.1 1:10	33.1277	33.00648	0.1714305	2.3234928	2.5280452	0.28928062	25.280452	
4-Jun-13	rpoB on aTBassay17	AD145	AD145.1	AD145.1 1:10	40.2732	39.78176	0.6949982	0.01950764	0.028578307	0.012827868	0.28578307	
4-Jun-13	rpoB on aTBassay17	AD145	AD145.1	AD145.1 1:10	39.290325	39.78176	0.6949982	0.03764898	0.028578307	0.012827868	0.28578307	
4-Jun-13	rpoB on aTBassay17	AD227	AD227.3	AD227.3 1:10	30.95884	30.983559	0.0349569	9.913952	9.752701	0.22804277	97.52701	
4-Jun-13	rpoB on aTBassay17	AD227	AD227.3	AD227.3 1:10	31.008276	30.983559	0.0349569	9.591451	9.752701	0.22804277	97.52701	
4-Jun-13	rpoB on aTBassay17	AD263	AD263.1	AD263.1 1:10	31.966204	31.910252	0.0791295	5.0533996	5.249809	0.2777646	52.49809	
4-Jun-13	rpoB on aTBassay17	AD263	AD263.1	AD263.1 1:10	31.854298	31.910252	0.0791295	5.446218	5.249809	0.2777646	52.49809	

4-Jun-13	rpoB on aTBassay17	AD264	AD264.1	AD264.1 1:10	32.25419	32.17816	0.1075183	4.1678996	4.391029	0.31555244	43.91029		
4-Jun-13	rpoB on aTBassay17	AD264	AD264.1	AD264.1 1:10	32.102135	32.17816	0.1075183	4.614158	4.391029	0.31555244	43.91029		
4-Jun-13	rpoB on aTBassay17	AD52	AD52.1	AD52.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB on aTBassay17	AD52	AD52.1	AD52.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB on aTBassay17	AD54	AD54.1	AD54.1 1:10	32.48188	32.905262	0.5987523	3.5790474	2.805147	1.0944606	28.05147		
4-Jun-13	rpoB on aTBassay17	AD54	AD54.1	AD54.1 1:10	33.328644	32.905262	0.5987523	2.0312464	2.805147	1.0944606	28.05147		
4-Jun-13	rpoB on aTBassay17	AD82	AD82.5	AD82.5 1:1	28.336628	28.210512	0.1406128	57.28767	62.516018	5.9460297	62.516018		
4-Jun-13	rpoB on aTBassay17	AD82	AD82.5	AD82.5 1:1000	38.636055	40.861603	3.1474	0.05832224	0.03064588	0.039140288	30.64588		
4-Jun-13	rpoB on aTBassay17	AD82	AD82.5	AD82.5 1:1000	43.08715	40.861603	3.1474	0.00296952	0.03064588	0.039140288	30.64588		
4-Jun-13	rpoB on aTBassay17	AD82	AD82.5	AD82.5 1:1000	Undetermined	40.861603	3.1474	0	0.03064588	0.039140288	30.64588		
4-Jun-13	rpoB on aTBassay17	AD82	AD82.5	AD82.5 1:100	36.415077	35.442463	0.8633469	0.25768003	0.5454288	0.26390845	54.54288		
4-Jun-13	rpoB on aTBassay17	AD82	AD82.5	AD82.5 1:100	35.145573	35.442463	0.8633469	0.60242337	0.5454288	0.26390845	54.54288		
4-Jun-13	rpoB on aTBassay17	AD82	AD82.5	AD82.5 1:100	34.76673	35.442463	0.8633469	0.7761832	0.5454288	0.26390845	54.54288		
4-Jun-13	rpoB on aTBassay17	AD82	AD82.5	AD82.5 1:1	28.058893	28.210512	0.1406128	68.98422	62.516018	5.9460297	62.516018		
4-Jun-13	rpoB on aTBassay17	AD82	AD82.5	AD82.5 1:1	28.236015	28.210512	0.1406128	61.276157	62.516018	5.9460297	62.516018		
4-Jun-13	rpoB on aTBassay17	AD82	AD82.3	AD82.3 1:10	29.957504	30.229752	0.3850159	19.371283	16.414474	4.181557	164.14474		
4-Jun-13	rpoB on aTBassay17	AD82	AD82.3	AD82.3 1:10	30.501999	30.229752	0.3850159	13.457667	16.414474	4.181557	164.14474		
4-Jun-13	rpoB2 on aTBassay17	AD1	AD1.2	AD1.2 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD1	AD1.2	AD1.2 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD105	AD105.1	AD105.1 1:10	38.243378			0.08446891			0.8446891		
4-Jun-13	rpoB2 on aTBassay17	AD105	AD105.1	AD105.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD106	AD106.1	AD106.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD106	AD106.1	AD106.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD11	AD11.1	AD11.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD11	AD11.1	AD11.1 1:10	44.328823			0.00133665			0.013366452	Yes	false-pos
4-Jun-13	rpoB2 on aTBassay17	AD114	AD114.1	AD114.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD114	AD114.1	AD114.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD127	AD127.1	AD127.1 1:10	40.353428	39.653374	0.9900261	0.02006014	0.03606734	0.0226376	0.3606734		
4-Jun-13	rpoB2 on aTBassay17	AD127	AD127.1	AD127.1 1:10	38.95332	39.653374	0.9900261	0.05207454	0.03606734	0.0226376	0.3606734		
4-Jun-13	rpoB2 on aTBassay17	AD128	AD128.1	AD128.1 1:10	41.776337	42.933643	1.6366788	0.00760845	0.004590133	0.004268539	0.045901327		
4-Jun-13	rpoB2 on aTBassay17	AD128	AD128.1	AD128.1 1:10	44.09095	42.933643	1.6366788	0.00157182	0.004590133	0.004268539	0.045901327	Yes	false-pos
4-Jun-13	rpoB2 on aTBassay17	AD134	AD134.1	AD134.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD134	AD134.1	AD134.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD145	AD145.1	AD145.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD145	AD145.1	AD145.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD227	AD227.3	AD227.3 1:10	44.448196			0.00123224			0.012322355	Yes	false-pos
4-Jun-13	rpoB2 on aTBassay17	AD227	AD227.3	AD227.3 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD263	AD263.1	AD263.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD263	AD263.1	AD263.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD264	AD264.1	AD264.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD264	AD264.1	AD264.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD52	AD52.1	AD52.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD52	AD52.1	AD52.1 1:10	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD54	AD54.1	AD54.1 1:10	33.656574	33.629288	0.038589	1.9227823	1.959202	0.051505387	19.59202		
4-Jun-13	rpoB2 on aTBassay17	AD54	AD54.1	AD54.1 1:10	33.602	33.629288	0.038589	1.9956219	1.959202	0.051505387	19.59202		

4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.5	AD82.5 1:10	28.170084	30.268673	1.7935151	80.794205	33.66301	34.95918	336.6301		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.5	AD82.5 1:1000	36.985107			0.19907585			199.07585		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.5	AD82.5 1:1000	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.5	AD82.5 1:1000	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.5	AD82.5 1:100	34.97219	35.329556	0.4477401	0.7845845	0.63340664	0.17828679	63.340664		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.5	AD82.5 1:100	35.831795	35.329556	0.4477401	0.4368007	0.63340664	0.17828679	63.340664		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.5	AD82.5 1:10	30.816189	30.268673	1.7935151	13.3167715	33.66301	34.95918	336.6301		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.5	AD82.5 1:10	28.573572	30.268673	1.7935151	61.374546	33.66301	34.95918	336.6301		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.5	AD82.5 1:10	31.993553	30.268673	1.7935151	5.9705915	33.66301	34.95918	336.6301		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.5	AD82.5 1:10	31.789974	30.268673	1.7935151	6.858928	33.66301	34.95918	336.6301		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.3	AD82.3 1:1	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.3	AD82.3 1:1	Undetermined			0			0		
4-Jun-13	rpoB2 on aTBassay17	AD82	AD82.5	AD82.5 1:100	35.18468	35.329556	0.4477401	0.6788347	0.63340664	0.17828679	63.340664		
5-Jun-13	rpoB on aTBassay18	AD216	AD216.2	AD216.2 1:10	34.7916	35.58311	0.7879728	0.27564484	0.17578085	0.09255814	1.7578085		
5-Jun-13	rpoB on aTBassay18	AD216	AD216.2	AD216.2 1:100	36.10375	36.961315	0.9700447	0.11141992	0.07058489	0.040809773	7.058489		
5-Jun-13	rpoB on aTBassay18	AD216	AD216.2	AD216.2 1:10	36.367496	35.58311	0.7879728	0.09287314	0.17578085	0.09255814	1.7578085		
5-Jun-13	rpoB on aTBassay18	AD216	AD216.2	AD216.2 1:10	35.59023	35.58311	0.7879728	0.15882461	0.17578085	0.09255814	1.7578085		
5-Jun-13	rpoB on aTBassay18	AD216	AD216.2	AD216.2 1:100	38.014133	36.961315	0.9700447	0.02980042	0.07058489	0.040809773	7.058489		
5-Jun-13	rpoB on aTBassay18	AD216	AD216.2	AD216.2 1:1	31.914865	32.356174	0.6241065	2.0081918	1.5500581	0.647899	1.5500581		
5-Jun-13	rpoB on aTBassay18	AD216	AD216.2	AD216.2 1:100	36.766056	36.961315	0.9700447	0.07053434	0.07058489	0.040809773	7.058489		
5-Jun-13	rpoB on aTBassay18	AD216	AD216.2	AD216.2 1:1	32.797485	32.356174	0.6241065	1.0919243	1.5500581	0.647899	1.5500581		
5-Jun-13	rpoB on aTBassay18	AD248	AD248.3	AD248.3 1:10	44.970135	39.467175	4.9217978	0.00024479	0.06738914	0.090780616	0.6738914	Yes	barely crosses MC
5-Jun-13	rpoB on aTBassay18	AD248	AD248.3	AD248.3 1:100	Undetermined	38.17028	3.6575832	0	0.08199966	0.10961842	8.199966		
5-Jun-13	rpoB on aTBassay18	AD248	AD248.3	AD248.3 1:1	35.602985	35.51962	0.1178979	0.15743212	0.16703433	0.013579557	0.16703433		
5-Jun-13	rpoB on aTBassay18	AD248	AD248.3	AD248.3 1:10	37.94539	39.467175	4.9217978	0.03124872	0.06738914	0.090780616	0.6738914	Yes	barely crosses MC
5-Jun-13	rpoB on aTBassay18	AD248	AD248.3	AD248.3 1:10	35.485996	39.467175	4.9217978	0.17067392	0.06738914	0.090780616	0.6738914		
5-Jun-13	rpoB on aTBassay18	AD248	AD248.3	AD248.3 1:100	35.583977	38.17028	3.6575832	0.15951158	0.08199966	0.10961842	8.199966		
5-Jun-13	rpoB on aTBassay18	AD248	AD248.3	AD248.3 1:100	40.75658	38.17028	3.6575832	0.00448773	0.08199966	0.10961842	8.199966	Yes	barely crosses MC
5-Jun-13	rpoB on aTBassay18	AD248	AD248.3	AD248.3 1:1	35.436253	35.51962	0.1178979	0.17663652	0.16703433	0.013579557	0.16703433	Yes	barely crosses MC
5-Jun-13	rpoB on aTBassay18	AD249	AD249.2	AD249.2 1:100	Undetermined	38.55922	5.2419457	0	0.13293047	0.18575303	13.293047		
5-Jun-13	rpoB on aTBassay18	AD249	AD249.2	AD249.2 1:1	37.007393	38.27552	1.7934058	0.05970997	0.035038657	0.03489051	0.035038657	Yes	
5-Jun-13	rpoB on aTBassay18	AD249	AD249.2	AD249.2 1:100	42.265835	38.55922	5.2419457	0.00158324	0.13293047	0.18575303	13.293047	Yes	
5-Jun-13	rpoB on aTBassay18	AD249	AD249.2	AD249.2 1:100	34.852604	38.55922	5.2419457	0.2642777	0.13293047	0.18575303	13.293047		
5-Jun-13	rpoB on aTBassay18	AD249	AD249.2	AD249.2 1:10	37.56033	36.50885	1.5085266	0.04076379	0.123777755	0.13350402	1.23777755	Yes	
5-Jun-13	rpoB on aTBassay18	AD249	AD249.2	AD249.2 1:10	37.1858	36.50885	1.5085266	0.05279103	0.123777755	0.13350402	1.23777755	Yes	
5-Jun-13	rpoB on aTBassay18	AD249	AD249.2	AD249.2 1:10	34.78043	36.50885	1.5085266	0.27777845	0.123777755	0.13350402	1.23777755		
5-Jun-13	rpoB on aTBassay18	AD249	AD249.2	AD249.2 1:1	39.54365	38.27552	1.7934058	0.01036734	0.035038657	0.03489051	0.035038657	Yes	
5-Jun-13	rpoB on aTBassay18	AD250	AD250.2	AD250.2 1:10	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD250	AD250.2	AD250.2 1:1	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD250	AD250.2	AD250.2 1:1	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD250	AD250.2	AD250.2 1:10	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD250	AD250.2	AD250.2 1:100	Undetermined			0			0		

5-Jun-13	rpoB on aTBassay18	AD250	AD250.2	AD250.2 1:100	44.34096			0.00037793			0.03779342		
5-Jun-13	rpoB on aTBassay18	AD250	AD250.2	AD250.2 1:10	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD250	AD250.2	AD250.2 1:100	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD251	AD251.2	AD251.2 1:100	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD251	AD251.2	AD251.2 1:100	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD251	AD251.2	AD251.2 1:10	35.58895	38.80198	4.543905	0.15896478	0.08042367	0.11107391	0.8042367		
5-Jun-13	rpoB on aTBassay18	AD251	AD251.2	AD251.2 1:10	Undetermined	38.80198	4.543905	0	0.08042367	0.11107391	0.8042367		
5-Jun-13	rpoB on aTBassay18	AD251	AD251.2	AD251.2 1:10	42.015003	38.80198	4.543905	0.00188255	0.08042367	0.11107391	0.8042367		
5-Jun-13	rpoB on aTBassay18	AD251	AD251.2	AD251.2 1:100	39.94834			0.00784043			0.78404285		
5-Jun-13	rpoB on aTBassay18	AD251	AD251.2	AD251.2 1:1	37.99765	35.5524	3.4581077	0.03014145	0.45595518	0.60219157	0.45595518	Yes	positive
5-Jun-13	rpoB on aTBassay18	AD251	AD251.2	AD251.2 1:1	33.107147	35.5524	3.4581077	0.8817689	0.45595518	0.60219157	0.45595518	Yes	positive
5-Jun-13	rpoB on aTBassay18	AD252	AD252.3	AD252.3 1:100	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD252	AD252.3	AD252.3 1:100	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD252	AD252.3	AD252.3 1:10	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD252	AD252.3	AD252.3 1:10	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD252	AD252.3	AD252.3 1:10	36.53885			0.08251219			0.82512185		
5-Jun-13	rpoB on aTBassay18	AD252	AD252.3	AD252.3 1:1	Undetermined			0			0	Yes	positive
5-Jun-13	rpoB on aTBassay18	AD252	AD252.3	AD252.3 1:1	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD252	AD252.3	AD252.3 1:100	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD255	AD255.1	AD255.1 1:10	36.040184			0.11641787			1.1641787		
5-Jun-13	rpoB on aTBassay18	AD255	AD255.1	AD255.1 1:1	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD255	AD255.1	AD255.1 1:100	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD255	AD255.1	AD255.1 1:100	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD255	AD255.1	AD255.1 1:100	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD255	AD255.1	AD255.1 1:10	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD255	AD255.1	AD255.1 1:10	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD255	AD255.1	AD255.1 1:1	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD257	AD257.4	AD257.4 1:1	34.659157	35.566925	1.2837754	0.30203456	0.19414172	0.15258354	0.19414172		
5-Jun-13	rpoB on aTBassay18	AD257	AD257.4	AD257.4 1:100	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD257	AD257.4	AD257.4 1:100	38.581615			0.02014138			2.0141378		
5-Jun-13	rpoB on aTBassay18	AD257	AD257.4	AD257.4 1:100	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD257	AD257.4	AD257.4 1:10	40.5536	39.061882	2.109611	0.00516275	0.022825422	0.02497879	0.22825422		
5-Jun-13	rpoB on aTBassay18	AD257	AD257.4	AD257.4 1:10	37.57016	39.061882	2.109611	0.04048809	0.022825422	0.02497879	0.22825422		
5-Jun-13	rpoB on aTBassay18	AD257	AD257.4	AD257.4 1:1	36.47469	35.566925	1.2837754	0.08624886	0.19414172	0.15258354	0.19414172		
5-Jun-13	rpoB on aTBassay18	AD257	AD257.4	AD257.4 1:10	Undetermined	39.061882	2.109611	0	0.022825422	0.02497879	0.22825422		
5-Jun-13	rpoB on aTBassay18	AD258	AD258.1	AD258.1 1:1	40.337753	39.4561	1.2468454	0.00599229	0.013116789	0.010075566	0.013116789	Yes	
5-Jun-13	rpoB on aTBassay18	AD258	AD258.1	AD258.1 1:100	39.210037	40.905045	2.397105	0.01305227	0.007154651	0.008340494	0.7154651		
5-Jun-13	rpoB on aTBassay18	AD258	AD258.1	AD258.1 1:100	42.600056	40.905045	2.397105	0.00125703	0.007154651	0.008340494	0.7154651	Yes	
5-Jun-13	rpoB on aTBassay18	AD258	AD258.1	AD258.1 1:100	Undetermined	40.905045	2.397105	0	0.007154651	0.008340494	0.7154651		
5-Jun-13	rpoB on aTBassay18	AD258	AD258.1	AD258.1 1:10	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD258	AD258.1	AD258.1 1:10	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD258	AD258.1	AD258.1 1:10	Undetermined			0			0		
5-Jun-13	rpoB on aTBassay18	AD258	AD258.1	AD258.1 1:1	38.574448	39.4561	1.2468454	0.02024129	0.013116789	0.010075566	0.013116789	Yes	
5-Jun-13	rpoB2 on aTBassay18	AD216	AD216.2	AD216.2 1:10	40.8139			0.00837253			0.08372532		
5-Jun-13	rpoB2 on aTBassay18	AD216	AD216.2	AD216.2 1:1	42.93358	42.68216	0.3555576	0.00192922	0.002331004	0.000568204	0.002331004		

5-Jun-13	rpoB2 on aTBassay18	AD216	AD216.2	AD216.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD216	AD216.2	AD216.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD216	AD216.2	AD216.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD216	AD216.2	AD216.2 1:1	42.430744	42.68216	0.3555576	0.00273279	0.002331004	0.000568204	0.002331004
5-Jun-13	rpoB2 on aTBassay18	AD216	AD216.2	AD216.2 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD216	AD216.2	AD216.2 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD248	AD248.3	AD248.3 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD248	AD248.3	AD248.3 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD248	AD248.3	AD248.3 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD248	AD248.3	AD248.3 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD248	AD248.3	AD248.3 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD248	AD248.3	AD248.3 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD248	AD248.3	AD248.3 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD248	AD248.3	AD248.3 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD249	AD249.2	AD249.2 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD249	AD249.2	AD249.2 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD249	AD249.2	AD249.2 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD249	AD249.2	AD249.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD249	AD249.2	AD249.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD249	AD249.2	AD249.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD249	AD249.2	AD249.2 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD249	AD249.2	AD249.2 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD250	AD250.2	AD250.2 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD250	AD250.2	AD250.2 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD250	AD250.2	AD250.2 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD250	AD250.2	AD250.2 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD250	AD250.2	AD250.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD250	AD250.2	AD250.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD250	AD250.2	AD250.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD250	AD250.2	AD250.2 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD251	AD251.2	AD251.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD251	AD251.2	AD251.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD251	AD251.2	AD251.2 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD251	AD251.2	AD251.2 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD251	AD251.2	AD251.2 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD251	AD251.2	AD251.2 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD251	AD251.2	AD251.2 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD251	AD251.2	AD251.2 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD252	AD252.3	AD252.3 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD252	AD252.3	AD252.3 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD252	AD252.3	AD252.3 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD252	AD252.3	AD252.3 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD252	AD252.3	AD252.3 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD252	AD252.3	AD252.3 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD252	AD252.3	AD252.3 1:100	Undetermined			0			0

5-Jun-13	rpoB2 on aTBassay18	AD252	AD252.3	AD252.3 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD255	AD255.1	AD255.1 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD255	AD255.1	AD255.1 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD255	AD255.1	AD255.1 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD255	AD255.1	AD255.1 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD255	AD255.1	AD255.1 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD255	AD255.1	AD255.1 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD255	AD255.1	AD255.1 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD255	AD255.1	AD255.1 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD257	AD257.4	AD257.4 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD257	AD257.4	AD257.4 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD257	AD257.4	AD257.4 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD257	AD257.4	AD257.4 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD257	AD257.4	AD257.4 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD257	AD257.4	AD257.4 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD257	AD257.4	AD257.4 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD257	AD257.4	AD257.4 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD258	AD258.1	AD258.1 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD258	AD258.1	AD258.1 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD258	AD258.1	AD258.1 1:10	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD258	AD258.1	AD258.1 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD258	AD258.1	AD258.1 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD258	AD258.1	AD258.1 1:1	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD258	AD258.1	AD258.1 1:100	Undetermined			0			0
5-Jun-13	rpoB2 on aTBassay18	AD258	AD258.1	AD258.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD102	AD102.1	AD102.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD102	AD102.1	AD102.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD102	AD102.1	AD102.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD102	AD102.1	AD102.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD102	AD102.1	AD102.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD102	AD102.1	AD102.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD109	AD109.1	AD109.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD109	AD109.1	AD109.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD109	AD109.1	AD109.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD109	AD109.1	AD109.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD109	AD109.1	AD109.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD109	AD109.1	AD109.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD15	AD15.1	AD15.1 1:1	37.17461	36.918232	1.0000224	6.001645	7.534504	4.1094327	7.534504
5-Oct-13	rpoB2 on aTBassay22	AD15	AD15.1	AD15.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD15	AD15.1	AD15.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD15	AD15.1	AD15.1 1:1	37.76511	36.918232	1.0000224	4.411821	7.534504	4.1094327	7.534504
5-Oct-13	rpoB2 on aTBassay22	AD15	AD15.1	AD15.1 1:1	35.814983	36.918232	1.0000224	12.190045	7.534504	4.1094327	7.534504
5-Oct-13	rpoB2 on aTBassay22	AD15	AD15.1	AD15.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD171	AD171.1	AD171.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD171	AD171.1	AD171.1 1:10	Undetermined			0			0

5-Oct-13	rpoB2 on aTBassay22	AD171	AD171.1	AD171.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD171	AD171.1	AD171.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD171	AD171.1	AD171.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD171	AD171.1	AD171.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD172	AD172.1	AD172.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD172	AD172.1	AD172.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD172	AD172.1	AD172.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD172	AD172.1	AD172.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD172	AD172.1	AD172.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD172	AD172.1	AD172.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD209	AD209.1	AD209.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD209	AD209.1	AD209.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD209	AD209.1	AD209.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD209	AD209.1	AD209.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD209	AD209.1	AD209.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD209	AD209.1	AD209.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD211	AD211.1	AD211.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD211	AD211.1	AD211.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD211	AD211.1	AD211.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD211	AD211.1	AD211.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD211	AD211.1	AD211.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD211	AD211.1	AD211.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD28	AD28.1	AD28.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD28	AD28.1	AD28.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD28	AD28.1	AD28.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD28	AD28.1	AD28.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD28	AD28.1	AD28.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD28	AD28.1	AD28.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD281	AD281.1	AD281.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD281	AD281.1	AD281.1 1:1	36.670547	36.09781	0.5162116	7.804745	10.765933	2.7214847	10.765933
5-Oct-13	rpoB2 on aTBassay22	AD281	AD281.1	AD281.1 1:1	35.954445	36.09781	0.5162116	11.33548	10.765933	2.7214847	10.765933
5-Oct-13	rpoB2 on aTBassay22	AD281	AD281.1	AD281.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD281	AD281.1	AD281.1 1:10	36.654583			7.8699527			78.699527
5-Oct-13	rpoB2 on aTBassay22	AD281	AD281.1	AD281.1 1:1	35.66843	36.09781	0.5162116	13.157573	10.765933	2.7214847	10.765933
5-Oct-13	rpoB2 on aTBassay22	AD31	AD31.1	AD31.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD31	AD31.1	AD31.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD31	AD31.1	AD31.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD31	AD31.1	AD31.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD31	AD31.1	AD31.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD31	AD31.1	AD31.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD44	AD44.1	AD44.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD44	AD44.1	AD44.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD44	AD44.1	AD44.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD44	AD44.1	AD44.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD44	AD44.1	AD44.1 1:1	Undetermined			0			0

5-Oct-13	rpoB2 on aTBassay22	AD44	AD44.1	AD44.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD51	AD51.1	AD51.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD51	AD51.1	AD51.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD51	AD51.1	AD51.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD51	AD51.1	AD51.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD51	AD51.1	AD51.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD51	AD51.1	AD51.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD58	AD58.1	AD58.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD58	AD58.1	AD58.1 1:1	35.854343	37.15626	1.2554954	11.942536	6.978235	4.479916	6.978235
5-Oct-13	rpoB2 on aTBassay22	AD58	AD58.1	AD58.1 1:1	38.359512	37.15626	1.2554954	3.2365487	6.978235	4.479916	6.978235
5-Oct-13	rpoB2 on aTBassay22	AD58	AD58.1	AD58.1 1:1	37.254925	37.15626	1.2554954	5.7556195	6.978235	4.479916	6.978235
5-Oct-13	rpoB2 on aTBassay22	AD58	AD58.1	AD58.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD58	AD58.1	AD58.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD60	AD60.1	AD60.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD60	AD60.1	AD60.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD60	AD60.1	AD60.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD60	AD60.1	AD60.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD60	AD60.1	AD60.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD60	AD60.1	AD60.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD65	AD65.1	AD65.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD65	AD65.1	AD65.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD65	AD65.1	AD65.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD65	AD65.1	AD65.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD65	AD65.1	AD65.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD65	AD65.1	AD65.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD79	AD79.1	AD79.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD79	AD79.1	AD79.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD79	AD79.1	AD79.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD79	AD79.1	AD79.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD79	AD79.1	AD79.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD79	AD79.1	AD79.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD85	AD85.1	AD85.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD85	AD85.1	AD85.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD85	AD85.1	AD85.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD85	AD85.1	AD85.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD85	AD85.1	AD85.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD85	AD85.1	AD85.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD87	AD87.1	AD87.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD87	AD87.1	AD87.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD87	AD87.1	AD87.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD87	AD87.1	AD87.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD87	AD87.1	AD87.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD87	AD87.1	AD87.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD90	AD90.1	AD90.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD90	AD90.1	AD90.1 1:10	Undetermined			0			0

5-Oct-13	rpoB2 on aTBassay22	AD90	AD90.1	AD90.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD90	AD90.1	AD90.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD90	AD90.1	AD90.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD90	AD90.1	AD90.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD93	AD93.1	AD93.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD93	AD93.1	AD93.1 1:1	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD93	AD93.1	AD93.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD93	AD93.1	AD93.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD93	AD93.1	AD93.1 1:10	Undetermined			0			0
5-Oct-13	rpoB2 on aTBassay22	AD93	AD93.1	AD93.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD102	AD102.1	AD102.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD102	AD102.1	AD102.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD102	AD102.1	AD102.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD109	AD109.1	AD109.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD109	AD109.1	AD109.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD109	AD109.1	AD109.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD109	AD109.1	AD109.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD109	AD109.1	AD109.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD109	AD109.1	AD109.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD15	AD15.1	AD15.1 1:1	34.13626	34.205982	0.0903891	58.11252	56.12376	2.5558896	56.12376
18-Oct-13	IS6110 on aTBassay23	AD15	AD15.1	AD15.1 1:10	37.222855	37.13968	0.7938459	12.059094	13.291068	5.4230213	132.91068
18-Oct-13	IS6110 on aTBassay23	AD15	AD15.1	AD15.1 1:10	36.30752	37.13968	0.7938459	19.224089	13.291068	5.4230213	132.91068
18-Oct-13	IS6110 on aTBassay23	AD15	AD15.1	AD15.1 1:10	37.888664	37.13968	0.7938459	8.590022	13.291068	5.4230213	132.91068
18-Oct-13	IS6110 on aTBassay23	AD15	AD15.1	AD15.1 1:1	34.173588	34.205982	0.0903891	57.017826	56.12376	2.5558896	56.12376
18-Oct-13	IS6110 on aTBassay23	AD15	AD15.1	AD15.1 1:1	34.30811	34.205982	0.0903891	53.240944	56.12376	2.5558896	56.12376
18-Oct-13	IS6110 on aTBassay23	AD171	AD171.1	AD171.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD171	AD171.1	AD171.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD171	AD171.1	AD171.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD171	AD171.1	AD171.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD171	AD171.1	AD171.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD171	AD171.1	AD171.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD172	AD172.1	AD172.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD172	AD172.1	AD172.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD172	AD172.1	AD172.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD172	AD172.1	AD172.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD172	AD172.1	AD172.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD172	AD172.1	AD172.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD209	AD209.1	AD209.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD209	AD209.1	AD209.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD209	AD209.1	AD209.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD209	AD209.1	AD209.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD209	AD209.1	AD209.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD209	AD209.1	AD209.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD211	AD211.1	AD211.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD211	AD211.1	AD211.1 1:1	Undetermined			0			0

18-Oct-13	IS6110 on aTBassay23	AD211	AD211.1	AD211.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD211	AD211.1	AD211.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD211	AD211.1	AD211.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD211	AD211.1	AD211.1 1:1	47.85154			0.05364864			0.05364864	Yes	positive
18-Oct-13	IS6110 on aTBassay23	AD28	AD28.1	AD28.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD28	AD28.1	AD28.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD28	AD28.1	AD28.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD28	AD28.1	AD28.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD28	AD28.1	AD28.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD28	AD28.1	AD28.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD281	AD281.1	AD281.1 1:1	36.93392	36.776073	0.9251052	13.971573	16.328316	7.885576	16.328316		
18-Oct-13	IS6110 on aTBassay23	AD281	AD281.1	AD281.1 1:10	37.081852	37.052937	0.4480214	12.957271	13.380628	3.067163	133.80628		
18-Oct-13	IS6110 on aTBassay23	AD281	AD281.1	AD281.1 1:10	37.485798	37.052937	0.4480214	10.547134	13.380628	3.067163	133.80628		
18-Oct-13	IS6110 on aTBassay23	AD281	AD281.1	AD281.1 1:10	36.591156	37.052937	0.4480214	16.637476	13.380628	3.067163	133.80628		
18-Oct-13	IS6110 on aTBassay23	AD281	AD281.1	AD281.1 1:1	37.6121	36.776073	0.9251052	9.889823	16.328316	7.885576	16.328316		
18-Oct-13	IS6110 on aTBassay23	AD281	AD281.1	AD281.1 1:1	35.7822	36.776073	0.9251052	25.123552	16.328316	7.885576	16.328316		
18-Oct-13	IS6110 on aTBassay23	AD31	AD31.1	AD31.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD31	AD31.1	AD31.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD31	AD31.1	AD31.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD31	AD31.1	AD31.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD31	AD31.1	AD31.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD31	AD31.1	AD31.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD44	AD44.1	AD44.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD44	AD44.1	AD44.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD44	AD44.1	AD44.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD44	AD44.1	AD44.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD44	AD44.1	AD44.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD44	AD44.1	AD44.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD51	AD51.1	AD51.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD51	AD51.1	AD51.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD51	AD51.1	AD51.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD51	AD51.1	AD51.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD51	AD51.1	AD51.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD51	AD51.1	AD51.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD58	AD58.1	AD58.1 1:1	38.4001	40.151306	2.0230067	6.619593	3.6439161	2.8766499	3.6439161		
18-Oct-13	IS6110 on aTBassay23	AD58	AD58.1	AD58.1 1:1	42.365757	40.151306	2.0230067	0.8777491	3.6439161	2.8766499	3.6439161		
18-Oct-13	IS6110 on aTBassay23	AD58	AD58.1	AD58.1 1:1	39.688057	40.151306	2.0230067	3.4344063	3.6439161	2.8766499	3.6439161		
18-Oct-13	IS6110 on aTBassay23	AD58	AD58.1	AD58.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD58	AD58.1	AD58.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD58	AD58.1	AD58.1 1:10	45.186924			0.20851865			2.0851865		
18-Oct-13	IS6110 on aTBassay23	AD60	AD60.1	AD60.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD60	AD60.1	AD60.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD60	AD60.1	AD60.1 1:10	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD60	AD60.1	AD60.1 1:1	Undetermined			0			0		
18-Oct-13	IS6110 on aTBassay23	AD60	AD60.1	AD60.1 1:1	Undetermined			0			0		

18-Oct-13	IS6110 on aTBassay23	AD60	AD60.1	AD60.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD65	AD65.1	AD65.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD65	AD65.1	AD65.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD65	AD65.1	AD65.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD65	AD65.1	AD65.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD65	AD65.1	AD65.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD65	AD65.1	AD65.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD79	AD79.1	AD79.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD79	AD79.1	AD79.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD79	AD79.1	AD79.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD79	AD79.1	AD79.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD79	AD79.1	AD79.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD79	AD79.1	AD79.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD85	AD85.1	AD85.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD85	AD85.1	AD85.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD85	AD85.1	AD85.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD85	AD85.1	AD85.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD85	AD85.1	AD85.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD85	AD85.1	AD85.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD87	AD87.1	AD87.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD87	AD87.1	AD87.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD87	AD87.1	AD87.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD87	AD87.1	AD87.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD87	AD87.1	AD87.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD87	AD87.1	AD87.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD90	AD90.1	AD90.1 1:1	39.675674			3.4561415			3.4561415
18-Oct-13	IS6110 on aTBassay23	AD90	AD90.1	AD90.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD90	AD90.1	AD90.1 1:1	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD90	AD90.1	AD90.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD90	AD90.1	AD90.1 1:10	Undetermined			0			0
18-Oct-13	IS6110 on aTBassay23	AD90	AD90.1	AD90.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD109	AD109.1	AD109.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD109	AD109.1	AD109.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD109	AD109.1	AD109.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD15	AD15.1	AD15.1 1:10	33.35128	33.361496	0.3282938	5.086708	5.134239	1.1228709	51.34239
5-Nov-13	rpoB on aTBassay24	AD15	AD15.1	AD15.1 1:1	31.928614	31.544443	0.349651	13.263802	17.492334	3.9269748	17.492334
5-Nov-13	rpoB on aTBassay24	AD15	AD15.1	AD15.1 1:1	31.244814	31.544443	0.349651	21.024614	17.492334	3.9269748	17.492334
5-Nov-13	rpoB on aTBassay24	AD15	AD15.1	AD15.1 1:1	31.459904	31.544443	0.349651	18.18859	17.492334	3.9269748	17.492334
5-Nov-13	rpoB on aTBassay24	AD15	AD15.1	AD15.1 1:10	33.03843	33.361496	0.3282938	6.2801204	5.134239	1.1228709	51.34239
5-Nov-13	rpoB on aTBassay24	AD15	AD15.1	AD15.1 1:10	33.69478	33.361496	0.3282938	4.035888	5.134239	1.1228709	51.34239
5-Nov-13	rpoB on aTBassay24	AD171	AD171.1	AD171.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD171	AD171.1	AD171.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD171	AD171.1	AD171.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD172	AD172.1	AD172.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD172	AD172.1	AD172.1 1:10	Undetermined			0			0

5-Nov-13	rpoB on aTBassay24	AD172	AD172.1	AD172.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD209	AD209.1	AD209.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD209	AD209.1	AD209.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD209	AD209.1	AD209.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD211	AD211.1	AD211.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD211	AD211.1	AD211.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD211	AD211.1	AD211.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD28	AD28.1	AD28.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD28	AD28.1	AD28.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD28	AD28.1	AD28.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD28	AD28.1	AD28.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD28	AD28.1	AD28.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD28	AD28.1	AD28.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD281	AD281.1	AD281.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD281	AD281.1	AD281.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD281	AD281.1	AD281.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD31	AD31.1	AD31.1 1:1	42.857502			0.00841765			0.008417646
5-Nov-13	rpoB on aTBassay24	AD31	AD31.1	AD31.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD31	AD31.1	AD31.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD31	AD31.1	AD31.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD31	AD31.1	AD31.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD31	AD31.1	AD31.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD44	AD44.1	AD44.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD44	AD44.1	AD44.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD44	AD44.1	AD44.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD44	AD44.1	AD44.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD44	AD44.1	AD44.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD44	AD44.1	AD44.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD51	AD51.1	AD51.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD51	AD51.1	AD51.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD51	AD51.1	AD51.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD51	AD51.1	AD51.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD51	AD51.1	AD51.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD51	AD51.1	AD51.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD58	AD58.1	AD58.1 1:1	35.598057	36.263184	0.5925174	1.1196823	0.75605774	0.31943095	0.75605774
5-Nov-13	rpoB on aTBassay24	AD58	AD58.1	AD58.1 1:10	37.476383	39.288383	2.5625553	0.31590077	0.1716985	0.20393282	1.716985
5-Nov-13	rpoB on aTBassay24	AD58	AD58.1	AD58.1 1:10	Undetermined	39.288383	2.5625553	0	0.1716985	0.20393282	1.716985
5-Nov-13	rpoB on aTBassay24	AD58	AD58.1	AD58.1 1:10	41.100384	39.288383	2.5625553	0.02749621	0.1716985	0.20393282	1.716985
5-Nov-13	rpoB on aTBassay24	AD58	AD58.1	AD58.1 1:1	36.734608	36.263184	0.5925174	0.52068233	0.75605774	0.31943095	0.75605774
5-Nov-13	rpoB on aTBassay24	AD58	AD58.1	AD58.1 1:1	36.456882	36.263184	0.5925174	0.6278085	0.75605774	0.31943095	0.75605774
5-Nov-13	rpoB on aTBassay24	AD60	AD60.1	AD60.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD60	AD60.1	AD60.1 1:1	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD60	AD60.1	AD60.1 1:10	48.76674			0.00015715			0.001571514
5-Nov-13	rpoB on aTBassay24	AD60	AD60.1	AD60.1 1:10	Undetermined			0			0
5-Nov-13	rpoB on aTBassay24	AD60	AD60.1	AD60.1 1:1	Undetermined			0			0

5-Nov-13	rpoB on aTBassay24	AD60	AD60.1	AD60.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD65	AD65.1	AD65.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD65	AD65.1	AD65.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD65	AD65.1	AD65.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD65	AD65.1	AD65.1 1:10	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD65	AD65.1	AD65.1 1:10	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD65	AD65.1	AD65.1 1:10	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD79	AD79.1	AD79.1 1:10	42.820297	43.492054	0.8303189	0.00863129	0.006031294	0.002878995	0.06031294	
5-Nov-13	rpoB on aTBassay24	AD79	AD79.1	AD79.1 1:10	44.420372	43.492054	0.8303189	0.00293724	0.006031294	0.002878995	0.06031294	
5-Nov-13	rpoB on aTBassay24	AD79	AD79.1	AD79.1 1:1	47.884197	45.40043	3.5125763	0.00028479	0.004186771	0.005518234	0.004186771	
5-Nov-13	rpoB on aTBassay24	AD79	AD79.1	AD79.1 1:10	43.235485	43.492054	0.8303189	0.00652535	0.006031294	0.002878995	0.06031294	
5-Nov-13	rpoB on aTBassay24	AD79	AD79.1	AD79.1 1:1	Undetermined	45.40043	3.5125763	0	0.004186771	0.005518234	0.004186771	
5-Nov-13	rpoB on aTBassay24	AD79	AD79.1	AD79.1 1:1	42.916664	45.40043	3.5125763	0.00808875	0.004186771	0.005518234	0.004186771	
5-Nov-13	rpoB on aTBassay24	AD82	AD82.9	AD82.9 1:1	31.95381	32.95332	0.8783079	13.040564	7.5291867	4.796784	7.5291867	
5-Nov-13	rpoB on aTBassay24	AD82	AD82.9	AD82.9 1:1	33.30419	32.95332	0.8783079	5.250656	7.5291867	4.796784	7.5291867	
5-Nov-13	rpoB on aTBassay24	AD82	AD82.9	AD82.9 1:1	33.601948	32.95332	0.8783079	4.296341	7.5291867	4.796784	7.5291867	
5-Nov-13	rpoB on aTBassay24	AD85	AD85.1	AD85.1 1:10	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD85	AD85.1	AD85.1 1:1	47.312363			0.00041863			0.000418627	
5-Nov-13	rpoB on aTBassay24	AD85	AD85.1	AD85.1 1:10	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD85	AD85.1	AD85.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD85	AD85.1	AD85.1 1:10	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD85	AD85.1	AD85.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD87	AD87.1	AD87.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD87	AD87.1	AD87.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD87	AD87.1	AD87.1 1:10	39.480175	39.49075	0.156899	0.08190286	0.08162353	0.008569653	0.8162353	
5-Nov-13	rpoB on aTBassay24	AD87	AD87.1	AD87.1 1:10	39.652668	39.49075	0.156899	0.07291763	0.08162353	0.008569653	0.8162353	
5-Nov-13	rpoB on aTBassay24	AD87	AD87.1	AD87.1 1:10	39.339405	39.49075	0.156899	0.09005011	0.08162353	0.008569653	0.8162353	
5-Nov-13	rpoB on aTBassay24	AD87	AD87.1	AD87.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD90	AD90.1	AD90.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD90	AD90.1	AD90.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD90	AD90.1	AD90.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD90	AD90.1	AD90.1 1:10	28.89057	28.831108	0.1212672	102.68619	107.12451	8.9409895	1071.2451	
5-Nov-13	rpoB on aTBassay24	AD90	AD90.1	AD90.1 1:10	28.91117	28.831108	0.1212672	101.27105	107.12451	8.9409895	1071.2451	
5-Nov-13	rpoB on aTBassay24	AD90	AD90.1	AD90.1 1:10	28.691587	28.831108	0.1212672	117.4163	107.12451	8.9409895	1071.2451	
5-Nov-13	rpoB on aTBassay24	AD93	AD93.1	AD93.1 1:10	44.39805	45.02047	0.8802393	0.00298175	0.002135393	0.00119693	0.021353932	
5-Nov-13	rpoB on aTBassay24	AD93	AD93.1	AD93.1 1:10	45.642895	45.02047	0.8802393	0.00128904	0.002135393	0.00119693	0.021353932	
5-Nov-13	rpoB on aTBassay24	AD93	AD93.1	AD93.1 1:10	Undetermined	45.02047	0.8802393	0	0.002135393	0.00119693	0.021353932	Yes
5-Nov-13	rpoB on aTBassay24	AD93	AD93.1	AD93.1 1:1	Undetermined			0			0	
5-Nov-13	rpoB on aTBassay24	AD93	AD93.1	AD93.1 1:1	Undetermined			0			0	
28-Jan-14	IS1081 on aTBassay29	AD112	AD112Tu	AD112Tu 1:1	Undetermined			0			0	
28-Jan-14	IS1081 on aTBassay29	AD112	AD112Tu	AD112Tu 1:10	Undetermined			0			0	
28-Jan-14	IS1081 on aTBassay29	AD112	AD112Tu	AD112Tu 1:1	Undetermined			0			0	
28-Jan-14	IS1081 on aTBassay29	AD112	AD112Tu	AD112Tu 1:10	Undetermined			0			0	
28-Jan-14	IS1081 on aTBassay29	AD112	AD112Tu	AD112Tu 1:10	Undetermined			0			0	
28-Jan-14	IS1081 on aTBassay29	AD160	AD160Tu	AD160Tu 1:10	Undetermined			0			0	

28-Jan-14	IS1081 on aTBassay29	AD160	AD160Tu	AD160Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD160	AD160Tu	AD160Tu 1:1	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD160	AD160Tu	AD160Tu 1:1	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD160	AD160Tu	AD160Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD250	AD250Tu	AD250Tu 1:1	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD250	AD250Tu	AD250Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD250	AD250Tu	AD250Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD250	AD250Tu	AD250Tu 1:1	39.81173			0.37434202			0.37434202		
28-Jan-14	IS1081 on aTBassay29	AD250	AD250Tu	AD250Tu 1:1	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD250	AD250Tu	AD250Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD258	AD258Tu	AD258Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD258	AD258Tu	AD258Tu 1:1	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD258	AD258Tu	AD258Tu 1:1	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD258	AD258Tu	AD258Tu 1:1	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD258	AD258Tu	AD258Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD258	AD258Tu	AD258Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD262	AD262Tu	AD262Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD262	AD262Tu	AD262Tu 1:1	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD262	AD262Tu	AD262Tu 1:1	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD262	AD262Tu	AD262Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD262	AD262Tu	AD262Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD3	AD3Tu	AD3Tu 1:1	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD3	AD3Tu	AD3Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD3	AD3Tu	AD3Tu 1:1	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD3	AD3Tu	AD3Tu 1:10	Undetermined			0			0		
28-Jan-14	IS1081 on aTBassay29	AD3	AD3Tu	AD3Tu 1:10	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD112	AD112Tu	AD112Tu 1:10	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD112	AD112Tu	AD112Tu 1:10	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD112	AD112Tu	AD112Tu 1:10	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD160	AD160Tu	AD160Tu 1:1	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD160	AD160Tu	AD160Tu 1:1	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD160	AD160Tu	AD160Tu 1:10	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD160	AD160Tu	AD160Tu 1:10	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD160	AD160Tu	AD160Tu 1:10	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD262	AD262Tu	AD262Tu 1:10	38.886375	40.50425	1.5612614	1.6773624	0.8276239	0.7538148	8.276239	Yes	
28-Jan-14	rpoB on aTBassay29	AD262	AD262Tu	AD262Tu 1:1	46.236916	46.552345	0.4460822	0.01697102	0.014206421	0.003909734	0.014206421	Yes	contam EB
28-Jan-14	rpoB on aTBassay29	AD262	AD262TuB	AD262TuB 1:10	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD262	AD262Tu	AD262Tu 1:1	46.86777	46.552345	0.4460822	0.01144182	0.014206421	0.003909734	0.014206421	Yes	contam EB
28-Jan-14	rpoB on aTBassay29	AD262	AD262TuB	AD262TuB 1:10	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD262	AD262TuB	AD262TuB 1:1	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD262	AD262TuB	AD262TuB 1:1	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD262	AD262TuB	AD262TuB 1:1	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD262	AD262Tu	AD262Tu 1:10	42.001953	40.50425	1.5612614	0.23936827	0.8276239	0.7538148	8.276239	Yes	contam EB
28-Jan-14	rpoB on aTBassay29	AD262	AD262TuB	AD262TuB 1:10	Undetermined			0			0		
28-Jan-14	rpoB on aTBassay29	AD262	AD262Tu	AD262Tu 1:10	40.624424	40.50425	1.5612614	0.5661411	0.8276239	0.7538148	8.276239	Yes	contam EB

29-Jan-14	rpoB2 on aTBassay28.2	AD112	AD112Tu	AD112Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD112	AD112Tu	AD112Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD112	AD112Tu	AD112Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD114	AD114Tu	AD114Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD114	AD114Tu	AD114Tu 1:1	45.55406			0.09442996			0.09442996
29-Jan-14	rpoB2 on aTBassay28.2	AD114	AD114Tu	AD114Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD114	AD114Tu	AD114Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD114	AD114Tu	AD114Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD151	AD151Tu	AD151Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD151	AD151Tu	AD151Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD151	AD151Tu	AD151Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD151	AD151Tu	AD151Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD151	AD151Tu	AD151Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD160	AD160Tu	AD160Tu 1:10	Undetermined	40.748158	1.6764007	0	1.1608646	0.8575566	11.608646
29-Jan-14	rpoB2 on aTBassay28.2	AD160	AD160Tu	AD160Tu 1:10	41.93355	40.748158	1.6764007	0.55448043	1.1608646	0.8575566	11.608646
29-Jan-14	rpoB2 on aTBassay28.2	AD160	AD160Tu	AD160Tu 1:10	39.562763	40.748158	1.6764007	1.7672486	1.1608646	0.8575566	11.608646
29-Jan-14	rpoB2 on aTBassay28.2	AD162	AD162Tu	AD162Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD162	AD162Tu	AD162Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD162	AD162Tu	AD162Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD162	AD162Tu	AD162Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD162	AD162Tu	AD162Tu 1:1	47.059822			0.04522493			0.045224927
29-Jan-14	rpoB2 on aTBassay28.2	AD235	AD235Tu	AD235Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD235	AD235Tu	AD235Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD235	AD235Tu	AD235Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD235	AD235Tu	AD235Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD235	AD235Tu	AD235Tu 1:1	44.581223			0.15194339			0.15194339
29-Jan-14	rpoB2 on aTBassay28.2	AD236	AD236Tu	AD236Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD236	AD236Tu	AD236Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD236	AD236Tu	AD236Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD236	AD236Tu	AD236Tu 1:1	38.03735			3.7256565			3.7256565
29-Jan-14	rpoB2 on aTBassay28.2	AD236	AD236Tu	AD236Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD240	AD240Tu	AD240Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD240	AD240Tu	AD240Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD240	AD240Tu	AD240Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD240	AD240Tu	AD240Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD240	AD240Tu	AD240Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD256	AD256Tu	AD256Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD256	AD256Tu	AD256Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD256	AD256Tu	AD256Tu 1:1	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD256	AD256Tu	AD256Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD256	AD256Tu	AD256Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD262	AD262Tu	AD262Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD262	AD262Tu	AD262Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD262	AD262Tu	AD262Tu 1:10	Undetermined			0			0
29-Jan-14	rpoB2 on aTBassay28.2	AD264	AD264Tu	AD264Tu 1:10	Undetermined			0			0

29-Jan-14	rpoB2 on aTBassay28.2	AD264	AD264Tu	AD264Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD264	AD264Tu	AD264Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD264	AD264Tu	AD264Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD264	AD264Tu	AD264Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD282	AD282Tu	AD282Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD282	AD282Tu	AD282Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD282	AD282Tu	AD282Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD282	AD282Tu	AD282Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD282	AD282Tu	AD282Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD3	AD3Tu	AD3Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD3	AD3Tu	AD3Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD3	AD3Tu	AD3Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD3	AD3Tu	AD3Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD3	AD3Tu	AD3Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD6	AD6Tu	AD6Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD6	AD6Tu	AD6Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD6	AD6Tu	AD6Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD6	AD6Tu	AD6Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD6	AD6Tu	AD6Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD76	AD76Tu	AD76Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD76	AD76Tu	AD76Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD76	AD76Tu	AD76Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD76	AD76Tu	AD76Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD76	AD76Tu	AD76Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD81	AD81Tu	AD81Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD81	AD81Tu	AD81Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD81	AD81Tu	AD81Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD81	AD81Tu	AD81Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD81	AD81Tu	AD81Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD81	AD81Tu	AD81Tu 1:1	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD81	AD81Tu	AD81Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD81	AD81Tu	AD81Tu 1:10	Undetermined	0	0
29-Jan-14	rpoB2 on aTBassay28.2	AD81	AD81Tu	AD81Tu 1:10	Undetermined	0	0

APPENDIX B

SUPPLEMENTARY INFORMATION AND EXTENDED DATA FOR "PRE-COLUMBIAN GENOMES REVEAL SEALS AS SOURCE OF NEW WORLD HUMAN TUBERCULOSIS"

Appendix B

Table S1 - Screening data for all 68 samples and negative controls

					SHOTGUN					BEAD CAPTURE							
										gyrA/B		rpoB		katG		mpt40	
Sample Name	Archaeological provenience	Continent	Burial number	Archaeological time period	No. merged reads for shotgun	No. reads mapping to human genome post rmdup	% merged reads mapping to human genome	No. reads mapping to TB complex post rmdup	% merged reads mapping to TB complex	No. merged reads mapping	% of region with ≥ 1X coverage	No. merged reads mapping	% of region with ≥ 1X coverage	No. merged reads mapping	% of region with ≥ 1X coverage	No. merged reads mapping	% of region with ≥ 1X coverage
Sample_1	Arikara, Leavenworth (39CO9)	North America	32534 1	AD 1725-1775	126904	1066	0.84%	609	0.48%	163	30.17	92 1	50.72	37	1 1.1 1	0	0
Sample_102	Schild	North America	SB-262A	AD 1040-1407	45284	475	1.05%	140	0.31%	167	30.29	685	48.76	34	16.47	0	0
Sample_105	Chirikov	North America	62CF/BSc-18	AD 1750-1899	75047	1944	2.59%	345	0.46%	1216	30.17	538	46.29	13	11.43	0	0
Sample_106	Chirikov	North America	62CF/AE3-14	AD 1750-1899	98088	1010	1.03%	235	0.24%	1175	31.30	1255	55.54	34	16.11	0	0
Sample_109	Yokem	North America	Md 2-5	AD 1039-1401	146843	2041	1.39%	367	0.25%	75	14.79	201	32.16	39	14.43	0	0
Sample_11	Arikara, Cheyenne River village (39STI)	North America	382713	AD 1750-1775	47578	566	1.19%	309	0.65%	82	24.52	1096	47.49	42	13.59	0	0
Sample_112	Gooden Mds	North America	F85-48	AD 800-1000	39415	1758	4.46%	150	0.38%	192	29.58	853	43.00	30	11.47	0	0
Sample_114	Highland Park	North America	16a	AD 1847-1850	78668	8331	10.59%	205	0.26%	264	26.45	777	46.94	19	14.75	4	11.36
Sample_11B	neg control				37782	540	1.43%	8	0.02%	0	0	1	0.68	1	0.84	0	0
Sample_127	Highland Park	North America	138a	AD 1847-1850	111331	1225	1.10%	356	0.32%	354	44.56	1150	58.80	42	17.51	8	23.03
Sample_128	Highland Park	North America	141a	AD 1847-1850	68629	17006	24.78%	178	0.26%	623	53.79	1030	51.21	33	14.19	10	24.55
Sample_134	Highland Park	North America	225a	AD 1847-1850	77283	726	0.94%	185	0.24%	67	22.99	189	34.45	7	10.07	0	0
Sample_139B	neg control				31590	515	1.63%	6	0.02%	0	0	0	0	0	0	0	0
Sample_145	Norris Farms	North America	21	AD 1100-1250	82507	1601	1.94%	231	0.28%	366	31.00	990	47.98	80	16.75	0	0
Sample_15	Arikara, Mobridge (39WW1)	North America	382990	AD 1500-1750	137986	1863	1.35%	980	0.71%	7	5.52	25	13.81	0	0	0	0
Sample_151	Norris Farms	North America	95	AD 1100-1250	53184	792	1.49%	160	0.30%	385	32.45	1093	48.03	77	16.87	0	0
Sample_156	Norris Farms	North America	160	AD 1100-1250	76544	1393	1.82%	199	0.26%	240	25.82	1070	49.44	113	16.43	0	0
Sample_160	Norris Farms	North America	228	AD 1100-1250	34078	4641	13.62%	123	0.36%	130	25.70	603	51.73	12	9.99	0	0
Sample_160B	neg control				34090	505	1.48%	0	0.00%	1	0.46	0	0	0	0	0	0
Sample_162	San Cristobal	North America	8708	1598)	39200	698	1.78%	243	0.62%	1638	54.67	3939	66.22	99	15.43	0	0
Sample_171	Orendorf	North America	Burial 116	AD 1150-1300	119724	3256	2.72%	335	0.28%	7	4.07	33	10.71	4	4.08	0	0
Sample_172	Orendorf	North America	Burial 153	AD 1150-1300	71967	885	1.23%	144	0.20%	6	3.02	24	14.99	1	0.80	0	0
Sample_19	Arikara, Sully School Village (39SL4)	North America	388120	AD 1550-1700	104257	792	0.76%	803	0.77%	188	34.51	2262	57.57	49	19.38	0	0
Sample_194	Moquegua (Chen Chen)	South America	M1-3353-1	AD 500-700	62036	409	0.66%	335	0.54%	397	27.90	699	50.35	29	11.35	0	0
Sample_19B	neg control				25488	647	2.54%	0	0.00%	0	0	0	0.00	0	0	0	0
Sample_209	Puruchuco	South America	Tomb 5, Indv. 2, Sector 15	AD 1470-1540	61584	480	0.78%	351	0.57%	112	8.60	1237	27.73	0	0	0	0
Sample_211	Puruchuco	South America	Tomb 49, Sector 15	AD 1470-1540	125040	950	0.76%	413	0.33%	26	6.04	518	22.83	0	0	0	0
Sample_216	Sitio Myren-7	South America	T.D.F. 30477	AD 1310+-20	147630	38650	26.18%	561	0.38%	257	33.67	3041	56.55	50	16.87	0	0
Sample_22	Arikara, Sully School Village (39SL4)	North America	388403	AD 1550-1700	95895	767	0.80%	595	0.62%	160	26.49	756	53.84	43	17.47	0	0
Sample_227	Lambayeque, Peru	South America	ILL-22	AD 900-1750	43886	421	0.96%	162	0.37%	117	15.25	760	47.51	159	24.30	0	0
Sample_23	Uxbridge	North America	P1 U1 L15 S3 F1	AD 1410-1570	N/A	N/A	N/A	N/A	N/A	3	1.66	9	6.46	2	1.88	0	0
Sample_24	Uxbridge	North America	P5 U2 L3	AD 1410-1570	32687	438	1.34%	118	0.36%	104	18.02	166	33.07	5	2.64	0	0
Sample_25	Uxbridge	North America	P6 U2 L6 S5	AD 1410-1570	N/A	N/A	N/A	N/A	N/A	0	0	0	0	0	0	0	0
Sample_250	Misión Salesiana	South America	E 14-15 (1)	AD 1870-1930	134877	688	0.51%	405	0.30%	6	1.41	36	10.24	0	0	0	0
Sample_250B_262B	neg control				72217	1040	1.44%	0	0.00%	0	0	1	0.63	0	0	0	0
Sample_252	Misión Salesiana	South America	C 7-8	AD 1870-1930	99528	13805	13.87%	299	0.30%	2	0.82	258	16.03	3	2.24	0	0
Sample_256	Misión Salesiana	South America	D-E 14	AD 1870-1930	125684	628	0.50%	465	0.37%	7	2.62	49	18.48	0	0	0	0
Sample_256B	neg control				83526	1462	1.75%	0	0.00%	0	0	0	0	1	1.60	0	0
Sample_258	Misión Salesiana	South America	D 16 (Bis)	AD 1870-1930	179828	106980	59.49%	198	0.11%	37	10.82	92	25.20	0	0	0	0
Sample_26	Uxbridge	North America	P7 U2 L5 S6	AD 1410-1570	90594	1522	1.68%	344	0.38%	120	11.14	358	34.04	22	7.03	0	0
Sample_262	Misión Salesiana	South America	D 15-16	AD 1870-1930	109119	524	0.48%	284	0.26%	17	7.82	97	27.29	0	0	0	0
Sample_263	Misión Salesiana	South America	C 15-16	AD 1870-1930	136855	23416	17.11%	315	0.23%								
Sample_264	Misión Salesiana	South America	D 14	AD 1870-1930	131461	1393	1.06%	421	0.32%	0	0	0	0	0	0	0	0
Sample_28	Uxbridge	North America	P16B U1 L17 S2 F1	AD 1410-1570	90839	1208	1.33%	554	0.61%	266	7.99	486	27.34	1	1.48	0	0
Sample_281	Las Delicias	South America	LD-90-1X-1 1	AD 770-990	132213	2234	1.69%	450	0.34%	0	0	0	0	0	0	0	0
Sample_29	Uxbridge	North America	P27 U1 South Wall	AD 1410-1570	105439	1476	1.40%	590	0.56%	186	13.68	138	32.60	4	3.80	0	0
Sample_3	Arikara, Sully School Village (39SL4)	North America	381390	AD 1550-1700	114969	6254	5.44%	908	0.79%	410	32.81	2440	56.24	133	16.63	0	0
Sample_31	Uxbridge	North America	P47 U3 L10 S7	AD 1410-1570	59873	886	1.48%	228	0.38%	201	22.70	326	40.03	13	8.11	1	4.55
Sample_32	Uxbridge	North America	P48 U1 L17 F1	AD 1410-1570	178627	3948	2.21%	982	0.55%	193	9.75	408	21.74	1	1.36	0	0
Sample_33	Uxbridge	North America	P55 U2 L9 S4	AD 1410-1570	90563	1349	1.49%	498	0.55%	237	11.73	384	33.07	6	3.36	0	0
Sample_35	Uxbridge	North America	P57 U3 L8 S7	AD 1410-1570	140267	1866	1.33%	673	0.48%	265	23.66	442	43.05	10	5.08	0	0
Sample_36	Uxbridge	North America	P132 U2 L8 S5	AD 1410-1570	36577	530	1.45%	139	0.38%	113	11.22	109	23.64	3	3.60	0	0
Sample_37	Uxbridge	North America	P135 U2 L8 S5	AD 1410-1570	127655	1685	1.32%	664	0.52%	245	13.76	654	36.17	17	6.31	0	0
Sample_39	Uxbridge	North America	P165 U3 L11 S7	AD 1410-1570	118583	1601	1.35%	534	0.45%	268	21.94	415	37.79	15	8.39	0	0
Sample_41	Uxbridge	North America	P219 U2 L7 S4	AD 1410-1570	104054	1384	1.33%	416	0.40%	258	16.93	582	35.89	19	8.91	0	0
Sample_43	Uxbridge	North America	P253 U1 L15 S2 F1	AD 1410-1570	76613	858	1.12%	498	0.65%	300	11.75	276	27.94	1	1.24	0	0
Sample_44	Uxbridge	North America	P280 U3 L10 S8	AD 1410-1570	182908	2268	1.24%	1061	0.58%	284	9.88	564	32.79	3	4.52	0	0
Sample_48	Uxbridge	North America	No Pit U3 L7 S8	AD 1410-1570	116728	1984	1.70%	479	0.41%	143	18.57	203	35.39	7	4.44	0	0
Sample_49	Uxbridge	North America	No Pit U3 L8 S7	AD 1410-1570	51859	928	1.79%	176	0.34%	90	22.68	166	37.16	8	5.32	0	0
Sample_51	Chiribaya Alta, Cemetery 1	South America	196-23-1	AD 772-1350	125236	4195	3.35%	864	0.69%	65	9.63	152	21.24	7	2.92	0	0
Sample_54	Chiribaya Alta, Cemetery 1	South America	2069-103-1	AD 1028-1224	70875	3260	4.60%	928	1.31%	3579	94.27	5312	92.34	2022	84.49	853	55.45
Sample_58	El Algodonal (AL)	South America	386-1	AD 1023-1211	112428	5104	4.54%	2080	1.85%	722	87.25	837	89.26	396	85.05	317	59.85

Sample_6	Arikara, Cheyenne River village (39ST1)	North America	382669	AD 1750-1775	130363	1747	1.34%	900	0.69%	177	35.66	2202	59.53	94	18.59	0	0
Sample_60	El Yaral, Cemetery 1	South America	59-1-E	AD 900 - 1200	96901	1521	1.57%	300	0.31%	65	3.94	180	18.56	1	1.20	0	0
Sample_64	El Yaral, Cemetery 2	South America	10291-205-1	AD 1158-1266	34780	539	1.55%	181	0.52%	2150	86.2597	398	66.2757	140	60.8713	59	50.1515
Sample_65	El Yaral, Cemetery 2	South America	10335-212-1	AD 900 - 1200	93026	1181	1.27%	577	0.62%	271	6.86	10106	39.74	0	0.00	0	0
Sample_76	Estuquina	South America	2279b	AD 1000-1476	63077	429	0.68%	524	0.83%	272	30.08	4526	66.17	735	21.86	0	0
Sample_79	Estuquina	South America	3215	AD 1000-1476	64079	628	0.98%	487	0.76%	239	19.99	5045	43.81	1043	16.23	0	0
Sample_81	Estuquina	South America	3644a	AD 1000-1476	117359	1397	1.19%	270	0.23%	62	11.08	439	38.36	27	9.59	0	0
Sample_82	Estuquina	South America	4165	AD 1000-1476	77640	893	1.15%	629	0.81%	438	69.46	1560	49.67	301	34.33	5	28.03
Sample_85	Estuquina	South America	4256	AD 1000-1476	55263	1299	2.35%	39	0.07%	160	5.08	246	19.39	3	4.28	0	0
Sample_87	Estuquina	South America	5859	AD 1000-1476	111009	2098	1.89%	644	0.58%	874	16.28	1995	45.17	32	2.76	0	0
Sample_89	Estuquina	South America	99407a	AD 1000-1476	60302	374	0.62%	84	0.14%	104	15.73	786	41.10	5	3.44	0	0
Sample_93	Schild	North America	SA-81	AD 1040-1407	134695	943	0.70%	364	0.27%	1	0.76	10	7.38	0	0	0	0
Sample_EB1.2	neg control				23270	391	1.68%	0	0.00%	3	1.24	5	5.13	1	1.44	1	6.52
Sample_EB1.25	neg control				70903	213	0.30%	0	0.00%	0	0	1	1.04	0	0	0	0
Sample_EB1.27	neg control				73996	969	1.31%	7	0.01%	3	2.16	1	1.33	0	0	0	0
Sample_EB1.3	neg control				51298	410	0.80%	0	0.00%	1	0.57	1	0.65	0	0	0	0
Sample_EB1.31	neg control				16371	345	2.11%	2	0.01%	1	0.65	2	1.15	0	0	0	0
Sample_EB2.2	neg control				20855	503	2.41%	0	0.00%	1	0.48	2	1.28	1	0.84	0	0
Sample_EB2.25	neg control				93591	449	0.48%	0	0.00%	0	0	0	0	0	0	0	0
Sample_EB2.27	neg control				49283	848	1.72%	5	0.01%	0	0	1	0.52	0	0	0	0
Sample_EB2.3	neg control				61711	506	0.82%	0	0.00%	0	0	2	1.75	0	0	0	0
Sample_EB2.31	neg control				16314	302	1.85%	3	0.02%	3	1.09	0	0	0	0	0	0
Sample_LB1.25	neg control				146552	747	0.51%	0	0.00%	1	0.46	6	3.73	1	0.76	0	0
Sample_LB1.27	neg control				36179	485	1.34%	4	0.01%	1	0.69	5	4.85	0	0	0	0
Sample_LB1_8_9_a_m	neg control				5625	135	2.40%	0	0.00%	0	0	0	0	0	0	0	0
Sample_LB1_8_9_Eve	neg control				56556	339	0.60%	0	0.00%	0	0	1	0.52	0	0	0	0
Sample_LB2.25	neg control				14925	54	0.36%	0	0.00%	0	0	0	0	0	0	0	0
Sample_LB2.27	neg control				26423	367	1.39%	3	0.01%	0	0	0	0	0	0	0	0
Sample_LB2_8_9_a_m	neg control				30809	471	1.53%	0	0.00%	1	0.42	4	3.96	0	0	0	0
Sample_LB2_8_9_Eve	neg control				51073	409	0.80%	0	0.00%	1	0.42	1	0.70	0	0	0	0

Appendix B
Table S5 - Information for all strains used in our phylogenetic and dating analyses

sample	SNP Calls	coverage(fold)	coverage(percent)	Source	Filename/Accession number	Lineage	Place of birth of the patient	Place of isolation of the strain
54U	756	31.35	87.35	This study		Lineage 6		
58U	697	20.45	83.05	This study		Lineage 6		
64U	740	22.73	90.41	This study		Lineage 6		
Pinipedi7011	944	54.25	96.53	This study		Lineage 6	-	animals
Pinipedi7739	946	53.38	96.18	This study		Lineage 6	-	animals
Pinipedi_G01222	949	14.41	91.19	This study		Lineage 6	-	animals
Pinipedi_G01491	1019	83.51	97.23	This study		Lineage 6	-	animals
Pinipedi_G01492	1013	134.06	97.69	This study		Lineage 6	-	animals
Pinipedi_G01498	835	48.32	97.2	This study		Lineage 6	-	animals
MicrotiERR027294	520	14.97	78.04	Sanger Institute	ERR027294	Lineage 6	-	animals
M.orygis	1415	129.04	95.59	NCBI SRA	ERS001592	Lineage 6	-	animals
capraeD028	956	10.21	80.31	Domogalla et al. 2013	SRR650227	Lineage 6	-	animals
capraeRW044	1171	19.53	90.63	Domogalla et al. 2013	SRR650219	Lineage 6	-	animals
capraeRW079	1145	17.39	88.58	Domogalla et al. 2013	SRR650221	Lineage 6	-	animals
bovis_ravenel	1087	25.98	95.29	NCBI SRA	SRS004666	Lineage 6	-	animals
chimpanzee_bacillus	1350	241.79	97.43	Coscolla et al. 2013	pending	Lineage 6	-	animals
Body68	484	22.28	96.16	Chan et al. 2013	SRP018736	Lineage 4	-	Hungarian mummy
L7_BTBS746	1038	43.95	96.88	Comas et al., Nat Genet, 2013	NG-5604_BTBS-746_sequence.fastq (ERP001731)	Lineage 7	Ethiopia	Ethiopia
L7_BTBS610	1011	54.24	97.14	Comas et al., Nat Genet, 2013	NG-5604_BTBS-610_sequence.fastq (ERP001731)	Lineage 7	Ethiopia	Ethiopia
L7_BTBH-935	1003	45.33	96.98	Comas et al., Nat Genet, 2013	NG-5604_BTBH-935_sequence.fastq (ERP001731)	Lineage 7	Ethiopia	Ethiopia
L7_BTBH-1012	1037	45.4	97	Comas et al., Nat Genet, 2013	NG-5604_BTBH-1012_sequence.fastq (ERP001731)	Lineage 7	Ethiopia	Ethiopia
L6_N0115	1072	169.81	95.58	Comas et al., Nat Genet, 2013	N0115b4_1/2.fastq (ERP001731)	Lineage 6	The Gambia	The Gambia
L6_N0099	1116	130.78	95.79	Comas et al., Nat Genet, 2013	N0099_1/2.fastq (ERP001731)	Lineage 6	-	The Netherlands
L6_N0098	1052	46.05	96.65	Comas et al., Nat Genet, 2013	NG-5604_N0098_sequence.fastq (ERP001731)	Lineage 6	-	The Netherlands
L6_N0092b4	1043	135.52	95.82	Comas et al., Nat Genet, 2013	N0092b4_1/2.fastq (ERP001731)	Lineage 6	The Gambia	The Gambia
L6_N0091	1125	160.3	96.43	Comas et al., Nat Genet, 2013	N0091b4_1/2.fastq (ERP001731)	Lineage 6	The Gambia	The gambia
L6_N0090	600	51.95	97.46	Comas et al., Nat Genet, 2013	NG-5604_N0090_sequence.fastq (ERP001731)	Lineage 6	The Gambia	The Gambia
L6_N0089	966	57.26	96.75	Comas et al., Nat Genet, 2013	NG-5604_N0089_sequence.fastq (ERP001731)	Lineage 6	The Gambia	The Gambia
L6_N0060	1013	54.54	96.53	Comas et al., Nat Genet, 2013	NG-5604_N0060_sequence.fastq (ERP001731)	Lineage 6	Senegal	San Francisco
L6_GM0981	944	29.93	94.8	Comas et al., Nat Genet, 2010	SRS004760	Lineage 6	The Gambia	The Gambia
L6_ABCG	1121	64.52	92.17	Comas et al., Nat Genet, 2013	MBO_BCGSSI1331_1/2.fastq (ERP001731)	Lineage 6	-	animals
L6_823602	906	39.51	84.16	Comas et al., Nat Genet, 2013	MAF_823602_1/2.fastq (ERP001731)	Lineage 6	Ghana	Ghana
L6_5468_02	1049	63.38	95.05	Comas et al., Nat Genet, 2013	MTB_5468_02_1/2.fastq (ERP001731)	Lineage 6	Ghana	Ghana
L6_541504	1072	103.18	95.77	Comas et al., Nat Genet, 2013	541504_1/2.fastq (ERP001731)	Lineage 6	Ghana	Ghana
L6_538302	1065	38.02	93.69	Comas et al., Nat Genet, 2013	MTB_5383_02_1/2.fastq (ERP001731)	Lineage 6	Ghana	Ghana
L6_533604	942	52	89.43	Comas et al., Nat Genet, 2013	MAF_533604_1/2.fastq (ERP001731)	Lineage 6	Ghana	Ghana
L6_414104	930	30.4	94.46	Comas et al., Nat Genet, 2010	SRS004764	Lineage 6	Sierra Leone	Sierra Leone
L5_DY26	942	44.05	95.3	Comas et al., Nat Genet, 2013	MTB_DY_26_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_DY21	930	54.4	96.37	Comas et al., Nat Genet, 2013	DY21_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_DY20	883	23.16	93.37	Comas et al., Nat Genet, 2013	MTB_DY_20_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_DY_135	934	101.63	96.03	Comas et al., Nat Genet, 2013	MTB_DY_135_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_553604	886	62.09	90.86	Comas et al., Nat Genet, 2013	MAF_553604_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_544404	865	36.24	95.16	Comas et al., Nat Genet, 2010	SRS004763	Lineage 5	Ghana	Ghana
L5_533304	841	51.96	90.32	Comas et al., Nat Genet, 2013	MAF_533304_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_349404	846	48.19	88.01	Comas et al., Nat Genet, 2013	MAF_349404_1/2.fastq (ERP001731)	Lineage 5	Sierra Leone	Sierra Leone
L5_348203	752	37.11	82.69	Comas et al., Nat Genet, 2013	MAF_348203_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_257702	920	101.98	95.39	Comas et al., Nat Genet, 2013	MTB_2577_02_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_256902	848	49.41	87.82	Comas et al., Nat Genet, 2013	MAF_256902_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_1449_02	916	73.88	95.36	Comas et al., Nat Genet, 2013	MTB_1449_02_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_1182103	866	32.94	95.13	Comas et al., Nat Genet, 2010	SRS004762	Lineage 5	Sierra Leone	Sierra Leone
L5_1048001	829	43.15	86.16	Comas et al., Nat Genet, 2013	MAF_1048001_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_1047301	831	49.33	85.91	Comas et al., Nat Genet, 2013	MAF_1047301_1/2.fastq (ERP001731)	Lineage 5	Ghana	Ghana
L5_1001003	931	102.79	96.28	Comas et al., Nat Genet, 2013	1001003_1/2.fastq (ERP001731)	Lineage 5	Sierra Leone	Sierra Leone
L4_X721	863	186.7	94.93	Comas et al., Nat Genet, 2012	SRS005448	Lineage 4	Mexico	Mexico
L4_X632	836	71.21	93.93	Comas et al., Nat Genet, 2012	SRS004831	Lineage 4	Mexico	Mexico
L4_X581	812	164.39	95.03	Comas et al., Nat Genet, 2012	SRS004841	Lineage 4	Mexico	Mexico
L4_V639EA	902	105.16	96.3	Comas et al., Nat Genet, 2013	MTB_V639EA_1/2.fastq (ERP001731)	Lineage 4	Vietnam	Vietnam
L4_V440EA	864	95.48	96.06	Comas et al., Nat Genet, 2013	MTB_V440EA_1/2.fastq (ERP001731)	Lineage 4	Vietnam	Vietnam
L4_V367IO	708	182.86	95.24	Comas et al., Nat Genet, 2013	MTB_5_V367IO_1/2.fastq (ERP001731)	Lineage 4	Vietnam	Vietnam

L4_V355EA	848	176.77	95.11	Comas et al., Nat Genet, 2013	MTB_1_V355EA_1/2.fastq (ERP001731)	Lineage 4	Vietnam	Vietnam
L4_V318EA	809	113.72	92.76	Comas et al., Nat Genet, 2013	MTB_V318EA_1/2.fastq (ERP001731)	Lineage 4	Vietnam	Vietnam
L4_V293EA	873	192.99	95.83	Comas et al., Nat Genet, 2013	MTB_3_V293EA_1/2.fastq (ERP001731)	Lineage 4	Vietnam	Vietnam
L4_V173IO	809	157.16	95.27	Comas et al., Nat Genet, 2013	MTB_2_V173IO_1/2.fastq (ERP001731)	Lineage 4	Vietnam	Vietnam
L4_N1057	800	62.83	97.63	Comas et al., Nat Genet, 2013	NG-5604_N1057_sequence.fastq (ERP001731)	Lineage 4	Nepal	Nepal
L4_N1052	838	60.85	97.68	Comas et al., Nat Genet, 2013	NG-5604_N1052_sequence.fastq (ERP001731)	Lineage 4	Nepal	Nepal
L4_N1015	808	60.06	97.55	Comas et al., Nat Genet, 2013	NG-5604_N1015_sequence.fastq (ERP001731)	Lineage 4	Nepal	Nepal
L4_N1008	784	49.27	97.45	Comas et al., Nat Genet, 2013	NG-5604_N1008_sequence.fastq (ERP001731)	Lineage 4	Nepal	Nepal
L4_N0185	788	60.1	97.72	Comas et al., Nat Genet, 2013	NG-5604_N0185_sequence.fastq (ERP001731)	Lineage 4	United Kingdom	London
L4_N0163	838	129.93	96.28	Comas et al., Nat Genet, 2013	MTB_N0163_1/2.fastq (ERP001731)	Lineage 4	Iran	San Francisco
L4_N0149	853	157.19	96.16	Comas et al., Nat Genet, 2013	MTB_N0149_1/2.fastq (ERP001731)	Lineage 4	USA	San Francisco
L4_N0148	825	61.45	96.59	Comas et al., Nat Genet, 2013	N0148mtb_1/2.fastq (ERP001731)	Lineage 4	USA	San Francisco
L4_N0146	894	164.39	96.11	Comas et al., Nat Genet, 2013	N0146b4_1/2.fastq (ERP001731)	Lineage 4	USA	San Francisco
L4_N0144	754	40.38	91.55	Comas et al., Nat Genet, 2013	MTB_N0144_1/2.fastq (ERP001731)	Lineage 4	China	San Francisco
L4_N0143	818	106.06	95.72	Comas et al., Nat Genet, 2013	MTB_N0143_1/2.fastq (ERP001731)	Lineage 4	USA	San Francisco
L4_N0142	767	77.12	93.44	Comas et al., Nat Genet, 2013	MTB_N0142_1/2.fastq (ERP001731)	Lineage 4	USA	San Francisco
L4_N0138	835	63.14	97.74	Comas et al., Nat Genet, 2013	NG-5604_N0138_sequence.fastq (ERP001731)	Lineage 4	USA	San Francisco
L4_N0137	850	24.63	94.84	Comas et al., Nat Genet, 2013	MTB_N0137_1/2.fastq (ERP001731)	Lineage 4	Mexico	San Francisco
L4_N0136	868	159.6	96.78	Comas et al., Nat Genet, 2013	N0136b4_1/2.fastq (ERP001731)	Lineage 4	USA	San Francisco
L4_N0135	791	45.81	94.31	Comas et al., Nat Genet, 2013	MTB_N0135_1/2.fastq (ERP001731)	Lineage 4	USA	San Francisco
L4_N0131	801	92.35	95.61	Comas et al., Nat Genet, 2013	MTB_N0131_1/2.fastq (ERP001731)	Lineage 4	USA	San Francisco
L4_N0126	811	159.8	95.95	Comas et al., Nat Genet, 2013	MTB_N0126_1/2.fastq (ERP001731)	Lineage 4	Colombia	San Francisco
L4_N0125	808	51.2	97.44	Comas et al., Nat Genet, 2013	NG-5604_N0125_sequence.fastq (ERP001731)	Lineage 4	Ethiopia	San Francisco
L4_N0120	850	127.5	96.07	Comas et al., Nat Genet, 2013	MTB_N0120_1/2.fastq (ERP001731)	Lineage 4	Puerto Rico	San Francisco
L4_N0109	871	139.72	96.64	Comas et al., Nat Genet, 2013	N0109b4_1/2.fastq (ERP001731)	Lineage 4	El Salvador	San Francisco
L4_N0107	819	142.26	96.7	Comas et al., Nat Genet, 2013	N0107b4_1/2.fastq (ERP001731)	Lineage 4	Guatemal	San Francisco
L4_N0103	850	97.61	96.46	Comas et al., Nat Genet, 2013	N0103mtb_1/2.fastq (ERP001731)	Lineage 4	Nicaragua	San Francisco
L4_N0101	843	144.55	95.8	Comas et al., Nat Genet, 2013	MTB_N0101_1/2.fastq (ERP001731)	Lineage 4	Nicaragua	San Francisco
L4_N0046b4	834	148.7	96.28	Comas et al., Nat Genet, 2013	N0046b4_1/2.fastq (ERP001731)	Lineage 4	Mexico	San Francisco
L4_N0011	757	35.81	96.92	Comas et al., Nat Genet, 2013	NG-5604_N0011_sequence.fastq (ERP001731)	Lineage 4	Mexico	San Francisco
L4_MT0005	875	153.84	97.16	Gardy at al., N Engl J Med, 2011	SRS084142	Lineage 4	-	Canada
L4_MT0001	871	136.28	97.04	Gardy at al., N Engl J Med, 2011	SRS074557	Lineage 4	-	Canada
L4_K37	737	41.56	95.03	Comas et al., Nat Genet, 2010	SRX002003	Lineage 4	Uganda	Uganda
L4_H37Rv	783	49.09	88.21	Comas et al., Nat Genet, 2013	MTB_H37RvDY_1/2.fastq (ERP001731)	Lineage 4	-	-
L4_erdman	489	20.32	89.73	NCBI SRA	SRS003328	Lineage 4	-	-
L4_DY8	861	63.46	96.53	Comas et al., Nat Genet, 2013	MTB_DY_8_1/2.fastq (ERP001731)	Lineage 4	Ghana	Ghana
L4_DY22	803	56.18	94.7	Comas et al., Nat Genet, 2013	MTB_DY_22_1/2.fastq (ERP001731)	Lineage 4	Ghana	Ghana
L4_DY195	863	147.66	96.65	Comas et al., Nat Genet, 2013	MTB_DY_195_1/2.fastq (ERP001731)	Lineage 4	Ghana	Ghana
L4_DY131	839	63.13	96.69	Comas et al., Nat Genet, 2013	MTB_DY_131_1/2.fastq (ERP001731)	Lineage 4	Liberia	Ghana
L4_DY_167	824	75.78	95.5	Comas et al., Nat Genet, 2013	MTB_DY_167_1/2.fastq (ERP001731)	Lineage 4	Ghana	Ghana
L4_CDC1551E	831	96.67	95.35	Comas et al., Nat Genet, 2012	SRS005450	Lineage 4	USA	United States
L4_BTBS458	775	51.15	96.89	Comas et al., Nat Genet, 2013	NG-5604_BTBS-458_sequence.fastq (ERP001731)	Lineage 4	Ethiopia	Ethiopia
L4_BTBS101	787	55.91	97.15	Comas et al., Nat Genet, 2013	NG-5604_BTBS-101_sequence.fastq (ERP001731)	Lineage 4	Ethiopia	Ethiopia
L4_BTBH587	773	42.71	97.42	Comas et al., Nat Genet, 2013	NG-5604_BTBH-587_sequence.fastq (ERP001731)	Lineage 4	Ethiopia	Ethiopia
L4_549204	751	45.78	88.47	Comas et al., Nat Genet, 2013	MTB_549204_1/2.fastq (ERP001731)	Lineage 4	Ghana	Ghana
L4_478304	765	31.36	94.93	Comas et al., Nat Genet, 2010	SRS004761	Lineage 4	Sierra Leone	Sierra Leone
L4_267903	852	74.5	95.94	Comas et al., Nat Genet, 2013	MTB_2679_03_1/2.fastq (ERP001731)	Lineage 4	Germany	Hamburg
L4_233602	815	111.45	95.96	Comas et al., Nat Genet, 2013	MTB_2336_02_1/2.fastq (ERP001731)	Lineage 4	Germany	Hamburg
L4_231806	821	69.53	96	Comas et al., Nat Genet, 2013	MTB_2318_06_1/2.fastq (ERP001731)	Lineage 4	Germany	Germany
L4_2307_99	810	102.3	96.42	Comas et al., Nat Genet, 2013	MTB_2307_99_1/2.fastq (ERP001731)	Lineage 4	Uganda	Uganda
L4_219799	795	91.45	95.4	Comas et al., Nat Genet, 2013	MTB_2197_99_1/2.fastq (ERP001731)	Lineage 4	Uganda	Uganda
L4_217399	808	104.55	95.69	Comas et al., Nat Genet, 2013	MTB_2173_99_1/2.fastq (ERP001731)	Lineage 4	Uganda	Uganda
L4_157108	838	116.71	96.47	Comas et al., Nat Genet, 2013	MTB_1571_08_1/2.fastq (ERP001731)	Lineage 4	Tanzania	Tanzania
L4_155008	829	44.71	94.35	Comas et al., Nat Genet, 2013	MTB_1550_08_1/2.fastq (ERP001731)	Lineage 4	Tanzania	Tanzania
L4_141702	692	49.88	82.87	Comas et al., Nat Genet, 2013	MTB_141702_1/2.fastq (ERP001731)	Lineage 4	Ghana	Ghana
L4_10264_03	854	117.35	96.42	Comas et al., Nat Genet, 2013	MTB_10264_03_1/2.fastq (ERP001731)	Lineage 4	Germany	Hamburg
L3_SG1	715	19.84	85.76	Comas et al., Nat Genet, 2013	MTB_SG1_12.fastq (ERP001731)	Lineage 3	India	San Francisco
L3_rus15	962	235.59	97.86	Casali et al., Genome Res, 2012	ERS003250	Lineage 3	-	Russia
L3_N37	821	48.21	91.01	Comas et al., Nat Genet, 2013	MTB_N37_1/2.fastq (ERP001731)	Lineage 3	Pakistan	San Francisco
L3_N24	921	62.95	96.89	Comas et al., Nat Genet, 2013	MTB_N24_read_1/2.fastq (ERP001731)	Lineage 3	India	San Francisco
L3_N1058	871	59.08	97.8	Comas et al., Nat Genet, 2013	NG-5604_N1058_sequence.fastq (ERP001731)	Lineage 3	Nepal	Nepal
L3_N1040	876	50.79	97.57	Comas et al., Nat Genet, 2013	NG-5604_N1040_sequence.fastq (ERP001731)	Lineage 3	Mumbai, India	India
L3_N1032	875	60.74	97.69	Comas et al., Nat Genet, 2013	NG-5604_N1032_sequence.fastq (ERP001731)	Lineage 3	Nepal	Nepal

L3_N1024	913	63.87	97.54	Comas et al., Nat Genet, 2013	NG-5604_N1024_sequence.fastq (ERP001731)	Lineage 3	Portugal	Switzerland
L3_N1022	886	64.07	97.38	Comas et al., Nat Genet, 2013	NG-5604_N1022_sequence.fastq (ERP001731)	Lineage 3	Nepal	Nepal
L3_N1014	889	62.81	97.85	Comas et al., Nat Genet, 2013	NG-5604_N1014_sequence.fastq (ERP001731)	Lineage 3	Nepal	Nepal
L3_N1007	874	63.66	97.77	Comas et al., Nat Genet, 2013	NG-5604_N1007_sequence.fastq (ERP001731)	Lineage 3	Nepal	Nepal
L3_N04	917	54.08	96.49	Comas et al., Nat Genet, 2013	MTB_N04_read_1/2.fastq (ERP001731)	Lineage 3	India	San Francisco
L3_N0197	887	122.05	97.21	Comas et al., Nat Genet, 2013	N0197_1/2.fastq (ERP001731)	Lineage 3	United Kingdom	London
L3_N0056	929	118.32	95.64	Comas et al., Nat Genet, 2013	MTB_N0056_1/2.fastq (ERP001731)	Lineage 3	Vietnam	San Francisco
L3_N0054	909	125.28	95.57	Comas et al., Nat Genet, 2013	MTB_N0054_1/2.fastq (ERP001731)	Lineage 3	Ethiopia	San Francisco
L3_N0033	886	53.43	97.2	Comas et al., Nat Genet, 2013	NG-5604_N0033_sequence.fastq (ERP001731)	Lineage 3	Pakistan	San Francisco
L3_K49	846	49.01	96.27	Comas et al., Nat Genet, 2010	SRX002002	Lineage 3	Tanzania	San Francisco
L3_BTBH273	891	53.87	97.67	Comas et al., Nat Genet, 2013	NG-5604_BTBH-273_sequence.fastq (ERP001731)	Lineage 3	Ethiopia	Ethiopia
L3_910079	744	24.24	94.93	Comas et al., Nat Genet, 2010	SRS004758	Lineage 3	Ethiopia	San Francisco
L3_858405	892	58.64	96.16	Comas et al., Nat Genet, 2013	MTB_8584_05_1/2.fastq (ERP001731)	Lineage 3	Afghanistan	Hamburg
L3_855B	930	155.59	96.98	Comas et al., Nat Genet, 2013	855Bb4_1/2.fastq (ERP001731)	Lineage 3	Nepal	Nepal
L3_793601	801	56.07	91.11	Comas et al., Nat Genet, 2013	MTB_793601_1/2.fastq (ERP001731)	Lineage 3	Turkmenistan/Uzbekistan	Turkmenistan/Uzbekistan
L3_751B	824	147.26	96.54	Comas et al., Nat Genet, 2013	751Bb4_1/2.fastq (ERP001731)	Lineage 3	Nepal	Nepal
L3_657005	881	81.82	95.93	Comas et al., Nat Genet, 2013	MTB_6570_05_1/2.fastq (ERP001731)	Lineage 3	Afghanistan	Hamburg
L3_597805	777	51.53	86.99	Comas et al., Nat Genet, 2013	MTB_597805_1/2.fastq (ERP001731)	Lineage 3	Afghanistan	Hamburg
L3_580605	756	44.85	88.48	Comas et al., Nat Genet, 2013	MTB_580605_1/2.fastq (ERP001731)	Lineage 3	Sri Lanka	Hamburg
L3_4497_06	939	73.03	96.07	Comas et al., Nat Genet, 2013	MTB_4497_06_1/2.fastq (ERP001731)	Lineage 3	Pakistan	Hamburg
L3_296904	712	40.46	83.8	Comas et al., Nat Genet, 2013	MTB_296904_1/2.fastq (ERP001731)	Lineage 3	Afghanistan	Hamburg
L3_282801	735	30.48	84.55	Comas et al., Nat Genet, 2013	MTB_282801_1/2.fastq (ERP001731)	Lineage 3	Turkei	Hamburg
L3_173305	881	65	95.65	Comas et al., Nat Genet, 2013	MTB_1733_05_1/2.fastq (ERP001731)	Lineage 3	Afghanistan	Hamburg
L3_162908	909	41.14	96.48	Comas et al., Nat Genet, 2013	MTB_1629_08_1/2.fastq (ERP001731)	Lineage 3	Tanzania	Tanzania
L3_159008	894	37.72	95.56	Comas et al., Nat Genet, 2013	MTB_1590_08_1/2.fastq (ERP001731)	Lineage 3	Tanzania	Tanzania
L3_1558_08	932	124.42	96.5	Comas et al., Nat Genet, 2013	MTB_1558_08_1/2.fastq (ERP001731)	Lineage 3	Tanzania	Tanzania
L3_138001	771	38.05	85.55	Comas et al., Nat Genet, 2013	MTB_138007_1/2.fastq (ERP001731)	Lineage 3	Iran	Hamburg
L3_1180Cb4	906	174.71	96.93	Comas et al., Nat Genet, 2013	1180Cb4_1/2.fastq (ERP001731)	Lineage 3	Nepal	Nepal
L3_1016204	812	62.72	88.34	Comas et al., Nat Genet, 2013	MTB_1016204_1/2.fastq (ERP001731)	Lineage 3	Eritrea	Hamburg
L2_V374BJ	909	45.24	95.47	Comas et al., Nat Genet, 2013	MTB_V374BJ_1/2.fastq (ERP001731)	Lineage 2	Vietnam	Vietnam
L2_V212BJ	941	174.46	95.4	Comas et al., Nat Genet, 2013	MTB_6_V212BJ_1/2.fastq (ERP001731)	Lineage 2	Vietnam	Vietnam
L2_V119BJ	919	150.23	95.16	Comas et al., Nat Genet, 2013	MTB_7_V119BJ_1/2.fastq (ERP001731)	Lineage 2	Vietnam	Vietnam
L2_rus16	988	235.97	96.88	Casali et al., Genome Res, 2012	ERS003237	Lineage 2	-	Russia
L2_rus14	955	237.07	96.63	Casali et al., Genome Res, 2012	ERS003236	Lineage 2	-	Russia
L2_N1051	888	66.81	96.95	Comas et al., Nat Genet, 2013	NG-5604_N1051_sequence.fastq (ERP001731)	Lineage 2	Nepal	Nepal
L2_N1037	900	52.35	96.97	Comas et al., Nat Genet, 2013	NG-5604_N1037_sequence.fastq (ERP001731)	Lineage 2	Mumbai, India	India
L2_N1001	891	57.3	97.16	Comas et al., Nat Genet, 2013	NG-5604_N1001_sequence.fastq (ERP001731)	Lineage 2	Nepal	Nepal
L2_N0158	948	122.02	96.23	Comas et al., Nat Genet, 2013	N0158_1/2.fastq (ERP001731)	Lineage 2	China	San Francisco
L2_N0151	936	139.62	96.62	Comas et al., Nat Genet, 2013	N0151b4_1/2.fastq (ERP001731)	Lineage 2	China	San Francisco
L2_N0150	973	151.82	96.31	Comas et al., Nat Genet, 2013	N0150_1/2.fastq (ERP001731)	Lineage 2	China	San Francisco
L2_N0130b4	957	184.21	96.31	Comas et al., Nat Genet, 2013	N0130b4_1/2.fastq (ERP001731)	Lineage 2	Cambodia	San Francisco
L2_N0128	874	52.43	96.88	Comas et al., Nat Genet, 2013	NG-5604_N0128_sequence.fastq (ERP001731)	Lineage 2	Vietnam	San Francisco
L2_N0094	950	161.51	96.58	Comas et al., Nat Genet, 2013	N0094b4_1/2.fastq (ERP001731)	Lineage 2	China	San Francisco
L2_N0053	981	161	96.19	Comas et al., Nat Genet, 2013	N0053b4_1/2.fastq (ERP001731)	Lineage 2	China	San Francisco
L2_N0051	947	185.76	95.9	Comas et al., Nat Genet, 2013	N0051b4_1/2.fastq (ERP001731)	Lineage 2	China	San Francisco
L2_N0050	878	51.74	96.71	Comas et al., Nat Genet, 2013	NG-5604_N0050_sequence.fastq (ERP001731)	Lineage 2	China	San Francisco
L2_N0044	949	156.61	96.76	Comas et al., Nat Genet, 2013	N0044b4_1/2.fastq (ERP001731)	Lineage 2	Singapore	San Francisco
L2_N0041	926	158.59	96.81	Comas et al., Nat Genet, 2013	N0041b4_1/2.fastq (ERP001731)	Lineage 2	Vietnam	San Francisco
L2_N0039	891	60.38	95.42	Comas et al., Nat Genet, 2013	MTB_N0039_1/2.fastq (ERP001731)	Lineage 2	Laos	San Francisco
L2_N0034	923	129.28	96.37	Comas et al., Nat Genet, 2013	N0034b4_1/2.fastq (ERP001731)	Lineage 2	China	San Francisco
L2_N0020	918	22.83	95.33	Comas et al., Nat Genet, 2013	MTB_N0020_1/2.fastq (ERP001731)	Lineage 2	Mongolia	San Francisco
L2_N0017	833	54.91	96.79	Comas et al., Nat Genet, 2013	NG-5604_N0017_sequence.fastq (ERP001731)	Lineage 2	South Korea	San Francisco
L2_N0010	912	40.67	95.33	Comas et al., Nat Genet, 2013	MTB_N0010_1/2.fastq (ERP001731)	Lineage 2	Indonesia	San Francisco
L2_N0008	882	50.11	96.96	Comas et al., Nat Genet, 2013	NG-5604_N0008_sequence.fastq (ERP001731)	Lineage 2	Vietnam	San Francisco
L2_N0005	928	123.27	96.15	Comas et al., Nat Genet, 2013	N0005_1/2.fastq (ERP001731)	Lineage 2	Indonesia	San Francisco
L2_N0003	933	132.96	96.88	Comas et al., Nat Genet, 2013	N0003b4_1/2.fastq (ERP001731)	Lineage 2	Vietnam	San Francisco
L2_N0002	957	136.5	96.61	Comas et al., Nat Genet, 2013	N0002b4_1/2.fastq (ERP001731)	Lineage 2	Cambodia	San Francisco
L2_M4100A	755	21.27	94.2	Comas et al., Nat Genet, 2010	SRS004757	Lineage 2	South Korea	San Francisco
L2_GT649	867	196.29	96.02	Comas et al., Nat Genet, 2013	MTB_GT_649_1/2.fastq (ERP001731)	Lineage 2	Vietnam	Vietnam
L2_GT345	859	47.51	94.81	Comas et al., Nat Genet, 2013	MTB_GT_345_1/2.fastq (ERP001731)	Lineage 2	Vietnam	Vietnam
L2_GT333	935	57.53	95.82	Comas et al., Nat Genet, 2013	MTB_GT_333_1/2.fastq (ERP001731)	Lineage 2	Vietnam	Vietnam
L2_GT_411	932	63.04	95.3	Comas et al., Nat Genet, 2013	MTB_GT_411_1/2.fastq (ERP001731)	Lineage 2	Vietnam	Vietnam
L2_GQ904	896	58.41	95.68	Comas et al., Nat Genet, 2013	GQ904_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai

L2_GQ859	946	151.21	96.08	Comas et al., Nat Genet, 2013	GQ859_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ856	709	22.71	80.76	Comas et al., Nat Genet, 2013	GQ856_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ762	863	133.79	95.16	Comas et al., Nat Genet, 2013	GQ762_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ74-3	911	43.52	95.17	Comas et al., Nat Genet, 2013	GQ74-3_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ73	882	80.36	95.8	Comas et al., Nat Genet, 2013	GQ73_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ657	918	140.28	96.04	Comas et al., Nat Genet, 2013	GQ657_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ366	919	228.26	96.8	Comas et al., Nat Genet, 2013	GQ366_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ254	932	135.1	96.19	Comas et al., Nat Genet, 2013	GQ254_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ229	932	105.34	95.81	Comas et al., Nat Genet, 2013	GQ229_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ1973	940	124.58	96.1	Comas et al., Nat Genet, 2013	GQ1973_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ1885	819	60.31	92.52	Comas et al., Nat Genet, 2013	GQ1885_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ1645	746	22.39	83.41	Comas et al., Nat Genet, 2013	GQ1645_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ1605	832	52.88	92.8	Comas et al., Nat Genet, 2013	GQ1605_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ1597	852	45.51	91.48	Comas et al., Nat Genet, 2013	GQ1597_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ1580	941	118.56	95.99	Comas et al., Nat Genet, 2013	GQ1580_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ1439	959	137.23	95.85	Comas et al., Nat Genet, 2013	GQ1439_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ1335	975	212.51	96.56	Comas et al., Nat Genet, 2013	GQ1335_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ1331	976	213.63	96.4	Comas et al., Nat Genet, 2013	GQ1331_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ1164	915	152.1	95.82	Comas et al., Nat Genet, 2013	GQ1164_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ1020	927	144.11	95.92	Comas et al., Nat Genet, 2013	GQ1020_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ10-3	921	66.31	95.92	Comas et al., Nat Genet, 2013	GQ10-3_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ-50	974	202.62	96.38	Comas et al., Nat Genet, 2013	GQ-50_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ-1972	932	96.16	95.93	Comas et al., Nat Genet, 2013	GQ-1972_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ-1438	933	123.07	96.45	Comas et al., Nat Genet, 2013	GQ-1438_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ-1343	939	109.38	96.07	Comas et al., Nat Genet, 2013	GQ-1343_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ-1168	954	203.51	96.52	Comas et al., Nat Genet, 2013	GQ-1168_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_GQ-1165	967	213.69	96.84	Comas et al., Nat Genet, 2013	GQ-1165_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_861	952	162.45	96.36	Comas et al., Nat Genet, 2013	861_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_759	934	192.16	96.64	Comas et al., Nat Genet, 2013	759_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_72	966	206.79	96.62	Comas et al., Nat Genet, 2013	72_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_1864	963	207.32	96.65	Comas et al., Nat Genet, 2013	1864_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_1336	994	156.23	96.27	Comas et al., Nat Genet, 2013	1336_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L2_1027	934	143.02	96.19	Comas et al., Nat Genet, 2013	1027_1/2.fastq (ERP001731)	Lineage 2	China	Shanghai
L1_V372IO	879	84.1	93.28	Comas et al., Nat Genet, 2013	MTB_V372IO_1/2.fastq (ERP001731)	Lineage 1	Vietnam	Vietnam
L1_V346IO	874	111.72	92.82	Comas et al., Nat Genet, 2013	MTB_V346IO_1/2.fastq (ERP001731)	Lineage 1	Vietnam	Vietnam
L1_V232IO	789	28.59	84.36	Comas et al., Nat Genet, 2013	MTB_V232IO_1/2.fastq (ERP001731)	Lineage 1	Vietnam	Vietnam
L1_T92	746	21.72	92.9	Comas et al., Nat Genet, 2010	SRS004759	Lineage 1	The Philippines	San Francisco
L1_T83	831	54.33	92.46	Comas et al., Nat Genet, 2013	MTB_T83_12.fastq (ERP001731)	Lineage 1	Vietnam	San Francisco
L1_N73	849	65.88	93.89	Comas et al., Nat Genet, 2013	MTB_N73_1/2.fastq (ERP001731)	Lineage 1	India	San Francisco
L1_N72	927	48.39	96.04	Comas et al., Nat Genet, 2013	MTB_N72_read_1/2.fastq (ERP001731)	Lineage 1	India	San Francisco
L1_N70	924	47.18	96.14	Comas et al., Nat Genet, 2013	MTB_N70_read_1/2.fastq (ERP001731)	Lineage 1	India	San Francisco
L1_N1009	911	54.4	97.56	Comas et al., Nat Genet, 2013	NG-5604_N1009_sequence.fastq (ERP001731)	Lineage 1	Nepal	Nepal
L1_N1006	890	63.27	97.62	Comas et al., Nat Genet, 2013	NG-5604_N1006_sequence.fastq (ERP001731)	Lineage 1	Nepal	Nepal
L1_N1004	890	49.88	96.92	Comas et al., Nat Genet, 2013	NG-5604_N1004_sequence.fastq (ERP001731)	Lineage 1	Nepal	Nepal
L1_N0203	888	63.48	97.59	Comas et al., Nat Genet, 2013	NG-5604_N0203_sequence.fastq (ERP001731)	Lineage 1	Sri Lanka	London/Homerton
L1_N0196	907	106.05	96.56	Comas et al., Nat Genet, 2013	N0196_1/2.fastq (ERP001731)	Lineage 1	Somalia	London/Homerton
L1_N0182	859	65.4	97.53	Comas et al., Nat Genet, 2013	NG-5604_N0182_sequence.fastq (ERP001731)	Lineage 1	Malaysia	London/Homerton
L1_N0141b4	959	179.1	96.85	Comas et al., Nat Genet, 2013	N0141b4_1/2.fastq (ERP001731)	Lineage 1	The Philippines	San Francisco
L1_N0134	941	163.9	96.64	Comas et al., Nat Genet, 2013	N0134b4_1/2.fastq (ERP001731)	Lineage 1	The Philippines	San Francisco
L1_N0132	933	35.54	95.86	Comas et al., Nat Genet, 2013	MTB_N0132_1/2.fastq (ERP001731)	Lineage 1	The Philippines	San Francisco
L1_N0127	957	170.68	96.58	Comas et al., Nat Genet, 2013	N0127b4_1/2.fastq (ERP001731)	Lineage 1	China	San Francisco
L1_N0079	931	41.06	96.67	Comas et al., Nat Genet, 2013	NG-5604_N0079_sequence.fastq (ERP001731)	Lineage 1	Vietnam	San Francisco
L1_N0067	920	72.61	97.38	Comas et al., Nat Genet, 2013	NG-5604_N00067_sequence.fastq (ERP001731)	Lineage 1	China	San Francisco
L1_N0065	895	60.69	97.14	Comas et al., Nat Genet, 2013	NG-5604_N0065_sequence.fastq (ERP001731)	Lineage 1	China	San Francisco
L1_N0062	864	58.06	97.11	Comas et al., Nat Genet, 2013	NG-5604_N0062_sequence.fastq (ERP001731)	Lineage 1	Cambodia	San Francisco
L1_N0043	842	16.54	90.12	Comas et al., Nat Genet, 2013	MTB_N0043_1/2.fastq (ERP001731)	Lineage 1	Burma	San Francisco
L1_N0014	880	70.66	94.35	Comas et al., Nat Genet, 2013	MTB_N0014_1/2.fastq (ERP001731)	Lineage 1	Thailand	San Francisco
L1_K93	864	44.44	96.18	Comas et al., Nat Genet, 2010	SRX002005	Lineage 1	Tanzania	San Francisco
L1_K67	904	49.08	95.82	Comas et al., Nat Genet, 2010	SRX002004	Lineage 1	Comoro Islands	San Francisco
L1_K21	929	51.14	96.77	Comas et al., Nat Genet, 2010	SRX002001	Lineage 1	Tanzania	San Francisco
L1_GT281	908	157.99	96.36	Comas et al., Nat Genet, 2013	MTB_GT_281_1/2.fastq (ERP001731)	Lineage 1	Vietnam	Vietnam
L1_DY28	953	175.19	96.55	Comas et al., Nat Genet, 2013	MTB_DY_28_1/2.fastq (ERP001731)	Lineage 1	Ghana	Ghana
L1_BTBS493	906	55.35	97.13	Comas et al., Nat Genet, 2013	NG-5604_BTBS-493_sequence.fastq (ERP001731)	Lineage 1	Ethiopia	Ethiopia

L1_BTBS280	928	50.66	97.3	Comas et al., Nat Genet, 2013	NG-5604_BTBS-280_sequence.fastq (ERP001731)	Lineage 1	Ethiopia	Ethiopia
L1_950545	545	30.07	95.26	Comas et al., Nat Genet, 2010	SRS004759	Lineage 1	Laos	San Francisco
L1_94701	933	96.1	96.01	Comas et al., Nat Genet, 2013	MTB_947_01_1/2.fastq (ERP001731)	Lineage 1	Afghanistan	Hamburg
L1_745902	966	79.74	95.44	Comas et al., Nat Genet, 2013	MTB_7459_02_1/2.fastq (ERP001731)	Lineage 1	Burkina Faso	Hamburg
L1_600603	933	74.63	96.57	Comas et al., Nat Genet, 2013	MTB_6006_03_1/2.fastq (ERP001731)	Lineage 1	Sri Lanka	Hamburg
L1_539606	885	25.7	94.44	Comas et al., Nat Genet, 2013	MTB_5396_06_1/2.fastq (ERP001731)	Lineage 1	Singapore	Hamburg
L1_468303	924	69.92	96.21	Comas et al., Nat Genet, 2013	MTB_4683_03_1/2.fastq (ERP001731)	Lineage 1	Serbia	Hamburg
L1_179703	855	52.1	91.42	Comas et al., Nat Genet, 2013	MTB_179703_1/2.fastq (ERP001731)	Lineage 1	India	Hamburg
L1_157508	932	121.61	95.63	Comas et al., Nat Genet, 2013	MTB_1575_08_1/2.fastq (ERP001731)	Lineage 1	Tanzania	Tanzania
L1_1571_99	944	87.41	95.97	Comas et al., Nat Genet, 2013	MTB_1571_99_1/2.fastq (ERP001731)	Lineage 1	Uganda	Uganda
L1_1277803	884	87.72	95.51	Comas et al., Nat Genet, 2013	MTB_12778_03_1/2.fastq (ERP001731)	Lineage 1	Vietnam	Hamburg
L1_1105103	895	74.06	95.86	Comas et al., Nat Genet, 2013	MTB_11051_03_1/2.fastq (ERP001731)	Lineage 1	Thailand	Hamburg
L1_10786_02	909	98.49	95.92	Comas et al., Nat Genet, 2013	MTB_10786_02_1/2.fastq (ERP001731)	Lineage 1	Thailand	Hamburg
Outgroup (M. canettii)	10821	62.24	92.98	Comas et al., Nat genetics, 2010	SRX002429	Outgroup (M. canetti)	Djibouti	Djibouti

Appendix B
Table S6 - mapping summary
All mappings reported here were performed with the stricter parameters, (see Supplementary text), and are mapped against the hypothetical ancestor (Comas et al, 2010)

Array captured libraries

Sample	Reference	#reads	#mapped reads	% reads mapped	#mapped reads (post rmdup)	av. dup. factor	fold coverage	% of H37rV reads in shotgun	Enrichment efficiency
54	mtbc_anc	81257798	30080214	37.02%	2980807	10.09	31.35	0.38%	98.66
58	mtbc_anc	20573490	5999216	29.16%	1944666	3.08	20.45	0.32%	91.27
64	mtbc_anc	50483372	8683999	17.20%	2479743	3.50	22.73	0.53%	32.22

Negative controls mapped against the constructed ancestor (Comas et al, 2010)

Sample	Reference	#reads	#mapped reads	%mapped reads	#mapped reads rmdup	dup. Factor
EB1.20U	mtbc_anc	110088	224	0.20%	215	1.04
EB1.31U	mtbc_anc	221273	728	0.33%	614	1.19
EB2.20U	mtbc_anc	177063	502	0.28%	426	1.18
EB2.31U	mtbc_anc	409545	967	0.24%	881	1.10
LB1U	mtbc_anc	298189	686	0.23%	590	1.16
LB2U	mtbc_anc	274185	515	0.19%	506	1.02

Appendix B

Table S8 - SNPs, and coverages for each genome when mapped to the constructed common ancestor (Comas et al, 2010) with "strict" mapping (see suppl text)

sample	SNP Calls	coverage (fold)	coverage (percent)	sample	SNP Calls	coverage (fold)	coverage (percent)
54U	756	31.35	87.35	L6_538302	1065	38.02	93.69
58U	697	20.45	83.05	L6_533604	942	52	89.43
64U	740	22.73	90.41	L6_414104	930	30.4	94.46
Pinipedi7011	944	54.25	96.53	L5_DY26	942	44.05	95.3
Pinipedi_G01222	949	14.41	91.19	L5_DY21	930	54.4	96.37
Pinipedi7739	946	53.38	96.18	L5_DY20	883	23.16	93.37
Pinipedi_G01491	1019	83.51	97.23	L5_DY_135	934	101.63	96.03
Pinipedi_G01498	835	48.32	97.2	L5_553604	886	62.09	90.86
Pinipedi_G01492	1013	134.06	97.69	L5_544404	865	36.24	95.16
MicrotiERR027294	520	14.97	78.04	L5_533304	841	51.96	90.32
M.orygys	1415	129.04	95.59	L5_349404	846	48.19	88.01
capraeD028	956	10.21	80.31	L5_348203	752	37.11	82.69
capraeRW044	1171	19.53	90.63	L5_257702	920	101.98	95.39
capraeRW079	1145	17.39	88.58	L5_256902	848	49.41	87.82
bovis_ravenel	1087	25.98	95.29	L5_1449_02	916	73.88	95.36
L6_ABCG	1121	64.52	92.17	L5_1182103	866	32.94	95.13
chimpanzee_bacillus	1350	241.79	97.43	L5_1048001	829	43.15	86.16
Body68	484	22.28	96.16	L5_1047301	831	49.33	85.91
L7_BTBS746	1038	43.95	96.88	L5_1001003	931	102.79	96.28
L7_BTBS610	1011	54.24	97.14	L4_X721	863	186.7	94.93
L7_BTBH-935	1003	45.33	96.98	L4_X632	836	71.21	93.93
L7_BTBH-1012	1037	45.4	97	L4_X581	812	164.39	95.03
L6_N0115	1072	169.81	95.58	L4_V639EA	902	105.16	96.3
L6_N0099	1116	130.78	95.79	L4_V440EA	864	95.48	96.06
L6_N0098	1052	46.05	96.65	L4_V367IO	708	182.86	95.24
L6_N0092b4	1043	135.52	95.82	L4_V355EA	848	176.77	95.11
L6_N0091	1125	160.3	96.43	L4_V318EA	809	113.72	92.76
L6_N0090	600	51.95	97.46	L4_V293EA	873	192.99	95.83
L6_N0089	966	57.26	96.75	L4_V173IO	809	157.16	95.27
L6_N0060	1013	54.54	96.53	L4_N1057	800	62.83	97.63
L6_GM0981	944	29.93	94.8	L4_N1052	838	60.85	97.68
L6_823602	906	39.51	84.16	L4_N1015	808	60.06	97.55
L6_5468_02	1049	63.38	95.05	L4_N1008	784	49.27	97.45

L6_541504	1072	103.18	95.77		L4_N0185	788	60.1	97.72
L4_N0163	838	129.93	96.28		L4_233602	815	111.45	95.96
L4_N0149	853	157.19	96.16		L4_231806	821	69.53	96
L4_N0148	825	61.45	96.59		L4_2307_99	810	102.3	96.42
L4_N0146	894	164.39	96.11		L4_219799	795	91.45	95.4
L4_N0144	754	40.38	91.55		L4_217399	808	104.55	95.69
L4_N0143	818	106.06	95.72		L4_157108	838	116.71	96.47
L4_N0142	767	77.12	93.44		L4_155008	829	44.71	94.35
L4_N0138	835	63.14	97.74		L4_141702	692	49.88	82.87
L4_N0137	850	24.63	94.84		L4_10264_03	854	117.35	96.42
L4_N0136	868	159.6	96.78		L3_SG1	715	19.84	85.76
L4_N0135	791	45.81	94.31		L3_rus15	962	235.59	97.86
L4_N0131	801	92.35	95.61		L3_N37	821	48.21	91.01
L4_N0126	811	159.8	95.95		L3_N24	921	62.95	96.89
L4_N0125	808	51.2	97.44		L3_N1058	871	59.08	97.8
L4_N0120	850	127.5	96.07		L3_N1040	876	50.79	97.57
L4_N0109	871	139.72	96.64		L3_N1032	875	60.74	97.69
L4_N0107	819	142.26	96.7		L3_N1024	913	63.87	97.54
L4_N0103	850	97.61	96.46		L3_N1022	886	64.07	97.38
L4_N0101	843	144.55	95.8		L3_N1014	889	62.81	97.85
L4_N0046b4	834	148.7	96.28		L3_N1007	874	63.66	97.77
L4_N0011	757	35.81	96.92		L3_N04	917	54.08	96.49
L4_MT0005	875	153.84	97.16		L3_N0197	887	122.05	97.21
L4_MT0001	871	136.28	97.04		L3_N0056	929	118.32	95.64
L4_K37	737	41.56	95.03		L3_N0054	909	125.28	95.57
L4_H37Rv	783	49.09	88.21		L3_N0033	886	53.43	97.2
L4_erdman	489	20.32	89.73		L3_K49	846	49.01	96.27
L4_DY8	861	63.46	96.53		L3_BTBH273	891	53.87	97.67
L4_DY22	803	56.18	94.7		L3_910079	744	24.24	94.93
L4_DY195	863	147.66	96.65		L3_858405	892	58.64	96.16
L4_DY131	839	63.13	96.69		L3_855B	930	155.59	96.98
L4_DY_167	824	75.78	95.5		L3_793601	801	56.07	91.11
L4_CDC1551E	831	96.67	95.35		L3_751B	824	147.26	96.54
L4_BTBS458	775	51.15	96.89		L3_657005	881	81.82	95.93
L4_BTBS101	787	55.91	97.15		L3_597805	777	51.53	86.99
L4_BTBH587	773	42.71	97.42		L3_580605	756	44.85	88.48
L4_549204	751	45.78	88.47		L3_4497_06	939	73.03	96.07
L4_478304	765	31.36	94.93		L3_296904	712	40.46	83.8

L4_267903	852	74.5	95.94	L3_282801	735	30.48	84.55
L3_173305	881	65	95.65	L2_GT333	935	57.53	95.82
L3_162908	909	41.14	96.48	L2_GT_411	932	63.04	95.3
L3_159008	894	37.72	95.56	L2_GQ904	896	58.41	95.68
L3_1558_08	932	124.42	96.5	L2_GQ859	946	151.21	96.08
L3_138001	771	38.05	85.55	L2_GQ856	709	22.71	80.76
L3_1180Cb4	906	174.71	96.93	L2_GQ762	863	133.79	95.16
L3_1016204	812	62.72	88.34	L2_GQ74-3	911	43.52	95.17
L2_V374BJ	909	45.24	95.47	L2_GQ73	882	80.36	95.8
L2_V212BJ	941	174.46	95.4	L2_GQ657	918	140.28	96.04
L2_V119BJ	919	150.23	95.16	L2_GQ366	919	228.26	96.8
L2_rus16	988	235.97	96.88	L2_GQ254	932	135.1	96.19
L2_rus14	955	237.07	96.63	L2_GQ229	932	105.34	95.81
L2_N1051	888	66.81	96.95	L2_GQ1973	940	124.58	96.1
L2_N1037	900	52.35	96.97	L2_GQ1885	819	60.31	92.52
L2_N1001	891	57.3	97.16	L2_GQ1645	746	22.39	83.41
L2_N0158	948	122.02	96.23	L2_GQ1605	832	52.88	92.8
L2_N0151	936	139.62	96.62	L2_GQ1597	852	45.51	91.48
L2_N0150	973	151.82	96.31	L2_GQ1580	941	118.56	95.99
L2_N0130b4	957	184.21	96.31	L2_GQ1439	959	137.23	95.85
L2_N0128	874	52.43	96.88	L2_GQ1335	975	212.51	96.56
L2_N0094	950	161.51	96.58	L2_GQ1331	976	213.63	96.4
L2_N0053	981	161	96.19	L2_GQ1164	915	152.1	95.82
L2_N0051	947	185.76	95.9	L2_GQ1020	927	144.11	95.92
L2_N0050	878	51.74	96.71	L2_GQ10-3	921	66.31	95.92
L2_N0044	949	156.61	96.76	L2_GQ-50	974	202.62	96.38
L2_N0041	926	158.59	96.81	L2_GQ-1972	932	96.16	95.93
L2_N0039	891	60.38	95.42	L2_GQ-1438	933	123.07	96.45
L2_N0034	923	129.28	96.37	L2_GQ-1343	939	109.38	96.07
L2_N0020	918	22.83	95.33	L2_GQ-1168	954	203.51	96.52
L2_N0017	833	54.91	96.79	L2_GQ-1165	967	213.69	96.84
L2_N0010	912	40.67	95.33	L2_861	952	162.45	96.36
L2_N0008	882	50.11	96.96	L2_759	934	192.16	96.64
L2_N0005	928	123.27	96.15	L2_72	966	206.79	96.62
L2_N0003	933	132.96	96.88	L2_1864	963	207.32	96.65
L2_N0002	957	136.5	96.61	L2_1336	994	156.23	96.27
L2_M4100A	755	21.27	94.2	L2_1027	934	143.02	96.19
L2_GT649	867	196.29	96.02	L1_V372IO	879	84.1	93.28

L2_GT345	859	47.51	94.81	L1_V346IO	874	111.72	92.82
L1_V232IO	789	28.59	84.36	L1_N0014	880	70.66	94.35
L1_T92	746	21.72	92.9	L1_K93	864	44.44	96.18
L1_T83	831	54.33	92.46	L1_K67	904	49.08	95.82
L1_N73	849	65.88	93.89	L1_K21	929	51.14	96.77
L1_N72	927	48.39	96.04	L1_GT281	908	157.99	96.36
L1_N70	924	47.18	96.14	L1_DY28	953	175.19	96.55
L1_N1009	911	54.4	97.56	L1_BTBS493	906	55.35	97.13
L1_N1006	890	63.27	97.62	L1_BTBS280	928	50.66	97.3
L1_N1004	890	49.88	96.92	L1_950545	545	30.07	95.26
L1_N0203	888	63.48	97.59	L1_94701	933	96.1	96.01
L1_N0196	907	106.05	96.56	L1_745902	966	79.74	95.44
L1_N0182	859	65.4	97.53	L1_600603	933	74.63	96.57
L1_N0141b4	959	179.1	96.85	L1_539606	885	25.7	94.44
L1_N0134	941	163.9	96.64	L1_468303	924	69.92	96.21
L1_N0132	933	35.54	95.86	L1_179703	855	52.1	91.42
L1_N0127	957	170.68	96.58	L1_157508	932	121.61	95.63
L1_N0079	931	41.06	96.67	L1_1571_99	944	87.41	95.97
L1_N0067	920	72.61	97.38	L1_1277803	884	87.72	95.51
L1_N0065	895	60.69	97.14	L1_1105103	895	74.06	95.86
L1_N0062	864	58.06	97.11	L1_10786_02	909	98.49	95.92
L1_N0043	842	16.54	90.12	Outgroup (M. canettii)	10821	62.24	92.98

Appendix B
Table S9 - Information on SNPs identified in the animal cluster

Position	Ref	SNP	M. pinnipedii	M microtii	M. orygis	bovis/caprae	L6/chimpanzee	SNP Effect	Gene ID	Gene name	Gene function	old_AA_new_AA	Old_codon_New_codon	Codon_Num(CDS)	CDS_size
3139	C	G	X					NON_SYNONYMOUS_CODING	Rv0002	dnaN	DNA polymerase III (beta chain) DnaN (DNA nucleotidyltransferase	A_G	gCc_gGc	363	1209
10049	T	C	X	N				NON_SYNONYMOUS_CODING	Rv0007	Rv0007	Possible conserved membrane protein	W_R	Tgg_Cgg	46	915
26053	C	G	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv0021c	Rv0021c	hypothetical protein	A_P	Gct_Cct	277	969
38121	A	C	X	N				NON_SYNONYMOUS_CODING	Rv0035	fadD34	Probable fatty-acid-CoA ligase FadD34 (fatty-acid-CoA synthetase)	E_A	gAa_gCa	288	1689
47899	C	G	X	X				NON_SYNONYMOUS_CODING	Rv0043c	Rv0043c	Probable transcriptional regulatory protein (probably GntR-family)	G_R	Ggc_Cgc	68	735
70838	C	A	X					NON_SYNONYMOUS_CODING	Rv0064	Rv0064	Probable conserved transmembrane protein	P_Q	cCg_cAg	740	2940
71551	G	A	X	N				NON_SYNONYMOUS_CODING	Rv0064	Rv0064	Probable conserved transmembrane protein	G_S	Ggt_Agt	978	2940
74161	C	G	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0066c	icd2	Probable isocitrate dehydrogenase [NADP] Icd2 (oxalosuccinate de	K_N	aaG_aaC	117	2238
78103	T	C	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0070c	glyA2	Serine hydroxymethyltransferase GlyA2 (serine methylase 2) (SHM	E_G	gAg_gGg	265	1278
86864	A	G	X	N	X	X		NON_SYNONYMOUS_CODING	Rv0078	Rv0078	Probable transcriptional regulatory protein	I_V	Atc_Gtc	113	606
97085	A	T	X	N				NON_SYNONYMOUS_CODING	Rv0088	Rv0088	Possible polyketide cyclase/dehydrase	K_N	aaA_aaT	53	675
100589	A	G	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv0092	ctpA	Cation transporter P-type ATPase a CtpA	T_A	Acg_Gcg	3	2286
100766	G	A						NON_SYNONYMOUS_CODING	Rv0092	ctpA	Cation transporter P-type ATPase a CtpA	D_N	Gac_Aac	62	2286
126103	C	T	X	X				NON_SYNONYMOUS_CODING	Rv0107c	ctpI	Probable cation-transporter ATPase I CtpI	G_E	gGg_gAg	1480	4899
126563	C	A	X	N	X	X		NON_SYNONYMOUS_CODING	Rv0107c	ctpI	Probable cation-transporter ATPase I CtpI	A_S	Gcg_Tcg	1327	4899
135928	G	A	X	X				NON_SYNONYMOUS_CODING	Rv0111	Rv0111	Possible transmembrane acyltransferase	S_N	aGt_aAt	660	2058
136595	C	T	X	X				NON_SYNONYMOUS_CODING	Rv0112	gca	Possible GDP-mannose 4,6-dehydratase Gca (GDP-D-mannose de	R_C	Cgt_Tgt	103	957
148102	C	T	X	N				NON_SYNONYMOUS_CODING	Rv0121c	Rv0121c	hypothetical protein	D_N	Gac_Aac	81	435
170083	A	G	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0143c	Rv0143c	Probable conserved transmembrane protein	F_L	Ttc_Ctc	34	1479
201567	G	C	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0171	mce1C	Mce-family protein Mce1C	E_D	gaG_gaC	212	1548
209424	G	A	X	X				NON_SYNONYMOUS_CODING	Rv0178	Rv0178	Probable conserved Mce associated membrane protein	A_T	Gcc_Acc	163	735
210442	C	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0179c	lprO	Possible lipoprotein LprO	R_H	cGc_cAc	124	1110
218540	T	C	X					NON_SYNONYMOUS_CODING	Rv0186A	mymT	Metallothionein, MymT	I_M	atA_atG	4	162
220807	A	G	X					NON_SYNONYMOUS_CODING	Rv0189c	ilvD	Probable dihydroxy-acid dehydratase IlvD (dad)	V_A	gTc_gCc	306	1728
227860	C	T	X	N	X	X		NON_SYNONYMOUS_CODING	Rv0194	Rv0194	Probable transmembrane multidrug efflux pump	P_L	cCg_cTg	328	3585
231929	G	A	X	N				NON_SYNONYMOUS_CODING	Rv0196	Rv0196	Possible transcriptional regulatory protein	A_T	Gcc_Acc	95	585
245921	G	C	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0206c	mmpL3	Possible conserved transmembrane transport protein MmpL3	D_E	gaC_gaG	466	2835
246169	A	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0206c	mmpL3	Possible conserved transmembrane transport protein MmpL3	F_I	Ttc_Atc	384	2835
272161	C	A	X					NON_SYNONYMOUS_CODING	Rv0227c	Rv0227c	Probable conserved membrane protein	V_F	Gtc_Ttc	227	1266
275367	G	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0230c	php	Probable phosphotriesterase Php (parathion hydrolase) (PTE) (aryl	D_E	gaC_gaA	199	981
311493	C	T	X					NON_SYNONYMOUS_CODING	Rv0259c	Rv0259c	hypothetical protein	G_R	Gga_Aga	9	744
311728	C	T	X	X				NON_SYNONYMOUS_CODING	Rv0260c	Rv0260c	Possible transcriptional regulatory protein	G_E	gGa_gAa	311	1146
382243	C	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0312	Rv0312	Conserved hypothetical proline and threonine rich protein	P_L	cCg_cTg	563	1863
384248	G	C	X	X				NON_SYNONYMOUS_CODING	Rv0315	Rv0315	Possible beta-1,3-glucanase precursor	W_S	tGg_tCg	216	885
411100	A	G	X	X	X	X		NON_SYNONYMOUS_CODING	Rv0342	iniA	Isoniazid inductible gene protein IniA	N_S	aAt_aGt	88	1923
420100	C	T	X					NON_SYNONYMOUS_CODING	Rv0350	dnaK	Probable chaperone protein DnaK (heat shock protein 70) (heat sh	P_L	cCg_cTg	89	1878

440679	G	A	X	X				NON_SYNONYMOUS_CODING	Rv0362	mgtE	Possible Mg2+ transport transmembrane protein MgtE	G_D	gGt_gAt	270	1383
447042	G	A	X					NON_SYNONYMOUS_CODING	Rv0369c	Rv0369c	Possible membrane oxidoreductase	P_L	cCt_cTt	2	516
482237	A	G	X	X				NON_SYNONYMOUS_CODING	Rv0402c	mmpL1	Probable conserved transmembrane transport protein MmpL1	L_S	tTg_tCg	332	2877
510766	C	A	X	X				NON_SYNONYMOUS_CODING	Rv0425c	ctpH	Possible metal cation transporting P-type ATPase CtpH	R_M	aGg_aTg	1519	4620
519331	A	G	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv0431	Rv0431	Putative tuberculin related peptide	T_A	Aca_Gca	87	495
539799	G	A	X	X				NON_SYNONYMOUS_CODING	Rv0450c	mmpL4	Probable conserved transmembrane transport protein MmpL4	H_Y	Cac_Tac	565	2904
556201	T	C	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0465c	Rv0465c	Probable transcriptional regulatory protein	N_D	Aac_Gac	36	1425
598723	G	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0507	mmpL2	Probable conserved transmembrane transport protein MmpL2	D_Y	Gat_Tat	509	2907
599363	C	G	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0507	mmpL2	Probable conserved transmembrane transport protein MmpL2	S_C	tCt_tGt	722	2907
650685	C	T	X	N				NON_SYNONYMOUS_CODING	Rv0559c	Rv0559c	Possible conserved secreted protein	A_T	Gcg_Acg	21	339
658732	C	T	X					NON_SYNONYMOUS_CODING	Rv0567	Rv0567	Probable methyltransferase/methylase	L_F	Ctc_Ttc	138	1020
658923	C	G	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0567	Rv0567	Probable methyltransferase/methylase	F_L	ttC_ttG	201	1020
670674	A	G	X					NON_SYNONYMOUS_CODING	Rv0576	Rv0576	Probable transcriptional regulatory protein (possibly ArsR-family)	Q_R	cAg_cGg	276	1305
682151	C	A	X	X				NON_SYNONYMOUS_CODING	Rv0585c	Rv0585c	Probable conserved integral membrane protein	A_S	Gcc_Tcc	708	2388
693003	A	G	X	X				NON_SYNONYMOUS_CODING	Rv0593	lprL	Possible Mce-family lipoprotein LprL (Mce-family lipoprotein Mce2E	D_G	gAc_gGc	327	1209
713802	C	T	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv0620	galK	Probable galactokinase GalK (galactose kinase)	R_C	Cgt_Tgt	363	1092
715796	T	G	X					NON_SYNONYMOUS_CODING	Rv0622	Rv0622	Possible membrane protein	Y_D	Tac_Gac	143	948
736919	G	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0642c	mmaA4	Methoxy mycolic acid synthase 4 MmaA4 (methyl mycolic acid syn	F_L	ttC_ttA	95	906
738820	C	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv0644c	mmaA2	Methoxy mycolic acid synthase 2 MmaA2 (methyl mycolic acid syn	R_L	cGc_cTc	114	864
778854	C	G	X					NON_SYNONYMOUS_CODING	Rv0677c	mmpS5	Possible conserved membrane protein MmpS5	V_L	Gtc_Ctc	18	429
779685	A	G	X					NON_SYNONYMOUS_CODING	Rv0679c	Rv0679c	Conserved threonine rich protein	V_A	gTc_gCc	119	498
787352	C	T	X	X				NON_SYNONYMOUS_CODING	Rv0687	Rv0687	Probable short-chain type dehydrogenase/reductase	A_V	gCc_gTc	85	828
817299	G	A	X	X				NON_SYNONYMOUS_CODING	Rv0724	sppA	Possible protease IV SppA (endopeptidase IV) (signal peptide pepI	R_H	cGt_cAt	546	1872
818161	G	C	X	X				NON_SYNONYMOUS_CODING	Rv0725c	Rv0725c	hypothetical protein	T_S	aCt_aGt	95	906
850983	T	G	X	N				NON_SYNONYMOUS_CODING	Rv0756c	Rv0756c	hypothetical protein	T_P	Acg_Ccg	162	726
858531	C	T	X	X				NON_SYNONYMOUS_CODING	Rv0765c	Rv0765c	Probable oxidoreductase	A_T	Gca_Aca	112	828
866345	T	C	X	N				NON_SYNONYMOUS_CODING	Rv0773c	ggtA	Probable bifunctional acylase GgtA: cephalosporin acylase (GL-7A	I_V	Att_Gtt	349	1539
867574	C	T	X	N				NON_SYNONYMOUS_CODING	Rv0774c	Rv0774c	Probable conserved exported protein	A_T	Gca_Aca	260	912
874792	A	C	X	X				NON_SYNONYMOUS_CODING	Rv0781	ptrBa	Probable protease II PtrBa [first part] (oligopeptidase B)	E_A	gAg_gCg	187	711
879463	G	A	X	N				NON_SYNONYMOUS_CODING	Rv0785	Rv0785	hypothetical protein	D_N	Gac_Aac	42	1701
886295	G	A	X					NON_SYNONYMOUS_CODING	Rv0792c	Rv0792c	Probable transcriptional regulatory protein (probably GntR-family)	H_Y	Cac_Tac	118	810
898492	G	A	X	N				NON_SYNONYMOUS_CODING	Rv0804	Rv0804	hypothetical protein	A_T	Gcc_Acc	138	630
910002	G	A	X	X				NON_SYNONYMOUS_CODING	Rv0816c	thiX	Probable thioredoxin ThiX	T_I	aCa_aTa	11	423
917259	C	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv0823c	Rv0823c	Possible transcriptional regulatory protein	G_C	Ggc_Tgc	130	1170
942276	G	T	X	X				STOP_GAINED	Rv0845	Rv0845	Possible two component sensor kinase	E_*	Gaa_Taa	363	1278
944725	C	T	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv0847	lpqS	Probable lipoprotein LpqS	A_V	gCt_gTt	128	393
980058	C	G	X					NON_SYNONYMOUS_CODING	Rv0881	Rv0881	Possible rRNA methyltransferase (rRNA methylase)	L_V	Ctg_Gtg	233	867
1000732	A	G	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv0896	gltA2	Probable citrate synthase I GltA2	T_A	Acg_Gcg	421	1296
1012232	A	G	X					NON_SYNONYMOUS_CODING	Rv0908	ctpE	Probable metal cation transporter ATPase P-type CtpE	M_V	Atg_Gtg	168	2394
1018359	C	T	X					NON_SYNONYMOUS_CODING	Rv0913c	Rv0913c	Possible dioxygenase	E_K	Gaa_Aaa	123	1509
1030043	A	G	X	N				NON_SYNONYMOUS_CODING	Rv0923c	Rv0923c	hypothetical protein	S_P	Tcc_Ccc	179	1065
1043169	T	C	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv0934	pstS1	Periplasmic phosphate-binding lipoprotein PstS1 (PBP-1) (PstS1)	V_A	gTt_gCt	352	1125
1046908	C	T	X	X				NON_SYNONYMOUS_CODING	Rv0938	ligD	ATP dependent DNA ligase LigD (ATP dependent polydeoxyribonu	A_V	gCg_gTg	258	2280
1050212	A	C						NON_SYNONYMOUS_CODING	Rv0939	Rv0939	Possible bifunctional enzyme: 2-hydroxyhepta-2,4-diene-1,7-dioate	S_R	Agc_Cgc	601	1935
1051633	C	T	X	X				NON_SYNONYMOUS_CODING	Rv0941c	Rv0941c	hypothetical protein	V_I	Gtt_Att	229	774
1078209	A	T	X	X				NON_SYNONYMOUS_CODING	Rv0967	csoR	Copper-sensitive operon repressor CsoR	I_F	Atc_Ttc	79	360
1101375	C	G	X	X				NON_SYNONYMOUS_CODING	Rv0985c	mscL	Possible large-conductance ion mechanosensitive channel MscL	D_H	Gac_Cac	36	456
1115540	G	A	X	X				NON_SYNONYMOUS_CODING	Rv0998	Rv0998	hypothetical protein	G_S	Ggc_Agc	265	1002
1118314	C	T	X					NON_SYNONYMOUS_CODING	Rv1001	arcA	Probable arginine deiminase ArcA (adi) (ad) (arginine dihydrolase)	T_I	aCc_aTc	377	1209

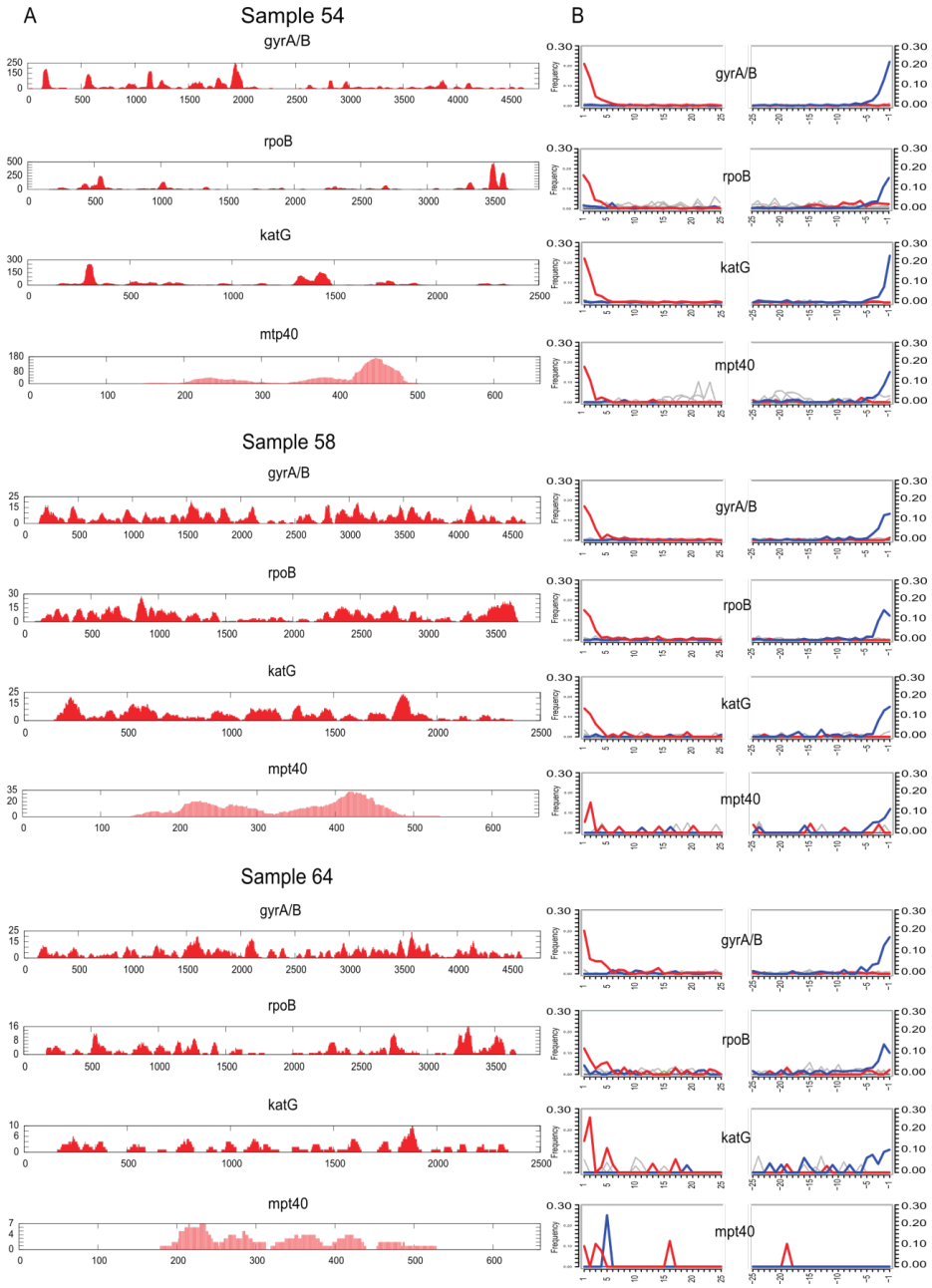
1119597	C	G	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1002c	Rv1002c	hypothetical protein	V_L	Gtc_Ctc	115	1512
1147542	C	T	X	N				NON_SYNONYMOUS_CODING	Rv1026	Rv1026	hypothetical protein	T_I	aCc_aTc	175	960
1150803	C	T	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1028c	kdpD	Probable sensor protein KdpD	G_D	gGc_gAc	295	2583
1153401	G	T	X					NON_SYNONYMOUS_CODING	Rv1029	kdpA	Probable potassium-transporting ATPase a chain KdpA (potassium	G_C	Ggt_Tgt	464	1716
1168515	G	C	X					NON_SYNONYMOUS_CODING	Rv1045	Rv1045	Hypothetical protein	Q_H	caG_caC	281	882
1180796	A	C	X	X				NON_SYNONYMOUS_CODING	Rv1058	fadD14	Probable medium chain fatty-acid-CoA ligase FadD14 (fatty-acid-C	H_P	cAc_cCc	38	1632
1221181	C	G	X					NON_SYNONYMOUS_CODING	Rv1093	glyA1	Serine hydroxymethyltransferase 1 GlyA1	A_G	gCg_gGg	203	1281
1225198	C	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv1096	Rv1096	Possible glycosyl hydrolase	P_S	Ccc_Tcc	272	876
1238483	G	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv1112	Rv1112	Probable GTP binding protein	V_I	Gtc_Atc	77	1074
1253655	G	A	X	N				NON_SYNONYMOUS_CODING	Rv1129c	Rv1129c	Probable transcriptional regulator protein	R_C	Cgt_Tgt	294	1461
1260134	C	T	X	N				NON_SYNONYMOUS_CODING	Rv1133c	metE	Probable 5-methyltetrahydropteroyltriglutamate--homocysteine met	A_T	Gcg_Acg	405	2280
1306281	A	C	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1175c	fadH	Probable NADPH dependent 2,4-dienoyl-CoA reductase FadH (2,4	I_S	aTc_aGc	649	2025
1313911	T	C	X	X				NON_SYNONYMOUS_CODING	Rv1180	pks3	Probable polyketide beta-ketoacyl synthase Pks3	S_P	Tcg_Ccg	63	1467
1352253	A	G	X					NON_SYNONYMOUS_CODING	Rv1208	gpgS	Probable glucosyl-3-phosphoglycerate synthase GpgS	E_G	gAa_gGa	37	975
1366736	T	C	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1223	htrA	Probable serine protease HtrA (DEGP protein)	F_L	Ttc_Ctc	288	1587
1367208	G	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv1223	htrA	Probable serine protease HtrA (DEGP protein)	S_N	aGc_aAc	445	1587
1387580	C	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1244	lpqZ	Probable lipoprotein LpqZ	Q_K	Caa_Aaa	242	861
1404944	A	G	X	X		N		NON_SYNONYMOUS_CODING	Rv1257c	Rv1257c	Probable oxidoreductase	Y_H	Tac_Cac	381	1368
1431918	T	C	X					NON_SYNONYMOUS_CODING	Rv1280c	oppA	Probable periplasmic oligopeptide-binding lipoprotein OppA	D_G	gAc_gGc	508	1776
1437945	T	C	X					NON_SYNONYMOUS_CODING	Rv1285	cysD	Probable sulfate adenylyltransferase subunit 2 CysD	F_L	Ttt_Ctt	13	999
1454047	C	A	X					NON_SYNONYMOUS_CODING	Rv1297	rho	Probable transcription termination factor Rho homolog	P_T	Ccc_Acc	282	1809
1457621	G	A	X	X				NON_SYNONYMOUS_CODING	Rv1301	Rv1301	hypothetical protein	G_R	Ggg_Agg	22	654
1465741	G	T	X	X				NON_SYNONYMOUS_CODING	Rv1309	atpG	Probable ATP synthase gamma chain AtpG	Q_H	caG_caT	286	918
1508682	T	G	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1343c	lprD	Probable conserved lipoprotein LprD	E_A	gAg_gCg	81	381
1527311	C	T	X	X				NON_SYNONYMOUS_CODING	Rv1358	Rv1358	Probable transcriptional regulatory protein	R_C	Cgc_Tgc	234	3480
1531861	G	T	X	X				STOP_GAINED	Rv1360	Rv1360	Probable oxidoreductase	E_*	Gaa_Taa	172	1023
1537362	C	T	X	X				NON_SYNONYMOUS_CODING	Rv1364c	Rv1364c	Possible sigma factor regulatory protein	D_N	Gat_Aat	95	1962
1537926	C	T	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1365c	rsfA	Anti-anti-sigma factor RsfA (anti-sigma factor antagonist) (regulato	E_K	Gag_Aag	82	387
1585574	G	T	X	N				NON_SYNONYMOUS_CODING	Rv1409	ribG	Probable bifunctional riboflavin biosynthesis protein RibG : diamino	Q_H	caG_caT	127	1020
1646982	C	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv1460	Rv1460	Probable transcriptional regulatory protein	A_V	gCg_gTg	266	807
1671329	C	T	X	N				NON_SYNONYMOUS_CODING	Rv1480	Rv1480	hypothetical protein	A_V	gCc_gTc	306	954
1674520	C	T	X					NON_SYNONYMOUS_CODING	Rv1484	inhA	NADH-dependent enoyl-[acyl-carrier-protein] reductase InhA (NADI	P_S	Ccg_Tcg	107	810
1690345	A	G	X	N	X	X	X	START_LOST	Rv1498A	Rv1498A	hypothetical protein	M_T	aTg_aCg	1	213
1694199	C	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv1503c	Rv1503c	hypothetical protein	E_K	Gaa_Aaa	116	549
1694300	C	A	X					NON_SYNONYMOUS_CODING	Rv1503c	Rv1503c	hypothetical protein	R_L	cGt_cTt	82	549
1694432	G	C	X					NON_SYNONYMOUS_CODING	Rv1503c	Rv1503c	hypothetical protein	P_R	cCc_cGc	38	549
1698720	A	G	X			N		NON_SYNONYMOUS_CODING	Rv1508c	Rv1508c	Probable membrane protein	I_T	aTc_aCc	392	1800
1712419	A	C	X	X				NON_SYNONYMOUS_CODING	Rv1521	fadD25	Probable fatty-acid-AMP ligase FadD25 (fatty-acid-AMP synthetase	T_P	Act_Cct	40	1752
1712522	G	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv1521	fadD25	Probable fatty-acid-AMP ligase FadD25 (fatty-acid-AMP synthetase	S_I	aGc_aTc	74	1752
1721814	C	T	X					NON_SYNONYMOUS_CODING	Rv1526c	Rv1526c	Probable glycosyltransferase	A_T	Gcc_Acc	83	1281
1742583	C	T	X	X				NON_SYNONYMOUS_CODING	Rv1539	lspA	Probable lipoprotein signal peptidase LspA	P_S	Ccg_Tcg	114	609
1753561	C	T	X					STOP_GAINED	Rv1549	fadD11.1	Possible fatty-acid-CoA ligase FadD11.1 (fatty-acid-CoA synthetas	R_*	Cga_Tga	18	528
1768141	A	G	X	N				NON_SYNONYMOUS_CODING	Rv1563c	treY	Maltooligosyltrehalose synthase TreY	V_A	gTc_gCc	431	2298
1831253	C	T	X	N				NON_SYNONYMOUS_CODING	Rv1629	polA	Probable DNA polymerase I PolA	R_C	Cgc_Tgc	197	2715
1845241	C	A	X					NON_SYNONYMOUS_CODING	Rv1638	uvrA	Probable excinuclease ABC (subunit A-DNA-binding ATPase) UvrA	L_M	Ctg_Atg	501	2919
1854695	C	T	X	N				NON_SYNONYMOUS_CODING	Rv1645c	Rv1645c	hypothetical protein	A_T	Gcg_Acg	254	1056
1877311	C	T	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1661	pks7	Probable polyketide synthase Pks7	R_W	Cgg_Tgg	670	6381
1893726	G	A	X					NON_SYNONYMOUS_CODING	Rv1667c	Rv1667c	Probable second part of macrolide-transport ATP-binding protein A	R_W	Cgg_Tgg	169	654
1912024	C	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1686c	Rv1686c	Probable conserved integral membrane protein ABC transporter	V_F	Gtc_Ttc	20	681

1952827	C	T	X	N	X	X		NON_SYNONYMOUS_CODING	Rv1726	Rv1726	Probable oxidoreductase	H_Y	Cac_Tac	326	1386
1968116	A	C	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv1741	vapC34	Possible toxin VapC34 Contains PIN domain	K_T	aAg_aCg	67	249
2004057	G	C	X	N				NON_SYNONYMOUS_CODING	Rv1770	Rv1770	hypothetical protein	E_D	gaG_gaC	60	1287
2007352	C	T	X	X				NON_SYNONYMOUS_CODING	Rv1773c	Rv1773c	Probable transcriptional regulatory protein	G_R	Ggg_Agg	139	747
2017861	G	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv1782	eccB5	ESX conserved component EccB5 ESX-5 type VII secretion systen	R_H	cGt_cAt	41	1521
2033021	G	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv1795	eccD5	ESX conserved component EccD5 ESX-5 type VII secretion systen	G_D	gGt_gAt	261	1512
2053155	G	A	X	X				NON_SYNONYMOUS_CODING	Rv1810	Rv1810	hypothetical protein	G_S	Ggc_Agc	75	357
2075778	C	G	X	X				NON_SYNONYMOUS_CODING	Rv1831	Rv1831	Hypothetical protein	L_V	Ctg_Gtg	70	258
2076502	A	C	X	N				NON_SYNONYMOUS_CODING	Rv1832	gcvB	Probable glycine dehydrogenase GcvB (glycine decarboxylase) (gl	D_A	gAc_gCc	209	2826
2083589	C	T	X	X				NON_SYNONYMOUS_CODING	Rv1836c	Rv1836c	hypothetical protein	A_T	Gcg_Acg	350	2034
2087585	G	T	X					NON_SYNONYMOUS_CODING	Rv1838c	vapC13	Possible toxin VapC13	A_E	gCg_gAg	23	396
2090776	A	G	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1842c	Rv1842c	hypothetical protein	L_S	tTg_tCg	437	1368
2095417	G	A	X	N				NON_SYNONYMOUS_CODING	Rv1845c	blaR	Possible sensor-transducer protein BlaR	A_V	gCg_gTg	251	951
2107511	C	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv1859	modC	Probable molybdenum-transport ATP-binding protein ABC transpor	T_M	aCg_aTg	313	1110
2136745	G	A	X	X				STOP_GAINED	Rv1887	Rv1887	Hypothetical protein	W_*	tGg_tAg	163	1143
2137267	C	T	X	X				NON_SYNONYMOUS_CODING	Rv1887	Rv1887	Hypothetical protein	T_M	aCg_aTg	337	1143
2156056	C	A	X	N				NON_SYNONYMOUS_CODING	Rv1908c	katG	Catalase-peroxidase-peroxynitritase T KatG	G_V	gGc_gTc	19	2223
2172012	G	C	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1920	Rv1920	Probable membrane protein	M_I	atG_atC	130	864
2225175	A	C	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv1981c	nrdF1	Ribonucleoside-diphosphate reductase (beta chain) NrdF1 (ribonuc	L_R	cTc_cGc	5	969
2243186	A	C	X	X				NON_SYNONYMOUS_CODING	Rv1998c	Rv1998c	hypothetical protein	L_R	cTc_cGc	179	777
2258216	A	G	X	X	X	X		NON_SYNONYMOUS_CODING	Rv2009	vapB15	Antitoxin VapB15	N_D	Aac_Gac	63	243
2274463	A	C	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv2027c	dosT	Two component sensor histidine kinase DosT	L_V	Ttg_Gtg	16	1722
2286850	G	T	X	X				NON_SYNONYMOUS_CODING	Rv2041c	Rv2041c	Probable sugar-binding lipoprotein	R_S	Cgc_Agc	333	1320
2313267	G	A	X					NON_SYNONYMOUS_CODING	Rv2054	Rv2054	hypothetical protein	R_H	cGt_cAt	48	714
2316624	G	A	X	X				NON_SYNONYMOUS_CODING	Rv2059	Rv2059	hypothetical protein	G_D	gGc_gAc	484	1536
2316624	G	A	X	X				NON_SYNONYMOUS_CODING	Rv2060	Rv2060	Possible conserved integral membrane protein	A_T	Gcc_Acc	116	402
2325725	G	A	X	N				NON_SYNONYMOUS_CODING	Rv2067c	Rv2067c	hypothetical protein	P_L	cCg_cTg	49	1224
2337373	T	C	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2080	lppJ	Lipoprotein LppJ	V_A	gTg_gCg	23	564
2351522	G	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2092c	helY	ATP-dependent DNA helicase HelY	T_M	aCg_aTg	178	2721
2364140	A	C	X					NON_SYNONYMOUS_CODING	Rv2103c	vapC37	Possible toxin VapC37 Contains PIN domain	S_R	agT_agG	127	435
2373439	C	A	X					NON_SYNONYMOUS_CODING	Rv2113	Rv2113	Probable integral membrane protein	N_K	aaC_aaA	270	1194
2400450	C	G	X	X				NON_SYNONYMOUS_CODING	Rv2141c	dapE2	hypothetical protein	V_L	Gtc_Ctc	425	1347
2437837	G	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv2175c	Rv2175c	Conserved regulatory protein	P_L	cCc_cTc	17	441
2438094	T	G	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2176	pknL	Probable transmembrane serine/threonine-protein kinase L PknL (p	S_A	Tct_Gct	52	1200
2448598	C	T	X	N				STOP_GAINED	Rv2187	fadD15	Long-chain-fatty-acid-CoA ligase FadD15 (fatty-acid-CoA synthetas	Q_*	Cag_Tag	147	1803
2461545	C	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv2197c	Rv2197c	Probable conserved transmembrane protein	A_S	Gct_Tct	202	645
2471816	A	G	X	N				NON_SYNONYMOUS_CODING	Rv2207	cobT	Probable nicotinate-nucleotide-dimethylbenzimidazol phosphoribos	T_A	Acc_Gcc	136	1086
2478967	C	G	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv2213	pepB	Probable aminopeptidase PepB	F_L	ttC_ttG	210	1548
2509362	G	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv2236c	cobD	Probable cobalamin biosynthesis transmembrane protein CobD	T_I	aCc_aTc	5	942
2527522	G	A	X	X				NON_SYNONYMOUS_CODING	Rv2252	Rv2252	Diacylglycerol kinase	M_I	atG_atA	178	930
2531767	G	A						NON_SYNONYMOUS_CODING	Rv2258c	Rv2258c	Possible transcriptional regulatory protein	P_L	cCg_cTg	44	1062
2535032	G	A	X	X				NON_SYNONYMOUS_CODING	Rv2262c	Rv2262c	hypothetical protein	A_V	gCg_gTg	174	1083
2542395	G	T	X	X				NON_SYNONYMOUS_CODING	Rv2267c	Rv2267c	hypothetical protein	A_E	gCa_gAa	139	1167
2547118	A	C	X	X				NON_SYNONYMOUS_CODING	Rv2275	Rv2275	hypothetical protein	E_A	gAa_gCa	79	870
2554391	A	C	X					NON_SYNONYMOUS_CODING	Rv2281	pitB	Putative phosphate-transport permease PitB	S_R	Agc_Cgc	407	1659
2576251	G	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2305	Rv2305	hypothetical protein	G_D	gGc_gAc	148	1290
2577994	G	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv2307c	Rv2307c	hypothetical protein	R_W	Cgg_Tgg	235	846
2590882	A	C	X					NON_SYNONYMOUS_CODING	Rv2318	uspC	Probable periplasmic sugar-binding lipoprotein UspC	D_A	gAc_gCc	122	1323
2631424	T	G	X	N	N	N		NON_SYNONYMOUS_CODING	Rv2351c	plcA	Membrane-associated phospholipase C 1 PlcA (MTP40 antigen)	S_R	Agt_Cgt	218	1539

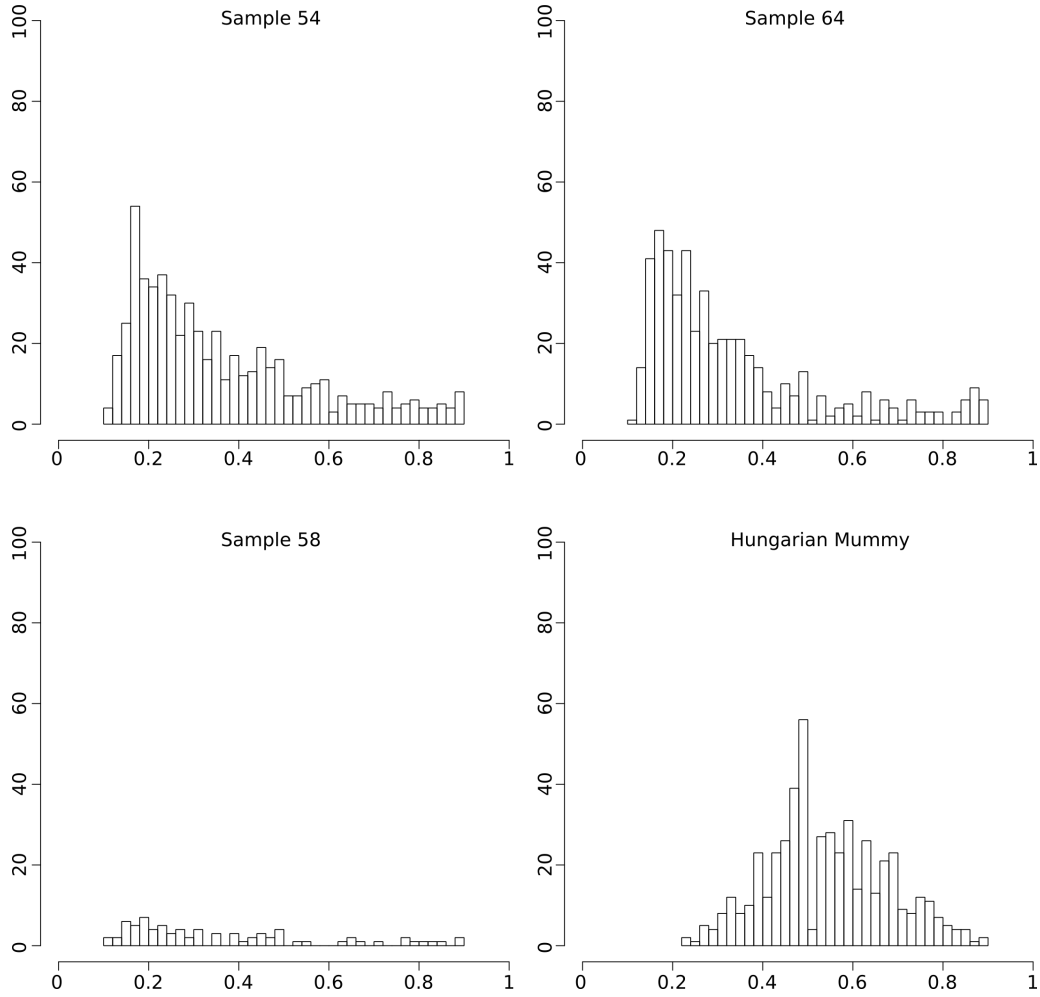
2654565	G	C	X					NON_SYNONYMOUS_CODING	Rv2374c	hrcA	Probable heat shock protein transcriptional repressor HrcA	Q_E	Cag_Gag	177	1032
2674180	G	A	X	N				NON_SYNONYMOUS_CODING	Rv2383c	mbtB	Phenyloxazoline synthase MbtB (phenyloxazoline synthetase)	A_V	gCg_gTg	553	4245
2689193	G	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2394	ggtB	Probable gamma-glutamyltranspeptidase precursor GgtB (gamma-	R_Q	cGg_cAg	395	1932
2750773	G	A	X	X				NON_SYNONYMOUS_CODING	Rv2449c	Rv2449c	hypothetical protein	A_V	gCc_gTc	267	1260
2762064	G	A	X	X				NON_SYNONYMOUS_CODING	Rv2459	Rv2459	Probable conserved integral membrane transport protein	G_D	gGc_gAc	404	1527
2772741	C	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2469c	Rv2469c	hypothetical protein	A_S	Gcg_Tcg	99	669
2772954	T	C	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv2469c	Rv2469c	hypothetical protein	S_G	Agt_Ggt	28	669
2773381	C	G	X	X				NON_SYNONYMOUS_CODING	Rv2470	glbO	Globin (oxygen-binding protein) GlbO	H_Q	caC_caG	68	387
2773381	C	A	X	X				NON_SYNONYMOUS_CODING	Rv2470	glbO	Globin (oxygen-binding protein) GlbO	H_Q	caC_caA	68	387
2792233	C	T	X	X				NON_SYNONYMOUS_CODING	Rv2484c	Rv2484c	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	D_N	Gac_Aac	88	1476
2798552	C	T	X	X	N			NON_SYNONYMOUS_CODING	Rv2488c	Rv2488c	Probable transcriptional regulatory protein (LuxR-family)	A_T	Gca_Aca	777	3414
2807597	A	C	X	X				NON_SYNONYMOUS_CODING	Rv2492	Rv2492	Hypothetical protein	E_A	gAa_gCa	107	753
2821049	G	A	X					NON_SYNONYMOUS_CODING	Rv2505c	fadD35	Probable fatty-acid-CoA ligase FadD35 (fatty-acid-CoA synthetase)	A_V	gCg_gTg	183	1644
2875808	C	G	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2555c	alaS	Probable alanyl-tRNA synthetase AlaS (alanine--tRNA ligase) (alar	K_N	aaG_aaC	226	2715
2881485	A	G	X	X				NON_SYNONYMOUS_CODING	Rv2561	Rv2561	hypothetical protein	H_R	cAc_cGc	26	294
2886400	G	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2566	Rv2566	Long conserved protein	G_S	Ggc_Agc	10	3423
2887180	G	T	X	N	X	X		NON_SYNONYMOUS_CODING	Rv2566	Rv2566	Long conserved protein	A_S	Gcc_Tcc	270	3423
2888055	G	T	X	N				NON_SYNONYMOUS_CODING	Rv2566	Rv2566	Long conserved protein	W_C	tgG_tgT	561	3423
2893507	G	T	X					NON_SYNONYMOUS_CODING	Rv2569c	Rv2569c	hypothetical protein	S_Y	tCt_tAt	301	945
2894853	G	A	X	N				NON_SYNONYMOUS_CODING	Rv2570	Rv2570	hypothetical protein	M_I	atG_atA	114	390
2896260	G	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2572c	aspS	Probable aspartyl-tRNA synthetase AspS (aspartate--tRNA ligase)	A_V	gCg_gTg	515	1791
2901924	G	A	X					NON_SYNONYMOUS_CODING	Rv2577	Rv2577	hypothetical protein	R_H	cGc_cAc	336	1590
2906715	C	T	X	X				NON_SYNONYMOUS_CODING	Rv2581c	Rv2581c	Possible glyoxalase II (hydroxyacylglutathione hydrolase) (GLX II)	V_M	Gtg_Atg	17	675
2911896	G	A	X					NON_SYNONYMOUS_CODING	Rv2585c	Rv2585c	Possible conserved lipoprotein	T_I	aCc_aTc	261	1674
2913111	G	T	X					NON_SYNONYMOUS_CODING	Rv2586c	secF	Probable protein-export membrane protein SecF	L_M	Ctg_Atg	301	1329
2921373	T	C	X					NON_SYNONYMOUS_CODING	Rv2590	fadD9	Probable fatty-acid-CoA ligase FadD9 (fatty-acid-CoA synthetase)	L_P	cTg_cCg	1168	3507
2948095	G	A	X					NON_SYNONYMOUS_CODING	Rv2621c	Rv2621c	Possible transcriptional regulatory protein	S_F	tCc_tTc	155	675
2955305	T	G	X	X				NON_SYNONYMOUS_CODING	Rv2628	Rv2628	Hypothetical protein	L_W	tTg_tGg	83	363
2963166	G	A	X	N				NON_SYNONYMOUS_CODING	Rv2636	Rv2636	hypothetical protein	V_I	Gta_Ata	152	678
2998012	C	T	X	X				NON_SYNONYMOUS_CODING	Rv2681	Rv2681	Conserved hypothetical alanine rich protein	T_I	aCa_aTa	425	1317
3000734	A	C	X	X				NON_SYNONYMOUS_CODING	Rv2684	arsA	Probable arsenic-transport integral membrane protein ArsA	T_P	Aca_Cca	41	1290
3011566	G	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2694c	Rv2694c	hypothetical protein	R_S	Cgc_Agc	68	369
3016735	C	G	X	N				NON_SYNONYMOUS_CODING	Rv2701c	suhB	Inositol-1-monophosphatase SuhB	V_L	Gtg_Ctg	1	873
3022136	C	T	X	N				STOP_GAINED	Rv2709	Rv2709	Probable conserved transmembrane protein	R_*	Cga_Tga	100	447
3063663	C	A	X	N				NON_SYNONYMOUS_CODING	Rv2751	Rv2751	hypothetical protein	T_N	aCc_aAc	9	891
3063872	C	T	X					NON_SYNONYMOUS_CODING	Rv2751	Rv2751	hypothetical protein	R_C	Cgc_Tgc	79	891
3083173	C	T	X	X				NON_SYNONYMOUS_CODING	Rv2775	Rv2775	GCN5-related N-acetyltransferase	L_F	Ctc_Ttc	89	462
3086424	G	T	X	X				NON_SYNONYMOUS_CODING	Rv2779c	Rv2779c	Possible transcriptional regulatory protein (probably Lrp/AsnC-famil	R_S	Cgc_Agc	111	540
3086728	C	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2779c	Rv2779c	Possible transcriptional regulatory protein (probably Lrp/AsnC-famil	M_I	atG_atA	9	540
3108272	A	G						NON_SYNONYMOUS_CODING	Rv2799	Rv2799	Probable membrane protein	K_E	Aag_Gag	169	630
3109886	G	A	X	N				NON_SYNONYMOUS_CODING	Rv2800	Rv2800	Possible hydrolase	A_T	Gcc_Acc	491	1650
3125235	G	A	X	N	X	X	X	STOP_GAINED	Rv2818c	Rv2818c	Hypothetical protein	Q_*	Caa_Taa	304	1149
3142098	G	A	X	X				NON_SYNONYMOUS_CODING	Rv2835c	ugpA	Probable Sn-glycerol-3-phosphate transport integral membrane prc	A_V	gCc_gTc	42	912
3142539	C	T	X	X				NON_SYNONYMOUS_CODING	Rv2836c	dinF	Possible DNA-damage-inducible protein F DinF	V_I	Gtt_Att	364	1320
3187792	T	C	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv2876	Rv2876	Possible conserved transmembrane protein	W_R	Tgg_Cgg	44	315
3217451	G	C	X					NON_SYNONYMOUS_CODING	Rv2909c	rpsP	30S ribosomal protein S16 RpsP	Q_E	Cag_Gag	65	489
3219973	G	A	X	X				NON_SYNONYMOUS_CODING	Rv2913c	Rv2913c	Possible D-amino acid aminohydrolase (D-amino acid hydrolase)	R_C	Cgc_Tgc	576	1836
3220433	G	T	X					NON_SYNONYMOUS_CODING	Rv2913c	Rv2913c	Possible D-amino acid aminohydrolase (D-amino acid hydrolase)	F_L	ttC_ttA	422	1836
3247579	G	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv2931	ppsA	Phenolpthiocerol synthesis type-I polyketide synthase PpsA	R_Q	cGg_cAg	712	5631

3272117	G	A	X				NON_SYNONYMOUS_CODING	Rv2935	ppsE	Phenolpthiocerol synthesis type-I polyketide synthase PpsE	E_K	Gag_Aag	1461	4467
3312169	G	C	X				NON_SYNONYMOUS_CODING	Rv2959c	Rv2959c	Possible methyltransferase (methylase)	H_D	Cat_Gat	224	738
3348258	G	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv2991	Rv2991	hypothetical protein	A_T	Gcc_Acc	93	492
3376061	G	A	X	X				NON_SYNONYMOUS_CODING	Rv3016	lpqA	Probable lipoprotein LpqA	E_K	Gaa_Aaa	102	630
3384939	C	A	X	N				NON_SYNONYMOUS_CODING	Rv3025c	iscS	Cysteine desulfurase IscS (NIFS protein homolog) (nitrogenase me	R_L	cGg_cTg	43	1182
3413842	C	A	X					NON_SYNONYMOUS_CODING	Rv3051c	nrdE	Ribonucleoside-diphosphate reductase (alpha chain) NrdE (ribonuc	V_L	Gtg_Ttg	109	2082
3444583	C	T	X	X				NON_SYNONYMOUS_CODING	Rv3080c	pknK	Serine/threonine-protein kinase transcriptional regulatory protein Pl	S_N	aGc_aAc	469	3333
3485271	A	G	X	X	N	N		NON_SYNONYMOUS_CODING	Rv3119	moaE1	Probable molybdenum cofactor biosynthesis protein E MoaE1 (mol	H_R	cAc_cGc	47	444
3488556	G	T	X	N	N	X	X	NON_SYNONYMOUS_CODING	Rv3122	Rv3122	Hypothetical protein	R_S	agG_agT	156	471
3543451	G	T	X	N				NON_SYNONYMOUS_CODING	Rv3175	Rv3175	Possible amidase (aminohydrolase)	G_C	Ggc_Tgc	198	1488
3564910	C	A	X	X				NON_SYNONYMOUS_CODING	Rv3195	Rv3195	hypothetical protein	L_M	Ctg_Atg	183	1419
3585433	A	G	X					NON_SYNONYMOUS_CODING	Rv3208	Rv3208	Probable transcriptional regulatory protein (probably TetR-family)	T_A	Acc_Gcc	144	687
3618932	T	C	X	N				NON_SYNONYMOUS_CODING	Rv3240c	secA1	Probable preprotein translocase SecA1 1 subunit	K_E	Aaa_Gaa	534	2850
3657048	A	G	X	X				NON_SYNONYMOUS_CODING	Rv3274c	fadE25	Probable acyl-CoA dehydrogenase FadE25	F_L	Ttc_Ctc	348	1170
3668591	C	A	X	X				STOP_GAINED	Rv3286c	sigF	Alternative RNA polymerase sigma factor SigF	E_*	Gaa_Taa	122	786
3673625	C	G	X	N				NON_SYNONYMOUS_CODING	Rv3293	pcd	Probable piperideine-6-carboxilic acid dehydrogenase Pcd (piperidi	I_M	atC_atG	8	1485
3674985	G	A	X	N				NON_SYNONYMOUS_CODING	Rv3293	pcd	Probable piperideine-6-carboxilic acid dehydrogenase Pcd (piperidi	E_K	Gag_Aag	462	1485
3702451	C	A	X	N				NON_SYNONYMOUS_CODING	Rv3314c	deoA	Probable thymidine phosphorylase DeoA (tdrpase) (pyrimidine pho	M_I	atG_atT	339	1284
3702543	G	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv3314c	deoA	Probable thymidine phosphorylase DeoA (tdrpase) (pyrimidine pho	R_W	Cgg_Tgg	309	1284
3707963	C	G	X	X				NON_SYNONYMOUS_CODING	Rv3320c	vapC44	Possible toxin VapC44 Contains PIN domain	W_C	tgG_tgC	36	429
3709604	T	A	X					NON_SYNONYMOUS_CODING	Rv3323c	moaX	Probable MoaD-MoaE fusion protein MoaX	Q_H	caA_caT	37	666
3773499	G	C	X	X				NON_SYNONYMOUS_CODING	Rv3361c	Rv3361c	hypothetical protein	H_Q	caC_caG	23	552
3786548	G	A	X	N				NON_SYNONYMOUS_CODING	Rv3372	otsB2	Trehalose 6-phosphate phosphatase OtsB2 (trehalose-phosphatas	E_K	Gaa_Aaa	79	1176
3790489	G	T	X					NON_SYNONYMOUS_CODING	Rv3376	Rv3376	hypothetical protein	V_F	Gtt_Ttt	112	654
3790688	T	A	X	X				NON_SYNONYMOUS_CODING	Rv3376	Rv3376	hypothetical protein	V_D	gTt_gAt	178	654
3807852	C	A	X					NON_SYNONYMOUS_CODING	Rv3392c	cmaA1	Cyclopropane-fatty-acyl-phospholipid synthase 1 CmaA1 (cyclopro	V_L	Gtg_Ttg	196	864
3866374	G	A	X	X				NON_SYNONYMOUS_CODING	Rv3447c	eccC4	ESX conserved component EccC4 ESX-4 type VII secretion systen	P_L	cCc_cTc	622	3711
3868175	G	A	X	X				NON_SYNONYMOUS_CODING	Rv3447c	eccC4	ESX conserved component EccC4 ESX-4 type VII secretion systen	R_W	Cgg_Tgg	22	3711
3882987	G	C	X					NON_SYNONYMOUS_CODING	Rv3465	rmlC	dTDP-4-dehydrorhamnose 3,5-epimerase RmlC (dTDP-4-keto-6-de	V_L	Gtg_Ctg	52	609
3888751	C	T	X	X				NON_SYNONYMOUS_CODING	Rv3470c	ilvB2	Probable acetolactate synthase (large subunit) IlvB2 (AHAS) (acet	V_I	Gtc_Atc	18	1659
3893290	G	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv3476c	kgtP	Probable dicarboxylic acid transport integral membrane protein Kgt	A_V	gCg_gTg	144	1350
3917241	G	A	X	N				NON_SYNONYMOUS_CODING	Rv3498c	mce4B	Mce-family protein Mce4B	T_M	aCg_aTg	253	1053
3937013	A	G	X	N				NON_SYNONYMOUS_CODING	Rv3509c	ilvX	Probable acetohydroxyacid synthase IlvX (acetolactate synthase)	V_A	gTc_gCc	471	1548
3952929	A	G	X	N				NON_SYNONYMOUS_CODING	Rv3516	echA19	Possible enoyl-CoA hydratase EchA19 (enoyl hydrase) (unsaturate	E_G	gAa_gGa	129	792
3952929	A	C	X	N				NON_SYNONYMOUS_CODING	Rv3516	echA19	Possible enoyl-CoA hydratase EchA19 (enoyl hydrase) (unsaturate	E_A	gAa_gCa	129	792
3964715	G	A	X	X				NON_SYNONYMOUS_CODING	Rv3528c	Rv3528c	hypothetical protein	H_Y	Cac_Tac	160	714
3987111	C	G	X	X				NON_SYNONYMOUS_CODING	Rv3547	ddn	Deazaflavin-dependent nitroreductase Ddn	L_V	Ctc_Gtc	90	456
3987939	C	T	X	X				NON_SYNONYMOUS_CODING	Rv3548c	Rv3548c	Probable short-chain type dehydrogenase/reductase	V_I	Gtc_Atc	120	915
3988401	G	A	X	N				NON_SYNONYMOUS_CODING	Rv3549c	Rv3549c	Probable short-chain type dehydrogenase/reductase	S_L	tCa_tTa	233	780
3997420	G	C	X	N				NON_SYNONYMOUS_CODING	Rv3557c	Rv3557c	Transcriptional regulatory protein (probably TetR-family)	A_G	gCc_gGc	71	603
4002847	C	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv3561	fadD3	Probable fatty-acid-CoA ligase FadD3 (fatty-acid-CoA synthetase)	T_N	aCc_aAc	404	1524
4028029	G	C	X					NON_SYNONYMOUS_CODING	Rv3586	Rv3586	hypothetical protein	E_Q	Gag_Cag	47	1077
4041838	T	C	X	X				NON_SYNONYMOUS_CODING	Rv3598c	lysS	Lysyl-tRNA synthetase 1 LysS (lysine--tRNA ligase 1) (LysRS 1) (l	H_R	cAc_cGc	368	1518
4047039	A	C	X	N	X	X		NON_SYNONYMOUS_CODING	Rv3604c	Rv3604c	Probable conserved transmembrane protein rich in alanine and arg	V_G	gTg_gGg	153	1194
4067044	G	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv3627c	Rv3627c	hypothetical protein	A_V	gCc_gTc	81	1386
4068804	A	C	X	N				NON_SYNONYMOUS_CODING	Rv3629c	Rv3629c	Probable conserved integral membrane protein	L_R	cTg_cGg	84	1098
4070961	G	A	X					NON_SYNONYMOUS_CODING	Rv3631	Rv3631	Possible transferase (possibly glycosyltransferase)	A_T	Gcg_Acg	150	726
4075152	C	T	X					NON_SYNONYMOUS_CODING	Rv3635	Rv3635	Probable conserved transmembrane protein	P_S	Ccg_Tcg	507	1776
4078292	T	G	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv3639c	Rv3639c	hypothetical protein	K_N	aaA_aaC	53	567

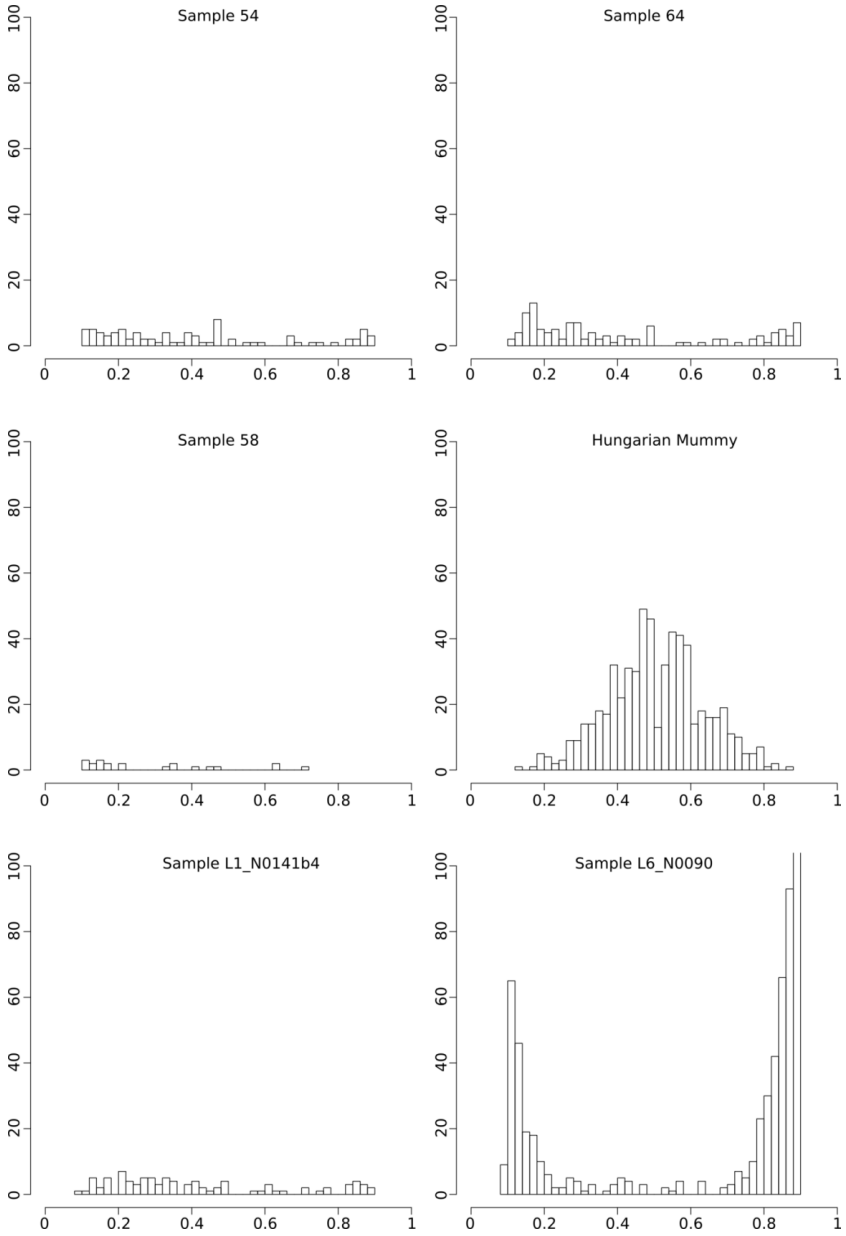
4099060	G	A	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv3660c	Rv3660c	hypothetical protein	P_L	cCg_cTg	30	1053
4106951	G	A	X	N				NON_SYNONYMOUS_CODING	Rv3666c	dppA	Probable periplasmic dipeptide-binding lipoprotein DppA	P_L	cCg_cTg	45	1626
4118789	C	T	X	N				NON_SYNONYMOUS_CODING	Rv3679	Rv3679	Probable anion transporter ATPase	T_I	aCa_aTa	5	1023
4119524	A	T	X					NON_SYNONYMOUS_CODING	Rv3679	Rv3679	Probable anion transporter ATPase	Q_L	cAg_cTg	250	1023
4133256	G	A	X	N				NON_SYNONYMOUS_CODING	Rv3691	Rv3691	hypothetical protein	G_S	Ggt_Agt	247	1002
4140694	A	G	X	N				NON_SYNONYMOUS_CODING	Rv3698	Rv3698	hypothetical protein	I_V	Atc_Gtc	68	1530
4154553	C	A	X					NON_SYNONYMOUS_CODING	Rv3710	leuA	2-isopropylmalate synthase LeuA (alpha-isopropylmalate synthase)	R_S	Cgc_Agc	272	1935
4158032	A	C	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv3712	Rv3712	Possible ligase	K_T	aAg_aCg	351	1242
4169587	G	A	X	N				NON_SYNONYMOUS_CODING	Rv3724A	cut5a	Probable cutinase precursor [first part] Cut5a	V_I	Gtc_Atc	41	243
4172173	A	C	X	N	X	X	X	NON_SYNONYMOUS_CODING	Rv3726	Rv3726	Possible dehydrogenase	R_S	agA_agC	251	1194
4179368	C	A	X	X				NON_SYNONYMOUS_CODING	Rv3729	Rv3729	Possible transferase	R_S	Cgc_Agc	362	2331
4190916	G	A	X	X	X	X		NON_SYNONYMOUS_CODING	Rv3740c	Rv3740c	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	R_C	Cgc_Tgc	422	1347
4214338	G	T						NON_SYNONYMOUS_CODING	Rv3768	Rv3768	hypothetical protein	S_I	aGc_aTc	90	360
4217456	G	A	X	X				NON_SYNONYMOUS_CODING	Rv3772	hisC2	Probable histidinol-phosphate aminotransferase HisC2 (imidazole a	R_H	cGc_cAc	108	1062
4220876	T	G	X	N				NON_SYNONYMOUS_CODING	Rv3775	lipE	Probable lipase LipE	Y_D	Tac_Gac	398	1248
4300626	C	A	X					NON_SYNONYMOUS_CODING	Rv3826	fadD23	Probable fatty-acid-AMP ligase FadD23 (fatty-acid-AMP synthetase	A_D	gCc_gAc	272	1755
4315537	G	C	X					NON_SYNONYMOUS_CODING	Rv3842c	glpQ1	Probable glycerophosphoryl diester phosphodiesterase GlpQ1 (gly	A_G	gCc_gGc	9	825
4330238	C	T	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv3858c	gltD	Probable NADH-dependent glutamate synthase (small subunit) Glt	R_H	cGt_cAt	423	1467
4339149	G	T	X	N				NON_SYNONYMOUS_CODING	Rv3863	Rv3863	Unknown alanine rich protein	G_C	Ggc_Tgc	101	1179
4369000	T	C	X					NON_SYNONYMOUS_CODING	Rv3886c	mycP2	Probable alanine and proline rich membrane-anchored mycosin My	M_V	Atg_Gtg	391	1653
4371675	G	C	X					NON_SYNONYMOUS_CODING	Rv3887c	eccD2	ESX conserved component EccD2 ESX-2 type VII secretion systen	P_A	Ccg_Gcg	4	1530
4376768	C	T	X	X				NON_SYNONYMOUS_CODING	Rv3894c	eccC2	ESX conserved component EccC2 ESX-2 type VII secretion systen	A_T	Gcc_Acc	1229	4191
4385917	G	A	X	N				NON_SYNONYMOUS_CODING	Rv3900c	Rv3900c	Conserved hypothetical alanine rich protein	A_V	gCg_gTg	131	936
4408570	G	A	X	X	X	X	X	NON_SYNONYMOUS_CODING	Rv3920c	Rv3920c	Conserved protein similar to jag protein	R_W	Cgg_Tgg	110	564



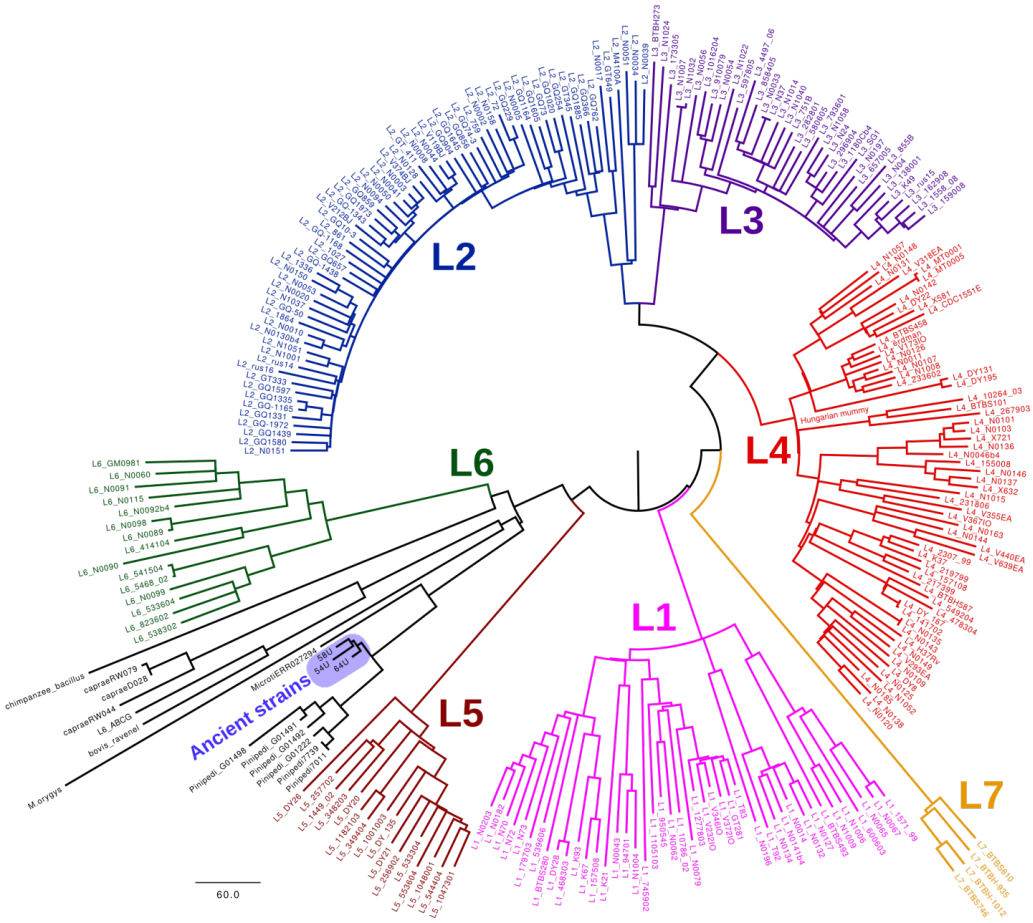
Extended data Figure 1 – Coverage and damage plots for the *M. tuberculosis* capture regions for samples 54, 58, and 64.



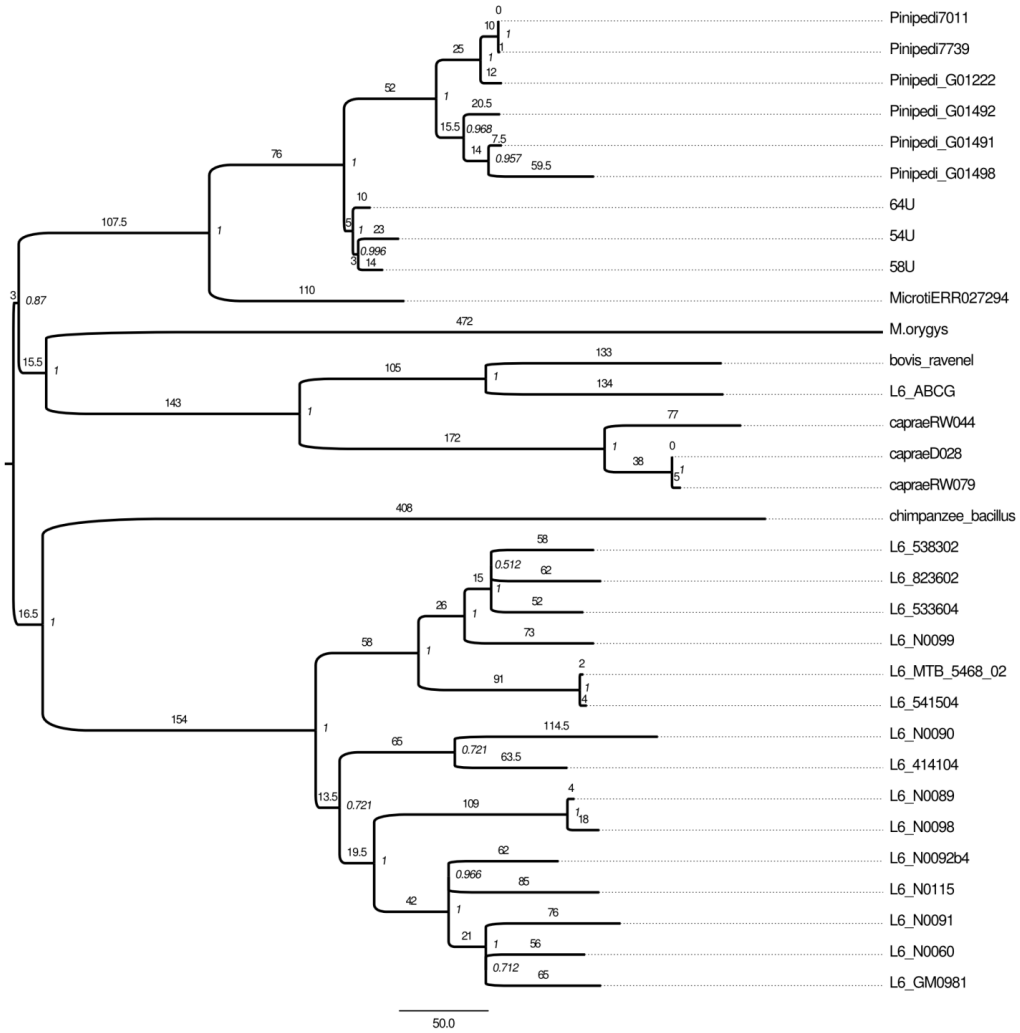
Extended Data Figure 2 - Histograms of SNP allele frequency distributions for the ancient samples and the Hungarian mummy sample using standard mapping parameters. The x-axis denotes the frequency of reads covering a SNP position in which the SNP was detected. The y-axis denotes the number of observed number of SNP calls with the respective frequency. All variants with a SNP allele frequency below 90% are shown.



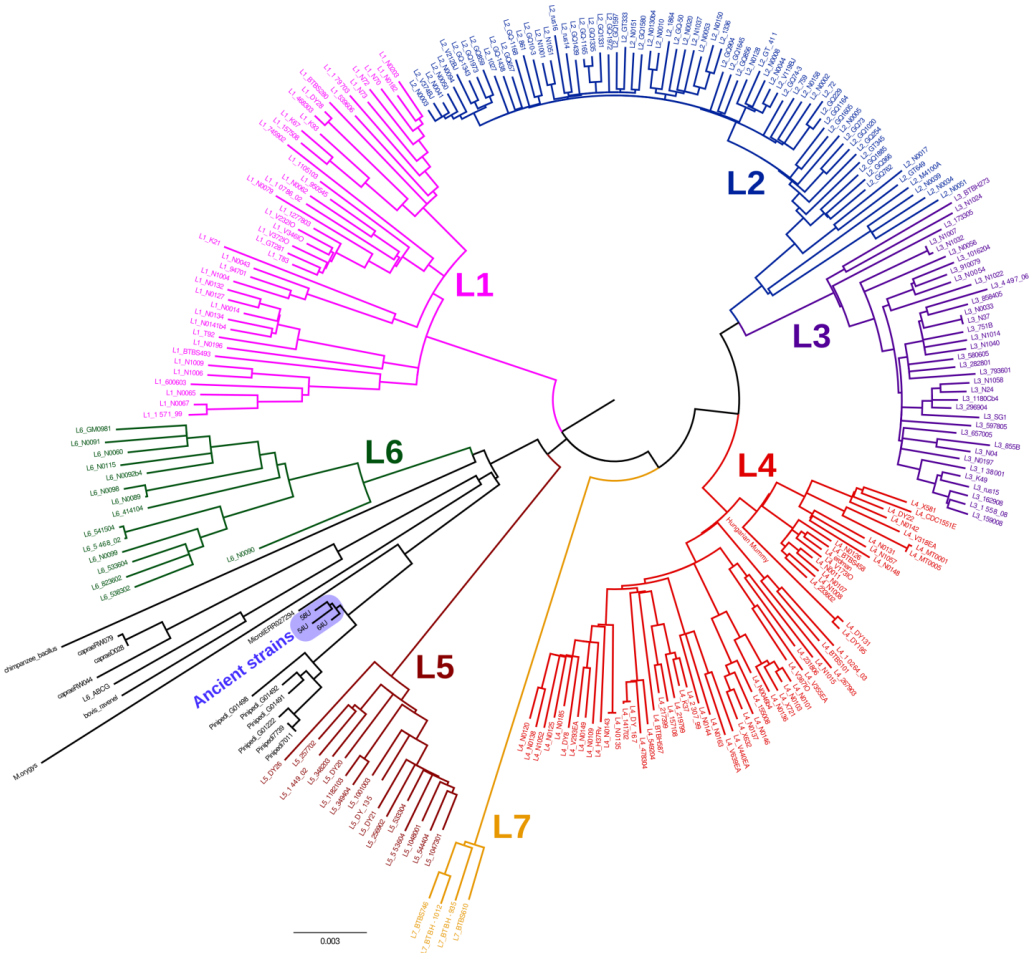
Extended Data Figure 3 -- Histograms of SNP allele frequency distributions for the ancient samples, the Hungarian mummy sample and two modern isolates using stricter mapping and filtering parameters. The x-axis denotes the frequency of reads covering a SNP position in which the SNP was detected. The y-axis denotes the number of observed number of SNP calls with the respective frequency. All variants with a SNP allele frequency below 90% are shown.



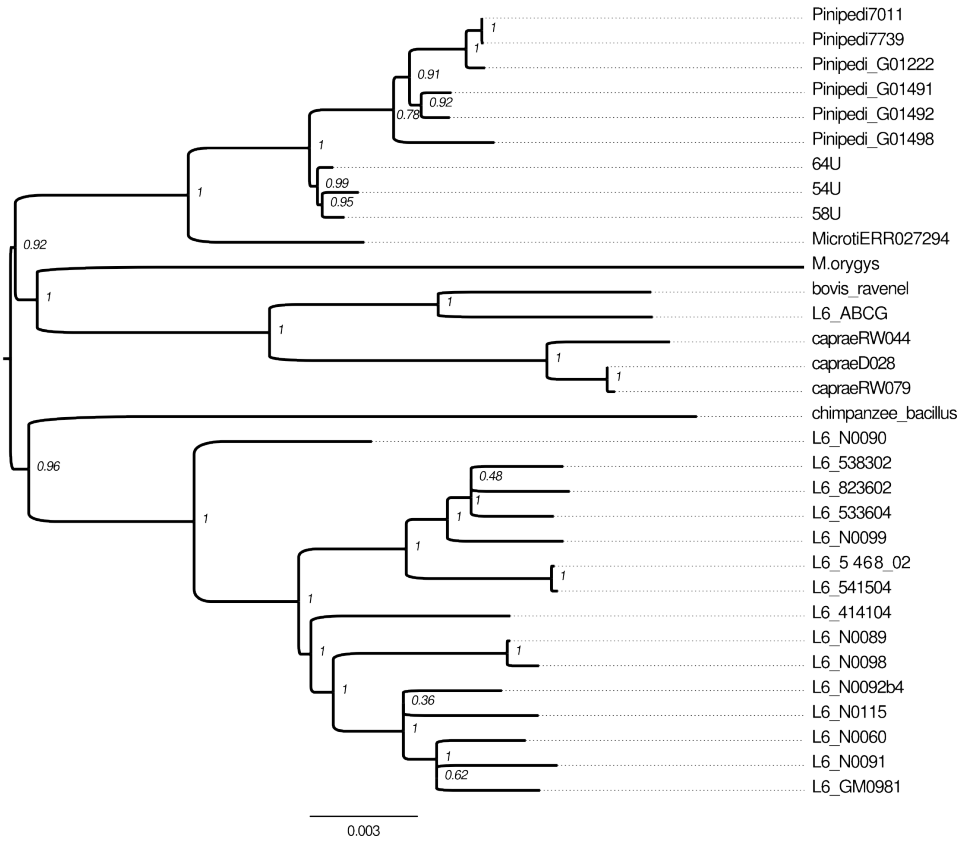
Extended Data Figure 4 - Maximum Parsimony tree of all 262 samples of the complete data set. Positions with missing data were excluded.



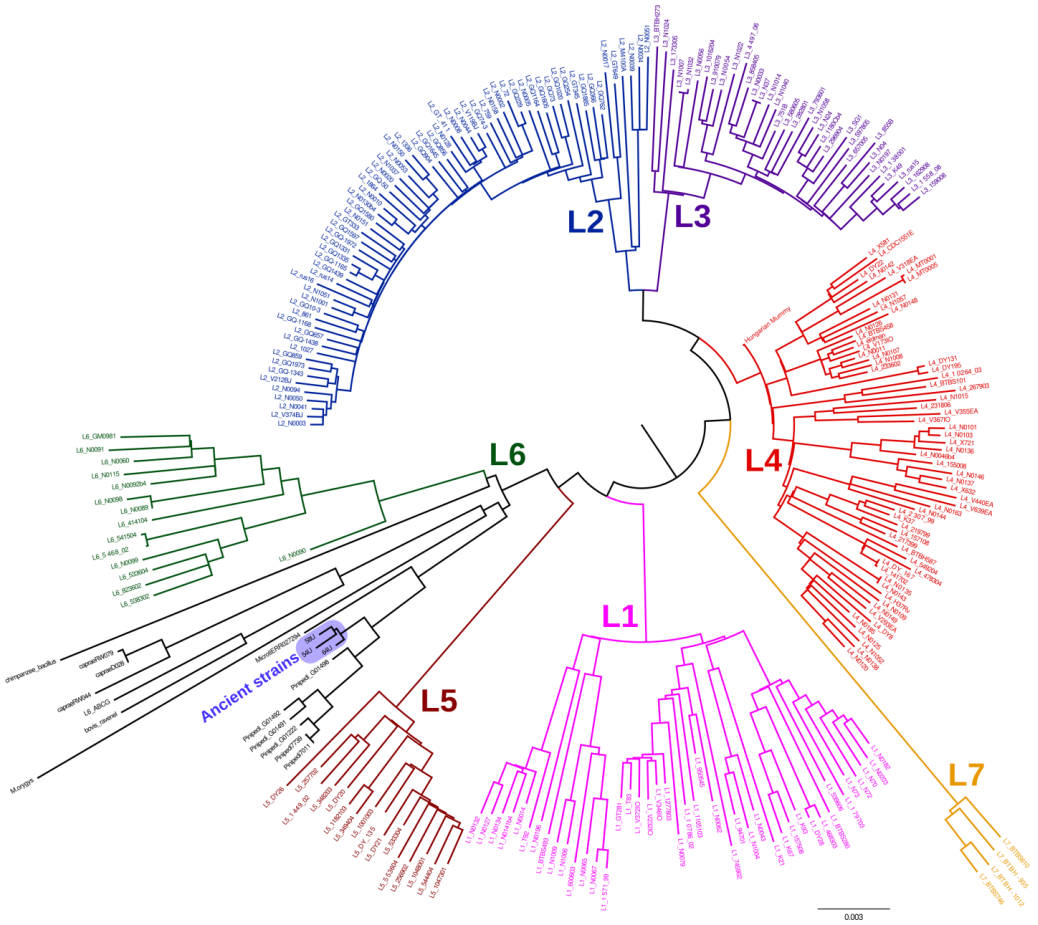
Extended Data Figure 5 - Subtree of the full Maximum Parsimony tree showing the lineage 6 and animal strains. Positions with missing data were excluded. Branches are labeled with the absolute number of substitutions. Internal nodes are labeled with bootstrap statistics obtained from 1000 replicates.



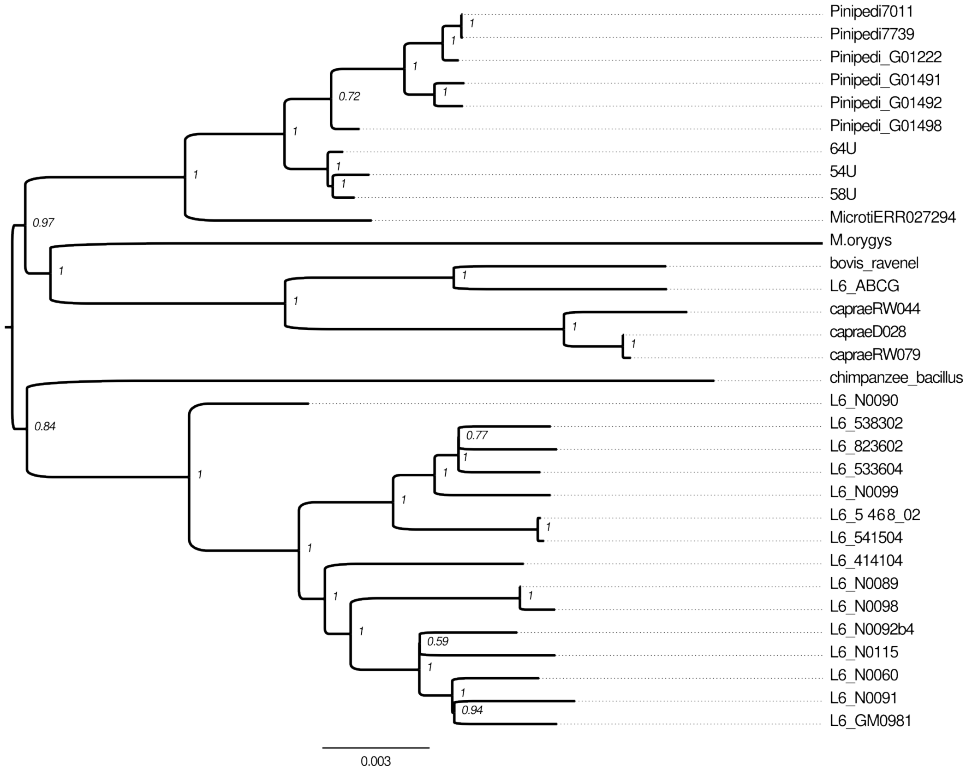
Extended Data Figure 6 - Maximum Likelihood tree of all 262 samples of the complete data set. Positions with missing data were excluded.



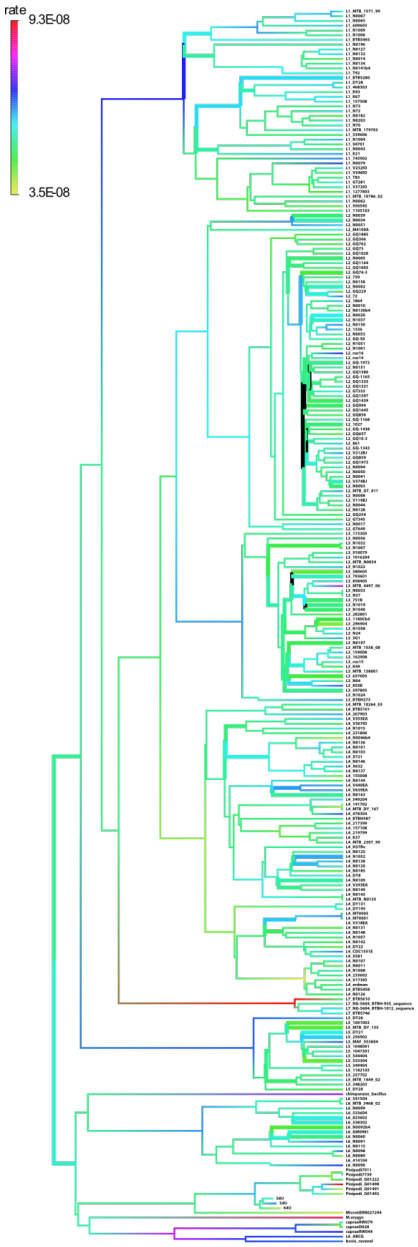
Extended Data Figure 7 – Maximum Likelihood subtree showing the lineage 6 and animal strains. Positions with missing data were excluded. Internal nodes are labeled with bootstrap statistics obtained from 200 replicates.



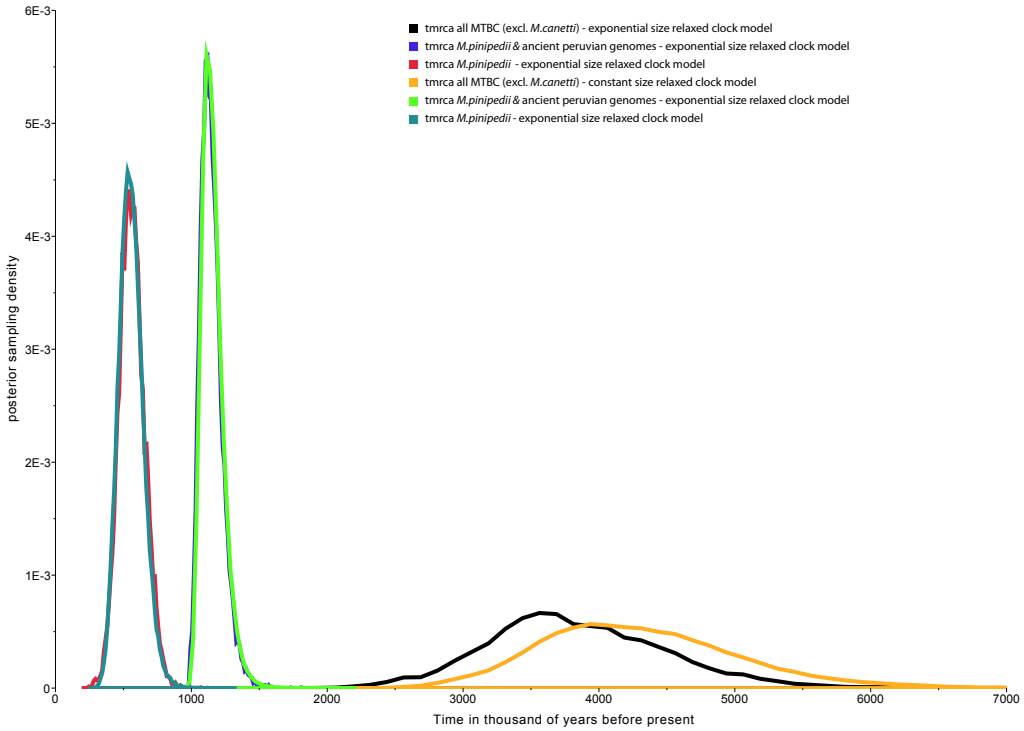
Extended Data Figure 8 – Neighbor Joining tree of all 262 samples of the complete data set. Positions with missing data were excluded.



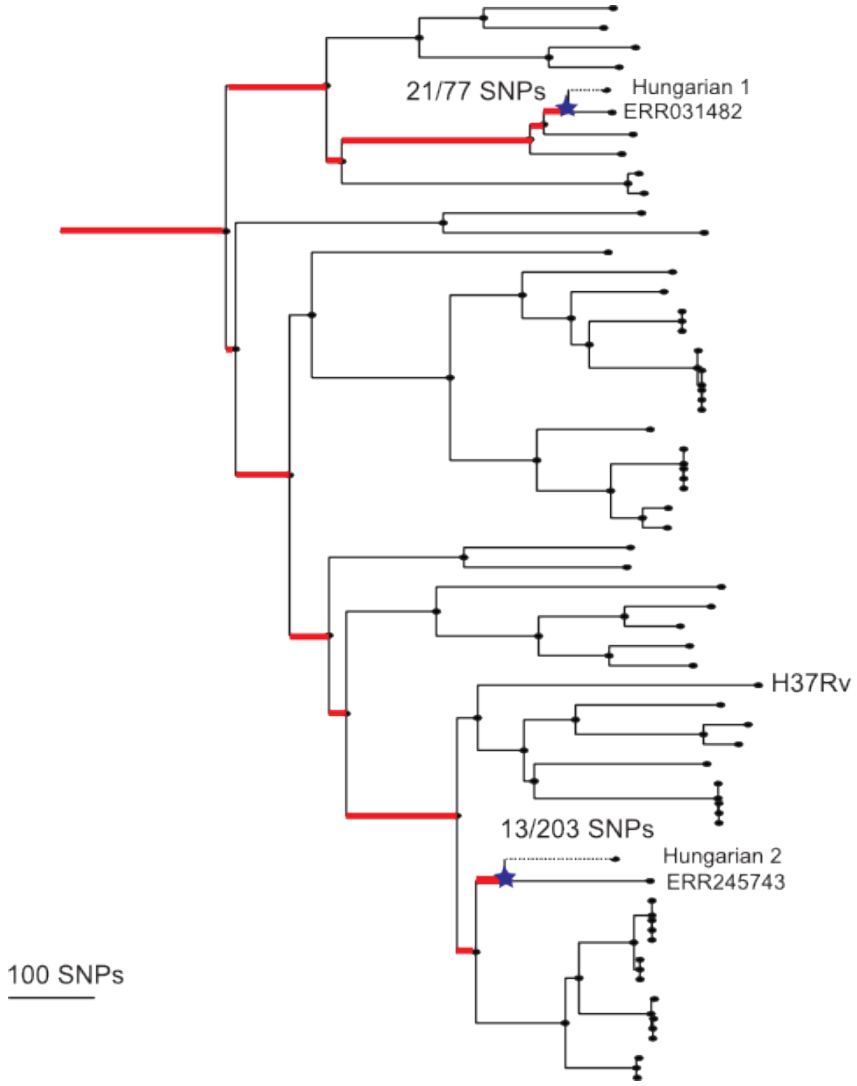
Extended Data Figure 9 – Neighbor Joining subtree showing the lineage 6 and animal strains. Positions with missing data were excluded. Internal nodes are labeled with bootstrap statistics obtained from 1000 replicates.



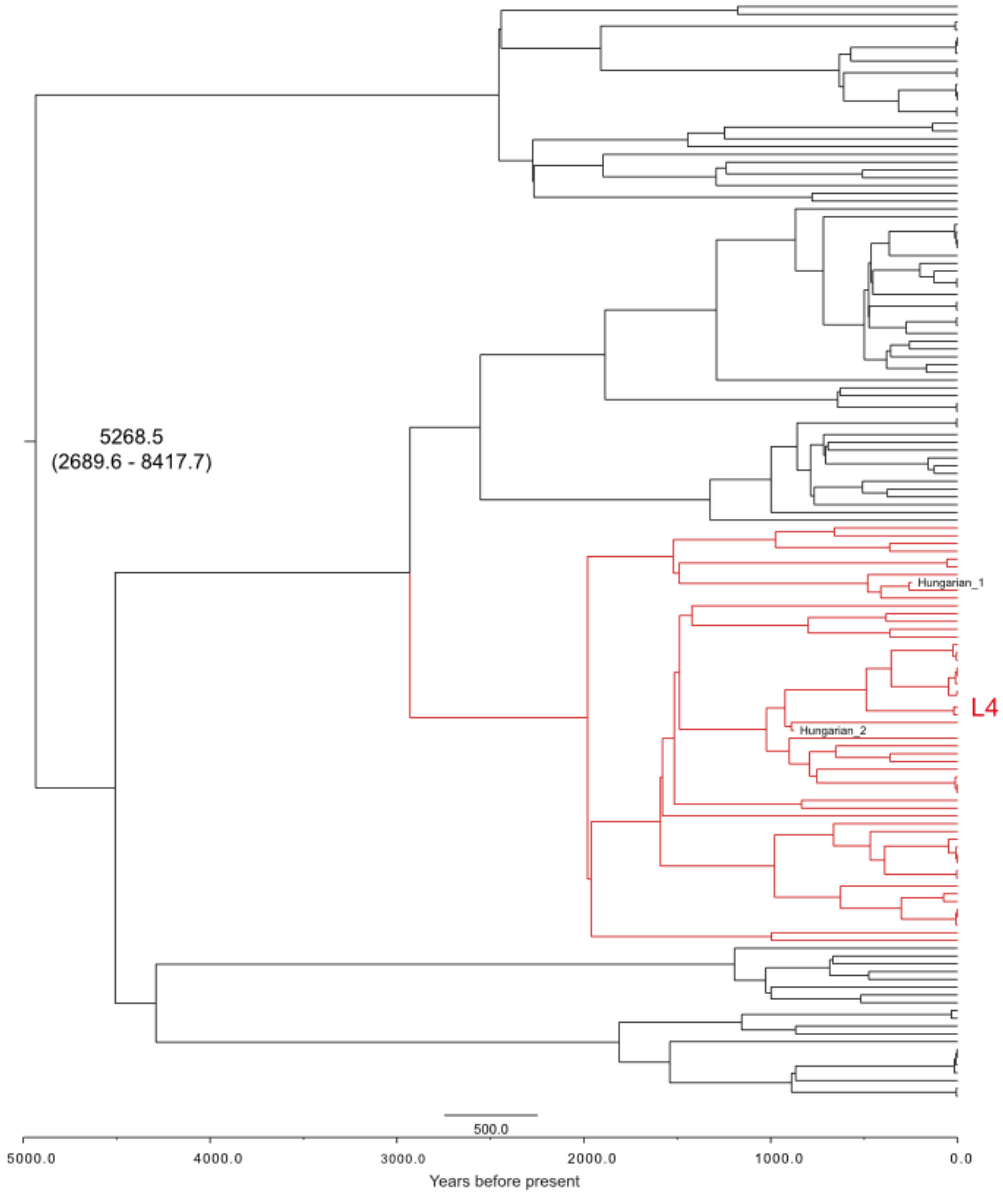
Extended Data Figure 10 - A maximum clade credibility tree of *M. tuberculosis*. The tree was estimated using the uncorrelated lognormal relaxed clock model in BEAST 1.7.5[74]. The radiocarbon dates of the ancient Peruvian strains were used as temporal estimates to date the tree. Branch lengths are scaled to years. Branch colors indicate the estimated branch substitution rate on the logarithmic scale shown in the legend at the left.



Extended Data Figure 11 – Posterior distributions of times to most recent common ancestor (TMRCA) for different MTBC branches and exponential growth and constant size models



Extended Data Figure 12 - Maximum likelihood phylogeny of L4 lineage including modern and ancient strains. The mixed samples are separated out into Hungarian 1 and 2. SNPs were mapped back onto the phylogeny, and branches marked in red are those that are defined by variants found to be mixed in the Hungarian sample. This allowed us to determine the ancestral nodes and branches for each of the two strains on the tree. The dotted lines represent the unknown length of the terminal branches with the stars representing the theoretical penultimate node for which age priors were determined.



Extended Data Figure 13 - Maximum clade credibility tree produced using BEAST (Drummond and Rambaut 2007). Produced using TreeAnnotator from 9,000 trees. Branch lengths are scaled by age. The mean age (ybp) of the MRCA plus 95% HPD, and the position of the separated Hungarian ancient strains are marked on the phylogeny.

APPENDIX C

PERMISSIONS

The collaborating labs and co-authors are in agreement about inclusion of this material in the dissertation.