Beyond Reductionism and Emergence: A Study of the Epistemic Practices in Gene Expression Research

by

Valérie Racine

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

Approved September 2016 by the
Graduate Supervisory Committee:

Jane Maienschein, Co-Chair
Manfred Laubichler, Co-Chair
Richard Creath
Michel Morange
Stuart Newfeld

ARIZONA STATE UNIVERSITY

December 2016

ABSTRACT

A central task for historians and philosophers of science is to characterize and analyze the epistemic practices in a given science. The epistemic practice of a science includes its explanatory goals as well as the methods used to achieve these goals. This dissertation addresses the epistemic practices in gene expression research spanning the mid-twentieth century to the twenty-first century. The critical evaluation of the standard historical narratives of the molecular life sciences clarifies certain philosophical problems with respect to reduction, emergence, and representation, and offers new ways with which to think about the development of scientific research and the nature of scientific change.

The first chapter revisits some of the key experiments that contributed to the development of the repression model of genetic regulation in the *lac* operon and concludes that the early research on gene expression and genetic regulation depict an iterative and integrative process, which was neither reductionist nor holist. In doing so, it challenges a common application of a conceptual framework in the history of biology and offers an alternative framework. The second chapter argues that the concept of emergence in the history and philosophy of biology is too ambiguous to account for the current research in post-genomic molecular biology and it is often erroneously used to argue against some reductionist theses. The third chapter investigates the use of network representations of gene expression in developmental evolution research and takes up some of the conceptual and methodological problems it has generated. The concluding comments present potential avenues for future research arising from each substantial chapter.

In sum, this dissertation argues that the epistemic practices of gene expression research are an iterative and integrative process, which produces theoretical representations of the complex interactions in gene expression as networks. Moreover, conceptualizing these interactions as networks constrains empirical research strategies by the limited number of ways in which gene expression can be controlled through general rules of network interactions. Making these strategies explicit helps to clarify how they can explain the dynamic and adaptive features of genomes.

ACKNOWLEDGMENTS

I am greatly indebted to my co-advisors, Jane Maienschein and Manfred Laubichler, not only for their critical feedback and encouragement as I wrote this dissertation, but for their enormous generosity and support they provided during my time at ASU. Not only are they both extremely talented scholars at the very top of their fields, they are also incredible mentors and people. They have both provided examples of the very best of the profession and I am so grateful for their insights into the discipline and academia as whole. Thank you so much! I have been extremely fortunate to have Richard Creath, Stuart Newfeld, and Michel Morange, all of whom are incredibly talented and distinguished in their respective disciplines, as members of my examination committee. I consider myself extremely lucky to have benefited from their critical feedback on my work. I am so very grateful for all of your help and your insights! I also owe many thanks to Charles Weijer for his continued support and encouragement over the years, and I hope to emulate the example he set for me to my future students. Thank you to Ehab Abouheif for his support, and for his generosity in inviting me into his lab during my summer in Montreal. Many thanks are due to Ben Hurlbut, Karin Ellison, and Joseph Herkert for the opportunities to work for them during my time at ASU. I am a better teacher and a better scholar because of those opportunities to learn from you. Thank you! I am also grateful for my extraordinary colleagues and friends in ASU's Center for Biology and Society, especially Steve Elliott, J.J. LaTourelle, Kate MacCord, and Erica O'Neil. I am very grateful to all of you for your critical feedback on my work, for your insights, and especially for contributing to making the Center such a great place to work. Thanks are also due to the staff members in SOLS and the Center for all the financial and

administrative support I received during my time at ASU. I would like to extend special thanks to Jessica Ranney, Andrea Cottrell, Yvonne Delgado, and Wendi Simonson, for their endless optimism and eagerness to help. Last but most importantly, I would like to express my gratitude to my family and to my husband. I am forever grateful for the endless love, support, and generosity of my parents, René and Lena. I could not have made it this far without them. Finally, I want to express my utmost gratitude to my husband and a brilliant philosopher, Wes, for his patience, encouragement, and love. I hope I have made you proud!

TABLE OF CONTENTS

Page

LIST OF FIGURES ix

INTRODUCTION 1

THE MECHANISTIC-HOLSITIC DIVIDE REVISTED: THE CASE OF THE LAC OPERON 9

Introduction 9

Revisiting the Development of the lac Operon Model 11

Jacob and Monod's Path to the Problem of Bacterial Enzyme Induction 12

The Pajama Experiments and the Inhibitor 18

The Operon as a Model for the Regulation of Gene Expression 23

Allen's Application of the Mechanistic-Holistic Dichotomy 27

Tracking Processes: Moving Beyond Dichotomies for Particular Episodes of Scientific Practice 33

Implications for Historiography: Mode and Tempo in the History of Science 36

Summary 40

THE MICRO-RNA WORLD: CONTINUITIES AND DISCONTINUITIES IN POST-GENOMIC RESEARCH 41

Introduction 41

From Pre-genomics to Post-genomics: A Gradual Continuation in the History of Molecular Biology 44

A Standard Narrative of the Conceptual Shift in Post-genomic Molecular Biology 44

Page

Woese 46
Keller 47
Morange 50
Burian 51
Micro-RNA: A Case Study of Scientific Practices in Post-genomics Molecular Biology 53
What are miRNAs? 55
miRNA Biogenesis and Functions 56
A Brief History of miRNA Discovery 58
Methodological Approaches and Strategies in miRNA Research 60
Post-genomic miRNA Research: A Continuation of Iterative and Integrative Research in Molecular Biology 64
A Case Against Anti-Reductionism (and Reductionism) in Post-genomic Molecular Biology 66
The Different Varieties of Reductionism & Emergence 66
Dupré's Arguments 71
Dupré's Reductionist Principle & Strong Emergence 72
Context Dependency in Molecular Genetics 74
Dupré's Anti-Reductionist Position 76
Context-Dependency Without Anti-Reductionism 77
Explanations and Representations of Interactions in Molecular Biology 77
miRNA Research in the Context of Cancer 79

Page

Context-Dependency does not Entail Anti-Reductionism (or Reductionism) 83

Conclusion 85

THE EPISTEMIC VALUE OF GENETIC NETWORKS IN DEVELOPMENTAL EVOLUTION 88

Introduction.....88

Developmental Evolution as a Research Program.....89

Evo-Devo and Developmental Constraints..... 90

Devo-Evo and the Mechanisms of Development 99

The Application of Developmental and Genetic Mechanisms in Devo-Evo Research 104

Patterns in Morphological Evolution 105

Convergent vs. Parallel Evolution 106

Beetle Horns & the Co-option of Developmental Genetic Pathways.....107

Between Homology & Homoplasy.....111

Homology & Evolutionary Novelty..... 115

Abouheif's Hierarchical Concept of Homology.....115

Wagner's ChIN Concept & the Genetic Theory of Homology 117

From Homology to Networks.....122

The Value of Network Strategies in Devo-Evo 124

The Challenges of Identifying New Genes and New Networks 125

Relational Properties of Regulatory Systems in Gene Networks: A Proposal for a Combined Strategy..... 130

Page

Conclusion 134

CONCLUSION 136

Summary 136

Future Avenues of Research 137

REFERENCES 145

LIST OF FIGURES

Figure Page

1: Results from the PaJaMo Experiment (Pardee, Jacob, & Monod 1959, 172). Reprinted from *Journal of Molecular Biology*, Vol. 1, Pardee, A. B., Jacob F., & Monod, J., "The Genetic Control and Cytoplasmic Expression of 'Inducibility' in the Synthesis of β-galactosidase by *E. coli*," 165-178, (1959), with permission from Elsevier..... 21

2: "Genetic Regulatory Mechanisms in the Synthesis of Proteins" (Jacob & Monod 1961, 344). Reprinted from *Journal of Molecular Biology,* Vol. 3, Jacob, F., & Monod, J, "Genetic Regulatory Mechanisms in the Synthesis of Proteins," 318-356, (1961), with permission from Elsevier. 26

3: Biogenesis of miRNA (He & Hannon 2004, 524). 57

4: Genetic and Genomic Approaches to miRNA Gene Discovery (Ambros 2004, 352) . 64

INTRODUCTION

"On se rend compte peu à peu que la molécule, c'est à la fois beaucoup et pas assez" (Gros 1986, 167).

This dissertation addresses some of the epistemic practices in gene expression and gene regulation research. The epistemic practices of a scientific field include its explanatory goals, approaches, and strategies. A science's approaches and strategies are used to achieve its explanatory goals. However, approaches and strategies are distinct. Throughout the dissertation, I use the term *approaches* to signify the set of methods in particular fields, or sub-fields, of molecular biology, such as biophysics, crystallography, molecular genetics, biochemistry, genomics, computational biology or bioinformatics, etc. I use the term *strategies* to denote conceptual frameworks or heuristics at work in the molecular life sciences; examples include top-down or bottom-up strategies, and mathematical/structuralist or biological/functionalist strategies.

There are two general questions motivating this dissertation. First, what is an adequate characterization of the approaches and strategies in research on gene expression and genetic regulation throughout the last half century? Second, how do representations of gene expression and genetic regulation constrain the strategies that aim to address the dynamic and adaptive features of genomes? To answer these historical and philosophical questions, I look closely at specific case studies, while focusing on the continuation of certain research strategies throughout the history of the molecular life sciences. In this project, I focus on models of genetic regulation, including the *lac* operon model in *E. coli*, the discovery of micro-RNA in *C. elegans* as well as subsequent research on the

different regulatory roles of micro-RNAs, and cases of changes in coordinated gene regulatory responses in the context of developmental evolutionary studies.

I have structured the dissertation into three relatively-independent chapters that investigate three research programs in the molecular life sciences – (1) bacterial genetic regulation in the mid-twentieth century, (2) post-genomic regulatory RNA biology, and (3) gene regulatory networks in developmental evolution. There is, however, a common thread weaving the project together under general themes in the philosophy of science. Before I address this common thread, I first briefly summarize the aims and arguments of each of the three chapters.

In the first chapter, I argue that the research on gene expression and genetic regulation during the dawn of molecular biology was an iterative and integrative process, which was neither reductionist nor holist, contrary to some historical and philosophical accounts of the history of molecular biology. To argue for this claim, I revisit some of the key experiments that contributed to the development of the repression model of genetic regulation in the *lac* operon (Jacob & Monod 1961). I present an analysis of the experiments that François Jacob and Jacques Monod designed to explain the regulation of enzyme synthesis in the bacteria, *E. coli*. I use this paradigmatic case in the history of molecular biology to argue against the application of a common conceptual framework in the history and philosophy of biology based on a dichotomy between mechanistic materialism and holistic materialism.

This dichotomy has been used to explain many of the scientific developments in twentieth century biology in terms of changes that began from a more practical mechanistic or reductionist strategy to a more complete and holistic one (Allen 1978). In

its place, I offer an alternative framework based on the notion of tracking processes to interpret and analyze particular episodes in the history of biology, and I consider the implications of this framework for the historiography of science. One of the upshots of this conceptual framework over one that emphasizes a dichotomy between reductionist and non-reductionist research strategies is that it is more useful for understanding the instances of convergence of the different concepts and experimental techniques from distinct scientific research traditions. That convergence has often proved to be crucial in generating and addressing empirically tractable research problems, as was the case in Jacob and Monod's fruitful collaboration on the problem of enzyme induction. This framework can better accommodate some of the details of scientific practice within particular historical episodes, which, in turn, can offer insights into the process of science at a finer grain of resolution.

In the second chapter, I investigate both the continuities and discontinuities between pre-genomic molecular biology and current post-genomic work. I argue first that many *strategies* and practices in post-genomic molecular biology are, in fact, continuous with earlier work in molecular biology, yet post-genomics incorporates many new tools and techniques, or more generally, *approaches*. Second, I argue that the concepts of emergence and context-dependency in the philosophy of biology are often too ambiguous to rigorously account for current research strategies in molecular biology and are often erroneously used to argue against certain forms of reductionism. Both theses are developed through the presentation of case studies in RNA biology.

The first half of chapter two addresses a standard narrative of the shift from the pre-genomic biology of the twentieth century to the current era of twenty-first post-

genomic biology, often presented as a shift from a reductionist science to a different, non-reductionist endeavor. Here, I outline some of the ways in which prominent biologists, philosophers, and historians have tried to make sense of some of the conceptual changes that post-genomic biology has brought about, with an emphasis on how they have defined and conceived of genes and genomes.

I next consider the ways in which some of these changes might have affected the epistemic practices in gene expression research. To do so, I present an overview of micro-RNA research to provide a concrete example of the kind of scientific research and practices that some biologists, philosophers, and historians have in mind when they describe the changing concepts in gene expression research in the post-genomic era. This case study illustrates several characteristics of the practices within current gene expression research. Within current micro-RNA research, biologists have adopted novel technologies, tools, and techniques previously unavailable to most biologists and embraced exploratory experimentation in addition to hypothesis-driven methods in their research. These approaches represent some discontinuities with pre-genomic molecular genetics of the twentieth century. However, the iterative and integrative use of approaches and strategies is continuous with previous episodes in twentieth century molecular biology, such as the one described in chapter one.

In the second half of the chapter, I argue against philosopher John Dupré's (2010) case for anti-reductionism in post-genomic biology. Much of Dupré's argument rests on the claim that the entities in post-genomic molecular genetics are inherently context-dependent. Because context-dependency relies on the emergent properties of a system, reductionism cannot make sense of these entities or their interactions. I argue that much

of current research in molecular genetics can and does take context-dependencies into account by providing representations of molecular entities and how they interact. Context ends up being represented as other entities that interact with the focal entities under investigation. Therefore, the thesis of emergence understood as context-dependency does not entail acceptance of anti-reductionism, as Dupré understands it. In fact, there seems to be no clear relationship (logical or otherwise) between claims of emergence as context-dependency, on the one hand, and claims about reductionism or anti-reductionism, on the other. The success of a strategy that attempts to take into account context-dependency has nothing to do with whether it is reductionist or not, and this has important consequences for how we understand and evaluate the development of research in molecular genetics.

In the third chapter, I investigate how representations of gene expression as networks are widespread and explanatorily fruitful in developmental evolutionary studies. Developmental evolution focuses on the role of development in phenotypic and morphological evolution and the generation of novel types. This chapter is divided into three sections. First, I discuss distinct narratives of the history of evolutionary biology in the twentieth century, and how development came to be seen as essential to the study of evolution. These historical narratives are useful for drawing a distinction between different research programs in evolutionary developmental biology (evo-devo) and developmental evolution (devo-evo). I next present examples of research in developmental evolution in order to address some conceptual and methodological problems this research has generated. These problems include (1) distinguishing between convergent and parallel evolution and drawing inferences about evolutionary processes from such patterns, and (2) defining homology and identifying novelty. And, third, I

distinguish between different strategies to studying gene regulatory networks in developmental evolution. Each strategy is explanatorily promising in some ways, yet limited in others. I focus on some challenges to the strategy based on the biological properties of gene regulatory networks, and propose a framework that combines that strategy with one that takes account of the relational properties of systems of regulation.

Throughout the chapters of this dissertation, I pay close attention to the role of historical narratives in the development of the life sciences over the last century. Such narratives sometimes obscure aspects of the scientific process, especially with respect to changes in the approaches and strategies of a science. Chapter one provides an example of how a narrative based on a dichotomy between reductionism and holism can conceal the iterative and integrative way research in molecular genetics developed within a particular research program.

At the same time, analyzing and comparing historical narratives, with particular attention to how they characterize scientific approaches and research strategies, can also illuminate certain aspects of the process of science, such as theory construction and theory change. Certain aspects of theory construction are revealed in the narratives of the origins of evolutionary developmental biology and developmental evolution, presented in chapter three, while other aspects of theory development are revealed in the narrative of the shift from pre-genomics to post-genomics and the evolving concepts of the gene and genome, presented in chapter two. In this sense, a common thread in the dissertation displays how attention to historical narratives of a science can provide philosophical insights into the process of science and the nature of scientific change.

By taking a closer look into the process of science and the nature of scientific change (both conceptual and experimental) via some leading historical reconstructions of the development of the molecular life sciences in particular, the dissertation also contributes to general issues in the philosophy of science, including the topics of reductionism and emergence, as well as scientific representation. With respect to reductionism, the dissertation addresses the nature and the legitimacy of reductionist strategies in the molecular life sciences. With respect to emergence and the emergent properties of systems, chapter two attempts to separate the concept of emergence from certain methodological and explanatory kinds of anti-reductionism.

In addition, the dissertation engages with how scientific representations of molecular or genetic interactions as regulatory networks have constrained certain research strategies in current molecular genetics and genomics. For example, in chapter two, I suggest that representing the context-dependency of biomolecules, such as regulatory, non-coding RNAs, by foregrounding their causal interactions with other molecular and genetic components represents a continuation with the complementary strategies of molecular genetics and biochemistry of the twentieth century. Similarly, in chapter three, I address the different research strategies used to study gene regulatory networks in developmental evolution. Representing the complex molecular and genetic interactions in development as networks has imposed certain structural, functional, and logical constraints on how such complex interactions can achieve stability over time, yet also how they are likely to evolve. These constraints provide convenient handles for researchers to manipulate or intervene on in experimental systems to test hypotheses about genomes and their evolution.

By critically evaluating the standard historical narratives of the molecular life sciences, and by addressing the philosophical issues of reduction, emergence, and representation, I provide the following answers to the two driving questions of the dissertation stated above. First, the approaches and strategies in gene expression research depict an iterative and integrative process, which is neither reductionist nor holistic, and which tends to produce representations of the complex interactions of gene expression as networks. Second, by representing the complex genetic interactions as networks, researchers can constrain their models by the limited number of ways in which gene expression can be controlled through general rules of network interactions.

THE MECHANISTIC-HOLSITIC DIVIDE REVISTED: THE CASE OF THE LAC OPERON

Introduction[1]

The history and philosophy of science requires a conceptual framework with which to interpret the significance of scientific events or periods, the emergence and acceptance of scientific theories, concepts, models, and metaphors, and the multiple lines of influence between scientists, institutions, and ideas. Most philosophers and historians adopt or develop conceptual frameworks and defend these frameworks as useful for understanding the nature of scientific change or for illuminating some aspect of scientific practice. That has been the case for many scholars addressing episodes in the history of molecular biology, especially the significance of the molecular turn in mid-twentieth century biology and the entrenchment of certain concepts, such as the genetic code and the genetic program. Some major works that have illuminated these aspects of molecular biology include Garland Allen (1978), François Jacob (1970), Horace Freeland Judson (1979), Lily E. Kay (2000), Michel Morange (1998), and Robert Olby (1974).

Often, in these works and elsewhere, episodes in the history of molecular biology are interpreted in terms of shifts from a reductionist or mechanistic approach to studying phenomena to a non-reductionist or holistic one. This framework has been used to explain transformative periods in the history of modern biology, such as the early contributions of biochemistry, physiology, and genetics to molecular biology (Allen 1978), and shifts to new problems in molecular biology, such as the role of genetic

[1] A version of this chapter appears in *Studies in the History and Philosophy of Science Part C: Studies in the History and Philosophy of Biological and Biomedical Science* (Racine 2016).

regulation in cellular differentiation and the developmental programs of multi-cellular eukaryotes (Morange 1997). The framework appears again after the completion of the Human Genome Project to address emerging questions about how to study the functions of non-coding DNA in the post-genomic era (Keller 2005, Woese 2004).

In this chapter, I focus on Allen's influential account of the key shifts in particular research traditions within twentieth century biology to illustrate the conceptual framework at work. Influenced by a dialectical materialist view of historical progression, Allen represents episodes in the history of science through dichotomies and views scientific change as a series of discontinuities and revolts. Allen uses a dichotomy between "mechanistic materialism" and "holistic materialism" to interpret the developments in the fields of physiology, biochemistry, and genetics, all of which converged to form the field of molecular biology in the mid-twentieth century. Imposing this lens on the developments in twentieth century biology leads Allen to make two general inferences about the nature of scientific practice; (1) that "the mechanistic approach has often been the only practical way to begin the study of a complex process," and (2) that the approach described as "holistic materialism" aims to provide a more complete and accurate description of the natural world (Allen 1978, 105-016). Because Allen's view has been influential and continues to receive attention, it is worth looking more closely at his framework in order to get at the implications of thinking about the history of the life sciences in the ways he suggests.

I argue that his generalized claims about scientific practice are doubtful when looking closely at the practices within a particular research program, such as those used by scientists working on the problem of enzyme induction in the labs of the Pasteur

Institute during the mid-twentieth century. To defend this claim, I first present an historical analysis of the development of the *lac* operon model in molecular biology. I then present a critical review of Allen's conceptual framework and his generalizations about scientific developments. I argue that a close look at the scientific practices in early gene expression research depicts an iterative and integrative process that does not progress from a mechanistic or reductionist approach towards a more holistic approach. And, it also presents a challenge to the assumption that the holistic or anti-reductionist approach to representing phenomena is inherently more complete or complex. I offer an alternative framework based on the notion of tracking processes to interpret and analyze episodes in the history of molecular biology in the following section. This alternative considers the ways in which biologists track and prioritize different aspects of biological phenomena over time. I argue that this framework is more useful for the purpose of shedding light on the nature of the conceptual and experimental practices in particular episodes within the history of biology. Finally, I consider some implications for the historiography of science by reflecting on the appropriateness of different conceptual frameworks for different grains of resolution in the history of biology.

Revisiting the development of the *lac* operon model

In this section, I present a study of the concepts and methods used by François Jacob and Jacques Monod (and their colleagues) to construct and justify their model of genetic regulation in bacteria. As is well documented, the Pasteurian scientists made use of many different experimental methods and techniques throughout their collaboration, including crossbreeding methods from classical genetics and induction techniques from biochemistry, as well as conceptual tools and metaphors from cybernetics. I first briefly

outline the research projects on which Jacob and Monod were working to explain how the problem of bacterial gene expression arose. I then show how their experimental designs resulted from a convergence of their previous research projects. Finally, I outline how they developed and justified their model to represent the regulation of the *lac* operon system in *E. coli*. The case is used to emphasize the integrative and iterative nature of the scientific practices involved, and to challenge Allen's general claims about the process of science and its goals.[2]

Jacob and Monod's path to the problem of bacterial enzyme induction

During the 1950s, many researchers at the Pasteur Institute focused their research on lysogenic bacteria (or lysogeny). A lysogenic bacterium is a bacterium infected by a phage, or virus, referred to as a bacteriophage. A phage infects a bacterium and inserts its genetic material, which is composed of either RNA or DNA, into the bacterial host cell's DNA. Within the host cell, the bacteriophage can have two different life cycles. During the lysogenic cycle, the bacteriophage that infects the bacterium is referred to as temperate, or non-virulent, because it does not immediately result in the lysing (destruction) of the bacterial cell. When the bacteriophage inserts its genetic material into the host cell's DNA, it is referred to as a prophage. The prophage's genetic material is then replicated with the rest of the genetic material of the host cell as the bacterium reproduces itself. However, the temperate bacteriophage has the ability to switch to the lytic, or virulent, state under certain conditions. The lytic cycle occurs when the

[2] Although Allen does not specifically address this episode in the history of molecular biology, it is an exemplar of the scientific practices adopted in molecular biology during this period and subsequent decades.

bacteriophage's genome directs the synthesis of enzymes that lyse the bacterial cell, essentially killing the host cell, allowing its progeny phages to disperse and infect other bacterial cells in the surrounding environment (Racine 2014).

When Jacob joined André Lwoff's laboratory at the Pasteur Institute in 1950, his main focus was the phenomenon of prophage induction, or how the phage shifts from the lysogenic to the lytic state (Jacob 1972 [1965], Kay 2000, 208).[3] While in Lwoff's lab, Jacob worked with Élie Wollman on bacteria and bacteriophages, with a focus on the temperate lambda phage (λ-phage) in a particular strain of *E. coli*, *E. coli* K-12. They studied the cellular and genetic properties of lysogenization and virulence, and mapped the *E. coli* K-12 genome using crossbreeding techniques, which they published in 1959, in *La sexualité des bactéries* (Jacob & Wollman 1959; Racine 2014, Wollman & Jacob 1956).

Within their research, Jacob and Wollman performed bacterial conjugation experiments (i.e. chromosomal transfers between bacteria), which led them to observe a phenomenon they called zygotic induction (Wollman & Jacob 1956). In their experiments, they transferred the chromosome from a donor lysogenic bacterium, referred to as male, into a receptive non-lysogenic bacterium, referred to as female. During this process, the receptive bacterium becomes temporarily partially diploid, referred to as a merozygote, because it possesses two copies of the chromosomal segments (Grmek & Fantini 1982, Kay 2000, Schaffner 1974, Wollman & Jacob 1956). Jacob and Wollman noticed that the transfer of genetic material induced the lytic cycle in

[3] Their research on lysogeny was influenced by the rich tradition in microbiological work conducted by their predecessors at the Pasteur Institute (See futher: Burian & Gayon 1999).

the receptive merozygotic bacterium, and so, they named the phenomena "zygotic induction" to draw attention to the *induced* lytic state in the *merozygotic* bacterium (Grmek & Fantini 1982, 200). From these experiments, Jacob and Wollman also established that the chromosome from the donor entered the receptive bacterium in a linear and unidirectional way, at constant rate (Wollman & Jacob 1956). They did this by interrupting the process of chromosome transfer between bacteria at different time intervals with the help of a Waring blender. These interventions would sever the chromosomes at different locations during the process at different time intervals, which served to create a genetic map of the bacterial chromosome. Thus, their series of experiments on lysogenic bacteria enabled the researchers to localize the phage's genetic material on a precise segment of the bacterial chromosome, and to establish that the phage's genes were immediately expressed after entering the receptive bacterial cell.

Jacob's work on lysogeny during the 1950s and his experimental knowledge of bacterial conjugation proved instrumental in his later experiments with Monod, which led to the establishment of the *lac* operon model. Jacob and Wollman's experimental system—merozygotic bacteria—provided a simple and effective way to study the problem which preoccupied Monod, the genetic control of enzymes in *E. coli*. During their collaboration, Jacob and Monod made use of the bacteria's temporary diploid state to track genetic dominance and used classical and reverse genetics methods to identify the function of particular genes or alleles in the bacterial and phage genomes. Their ability to engineer their experimental system in order to design experiments that could

test their working hypotheses was perhaps one of their greatest achievements and is now a commonplace practice in molecular genetics.[4]

At the other end of the attic of the Pasteur Institute, Monod worked on a different, but related problem: the synthesis of the enzyme beta-galactosidase (ß-gal) in *E. coli*. Monod's work with Melvin Cohn in 1953 had revealed some interesting features of enzyme synthesis in *E. coli* (Monod 1966). When *E.coli* is grown on lactose, the bacteria synthesize the enzyme, ß-gal, to metabolize, or break down, lactose. Lactose is considered to be an exogenous inducer of the enzyme. However, during their collaboration, Monod and Cohn noticed a result they found peculiar. They observed that other types of sugars could serve as exogenous inducers of the enzyme, ß-gal, even though the enzyme could not metabolize those other sugars. In other words, those other sugars acted as inducers, but not as substrates, to enzyme, ß-gal. This result led Monod to question the previously-held theory that a precursor molecule transformed into the functional enzyme, ß-gal, only after interaction with the substrate. The presence of these "free inducers" led Monod and Cohn to believe that the function of inducer must be distinct from that of the substrate (Grmek & Fantini 1982, 197; Monod & Cohn 1952,

[4] As Craver and Darden note: "[T]he logic of experimentation in contemporary biology often demands that the experimental system be isolated and prepared so that the experiments can shape the space of possible mechanisms. …In many cases, this engineering is not antecedent to the experiment, something readily cordoned off as a background against which the intervention takes place. Instead, the preparation of experimental systems is itself crucial to understanding just how the experiment works. It is only in the context of such contrived experimental systems that the proposed interventions can meaningfully be interpreted. Model organisms and model systems are rightly prized by biologists precisely because standard organisms and systems have been prepared in such a way as to afford the researcher a particular kind of leverage over the mechanism that could not be achieved or presumed without the active construction of the model system" (Craver & Darden 2013, 138-141).

Monod 1966).[5] Monod's studies with Cohn were the first step towards a clear articulation of the problem of the genetic control of enzyme synthesis in *E. coli* precisely because they were able to think about the inducing "factor" and the substrate separately.

The next step involved Monod's later collaboration with Georges Cohen, in which they investigated so-called "cryptic" bacterial mutants. One of these mutant strains, referred to as "Lac-," could synthesize ß-gal but could not metabolize lactose. They hypothesized that there existed another enzyme—a "permeation factor" or permease—in non-mutant strains. This additional enzyme was thought to allow lactose to permeate the cell so that the non-mutant strains could metabolize it. The idea, then, was that these "cryptic" mutants had dysfunctional permeases. The permease was later experimentally isolated by Eugene Kennedy (Monod 1966, Kay 2000, 206).

Their study of the permease led Monod and his collaborators towards the final piece of the puzzle. In particular, they next turned their attention towards strains of *E. coli* that synthesized ß-gal even in the absence of lactose. These bacteria were said to have phenotypes that were constitutive for ß-gal. They noticed also that the phenotypes that were constitutive for ß-gal were also constitutive for permease. From this observation, they inferred that the two enzymes might be genetically linked. Consequently, Monod

[5] The existence of "free inducers" led Monod to question the appropriateness of the term enzyme *adaptation* to describe the phenomenon (Cohn *et al.* 1953; Cohn 1980). The term *adaptation* carries with it connotations of an advantage or an increase in fitness. Yet it is not uniformly advantageous to synthesize an enzyme when induced to do so, for sometimes those inducers cannot even be metabolized. In light of this, Monod and Cohn, along with Martin R. Pollock, Sol Spiegelman, and Roger Y. Stainier wrote a note entitled, "Terminology of Enzyme Formation," published in *Nature* in 1953, to propose a change in terminology to replace enzyme *adaptation* with enzyme *induction*, which they regarded as a much more neutral term. Some have interpreted their proposal to replace the terms enzyme *adaptation* with enzyme *induction* as an effort to avoid confusion or implications of teleology, and to challenge vitalist and neo-Lamarckian thinking in French biology (Cohn 1980, 4-5; Kay 2000, 197-198).

came to understand the problem of enzyme induction as one of genetic control. He hypothesized that there must be a genetic factor that had a pleiotropic effect on the other genes determining the enzymes, ß-gal and permease. He constructed this hypothesis by classifying mutants by three variables, *z, y, i,* in their efforts to understand the synthesis of ß-gal (Grmek & Fantini 1982, 199; Jacob & Monod 1961, 328; Kay 2000, 206). Types *z*+ could synthesize ß-gal and types *z*- could not. Types *y*+ could synthesize permease and types *y*- could not. And, the inducible wild type *i*+ synthesized both ß-gal and permease in the presence of lactose, whereas the constitutive type *i*- could also synthesize both enzymes even in the absence of lactose. At that time, Monod held the "generalized induction theory," which predicted that the crossing of an i^+ type with a i^- type would result in the dominance of the constitutive type, as the latter variant, he thought, might be responsible for producing an endogenous inducer, which the inducible phenotype could not produce and therefore relied on exogenous inducers (Schaffner 1974, 368-369).

By 1957, Monod was set to tackle the problem of locating the genes involved in this lac system and discovering the mechanisms of the genetic control responsible for the observed phenotypes (Grmek & Fantini 1982, 199). To address this problem of enzyme induction in the lac system, Monod needed tools to adapt classical genetic methods to his investigation. Jacob and Wollman's merozygotic bacteria provided the perfect experimental system to pursue a deeper understanding of enzyme induction in the lac system because they could use the partially diploid systems to track gene function. Given this, the stage was set for the two Pasteurian scientists to tackle the problem. Jacob was an expert with an experimental system that was ideal to design and test genetic hypotheses, and he had some experience inducing the lytic cycle in lysogenic bacteria,

which (as we shall see) later helped them infer a generalized model of negative regulation. Monod had been preoccupied with the induction of a certain metabolic enzyme and, through a series of experimentation and re-conceptualizations, he had gained greater understanding of the phenomena, and he was prepared to address the genetic control of that phenomena.

The Pajama experiments and the inhibitor

In 1957, Jacob and Monod began to collaborate, along with Arthur Pardee, who was on a sabbatical at the Pasteur Institute from the University of California, Berkeley. They worked together on experiments with the *E. coli* K-12 system to better understand the mechanism of induction of the enzyme β-gal by lactose in *E. coli*. From this collaboration came the "PaJaMo" or "Pajama" experiment, named after the researchers' first letters of their last names (Pardee *et al.* 1959; Cobb 2015, 149). In these experiments, they introduced mutations into the genetic circuitry of their experimental system in order to see how each gene contributed to the functioning of the overall system. Their approach became associated with the typical research approach in molecular genetics, in contrast to the approach in biochemistry. This distinction is illustrated in the correspondence between geneticist William Sullivan and biochemist Douglas Kellogg (Stephens 2004). As a geneticist, Sullivan was particularly interested in the genes that are responsible for regulating cell growth and division. Kellogg, a biochemist, focused on proteins and how they carry out their regulatory functions in the cell. In their correspondence, the geneticist was likened to a car mechanic who wanted to understand the functions of all the different components involved in a functioning car. To do this, he intervened in the delivery of specific parts to a manufacturing facility (or intervened in

the order in which the parts were introduced in the car-building process) then waited to see the effects on the functioning of the cars that were produced when they left the plant. Such a process represents the top-down strategy in genetics, where researchers are interested in particular components of a system insofar as the parts are arranged in a particular way and that arrangement can influence or restrict the function of other components, as well as the activity of the entire system. In molecular genetics, this kind of strategy is exemplified in knock-out and knock-down experiments, in which a gene is either knocked out of the genome completely or its transcript is interfered with so that it is prevented from being translated. In contrast, the biochemist was compared to a curious mechanic with a penchant for a more bottom-up strategy, who instead started by taking apart the car and separating each component to study its individual properties. She then attempted to rebuild the car in a piecemeal manner by looking at how the specific components interacted with each other. The notion of the genetic method as a top-down approach is a plausible description of Pardee, Jacob, and Monod's work. They were interested in the control of enzyme induction in the lac system and adopted genetic methods in their research, rather than a strictly biochemical approach, which would have them attempting to isolate the component parts to analyze their properties.

In the Pajama experiments, Pardee, Jacob, and Monod's goal was to identify the genes involved in *E. coli's* ability or inability to synthesize the enzyme, β-gal (Pardee, Jacob, & Monod 1959).[6] They had identified several loci, including *i* and *z*, which they

[6] Pardee, Jacob, and Monod also looked at the effects of the gene locus, *y*, to better understand the regulation of the lactose permease. Another enzyme, galactoside transcaetylase, is also part of the *lac* operon. Hereafter, I focus on *i* and *z* for simplicity, and because they are the crucial elements in the Pajama experiments.

believed to be located in the part of *E. coli*'s genome that was previously established to be involved in lactose metabolism. There were two variants at each loci. Allele z^+ was the wild-type responsible for the synthesis of β-gal, whereas z^- was the mutant that lost the ability to produce the active enzyme. The i^+ allele was the wild-type responsible for the *inducible* phenotype, and the i^- mutant was the type with the continuous production of the enzyme, whether or not an inducer was present, referred to above as the *constitutive* phenotype. To better understand the effects of the genes on the different phenotypes, they performed experiments using Jacob's experimental techniques of bacterial mating. As before, these crosses created merozygotic bacteria. But now, the researchers were tracking the effects of the introduced gene variants on the production of β-gal.

Pardee, Jacob, and Monod, performed several crosses, but their most significant result occurred when they conjugated a wild-type male into a mutant female, that is, a cross between genotype [z^+ i^+] and genotype [z^- i^-]. Before the cross, both types could not synthesize β-gal. In the absence of any inducers, the male [z^+ i^+] could not synthesize the enzyme because of the i^+ variant, and the female [z^- i^-] could not synthesize it because of its z^- variant. Their initial hypothesis about the results of this crossing was that the merozygote would continually produce β-gal, because of both the introduction of the z^+ variant and because they thought that the constitutive type, i^-, was responsible for producing an endogenous inducer that induced the enzyme's continual synthesis. After the cross, they observed the immediate synthesis (i.e. 3 to 4 minutes after injection) of β-gal, which the researchers expected as a result of the introduction of the z^+ gene. However, two hours after the intervention, the synthesis of β-gal ceased. It could be re-initiated by introducing an exogenous inducer (Figure 1). This result was surprising, as it

was contrary to what the researchers were expecting on the basis of their initial hypothesis about the constitutive type. The results revealed that the inducible allele was indeed 'dominant' (albeit, after some time had passed) over the constitutive type.

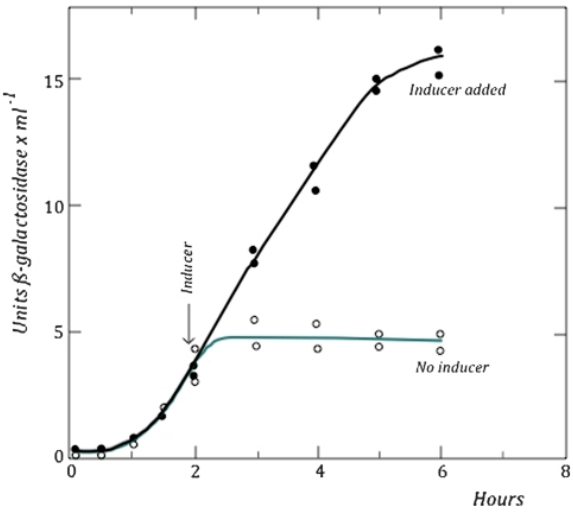

Figure 1: Results from the PaJaMo Experiment (Pardee, Jacob, & Monod 1959, 172). Reprinted from *Journal of Molecular Biology*, Vol. 1, Pardee, A. B., Jacob F., & Monod, J., "The Genetic Control and Cytoplasmic Expression of 'Inducibility' in the Synthesis of β-galactosidase by *E. coli*," 165-178, (1959), with permission from Elsevier.
This graph represents the different rates of enzyme synthesis over time that resulted from the experimenters' crossing between [z+ i+] and [z- i-]. It shows the synthesis of the enzyme, β-galactosidase, first, as a result of the introduction of the z+ variant (without inducer). The synthesis of the enzyme stopped approximately two hours after the intervention. After the addition of an exogenous inducer, the bacteria were able to synthesize β-galactosidase again, as a result of the i+ gene.

In their interpretation of the results, Pardee, Jacob, and Monod, rejected their initial hypothesis and eventually constructed an alternative account of enzyme induction, with the help of Leo Szilard (Monod 1966, Schaffner 1974). When Szilard visited the Pasteur Institute, he suggested to Monod that instead of inducers, either endogenous or exogenous, being directly responsible for the onset of enzyme synthesis, it might be the case that enzyme synthesis resulted from interference in a mechanism in which synthesis was being repressed or inhibited. Inducers, then, would be acting by interfering with the

mechanism of repression. This was directly contrary to Jacob and Monod's initial working hypothesis. When faced with their results, they had to flip their logic. It was not the case that the i^- allele coded for an endogenous inducer. Rather it was the case that the i^+ allele coded for the synthesis of a *repressor* that inhibited enzyme synthesis. The exogenous inducer then functioned by interfering with, or de-repressing, the repressor.[7]

In summary, the Pajama experiment helped expose how the *E. coli* genome itself regulates the synthesis of β-gal. After Szilard's visit, Pardee, Jacob, and Monod hypothesized that a protein, the lac repressor, binds to the gene that synthesizes β-gal when lactose is not present, suppressing the gene's expression. When the inducer is present, the lac repressor protein detaches from the gene, the gene's transcription and translation are activated, and β-gal is synthesized. In 1961, Jacob and Monod published "Genetic Regulatory Mechanisms in the Synthesis of Proteins," in which they explained in more depth the negative regulation of enzyme induction in *E. coli* and proposed a more general model of gene regulation (Jacob & Monod 1961).

During the 1950s, other groups of scientists in American research institutions were also studying the mechanisms involved in enzyme, or more broadly, protein synthesis (Burian 1993). For example, Paul Zamecnik and his colleagues studied tissue slices and developed a cell-free system in which to study the biochemical basis of protein

[7] While Szilard played a role in pushing Monod and Jacob to think about their experiments through a framework of negative regulation, both scientists, especially Monod, were already aware of the concepts and metaphors of cybernetics and they knew about the negative feedback control of the synthesis of precursors to the amino acid tryptophan by tryptophan itself observed by Szilard and Aaron Novick in 1953. Monod even prepared a manuscript in 1959, entitled *Cybernétique enzymatique*, which was never published (Peluffo 2015). I thank an anonymous reviewer for bringing attention to this point. As the reviewer pointed out, Monod's mind "was prepared" to reconceptualise the results of the experiments.

synthesis (Burian 1993, Morange 1998, Rheinberger 1997). They did so by separating the cellular components in an ultracentrifuge and analyzing the properties of the cellular components. Zamecnik's research program contributed to particular facets of the problem of protein synthesis; namely, the role of the ribosomes in translation. In contrast, whereas Jacob and Monod were likewise concerned with the mechanics of protein synthesis, they were also interested in the metabolic processes of the entire cell (not just the cell's individual components). Consequently, they conducted genetic experiments to discover the organization of the system and how it functioned (Racine 2014).

Jacob and Monod's research program, which led to the development of the operon model, is not a clear-cut example of a bottom-up strategy, or a reductionist analysis, of individual genes, enzymes, and their substrates. They did not gradually move from that approach to synthesize a more holistic account of the organization of the genome in the bacterial cell. In fact, as the later justification of their model suggested (see below), their presentation of the *lac* operon model was theoretical and lacked descriptions of some of the molecular entities and activities involved in the process of genetic regulation.

The operon as a model for the regulation of gene expression

Jacob and Monod continued to provide evidence for the negative control of enzyme synthesis and the existence of a protein they called the "lac repressor" (Jacob & Monod 1961). They conjectured that the repressor protein binds to another element—what they called the "operator" or the *o* gene—to suppress the transcription and translation of the metabolic enzymes, like β-gal and permease, when the bacteria exhibit the inducible phenotype. In other words, Jacob and Monod thought that the repressor protein, by either binding or not on the operator, acts as an on/off switch in gene

expression, controlling cell metabolism under different environmental conditions. Jacob and Monod called this "switch" system of gene expression an operon, and thus their model of the lac system in *E. coli* became known as the *lac* operon. In their most-cited 1961 paper, "Genetic Regulatory Mechanisms in the Synthesis of Proteins," their model was pictured graphically. The *i* gene was depicted as having a context-sensitive influence on the *o* gene. In the absence of lactose, this influence was negative. In the presence of lactose, the repressor gene had no influence on the operator gene. The *o* gene, in turn, was thought to have had a pleiotropic effect on both genes *z* and *y*.

In the fourth section, "The Operator and the Operon," of their 1961 paper, Jacob and Monod addressed the question of how the repressor was able to act on the operator in the lactose system. While they had previously established that the repressor kept β-gal from being synthesized until lactose interrupted repression, they next wanted to find out exactly how the repressor repressed the expression of the *o* gene in the wild-type. That is, they wanted to know the particular mechanical properties of the system that made it so that the expression of the repressor could do its job in the absence of lactose. To solve this problem, they once again turned to mating experiments with different strains of *E. coli*. Like the other genes in this system, the *o* gene has two variants: The o^+ variant of the *o* gene locus is the inducible type and the o^c variant is the constitutive type. Jacob and Monod conjugated a wild-type male [o^+ i^+] into a mutant female [o^c i^-].[8] Because they already knew that the i^+ allele was dominant, they figured that the outcome of the expression of the repressor gene in contexts of the o^c allele would help them understand

[8] Jacob and Monod also looked at various other crossings that I do not have the space to discuss in any great detail. For the purposes of this article, the crossings I describe are sufficient.

the conditions under which the repressor gene could properly repress the operator gene. From these conjugation experiments, the o^c mutants continued to produce β-gal in the absence of any inducer. They concluded that, "the $o^+ \rightarrow o^c$ mutations correspond to a modification of the specific, repressor-accepting, structure of the operator. This identifies the operator locus, i.e. the genetic segment responsible for the structure of the operator, but not the operator itself" (Jacob & Monod 1961, 342). The last sentence of the quotation is crucial, as it indicates that Jacob and Monod were unable to infer any more details about the conditions or mechanistic properties under which the repressor could function from their particular bacterial mating experiments. They remained uncertain about the nature of the operator molecule and, in particular, whether the repressor acted upon the o^+ variant by binding directly to a segment on the DNA molecule or whether the repressor acted upon the *o* gene's product, or what they called the "cytoplasmic messenger" (Jacob & Monod 1961, 344; see figure 2).[9]

[9] Jacob and Monod go on to address this very question in the following section of their paper, "The Kinetics of Expression of Structural Genes, and the Nature of the Structural Messenger." In brief, in that section, Jacob and Monod describe their work that provided evidential support for the "genetic operator model" and claimed that the system indeed functions at the genetic level "by governing the synthesis of the structural message," rather than "at the cytoplasmic level, by controlling the protein-synthesizing activity of the messenger" (Jacob & Monod 1961, 346-352).

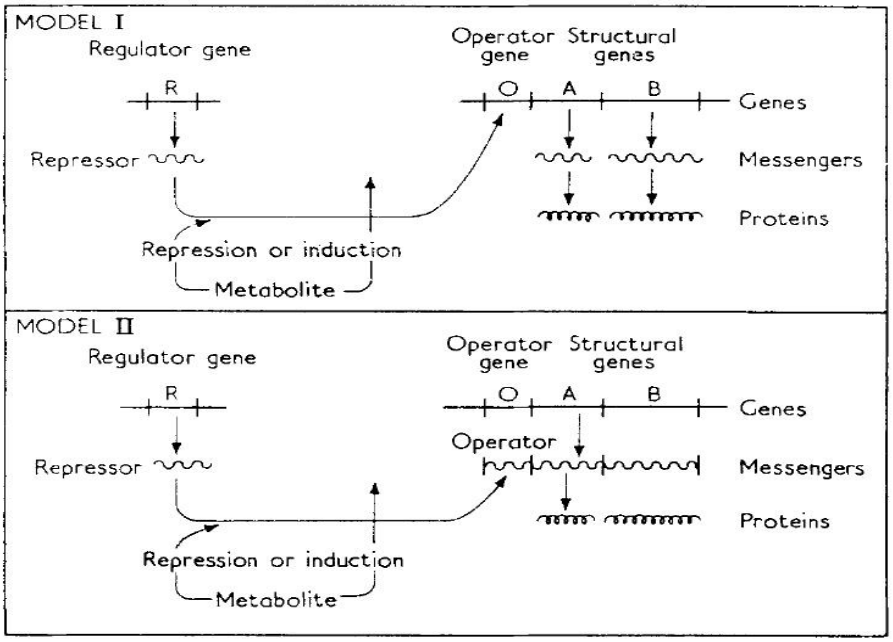

FIG. 6. Models of the regulation of protein synthesis.

Figure 2: "Genetic Regulatory Mechanisms in the Synthesis of Proteins" (Jacob & Monod 1961, 344). Reprinted from *Journal of Molecular Biology,* Vol. 3, Jacob, F., & Monod, J, "Genetic Regulatory Mechanisms in the Synthesis of Proteins," 318-356, (1961), with permission from Elsevier.
The top half of the figure, Model I, depicts the operator as a site on the DNA segment linked with the structural genes of the operon. Model II, in the bottom half of the figure, depicts the operator linked to a "cytoplasm messenger".

As they had foreseen, Jacob and Monod's 1961 paper had a large impact on subsequent research in molecular genetics. Some of their novel contribution to molecular biology was to show that genes could be switched on and off, and that enzymes could be synthesized *de novo*. However, it remained unclear exactly how inducers interact with the repressors, and how repressors acted upon operators. Their work encouraged researchers to try to isolate the repressor molecules, both in *E. coli* and the λ-phage genes. In 1966, Walter Gilbert and Benno Müller Hill achieved the biochemical isolation of the lac repressor (Gilbert & Müller Hill 1966). And, in 1967, Mark Ptashne successfully isolated the λ-phage repressor and its binding site on DNA (Ptashne 1967). Jacob and Monod's

model was relatively simple, but later biochemical isolation of the repressors in the experimental systems they studied and discoveries of the specific sites on DNA to which repressors bind allowed for the further development of models of genetic regulation (Morange 1997).

The previous presentation of research on the *lac* operon system suggests that biology does not always begin with a bottom-up strategy, or reductionist approach, by studying the properties of the smallest individual components of a system and then progress towards a more complete or holistic description of the phenomena under study. In fact, the case of the origins, the subsequent developments, and the justification of the *lac* operon model indicates that while Jacob and Monod were curious about the role of individual molecular components within bacterial cells, they were more concerned with explaining the components' role in the functions of the system. To do so, they used a hybrid of techniques in biochemistry and classical genetics and designed an experimental system, the merozygotic bacterium, which allowed them to impose a general organizational structure on the phenomena and to trace the processes of interest. The experimental techniques employed by Jacob and Monod provided the best handle with which to manipulate and understand the regulation of enzyme induction in *E. coli* at the time. It was only later that their model was further supported by knowledge of the biochemistry of its components.

Allen's application of the mechanistic-holistic dichotomy

In the previous section, I have presented the development of the *lac* operon model as an example of a particular episode in the history of biology that does not fit neatly within the mechanistic-holistic dichotomy. I now turn to Allen's account of developments

in physiology, biochemistry and molecular biology in early and mid-twentieth century biology to illustrate one way that the dichotomy between mechanistic and holistic approaches has been used to interpret historical developments in modern biology. Throughout his account, Allen suggests that the production of scientific knowledge often proceeds from a mechanist and reductionist account of phenomena to a more complete, holistic description of the phenomena. In fact, many other historians and philosophers have interpreted some of the transitions in twentieth century biology along these lines. Some have described the empirical success of early research in molecular biology as a result of adopting the mechanistic, or reductionist, approach (Griffiths & Stotz 2013, Morange 2006, Powell & Dupré 2009). However, in Allen's treatment of the history of the life sciences, he infers the following generalizations: (1) that the approach described as "holistic materialism" provides a more complete and accurate description of the natural world (Allen 1978, 105-106), and (2) that "the mechanistic approach has often been the only practical way to begin the study of a complex process" (Allen 1978, 106).

Allen explicitly articulates the difference between mechanistic materialism and holistic materialism in his chapter on physiology (Allen 1978, 103). There, he uses the dichotomy to make sense of the shift from Jacques Loeb's mechanistic physiology to Lawrence J. Henderson and Walter B. Cannon's work on the self-regulating buffer systems involved in certain physiological processes.[10] Allen explains that holistic materialists, like Henderson and Cannon, believed that looking at parts of a system and

[10] In brief, Henderson worked on the carbonic acid buffer system of blood, which regulates the blood's pH level (See further: Allen 1978, 95-100; Henderson 1928). Cannon, known for coining the term "homeostasis," worked on the sympathetic nervous system (See further: Allen 1978, 100-103; Cannon 1929).

their properties in isolation, as the mechanist materialist would, could not provide adequate explanations of the systems under study because "new characteristics emerge from the interaction of parts" and these new characteristics "are qualitatively different" (Allen 1978, 106). Allen understands the mechanistic tactic as a sort of decompositional approach to understanding biological processes and claims that, by the 1920s, physiologists were slowly moving away from that approach and towards holistic materialism (Allen 1978, 103). The mechanistic view upheld a materialist metaphysics, and so rejected idealist notions such as vitalism. However, it prioritized the events and processes between entities rather than the material entities themselves. For physiologists, "atoms and molecules as material entities [became] less important than the *interactions*… in which they [were] involved" (Allen 1978, 104). This shift from entities to processes or interactions is what Allen takes to represent the move from a mechanistic materialism to a holistic materialism.

Allen is careful not to equate either position with ontological or metaphysical claims. He claims: "The chief, or important, difference was *not* that, by and large, the two groups worked at different levels of organization within living systems" (Allen 1978, 104; my emphasis). Many of the scientists who Allen sees as occupying opposite sides of the divide were concerned with the same ontological level of organization. It was their outlook on which methodology, or strategy, would be more successful that differed. According to Allen, Loeb and other mechanistic materialists assumed that the whole was just the sum of its parts. The holistic materialists, in contrast, believed that the whole, consisting in the interactions between the system's components, exhibited emergent properties that mechanistic materialism would fail to countenance. For example,

mechanists would perhaps overlook, or fail to explain, phenomena like the control of antagonistic muscle contractions or the total buffering potential of the blood. These phenomena could not be explained by the additive effects of a system's components; e.g. single pathways of stimulation or the chemical components of blood (Allen 1978). Thus, if Allen's notion of mechanistic materialism has anything to do with how philosophers now discuss different varieties of reductionism, then his notion is clearly concerned with a methodological kind of reductionism, and not an ontological or metaphysical one.[11]

A methodological reductionist assumes that the most fruitful way to study living systems is to decompose these systems into lower-level components, then, ideally, to localize how structural entities perform certain functional activities (Bechtel & Richardson 1993; Griffiths & Stotz 2013). And, according to Allen, during the development of physiology, towards the advent of molecular biology, scientists were rejecting that thesis in favor of a methodology that embraced the emergent properties of complex phenomena. Allen summarizes the point of difference:

> To a mechanist, interaction does not impart new characteristics to any one component when it is interacting as part of a whole or when it is acting in isolation. A complete description of the characteristics of any part is possible by studying that part separately from others. Mechanists believe that studying interactions is also important. But the interactions are just quantitatively more complex situations. To holistic materialists, on the other hand, new characteristics emerge from the interaction of parts. These new characteristics are not merely

[11] See Griffiths and Stotz (2013) for an overview of the different types of reductionisms. I also offer an overview of reductionism in the second chapter of this dissertation.

> quantitatively more complex. They are qualitatively different. New characteristics of parts emerge during interactions. These new characteristics result from the parts affecting and altering each other. To holistic materialists, these new characteristics are interpretable in terms of rational laws of science, but they are not quantitatively extrapolatable from analytical studies alone (Allen 1978, 106).

Likewise, on the development of biochemistry in the early twentieth century, Allen emphasizes the move towards a holistic materialist approach to the study of biomolecules. When describing Otto Warburg's work on the problem of cellular respiration, he claims that Warburg's mature view was consistent with "the principles of mechanistic materialism" because he looked at enzymes *in vitro*, outside of the context of a living system, as "free-floating molecules carrying out a reaction whenever they encountered appropriate substrate molecules" (Allen 1978, 181). Later developments in biochemistry, according to Allen, progressed towards a holistic approach in order to produce a detailed picture of the structure and function of cells. In this case, he claims that the mechanistic work by Warburg and others was necessary to transition to the holistic phase of biochemistry: "The isolation studies came first out of necessity but, by the 1960s, biochemists were beginning to move away from the view that the characteristics of a reaction *in vivo* were necessarily identical to those *in vitro*" (Allen 1978, 182). He reiterates this point about the normal, and somewhat expected, transition from the mechanistic to the holistic, in his conclusion:

> The history of biochemistry, as with the other areas of biology discussed in this textbook, illustrates clearly the importance of the mechanistic materialist stage in the development of the modern life sciences: that it is often necessary to isolate

> individual components (in this case single enzymatic reactions) from a complex interacting system at the outset. The cast of characters, so to speak, must be identified. But the history of biochemistry also illustrates the limitations inherent in the mechanistic approach if applied as an overall philosophy of nature (Allen 1978, 183-184).

Finally, on the origins of molecular biology, Allen argues that there was a brief return to mechanistic thinking, when the molecular structure of DNA was discovered and explained in terms of its physical structure and biochemical properties. A central role was attributed to the gene, and to a particular molecule in the cell, that of DNA, in the 1940s and 1950s. However, soon after, according to Allen, molecular biologists recognized that gene expression involved complex control systems involving interactions that resulted in emergent properties of the whole. As Allen writes, "*in vivo* gene transcription and translation [was discovered to be] inordinately more involved than the systems studies *in vitro*," (Allen 1978, 223). This mediated against a simplistic view that the properties of individual molecules or the specific chemical reactions between enzymes and substrates could suffice to explain living systems. In the case of molecular biology, then, Allen maintains the narrative of a practice originating from a mechanistic materialist view evolving towards a more holistic view. While this interpretation could be consistent with the research on enzymes and their substrates preceding the research on the operon model, some of the research subsequent to the operon seems to flow in the opposite direction; that is, as scientists accepted the sketch of gene regulation from the operon model, some of the later research turned to achieving a better understanding of the molecular components involved in bacterial physiology in isolation. Given this, we might explore

the idea that the dichotomy itself is not the best lens with which to represent and interpret particular episodes, or the development of particular research programs, in the history of biology. Perhaps another perspective would be better suited for such ends, and perhaps the dichotomy itself (along with Allen's generalizations) might be better suited for different ends.

Tracking processes: Moving beyond dichotomies for particular episodes of scientific practice

As an alternative to the dichotomy, I submit that a better framework with which to represent and interpret some of the particular episodes in twentieth century molecular biology is to view the scientific practices involved as efforts to track certain biological processes. To defend this view as an alternative to the dichotomy, I make the plausible assumption that there is a fact of the matter concerning what molecular biologists track. That is, there is a fact of the matter about what a molecular biologist foregrounds and backgrounds in her models. To develop this framework, I make use of James Griesemer's notion of "tracking processes" in science, and, more specifically, his description of scientific representations (including models) as commitments to following a process in a particular way, foregrounding and backgrounding different elements of the process, or phenomena, under study (Griesemer 2006; 2007).

Tracking a process representative of a phenomenon is a common practice in science. It is used both as an exploratory tool (to observe where a process will lead) and as an intervention (to determine causal relationships) (Griesemer 2007). Moreover, the practice of following a process necessitates creating a representation of the phenomenon in such a way that some aspects of the process are foregrounded while others are

backgrounded (Griesemer 2007). Griesemer explains that the representation "…focuses attention on foregrounded elements as the significant and explanatory aspects of the process-as-followed. The result is constraint and guidance on how processes may be followed on other occasions, as well as what implications are (literally) drawn from reported work" (Griesemer 2007, 376).

Griesemer borrows from Hans Reichenbach's (1991) "mark principle" and Salmon's (1984) "mark transmission" to explain how scientists habitually follow processes to gain a causal understanding of phenomena (Griesemer 2007, 377). The marks that are tracked are thought to be factors that have causal relevance in the process of interest. The notion of mark transmissions can be operationalized such that the manipulation plays a theoretical role in determining the process as causal:

> In a manipulative marking intervention, the experimenter focuses on a 'target' of attention prior to the marking interaction and then introduces a mark that physically changes a property of the process in such a way that continuous mental attention is not required to track the process. This is a procedural benefit of the irreversibility of marks that Reichenbach required of causal processes. The mark can be tracked in intermittent 'checkups' via subsequent observation of or interventions in the process to see if the process still carries the mark. This operational notion of mark transmission thus also plays a theoretical role in identifying the process as causal. Theory and methodology are as intimately related as two sides of a coin" (Griesemer 2007, 380).

Gene knockout or knockdown methods provide a now commonplace example of this kind of process-following interventionist practice within the life sciences. Likewise, Jacob and

Monod's bacterial conjugation methods and experiments can be interpreted as this kind of intervention on the process of enzyme induction in *E. coli*. Additionally, in the case of the repression model of the *lac* operon, foregrounding the entities of a system of regulation means attending to the so-called "mark transmissions" in the genetic control and regulation of *E. coli's* metabolic enzymes. This is the research strategy or style that Jacob and Monod established in molecular genetics with their operon model. The later biochemical studies of the entities within this system of genetic regulation followed the same biological process as did Jacob and Monod; however, those studies re-oriented their research strategy in such a way that the structural and chemical properties of the entities in the system were foregrounded in a different manner. Nonetheless, the regulatory structure (i.e. the logic of genetic regulation) that Jacob and Monod imposed on the entities within the bacterial cell was crucial in making sense of one part of the process of enzyme synthesis, and furthermore, constrained subsequent research strategies in regulatory genetics and developmental biology.[12]

There are several upshots to adopting this conceptual framework for interpreting particular episodes in the history of science. By shifting our focus to the practices of tracking processes, we can better understand how and why much of molecular genetics, including Jacob and Monod's construction of the operon model, made use of the investigative methods of classical genetics (e.g. crossbreeding with bacterial conjugation). We can also see how, in this particular example from twentieth century biology, the scientific knowledge of transcriptional genetic regulation did not proceed

[12] See Morange (1997) for an account of the influence of the *lac* operon model on later research projects in regulatory genetics and embryology.

from knowledge about individual entities to a more holistic understanding of the functioning of the entire system, *contra* Allen's generalizations. The more gradualist picture of scientific practice that emerges from this conceptual framework blurs any clear distinctions between theory construction and experimentation, and offers a narrative of scientific change that pays attention to the origin and entrenchment of certain scientific practices and research strategies, and the way in which these constrain subsequent scientific developments. This framework also better encapsulates the usefulness of considering the details of scientific practices to understand episodes in the history of the life sciences.

Implications for historiography: mode and tempo in the history of science

There is a further point about understanding and interpreting the history of science with implications for both the history and the philosophy of science that I would like to make in this essay. As I mentioned in the introduction, Allen represents scientific change as a series of discontinuities and revolts. In a commentary on Allen's work, Jane Maienschein, Ronald Rainger, and Keith R. Benson claim instead that scientific change ought to be interpreted in a gradualist manner, as "complex changes that cannot be stated in oversimplified terms as dichotomies; [they] seek to understand continuities and gradual change" (Maienschein *et al*. 1981, 86).

In his response to the commentary by Maienschein, Rainger, and Benson, Allen claims that the distinction between the gradualist approach and the revolutionary (dichotomist, or saltationist) approach to the history of science represents "two sides of the same coin" (Allen 1981, 173):

> There is a constant interplay between quantitative (evolutionary) and qualitative (revolutionary) changes in the history of science. The rate of change may vary considerably from one period to another… to ignore differences in tempo, and the rather important differences in consequences between quantitative and qualitative change, is to miss an important aspect of the dynamics of history (Allen 1981, 174).

Instead, Maienschein, Rainger, and Benson suggest that the difference between their approaches might simply be a difference in emphasis or orientation, which offers a better way to understand the difference between an evolutionary/gradualist approach and a revolutionary/dichotomist approach to the history of science.

Rather than representing an actual difference in the rate of change or tempo in the historical development of twentieth century biology, I propose that both Allen's approach and my proposed framework represent shifts in the *grain of resolution* on historical timescales. Historians of science can zoom in or out of the timescale to focus on slightly different patterns emerging from a given time period. Jon F. Wilkins and Peter Godfrey-Smith first proposed that shifting the grain of resolution is a useful heuristic for understanding different facets of evolutionary change (Wilkins & Godfrey-Smith 2009; 2011). They argue that the debate in evolutionary biology over the significant factors responsible for evolutionary outcomes (e.g. ecological demands, developmental biases, or historical contingencies) can be deflated, to some extent, by paying sufficient attention to the grain of evolutionary analysis. For example, at the broadest, most "zoomed out" level of analysis, evolutionary patterns appear such that populations seem restricted by a very limited range of possible phenotypes and ways of life. Such macro-evolutionary patterns

bring attention to restrictions imposed by developmental biases on natural selection. When we zoom in on shorter timescales in the history of life, the restrictive features at the broader grain of analysis become less important, and they are replaced by factors that can explain different variants within populations and between several generations. At the mid-level analysis – that is, zooming somewhere in between microevolutionary and macroevolutionary patterns – the variations within any groups on the tree (genera or species) will become more apparent, but so will some of the traits or characters that become entrenched in a lineage. Wilkins and Godfrey-Smith suggest that this mid-way level is useful in understanding certain phenomena, such as cases of heterozygote superiority (Wilkins & Godfrey-Smith 2011, 198).

In any case, the main idea here is that different factors become appropriate, or relatively more significant, as *explanans* of evolutionary change according to both what is being explained and the chosen grain of analysis on evolutionary timescales. Likewise, for the history of science, it might be the case that looking at the history of science at a finer grain of resolution – for example, the development of the research programs of particular individuals or research groups at a particular time and place – will culminate in a gradualist picture of continuity and path-dependency; whereas looking at the history of biology at a more "zoomed-out" level – for example, the trends in the currents of thought within a certain cultural context, country, or continent during an extended period – will bring about a more saltationist, revolutionary view, which can be represented as shifts from a mechanistic approach to a holistic one. This view of the different ways to delineate periods and episodes in the history of science suggests a pluralist view of the historiography of science.

However, it is still worth arguing about which approach (or level of analysis) is better or more appropriate for any particular historical question, and about the potential benefits and challenges that different levels might bring. On the one hand, taking a too myopic view of the history of biology during the twentieth century might mask the extent to which the science changed during that period. That lens, or level of analysis, might not always be optimal to explain the sorts of Copernican revolutions – that is, the foundational changes which transform the way scientists conceive of their work and the phenomena they study – that have occurred in the history of science. Framing the history of science in terms of debates and oppositions can sometimes help bring out the fundamental issues and assumptions that are at odds within competing conceptual frameworks in science. On the other hand, looking at the history of molecular biology as a dichotomy between clear-cut methodologies or research strategies can misrepresent the range of scientific work being done at the time and conceal the continuation between research programs. For example, imposing apparent dichotomies between mechanistic materialism and holistic materialism, or reductionism and anti-reductionism, on the history of biology during the twentieth century might skew the facts about what scientists were tracking in their experimental and conceptual practices. Perhaps then, in some instances, taking a more gradualist, or "zoomed-in," view can help communicate the myriad of influences, both constitutive and contextual, which contribute to the uptake of scientific ideas and research agendas and constrain subsequent developments within a scientific domain.

Summary

Conceptual frameworks should be useful for shedding light on the process of science and its goals. The framework adopted by Allen leads him to make generalizations about the development of science that do not apply neatly to the details of particular episodes in the history of molecular biology. It becomes difficult to understand the instances of convergence of different concepts and experimental techniques, like that of Jacob and Monod's collaboration, in terms of a clear development from a mechanistic to a holistic approach to science. I have suggested that a framework based on the idea of tracking processes can better accommodate some of the details of scientific practices in particular historical episodes and can offer a better account of the scientific process at a finer grain of resolution.

THE MICRO-RNA WORLD: CONTINUITIES AND DISCONTINUITIES IN POST-GENOMIC RESEARCH

Introduction

In an opinion piece published in *Nature*, geneticist John Mattick suggested that research on non-protein coding RNA molecules, specifically their role in genetic regulation, is at the heart of post-genomic molecular biology in the twenty-first century (Mattick 2004; see also Mattick 2003). Historians and philosophers of biology are attending to what post-genomic biology means for historical and philosophical accounts of theory change and scientific methodology, particularly with respect to discussions about reductionism and anti-reductionism (Burian 2007; Dupré 2010, 2012; Keller 2005; O'Malley, Elliott & Burian 2010; Morange 2006; Powell & Dupré 2009; Richardson & Stevens 2015; Woese 2004). Many have represented the revised concepts of genes and genomes within post-genomic molecular biology as an acknowledgement that the so-called reductionist strategy, in which the whole system is explained by an analysis of its parts and their interactions, has now reached its limits (see, for example, Woese 2004). The general argument is that the study of these systems will have to be holistic, or anti-reductionist, for several reasons. Reductionism won't work, the argument goes, because genetic or genomic systems exhibit emergent behaviors, or because the functions of these systems' components are context-dependent (i.e. the context in which molecules interact makes a difference to the outcome of the system), or because these functions cannot be inferred solely from the components' structure. Though there are clear conceptual and experimental shifts between pre-genomics and post-genomics, there are also continuities in these research traditions. In this chapter, I begin to tease out these changes and

continuities. I argue first that many practices in post-genomic molecular biology are, in fact, continuous with earlier work in molecular biology, yet post-genomics incorporates many new tools and techniques within these practices. Second, I argue that some claims about the shortcomings of reductionism rest on a mistaken view of emergence and what it entails about anti-reductionism, and a misunderstanding about what it means to account for context dependencies.

I begin by demonstrating the presence of a standard narrative of the trajectory of molecular biology in the history and philosophy of science. This narrative traces the historical development of molecular biology from the twentieth century pre-genomics era to the current twenty-first century post-genomic period as a shift from reductionism to some form of anti-reductionism, or holism. To do this, I sketch some of the ways in which biologists, philosophers, and historians have described the conceptual change in post-genomic molecular biology and the ways in which some have called for novel, non-reductionist strategies to understand and explain gene expression.

I next present an overview of micro-RNA (henceforth, miRNA) research to provide a more concrete example of the kind of scientific research and practices that some biologists, philosophers, and historians have in mind when they describe the changing concepts in gene expression research in the post-genomic era. I first give a general description of the biogenesis and some molecular functions of miRNAs. I also briefly outline the history of miRNA discovery, bringing attention to some of the important contributions made by Victor Ambros and Gary Ruvkun, and their respective research teams. I finish the overview of miRNA research by describing some of the methods and techniques within miRNA research, with a particular focus on some of the

new approaches used to detect miRNA, their targets, and their biological functions.[13] This survey reveals several characteristics of the practices within current gene expression research. These characteristics indicate that the research programs in post-genomic molecular biology, especially in miRNA research, have adopted novel tools and techniques, yet the iterative and integrative processes between the approaches and strategies are continuous with previous episodes in twentieth century molecular biology.

In the second half of the paper I clarify different meanings of reductionism in the philosophical literature, with particular attention to the concept of emergence. Philosopher of science John Dupré argues that the context-dependency inherent in the conceptualization of the genome leads to a strong sense of emergence that precludes any sort of reductionism (Dupré 2010). Dupré's arguments concerning reductionism engage with exactly the sorts of research projects in post-genomic molecular biology that I discuss throughout this chapter. Against Dupré, I argue that the context-dependency of miRNAs does not necessarily preclude a reductionist methodology. In fact, context-dependencies are easily accounted for by the causal interactions between different macromolecules. The explanatory frameworks used to represent these kinds of interactions have a long and successful history in molecular biology. I illustrate my argument with appeal to current biomedical research focused on miRNA target recognition and regulation in the context of lung cancer.

[13] Recall I am using the terms *approaches* and *strategies* in a specified way throughout the dissertation. *Approaches* signify the set of methods in particular fields, or sub-fields, of molecular biology, including the new technologies in genomics, computational biology or bioinformatics, etc. *Strategies* represent conceptual frameworks used within these approaches, such as the top-down and the bottom-up strategies described in the previous chapter.

In summary, this chapter aims to provide a better understanding of the concepts and practices within a branch of current post-genomic molecular biology and to help clarify and deflate some of the tensions in philosophical debates between reductionism and anti-reductionism with respect to molecular biology.

From pre-genomics to post-genomics: A gradual continuation in the history of molecular biology

A standard narrative of the conceptual shift in post-genomic molecular biology

After the successful sequencing of the human genome, some biologists hoped, and often assumed, that knowledge about the underlying genetic and genomic make-up of organisms would lead to a clearer and more complete understanding of life, its development, and its evolution. But that type of optimism surrounding the promise of genomics was never quite realized, at least not in the way that many had hoped. Nonetheless, the assumption was reasonable because, during the previous decades, molecular biology had made considerable progress by looking deeper and deeper into the cell and identifying, classifying, and organizing the cell's components to explain gene expression, physiological and developmental processes, and to support phylogenetic reconstructions.

While molecular biologists continued to conduct research, philosophers became interested in the changing concept of the gene within this research tradition. They began to wonder about the relation between different gene concepts, such as the Mendelian concept of the gene, the molecular concept of the gene, and the post-genomic concept of the gene. The classical Mendelian notion of the gene signified a unit of transmission across generations. The molecular concept of the gene was associated with a physical bit

of DNA that determined a protein with a particular structure and function. The emerging post-genomic notion of genes includes protein-coding genes, as well as non-protein coding RNAs that have a variety of enzymatic and regulatory functions previously relegated as "junk" DNA. Many began questioning whether the simplicity and elegance of the Central Dogma – i.e. the idea that information cannot be transferred from proteins back to DNA – could adequately capture the complex processes of gene expression within an organism. The new discoveries of introns, exons, non-coding RNAs, such as miRNAs, as well as the phenomena of alternative splicing, alternative polyadenylation, and epigenetic regulation, all signaled a move towards discarding, or revising, the Central Dogma. Some theoretical updating was called for to better understand the genomic contributions to biological complexity, the development of organisms and their ability to reproduce, and the scope of phenotypic variation (Mattick 2003). Shortly after the beginning of the twenty-first century, historians, philosophers, and scientists began to try to make sense of the meaning of these changes. I now turn to some of these influential scholars in order to give textual evidence of what I take to be the standard narrative of the shift from twentieth century molecular genetics to twenty-first century genomics. This narrative is centered around the notion that pre-genomic research was driven by a reductionist approach to science and reaped its rewards because of that approach, whereas the dawn of the post-genomic era is characterized by a move towards a holistic, or anti-reductionist, science with a focus on the emergent properties of biological systems.

Woese

In 2004, Carl Woese, a microbiologist best known for distinguishing the group Archaea from Bacteria and Eukarya, wrote "A New Biology for a New Century." In this commentary, he argues that twenty-first century biology was in need of a new "guiding vision," and "a new, deeper, more invigorating representation of reality" (Woese 2004, 173). Woese addresses several research programs in the life sciences in that article, but he is particularly concerned about the effects that twentieth century molecular biology has had on the current state and the future directions of biological investigations. The molecular era, he claims, was centered on the "encapsulatable" problems of the gene and the cell, which meant that those biological entities could be studied in isolation. According to Woese, molecular biology was influenced by the metaphysics on which much of nineteenth century classical physics was founded. The central idea was that scientists gained true knowledge and a firm understanding of phenomena by studying its smallest underlying physical entities. Investigations of the cell and the gene, explains Woese, "were amenable to a reductionist approach" and "would benefit from the fresh, no-nonsense outlook and experimental power of molecular biology" (Woese 2004, 174). However, he continues, reductionism "is deeply woven into the fabric of modern biology, and biology today has hit the wall of biocomplexity, reductionism's nemesis" (Woese 2014, 174). He offers the following evaluation of the historical significance of the molecular era in biology:

> I think the twentieth century molecular era will come to be seen as a necessary and unavoidable transition stage in the overall course of biology: necessary because only by adopting a heavily reductionist orientation and the technology of

> classical physics could certain biological problems be brought to fruition and transitional because a biology viewed through the eyes of fundamentalist reductionism is an incomplete biology. Knowing the parts of isolated entities is not enough. A musical metaphor expresses it best: molecular biology could read notes in the score, but it couldn't hear the music… The molecular cup is now empty. The time has come to replace the purely reductionist "eyes-down" molecular perspective with a new and genuinely holistic, "eyes-up," view of the living world, one whose primary focus is on evolution, emergence, and biology's innate complexity (Woese, 2004, 175).

Woese is advocating for a shift in the methodological strategies to studying and representing biological phenomena because the reductionist strategy cannot address some of the most fundamental problems in biology, including the problem of the nature of complex organization. To be fair, Woese is not advocating for an end to molecular research in biology, but rather insists that the new representations in biology will need to be freed from the "procrustean reductionist perspective," borrowed from the metaphysics of nineteenth century science (Woese 2004, 175). Nonetheless, he is clearly presenting the shift from twentieth century molecular era to the twenty-first century as one that brings holistic problems to the fore, and describes his own scientific contributions as providing "the links between biology's reductionist past and its holistic future" (Woese 2004, 176).

Keller

In 2005, five years after the publication of *The Century of the Gene*, historian and philosopher of science Evelyn Fox Keller published a reflection paper entitled, "The

Century Beyond the Gene" (Keller 2000, 2005). In both works, Keller alludes to a move away from the reductionist paradigm of early molecular biology in current biology. In the latter work, she addresses, more specifically, what the United States Department of Energy (DOE) has called, "Bringing Genomes to Life," to understand the move "beyond 'reductionism' to 'systems biology'" (Keller 2005, 3). The project proposed by the DOE represents a broader turning point in the history of gene concepts, according to Keller. One in which the old paradigm of the genome as a collection of single genes and the reductionist approach to the study of gene structure and function were being replaced by a systems-level focus on the complexity of biological systems and their emergent properties. About the transition, she writes:

> The more we learn about how the parts work not only in interaction with each other, but also with the larger entities in which they are embedded, about the extraordinarily complex and versatile systems of gene regulation, about the signals mediating all the different levels of organization, and about the variety of epigenetic mechanisms of inheritance at play and the evolutionary feedback between the different mechanisms, the more compelling the need for an entire new lexicon, one that has the capacity for representing the dynamic interactivity of living systems, and for describing the kinds of inherently relational entities that can emerge from those dynamics (Keller 2005, 9).

Here, Keller emphasizes the emergent properties of the regulation of gene expression as *the* pressing problem of twenty-first century molecular biology. Keller has maintained this narrative in later work, and though she does not herself explicitly advocate for a move towards anti-reductionist strategies in her writing, she sometimes suggests how

alternative conceptions of the gene and the genome can lead to different strategies in certain research programs.

In a recent paper on the "postgenomic genome," Keller writes:

> I am proposing that today's genome, the postgenomic genome, looks more like an exquisitely sensitive reaction (or response) mechanism – a device for regulating the production of specific proteins in response to the constantly changing signals it receives from its environment – than it does the pregenomic picture of the genome as a collection of genes initiating causal chains leading to the formation of traits" (Keller 2015, 25).

She goes on to suggest her concept of the "postgenomic genome" can have implications on how certain research programs are pursued. For instance, she argues that biomedical research of disease that is focused on locating single nucleotide polymorphisms (SNPs) is limited because it is too reliant on the classical concept of mutation and gene variants (alleles). That research carries the implicit assumption that a gene is an identifiable and discrete unit within the genome, and masks other kinds of causes of diseases that are better characterized as "genomic disorders" (Keller 2005). These disorders have to do with structural features of the genome and its organization. Research on genomic disorders, according to Keller, should focus less on DNA sequences and more on genomic level phenomena, such as genomic inversion, genome duplication, and genome deletion. In this way, Keller thinks the conceptual shift in postgenomic biology, from the reductionist focus on DNA sequences and genes towards the properties of the genome, has implications for how research is pursued.

Morange

Like Keller, Michel Morange is attentive to the "reductionism to anti-reductionism" narrative in the history of molecular biology and uses these terms to describe some of the scientific developments in post-genomic biology. And also like Keller, Morange leaves any normative claims about whether post-genomic molecular biology *ought* to adopt research strategies that can be described as holistic or anti-reductionism as open-ended. In a short paper, published in 2006, Morange tackles the different meanings of post-genomics and of emergence and considers whether this new era in biology is a reductionist or holist project. He begins the paper by claiming:

> Reductionism, under the label of molecular biology, seemed to have definitively won at the end of the twentieth century: the characteristics of organisms were considered to be explained by the structural properties and enzymatic capabilities of their macromolecules, which could be assessed using genome sequences. This victory was only apparent and transient, and holistic models reemerged at the eve of the twenty-first century (Morange 2006, 355).

In this passage, Morange is describing the current view in biology that knowledge of gene sequences alone does not explain much about organisms. Morange goes on to argue that there are several ambiguous terms and metaphors in post-genomic biology that lie "at the border between reductionism and holism" (Morange 2006, 357). The ambiguous concepts generate confusion about what the terms, reductionism and holism, imply for scientific practice (Morange 2006, 357). An example of this is the concept of pleiotropy. Morange argues that this term is ambiguous because it is applicable in different ways. For example, a gene can be characterized as pleiotropic when its protein product has multiple

structural domains,[14] each of which have a specific function. Or, it could be pleiotropic when its protein product has structural components that associate with several targeted molecules in different cellular contexts. In the latter case, the pleiotropic effects of a gene are dependent on different contexts, and not simply on the structure of its protein.

Morange leaves open how one ought to interpret the meaning of this kind of context-dependency and how biologists go about studying, representing, and explaining the phenomena at hand. Biologists often continue to operate by explaining the properties of a pleiotropic protein, or gene, by characterizing its interactions with the molecular components involved in its different functional roles (Morange 2006, 358). This is, nonetheless, a departure from investigations reliant on the pre-genomic concepts of genes and genomes, according to Morange. Post-genomic biology addresses problems of gene expression and gene function as problems of complex systems of networks, rather than straightforward correspondence between a gene's structure and function.

Burian

The preceding authors have addressed the shift from pre-genomics to post-genomics in terms of the changes in the conceptualizations of the gene and the genome. Additionally, all have appealed to instances of research on the regulation of gene expression to illustrate the conceptual change happening in the new era of post-genomic molecular biology. Richard Burian has further probed into the characteristics of post-genomic molecular biology by paying close attention to one particular research program focused on specific non-coding regulatory RNAs and their roles in the regulation of gene

[14] For example, important protein domains include the homeodomains (to bind to DNA, such as transcription factors), RNA recognition motifs (to bind to RNA), and nuclear hormone receptors.

expression (Burian 2007). Burian addresses both the conceptualization of the genome's causal contribution to the development and maintenance of organisms, and the approaches and strategies available to biologists in the post-genomic era to study gene expression.

In brief, Burian argues that biologists working in the post-genomic era no longer attempt to infer the biological functions of regulatory molecules, such as miRNAs, from nucleotide sequence or structure alone. What is new in post-genomic molecular biology, according to Burian, is a conceptualization of the genome and of gene expression that takes into account the context in which macromolecules carry out certain biochemical functions. Burian's notion of context includes much more than the cellular context in which gene expression occurs. It also includes what he calls "epigenetic historicity" (Burian 2007, 286). By this, he means the evolutionary contingency of the entire organism, which has resulted from processes of co-adaptation at several levels of organization within the organism. He draws out this point with respect to his example of miRNA research, explaining:

> Part-by-part analysis of relevant biological structures plus the arrangement of their parts does not suffice to reveal their roles in integrated organisms or their functions… The context – e.g. what happens or happened in distant parts of an organism or its environment – may fundamentally alter the structure and/or function of a particular biomechanical structure, machine, organ, or process. The resulting contingency or historicity poses serious obstacles to any principled theory of function. The problem is amplified and clarified by co-evolutionary changes in structure-function correlations and is exceptionally clearly illustrated

> in the instance of miRNA… Accordingly, it may not be possible to derive relevantly significant properties of developing *systems* from the properties and relations of the fundamental units out of which they are built at a particular time (Burian 2007, 286-288).

Burian argues that this complexity in post-genomic research on miRNAs requires exploratory experimentation, which he defines as "situations in which experimental outcomes cannot be accurately predicted by available theories together with general background knowledge plus boundary conditions" (Burian 2007, 287). I take it that what Burian has in mind here is that much of post-genomic research proceeds by exploring genomes for recognizable patterns of interest. For Burian, exploratory experimentation represents a new research strategy in post-genomic biology. This strategy is required because there is no general theory about the rules of molecular interactions that can be used to assess the functions of miRNAs in different cells and under different circumstances.

Burian's account of current research on miRNA and his argument for exploratory experimentation is consistent with the general narrative about the shift from pre-genomics to post-genomics. He presents this research as a move away from the reductionist strategies of the twentieth century towards a renewed interest in studying the contingent contexts in which genes, genomes, organisms and environments interact.

Micro-RNA: A case study of scientific practices in post-genomics molecular biology

There is a standard narrative in the history and philosophy of biology about the significance of the shift from twentieth century pre-genomics to post-genomic molecular biology in the twenty-first century. The shift is portrayed as one that has moved from a

reductionist strategy to a holistic strategy of investigating the different phenomena in gene expression. Post-genomic molecular biologists realize that there is no straightforward path between genes (as specific sequences) and their phenotypic effects. They are concerned with the complexity, contingency, and interactions (e.g. between genes and environments) involved in the development of biological systems. Of course, there are some exceptions to this narrative. Gene sequences still occupy a central position in some approaches in post-genomic research and even the research program of epigenetics has been defended by some as a continuation of gene-centric research as it is primarily focused on how molecular mechanisms affect gene expression (Richardson & Stevens 2015, 4; Maderspacher 2010). It is therefore reasonable to state that, unlike Kuhn's notion of paradigms shifts, the transition from pre-genomics molecular biology to the era of post-genomics cannot be neatly represented in terms of a clean break or transition from a reductionist science to a holistic one, at least not without many caveats and qualifications.

But, questions about the nature of the continuities and discontinuities between these periods in molecular biology remains. The philosophers and historians of science cited above are correct in emphasizing the conceptual change that has occurred in how we define and conceive of genes and genomes. What I want to consider next is how this conceptual change has translated into some of the scientific practices in gene expression research. Morange has suggested that, despite the radical changes in our conceptualizations of genes, genomes, and genotype-phenotype maps, there is, perhaps, still room for some kind of fruitful reductionist strategy in post-genomics molecular biology (Morange 2006). Others have been more forceful in claiming that new

approaches and new strategies, particularly non-reductionist ones, are required to pursue research in post-genomic biology. In the following subsections, I consider the case of miRNA discovery and subsequent miRNA research to illustrate some of the continuities this research shares with earlier research in molecular biology, as well as its discontinuities with respect to novel approaches.

What are miRNAs?

Before identifying the class of RNAs referred to as micro-RNAs (miRNAs), molecular biologists were aware of numerous small non-coding or "non-messenger" RNAs (ncRNAs) present in cells, including small nucleolar RNAs (snoRNAs) and small nuclear RNAs (snRNAs). When Andrew Fire and Craig C. Mello determined the mechanism of RNA interference in the late 1990s, researchers were beginning to understand the identity and the functions of these small non-coding RNA molecules (Fire *et al.* 1998). RNA interference, in brief, is the process by which short ncRNAs mediate both transcriptional and post-transcriptional gene silencing via target degradation. Micro-RNAs were later identified as specific kinds of short ncRNA found in both plants and animals, involved in developmental, physiological and pathological processes via its regulation of gene expression. It is estimated that the human genome encodes approximately 1000 miRNAs that target about 30% of genes in many different cell types (Kuhn *et al.* 2008).

In contrast to small-interfering RNAs (siRNAs) involved in RNA interference, miRNAs are coded in the genome, and thus are produced endogenously. miRNA transcripts tend to fold back on themselves to form recognizable hairpin secondary structures, whereas other siRNA molecules are derived from longer hairpin structures or

from single-stranded precursors, such as piwi-interacting RNAs (piRNAs) (Bartel 2009). Most miRNA genes are transcribed by RNA polymerase II, which results in precursor transcripts with a 5` cap and a 3` polyadenylated tail. However, miRNAs can also be transcribed from introns or exons from their host, or target, genes. Some miRNAs can also be derived from polycistronic transcripts.

miRNA biogenesis and functions

Transcription of miRNAs occurs via two steps: First, primary transcripts, referred to as primary miRNAs, or *pri*-miRNAs, are transcribed either from genes, introns or exons of their target protein-coding genes, or polycistronic transcripts. Two proteins, Drosha and Pasha, then process the pri-mRNA. Drosha, an enzyme from the RNAse III family, crops the pri-miRNA into a ~70nt-long precursor transcript, referred to as *pre*-miRNA. The pre-miRNA transcripts feature a hairpin structure with some bumps or bulges in its stem. The precursors also have two unpaired nucleotides that overhang at the end of the hairpin, which help to interact with factors that are important for further processing. The dsRNA-binding protein Pasha interacts with Drosha by helping to orient Drosha's catalytic RNAse III domain to the appropriate locations on the pri-miRNA for cleavage. The pre-miRNA is then carried out of the nucleus to the cytoplasm by exportin-5 with the help of Ran GTPase. Once in the cytoplasm, the Dicer protein processes the precursor into a 22-nt RNA duplex (i.e. miRNA:miRNA*), which assembles with RISC (i.e. the effector complex) to become a mature miRNA (Figure 3).

Mature miRNAs are then used to regulate gene expression by silencing genes through post-transcriptional repression or degradation. To carry out their function, mature miRNAs associate with the Argonaute protein to recognize their target mRNA. In plants,

miRNAs tend to have perfect, or near perfect, complementarity with their target, resulting in the cleavage of the target and its subsequent degradation. In animals, the guiding miRNA, bound with the Argonaute protein, tends to have only partial complementarity with its target, usually on a ~7-9nt region of its target, referred to as the "seed region." This partial match allows for the recruitment of factors that inhibit translation during different stages of the process. Whereas perfect complementarity results in the cleavage and degradation of mRNA, partial matching allows for only temporary translational repression because the silenced mRNA transcripts can be kept intact in the cytoplasm (Figure 3).

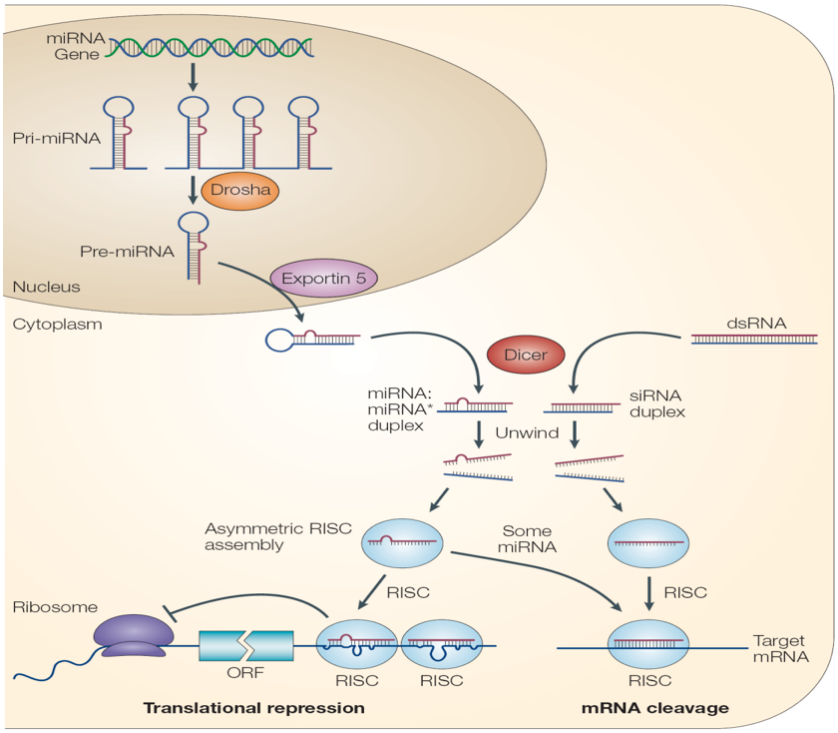

Figure 3: Biogenesis of miRNA (He & Hannon 2004, 524).
Model of the biogenesis of pri-miRNA transcripts to mature miRNA transcripts, with depiction of how miRNAs acts on its targets via translational repression or mRNA cleavage.

A brief history of miRNA discovery

The first miRNAs were detected in the nematode *C. elegans* by Victor Ambros and Gary Ruvkun, and their research associates, in the 1980s (O'Malley, Elliott, & Burian 2010). When constructing cell lineage maps to understand the normal development of the worms, the researchers used mutants to understand how each genetic "switch" turned on or off during development to develop different cell types in the organism during different stages of development. From their experiments, they discovered heterochronic mutants in the *lin* (lineage) gene, which resulted in abnormal cell lineages, and the *let* (lethal) gene, which resulted in arrested development.

Ambros and his colleagues showed that the gene *lin-4* downregulates another gene, *lin-14*, by some form of repression or silencing. *Lin-14* produces a protein that forces cuticle cells to remain in the first larval stage instead of turning into adult cells. When *lin-4* was not expressed, *lin-14* continued to express its protein and the organism stayed in an arrested state of development and failed to develop into an adult (Ambros & Horvitz 1987; O'Malley, Elliott, & Burian 2010). It was only when *lin-4* was expressed that *lin-14* was properly regulated and the organism could continue to develop in the next stages of larval development. The researchers then showed that *lin-4* regulated gene expression post-transcriptionally, because its target, the mRNAs of *lin-14*, were still present in the cells' cytoplasm when *lin-4* was expressed (Bartel 2004). Ruvkun and his team thus characterized the regulatory influence of *lin-4* by showing how it suppressed *translation* rather than *transcription*.

In the early 1990s, Ambros and Ruvkun exchanged notes on their research to discover that the transcript of the *lin-4* gene shared partial complementarity with 3' UTR

regions of the *lin-14* mRNA transcript (O'Malley, Elliott, & Burian 2010). Lee, Feinbaum, and Ambros also began to characterize the product of the *lin-4* gene. They indicated, at first, that the gene's product was approximately 22 to 61nt long, producing very short RNAs that do not encode amino acids for protein synthesis. They then suggested that *lin-4* regulates expression of *lin-14* via an anti-sense RNA-RNA interaction (Lee *et al.* 1993).

Seven years later, micro-RNA research took off when researchers discovered another miRNA gene in worms, *let-7*, and then found homologous versions of some of the miRNAs that were initially discovered in worms in other organisms, including vertebrates (Ruvkun 2008; Reinhart *et al.* 2000; Pasquinelli *et al.* 2000). The discovery of *let-7* by Reinhart *et al.* showed a gene whose product was a 21nt-long RNA that intervened on the 3'UTR of five different heterochronic genes, including *lin-4*. Base-pairing with sites from the 3'UTR of target genes showed that the miRNAs interacted with 39 sites on their mRNA targets, revealing the complexity of miRNA's involvement in post-transcriptional regulation of gene expression. Moreover, genomic comparisons revealed the deep conservation of *let-7* in worms, mammals, and flies (Pasquinelli *et al.* 2000). Meanwhile other miRNA genes were being identified in different model organisms. For example, researchers discovered the miRNA gene, *bantam,* in *Drosophila* (Brennecke *et al*. 2003). That gene codes for a miRNA which functions to downregulate the *Hid* gene by targeting the 3'UTR region of its mRNA. The *Hid* gene activates apoptosis in the cell and its regulation is crucial for normal development.

This deep conservation signaled that miRNAs are part of an ancient regulatory mechanism that has been present in bilateral animals for approximately 400 million years

(Pasquinelli *et al.* 2000; O'Malley, Elliott, & Burian 2010). Knowledge of their conservation triggered a search for the tiny molecules, their targets, and most importantly, their biological functions. As stated above, most miRNAs in animals do not have perfect complementarity with the sequences of their target sites. They also often combine with other elements or modes of gene regulation to carry out their functions. Because of these features, researchers realized that the traditional on/off switching model of gene regulation, depicted in Jacob and Monod's operon model, would be insufficient to understand miRNA interactions and functions. miRNAs have a subtler influence on gene regulation by fine-tuning differential gene expression in post-transcription. Because of the added degrees of complexity, both experimental genetic approaches and sequence-based approaches to predicting or identifying miRNAs are insufficient, and sometimes even unreliable, by themselves (Ambros 2004; Bartel 2009). Biologists and philosophers of biology have correctly pointed out that miRNA research requires "multiple levels of confirmation" from several approaches in molecular biology (O'Malley, Elliott, & Burian 2010). The next section gives an overview of these different methodological approaches and strategies used in miRNA research.

Methodological approaches and strategies in miRNA research

The history of micro-RNA discovery depicts a series of incremental discoveries that depended on the available technologies and experimental techniques of the time, as well as the accessibility of data from certain model organisms, such as *C. elegans*. Current research on miRNA identification in humans and different model organisms, such as *Drosophila*, and understanding the biological functions of these miRNAs once identified, depend on both genetic and genomic experimental methods and tools.

Experimental work in genetics includes two different methods (Ambros 2004). There is the *forward genetic* analysis of miRNA function and the *reverse genetics* approach to studying miRNA. The former kind of analysis proceeds by classical methods in genetics, as well as more modern techniques developed in molecular biology. Forward genetics, simply put, starts with observations of a phenotype and looks for the genes that act as difference-makers in the production of that observed phenotype. The classical methods involved Mendelian cross-breeding, with several methods of generating mutants using chemicals or radiation, but modern forward genetics involves additional laboratory techniques to introduce mutagens that help identify the gene(s) by attempting to "rescue" the phenotype. Within this approach, researchers try to identify the gene (or genes) thought to be a difference-maker for an observed phenotype, then clone it, and then, they look to see whether the gene encodes for a non-coding RNA product in the cell that is involved in mRNA silencing. The canonical miRNA genes discovered in *C. elegans*, *lin-4* and *let-7*, were discovered by the identification of loss-of-function mutations in genes, which played a role in the developmental timing of the sequential larval stages in worms. The miRNA gene, *bantam*, which regulates programmed cell death or cell proliferation in different contexts in *Drosophila*, was also discovered by forward genetic analysis (Hipfner *et al.* 2002). Moreover, the *lsy-6* gene, another miRNA gene found in worms that regulates neural development, was also specified by forward genetic methods. In fact, Ambros contends that this miRNA gene could not have been revealed by any other method (Ambros 2004, 351). Because, he argues, the number of *lsy-6* miRNAs in specific cells was so low, the researchers could not get any clear confirmation of the

gene's expression from northern-blot hybridization procedures[15] (Ambros 2004). They introduced different mutations in the *lsy-6* locus to see if their interventions would result in a rescued phenotype. They tested rescuing activity by looking for disruptions to the distinctive hairpin structure of the gene's transcript. Their results showed that a single point mutation could eliminate rescuing activity by disrupting the hairpin required for the processing of the mature miRNA.

Reverse genetics, as Ambros describes it, is a method whereby a specific gene is knocked-out, knocked-down, or over-expressed to identify the gene's function or its contribution to a phenotype (Ambros 2004, 350). However, testing for gene functions by the method of reverse genetics is often dependent on computational methods to first identify potential miRNA genes. The reverse genetic approach led to the discovery of the *mir-273* gene, a miRNA gene associated with the same pathway as the *lsy-6* gene in worms, which regulates neuron asymmetry by repressing *die-1* mRNA (Ambros 2004, 351). Researchers came to understand the function of this miRNA by ectopic expression of the gene. That is, instead of conducting a knock-out experiment by introducing a loss-of-function mutation, they interfered in the expression of the gene by introducing a transgene with an inactive promoter.

The genomic approach in miRNA research can be illustrated by two different methods, including cDNA cloning and search algorithms in bioinformatics designed for generating computational predictions. cDNA cloning is used by looking at isolated samples of small RNA molecules in the cell as candidates of potential miRNAs in order

[15] Northern-blot hybridization procedure is a common laboratory technique used to detect mRNA transcripts in a sample of cells.

to detect new miRNAs. The small RNAs in the sample are then joined to adaptor molecules and amplified by PCR (Elliott & Ladomery 2011, 404). There are potential drawbacks to using this method, such as producing clones that correspond to other RNA molecules; e.g. rRNA and tRNA. But, such errors can be detected and corrected. However, cDNA cloning methods can fail to clearly reveal the presence of miRNAs in a cell because, like the example of the miRNA transcribed from the *lys-6* gene, miRNAs are sometimes not very abundant in a sample (Elliott & Ladomery 2011, 404).

Finally, bioinformatics search methods play a crucial role in current miRNA research. Using this method, algorithms are designed to search for and identify potential miRNA genes by focusing on some of the familiar properties of miRNA. For instance, because miRNA have precursor transcripts with hairpin structures, algorithms can be generated to search for sequences that give rise to such structures (Elliott & Ladomery 2011, 404). In addition, algorithms can track gene clusters, as miRNAs are thought to be found in clusters as a result of duplication, and they can look for homologues to already known miRNA genes. The results from these searches must then be verified experimentally. Computational methods are also used to discern the targets of miRNA on mRNA transcripts. miRNA targets are often located in the 3'UTR of mRNAs, so searches target the phylogenetic conservation of the 3`UTR of orthologous genes. This kind of search is difficult because a given miRNA can target and regulate more than one protein-coding gene, and a given mRNA transcript can be regulated by more than one miRNA. In most animal miRNA, the seed region – the guide sequence of the miRNA at the 5' end and the 3' end of the targeted mRNA, usually between 6 to 8 nucleotides long – is the standard used to search for and identify miRNA targets. Figure 4 illustrates the

way that some genomic and computational approaches have been used alongside genetic approaches to discover and study particular miRNAs. These approaches continue to be "refined through an iterative process of *in silico* prediction and *in vivo* experimentation" (Ambros 2004, 354).

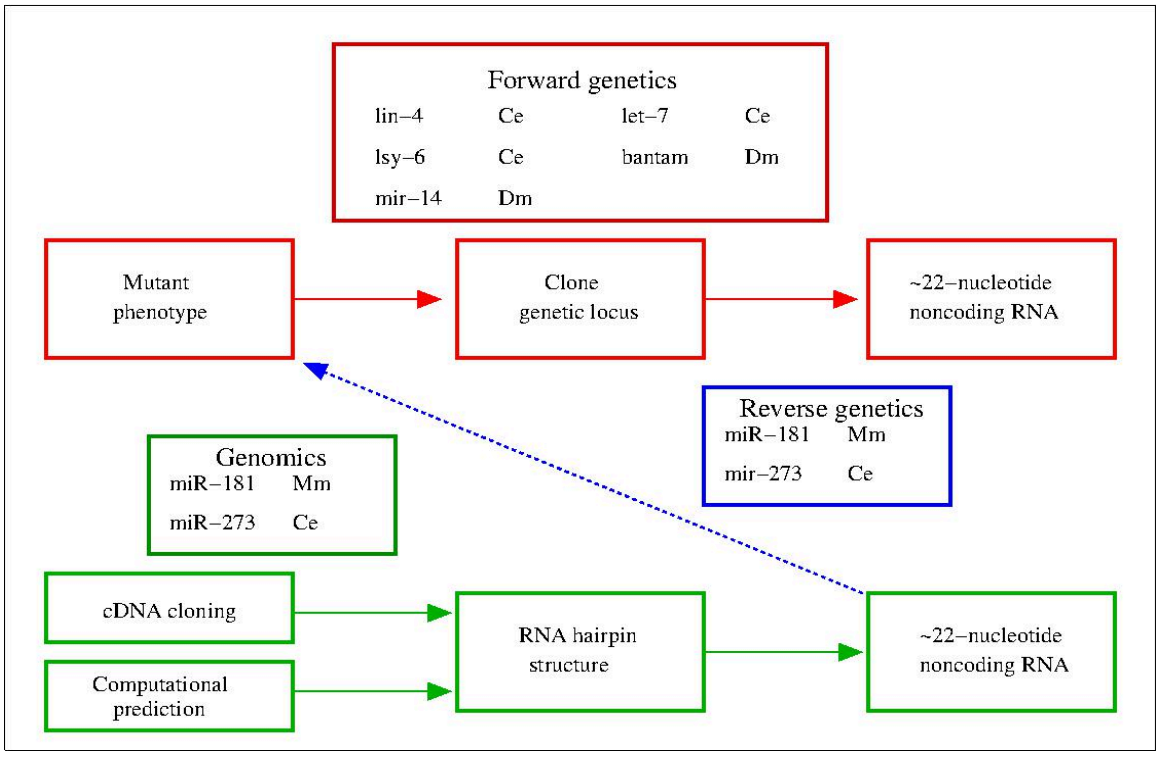

Figure 4: Genetic and Genomic Approaches to miRNA Gene Discovery (Ambros 2004, 352)

Post-genomic miRNA research: A continuation of iterative and integrative research in molecular biology

The case of the discovery of miRNA, and the subsequent research program it has generated, allows us to see that several research strategies and novel experimental approaches in genetics and genomics were required to understand the biological functions of these tiny regulatory molecules. Because miRNAs work to "actively sculpt expression domains through a combination of tuning and classical switch targeting," their study is

crucial for understanding the regulation of gene expression in a variety of contexts, from development to metabolism (Bartel 2009). miRNA research is continuous with earlier research on genetic regulation, such as the development of the *lac* operon model, in the way it displays the iterative and integrative use of different approaches and strategies used to trace the process of genetic regulation. That is, miRNA research presents an example of the iterative and integrative process between genetic and genomic methods. As genetic experimentation accumulates evidence of these molecules and their regulatory functions, better bioinformatics tools are designed to track new potential candidates of miRNAs and their regulatory sites, and then further genetic experimental manipulations are preformed to reach a better understanding of the regulation of gene expression.

miRNA studies also reveal that there are new conceptualizations of genes and the genome at play in post-genomic molecular biology that differ from the assumptions about genes in pre-genomic biology. For example, biologists know that the regulatory function of any given miRNA gene cannot be determined simply by studying its sequence or structure. The functions of miRNA genes are known to be context-dependent, as they depend on the presence of specific molecular complexes to carry out their functions, and on the presence of specific mRNA targets and their particular context. Given this, the biological functions of miRNAs might be said to depend on the emergent state of the entire cell. However, whether the idea of emergence as a cell's molecular and genetic interactions precludes any sort of reductionist strategy in post-genomic molecular genetics depends on how we represent and understand the context-dependencies of miRNA function. The next section of this chapter addresses that issue.

A case against anti-reductionism (and reductionism) in post-genomic molecular biology

Much of the philosophical debate surrounding discussions about how to understand genes and their functions depends on particular theses about reductionism and emergence. The terms *reductionism* and *emergence* have taken on several different meanings in the history of philosophy of science. I begin with a survey of these meanings to indicate how some theses about reductionism and emergence in science are related to each other, and to prevent confusion about the use of the terms in the following sub-sections. I then present Dupré's arguments in defense of anti-reductionism and strong emergence in the molecular life sciences. Last, I then present my argument against Dupré's position by appealing to how diagrammatic representations of interaction between molecular components in the cell in current biomedical research on miRNAs can account for context-dependency.

The different varieties of reductionism & emergence

Philosophers of biology were initially interested in questions about reductionism with respect to the relation between the classical Mendelian concept of the gene and the molecular concept of the gene. What motivated this research question was the idea that perhaps the theory of Mendelian genetics could be reduced to the theory of molecular genetics. That philosophical problem is concerned with *epistemic* reductionism, which characterizes the relation between two scientific theories or entire scientific domains.

Ernst Nagel first proposed this model of reduction to demonstrate how classical thermodynamics can be reduced to statistical mechanics. Nagel's schema is that a reducing theory T_1 reduces the reduced theory T_2 if and only if the laws of T_2 are

derivable from those of T_1 (Nagel 1961). Nagel's type of reductionism assumes formal relations between the theoretical terms or the laws of the reduced and reducing theories, which are described by bridge laws or coordinating definitions. But, some philosophers of biology came to agree that this type of reductionism was inapplicable to the two genetic theories because they lacked laws or a set of statements that could be expressed in a formal language, and so they could not partake in formal deductions (Hull 1979; Kitcher 1984).

Shifting to think about the semantic relation between the theories or theoretical terms, others also argued against the possibility of reducing the Mendelian gene to the molecular gene. The problem was that biological phenomena described in Mendelian genetics terms are multiply realizable by underlying molecular mechanisms. That is, there are multiple kinds of molecular mechanisms described by a molecular theory that can produce a phenomenon described in classical genetics (Griffiths & Stotz 2013, 59). One example is the phenomenon described by the concept of dominance. In classical genetics, dominance occurs when the phenotype of a heterozygote at a particular locus, with one recessive allele, has the same phenotype as a homozygote at that locus because the dominant allele suffices to instantiate the particular functional phenotype. At the molecular level, there is no one kind of mutation or molecular mechanism that underlies the phenomenon of dominance. Given these peculiarities in the study of genetics, most philosophers, with some exceptions,[16] have reached an anti-reductionist consensus with respect to epistemic reductionism (Waters 1990).

[16] C.f. Rosenberg, 2008.

However, philosophers became more interested in *explanatory* and *methodological* reductionism, rather than theory reduction, after developments in molecular biology during the 1970s and 1980s. *Explanatory* reductionism is a form of epistemic reductionism in which lower-level explanations, such as those offered in molecular biology and biochemistry can explain phenomena that is described at a higher level of biological organization (Sarkar 2005). *Methodological* reductionism is associated with *explanatory* reductionism as it prescribes the methods or research strategies that can reliably lead to reductionist explanations.

There are two further senses of reductionism—ontological and metaphysical reductionism. In contemporary philosophy of biology, most philosophers begin their discussions of reductionism by clarifying their stance on *ontological* reductionism with respect to organisms and the life sciences. This is a stance about what exists, or about what organisms consist of. And, most accept a materialist position (e.g. Dupré 2010; Keller 2010). That position stands in contrast to the belief in vitalism; i.e. the belief that all life forms must be animated by some kind of non-physical *élan vital* (vital force).

Metaphysical reductionism, in turn, addresses the hard philosophical question about the nature of the connection between the higher-level biological phenomenon and the lower-level physical one assumed by ontological reductionists. A common concept related to this type of reductionism (or, more accurately, related to metaphysical *anti*-reductionism) is downward causation, which, in short, is the idea that the properties and behavior of the individual components in a system are caused by higher level properties of the whole system. A common example used to illustrate the idea of downward causation is the idea that mental states can cause physical states. The idea of downward

causation is a contentious one, and I return to it in the following section on Dupré's argument for anti-reductionism. But, before I turn our attention to his position, I want to briefly discriminate between the two different senses of emergence found in the philosophical literature.

Philosophers have drawn a distinction between weak emergence and strong emergence (Bedau & Humphreys 2008; Dupré 2010, 2012; Pigliucci 2014). *Weak* emergence is the view that the emergent properties of a system are completely dependent on (and determined by) its parts, their properties and interactions, but because there might be a lack of general laws or principles at the lower-level that explain certain kinds of emergent properties, the more practical way to explain them might be by appeal to higher-level properties or through simulations of the interactions between the system's components. This understanding of emergence captures the idea that it might be too practically cumbersome to study or explain the behavior of certain complex systems by appealing to their lowest-level components (Pigliucci 2014). Such a weak concept of emergence is therefore consistent with certain kinds of reductionist theses. In contrast, the concept of *strong* emergence holds that complex systems, specifically in biology, are not determined by the properties of their parts, or their interactions. Strong emergence requires that emergent properties of a system have a causal influence on a system's components.

There is an additional, but related, distinction that philosophers have used to discriminate between epistemic theses about emergence and their metaphysical implications. O'Connor and Wong (2015) present a distinction between *predictive* emergence and *irreducible-pattern* emergence. The former concept is defined as:

"properties [that] are systemic features of complex systems which could not be predicted (practically speaking; or for any finite knower; or for even an ideal knower) from the standpoint of a pre-emergent stage, despite a thorough knowledge of the features of, and laws governing, their parts" (O'Connor & Wong 2015). The idea of predictive emergence is similar to that of weak emergence, and emphasizes the pragmatic challenges of studying and explaining complex phenomena. The latter concept of *irreducible-pattern* emergence is defined as: "properties and laws [that] are systemic features of complex systems governed by true, law-like generalizations within a special science [and] irreducible to fundamental physical theory for conceptual reasons. The macroscopic patterns in question cannot be captured in terms of the concepts and dynamics of physics" (O'Connor & Wong 2015). The idea of irreducible-pattern emergence emphasizes the conceptual challenges of explaining certain phenomena using the theoretical terms and concepts of a more fundamental theory. The example of irreducible-pattern emergence most cited in the literature, offered by Fodor, is that of the struggles of an "immortal economist" explaining the relation of supply and demand, or boom-and-bust cycles, in terms of the properties of quarks or the principle of indeterminacy (Fodor 1974; Pigliucci 2014). Fodor holds that this is impossible not because it is impractical to do so, but because the concepts and theoretical terms in physics are just not useful to explain the phenomena and the regularities which economists seek to understand.

While both concepts of emergence here are epistemic in nature, they can also have metaphysical implications. For example, if we think that "metaphysical statements ought to be evaluated in terms of our epistemic access to the world, meaning that what we can know empirically should constrain how we think metaphysically," then we should

adopt a metaphysical position about the *irreducible-pattern* type of emergence (Pigliucci 2014, 263; see also Ladyman & Ross 2009).

Given the general landscape of the different categorizations of reductionism and emergence just described, there remain many nuances in what can be implied by specific reductionist and anti-reductionist positions. I next provide a brief re-construction of Dupré's argument for anti-reductionism in the context of post-genomic biology to tease apart some of its implications.

Dupré's arguments

In a recent edited collection of current debates in the philosophy of biology, Dupré was tasked to answer the question: "Is it possible to reduce biological explanations to explanations in chemistry and/or physics?" and his response was negative (Dupré 2010). Dupré is somewhat ambiguous about the exact meaning of the question and about what exactly the phrase "physics and/or chemistry" means. He claims that he is in agreement with Keller (who defends a reductionist position) that biology cannot be derived from the theories or laws of physics and chemistry (Dupré 2010, 33). And, like Keller, he is a materialist about the ontology of life. In later parts of the discussion, he describes the *reductionist principle* to outline what he takes to be the crucial disagreement between reductionists and anti-reductionists. His principle seems to imply that all explanations and representations in molecular genetics and biochemistry below the level of the cell are reductionist explanations. In short, his arguments are targeted towards the relative merits of a more general, *epistemological* kind of reductionism, with clear implications for explanatory, and perhaps methodological, reductionism. However,

Dupré also defends a strong notion of emergence, which, he claims, has certain metaphysical implications.

Dupré's reductionist principle & strong emergence

Dupré begins by clarifying what he takes to be implied by reduction in the question posed above. He does so by articulating the *reductionist principle*:

> The reductionist's claim should be that [the organism] is nothing but a collection of physical parts assembled in a certain way. So here, finally, is a proposition that we might expect the reductionist and the anti-reductionist to disagree about: if we knew everything about the chemicals [or molecules] that make up [an organism], and the way they are assembled into cells, organs, and so on, we would in principle, know everything about the [organism]. Reductionists will generally endorse something like this, whereas anti-reductionists will deny it. Let me call this… the reductionist principle (RP) (Dupré 2010, 34).

The passage seems to indicate that the reductionist's explanatory toolkit includes the cell's molecular and genetic components, and the "way they are assembled," or the way they interact. What is more, the "in principle" part of the reductionist principle is important here because it stops short of dismissing the work of other life sciences (as well as all the other special sciences that are not part of the physical sciences) as misguided or superfluous. That is, the reason there remain non-molecular explanations in biology could be because our current state of knowledge of the molecular foundations of life is underdeveloped, or we might concede that non-molecular explanations are a pragmatic short-hand to explain and predict some phenomena. So, the methodological and explanatory reductionist's claims are not necessarily normative. That is, the reductionist

might not think that we *should* explain everything about life in terms of its molecular structure. But, descriptively, and *in principle*, it is possible to do so.

If we limit our characterization of reductionism in this discussion to the reductionist principle articulated above, then it follows that Dupré's anti-reductionist's objection to the "in principle" argument relies on a strong notion of emergence. Recall that, unlike a weak notion of emergence, the strong notion of emergence precludes the possibility of explaining the behavior or properties of a system by appealing to its components, and not only because doing so might be practically unmanageable. Dupré explicitly claims to defend strong emergence and argues that there are systems, specifically in biology, that are not determined by the properties of their parts, or their interactions (Dupré 2010, 35). For Dupré, strong emergence in biology entails some form of downward causation.

Dupré admits that most reductionists feel uneasy about downward causation because it seems to invoke a mysterious force. However, he thinks that it is a natural way to think about biology (Dupré 2010, 42). For instance, he argues that phenomena, such as genome behavior (e.g. the transcription of particular sequences) and protein-folding, are explained "by appeal to features of the whole" and not merely additional interactions (Dupré 2010, 42). With respect to protein-folding, he writes:

> There is a very specific environment, in this case one replete with appropriate chaperones, which endows the amino acid sequence with the capacity, or disposition, to fold in a particular, functionally desirable way. Still more it is a specific environment *that disposes* the various relevant parts of the genome to produce, in the end, an appropriately folded protein. And again, this environment

> is not something that could possibly be generated *de novo* by the genome but, on the contrary, it is one that took a few billion years to evolve. The cell, I think we must say, with all its intricate structure and diverse contents, *is what causes* these contents to behave in these life-sustaining ways (Dupré 2010, 42; my emphasis).

Dupré is arguing that explanations about certain macromolecular mechanisms, such as protein-folding, refer to the genome or the entire cell (the complex systems in which the molecular components operate). But, more strongly, he is also arguing that these systems have a causal influence on the behavior of the molecules and their interactions. The point Dupré is making here is stronger than merely pointing out that many molecular-level phenomena in biology are described in terms of their functional capacities or dispositions within a particular system. His point is not merely semantic. He is presenting a metaphysical argument for strong emergence and downward causation, and against reductionism.[17]

Context dependency in molecular genetics

Dupré further argues for downward causation by appealing to the fact that the main theoretical terms and entities in molecular biology are also dependent on context, in a similar way to descriptions of molecular phenomena, such as protein-folding described above. To make this argument, Dupré attends to the changing concept of the molecular gene to illustrate this idea of context-dependency. In biology, it might be assumed that a

[17] In some parts of the paper, Dupré seems to be presenting a semantic argument for anti-reductionism. For example, he writes: "…we are talking about science, not Nature. Biological explanations are part of biology, not part of the world, and biology, like any other science, is an articulated conceptual structure not a repository of things-in-themselves…and the fact that biology—a science—works with *concepts* that depend on the larger systems of which they are a part, as well as on their constituents, is a fatal objection to the [reductionist] claim" (Dupré 2010, 37-38).

gene simply corresponds to a particular sequence of nucleotides. However, in addition to a specified sequence of nucleotides, scientists sometimes use the term to refer to a unit that tracks inherited phenotypes in organisms, a gene's protein product, or sometimes its primary transcript. And, other times, a gene can refer to a complex set of entities, including a transcript's exons and introns, alternative mRNA isoforms, and other regulatory RNAs that take part in a gene transcript's expression. Stotz and Griffiths (2004) have investigated how scientists actually use the term 'gene' and showed that there is little consensus on an accepted definition. Though this view does not entail that there are no such things as genes, according to Dupré, it does suggest that there is no one true way to carve up the genome into genes. Partitioning the genome into genes is only possible by categorizing a particular sequence of nucleotides according to its functional role in the cell. He writes: "…the conceptualization of the genome, as an object of study and as divisible into discrete functional constituents, requires that it be placed in the wider *context* with which it interacts" (Dupré 2010, 42). Dupré, therefore, takes the cell as the complex system that is required to understand the meaning of the entities to which molecular geneticists refer.

He then argues that the context-dependency of genetic entities and their interactions also captures a metaphysical notion of downward causation. He acknowledges that "reductionist methods" have been useful in molecular biology, but claims that they cannot address or explain some of the features of biological systems, such as their stability, order, and function (Dupré 2010, 44). He argues that it is those *emergent* properties of the systems that in fact "constrain and causally influence the behavior of their molecular constituents" (Dupré 2010, 44). According to Dupré, the

reason that a reductionist strategy cannot explain that particular set of behaviors or properties is because biological complex systems, like cells and organisms, are order-preserving, evolved systems. He states: "Biological order is the extraordinary achievement of systems honed by billions of years of evolution. It is not something that comes for free with the determinism of the physical and chemical worlds" (Dupré 2010, 44). What I take Dupré to be saying here is that, unlike physical systems (according to him), there is some degree of contingency by which biological systems, such as cells, behave the way they do and display the particular properties they do. Because of that fact of contingency, it is the higher, systems-level, properties or behaviors, like order and stability, that cause, quite literally, the molecular and genetic interactions within cells.

Dupré's anti-reductionist position

Dupré is right to emphasize the context-dependency required to identify molecular and genetic entities and to study their functions in post-genomic molecular biology. But, I suspect that he is wrong to think that context-dependency entails a strong notion of emergence. In fact, there is much that molecular explanations can explain about biological systems, such as cells, by explaining what Dupré calls context in terms of the interactions between the cell's molecular and genetic components. However, before I get to my argument against what Dupré infers from the context-dependency of molecular entities, I want to clearly summarize what I take to be his argument against reductionism.

Dupré argues that we cannot fully understand complex biological systems by a detailed knowledge of their constitutive parts, their properties, and their interactions. The main reason for his claim against the *reductionist principle* is that identifying and characterizing the molecular or genetic components of complex systems requires the

context of the larger system of which they are parts. Dupré takes strong emergence to imply downward causation, and the mark of downward causation is the context-dependency of molecular mechanisms and genetic entities. So, because there is context-dependency in the domain of molecular biology, reductionist strategies are not optimal.

Context-dependency without anti-reductionism

Dupré's position seems to carry the normative implication that molecular biologists ought to do something different from merely elucidating the ways in which molecular constituents interact with each other. Yet, much of current work in molecular biology aims to provide representations of molecular constituents and how they interact to explain certain phenomena. Context is represented as other (collections of) entities that interact with the focal entities in the investigation. Hence, I argue that Dupré's ideas of strong emergence and context-dependency do not entail, or lead to, an acceptance of anti-reductionism across the board. In fact, I question whether there is any clear conceptual relation between the context-dependency of the genetic and molecular elements within a cell, on the one hand, and any normative claims about reductionism or anti-reductionism, on the other.

Explanations and representations of interactions in molecular biology

The methods used in post-genomic miRNA research, as I have sketched above, are numerous and inter-dependent. One of the strategies, however, can be described according to Dupré's account of reductionist explanation insofar as its aim is to elucidate the mechanisms in which miRNAs play a causal role. To help make my argument, I rely on the difference between approaches and research strategies. Approaches include the set of experimental tools and techniques that researchers use in experiments, such as the

computational search processes in genomics and knock-down experiments in reverse genetics. The explanatory strategy in much of post-genomic miRNA research accounts for the way researchers conceive of the phenomena of interest, as well as how they conventionally understand, represent, and explain them.

Biologists and philosophers have given several accounts of post-genomic molecular biology, by focusing on its new approaches. Some have divided the approaches in current molecular biology into different categories of strategies, like data-driven methods, hypothesis-driven methods, and exploratory experimentation (Burian 2007; O'Malley and Soyer 2012). Exploratory experimentation is needed when there is not a clear theoretical framework that can provide background assumptions and boundary conditions to guide experimental design so that a set of outcomes can be predicted, or expected. Exploratory experimentation describes the process by which researchers aim to discover patterns or regularities that call out for further investigation and explanation by varying parameters and conditions. Some researchers adopt this strategy of exploratory experimentation within the context of miRNA studies. However, here I focus on miRNA target validation studies because it illustrates a research strategy that attempts to provide mechanistic descriptions of miRNA functions by representing and explaining how different molecular constituents causally interact in cells.

In the philosophical literature, a mechanism is defined as a group of "entities and activities organized such that they are productive of regular changes from start or set-up to termination conditions" (Machamer, Darden, & Craver 2000). Scientists attempt to provide *descriptions* of mechanisms via schematic, diagrammatic, or mathematical representations – i.e. *models* – to explain the phenomena (Tabery 2004). Philosophers

have argued that one of the main epistemic goals of molecular biology is to "elucidate" models of mechanisms (Craver & Darden 2013). Mechanistic models explain by more than merely describing the regularity of certain phenomena. Models of mechanisms provide researchers with control over the phenomena by identifying the causal links in a system, and picking out the loci of control in these systems. Researchers can then systematically intervene in the system to change its outcome. It is important to note that mechanisms, so understood, are neither necessarily reductionist or anti-reductionist.[18]

miRNA research in the context of cancer

To illustrate the research strategy used to elucidate the mechanisms and causal interactions between molecular genetic components, I present a case of biomedical research on the function of miRNA in cancer. The epistemic goal of providing a mechanistic description of miRNA function is evident in the biomedical context. One of the most pressing challenges in miRNA biology is to discover the mRNA targets of mammalian miRNAs. Doing so is often a preliminary step towards understanding the regulatory structures in which miRNA are involved, and inferring their larger biological functions. In the biomedical context, target validation studies have been important for determining the role of miRNAs in cancer. *miR-15-16* was the first miRNA found to be implicated in cancer. But, since then, many other miRNAs involved in cancer have been discovered and shown to regulate processes such as cell proliferation, cell differentiation, and apoptosis (Garzon *et al.* 2006). miRNAs can act either as oncogenes, which are

[18] While there is an extensive literature on mechanisms and mechanistic explanations in the philosophy of biology, for the purposes of this paper, it suffices to summarize this general description of the practices involved in searching for and elucidating mechanisms in molecular biology.

typically upregulated in tumors and promote oncogenesis, or as tumor suppressors, which are typically downregulated and target the mRNAs of oncogenes. They carry out these functions in a variety of ways, including sustaining proliferative signaling, evading growth suppressors, activating invasion and metastasis, enabling replicative immortality, inducing angiogenesis, and resisting cell death (Esquela-Kerscher 2006).

Researchers have found differences between expression profiles of miRNAs in cancer cells and those of normal tissue (Calin & Croce 2006; Elliott & Ladomery 2011). For instance, in a significant and oft-cited study by Johnson *et al.* (2005), researchers found that the expression of the *let-7* homologue in humans was comparatively lower in patients with lung cancer and inferred that the miRNA played a regulatory role in the incidence of lung cancer (Johnson *et al.* 2005; Elliott & Ladomery 2011, 408). The researchers thought that the *let-7* homologue in humans was involved in the regulation of the expression of *Ras*, a family of oncogenes involved in cell-signaling whose mutations often lead to cancer (Elliott & Ladomery 2011, 408). In their investigations, they used a variety of methods, including animal experiments, in vitro studies with HeLa cells, miRNA microarray analysis, and Northern analysis to provide supporting evidence for their hypothesis about the regulatory function of the miRNAs. The researchers used *let-7* in a model organism, *C. elegans*, to hypothesize about potential complementary sites on the 3'UTR regions of target transcripts. Using computational search tools, they identified *let-60/Ras* as a potential target. From this comparative work (along with other analyses), the researchers found that human *Ras* transcripts contain similar *let-7* complementary sites (LCSs), which allows the *let-7* homologue to regulate *Ras* expression in a similar way to the regulation that occurs in *C. elegans*. The researchers also produced data that

showed that *let-7* expression is down-regulated in lung cancer tissue, while the RAS protein is over-expressed in these same samples, leading the researchers to infer a possible causal relationship between the two gene products. They then showed, in *in vitro* experimentation, that the overexpression of the miRNA gene inhibited the growth of a lung cancer cell line. They inferred from all of these lines of evidence that the *let-7* miRNA gene functioned as a tumor suppressor.

In this case, the identification of the mRNA target was crucial in determining the role of miRNA as a tumor suppressor because the mRNA target is part of the *Ras* pathway. That pathway is part of a well-studied regulatory network involving other molecular genetic components in the cell (Cox & Der 2010; Malumbres & Barbacid 2003). Scientists are able to take into account the relevant pieces of the cellular and inter-cellular context by extending the interactions of the miRNA with a previously established model of the interactions involved in the *Ras* pathway. The explanation of the functional, regulatory role of the miRNA in the context of lung cancer remains "reductionist," according to Dupré's sense of the term, yet it takes into account the context of the system by including an extension of the miRNA's interactions with other molecular components into the foreground of the research.[19]

Moreover, researchers have also demonstrated that miRNAs play a role in the control of the expression of the protein products of proto-oncogenes, which contribute to tumor growth, without any genetic mutations or alterations to the proto-oncogenes

[19] My claim here is similar to Delehanty's (2005) description of a "reductive" molecular explanation of the behavior of the slime mold, *Dictyostelium discoideum*. In brief, Delehanty argues that emergent properties are susceptible to reductive explanations by extending mechanisms to incorporate the required context.

themselves (Mayr & Bartel 2009). In this case, miRNAs are believed to play the role of repressive mediator in the expression of these genes. The mRNA transcripts of these proto-oncogenes can be instantiated in different isoforms, some of which have shorter 3'UTR regions because of alternative cleavage and polyadenylation (APA) mechanisms (Shi 2012; Kornblihtt *et al.* 2013). Because miRNAs generally target the 3'UTR of mRNA transcripts, miRNAs cannot regulate the alternatively polyadenylated mRNA transcripts which lack the regions with the targeted sites. Consequently, the shorter mRNAs, because of the increased stability of the shorter isoforms, tend to produce ten-fold more protein, which in turn may play a role in oncogene activation. In this example, identifying the target genes and the target sites of miRNAs is a crucial step to understanding their regulatory functions. But, more importantly, an important part of the explanations offered is centered on the interactions between particular molecular constituents and their extended interactions with other pathways in the cell.

miRNA research, in the context of cancer, requires genomic studies and genetic experimentation to identify and understand the functional roles of particular miRNAs in specific kinds of cancers. That is, researchers begin with bioinformatics tools to search for candidates by looking for homologues, clusters, or hairpin structures to identify miRNA genes and by looking for sequence complementarity, thermodynamic stability, or evolutionary conservation to identify their targets genes and sites. They must then investigate the implications of repression or activation of miRNA molecules on the function of particular genes and their expression profiles within a cell or tissue through further genetic interventions. The kind of strategy within this research is consistent with the kinds of explanations that Dupré deems to be reductionist. Its aim is to provide

representations and explanations of the causal interactions between different molecular genetic components within the cellular environment in order to explain or infer the biological functions of miRNAs in the cell. In the cases of post-genomic miRNA cancer research, which I have described above, the miRNAs' context ends up being represented as additional entities that interact with other known entities and their interactions.

Context-dependency does not entail anti-reductionism (or reductionism)

Post-genomic research on miRNA can and does take into account the context in which genetic and molecular components interact. Dupré cannot defend a strong anti-reductionist position and strong emergence on the basis that the biological functions of genetic and molecular elements of a cell are dependent on context, if context amounts to the elements' causal interactions with other molecular components and networks within cells or tissues. There seems to be no logical connection between Dupré's idea of context-dependency in post-genomic biology and any forms of reductionist or anti-reductionist theses. Perhaps, then, we are better to move beyond the reductionist/anti-reductionist debate when thinking about post-genomic molecular biology in the history and philosophy of science, and to conceive of the properties of complex biological systems, such as genomes, cells, and organisms, instead in terms of the causal interactions of their components.

Lastly, I want to briefly address a possible rebuttal to my argument against Dupré's position on emergence. Recall that part of Dupré's argument rests on the idea that the reductionist strategy, as he sees it, cannot adequately explain particular kinds of emergent properties of biological systems, such as stability, order, and function. So, it seems that Dupré could accept my claim that certain functional properties, such as the

functions of particular miRNA within networks of molecular interactions, can be explained by the research strategy that I describe. However, he might insist that those types of properties are not the only, or the most important, properties in the explanatory domain of molecular biology. He writes: "A science such as molecular biology tells us not only how particular entities come to have the complex capacities they do, but also how complex systems enroll some of these capacities to create stability, order, and function" (Dupré 2010, 44). Those latter sorts of properties might still preclude what he takes to be a reductionist strategy.

It might seem that Dupré thinks that accounting for those properties involves identifying laws, or generalizable regularities, that govern complex biological systems. Yet, he claims that the non-reductionist, "top-down approach" is required to explain the "*actual* behavior" of systems and offer "higher-level description of *particular* systems" (Dupré 2010, 43; my emphasis). The practices in biomedical research show that researchers are, in fact, trying to understand the *actual* causal interactions between the molecular genetic components of cells and providing mechanistic representations of behaviors of *particular* systems. Doing so remains a significant goal in biomedical and biological research because it gives researchers insight into where they might intervene to change certain outcomes.

However, molecular biologists are often not only after one-off models or representations of particular systems. Like any science, they also aim for models that can be generalizable, though perhaps not universal. Providing causal representations of the interactions between genetic molecular components, in terms of mechanisms or networks, can also help to fulfill that goal. If there are particular causal dependencies or structures,

such as certain types of regulatory network systems that reoccur again and again in molecular biology, then explicitly modelling and representing these systems in terms of the variables that correspond to particular genetic or molecular features of the system will provide insight into the causal dependencies within biological systems. It may also give us insight into what kinds of causal dependencies at the molecular genetic level generate the patterns that are observed at higher levels of biological organization. Those patterns are what I take Dupré to be referring to when he mentions the properties of order and stability. If certain types of causal interactions reliably produce these patterns, and that can be established only with further empirical investigations, then, once again, Dupré's metaphysical notion of downward causation would seem superfluous. So, even here, Dupré's arguments for strong emergence seem to lose their bite.

Conclusion

In this chapter, I first described a standard narrative of the historical trajectory of molecular biology in the history and philosophy of science. This narrative presents a shift from reductionism to some form of anti-reductionism in biology, while highlighting the changing concepts of the gene and the genome in molecular biology. I then considered how that conceptual change has translated to the scientific practices in post-genomic biology. Using miRNA research as an example, I argue that there are many new approaches used to detect miRNA, their targets, and their biological functions. However, despite the adoption of many new tools and techniques, the iterative and integrative processes between the approaches and strategies are continuous with previous episodes in twentieth century molecular biology.

Next, I showed that research strategies in post-genomic molecular biology can take into account the context dependencies of molecular genetics. Many miRNA studies in biomedicine seek mechanistic representations of interactions between the genetic and molecular components of the cell to understand the regulatory role of miRNAs in cancer. The claim that the biological functions of molecules, like miRNAs, cannot be inferred without taking into account the cellular context does not preclude molecular-level explanations. These explanatory strategies illustrated in the case of miRNA cancer research are also in part continuous with earlier research strategies of twentieth century molecular biology, such as the research guiding the development of the repression model of regulation of the *lac* genes in *E. coli*, presented in the previous chapter (see Figure 2).

An adequate account of the epistemic practices in post-genomic molecular biology is important for epistemological reasons within philosophy of science and for gene expression research, as well as for broader sociological issues concerning scientific practice. First, understanding the epistemic aims and constraints of a given research project is crucial for designing successful experiments, evaluating its success and failures, and creating theoretical frameworks to guide future work. Thinking about post-genomic biology through the lens of the reductionist/anti-reductionist framework seems to blur these epistemic aims rather than illuminate them. Second, evaluating the promise of research proposals based on their methodology and experimental design is an important task of funding agencies. With increasing cuts to funds appropriated for basic research, promising programs that are described as "reductionist" might be overlooked due to the opinion that post-genomic biology is so radically different in nature than previous research in molecular biology and that the so-called reductionist biology of the

twentieth century is passé. On the other hand, research described as “anti-reductionist” might be dismissed as non-tractable. In either case, these are reasons to be wary of using these labels to describe current research in post-genomic biology.

THE EPISTEMIC VALUE OF GENETIC NETWORKS IN DEVELOPMENTAL EVOLUTION

Introduction

In recent years, scientists, historians, and philosophers have discussed the prospects of extending or reconceptualizing some aspects of the Modern Synthesis paradigm of evolutionary theory (Pigliucci & Müller 2010; Laland *et al*. 2015; Laubichler & Renn 2015). The old paradigm of the twentieth century has been criticized for being too narrow, or sometimes too reductionist in its gene-centered approach, to fully account for the richness of evolutionary phenomena or to track the many factors that contribute to them (e.g. Oyama, Griffiths, & Gray 2003). Among the new empirical research programs in evolutionary biology that have challenged some previous assumptions and expectations about evolution, many philosophers have focused on the recent convergence between development and evolution (Amundson 2005; Hall 2000, 2012a; Laubichler 2010; Laubichler & Maienschein 2007, 2013; Maienschein & Laubichler 2014; Sansom & Brandon 2007).

In this chapter, I focus on one of these research programs referred to as developmental evolution. Developmental evolution is focused on the role of developmental mechanisms in phenotypic/morphological evolution, the generation of novel types, and the origin of variation. I demonstrate how research on the regulation of gene expression in developmental evolution has generated interesting philosophical problems regarding the patterns of morphological evolution, as well as the concepts of homology and evolutionary novelty. The focus on gene regulatory networks has been fruitful in developmental evolution, and models centered on gene networks can provide

valuable explanations of evolutionary phenomena such as the generation and origin of variation, but there remain interesting conceptual and empirical challenges to some investigative strategies reliant on gene networks.

I distinguish between two strategies using networks to conceptualize gene expression. There is a mathematical strategy to study gene networks wherein the focus is on abstract topological features of networks. And, there is a biological strategy to studying gene networks wherein the focus is on the biological properties of the molecular or genetic constituents of gene expression. I point out some of the challenges that the strategy based on the biological properties of networks might face. I then propose a framework to combine what I call the relational or logical properties of certain types of regulatory systems with the biological strategy. This might serve as a basis for developing an integrative research strategy for studying networks in development evolution.

Developmental evolution as a research program

There have been several recent attempts to draw epistemological and methodological distinctions between the research programs labeled as evolutionary developmental biology and developmental evolution (Hall 2000; Laubichler 2010; Laubichler & Maienschein 2007, 2013; Maienschein & Laubichler 2014; Wagner, Chiu, & Laubichler 2000). Most notably, Manfred Laubichler and Jane Maienschein have provided a critical look at different historical trajectories tracing the convergence of developmental and evolutionary biology in recent decades. Their work has been crucial for shedding light on different epistemologies at work within these research traditions. It has also elucidated the ways in which some research within this convergence can be

interpreted as a synthesis of evolution and development, whereas other research might offer a significant challenge to the explanatory structure of the Modern Synthesis paradigm of evolutionary theory. To better understand the nature and origins of the research problems in development evolution, to which I attend in the following section of this chapter, it is worthwhile to begin with a brief exposition of the different epistemologies at work in these research traditions. Evolutionary developmental biology is concerned with elucidating the genotype-phenotype map. The concept of developmental constraints has played an influential role in creating a theoretical space for development in evolutionary theory within this research tradition. Developmental evolution is concerned with explaining the nature and origin of variation, and thus the focus is on the ways in which developmental and genetic mechanisms generate novel phenotypes and phenotypic change. In what follows, I present parts of the historical narratives presented by Maienschien and Laubichler and the distinctions they have drawn between evo-devo and devo-evo, but I also emphasize how the diverging epistemologies have been represented and taken up in the philosophy of biology.

Evo-devo and developmental constraints

As Laubichler and Maienschein have noted, the history of a science is often presented as a succession of changes and developments that have progressively led to the theories and beliefs that make up the current state of knowledge. They have labelled this standard narrative, "From Darwin to Evo-Devo," to emphasize the way it portrays its development as "a story of completions and syntheses that not only celebrates Darwin's genius but also implies an implicit progression of ideas, with inclusion of new empirical facts and methodological approaches within the general framework of Darwinism leading

to an increasingly more complete understanding of the evolutionary process" (Laubichler & Maienschein 2013, 375).

That narrative's defining character is its depiction of an implicit progression of ideas in evolutionary biology while bracketing ontogeny as a related but independent process. Darwin's main achievement was the explanatory unification, or the consilience of explanations, provided by his theory of evolution by natural selection. Natural selection could explain a vast set of observations recorded by Darwin of the distribution and diversity of species, the adaptations of organisms to their environments, and the numerous homologies found in the structures of different species. The explanatory framework of Darwin's theory is simple and elegant: (1) there is variation between organisms; (2) those variations are heritable across generations; (3) resources are limited, so organisms compete for them; and (4) those variants more successful at competing for resources leave more offspring. Though Darwin's theory explained many observations of species' characters, it did not provide a theory of inheritance or an account of variation.[20] In the decades following Darwin's theory of natural selection, many different theories of inheritance were proposed and rejected. Yet, during this time, there was a somewhat unified view of development, inheritance, and evolution.

August Weismann's research into the material continuity between generations that led him to posit the separation between the soma and germ cells is often interpreted

[20] In his *Variation of Animals and Plants under Domestication*, published in 1868, Darwin proposed a hypothesis to explain inheritance, which he called pangenesis. Darwin suggested that tiny particles, called pangenes, located around the reproductive organs, were transmitted from parent to offspring through the egg or the sperm. These particles carried the characteristics of the parents that had been altered or acquired during the parents' lifetime as a result of the external conditions impinging on the reproductive organs (Young 2007, 155-156). Darwin's cousin, Francis Galton, later refuted Darwin's theory of pangenesis.

as the next big step in the progressivist story. Although Weismann began his research partial to a theory of epigenesis, which understood development to be the gradual emergence of complex form from zygote to embryo to adult form, he settled on a theoretical framework that proposed a certain deterministic character of the germ cells, separated from the differentiated soma lineages (Churchill 1987; 2015). Weismann's initial theoretical separation led to the study of development, inheritance, and evolution as separate experimental and conceptual problems (Laubichler & Maienschein 2007). The study of inheritance became focused on tracking the discrete factors of heredity, exemplified in the breeding experiments of William Bateson and Hugo de Vries, and the re-discovery of Mendel's work in 1900. Thomas Hunt Morgan's research on *Drosophila* established both the transmission rules of these hereditary factors and the location of these factors on the chromosomes. Thus, within a progressive type of narrative, the chromosomal theory of heredity can be interpreted as the culmination of the separation of the problem of inheritance from the problem of development:

> Heredity was now a problem of transmission rules; genes, still identified by their phenotypic effects, were localized on chromosomes; and complications that arose due to development (the genotype-phenotype mapping problem) were soon hidden behind conceptual innovations designed to insulate the core assumptions of transmission genetics from all potential threats to the theory. Concepts such as 'penetrance' and 'expressivity' allowed researchers to maintain a simple model of genetic determinism, while paying lip-service to the intricate process of development (Laubichler & Maienschein 2007, 15).

The standard narrative continues with the Modern Synthesis of Darwinian evolution and Mendelian transmission genetics in the following decades, and eventually, to the mathematical and statistical representations of evolution in population genetics (Mayr 1982). Population geneticists, for example, estimate changes in the frequencies of alleles in populations. They do this, in turn, by estimating parameters such as fitness, mutation rate, migration rate, and effective population size and writing some general equation for changes in allele frequencies given these population parameters. The apparent explanatory success and tractability of these models—despite their abstract and idealized nature—led to a constricted definition of evolution as allele changes across generations, and established certain "forces," such as selection, mutation, migration (or gene flow), and drift, as estimable contributors of evolutionary change.

As a consequence, the metaphor of developmental biology as a black box emerged, relegating development to a particular place in the framework of evolutionary biology. According to some historians and philosophers, the parting of developmental biology from evolutionary research was also the result of technical and experimental challenges within the life sciences during the first half of the twentieth century. In his work on the concept of the gene, Michel Morange argues that because many of the mechanisms of development, including the role of genes in development, were beyond experimental reach at that time, it was to be expected that biologists came to focus on the consequences of evolution and the evolutionary dynamics of populations, rather than attempt to account for the entire complexity of both organisms and populations, and their evolution. Reflecting on this period in the history of biology, he comments:

> The progress of knowledge is often, if not always, the result of a renunciation. Seeking to know is to choose amongst the complexity of the real world what we want to know, and thus accepting to not know about the rest; everything that is neglected, forgotten. A choice that is unconscious more than conscious, social more than individual, and that only a study of the context could justify (Morange 1998, 23; my translation).[21]

Philosopher Ron Amundson presents a somewhat similar interpretation, by drawing a distinction between realism and phenomenalism in scientific methodology (Amundson 2008, 254-255). The main difference between a realist and a phenomenalist lies in the inferences they make from a set of observations. A realist will infer that the unobservable, or theoretical, entities in a scientific theory really exist. Thus, a realist will be more likely to think that investigation into the nature of these entities, and to any other unobserved processes related to them, is legitimate and even necessary. A phenomenalist, alternatively, will be content with accounting for the patterns and variations in the observed phenomena. The aim of their investigations is restricted to discovering laws, or fashioning general rules, that can explain and predict the phenomena. Amundson understands Morgan's contribution to the history of modern evolutionary biology through this distinction. Because most of the chemical and physical causes of development remained a mystery at the beginning of the 1900s, according to Amundson, Morgan successfully split heredity from embryology by adopting a phenomenalist methodology

[21] "Le progrès de la connaissance est souvent, sinon toujours, le fruit d'un renoncement. Chercher à connaître, c'est choisir, parmi la complexité du reel, ce que l'on veut connaître, et donc accepter de ne pas connaître le reste, tout ce qui sera négligé, oublié. Choix inconscient plus que conscient, social plus qu'individuel, que seul l'étude du contexte permet de justifier" (Morange 1998, 23).

regarding the problem of inheritance. His elaboration of transmission genetics successfully replaced the need for a causal understanding of the role of genes in development with an attempt to represent the correlations between parental traits and offspring traits. As Amundson notes: "Population genetics was based on transmission genetics, which was *defined* in terms of the Mendelian patterns of correlation of phenotypic traits between generations. Embryological development had been black-boxed by transmission genetics. When transmission genetics was incorporated into the Evolutionary Synthesis, development remained in its black box" (Amundson 2008, 256).

The re-telling of the succession of select episodes overlooks many contributions to the life sciences at the turn of the century. However, the popular narrative of evolutionary biology captures how the perceived separation of developmental biology from evolutionary biology, for some, came to represent much more than a merely historical or sociological accident in the history of evolutionary biology. Rather, it came to be viewed as an epistemological separation in which evolutionary biology could develop into a mature and empirically successful science. That divergence has implications for the way in which the origins of evolutionary developmental biology came to be represented, as is made evident by certain lines of criticisms against the Modern Synthesis paradigm that emerged in the 1970s. Here, I address one of these critiques and how it has contributed to generating certain problems in the philosophy of biology.

Stephen Jay Gould and Richard Lewontin's (1979) critique of the adaptationist program had significant uptake in the philosophy of biology. Gould and Lewontin's central concern with what they considered as the dominant adaptationist framework in

evolutionary biology at the time was its potential for generating badly designed hypotheses and insufficiently verified claims, and, perhaps most importantly neglecting the potential contributions of other factors in explaining evolution change. Although there are valid criticisms of some of Gould and Lewontin's argument structure and style,[22] their arguments were thought to have provided a "useful corrective to naïve adaptationist assumptions" and inaugurated a shift towards a clearer methodological awareness in evolutionary biology (Okasha 2006). Adaptationists, according to Gould and Lewontin, often conceptualized individual organisms as conglomerates of disparate parts, abstracting away from the development and correlated growth of organisms, and the environment as an unchanging, pre-existent condition in which organisms are immersed. This oversimplified externalist understanding of the process of evolution by natural selection represents a static selecting environment acting upon passive individual organisms with separately optimizable parts.[23] Gould and Lewontin's critical reflections launched a series of attempts to defend or correct the adaptationist strategy in evolutionary biology. These responses often questioned the primacy of natural selection in evolution, which in turn, elicited reflections on the importance of developmental constraints.

[22] Some responses to Gould and Lewontin's critique have been to point out that working biologists are much more sophisticated and rigorous in their methods and experimental designs when testing specific adaptationist hypotheses that Gould and Lewontin let on (Maynard-Smith 1978; Harvey & Pagel 1991.)

[23] This point has led to a re-conceptualization of organism-environment interactions – including the concepts of biological levers, plasticity, and niche construction, for example – with a focus on the life cycles of organisms. C.f. Barker (2008); Glymour (2011); Godfrey-Smith (1996); Laland, Odling-Smee & Feldman (2001); Pigliucci (2001); West-Eberhard (2003).

Philosophers of biology started to quarrel over whether the notion of developmental constraints can have both negative and positive connotations and what theoretical role the concept could play in evolutionary biology. At first, developmental constraints took on a negative connotation, as constraining adaptations. Developmental constraints, in this sense, were seen as factors that limit or impede adaptive perfection; in other words, as design constraints. Those with stakes in this debate argued over whether adaptationists could properly account for developmental constraints in their use of optimality models to test adaptationist hypotheses. Steven H. Orzack and Elliott Sober, for example, defended the use of these models and argued that constraints are discoverable and explainable only by way of the adaptationist method (Orzack & Sober 1994; 1996). Development, under this view, was treated as a background condition and developmental constraints were thought to be passive by-products of ontogeny, not contributing forces—like migration, drift, and selection—to evolutionary dynamics.

Alternatively, other philosophers defended a notion of developmental constraints as a positive influence, preventing certain evolutionary trajectories and influencing other towards a certain range of possible outcomes (Amundson 1994, 2008; Gould 2002; Orzack & Forber 2011). Under this view of constraints, ontogenetic processes were interpreted as positive contributors, or factors, to evolutionary change. Some philosophers have presented the conceptual shift from the negative connotation of developmental constraints on adaptation to a positive notion as a reflection of the growth of a new explanatory agenda in evolutionary developmental biology (Brigandt 2015). Philosopher Ingo Brigandt, for example, claims that the positive view of development as

constraining or biasing certain outcomes over others was expressed in the notion of evolvability. Brigandt states:

> [W]hile considerations about development are essential to an account of evolvability, unlike developmental constraints, evolvability is not set in opposition to selection, but, in fact, operates on a *different dimension* than selection… Selection presupposed the availability of phenotypic variation, and therefore evolvability, which means that an account of evolvability need not be in conflict with an evolutionary theory centered on natural selection; instead, a theory of evolvability *completes* evolutionary theory (Brigandt 2015, 309).

The idea of evolvability, according to Brigandt, carved out a theoretical space for developmental biology in the Modern Synthesis paradigm of evolutionary theory. It fit perfectly with the idea that evolutionary developmental biology was to complete a theory of evolution that black-boxed the connection between genotype and phenotype. Thus, this positive construal of developmental constraints and the notion of evolvability were seen as orthogonal, rather than oppositional, to paradigmatic evolutionary explanations. This philosophical narrative of the origins of evo-devo structured subsequent philosophical discussions about the convergence of evolution and development as a problem of synthesis, and sustained philosophical attention to specific sorts of questions.

The focus on developmental constraints and the portrayal of the epistemology of evolutionary developmental biology had consequences for what sorts of conceptual problems were taken up in the philosophy of biology. For example, philosophers debated over the relative significance of developmental factors in evolution and whether development can be understood as a probabilistic contribution to evolutionary outcomes

(e.g. Amundson 1994). Still others argued that these questions can be resolved empirically by using comparative methods to test adaptationist hypotheses versus constraint hypotheses in particular cases (Sansom 2003). In many of these cases, development has been represented as an additional factor that can complement the existing accounts of evolutionary outcomes.

These kinds of philosophical problems emerged out of what Laubichler and Maienschein have called the standard narrative of the convergence between evolution and development. The philosophers cited above seemed to assume that the theoretical bracketing of development was necessary for the continued progress of evolutionary biology during the first half of the twentieth century. Further, they assumed that the field of evolutionary developmental biology progressed in accordance with an agenda focused on constraints, evolvabilty, and, ultimately, on how this research could fill out the genotype-phenotype map.

Devo-evo and the mechanisms of development

Laubichler and Maienschien have questioned the standard narrative recounting the linear progression from Darwin's theory to the latest synthesis of evolutionary developmental biology as the one and only interpretation of recent convergence between evolution and development (Laubichler & Maienschien 2013; Maienschein & Laubichler 2014). They argue that this traditional narrative leaves out many substantial research contributions since Darwin. Moreover, the alternative history that they offer (which I sketch, in part, in this section) presents a different understanding of the current convergence of evolution and development in modern biology; one, they argue, that

explains developmental evolution as a distinct research program from evolutionary developmental biology (Laubichler & Maienschein 2013).

In the aftermath of Darwin's synthesis, and the theoretical separation of the problems of development, inheritance and evolution, some biologists continued to study the processes of differentiation and morphogenesis within an evolutionary context, including Morgan, as well as Theodor Boveri and Alfred Kühn, amongst others. Laubichler considers Boveri's experimental and theoretical work in cell biology as an early representative of this alternative tradition of developmental evolution (Laubichler & Davidson 2008; Laubichler & Maienschein 2013). In his experimental work on the process of fertilization and inheritance, Boveri helped to establish the chromosomal theory of inheritance, later developed by Morgan, and the functional roles of the nucleus and the cytoplasm in the egg during development. He also offered an important conceptual insight about the practice of evolutionary biology in a 1906 speech, "Organisms as Historical Beings," which clearly outlined the different focus of developmental evolution. He argued that to explain the evolution of organisms, one must first understand the constructive mechanisms that generate their forms. Moreover, Boveri claimed that because it was known that these mechanisms of development are controlled by the material found within the nucleus, the constructive processes ought to be studied experimentally via manipulation of the components of the cell.

> Experimentation - that will be the watchword for additional work in our field. But the most important experiment of all will be the attempt to modify organisms before our very eyes. After all, it seems inconceivable to me that we can proceed in a precise manner without devoting ourselves to the task of exposing organisms

> to new conditions where they must perform a function unusual to them or no longer perform a familiar function. In addition, we must then record the resulting responses (Boveri 1906).

According to Laubichler and Maienschein, Boveri's methodological insights "mapped out the research program of experimental developmental evolution" (Laubichler & Maienschein 2013, 379).

The experimental approach of developmental evolution continued in molecular biology and developmental genetics during the latter half of the twentieth century. François Jacob and Jacques Monod were among the first, in 1961, to contribute to understanding the logic of gene regulation and expression, by discovering a process of enzyme induction in *Escherichia coli* (*E. coli*) (Jacob & Monod 1961). While Jacob and Monod were cautious in extrapolating evolutionary consequences from this discovery, others like Roy Britten and Eric Davidson, and later, Allan Wilson and Mary-Claire King, considered whether mutations in regulatory genes, rather than structural genes, were more important for phenotypic evolution.

In 1969, Britten and Davidson published *Gene Regulation for Higher Cells: A Theory*, in which they proposed a theory suggesting how the embryological processes of differentiation were controlled by the coordinated regulation of gene activity. The evolutionary implications of their theory suggested what King and Wilson (1975) later indicated in their comparative work between human beings and chimpanzees, i.e. because there was very little variation in the protein-coding genes in humans and chimpanzees, the major changes observed in phenotypes were more likely due to changes in the regulatory structure of the genome rather than to the addition of new genes. This belief

grew stronger with the discovery of the Hox genes in the 1980s and the surprising conservation of genetic material in evolutionary history.

Developmental geneticists continued to detect deep patterns of conservation in enhancers, transcription factors, and entire signaling pathways in the genetic mechanisms of development (Morange 2011). The development of this research contributed to a distinct research program in which the emphasis was on the genomic control of gene expression during development and how those mechanisms of control mapped on to the patterns of phenotypic evolution.

There are important conceptual differences between developmental evolution (and the research tradition that influenced it) and the research program of evolutionary developmental biology represented by the standard narrative described in the previous section (Laubichler & Maienschein 2013, 380). First, developmental evolution focuses on entire genomes, as integrated regulatory systems, rather than on allele frequencies of particular gene loci. Second, it focuses on the generation of variation, rather than the resulting distribution of types. And, third, it introduces experimental and causal-mechanical thinking into evolution theory. In terms of approaches and strategies to studying evolution, developmental evolution offers a new *manière de faire*. Instead of the manipulation of phenotypic characters to study fitness consequences, the design of selection experiments tracking changes in allele frequencies in different populations, or the introduction of mutagens into different population strains, developmental evolution proceeds by studying the generation of novel variation and the way it is structured by investigating the control of genetic regulation in development. With developments in synthetic experimental evolution, it may also be possible to intervene in genomes (and

gene regulatory networks), via *in silico* and *in vivo* methods, by attempting to re-engineer changes that may have occurred in evolutionary history (Wagner, Chiu, Laubichler 2000; Laubichler & Maienschein 2013). Because of developmental evolution's particular explanatory domain, the concepts of developmental and genetic mechanisms, including gene regulatory networks, play a central theoretical and experimental role. They are likely to be the key to understanding how novel variation is generated.

As with the discussion on the origins of evolutionary developmental biology, the tradition of developmental evolution likewise raises historical and philosophical questions concerning the nature of the convergence between its explanatory domain and the explanatory domain of population and quantitative genetics in evolutionary biology. Some theoretical biologists have addressed questions about whether developmental evolution challenges the structure of evolutionary explanations within the Modern Synthesis paradigm (Laubichler 2010; Laubichler & Renn 2015; Wagner, Chiu, Laubichler 2000). However, another way to think about the history and research program of developmental evolution is one that might eventually lead to a truly integrative model of evolutionary theory, which would bring together two types of genomic changes – i.e. changes in the regulatory systems that control development *and* changes in quantitative traits controlled by single- or multi-loci polymorphisms (Maienschein & Laubichler 2014, 166-167).

While the questions of whether evo-devo presents an extended synthesis of evolutionary theory or whether devo-evo offers a new avenue towards an integrative theory of evolution provide interesting material for historical and philosophical reflection, an analysis of these questions is beyond the scope of this project. Instead, my

intent in this section was to show that developmental evolution has a particular explanatory domain. That domain includes the origin and structure of variation in evolution, and the generation of novel types.

The application of developmental and genetic mechanisms in Devo-Evo research

Developmental evolutionists think that causal-mechanistic explanations of developmental and genetic mechanisms can contribute something different from the explanation of how a phenotypic trait or organismal form tends to result from a given genotype. They can also provide novel insights into patterns in morphological evolution, homology, and evolutionary novelty, by investigating the generation of variation. In this section, I describe recent cases that have done that.

The two following case studies are paradigmatic illustrations of developmental evolutionary studies. The first case addresses the findings of Armin Moczek's lab concerning the origin and diversification of beetle horns in several related species. This case is illustrative of how approaches in developmental evolution can provide new insights on the patterns in morphological evolution. The second case addresses Ehab Abouheif's and Günter Wagner's conceptual work on homology, as well as experimental work from Wagner's lab on homology and the origin of novelty. The second case is illustrative of how developmental evolution can provide explanations about genetic mechanisms that determine homology, or character identity, and produce novelties. In both these cases, genetic networks are used to explore the nature and the origin of heritable variation often assumed in evolutionary genetics, and provide different ways to think about morphological diversification.

Patterns in morphological evolution

There have been several attempts to make sense of the evidence of highly conserved genetic mechanisms in developmental, and how they affect our understanding of the observed patterns in morphological evolution. Before evidence of deeply homologous genetic mechanisms in development, biologists often assumed that many similar morphological structures were the result of independent, convergent evolution. In other words, they assumed that some observed similarities were the outcomes of finding similar solutions to similar ecological problems. The fact that deep homologies exist at the level of the regulatory genetic circuits deployed during development provided the insight that novel structures and adaptations do not evolve from scratch. But, that insight was not entirely new. Biologists, even those who adopted the Modern Synthesis paradigm, acknowledged, to some extent, the interplay between conservation and innovation in evolution – i.e. the process of *bricolage* (tinkering) – in which evolutionary innovations and adaptations result from modifications of previous structures. However, the way in which these conserved developmental pathways could give glimpses into the patterns in evolutionary history provided new possibilities for research in developmental evolution.

Shubin, Tabin, and Carroll (2009) have remarked that these discoveries show that the same elements and tools have been used and re-used to create similar phenotypes that were formerly believed to have completely independent histories. They write: "If the mechanisms behind the formation of diverse organs are ancient and highly conserved, then parallel evolution must be considered a fact of life in the phylogenetic history of animals" (Shubin, Tabin, & Carroll 2009, 822). But, it has not been always clear what is

meant by parallel evolution, nor how biologists distinguish between parallel and convergent evolution when reconstructing phylogenetic histories.

Convergent vs. Parallel Evolution

There are three general patterns in morphological evolution. Convergent evolution occurs when similar morphologies evolved from different or unrelated common ancestors. The resulting forms are sometimes described as analogies or, in cladistics, homoplastic character states. These analogous structures often result when unrelated species occupy similar ecological niches. In contrast, divergent evolution occurs when an ancestor group accumulates differences that result in diverging species, which share similar character traits. In cladistics, these characters are called homologies. The different morphologies of Darwin's finches are a paradigm example of divergent evolution. Similarities in characters that result from convergent evolution are thought to provide evidence of the influence of natural selection, and, to some extent, developmental constraints and the limits of chemical/physical possibilities. On the other hand, similarities that result from divergence are thought to provide evidence of the historical traces of common ancestry.

Parallel evolution, in the general sense of thinking about phenomenal patterns rather than the developmental mechanisms that might be giving rise to them, describes a pattern in which similar characters have distantly-related ancestors, but are found in different clades.[24] However, those concerned to elucidate the mechanistic underpinning of these patterns have given different understandings of parallel evolution because of the

[24] A clade is defined as a group of biological individuals that includes a common ancestor and all of its descendants, both extinct and extant.

evidence of the deep conservation of developmental genetic mechanisms. For example, Abouheif's account of parallelism is "the convergent or independent evolution of similar morphological characters that share a common developmental basis," which emphasizes a distinction between different levels of biological organization (Abouheif 2008, 3). This account might suggest that parallel evolution really depends on which level of biological organization is being compared, and on whether the "common developmental basis" represents truly homologous structures. Consequently, biologists have come to different ways of accounting for observed cases of parallelism (Hall 2012b, 29).

To consider this problem of parallel evolution, I next present research on the developmental genetic pathways involved in the formation of beetle horns. This research indicates the usefulness of studying developmental genetic mechanisms not only for insights into the processes responsible for the diversification of morphological characters, but also for studying the extent to which the same genetic mechanisms can be reused or redeployed in new contexts. Those studies can offer insights into whether similar generative developmental genetic pathways ought to count as truly deep homologies, or as having had separate origins.

Beetle horns & the co-option of developmental genetic pathways

The origin and evolution of beetle horns, weapons used in male combat to gain access to females, presents a felicitous model for investigating whether such anatomical structures arose *de novo* on multiple occasions or whether they formed from pre-existing structures co-opted for new use. Beetle horns are cuticular and tubular projections from the body wall, similar in some respects to other insect appendages like antennas and legs (Shubin, Tabin & Carroll 2009). However, unlike other appendages in insects, beetle

horns typically lack joints, nerves and muscles, and grow from different anatomical sites like the head and the pronotum (the dorsal surface of the first thoracic segment), which generally do not develop outgrowth structures. The majority of the super-family of scarab beetles are hornless, including the family of dung beetles (*Scarabaeidae*). However, there are several genera of this family that have evolved a great diversity of horns, such as the genus *Onthophagus* on which several studies of beetle horn formation have focused. The beetle horns in *Onthophagus* are considered a prime example of evolutionary innovations, as they lack obvious homology to other morphological structures in insects.

Beetle horns are unevenly distributed among several species and within different members of a species, and display great diversity in morphological structure. Prior analyses of horn distribution in beetle species have supported the hypothesis that they have arisen independently. However, recent studies by Moczek and colleagues have shown that beetle horns in the genus *Onthophagus* likely originated from the co-option of an ancient developmental limb-outgrowth program (Moczek & Nagy 2005; Moczek et al. 2006a, 2006b; Moczek 2006). The horns form in the late larval stage of the insect's development, within compact discs of epidermal cells. The cells proliferate and the discs then turn inside out during the pupal molt and grow to their full length. Not only is the developmental pathway of horn formation similar to the development of most body appendages in insects, it also uses the familiar suite of genes in the development of insect limbs, which subdivide the limbs' proximodistal (p/d) axis. Several important genes are expressed during the development of the beetle horn in members of *Onthophagus*, including the Distal-less (*Dll*) gene found in the distal tip of the horn, the Homothorax

(*hth*) gene, and the Extradenticle (*exd*) gene expressed in the proximal area of the horn (Moczek et al. 2006a; Shubin, Tabin & Carroll 2009).

In a comparative study on *Onthophagus taurus* and *Onthophagus nigriventris*, Moczek and Nagy found similar expression patterns of transcription factors, Distal-less (*Dll*) and aristaless (*al*) in the prepupal horn growth in male morphs (Moczek & Nagy 2005). The expression of the suite of genes that determine the proximodistal axis was found in the developmental programs of both species, which develop horns on different locations; namely, the head and the pronotum. These expression patterns were similar to the gene patterning of appendages in other arthropods.

Moczek and Nagy found evidence to confirm their hypothesis that the hornless male morphs and females in these two polyphenic and sexually dimorphic species would have reduced or absent expression domains of the transcription factors (Moczek & Nagy 2005, 177). They found that *O. nigriventris* males expressed *Dll* in the distal part of the developing pronotal horn axis, but the hornless males (or minor male morphs) expressed a smaller domain of the same transcription factor compared to the horned males. In contrast, the hornless females expressed the *Dll* transcription factor in the proximal region of the axis, instead of the distal area of the horn primordia. Similar *Dll* expression patterns were detected in the developing head horns of *O. taurus*. However, there was no detection of the protein in the pronotal horn primordia in *O. taurus*, where the reabsorption of pre-pupal thoracic proto-horns occurs in both males and females, eliminating the initial sexual dimorphism in the thoracic horn at the larval stage (Moczek & Nagy 2005, 181).

The expression patterns of *al*, on the other hand, was found to be similar in the development of pronotal horns in both morphs and sexes of both species, but the protein was absent in the development of head horns in *O. Taurus* males (Moczek & Nagy 2005, 181). Moczek and Nagy concluded that the conservation of *Dll* in the patterning of horn development indicates that horn development might share a similar mechanism of arthropod appendage patterning (Moczek & Nagy 2005, 182). However, the origin of the sexual dimorphism of horns and male horn polyphenism in these species might be due to differences in the timing and exact location of the patterning genes (Moczek & Nagy 2005, 182-183).

Moczek and colleagues investigated further some of the differences in *Dll* expression profiles within the genus *Onthophagus* and within the different morphs of the same species within this genus (Moczek et al. 2006a). The main differences were found in the location and the domain size of *Dll* expression, and these differences were correlated with the degree to which the proto-horns were retained or re-absorbed via programmed cell death (PCD) in the later development of adult beetles. Similar re-absorption of appendage primordia has been recorded in ants, suggesting the developmental program of PCD is also an ancient and highly conserved developmental mechanism in arthropod appendage development. Thus, that evidence suggested that beetle horns might have originated in the co-option or re-deployment of several ancestral developmental pathways. Moreover, because the p/d axis-patterning genes regulating the growth of horns in the two species act at different anatomical sites, the two instances of horns in the different species may also represent two different instances of the co-option or redeployment of similar ancestral developmental genetic pathways.

In a related study, Moczek and colleagues also found evidence that the development of the proto-horns, or horn primordia, on the pronotum during the larval period was functional whether it was reabsorbed or not in later development, as it helped to split the head capsule during the pupal moult (Moczek et al. 2006b). This functional aspect of larval proto-horns offers a potential explanation for why the developmental capacity to form proto-horns has been maintained in many hornless members of the genus, despite the "loss" or "gains" of the anatomical structure in adults of the genus. The function of the proto-horns at the larval stage suggests that some beetle horns may be an exaptation.

The conclusions supported by Moczek and colleagues' analyses of the development genetic pathways in horn formation at different anatomical sites are different than what have been inferred from other phylogenetic analyses of the genus *Onthophagus*, which have previously proposed three independent origins for horn formation (Moczek *et al.* 2006a; 2006b). However, the authors point out, these previous phylogenetic analyses were based on adult morphologies and ignored the development of pre-pupal pronotal horn development and the later re-absorption of these outgrowths during the pupal stage. The incongruence between the phylogenetic studies based on morphological data and Moczek *et al.*'s comparative studies of the developmental mechanisms involved in horn formation signal an important role for the study of the regulation of gene mechanisms during development in evolutionary studies.

Between homology & homoplasy

Given evidence of the deep conservation of developmental genetic pathways and their co-option throughout a lineage's evolutionary history, there remains a challenge of

defining homology at the level of developmental genetic mechanisms, and relatedly, of identifying patterns of evolution as parallelism or convergence. Moczek's group's experimental and comparative studies of the developmental genetic mechanisms of the genus *Onthophagus* raise questions about the role of conserved genetic pathways in both the origin of evolutionary novelties and the processes of phenotypic evolution. The results of these studies have shown several developmental mechanisms involved in the development of horns, including the function of proximodistal (p/d) axis patterning genes during prepupal growth, involving the *Dll* transcription factor, and the regulation of pupal re-modeling and reabsorption of horns via the activation of programmed cell death (PDC).

Evidence of the first pathway supports the hypothesis that beetle horn formation in the prepupal developmental stage involves the co-option of some p/d patterning genes similarly involved in arthropod appendage formation. The results indicate that these new morphological structures do not require new genes or new developmental pathways. Instead, pre-existing developmental pathways were most likely recruited into new contexts. But, the results also indicate that beetle horns have an "unexpected degree of evolutionary lability," suggesting that both anatomically different, as well as (morphologically) similar, horns may be regulated differently, and have distinct evolutionary origins and histories (Moczek 2009, 145-147). The re-occurrence of these conserved transcription factors at different locations and times during horn development might indicate that the same developmental pathways were co-opted numerous times in different ways.

The same kinds of implications arise from the discovery of the developmental mechanisms responsible for the regulation of pupal re-modeling and reabsorption of horns in some *Onthophagus* species. During the pupal stage of adult horned phenotypes, the beetles undergo the apolysis of the epidermis (the process of separating the cuticula from the epidermis), the secretion of a new cuticle, and eclosion to the next developmental stage (Moczek 2009, 147). However, in some species, such as *O. taurus*, there is no proliferation of horn primordia during the pupal stage. Instead, there is reabsorption of the horn primordial tissue via PCD. As mentioned above, PCD is an ancient and conserved cellular physiological process found in the cells of all metazoan life during development. Thus, again in this case, experimental results indicate that the pupal reabsorption of horn primordia is based on an ancient developmental mechanism that was recruited into a new developmental context. But, the results also indicate that the exact timing and position of PCD-mediated horn remodeling – hypothesized to be regulated by endocrine factors – as well as the degree to which re-modeling or re-absorption occurs, can differ between morphs, sexes, populations, and species. This variation implies that there are modifier mechanisms regulating these processes (Moczek 2009, 148). The modifications of genetic regulatory mechanisms in horned beetles may represent how species in the genus can diversify quickly within lineages (Moczek *et al.* 2006a).

According to Moczek, the co-option of highly-conserved developmental mechanisms in the formation of horns in *Onthophagus* reveals a certain tinkering pattern in the evolution of development: “The evolution of beetle horns involved the recruitment

of conserved developmental mechanisms into new contexts enriched by novel regulatory interactions acquired by pre-existing regulatory genes" (Moczek 2009, 153).

The discovery of highly conserved mechanisms and the great variation of horn morphology in dimorphic males, sexes, and species present the following challenges (Moczek 2009, 153). First, they bring into question the assumed entrenchment and unmalleability of upstream regulators in evolution, such as the p/d patterning genes, and indicate the possibility that such regulators can acquire new additional functions while maintaining their highly-conserved properties. Second, they indicate that even these highly-conserved upstream regulatory genes and their networks can result in relatively quick adaptive radiation and diversification of morphological structures within species and even populations.

The focus on developmental genetic regulation in the morphogenesis of beetle horns in *Onthophagus* challenges some previous assumptions about evolutionary trees, as well as the process of evolution, with respect to innovation and diversification. These studies might not settle the question of whether seemingly homoplastic character traits that share underlying genetic mechanisms ought to be considered homologous in some sense, but they do bring into question the assumed ubiquity of convergent evolution conceived as a process primarily driven by selective pressures external to organisms. The studies provide the kind of evidence that might be used to shed light on the grey areas between these patterns, and the genetic, developmental, and ecological processes responsible for how novelties can originate and how they can diversify within a lineage.

Homology & Evolutionary Novelty

In the above case, the studies revealed that distinguishing homology from homoplasy and the evolutionary processes responsible for those respective character types is not always clear-cut in biological practice. However, Wagner thinks that the concept of homology in evolutionary biology should not be so vague as to be merely a coherent organizing principle used to describe shared similarities between different phenomena. Rather, the concept of homology should pick out a natural kind (Wagner 2014). A concept of homology, he argues, is important to guide a certain empirical research program that can pick out characters, distinguish character identity from different character states, and explain evolutionary novelty. According to Wagner, the way to do so is to develop a genetic theory of homology (Wagner 2014). Before addressing Wagner's genetic theory of homology in the following section, I first outline some of the challenges of identifying patterns of similarities across several levels of biological organization, and present Abouheif's proposal of a hierarchical concept of homology to address these challenges. Wagner is concerned with some of the same challenges Abouheif outlines. However, Wagner is more explicit about the role that developmental genetic mechanisms play in generating structured variation.

Abouheif's hierarchical concept of homology

Homology is typically applied to similar characters derived from a common ancestor. However, morphological structures are sometimes not found in the immediate common ancestor of a group, but the genetic and development potential to produce the character is retained throughout the taxonomic group and can be traced back to some

shared ancestor. In that case, there is a sort of continuity at the developmental genetic level, even though the character or trait is not continuously expressed in all taxa within a group.

According to Abouheif, attention to hierarchical levels of organization – i.e. genes, gene networks, embryonic origins, morphological structures – can allow for the integration of data from developmental genetics into an explicit comparative method to determine homology at different levels of organization and to reconstruct phylogenies (Abouheif 1997). Abouheif appeals to the debate between Walter Gehring and W.J. Dickinson about the evolution of arthropod and vertebrate eyes to illustrate how focusing on only one level of organization may lead to problematic inferences (Abouheif 1997, 407). He considers the example of the discovery of the similarities between the *eyeless* gene in Drosophila and the *Pax-6* gene in mice and their roles in eye morphogenesis. Gehring and his colleagues had taken the similarities between the two genes as a reason to reconsider whether the compound insect eye and the vertebrate eye really evolved independently. Dickinson, on the contrary, argued that more information was needed about the genes in question to reasonably doubt the independent evolution of the different morphological structures generated by *eyeless* and *Pax-6*. That is, there needed to be evidence that a common ancestor of vertebrates and insects had an orthologous gene that functioned similarly in the morphogenesis of the eye.

Given this example, Abouheif suggests that it is consistent to say that the genes responsible for the morphogenesis of some structure are homologous but the morphological structures generated by such genes themselves are non-homologous. Furthermore, according to Abouheif, the example suggests something important about the

role of regulatory genes that act as transcription factors, like *eyeless* and *Pax-6*, in evolution. He suggests that regulatory genes may be more likely to be co-opted or re-deployed in evolution in such a way as to constrain evolutionary trajectories or reveal evolutionary opportunities in morphological structures that have evolved independently.

Abouheif's hierarchical concept of homology is designed to provide a principled way to make phylogenetic inferences based on evidence from genes, gene expression patterns, developmental genetic pathways, embryonic origins, and morphological structures. But, for Wagner, the fact that developmental genetic mechanisms can be conserved and generate different structures is something that requires explanation in itself. As Wagner insists, "Only when we can understand both the conservation and the variation of the development of homologous characters can we have a chance at successfully integrating developmental biology into evolutionary theory" (Wagner 2014, 38). I now turn to Wagner's proposal of a genetic theory of homology.

Wagner's ChIN concept & the genetic theory of homology

Wagner insists that his genetic theory of homology is not hierarchical because the hierarchical concept of homology cannot distinguish between character identity and character states (Wagner 2014, 420). Wagner appeals to the example of the similarities between a bird wing and a bat wing to illustrate his point. He claims that while the wings are usually considered to be non-homologous "as wings," they are homologous "as forelimbs." His account, Wagner insists, can differentiate between this conflating sense of homology, by proposing that the bat wing and the bird wing are homologous characters (the homologue they share is that of forelimb), but their character states as

functional wings are divergent. He argues that statements about character identity and character states are not "conceptually equivalent homology statements" and that to talk about character states as homologous misses the point. Instead, Wagner introduces the concept of character identity networks, or ChINs, to explain homology as characters with shared identity.

Wagner's model of ChINs is described as "genes and gene regulatory networks that interpret the positional information signal and activate position-specific developmental programs" (Wagner 2014, 97). His account of ChINs involves important genetic characteristics: (1) the information for the developmental fate of a cell is contained within the responding cell itself and not in the inductive signals, (2) the genes involved in the character identity network are different from and control the downstream genes that are responsible for realizing a specific character state, and (3) these networks are often governed by transcription factor protein-protein interactions, in which the members of the network integrate multiple signals and sustain each other's expression. These characteristics are important to the ChIN model because they indicate why the mere continuity of genetic information (genes or gene expression profiles) is not enough to identify or explain the structure of characters, and to differentiate between characters and their states.

Wagner first explains why embryonic induction, or inductive signals more generally, cannot be a part of the causes of character identity. Since Spemann's discovery of embryonic induction, biochemists have tried to discover the molecular or chemical basis for induction. Biologists could induce developmental processes with specific molecules, but they could also succeed with other chemical agents and even pH changes,

in *in vitro* bioassays. This proved to be both a technical and a conceptual difficulty until the rise of developmental genetics in the 1980s. Wagner explains that "the information for the developmental fate of a cell is not contained in the inductive signal itself; rather, it is in the responding cells. It turns out that, during each stage of development, cells have a limited number of possible fates, and the inductive signals simply choose between them. If left unperturbed, most cells have a default developmental pathway that they will pursue, as for example ectodermal cells become skin if not told otherwise" (Wagner 2014, 93). Inductive signals, then, are only permissive, while the limited number of developmental fates is intrinsic to the cell's genetic information. Thus, the genetic information for character identity is found within the cell.

To unpack the second characteristic of ChINs, Wagner offers well-studied examples of arthropod development to illustrate that the genes in ChINs are different from the genes responsible for character state modification. Studies of homeotic genes provide a good empirical test for identifying the networks responsible for the determination of characters, and how to distinguish them from the downstream genes responsible for the morphogenesis of character states. More specifically, studies on *ultra bithorax* (*Ubx*) genes in Drosophila (and other arthropods) provide a felicitous example. *Ubx* codes for a transcription factor protein containing a homeodomain. Wagner explains that there are two possible hypotheses about the role of *Ubx*: it can be either a determinant of a derived character state (e.g. the haltere in the hindwing), or it can be a determinant of character identity, regardless of the state of the character (e.g. a factor necessary for the identity of hind wings). With the ChIN model, this difference becomes biologically meaningful, rather than merely semantic. Wagner cites studies conducted by

Warren *et al.* (1994) and Tomoyasu *et al.* (2005) to explain the evidence that show *Ubx* is one of the determinants of character identity, rather than character state. In the 1994 study, Warren *et al.* hypothesized that if *Ubx* was involved in the determination of the character state of an haltere, then it would not be present in the third thoracic segment of the butterfly *Junonia*, as it has four wings. But, they found that, like Drosophila, it is expressed in the third thoracic segment in the butterfly *Junonia*. In the 2005 study, the researchers knocked down the expression of *Ubx* in the hind wing of the beetle, *Tribolium.* In that species, the hind wing has the character state of a wing blade, whereas the forewing has the state of the elytra (a hard protective cover). When *Ubx* expression was knocked down during the development of the hind wing, it developed into the elytra. Wagner interprets this result as a transformation of characters (or character identity) where the hind wing (a character) took on the identity of the forewing (also a character). For Wagner, these examples demonstrate the genetic developmental basis for distinguishing between character states and character identity. In this case, the character state determining genes have to be different from *Ubx* and act downstream from the character identity determining genes (Wagner 2014, 96).

Third, Wagner suggests that there are important genomic processes other than changes to cis-regulatory elements, such as transcription factor protein-protein interactions and the role of transposable elements in changes to genomic architecture, which provide key mechanisms for the origin of characters. (Wagner & Lynch 2010; Wagner 2014). For instance, Wagner emphasizes protein-protein interactions as essential for character determination because they can integrate multiple signals to form a coherent gene regulatory response. Moreover, these interactions may play an important causal role

in the origin of novel regulatory responses, such as the gene network involved in the origin of endometrial stromal cells (ESCs) in placental mammals.

Endometrial stromal cells in placental mammals work by suppressing the immune reaction (and inflammation) against the invading fetus and the fetal placenta. An essential gene regulatory network, in this case, activates and regulates the expression of prolactin in the cells by forming a transcription factor complex to bind to an alternative promoter located about 6kb from the transcriptional start site (Wagner 2014, 115; Lynch *et al.* 2009; Brayer *et al.* 2011). Wagner and his colleagues have hypothesized that the evolution of a protein-protein interaction between Homeobox A11 (HoxA11) and Forkhead box 01A (FOXO1A) was important for the origin of a novel network responsible for the expression of prolactin in ESCs (Brayer *et al.* 2011; Lynch *et al.* 2009; Nnamani *et al.* 2016). The functional cooperation between the two transcription factors was derived from its initial physical interaction and resulted in a neo-allosteric switch that, with other factors, could specify the upregulation of the expression of prolactin (Nnamani *et al.* 2016). These studies suggest that transcription factor complexes may evolve new functional or regulatory roles with significant effects. They also illustrate the logic of transcription factor complexes, emphasized by Wagner. The formation of transcription factor complexes can integrate multiple signals into one transcriptional complex. In this way, they combine to regulate specific target genes and act like an "AND" gate, such that the response is all-or-nothing rather than graded, as would be the case if transcription factors bound to enhancers or promoters independently and acted additively (Wagner 2014, 117). The all-or-nothing response is essential for networks that are responsible for determining cell identity (or character identity more

generally), and correlatively, for the origin of novel cells/characters. It is this characteristic of the genetic mechanism in question which makes the origin and determination of character identity different from character modification. While mechanisms responsible for mutations in cis-regulatory elements might explain the modification of characters, different mechanisms, like that of protein-protein interactions, as well as changes to genome architecture, such as transposable elements and gene duplication, are more likely responsible for generating novel characters (Wagner & Lynch 2010).

Wagner's ChIN model can be used to identify patterns of continuity (homology) in different lineages and explain certain kinds of changes in morphological evolution. Characters are determined by the ability of certain gene regulatory networks to activate one differential state out of the intrinsic possibilities within the cell. And, novel characters originate when ChINs form new combinatorial transcription factor complexes that yield new intrinsic possibilities within a cell.

From homology to networks

Although Wagner and Abouheif seem to offer different theories or concepts of homology, they both focus on a crucial aspect of the conservation of developmental genetic mechanisms in evolution. Wagner argues that evidence of similar genes or gene expression patterns is not enough to identify homology (although it can be suggestive). We also need to know the developmental role of these genes – i.e. "what the genes are doing and whether they convey developmental individuality" – to identify characters (Wagner 2014, 420).

Similarly, Abouheif accepts the limitations of inferring anything about similar genes or gene expression patterns alone. However, he appreciates the unique role that certain genetic elements, such as transcription factors, play in controlling gene expression and suggests that because these genes perform important regulatory functions, it is not so surprising that they are highly conserved, yet have evolved new roles in new contexts. Returning to the debate between Gehring and Dickinson about the role of *eyeless* and *Pax-6* in eye development (whether similar gene expression patterns can provide a basis for homology), Abouheif brings attention to the distinction between the biochemical function of a gene/protein and the developmental role of regulatory genes (Abouheif 1997, 407). Biochemically, both genes in the example function as general transcription factors to activate downstream genes. However, the developmental role of transcription factors can vary in the development of a complex structure, such as eye morphogenesis. So, a transcription factor gene's biochemical function can be highly conserved while its developmental role can be co-opted to perform new functions in morphogenesis (Abouheif 1997, 407). This distinction leads Abouheif to conclude that the similarities between *eyeless* and *Pax-6* present a likely scenario of developmental opportunity, in which "homologous regulatory developmental genes can be potentially co-opted to function in the origin of new traits through evolutionary time" (Abouheif 1997, 406).

Yet, while Abouheif singles out transcription factors, he skims over exactly how or why their developmental role can be easily co-opted in evolution or how they are "relatively free to vary," apart from stating that they possess an "inherent property" to do so (Abouheif 1997, 407). That is what Wagner's account attempts to do. Wagner's theory aims to explain how transcription factors display the properties Abouheif picks out. At

the foundation of Wagner's thinking is a conceptualization of the regulation of gene expression during development in terms of gene regulatory networks. That is, Wagner's uses gene regulatory networks as a strategy to get a handle on the origin of variation and its structure in evolution.

There are a number of conceptual and methodological issues regarding the study of gene networks, their structures, their properties, and their evolution. The last section of this chapter provides a partial geography of research strategies on networks and propose one way to think about the structure and components of certain types of gene regulatory networks and their evolution. Making explicit the various research strategies used to study gene regulatory networks can provide tools for evaluating research that Wagner and others are conducting.

The value of network strategies in Devo-Evo

In the previous sections of this chapter, I attempted to establish what developmental evolution is and what it is trying to explain by looking at different research traditions in the history of evolutionary biology. In brief, developmental evolution is a framework focused on the generations of novel types, and the origin of structured variation. I next provided cases of research within developmental evolution to illustrate how it has contributed to conceptual problems in evolutionary biology, such as the distinction between convergent and parallel evolution, and the difficult problem of defining and identifying homology. What these cases of research in developmental evolution indicate is that much of this work centers around conceiving of the regulation of gene expression during development as gene regulatory networks.

There are several different research strategies to study gene regulatory networks in biology. One way has been to focus on the structural and mathematical properties of networks abstracted away from biological details. For instance, Stuart Kauffman, Albert-László Barabási, and many others have taken this route (e.g. Kauffman 1992; Barabási & Oltvai 2004). Many scientists are now using computational tools to further investigate topological features of networks and to create generating models, which they then use to hypothesize about possible mechanisms that can make sense of genetic and metabolic data (e.g. Peter, Faure, & Davidson 2012; Berry & Widder 2014). Others have focused on the biological properties of the nodes in these networks – i.e. properties of the molecular or genetic constituents of the cell (e.g. Alonso & Wilkins 2005). Another related strategy looks at certain general types of genetic and molecular mechanisms hypothesized to play a role in the restructuring of gene networks in evolution (Emera & Wagner 2012; Wagner & Lynch 2010). This kind of research has engendered certain epistemological and methodological challenges. I address some of the challenges of identifying new genes and new networks within the strategy focused on the biological properties of networks in the following section. I then propose a conceptual framework to incorporate what I call the relational properties of regulatory networks into a research strategy focused on biological properties.

The challenges of identifying new genes and new networks

Within developmental evolution, there have been several discussions about what types of molecular or genetic properties, and what sorts of molecular and genetic mechanisms, are more likely to be responsible for changes in gene networks. Some scientists have challenged the centrality of modifications to enhancers and promoters (or

cis-regulatory mutations, more generally) in developmental evolution. In the previous section, I mentioned Wagner's insistence that other genomic processes, such as protein-protein interactions, transcription factor evolution, and the effect of transposable elements on genomic architecture, provide promising avenues for research on character determination and origin, apart from looking for changes to enhancers (Wagner 2014; Wagner & Lynch 2010).

Similarly, Claudio R. Alonso and Adam S.Wilkins (2005) have argued that "alternative regulatory levels" (ARLs), comprised of different types of molecular elements involved in transcriptional and translational processing, might be at least as important to changes in networks of gene expression as enhancers, or other cis-regulatory changes. These alternatives include the processing of untranslated regions of transcripts, including UTR-dependent modulation of translation and alternative polyadenylation, the processing of exons in pre-mRNA, such as alternative splicing, as well as other regulatory functions of small, non-coding RNAs (Alonso & Wilkins 2005). This is not to say that changes to enhancers have not been effective in generating new variants in developmental evolution. The oft-cited study of the role of changes to the enhancer of *Pitx1* in pelvic reduction in sticklebacks is a well-documented case of that phenomena (Chan *et al.* 2010). Scientists like Alonso and Wilkins, and Wagner, recognize these cases, yet want to draw attention to other sorts of changes to the genomic architecture, which are likely to produce resources for developmental evolution.

Detecting genetic regulatory mechanisms, new genetic elements and new interactions between genes presents a challenge because of methodological constraints in molecular genetics and genomics. Gene networks can change in the following three ways.

A particular gene can be replaced by a new one that is functionally equivalent to the previous gene within a network. An established network may recruit novel genes into its existing regulatory network. Or, entire gene networks may by co-opted into new contexts to perform entirely different developmental functions. Biologists have tried to establish criteria to distinguish the similarities between gene networks, and by extension, between different kinds of changes to networks.

In a paper later than the one cited above, Abouheif (1999) proposes some criteria to "outline some evolutionary properties of regulatory gene networks, and to establish both similarity and phylogenetic criteria for recognizing whether two genetic networks are homologous" (Abouheif 1999, 208). First, he argues that it is crucial to explicitly define the boundaries of the networks, or sub-networks, being compared. This is a challenging task because networks are complex and often parts of nested hierarchies or continuous with other networks. Consequently, there can be components of networks that are highly conserved and other components that have undergone extensive modification over evolutionary time. Given this, Abouheif argues that biologists ought to choose several sets of boundaries for comparison in order to judge how sensitive their chosen boundaries are to conclusions drawn from phylogenetic comparisons and experimental manipulations.

Second, to identify whether two gene networks have similar regulatory genes, or whether a novel regulatory gene has been substituted in a derived network, similarities between genes must be established. This similarity can be established with sequence comparisons by determining whether the two genes are orthologues (i.e. gene copies produced through speciation), rather than paralogues (i.e. gene copies produced through

duplication events). However, this distinction cannot always be reliably inferred from sequence comparisons alone because sequence information cannot always clearly distinguish between orthologues and paralogues. While it is important to remain cautious in drawing conclusions from sequence comparisons, Abouheif suggests that biologists can reliably establish gene orthologues by reconstructing the history of the gene families being compared (Abouheif 1999, 213). However, duplication events can also occur after speciation, which can complicate straightforward comparisons (Abouheif 1999, 213). In such cases, biologists have to be careful with inferring similarities based on one-to-one comparisons.

In addition to the challenges of defining the boundaries of gene networks and comparing the similarity between regulatory genes of two networks, there is a third challenge. Abouheif explains that in order to establish the similarity of gene interactions, researchers must not only compare the biochemical functions of the interacting genes, but they must also compare their developmental function, and the relative spatial position in which they are expressed (Abouheif 1999, 214).

Recall that Abouheif's primary aim is to identify criteria for comparing gene regulatory networks in order to establish when two networks are homologous, and to point out some of the methodological challenges of meeting these criteria. He does so by emphasizing some of the biological properties of regulatory networks, such as whether genes from different networks are orthologous or paralogous, and what their biochemical and developmental functions are. Because his criteria are meant to guide biologists in establishing similarities between networks, they can be helpful at indicating the ways in

which networks might evolve to perform similar functions, or to perform new functions in different contexts.

Abouheif argues that given his set of criteria, it is reasonable to assume that convergent evolution of gene regulatory networks is much less probable than the evolutionary loss of a network. That is because, in the case of loss, all that is needed is a change that represses the expression of an upstream gene, whereas the independent construction of an entire new networks would require the evolution of numerous gene interactions (Abouheif 1999, 215). The idea here that certain kinds of changes to networks in evolution are more likely than others relies on the principle of parsimony. The principle of parsimony is a guiding principle in evolutionary biology that holds that, ceteris paribus, the best hypothesis will include the fewest evolutionary changes. While Abouheif explicitly appeals to the principle of parsimony, he is also appealing to the relational properties of regulatory systems in addition to the noted biological properties of gene networks. For instance, if some known gene regulatory systems display patterns of particular systems of regulation, such as feedback control, then studying the ways in which control systems can shift or change their regulatory states can provide insights into the evolution of these networks.

In the next section, I propose a conceptual framework that combines the relational properties of regulatory systems with the biological properties of gene networks as a research strategy. Such a framework can perhaps provide a promising research strategy to study the evolution of gene networks and the patterns of novel variation that these systems are likely to generate.

Relational properties of regulatory systems in gene networks: a proposal for a combined strategy

To help illustrate my proposed framework, which takes into consideration the relational properties of regulatory systems, as well as the biological properties of the molecular or genetic components of gene networks, I appeal to the simple feedback model of the *lac* operon, described in chapter one. Although Jacob and Monod's model was based on experimental evidence from a prokaryote model, they nonetheless speculated that their logic of genetic regulation in *E. coli* could be generalized to serve as a model for understanding the processes of genetic regulation in multi-cellular eukaryotes (Jacob & Monod 1961).

The discovery of a negative feedback circuit involved in enzyme synthesis presented an opportunity to look at the way in which both the control process of a simple genetic network operates and how it can be changed in ways that could either maintain its functional integrity or co-opt its causal structure to perform new functions. This opportunity is more perspicuous when one considers the abstract functional elements in the simplest model of negative feedback regulation. As the example of the *lac* system shows, negative regulation works by regulating some variable to ensure that some value or stable state is maintained in the system. Any disturbances to the system must be counteracted to ensure this stability. This kind of regulatory mechanism requires functional elements that can (1) identify the state of the regulated variable, (2) compare that value to the ideal reference state, and (3) cause the appropriate changes so that any differences between the identified state, or value, at any given instance, and the ideal reference, is minimized (Barker 2008, 10).

Philosopher of science Gillian Barker (2008) provides a taxonomy of the functional elements of a negative regulatory mechanism. She categorizes the following five elements in a system of negative regulation:

- A *reference* that fixes the predetermined 'goal' value for the regulated variable (the reference value may itself vary according to conditions outside the control process, but in the simplest case it is constant.)
- An *indicator* whose state is determined by (and which thus 'senses') the present values of the regulated variable.
- A *feedback loop* that communicates the state of the indicator to the comparator.
- A *comparator* that compares the state of the indicator with the reference. If a difference is detected, the comparator sends *an error signal* to the effector.
- An *effector* that is induced by the error signal to modify the regulated variable in such a way as to reduce the error (the difference between reference and indicator values) (Barker 2008, 10).

Here is how I think we can apply these abstract functional elements to the *lac* system of *E. coli*. When the *lac* genes are expressed, the *reference* value would indicate when the system is "switched on," so to speak. This occurs when there is lactose present in the bacterium's environment, which leads to the synthesis of β-gal and the other associated genes in the *lac* system (lactose permease and galactoside transcaetylase). The *indicator* could represent an environmental variable, tracking the level of lactose in the bacteria's environment, or perhaps the gene transcripts or the enzymes in the cell responsible for the metabolism of lactose. The mechanism by which the repressor protein interacts with the operator – in this case, the DNA-binding site – would be the *effector*. The feedback loop

communicates the state of the expression of the *lac* genes and the level of lactose in the environment to the state of the repressor – i.e. the *comparator* – and whether or not it is binding to the operator. Such a system of regulation can track a changing reference condition that allows it to meet the demands of a varying environment. In the case of the operon, that would consist of the varying metabolites and substrates present in the bacterium's environment that requires the synthesis of particular enzymes at particular times.

This kind of simple negative feedback system has critical points of causal action (e.g. the effector or the comparator) that can be exploited to modify the network's state or to modify the entire system of regulation. Recall from the descriptions of Jacob and Monod's experiments that the repressor protein can modify the system's state by repressing the expression of its genes or not, given the presence or absence of metabolites. Additionally, removing the repressor from the system completely also results in a constitutive phenotype wherein the enzymes are continuously synthesized. So, the comparator can be tinkered with or re-directed in different ways to modify the reference variable.

In our current example, the comparator is a regulatory gene or, more specifically, a repressor. It should come as no surprise that most comparators in simple sorts of feedback networks will be regulatory genes, like the *lac* repressor. Similarly, there may be other kinds of genetic and molecular components that are more likely to possess the biological properties required to play the role of effector in other kinds of networks, or sub-networks, found in genomes, such as enhancers for example. Thinking about gene regulatory networks as systems of regulation with particular functional elements, can

provide a heuristic or framework with which to categorize certain types of molecular genetic components of networks as having particular kinds of causal roles. Therefore, it might be worthwhile to attend to the different systems of control and regulation, and their functional elements, as a way to identify the critical points of causal action that will affect the reference variable within a system of interest.

Moreover, because identifying and analyzing the causal structures of regulatory systems could reveal which elements of gene regulatory networks might act as the comparator or the effector, they could also indicate how some elements are more likely to be available for co-option by other causal networks within a complex system. In other words, genetic and molecular elements that tend to act as comparators or effectors in regulatory control systems might be co-opted to function in other contexts as biological levers. Barker defines a "biological lever" as "a causal structure that transforms a small initial cause into a much larger effect" (Barker 2008, 12). The idea of a biological lever seems to represent well the fact that many transcription factor genes have been conserved in evolution, yet often co-opted or re-deployed in new developmental contexts.

Understanding the relational properties of gene regulatory network in terms of the functional elements of systems of regulation might help to identify and explain some aspects of the evolution of development, such as how a gene regulatory network might be re-deployed in a new context. This framework is thus a modest proposal for a combined research strategy that takes into account both the biological properties of the molecular and genetic constituents of gene networks as well as the relational properties of certain kinds of systems of regulation that occur in biology. This strategy might help to pick out the causal variables in particular gene regulatory networks, but also to identify whether

there are common causal structures in the regulation of gene expression that reliably cause the observable patterns at higher levels of biological organization.

Conclusion

In this chapter, I focused my attention on developmental evolution as a distinct research program within evolutionary biology. I argued that developmental evolution has a specific set of *explananda*, which includes the generation of novel types and the origin and structure of variation. Examples of this research have generated many new conceptual and methodological issues for philosophers and scientists to address. I outlined some of these issues, such as distinguishing between convergent and parallel evolution, as well as defining and identifying homology and novelty. Lastly, I addressed some of the research strategies to study gene regulatory networks in developmental evolution. There are several research strategies that take networks into account in developmental evolution and each of these have contributed to our understanding of genomic architecture and its evolution, yet each, in turn, face certain challenges. I focused on challenges faced by the strategy centered on the biological properties of gene networks and proposed a framework to combine this strategy with one that pays attention to the relational properties of systems of regulation.

Finally, while molecular models of gene expression and gene regulatory networks offer insights into developmental evolution, they are not naively reductionist nor are they akin to previous gene-centered approached to evolutionary biology. While developmental evolution often focuses on the level of the genome, it takes the genome to be a locus of causal interactions rather than the metaphysically prior or efficient cause of the phenotype and its evolution. This research tradition, however, also allows for

consideration of non-genetic causal factors, including those from ecology, developmental plasticity, epigenetics, and social evolution, within its framework (e.g. Abouheif 2014; Gilbert 2009).

CONCLUSION

Summary

This dissertation began with two driving questions and a promise of how a critical evaluation of prevalent historical narratives of the molecular life sciences during the last half century can offer philosophical insights into the process of science and scientific change. The philosophical contributions of this dissertation are therefore reliant on a serious engagement with both the science of molecular biology and its history. I have engaged with three related, but separate, research programs within modern biology. Chapter one dealt with the research on bacterial genetic regulation, especially the problem of enzyme induction, during the zenith of molecular biology in the mid-twentieth century. Chapter two addressed the research program on micro-RNAs in post-genomic gene expression research. Finally, chapter three focused on developmental evolution and its explanatory domain, and analyzed its use of gene regulatory networks in developmental evolution research.

Through my critical evaluation of historical narratives, as well as my investigation of the three research programs in genetic regulation stated above, I have reached the following conclusions about the epistemic practices within gene expression research. First, both past and current research programs engage with a multitude of approaches and research strategies in an iterative and integrative fashion, which cannot be adequately described by the labels of reductionism, anti-reductionism, or emergence. Second, the representations of gene expression as regulatory network interactions present several strategies for researchers to constrain their models and their experiments, and to control, manipulate, or intervene on the phenomena of gene expression. The dissertation

demonstrates the importance of distinguishing between the approaches within a science and its research strategies. A historical and philosophical study of the latter, rather than a reliance on labels like reductionism and emergence, can contribute to a better understanding of the historical and theoretical developments of a science. Additionally, my investigations in each substantial chapter have engendered future avenues for research contributions in the history and philosophy of science.

Future avenues of research

In chapter one, I proposed an alternative conceptual framework to the oft-used dichotomy between mechanistic materialism and holistic materialism, or reductionism and anti-reductionism, to understand and explain the developments in modern biology. I claimed that, for particular episodes in the history of science, a framework based on the notion of tracking processes would be better at picking out instances of conceptual and experimental convergence in science, as well as revealing how new research problems and questions come to the forefront within a research community. I appealed to a paradigmatic case study in the history of science to make a case for this alternative framework. However, this methodology faces the so-called dilemma of case studies within scholarship that attempts to integrate the history and philosophy of science (Burian 2001; Pitt 2001). In short, the dilemma signals, on the one hand, the worry that going from philosophical analysis to historical case studies can reflect a selection bias in presenting historical cases that fit with preferred philosophical positions. On the other hand, beginning with particular historical cases to philosophical theses can lead to hasty generalizations about science and scientific practice as a whole.

Recently, Raphael Scholl and Tim Räz (2016) have addressed this methodological skepticism concerning the dilemma of cases studies in HPS studies. They suggest ways in which both horns of the dilemma can by mitigated by, first, making explicit the criteria by which cases are chosen in order to evaluate the charge of selection bias, and second, bringing attention to the ways in which philosophical theses and historical narratives interact recursively over time to prevent assent to hasty or unwarranted generalizations. The dilemma, they argue, rests on outdated ideas about confirmation and falsification (Scholl & Räz 2016).

One of their proposed criteria for selecting historical cases is choosing "paradigmatic cases" as *concrete* cases of *abstract* philosophical principles or conceptual frameworks because they are typical of some particular aspect of science, in the same way that model organisms might represent something typical about particular organisms. The case of the *lac* operon is a clear-cut example of a paradigmatic episode in the history of molecular biology. Relatedly, Scholl and Räz further suggest that a reliable method to prevent problematic generalizations is to ensure that our philosophical theses continue to be confronted by further historical cases (Scholl & Räz 2016).[25]

Given Scholl and Räz's proposed criteria and methods for integrative work in HPS, one avenue for future research would be to consider the applicability of the framework I proposed in the first chapter based on the notion of tracking processes to additional historical episodes. A study of the development and creation of the CRISPR-

[25] Problematic generalizations may occur both in the sense of thinking that philosophical conclusions inferred from one case are true for all of science *and* thinking that one counter-example disproves a conclusion.

Cas9 complex in bacterial genetics can provide an apt case study with which to test my conceptual framework.

The CRISPR-Cas9 complex enables biologists to target a specific gene by binding and splicing the DNA at specific locations, and then replace or repair the segment by inserting another sequence in its place. Potential applications of this technology span a variety of contexts, including health, agriculture, and ecosystems engineering. As a result, CRISPR technology has already captured the attention of philosophers and ethicists concerned with the ethical issues of this emerging biotechnology. However, this well-documented case also represents an opportunity for historians and philosophers of science to analyze the epistemic practices of tracing certain kinds of biological processes and consider how these practices channeled researchers towards this breakthrough in biotechnology. One might do this by considering the ways in which several research projects into bacterial immune systems and genomic regulation developed and led to the creation of the bioengineered enzyme complex. Historians and philosophers can also use this case to address the epistemological implications of how biotechnological considerations have framed the development of genetic research in contemporary life sciences, as historian and philosopher of science Hans-Jörg Rheinberger has proposed (Rheinberger 2008, 2009).

Further, I claim in the second chapter that contemporary research in biomedical genomics and genetics often accounts for the context-dependency of molecular and genetic constituents, like micro-RNAs, by elucidating how they causally interact with other constituents within cell types. Some philosophers of science, in collaboration with computational biologists, have begun to explore how the tools of causal modelling, based

on an interventionist notion of causality, can provide useful methods and heuristics for causal inference in these sciences (e.g. Danks, Glymour, & Spirtes 2002). Yet, philosophers of science can also contribute to our understanding of the study of complex systems of gene expression by clarifying the metaphysical and epistemological challenges of choosing one set of variables over others when modelling these systems.

The problem of variable choice has been a persistent one in philosophy of science. Nelson Goodman's (1955) famous new riddle of induction outlined this problem. In brief, because the predicates we use to describe properties of entities can be logically equivalent (e.g. green or grue), a piece of evidence can be used to support two different claims or hypotheses under a syntactic view of induction. Goodman suggested that our inductive practices ought to be based on "projectable" predicates; i.e. those predicates or properties that are historically entrenched. While this proposal has been criticized for different reasons, Goodman was right to point out the need to go beyond the purely syntactic structure of predicates in our inductive practices and attend to the content of the predicates as well. Accounting for "projectable predicates" that attend to content can be reinterpreted as a problem of choosing the right variables to represent phenomena in the world, which, in turn, grounds our explanatory practices.

The problem has recently re-emerged in the context of a particular framework of causation and causal explanation. This widely adopted framework is based on James Woodward's theory of causal explanation (Woodward 2003). In brief, Woodward makes use of the interventionist notion of causation in his theory, defined as follows:

> A necessary and sufficient condition for X to be a direct cause of Y with respect to some variable set **V** is that there be a possible intervention on X that will change Y

(or the probability distribution of *Y*) when all other variables in **V** besides *X* and *Y* are held fixed at some value by interventions (Woodward 2003, 55).

Philosopher Laura Franklin-Hall has recently argued that effective constraints on choosing causal variables under Woodward's interventionist notion are lacking for high-level explanations, which may lead to negative consequences for the explanatory practices of the life sciences (Franklin-Hall 2014). In that paper, she argues that while Woodward's theory of causal explanation is typically used to defend explanations in the special sciences, it cannot be used to defend their superiority over physical-level explanations using his proposed criterion of proportionality. Proportionality, sometimes referred to as "causal fit," constrains the choice of causal variables to include those that produce a proportional change in an effect variable given some degree of change in the causal variable. She argues that this criterion is not adequate to defend the sorts of explanations found in the special sciences because it cannot adequately guide choices between explanations with coarse-grain variables from explanations with finer-grain variables. She concludes by urging philosophers to develop and evaluate principles by which variable choice can be better constrained.

In a recent paper, Woodward (2016) presents some criteria for variable choice. For example, he argues that stability, as well as the related criteria of specificity and sparsity, can provide a useful heuristic for choosing between variable sets. Stability signals a causal relationship between two variables that hold for a relatively significant range of background conditions. Stable causal relationships, in this sense, are preferable because they are more generalizable.

A non-specific causal relationship is one in which a causal variable *X* can have many other effects besides the effect of interest *Y*, and likewise, the effect *Y* may have many other causes besides *X*. Such causal relationships are also non-sparse because they tend to produce causal representation with many causally-connected variables. The idea then is that favoring specific and sparse causal variables that can adequately characterize an effect of interest is better than non-specific causal relationships. The criteria of specificity and sparsity, Woodward argues, provides better handles for manipulation to target the effect(s) of interest. For example, chemotherapy is a non-specific causal agent that targets tumor as well as healthy cells, whereas a more specific genetic manipulation might restrict the intervention to target the diseased cells (Woodward 2016, 1071). Woodward also claims that specific causes are more learnable and they reduce the chances of having several equivalent causal structures that are indistinguishable with available evidence.

Given these recent discussions about variable choice, a future project is to evaluate whether Woodward's criteria of stability, specificity, and sparsity actually do provide the benefits he claims in scientific practice. More specifically, the project can consider whether and to what extent Woodward's criteria apply to the causal structures of genomic regulation in RNA biology, and evaluate the extent to which they characterize the explanatory practices of the molecular life sciences.

Furthermore, the historical discussion of the differences between evo-devo and devo-evo in the third chapter raised some questions about whether developmental evolution can be integrated with the population-level framework of evolutionary genetics. Wagner's recent theoretical work in developmental evolution comments on this exact

issue. For Wagner, previous work in evolutionary biology – what he calls the population/functionalist strategy to studying evolution – was limited by its approach to variation. His aim has been to correct this limitation by studying the generation of novel variation in evolution through developmental genetic mechanisms. Wagner suggest that "the realization that complex organisms / systems have unique and historically contingent variational constraints and biases paves the way for a seamless unification of functionalist [traditional] and structuralist [developmental] agendas… (Wagner 2014, 19). How such a "seamless unification," or perhaps, the integration, of these research traditions can be articulated and carried out remains an interesting conceptual problem for scientists, historians, and philosophers.

Wagner's suggestion also invokes interesting philosophical questions about the nature of unification and integration in science and the relationship between these concepts more generally. There has been a lot of recent work on these issues (c.f. Brigandt 2010; Mitchell 2003, 2004; Mitchell & Dietrich 2006; O'Malley & Soyer 2012, Plutynski 2013). For example, Ingo Brigandt has argued that it is the particular research projects, or 'problem agendas,' in the life sciences which permit some degree of local integration between distinct disciplinarian approaches (Brigandt 2010). Others, like William Bechtel, have argued that research projects that begin with a seemingly unified or integrated research agenda can end up being fragmented into new sub-disciplines, which demand specialized technical expertise, instead of providing a lasting integration of disciplines or theories (Bechtel 1993).

A feature of Bechtel's approach is that he examines integration as a process rather than an end goal. Similarly, Anya Plutynski interprets unification and integration in the

life sciences as parts of a process (Plutynski 2013). Using the historical development of the multistage model of cancer as a case study, she argues that integration is a process, which originates from a unifying theory or model. Integration occurs when there is a pressing question or problem which requires multiple sources of evidence from different fields of inquiry. Plutynski draws attention to the fact that unification is often thought to provide simplified or idealized explanations, whereas integration is thought to offer more complete, or comprehensive, explanations. But, she warns that thinking about integration as an attempt to be complete is misleading because integration in science is not singular in type and admits of degrees (Plutynski 2013, 474).

Given these considerations about the nature of integration and unification, a future avenue of research is to test these varying philosophical accounts with respect to research programs described in this dissertation. I suspect that applying the varying concepts of integration and unification in the philosophical literature to these research programs will give similar insights into the process of science and the sorts of interactions between disciplines at play in the life sciences. However, another avenue of research is to focus on the normative aspects of unification and integration as regulative ideals in science. One way to address how these might be considered as scientific virtues is to consider further the research strategies that investigate the general patterns of causal dependencies in the regulation of gene expression. I take these strategies to be motivated in part by particular explanatory goals, which aim to unify some set of phenomena under shared causal patterns.

REFERENCES

Abouheif, Ehab. “Developmental genetics and homology: a hierarchical approach.” *Trends in ecology & evolution* 12, no. 10 (1997): 405-408.

Abouheif, Ehab. “Establishing homology criteria for regulatory gene networks: prospects and challenges.” In *Homology*, edited by Gregory R. Bock and Gail Cardew, 207-221. John Wiley & Sons, 1999.

Abouheif, Ehab. “Parallelism as the pattern and process of mesoevolution.” *Evolution and Development* 10, no. 1 (2008): 3-5.

Abouheif, Ehab, Marie-Julie Favé, Ana Sofia Ibarrarán-Viniegra, Maryna P. Lesoway, Ab Matteen Rafiqi, and Rajendhran Rajakumar. “Eco-evo-devo: The time has come.” In *Ecological genomics: Ecology and the Evolution of Genes and Genomes*, edited by Christian R. Landry and Nadia Aubin-Horth, 107-125. Springer Netherlands, 2014.

Allen, Garland. *Life Science in the Twentieth Century*. Cambridge: Cambridge University Press, 1978.

Allen, Garland. “Morphology and Twentieth-Century Biology: A Response.” *Journal of the History of Biology* 14, no. 1 (1981): 159-176.

Alonso, Claudio R., and Adam S. Wilkins. “The molecular elements that underlie developmental evolution.” *Nature Reviews Genetics* 6, no. 9 (2005): 709-715.

Ambros, Victor, and H. Robert Horvitz. “The *lin-14* locus of *Caenorhabditis elegans* controls the time of expression of specific postembryonic developmental events.” *Genes and Development* 1 (1987): 398-414.

Ambros, Victor. “The Functions of Animal microRNAs.” *Nature* 431 (2004) 350-355.

Amundson, Ron. “Two Concepts of Constraint: Adaptationism and the Challenge from Developmental Biology.” *Philosophy of Science* 61 (1994): 556-578.

Amundson, Ron. *The changing role of the embryo in evolutionary thought: roots of evo-devo*. Cambridge University Press, 2005.

Amundson, Ron. “Development and Evolution.” In *A Companion to the Philosophy of Biology*, edited by Sahotra Sarkar and Anya Plutynski, 248-278. Wiley-Blackwell Publishing, 2008.

Barabási, Albert-László, and Zoltán N. Oltvai. “Network biology: understanding the cell's functional organization.” *Nature reviews genetics* 5, no. 2 (2004): 101-113.

Barker, Gillian. “Biological Levers and Extended Adaptationism” *Biology & Philosophy* 23, no. 1 (2008): 1-25.

Bartel, David P. “MicroRNAs: Genomic, Biogenesis, Mechanism, and Function.” *Cell* 116 (2004): 281-297.

Bartel, David P. “MicroRNAs: Target Recognition and Regulatory Functions.” *Cell* 136, no. 2 (2009): 215-233.

Bechtel, William. “Integrating Sciences by Creating New Disciplines: The Case of Cell Biology,” *Biology and Philosophy* 8 (1993): 277-299.

Bechtel, William, and Robert C. Richardson. *Discovering complexity: Decomposition and Localization as Strategies in Scientific Research.* Princeton, NJ: Princeton University Press, 1993.

Bedau, Mark A., and Paul Ed Humphreys. *Emergence: Contemporary readings in philosophy and science*. MIT press, 2008.

Berry, David, and Stefanie Widder. “Deciphering microbial interactions and detecting keystone species with co-occurrence networks.” *Frontiers in microbiology* 5 (2014): 219.

Boveri, Theodor. “Organisms as Historical Beings.” In *Speech on the Occasion of the 324th Anniversary Celebration of the Royal Julius Maximilian University of Würzburgm*, translated by Manfred D. Laubichler, delivered on May 11, 1906.

Brayer, Kathryn J., Vincent J. Lynch, and Günter P. Wagner. “Evolution of a derived protein–protein interaction between HoxA11 and Foxo1a in mammals caused by changes in intramolecular regulation.” *Proceedings of the National Academy of Sciences* 108, no. 32 (2011): E414-E420.

Brennecke, Julius, David R. Hipfner, Alexander Stark, Robert B. Russell, and Stephen M. Cohen. “*bantam* Encodes a Developmentally Regulated microRNA that Controls Cell Proliferation and Regulates the Proapoptotic Gene *hid* in *Drosophila*.” Cell 113, no. 1 (2003): 25-36.

Brigandt, Ingo. “Beyond Reduction and Pluralism: Toward an Epistemology of Explanatory Integration in Biology.” *Erkenntnis* 73 (2010): 295-311.

Brigandt, Ingo. “From developmental constraint to evolvability: how concepts figure in explanation and disciplinary identity.” In *Conceptual Change in Biology*, edited by Alan C. Love, 305-325. Springer Netherlands, 2015.

Britten, Roy J., and Eric H. Davidson. “Gene regulation for higher cells: a theory.” *Science* 165, no. 3891 (1969): 349-357.

Burian, Richard M. “Technique, task definition, and the transition from genetics to molecular genetics: Aspects of the work on protein synthesis in the laboratories of J. Monod and P. Zamecnik.” *Journal of the History of Biology* 26, no. 3 (1993): 387-407.

Burian, Richard M. “On microRNA and the need for exploratory experimentation in post-genomic molecular biology.” *History and Philosophy of the Life Sciences* 29 (2007): 285-311.

Burian, Richard M. “The dilemma of case studies resolved: The virtues of using case studies in the history and philosophy of science.” *Perspectives on Science* 9, no. 4 (2001): 383-404.

Burian, Richard M., and Jean Gayon. “The French School of Genetics: From Physiological and Population Genetics to Regulatory Molecular Genetics.” *Annual Review of Genetics* 33 (1999): 313-349.

Calin, George A., and Carlo M. Croce. “MicroRNA signatures in human cancers.” *Nature Reviews Cancer* 6, no. 11 (2006): 857-866.

Cannon, Walter. B. “Organization for physiological homeostasis.” *Physiological reviews* 9, no. 3 (1929): 399-431.

Chan, Yingguang Frank, Melissa E. Marks, Felicity C. Jones, Guadalupe Villarreal, Michael D. Shapiro, Shannon D. Brady, Audrey M. Southwick, *et al.* “Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer.” *Science* 327, no. 5963 (2010): 302-305.

Churchill, Frederick B. “From heredity theory to Vererbung: The transmission problem, 1850-1915.” *Isis* 78, no. 3 (1987): 337-364.

Churchill, Frederick B. *August Weismann*. Harvard University Press, 2015.

Cobb, Matthew. *Life's greatest secret: the race to crack the genetic code*. Basic Books, 2015.

Cohn, Melvin, Jacques Monod, Martin R. Pollock, Sol Spiegelman, and Roger Y. Stanier. “Terminology of enzyme formation.” *Nature* 172 (1953): 1096-1096.

Cohn. Melvin. “In Memorium.” In *The Operon*, edited by Jeffrey H. Miller and William S. Reznikoff, 1-9. New York: Cold Spring Harbor Laboratory, 1980.

Cox, Adrienne D., and Channing J. Der. “Ras history: The saga continues.” *Small GTPases* 1, no. 1 (2010): 2-27.

Craver, Carl F., and Lindley Darden. *In Search of Mechanisms: Discoveries Across the Life Sciences* University of Chicago Press, 2013.

Danks, David, Clark Glymour, and Peter Spirtes. "Confirmation versus Search in Gene Regulation: The Complexity of Gene Perturbation Experimentation as a Search Method." IJCAI Bioinformatics Workshop, 2002.

Davidson, Eric H., and Douglas H. Erwin. "Gene regulatory networks and the evolution of animal body plans." *Science* 311, no. 5762 (2006): 796-800.

Delehanty, Megan. "Emergent properties and the context objection to reduction." *Biology & Philosophy* 20, no. 4 (2005): 715-734.

Dupré, John. "It is not possible to reduce biological explanations to explanations in chemistry and/or physics." In *Contemporary debates in philosophy of biology,* edited by Francisco J. Ayala and Robert Arp, 32-47. Wiley-Blackwell, 2010.

Dupré, John. *Processes of life: essays in the philosophy of biology*. Oxford University Press, 2012.

Elliott, David, and Michael Ladomery. *Molecular Biology of RNA*. Oxford: Oxford University Press, 2011.

Emera, Deena, and Günter P. Wagner. "Transposable element recruitments in the mammalian placenta: impacts and mechanisms." *Briefings in functional genomics* (2012): els013.

Erwin, Douglas H., and Eric H. Davidson. "The evolution of hierarchical gene regulatory networks." *Nature Reviews Genetics* 10, no. 2 (2009): 141-148.

Esquela-Kerscher, Aurora, and Frank J. Slack. "Oncomirs—microRNAs with a role in cancer." *Nature Reviews Cancer* 6, no. 4 (2006): 259-269.

Fire, Andrew, and SiQun Xu, Mary K. Montgomery, Steven A. Kostas, Samuel E. Driver, and Craig C. Mello. "Potent and Specific Genetic Interference in Double-Stranded RNA in *Caenorhabditis elegans*." *Nature* 391 (1998): 806-811.

Fodor, Jerry A. "Special sciences (or: the disunity of science as a working hypothesis)." *Synthese* 28, no. 2 (1974): 97-115.

Franklin-Hall, Laura R. "High-Level Explanation and the Interventionist's 'Variables Problem'." *The British Journal for the Philosophy of Science* (2014): 1-25.

Garzon, Ramiro, Muller Fabbri, Amelia Cimmino, George A. Calin, and Carlo M. Croce. "MicroRNA expression and function in cancer." *Trends in molecular medicine* 12, no. 12 (2006): 580-587.

Gilbert, Scott F. "Ecological Development Biology." *eLS* (2009): 1-8.

Gilbert, Walter, and Benno Müller-Hill. "Isolation of the lac repressor." *Proceedings of the National Academy of Sciences of the United States of America* 56, no. 6 (1966): 1891-1898.

Glymour, Bruce. "Modeling environments: Interactive causation and adaptations to environmental conditions." *Philosophy of Science* 78, no. 3 (2011): 448-471.

Godfrey-Smith, Peter. *Complexity and the Function of Mind in Nature*. Cambridge University Press, 1998.

Goodman, Nelson. *Fact, fiction and forecast*. Vol. 74. Cambridge, MA: Harvard University Press, 1955.

Gould, Stephen Jay, and Richard C. Lewontin. "The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme." *Proceedings of the Royal Society of London B: Biological Sciences* 205, no. 1161 (1979): 581-598.

Gould, Stephen Jay. *The structure of evolutionary theory*. Harvard University Press, 2002.

Griesemer, James. "Theoretical integration, cooperation, and theories as tracking devices." *Biological Theory* 1, no. 1 (2006): 4-7.

Griesemer, James. "Tracking Organic Processes: Representations and Research Styles in Classical Embryology and Genetics." In *From Embryology to Evo-Devo: A History of Developmental Evolution*, edited by Manfred D. Laubichler and Jane Maienschein, 375-433. Cambridge & London: The MIT Press, 2007.

Griffiths, Paul, and Karola Stotz. *Genetics and Philosophy: An Introduction*. Cambridge University Press, 2013.

Grmek, Mirko D., and Bernardino Fantini. "Le rôle du hasard dans la naissance du modèle de l'opéron." *Revue d'histoire des sciences* 35 (1982): 193-215.

Gros, François. *Les secrets du gène*. Odile Jacob, 1986.

Hall, Brian K. "Evo-Devo or Devo-Evo: Does it Matter?" *Evolution and Development* 2 (2000): 177-178.

Hall, Brian K. “Evolutionary developmental biology (Evo-Devo): Past, present, and future.” *Evolution: Education and outreach* 5, no. 2 (2012a): 184-193.

Hall, Brian K. “Parallelism, deep homology, and evo-devo.” *Evolution & development* 14, no. 1 (2012b): 29-33.

Harvey, Paul H., and Mark D. Pagel. *The comparative method in evolutionary biology*. Vol. 239. Oxford: Oxford university press, 1991.

He, Lin, and Gregory J. Hannon. “MicroRNAs: small RNAs with a big role in gene regulation.” *Nature Reviews Genetics* 5, no. 7 (2004): 522-531.

Henderson, Lawrence J. *Blood: a study in general physiology*. New Haven: Yale University Press, 1928.

Hipfner, David R., Katrin Weigmann, and Stephen M. Cohen. “The bantam gene regulates Drosophila growth.” *Genetics* 161, no. 4 (2002): 1527-1537.

Hull, David L. “Reduction in genetics.” *Philosophy of Science* 46, no. 2 (1979): 316-320.

Jacob, François, and Jacques Monod. “Genetic Regulatory Mechanisms in the Synthesis of Proteins.” *Journal of Molecular Biology* 3 (1961): 318–356.

Jacob, François, and Élie Léo Wollman. *La sexualité des bactéries*. Librairies de L'Académie de Médecine, 1959.

Jacob, François. *La Logique du vivant, une histoire de l’hérédité.* Paris: Gallimard, 1970.

Jacob, François. “Genetics of the bacterial cell. *Nobel Lecture, December 11, 1965*.” In *Nobel Lectures, Physiology or Medicine 1963-1970*, Elsevier Publishing Company, Amsterdam, 1972.

Johnson, Steven M., Helge Grosshans, Jaclyn Shingara, Mike Byrom, Rich Jarvis, Angie Cheng, Emmanuel Labourier, Kristy L. Reinert, David Brown, and Frank J. Slack. “*RAS* is regulated by the *let-7* microRNA family.” *Cell* 120, no. 5 (2005): 635-647.

Judson, Horace F. *The eight days of creation*. New York: Simon and Schuster, 1979.

Kauffman, Stuart A. “Origins of order in evolution: Self-organization and selection.” In *Understanding Origins*, edited by Jean-Pierre Dupuy, and Francisco J. Varela, 153-181. Springer Netherlands, 1992.

Kay, Lily E. *Who wrote the book of life?: A history of the genetic code.* Stanford University Press, 2000.

Keller, Evelyn Fox. *The century of the gene*. Harvard University Press, 2000.

Keller, Evelyn Fox. "The century beyond the gene." *Journal of Biosciences* 30, no. 1 (2005): 3-10.

Keller, Evelyn Fox. "It is possible to reduce biological explanations to explanations in chemistry and/or physics." In *Contemporary debates in philosophy of biology,* edited by Francisco J. Ayala and Robert Arp, 19-31. Wiley-Blackwell, 2010.

Keller, Evelyn Fox "The Postgenomic Genome." In *Postgenomics: Perspectives on Biology After the Genome*, edited by Sarah S. Richardson and Hallam Stevens, 9-31. Duke University Press, 2015.

King, Mary-Claire, and Allan C. Wilson. "Evolution at two levels in humans and chimpanzees." *Science* 188, no. 4184 (1975): 107-116.

Kitcher, Philip. "1953 and all that: A tale of two sciences." *The Philosophical Review* 93, no. 3 (1984): 335-373.

Kornblihtt, Alberto R., Ignacio E. Schor, Mariano Alló, Gwendal Dujardin, Ezequiel Petrillo, and Manuel J. Muñoz. "Alternative splicing: a pivotal step between eukaryotic transcription and translation." *Nature Reviews Molecular Cell Biology* 14, no. 3 (2013): 153-166.

Kuhn, Donald E., Mickey M. Martin, David S. Feldman, Alvin V. Terry, Gerard J. Nuovo, and Terry S. Elton. "Experimental validation of miRNA targets." *Methods* 44 (2008): 47-54.

Ladyman, James, and Don Ross. *Everything must go: Metaphysics naturalized*. Oxford: Oxford University Press, 2009.

Laland, Kevin N., F. John Odling-Smee, and Marcus W. Feldman. "Niche construction, ecological inheritance, and cycles of contingency in evolution." In *Cycles of contingency: Developmental systems and evolution*, edited by Susan Oyama, Paul E. Griffiths, and Russell D. Gray, 117-126. MIT Press, 2001.

Laland, Kevin N., Tobias Uller, Marcus W. Feldman, Kim Sterelny, Gerd B. Müller, Armin Moczek, Eva Jablonka, and John Odling-Smee. "The extended evolutionary synthesis: its structure, assumptions and predictions." *Proceedings of the Royal Society B* 282, no. 1813 (2015): 1-14.

Laubichler, Manfred D. "Does History Recapitulate Itself? Epistemological Reflections on the Origins of Evolutionary Developmental Biology?" In *From Embryology to Evo-Devo: A History of Developmental Evolution*, edited by Manfred D. Laubichler and Jane Maienschein, 13-33. MIT Press, 2007.

Laubichler, Manfred D. “Evolutionary Developmental Biology offers a Significant Challenge to the Neo-Darwin Paradigm.” In *Contemporary Debates in Philosophy of Biology*, edited by Francisco J. Ayala and Robert Arp, 199-212. Wiley Blackwell, 2010.

Laubichler, Manfred D., and Eric H. Davidson. “Boveri's long experiment: Sea urchin merogones and the establishment of the role of nuclear chromosomes in development.” *Developmental biology* 314, no. 1 (2008): 1-11.

Laubichler, Manfred D., and Jane Maienschein, eds. *From Embryology to Evo-Devo: A History of Developmental Evolution*. MIT Press, 2007.

Laubichler, Manfred D., and Jane Maienschein, eds. *Form and Function in Developmental Evolution*. Cambridge University Press, 2009.

Laubichler, Manfred D., and Jane Maienschein. “Developmental evolution.” In *The Cambridge encyclopedia of Darwin and evolutionary thought*, edited by Michael Ruse, 375-382. Cambridge University Press, 2013.

Laubichler, Manfred D., and Jürgen Renn. “Extended evolution: A conceptual framework for integrating regulatory networks and niche construction.” *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 324, no. 7 (2015): 565-577.

Lee, Rosalind C., Rhonda L. Feinbaum, and Victor Ambros. “The C. elegans heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*.” Cell 75, no. 5 (1993): 843-854.

Lynch, Vincent J., Kathryn Brayer, Birgit Gellersen, and Günter P. Wagner. “HoxA-11 and FOXO1A cooperate to regulate decidual prolactin expression: towards inferring the core transcriptional regulators of decidual genes.” *PLoS One* 4, no. 9 (2009): e6845.

Machamer, Peter, Lindley Darden, and Carl F. Craver. “Thinking about mechanisms.” *Philosophy of Science* (2000): 1-25.

Maderspacher, Florian. "Lysenko rising." *Current Biology* 20, no. 19 (2010): R835-R837.

Maienschein, Jane, Ronald Rainger, and Keith R. Benson. “Introduction: Were American morphologists in revolt?” *Journal of the History of Biology* 14, no. 1 (1981): 83-87.

Maienschein, Jane, and Manfred D. Laubichler. “Exploring Development and Evolution on the Tangled Bank.” In *Evolutionary Biology: Conceptual, Ethical, and Religious Issues*, edited by Paul Thompson and Denis Walsh, 151-171. Cambridge University Press, 2014.

Malumbres, Marcos, and Mariano Barbacid. “RAS oncogenes: the first 30 years.” *Nature Reviews Cancer* 3, no. 6 (2003): 459-465.

Mattick, John S. “RNA regulation: a new genetics?” *Nature Reviews Genetics* 5, no. 4 (2004): 316-323.

Mattick, John S. “Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms.” *Bioessays* 25, no. 10 (2003): 930-939.

Maynard-Smith, John. “Optimization theory in evolution.” *Annual Review of Ecology and Systematics* 9 (1978): 31-56.

Mayr, Christine, and David P. Bartel. “Widespread shortening of 3′ UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells.” *Cell* 138, no. 4 (2009): 673-684.

Mitchell, Sandra D. *Biological Complexity and Integrative Pluralism*. Chicago: University of Chicago Press, 2003.

Mitchell, Sandra D. “Why Integrative Pluralism?” *Emergence, Complexity and Organization* 6 (2004): 81-91.

Mitchell, Sandra D., and Michael R. Dietrich. “Integration without unification: An argument for pluralism in the biological sciences.” *The American Naturalist* 168, no. S6 (2006): S73-S79.

Moczek, Armin P., and Lisa M. Nagy. “Diverse developmental mechanisms contribute to different levels of diversity in horned beetles.” *Evolution & Development* 7, no. 3 (2005): 175-185.

Moczek, Armin P., Debra Rose, William Sewell, and Bethany R. Kesselring. “Conservation, innovation, and the evolution of horned beetle diversity.” *Development Genes and Evolution* 216, no. 11 (2006a): 655-665.

Moczek, Armin P., Tami E. Cruickshank, and Andrew Shelby. “When ontogeny reveals what phylogeny hides: gain and loss of horns during development and evolution of horned beetles.” *Evolution* 60, no. 11 (2006b): 2329-2341.

Moczek, Armin P. “Integrating micro-and macroevolution of development through the study of horned beetles.” *Heredity* 97, no. 3 (2006): 168-178.

Moczek, Armin P. “The Origin and Diversification of Complex Traits Through Micro- and Macroevolution of Development: Insights from Horned Beetles.” *Current Topics in Developmental Biology* 86 (2009): 135-162.

Monod, Jacques. “From enzymatic adaptation to allosteric transitions.” *Science* 154 (1966): 475-483.

Monod, Jacques, and Melvin Cohn. “La Biosynthèse induite des enzymes (adaptation enzymatique).” *Advances in Enzymology* 13 (1952): 67-119.

Morange, Michel. (1997). “The Transformation of Molecular Biology on Contact with Higher Organisms, 1960-1980: from a Molecular Description to a Molecular Explanation.” *History and Philosophy of the Life Sciences*, 19, no. 3 (1997): 369-393.

Morange, Michel. *La Part des Gènes*. Éditions Odile Jacob, 1998.
Morange, Michel. *A History of Molecular Biology*. Cambridge & London: Harvard University Press, 1998.

Morange, Michel. “Post-genomics, between reduction and emergence.” *Synthèse* 151 (2006): 355-360.

Morange, Michel. “Evolutionary Developmental Biology: Its Roots and Characteristics.” *Developmental Biology* 357, no. 1 (2011): 13-16.

Nagel, Ernest. *The structure of science: Problems in the logic of scientific explanation*. New York: Harcourt, Brace & World, 1961.

Nnamani, Mauris C., Soumya Ganguly, Eric M. Erkenbrack, Vincent J. Lynch, Laura S. Mizoue, Yingchun Tong, Heather L. Darling, Monika Fuxreiter, Jens Meiler, and Günter P. Wagner. “A Derived Allosteric Switch Underlies the Evolution of Conditional Cooperativity between HOXA11 and FOXO1.” *Cell reports* 15, no. 10 (2016): 2097-2108.

O'Connor, Timothy and Wong, Hong Yu, “Emergent Properties.” *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2015/entries/properties-emergent/>.

Okasha, Samir. “Adaptation.” In *Routledge Encyclopedia of Philosophy of Science*, edited by Sahotra Sarkar and Jessican Pfeiffer, 732-737. Routledge, 2006.

Olby, Robert. C. *The path to the double helix: the discovery of DNA*. Courier Corporation, 1974.

O’Malley, Maureen A., Kevin C. Elliott, and Richard M. Burian. “From Genetic to Genomic Regulation: Iterativity in microRNA Research.” *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 41, no. 4 (2010): 407-417.

O’Malley, Maureen A., and Orkun S. Soyer. “The roles of integration in molecular systems biology.” *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 43, no. 1 (2012): 58-68.

Orzack, Steven Hecht, and Elliott Sober. "Optimality models and the test of adaptationism." *American Naturalist* 143, no. 3 (1994): 361-380.

Orzack, Steven Hecht, and Elliot Sober. "How to formulate and test adaptationism." *The American Naturalist* 148, no. 1 (1996): 202-210.

Orzack, Steven Hecht, and Patrick Forber. "Adaptationism." *The Stanford Encyclopedia of Philosophy* (Summer 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2011/entries/adaptationism/>.

Oyama, Susan, Paul E. Griffiths, and Russell D. Gray. *Cycles of contingency: Developmental systems and evolution*. The MIT Press, 2003.

Pardee, Arthur B., François Jacob, and Jacques Monod. "The Genetic Control and Cytoplasmic Expression of 'Inducibility' in the Synthesis of β-galactosidase by *E. coli*." *Journal of Molecular Biology* 1 (1959): 165-178.

Pasquinelli, Amy E., Brenda J. Reinhart, Frank Slack, Mark Q. Martindale, Mitzi I. Kuroda, Betsy Maller, David C. Hayward et al. "Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA." *Nature* 408, no. 6808 (2000): 86-89.

Peluffo, Alexandre E. "The 'Genetic Program': Behind the Genesis of an Influential Metaphor." *Genetics* 200, no. 3 (2015): 685-696.

Peter, Isabelle S., Emmanuel Faure, and Eric H. Davidson. "Predictive computation of genomic logic processing functions in embryonic development." *Proceedings of the National Academy of Sciences* 109, no. 41 (2012): 16434-16442.

Pigliucci, Massimo. *Phenotypic plasticity: beyond nature and nurture*. John Hopkins University Press, 2001.

Pigliucci, Massimo. "Between holism and reductionism: a philosophical primer on emergence." *Biological Journal of the Linnean Society* 112, no. 2 (2014): 261-267.

Pigliucci, Massimo, and Gerd B. Müller. *Evolution: The Extended Synthesis*. MIT Press, 2010.

Pitt, Joseph C. "The dilemma of case studies: Toward a Heraclitian philosophy of science." *Perspectives on Science* 9, no. 4 (2001): 373-382.

Plutynski, Anya. "Cancer and the goals of integration." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 44, no. 4 (2013): 466-476.

Powell, Alexander, and John Dupré. "From Molecules to Systems: The Importance of Looking Both Ways." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 40 (2009): 54-64.

Ptashne, Mark. "Isolation of the λ-phage repressor." *Proceedings of the National Academy of Sciences* 57 (1967a): 306–313.

Ptashne, Mark. "Specific binding of the λ-phage repressor to λ-DNA." *Nature* 214 (1967b): 232-234.

Racine, Valérie. "Lysogenic Bacteria as an Experimental Model at the Pasteur Institute (1915-1965)" *Embryo Project Encyclopedia*, 2014. ISSN: 1940-5030 http://embryo.asu.edu/handle/10776/8223

Racine, Valérie. "The mechanistic-holistic divide revisited: The case of the *lac* operon." *Studies in the History and Philosophy of Science Part C: Studies in the History and Philosophy of Biological and Biomedical Science* 59 (2016): 1-10.

Reichenbach, Hans, and Maria Reichenbach. *The Direction of Time*. Berkeley: University of California Press, 1991.

Reinhart, Brenda J., Frank J. Slack, Michael Basson, Amy E. Pasquinelli, Jill C. Bettinger, Ann E. Rougvie, H. Robert Horvitz, and Gary Ruvkun. "The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans." *Nature* 403, no. 6772 (2000): 901-906.

Rheinberger, Hans-Jörg. *Toward a history of epistemic things: Synthesizing proteins in the test tube*. Stanford: Stanford University Press, 1997.

Rheinberger, Hans-Jörg. "What happened to molecular biology?" *BioSocieties* 3, no. 3 (2008): 303-310.

Rheinberger, Hans-Jörg. "Recent science and its exploration: the case of molecular biology." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 40, no. 1 (2009): 6-12.

Richardson, Sarah S., and Hallam Stevens, eds. *Postgenomics: Perspectives on Biology After the Genome*. Duke University Press, 2015.

Rosenberg, Alexander. *Darwinian reductionism: Or, how to stop worrying and love molecular biology*. University of Chicago Press, 2008.

Ruvkun, Gary. "The perfect storm of tiny RNAs." *Nature medicine* 14, no. 10 (2008): 1041-1045.

Salmon, Wesley C. *Scientific explanation and causal structure of the world*. Princeton, New Jersey: Princeton University Press, 1984.

Sansom, Roger. "Constraining the adaptationism debate." *Biology and Philosophy* 18, no. 4 (2003): 493-512.

Sansom Roger, and Robert N. Brandon. *Integrating Evolution and Develoment*. MIT Press, 2007.

Sarkar, Sahotra. *Molecular models of life: philosophical papers on molecular biology*. MIT Press, 2005.

Schaffner, Kenneth. "Logic of discovery and justification in regulatory genetics." *Studies in History and Philosophy of Science, Part A* 4 (1974): 349-385.

Scholl, Raphael, and Tim Räz. "Towards a methodology for integrated history and philosophy of science." In *The Philosophy of Historical Case Studies*, edited by Tilman Sauer and Raphael Scholl, 69-91. Springer International Publishing, 2016.

Shi, Yongsheng. "Alternative polyadenylation: New insights from global analyses." *RNA* 18, no. 12 (2012): 2105-2117.

Shubin, Neil, Cliff Tabin, and Sean Carroll. "Deep homology and the origins of evolutionary novelty." *Nature* 457, no. 7231 (2009): 818-823.

Stephens, Tim. "The Geneticist & the Biochemist." *UC Santa Cruz Review*, Spring 2004.

Stotz, Karola, and Paul Griffiths. "Genes: Philosophical analyses put to the test." *History and Philosophy of the Life Sciences* 26 (2004): 5-28.

Tabery, James G. "Synthesizing Activities and Interactions in the Concept of a Mechanism." *Philosophy of Science* 71, no.1 (2004): 1-15.

Tomoyasu, Yoshinori, Scott R. Wheeler, and Robin E. Denell. "Ultrabithorax is required for membranous wing identity in the beetle Tribolium castaneum." *Nature* 433, no. 7026 (2005): 643-647.

Wagner, Günter P. "The biological homology concept." *Annual Review of Ecology and Systematics* 20 (1989): 51-69.

Wagner, Günter P. "Homology and the mechanisms of development." In *Homology: The hierarchical basis of comparative biology*, edited by Brian K. Hall, 273-299. Academic Press, 1994.

Wagner, Günter P. *Homology, genes, and evolutionary innovation*. Princeton University Press, 2014.

Wagner, Günter P., Chi-Hua Chiu, and Manfred Laubichler. “Developmental evolution as a mechanistic science: the inference from developmental mechanisms to evolutionary processes.” *American Zoologist* 40, no. 5 (2000): 819-831.

Wagner, Günter P., and Vincent J. Lynch. “Evolutionary novelties.” *Current Biology* 20, no. 2 (2010): R48-R52.

Warren, Robert W., Lisa Nagy, Jane Selegue, Julie Gates, and Sean Carroll. “Evolution of homeotic gene regulation and function in flies and butterflies.” *Nature* 372, no. 6505 (1994): 458-461.

Waters, C. Kenneth. “Why the anti-reductionist consensus won't survive: The case of classical Mendelian genetics.” In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 125-139. Philosophy of Science Association, 1990.

West-Eberhard, Mary Jane. *Developmental plasticity and evolution*. Oxford University Press, 2003.

Wilkins, Jon F., and Peter Godfrey-Smith. “Adaptationism and the adaptive landscape.” *Biology & Philosophy* 24, no. 2 (2009): 199-214.

Wilkins, Jon F., and Peter Godfrey-Smith “Adaptationism.” In *A Companion to the Philosophy of Biology*, edited by Sahotra Sarkar and Anya Plutynski, 186-201. West Sussex: Wiley-Blackwell, 2011.

Woese, Carl R. “A new biology for a new century.” *Microbiology and molecular biology reviews* 68, no. 2 (2004): 173-186.

Wollman, Élie Léo, and François Jacob. “Sexuality in Bacteria.” *Scientific American*, 195 (1956): 109-118.

Woodward, James. *Making things happen: A theory of causal explanation*. Oxford University Press, 2003.

Woodward, James. “The problem of variable choice.” *Synthese* 193, no. 4 (2016): 1047-1072.

Young, David. *The Discovery of Evolution*. (2nd Ed.) Cambridge University Press, 2007.

Zhao, Yong, and Deepak Srivastava. “A developmental view of microRNA function.” *Trends in biochemical sciences* 32, no. 4 (2007): 189-197.