Detection and Manipulation of Extrachromosomal Circular DNA in Yeast

by

Tula Keal

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved April 2022 by the
Graduate Supervisory Committee:

Xiao Wang, Chair
Xiaojun Tian
Christopher Plaisier

ARIZONA STATE UNIVERSITY

May 2022

ABSTRACT

Extrachromosomal circular DNA (eccDNA) has become an increasingly popular subject of study in eukaryotic cell biology due to its prevalence in human cancer. Though the literature reports a consensus regarding DNA break repair as a driver of eccDNA formation, there remains a lack of knowledge surrounding the exact mechanisms for eccDNA formation and the selective dynamics that promote their retainment in a cell or population. A central issue to studying eccDNA is the inability to distinguish between linear and circular DNA of homologous sequence. The work presented here describes an adapted eccDNA enrichment and detection assay, specifically for investigating the effects of manipulating a known eccDNA-forming locus in the budding yeast *Saccharomyces cerevisiae*. First, a galactose inducible GFP reporter was integrated within the copper inducible *CUP1* tandem repeat locus of yeast cells. The eccDNA enrichment and detection assay was first applied to wildtype yeast to demonstrate the presence of *CUP1* eccDNA in copper induced cells by qPCR. Although subsequent sequencing analysis failed to validate this result, it indicated the presence of various other known and previously un-reported eccDNA species. Finally, application of the enrichment protocol and qPCR detection assay to *CUP1-GFP* reporter cells yielded inconclusive results, suggesting the assay requires further optimization to sensitively detect eccDNA from this altered locus. While more work is necessary to draw conclusions regarding the limits of eccDNA production at a manipulated eccDNA-forming locus, this knowledge will lend to the potential for therapeutically targeting eccDNA at the point of *de novo* formation.

DEDICATION

This thesis is dedicated to my friends and family who have offered immense emotional and financial support to me as I completed my undergraduate and master's degrees. In particular, my mother Farrell Keal and partner Jared Mazzarella were instrumental to maintaining the motivation necessary for me to complete this work. Finally, this work is dedicated to my grandmother who passed before its completion.

ACKNOWLEDGMENTS

I am deeply grateful for my supervisor, Dr. Xiao Wang, who welcomed me into his lab and provided space for me to explore a new question. Dr. Wang granted me a uniquely independent and enriching master's research experience, and I am confident that his training has adequately prepared me for a career in academic research. I greatly appreciated his mentorship and close guidance as I explored career options and made the decision to enroll in the Biomedical Sciences PhD program at UC San Diego next Fall.

I am also grateful for my peers in the Wang and Tian labs who assisted in my transition to graduate-level research during the height of the pandemic. Specifically, Kylie Standage-Beier has been an indispensable asset to my work on this project and I have greatly valued his advice throughout these two years.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In eukaryotes the genomic content of a cell is primarily organized into linear chromosomes, however a growing body of evidence suggests the significance of extrachromosomal circular DNA (eccDNA) for genome expression and stability. eccDNA was first observed in boar sperm and wheat embryo in 1965, yet its ubiquity among eukaryotes went largely unnoticed until the advent of more powerful sequencing and microscopy techniques (Hotta and Bassel, 1965). Since then, eccDNA has also been identified in the germline of humans, as well as in both healthy and cancerous somatic tissues (Henriksen et al., 2022; Møller et al., 2018; Kim et al., 2020). Moreover, eccDNA has been characterized in other model eukaryotes including yeast, *Drosophila*, mouse, and the flowering plant *A. palmeri* (Møller et al., 2015; Stanfield et al., 1979; Sunnerhagen et al., 1986; Koo et al., 2018). While eccDNA is present in many organisms, its frequency and size are highly variable. As a result, eccDNA can be sub-classified into microDNA for the more commonly observed circles less than 500bp in size, or ecDNA for circles that are observed in cancer and multiple-megabases long, also referred to as double minutes (Liao et al., 2020).

The exact mechanism for eccDNA formation remains unknown, but its small size in healthy tissues and relative abundance in cancer suggests it results from genome instability and aberrant DNA repair. Multiple models have been proposed to describe eccDNA formation from the linear chromosome, and there is contention in the literature regarding the molecular components of such events. One study suggests chromothripsis as a primary driver of ecDNA formation and cancer progression, while another proposes

that oncogenic DNA is excised from the chromosome as an episome which self-amplifies and enlarges (Shoshani et al., 2021; Storlazzi et al. 2006). Early molecular analysis of ligase-knockout mice suggested that eccDNA formation in healthy cells is dependent on non-homologous end-joining (NHEJ), but a more recent study in human and chicken cells demonstrated the opposite (Cohen et al., 2006; Paulsen et al. 2021). Indeed, the majority of eccDNA in healthy human cells contain short repeats or derive from repetitive elements in the genome, suggesting a role for the microhomology-mediated repair (MMR) and homology directed repair (HDR) pathways (Møller et al., 2019). Molecular validation of MMR- and HDR-mediated eccDNA formation have both been reported in the literature, thus eccDNA formation is likely a result of various DNA repair-related mechanisms depending on cellular context (Paulsen et al., 2021; Cohen and Mechali, 2001).

Though most eccDNA are too small to contain any genetic elements and are therefore quickly lost from the population, some contain an origin of replication to promote their amplification (Molin et al., 2020). In addition to amplification through self-replication, eccDNA have been shown to reintegrate into the chromosomal genome via homologous recombination (HR) (Brewer et al., 2015). In cancer, this function of eccDNA results in novel interactions between oncogenes and regulatory elements, effectively remodeling the genomes of tumor cells (Koche et al., 2020). While eccDNA can contain a replication origin they generally do not harbor a centromere; only 2% of eccDNA derived from healthy human muscle cells were shown to contain centromeric sequences (Møller et al., 2018). The effect of this is random and unequal segregation of eccDNA between dividing cells, which has been thoroughly demonstrated in multiple

cancer cell lines (Yi et al., 2022; Lange et al. 2021). Ultimately, the amplification, reintegration, and random segregation of eccDNA in both healthy and cancerous tissues contribute to genetic copy number variation (CNV). In turn, this CNV supports the rapid adaptation of cells in response to environmental changes, such as the acquisition of drug resistance demonstrated in both human cancer and glyphosate-resistant *A. palmeri* (Nathanson et al., 2014; Molin et al., 2020).

The budding yeast *Saccharomyces cerevisiae* has proven to be an efficient model for studying the molecular mechanisms and functionality of eccDNA, as 23% of the yeast genome was found to reside also on eccDNA (Møller et al., 2015). Furthermore, 80% of eccDNA in yeast contain an autonomously replicating sequence (ARS), and some eccDNA species are highly enriched across yeast populations (Møller et al., 2015). Of these, extrachromosomal ribosomal circles (ERCs) are frequently observed in yeast, specifically in aged cells (Prada-Luengo et al., 2020). It was previously shown that 94% of the eccDNA present in young yeast cells are lost as cells age; instead, these aged cells contain enriched populations of eccDNA from repetitive regions of the genome like rDNA (Hull and Houseley, 2020). This observation suggests an evolutionary function for eccDNA in which its accumulation in aged yeast cells confers an adaptive benefit to the population, since young cells avoid the metabolic burden of eccDNA amplification while taking advantage of the genome plasticity it confers (Prada-Luengo et al., 2020). Another example of an enriched eccDNA species in aged yeast is from the *CUP1* locus, a copper-inducible and tandemly repeated gene encoding for copper metallothionein. It was determined that transcriptional induction of the locus by copper results in site-specific

eccDNA formation, dependent on the double-stranded break repair proteins Sae2, Mre11, and Mus81 in a HR-directed mechanism (Hull et al., 2019).

The composition of the *CUP1* locus and its inducible activity makes it an ideal candidate for reporter construction to investigate the formation of eccDNA. Previous approaches using a fluorescent reporter to observe eccDNA have focused on visual tracking for the determination of segregation behavior (Yi et al., 2022; Lange et al., 2021). In the study presented here, a galactose inducible GFP cassette was integrated between tandem repeats of *CUP1* in *S. cerevisiae* to elucidate the dynamics of eccDNA formation under selective pressure. While it was previously reported that eccDNA forms from the *CUP1* locus in a site-specific manner regardless of coding sequence, insertion of a multiple-kilobase cassette was hypothesized to potentially inhibit eccDNA formation (Hull et al., 2019). Additionally, it was hypothesized that induction of *GFP* by galactose would result in eccDNA formation at the *CUP1* reporter locus, and furthermore, that induction of this locus might generate chimeric *CUP1-GFP* eccDNA of several kilobases in size. To test these hypotheses, a sensitive and quantitative method for the detection of target eccDNA species was developed using a previously described enrichment technique and a consequent qPCR assay (Møller et al., 2016). Molecular validation of eccDNA has been a commonly cited issue in the literature due to the sequence homology shared by eccDNA species and the regions of linear DNA from which they arise (Wang et al., 2021; Mouakkad-Montoya et al., 2021). I demonstrate here an approach for detecting the relative quantities of target eccDNA in an enriched sample and apply this method to investigate the formation of novel eccDNA from the *CUP1-GFP* reporter locus.

CHAPTER 2

METHODS

Culturing and Induction

Escherichia coli: Plasmid construction was performed in competent DH10β cells (NEB, C3019) cultured in LB media supplemented with 50 μg/mL of ampicillin.

Saccharomyces cerevisiae: The yeast strain used in this study was *S. cerevisiae* 834 (ATCC, 834). For cloning applications, competent cells were cultured in selective YPD media and on YPD agar dropout plates. The *CUP1-GFP* reporter and plasmid p415 contained histidine and leucine auxotrophic markers, respectively. For induction experiments, cells were grown in 3 mL selective YPD to the log phase then induced with 1 mM $CuSO_4$ and/or YP Galactose medium. Induction cultures grew 24 hours with dilutions to maintain the log phase prior to extraction.

CUP1-GFP Reporter

Molecular Cloning: The *CUP1-GFP* reporter was assembled in plasmid pMG containing a galactose inducible *GFP* cassette, a constitutive *mCherry* cassette, and *AmpR* and *His3* selective markers. Homology arms with restriction sites were amplified from genomic DNA using primers designed at the junction of two *CUP1* repeats, listed in Supplementary Table 1. Reactions were composed of 20 ng genomic DNA, 10 mM dNTPs, 10 μM primers, 1X HF Buffer, and 1 unit of Phusion High-Fidelity DNA polymerase (NEB, M0530). PCR products were purified using the GenElute PCR Clean-Up Kit (Sigma Aldrich, NA1020) following the manufacturer's protocols.

pMG and the *CUP1A* homology arm were first digested with SpeI-HF and SacI-HF (NEB, R3133 and R3156), and reactions contained 1X CutSmart buffer and 1 unit of each enzyme to a volume of 50 µL. The reactions contained either 1 µg of plasmid DNA or 300 ng of the amplified homology arm. Reactions were incubated at 37°C for an hour, and products were confirmed by gel electrophoresis then purified. The digested plasmid was dephosphorylated by adding rSAP (NEB, M0371) and 1X CutSmart buffer in a total volume of 50 µL at 37°C for an hour, then subsequently purified. Next, 40 ng of the dephosphorylated backbone was ligated with the digested *CUP1A* homology arm at a 3:1 insert:vector ratio in a 20 µL reaction containing 1X ligase buffer and 0.5 µL T4 DNA Ligase (NEB, M0202). Ligations were incubated for an hour at room temperature before 5 µL was added to 50 µL competent *E. coli* cells for heat shock transformation. Colonies were grown in selective media and subsequently underwent plasmid extraction following the GenElute Plasmid DNA Miniprep Kit (Sigma Aldrich, NA0200). Plasmid construction was confirmed by screening digestion repeating the previous SpeI/SacI digestion protocol. The assembled pCUP1a-G plasmid and the amplified *CUP1B* homology arm were then digested with SpeI-HF and EcoRI-HF (NEB, R3101), and all downstream cloning steps were repeated to assemble pCUP1-G. The plasmid was then linearized by digestion with SpeI-HF and transformed into competent *S. cerevisiae* cells using the Frozen-EZ Yeast Transformation II Kit (Zymo Research, T2001).

Reporter Validation: Integration of the *GFP* cassette at the *CUP1* locus was first confirmed by fluorescent microscopy. Cells were induced in selective YP Galactose media overnight, and 1 µL of each clone was loaded onto glass slides. Cells were observed at 20X on the Nikon Eclipse Ti20E with a 250 ms exposure time. The samples

were additionally analyzed on a BD Accuri C6 Plus Flow Cytometer to confirm GFP integration. For the copper induction time course experiment, cells were induced with copper and galactose as previously described in 3 mL of either selective or non-selective media and cultured for three days. Cultures were diluted daily and timepoints were taken by flow cytometry in triplicate. To facilitate the application of the eccDNA detection qPCR assay to the *CUP1-GFP* reporter, p415-mCherry was generated by substituting the *GFP* coding sequence on p415-GFP for that of *mCherry*. Briefly, *mCherry* was amplified from pMG using the primers provided by Supplementary Table 1, and both the amplicon and p415-GFP were digested with XhoI and BamHI-HF following the digestion protocol described previously. Subsequent dephosphorylation, ligation, and transformation steps were also performed as previously described. Clones were confirmed by colony PCR following the previous PCR protocol, and correct clones were transformed into 834 *CUP1-GFP* competent cells.

Extrachromosomal Circular DNA Enrichment

DNA Isolation: Yeast genomic DNA and circular DNA were extracted using the Zymo YeaStar Genomic DNA Kit (Zymo Research, D2002) and Zymoprep Yeast Plasmid Miniprep Kit (Zymo Research, D2001). One mL of log phase yeast cells was used for each extraction following the manufacturer's protocols. Total genomic DNA was eluted into 60 µL of ddH_2O while miniprep DNA extractions were eluted into 15 µL ddH_2O. The concentration of isolated DNA was taken on a Nanodrop Lite spectrophotometer.

Exonuclease Treatment: Digestion of linear DNA was performed using two exonucleases, T5 (NEB, M0663) and ExoV (RecBCD) (NEB, M0345). Both reactions included 10 units of enzyme, 1X reaction buffer, and 10 ng of DNA in a total volume of 50 µL. The ExoV reaction contained an additional 1 mM ATP. Reactions were incubated at 37°C overnight and terminated the following day. ExoV was heat inactivated at 70°C for 30 minutes, while T5 was inactivated by the addition of 1X Purple Gel Loading Dye containing SDS (NEB, B7024). Remaining DNA was purified and eluted into 30 µL ddH_2O.

Rolling Circle Amplification: Circular DNA present in the exonuclease-treated samples was enriched by phi29 polymerase in a rolling circle amplification reaction adapted from Wang et al., 2021. Briefly, 1 µL of the purified exonuclease-treated sample was added to a reaction containing 1X phi29 amplification buffer, 1X Exo-resistant Random Primer (ThermoFisher, SO181), 200 µM dNTPs, and ddH_2O to a volume of 18.6 µL. Prior to the addition of enzyme, the reaction was incubated at 95°C for 5 minutes before a 1 °C/s ramp down to 30℃. Next, 10 units of phi29 (NEB, M0269) and 0.1 mg/mL of recombinant albumin were added for a total reaction volume of 20 µL. The reaction was incubated at 30℃ overnight and heat inactivated the following day at 65°C for 10 minutes. The reaction products were then purified and eluted into 35 µL ddH_2O.

Extrachromosomal Circular DNA Validation

qPCR: Quantitative PCR assays were designed using iTaq Universal SYBR Green Supermix (Bio-Rad, 1725121). Reactions were composed of 1X supermix, 500 nM primers, and 1 ng template DNA to a total volume of 10 µL. Primer sequences are shown in Supplementary Table 1. Reactions were incubated in a Bio-Rad CFX384 Real-Time

PCR System with a 60℃ annealing/extension step and 40 cycles. Assays were repeated twice (three repeats for Figure 5B), and reactions were performed in triplicate. Data was analyzed using the ΔΔCt method in Microsoft Excel (v. 16.57). Outlying replicates which deviated from the average Ct of a sample by more than six cycles were removed prior to analysis. For samples in which all technical replicates failed to amplify, an average Ct of 40 was assigned. Standard deviation was calculated only for data plotted linearly.

Sequencing: Libraries were prepared using 1 ng of each eccDNA-enriched sample in the Nextera XT Library Prep Kit, which included enzymatic digestion to generate 150 bp reads (Illumina, FC-131-1024). Samples were run in parallel on the Illumina NovaSeq 6000 platform to obtain ~8 million reads per sample. Sequence reads were mapped to the reference S288C genome (SacCer3) and the p415-GFP plasmid sequence using Bowtie 2 (Langmead, B. and Salzburg, S.L., 2012). Aligned reads were visualized in Integrative Genomics Viewer v. 2.9.4.

CHAPTER 3

RESULTS

Construction of a *CUP1* eccDNA Fluorescent Reporter

The reference yeast genome describes the *CUP1* locus as containing two tandem repeats of the copper metallothionein coding sequence, however as many as 15 tandem repeats have been reported (Fogel and Welch, 1982). To confirm at least two copies are present in the *S. cerevisiae* 834 genome, primers were designed around the individual *CUP1* repeats, referred to here as *CUP1A* and *CUP1B* (Supplementary Table 1). The junction of these repeats was also amplified, and the entire 4 kb double-repeat was attempted. Due to the repetitive nature of this locus, the sequences complementary to the *CUP1A* reverse and *CUP1B* forward primers were repeated such that smaller, non-target amplicons were commonly observed. The gel image shown in Supplementary Figure 1 is representative of several attempts in which these inadvertent products of ~300 bp were preferentially amplified over the ~2 kb *CUP1A* and *CUP1B* repeats. However, the full *CUP1A* repeat was achieved as indicated by the band at 2300 bp, suggesting the presence of an adjacent *CUP1B* repeat. Furthermore, the ~4 kb double-repeat amplicon was never achieved experimentally which could indicate the presence of a third tandem repeat.

With at least two copies of *CUP1* in the *S. cerevisiae* 834 genome, it was possible to design an intervening fluorescent reporter for integration via HR. The cloning scheme for the *CUP1-GFP* reporter is shown by Figure 1A, in which homology arms were amplified from genomic DNA at the end of *CUP1A* and the beginning of *CUP1B*, 182 bp and 229 bp respectively. Figure 1B shows these amplicons (left), as well as the homology arms following digestion with their respective restriction enzymes (right). These arms

were cloned into plasmid pMG individually to substitute the *TEF1* promoter region for the *CUP1A* homology arm, and the *mCherry* coding sequence for the *CUP1B* homology arm. Figure 1C shows digestion of pMG to excise the 862bp *TEF1* promoter region (left), and a subsequent screening digestion to identify positive clones following ligation of the *CUP1A* homology arm (middle). The two positive clones were then digested to excise the 730 bp *mCherry* sequence from pCUP1a-G, also shown in Figure 1C (right). After ligating the *CUP1B* homology arm, clones underwent screening digestion to confirm the presence of both homology arms as shown in Figure 1D.

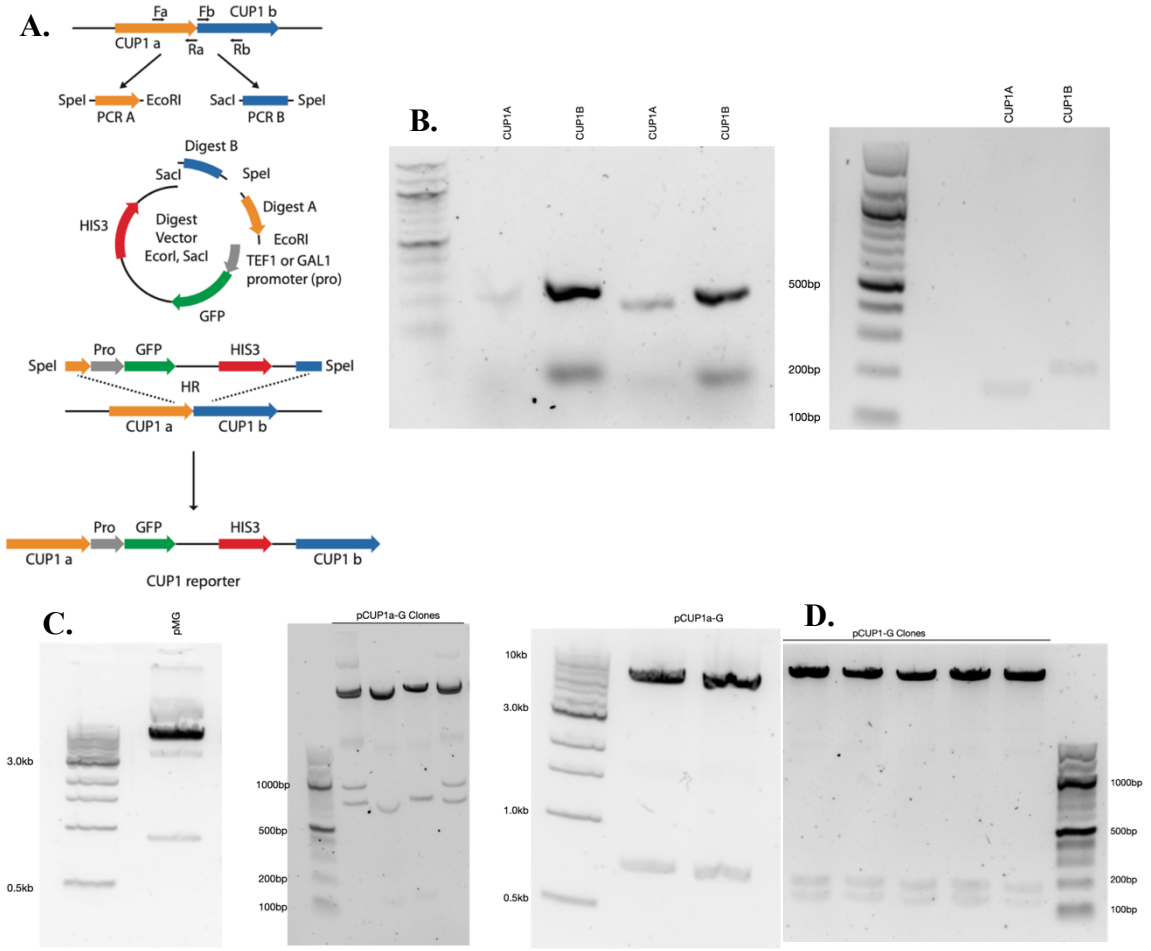


Figure 1. Construction of the *CUP1-GFP* reporter. **A.** Overview of the *CUP1-GFP* reporter cloning and integration scheme. *CUP1* homology arms were cloned into plasmid pMG containing a Gal-inducible *GFP* cassette and a *His3* auxotrophic marker. pCUP1-G was then linearized to promote homologous recombination of the plasmid into the *CUP1* repeat locus of *S. cerevisiae*. **B.** *CUP1* homology arms

following PCR amplification from genomic DNA (left) and following restriction digestion (right). *CUP1A*, 182 bp; *CUP1B*, 229 bp. **C.** Digested products of pMG (left), four clones following transformation with pCUP1a-G (middle), and two positive pCUP1a-G clones (left). The initial pMG digestion excised the 862 bp promoter region, and the subsequent pCUP1a-G digestion excised the 730 bp *mCherry* sequence. The intermediate screening digestion confirmed substitution of the promoter for the *CUP1A* homology arm. **D.** Screening digestion of six clones following transformation with pCUP1-G containing both homology arms.

Competent yeast cells were then transformed with linearized pCUP1-G to promote HR of the reporter cassette between *CUP1* repeats. Integration was confirmed by galactose induction and subsequent visualization by fluorescent microscopy, shown in Figure 2A. To initially test the responsiveness of the reporter to copper induction, cells were grown in galactose media with 1 mM $CuSO_4$ for up to three days. Furthermore, cells were cultured with and without histidine to observe the effect of selection on reporter activity. Figure 2B shows that cells cultured in media selecting for retainment of the *CUP1-GFP* reporter experienced a decrease in viability after 48 hours post-induction, but inversely exhibited a sharp increase in average GFP fluorescence. Contrastingly, cells cultured without selection had greater viability under long-term copper induction while average fluorescence declined by about 68%.

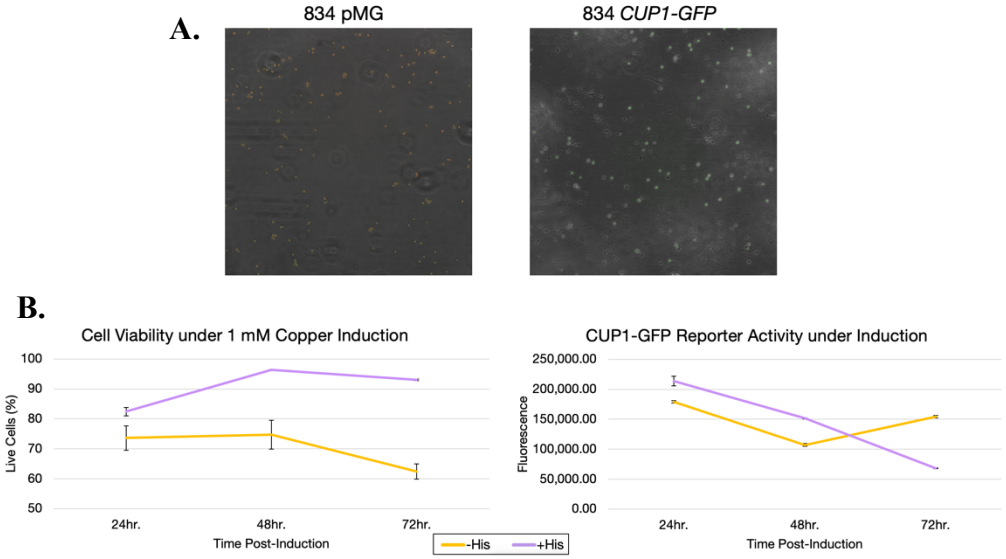


Figure 2. Initial testing of the *CUP1-GFP* reporter. **A.** Fluorescent micrographs of *S. cerevisiae* 834 cells transformed with pMG (left) and the 834 *CUP1-GFP* reporter strain (right). Cells were cultured in YP Galactose for 24 hours to induce *GFP* expression in both samples; *mCherry* expression was constitutive. Images were taken at 20X with a 250ms exposure time. **B.** Flow cytometry results following long-term

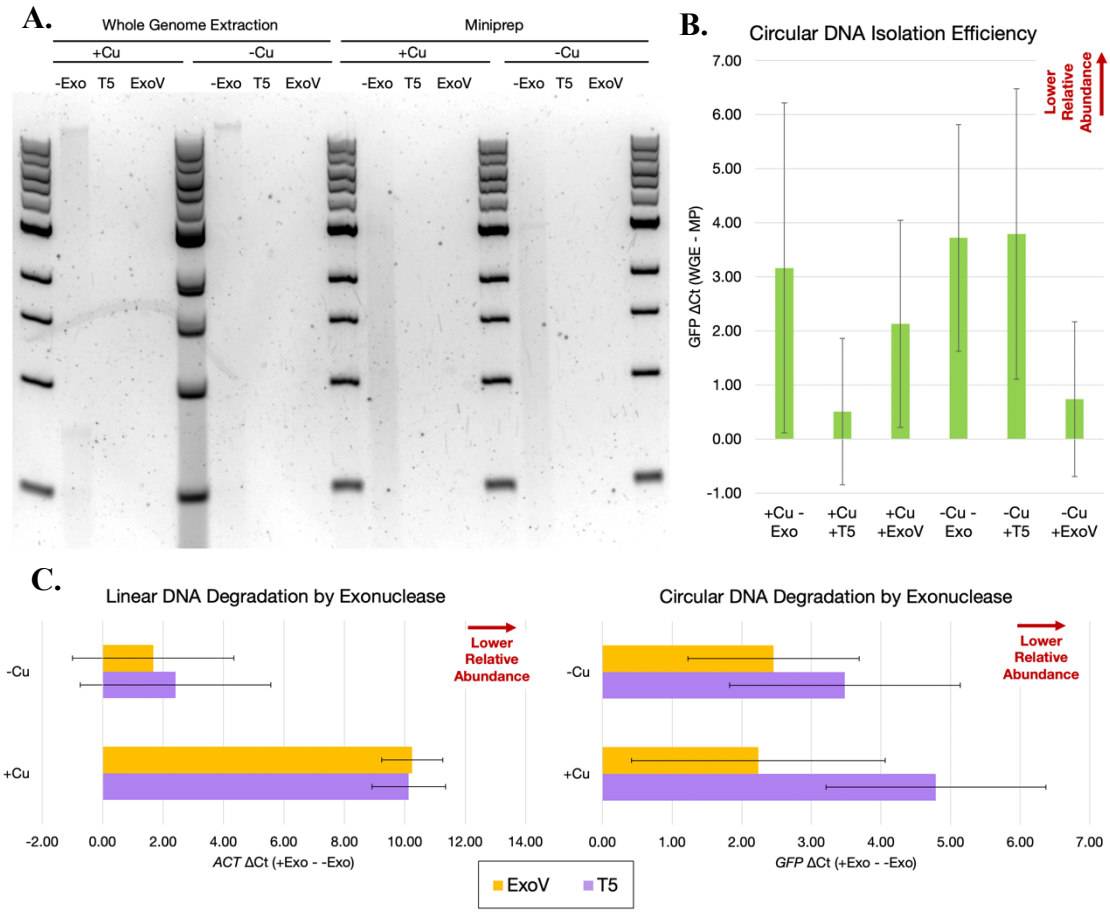
copper induction of 834 *CUP1-GFP* cells. Cells were cultured for three days in YP Galactose with (purple) and without (orange) histidine and were induced with 1 mM $CuSO_4$. The proportion of cells within the live gate were recorded to reflect cell viability following exposure to copper (left). GFP fluorescence was also recorded over time to reflect *CUP1-GFP* reporter activity under galactose and copper induction (right).

Development of an eccDNA Enrichment and Detection Assay

Before determining if eccDNA forms at the *CUP1-GFP* reporter locus, a protocol was first adapted to enrich and validate a target population of eccDNA. This assay was developed in wildtype yeast transformed with the single-copy plasmid p415-GFP, which served as a housekeeping eccDNA for the purpose of qPCR data analysis. The eccDNA enrichment assay entailed the following steps: DNA isolation, linear DNA degradation, circular DNA amplification, and consequent qPCR analysis. Circular DNA isolation efficiency was tested using whole genome extraction (WGE) and miniprep (MP) approaches, the latter of which was expected to perform better. Figure 3A shows the consistently positive ΔCt of *GFP* isolated by MP as compared to WGE, confirming more efficient circular DNA isolation by MP. Linear DNA degradation was also tested using variable methods, by treatment with either T5 exonuclease or ExoV (RecBCD). Figure 3B qualitatively illustrates the activity of each exonuclease on both WGE and MP prepared samples, in which starting DNA is only visible in the untreated lanes.

To quantitatively compare the sensitivity of each exonuclease for linear DNA, exonuclease-treated and -untreated copper induction samples prepared by MP were used in a qPCR assay measuring the relative abundance of *ACT* and *GFP*. Figure 3C shows the activity of each exonuclease on linear DNA (*ACT*; top panel) and circular DNA (*GFP*; bottom panel). Exonuclease T5 and ExoV showed similar levels of linear DNA degradation, and this activity was greater in copper-induced samples in which *ACT* was detected about 8 cycles later than in un-induced samples. Contrastingly, T5 degraded

more circular DNA than ExoV, with *GFP* detection around 1 and 2.6 cycles sooner in T5-treated samples than in ExoV-treated samples for un-induced and copper-induced cells, respectively. Due to the relatively non-specific exonuclease treatment step, a final enrichment step of rolling circle amplification (RCA) was implemented using the polymerase phi29. The activity of phi29 was first tested using various primers including those specific for *ACT* and *GFP*, as well as a random hexamer. Supplementary Figure 2 shows that phi29 has little specificity when using target-specific primers, and furthermore, that the polymerase is most active upon priming with a random hexamer.



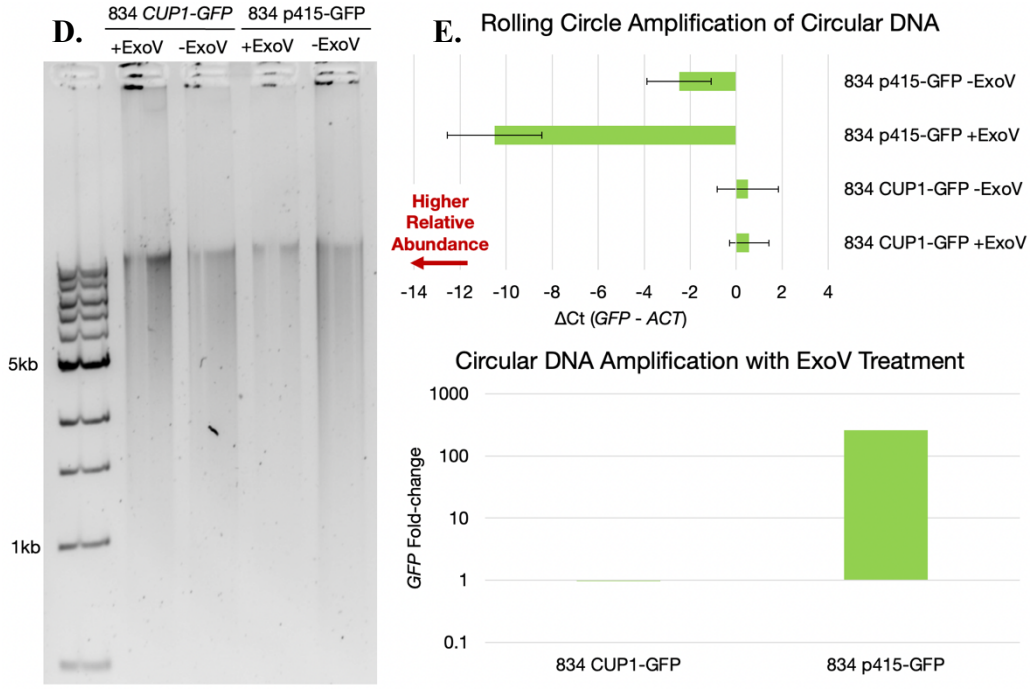


Figure 3. Workflow to develop an eccDNA enrichment and detection assay. **A.** Gel image of exonuclease-treated and -untreated copper-induction samples. 834 p415-GFP cells were cultured in YPD with (+Cu) or without (-Cu) 1 mM $CuSO_4$. Circular DNA was extracted by Whole Genome Extraction (WGE) or Miniprep (MP), and 10 ng of each sample were used in exonuclease reactions with either T5 or ExoV. **B.** Samples from A were used in a qPCR assay measuring relative *GFP* abundance between extraction methods. The cycle thresholds (Ct) of MP samples were subtracted from that of the corresponding WGE samples to obtain the plotted ΔCt values. **C.** MP samples from A were used in a qPCR assay measuring the relative abundance of linear *ACT* (left) and circular *GFP* (right) following exonuclease digestion by either T5 (purple) or ExoV (orange). The plotted ΔCt values were obtained by subtracting the Ct of each exonuclease-untreated sample from the Ct of corresponding samples treated with either enzyme. **D.** Gel image of products following rolling circle amplification (RCA) by phi29 polymerase. 834 p415-GFP and 834 *CUP1-GFP* DNA were extracted by WGE, and 10 ng of each were treated with ExoV prior to being added to individual phi29 RCA reactions. **E.** The RCA products from D were used in a qPCR assay to measure the activity of phi29 polymerase. The relative abundance of *GFP* to *ACT* was measured by obtaining the difference in Ct values between the two targets for each sample (top). Additionally, the effect of ExoV treatment prior to phi29 RCA was assessed by calculating the fold-change in *GFP* between exonuclease-treated and -untreated samples (bottom).

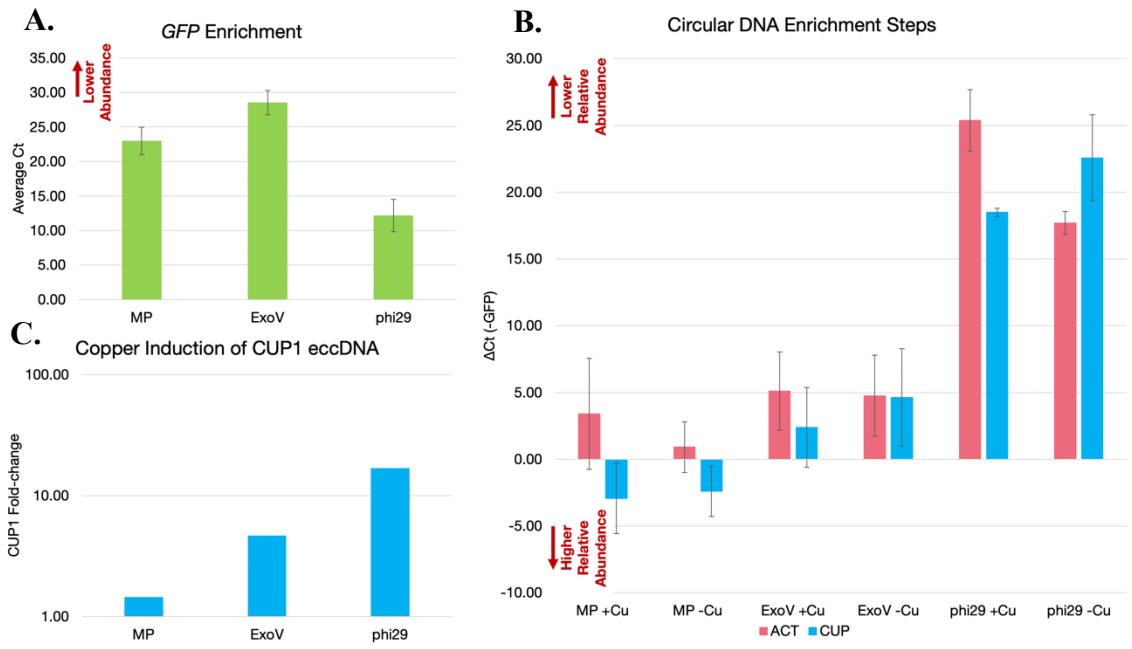
To next test the sensitivity of phi29 for circular DNA, an assay was designed using both the *CUP1-GFP* reporter yeast and wildtype yeast transformed with p415-GFP to compare the amplification of linear and circular *GFP* by RCA. Samples also underwent treatment with ExoV to observe the compounded effect of applying both enrichment steps. Figure 3D confirms the ability of phi29 to amplify both linear and circular DNA, but obscures quantitative differences in phi29 activity between the

samples. The results of subsequent qPCR analysis are shown in Figure 3E (top panel), which reveals a much higher abundance of *GFP* to *ACT* in the p415-GFP samples as compared to the *CUP1-GFP* samples. Furthermore, the preceding degradation of linear DNA by ExoV was shown to greatly improve the ability of phi29 to specifically amplify circular DNA (bottom panel), with a 258-fold-change in *GFP* between the ExoV-treated and -untreated RCA products. Taken together, these results demonstrate a robust approach for enriching and sensitively detecting populations of eccDNA in yeast cells.

Confirmation of Wildtype *CUP1* eccDNA Production

Next, the optimized assay was performed again in wildtype cells containing p415-GFP to measure the increase in *CUP1* abundance following copper induction. Figure 4A shows the average Ct of *GFP* between copper-induced and un-induced samples over the course of successive circular DNA enrichment steps. In Figure 4B, the average Ct values for *ACT* and *CUP* are normalized to those of *GFP*, with larger values indicating a lesser abundance relative to *GFP*. While *CUP1* is initially more abundant than *GFP* following MP extraction, consequent ExoV and phi29 enrichment steps result in a greater abundance of *GFP* than of *CUP1*. The disparity between *ACT* and *GFP* similarly increases with each circular DNA enrichment step. Following RCA, *CUP1* was detected about 4 cycles earlier in the copper-induced sample than in the un-induced sample. These data are also represented by Figure 4C, in which the fold-change of *CUP1* following induction with copper is shown for each eccDNA enrichment step. While *CUP1* had a slight fold-change of about 1.5 in the MP sample induced with copper, this increased to 4.7 and 17 in the subsequent ExoV-treated and phi29-amplified samples, respectively.

Whole genome sequencing was then performed to confirm the presence of *CUP1* eccDNA and identify other potential eccDNA species in the copper-induced and un-induced enriched samples. For the induced sample, 9.53% of reads aligned to the p415-GFP plasmid while 2.89% mapped to the S288C reference genome. For the un-induced sample, only 2.86% of reads mapped to p415-GFP while 1.95% mapped to S288C. Interestingly, most reads mapped to the endogenous 2µ yeast plasmid, with 22.7% of reads from the induced sample and 25.19% of reads from the un-induced sample aligning to the 2µ sequence. Figure 4D reveals that no reads mapped to *CUP1* in the copper-induced sample, however some low-confidence reads mapping to *CUP1* were obtained in the un-induced sample, contradictory to previous qPCR results. While no *CUP1* sequence was obtained from the induced sample, eccDNA was identified in this sample from another tandem repeat locus, namely the *ENA1/2/5* array, as shown in Figure 4E.



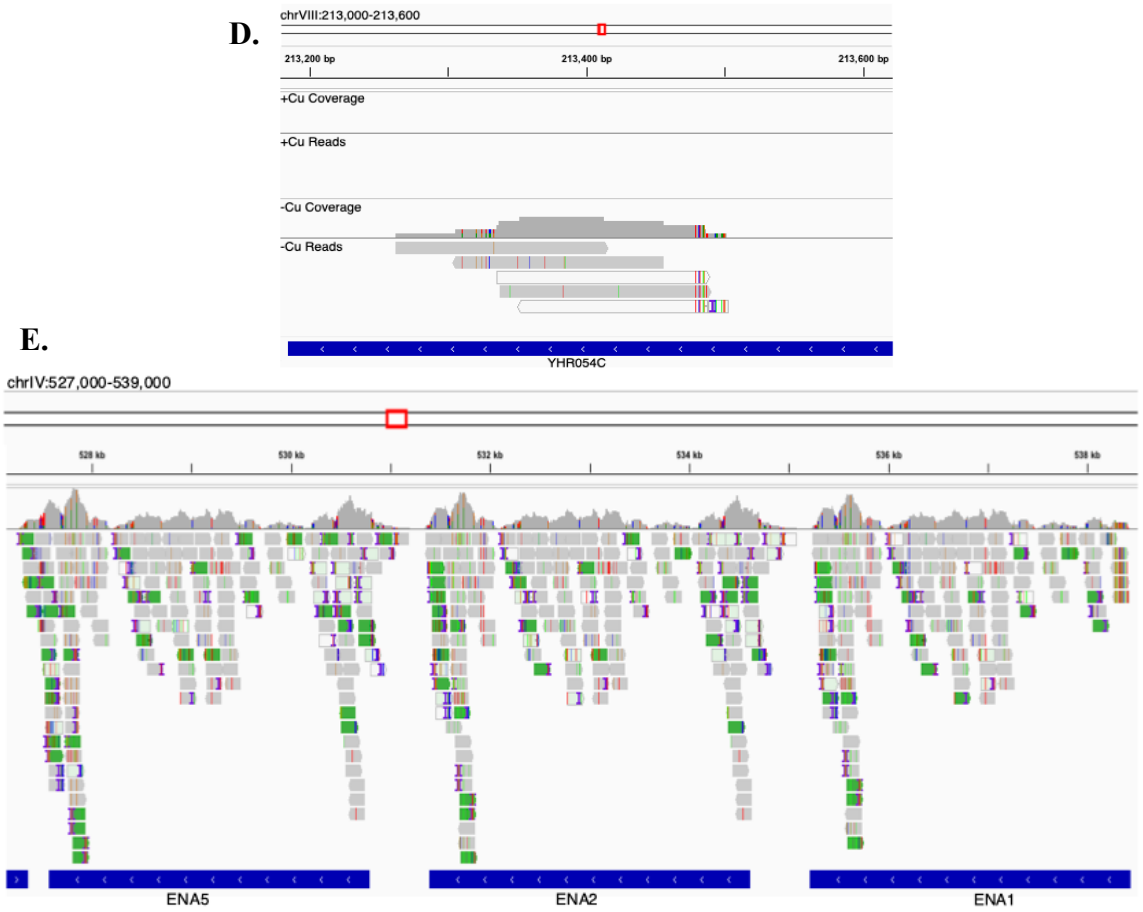


Figure 4. Detection of *CUP1* eccDNA. **A.** The efficacy of the eccDNA enrichment protocol was first analyzed in terms of *GFP*, which was present in wildtype cells on the single-copy p415-GFP plasmid. Each successive step for eccDNA enrichment is shown, beginning with Miniprep isolation (MP), then linear DNA degradation by ExoV, and subsequent rolling circle amplification (RCA) by phi29. The average Ct of *GFP* at each step was calculated using both + and -Cu samples since p415-GFP was unperturbed by copper induction. **B.** To infer the presence of *CUP1* eccDNA in copper-induced samples, the Ct values for *ACT* and *CUP1* were first normalized to the housekeeping gene *GFP* shown in A. This ΔCt value was plotted for + and -Cu samples at each stage in the enrichment process to visualize the abundance of each target relative to *GFP*. **C.** The fold-change of *CUP1* abundance relative to *GFP* upon copper induction is shown for each step in the circular DNA enrichment process. **D.** Sequencing coverage and reads mapping to the *CUP1* repeat locus (YHR054C) in copper-induced (upper tracks) and un-induced (lower tracks) samples. **E.** Sequencing coverage and reads mapping to the *ENA1/2/5* repeat locus in the copper-induced sample.

Assay Application to Detect Potential *CUP1-GFP* eccDNA

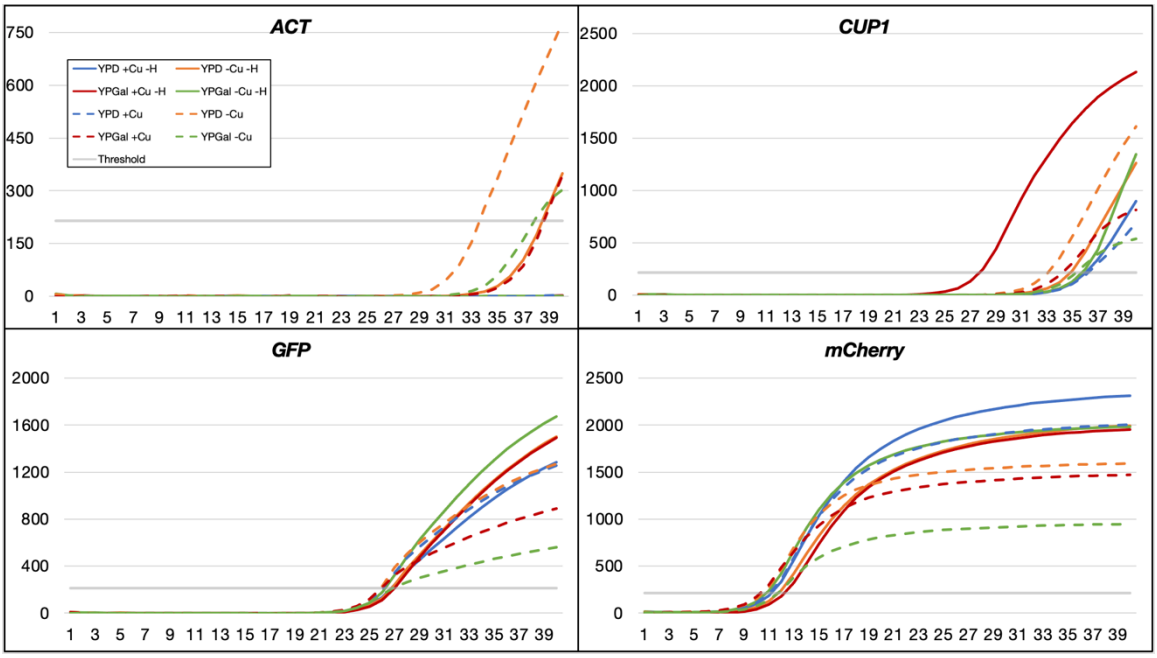
Despite discrepancies between the detection of *CUP1* eccDNA by qPCR and sequencing, the enrichment assay was next applied to the yeast strain containing a *CUP1-GFP* reporter cassette on Chr. VIII. Cells were cultured with induction by copper sulfate and/or galactose, and in histidine dropout media (-His) which selects for the reporter

cassette. Figures 5A-B show the amplification curves of all relevant gene targets in each of the described media conditions. Sample preparation was performed twice to give two trials of the experiment, which yielded differing results. In Trial 1 (Figure 5A), *ACT* was non-detectable except in the un-induced cultures and the non-selective cultures with galactose. Contrastingly, *CUP1* was observed in all samples, with the double-induction sample cultured under selection being the earliest detected at around 28 cycles. Both *GFP* and the *mCherry* housekeeping gene had little variation in Ct values between samples, with *GFP* becoming detectable at around 26 cycles and *mCherry* 15 cycles earlier.

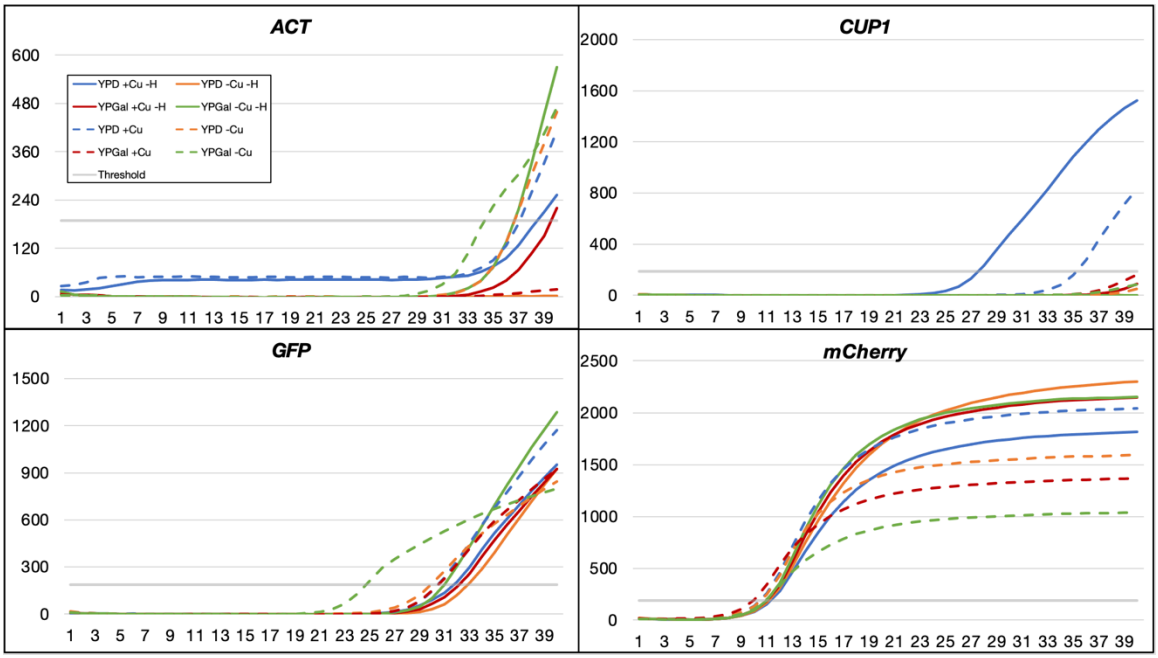
In Trial 2 (Figure 5B) *ACT* was instead detected in all but two samples, while *CUP1* was only detected before 40 cycles in the samples under copper induction in glucose media. While *mCherry* was again detected after about 11 amplification cycles for all samples, *GFP* seemed to be detected earlier in the samples that were cultured in non-selective media. Furthermore, *GFP* appeared significantly more abundant in the samples induced only with galactose. One consistent result between the trials was the higher abundance of *GFP* than *CUP1* relative to *mCherry*, as shown by the consistently smaller ΔCt of *GFP* in Figure 5C. In Trial 1 the relative abundance of *GFP* to *mCherry* was consistent between different selection and induction conditions; however, in Trial 2 *GFP* abundance was greater in response to galactose induction in the non-selective media by about 5 amplification cycles, but not in selective or copper-induced medias. An expected increase of *CUP1* in response to copper induction but not galactose was observed in Trial 2, with a cycle difference of about 10 in selective media and 3 in non-selective media. In Trial 1, *CUP1* abundance seemed to significantly increase in response to copper only in selective galactose media; *CUP1* in this sample was detected 10 cycles earlier than in

cells cultured in both selective galactose media without copper and selective glucose media with copper.

A.



B.



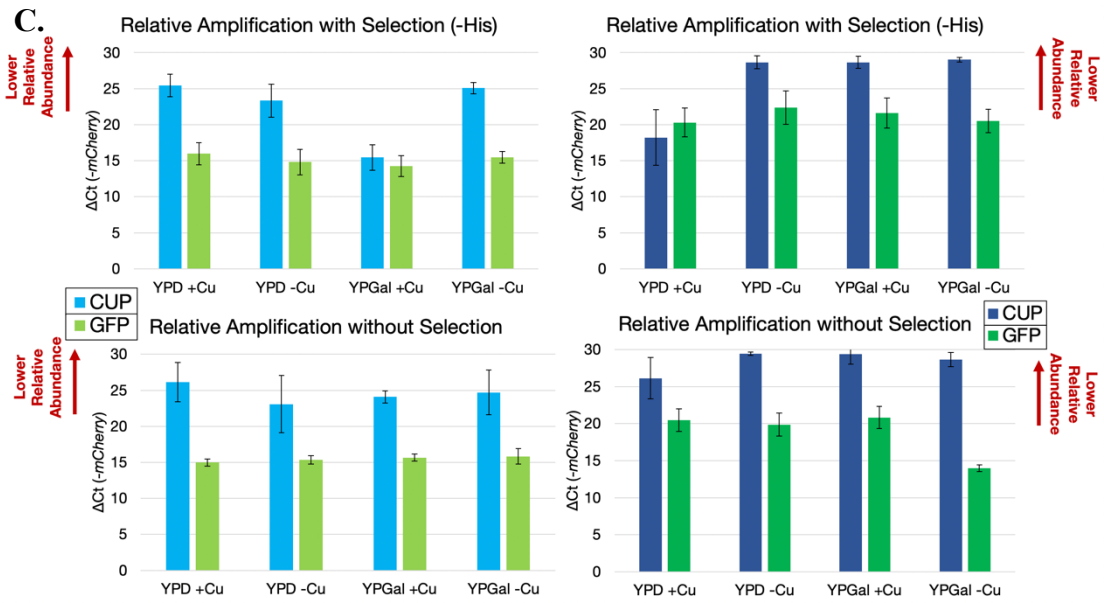


Figure 5. Detecting eccDNA from the *CUP1-GFP* reporter locus. **(A and B).** Average amplification of *ACT, CUP1*, *GFP*, and *mCherry* targets for Trial 1 (three separate runs with technical triplicates, A) and Trial 2 (four runs in triplicate, B). Solid lines denote samples cultured in histidine dropout media selective for the *GFP* cassette, while dotted lines denote non-selective media. Various induction conditions with 1 mM $CuSO_4$ and/or galactose are denoted by color. **(C.)** Amplification of *CUP1* and *GFP* relative to the housekeeping circular gene *mCherry* in selective (-His, top) and non-selective (bottom) media under various induction conditions. Trial 1 is shown in the left panel (light), Trial 2 is shown on the right (dark).

CHAPTER 4

DISCUSSION

CUP1-GFP Reporter is Lost under Non-selective Induction

The ability to accurately detect eccDNA is a commonly reported challenge due to the sequence homology shared by eccDNA and their linear origins, which can result in erroneous identification (Mouakkad-Montoya et al., 2021; Wang et al., 2021). This study sought to develop a robust method for target eccDNA detection, as well as a reporter system for investigating eccDNA formation. The *CUP1* locus was selected for this purpose, as it was previously reported to strongly induce eccDNA formation under exposure to copper (Hull et al., 2019). Furthermore, it was determined to contain multiple endogenous repeats, allowing for facile insertion of a reporter cassette via the integration of homology arms (Supplementary Figure 1; Figure 1).

Following its cloning into the *CUP1* repeat array of *S. cerevisiae*, the *CUP1-GFP* reporter was initially tested in the copper-induction time course experiment shown by Figure 2B. The integrated reporter cassette contains a *His3* auxotrophic marker, which enabled experimental selection for the reporter by culturing cells in histidine dropout media. As expected, cells cultured without selection appeared to lose the *CUP1-GFP* reporter as indicated by a linear decrease in GFP fluorescence. eccDNA from the *CUP1* locus has been proposed to form by homologous recombination between tandem repeats (Hull et al., 2019). Accordingly, it may be possible that copper induction results in excision of the *CUP1-GFP* reporter as eccDNA, which is then randomly lost during cell division. This hypothesis is supported by the apparent inverse correlation between average fluorescence and cell viability observed in the histidine dropout culture, most

prominent after 48 hours post-induction. Together, these results suggest that cells may continue to form eccDNA at the *CUP1* locus regardless of the inclusion of a reporter cassette. Auxotrophic selection for the reporter may reveal a population of cells containing eccDNA from the *CUP1-GFP* locus, or alternatively, inhibit the mechanism of eccDNA formation to prevent chromosomal excision of the cassette.

Circular DNA Enrichment Facilitates Sensitive Detection by qPCR

With evidence of eccDNA being formed at the *CUP1-GFP* reporter locus, it was necessary to devise an approach for sensitively detecting such eccDNA for later characterization. Exonuclease T5 is a highly active enzyme with specificity for linear DNA, but it is described by the manufacturer to also degrade nicked circular DNA (**NEB**). Contrastingly, ExoV (Exonuclease RecBCD) is reported to be highly specific for un-nicked linear dsDNA but requires supplementation with ATP (**NEB**). To optimize the eccDNA enrichment protocol, both enzymes were tested and were qualitatively shown to degrade linear DNA in Figure 3A. Additionally, DNA extraction methods were compared in Figure 3B, which quantitatively shows the superior efficiency of miniprep extraction for isolating circular DNA.

To observe a similar quantitative difference in the activities of exonuclease T5 and ExoV, the ΔCt between exonuclease-treated and -untreated samples was determined for each enzyme with respect to both linear *ACT* and circular *GFP*. Figure 3C shows the ΔCt for *ACT*, which describes the degree of linear DNA degradation by either T5 or ExoV. Although the activity of each enzyme appears comparable in this context, observing the ΔCt of *GFP* instead reveals a difference in their sensitivity towards circular DNA. While both exonucleases gave a positive ΔCt indicating some degree of circular

GFP degradation, this was much more pronounced for T5. This observation is in line with the ability of exonuclease T5 to degrade nicked circular DNA, which may result from MP isolation. As such, later experiments focused on the use of ExoV to degrade linear DNA and preserve circular DNA within a sample.

The final step in circular DNA enrichment implemented by this study was RCA by phi29 polymerase. To assess the specificity of phi29 for circular DNA, RCA reactions were composed using variable primers (Supplementary Figure 2) and either with or without treatment by ExoV (Figure 3D, E). After determining that RCA by phi29 is most efficient with the use of a random hexamer, reactions were performed on ExoV-treated and -untreated samples from both wildtype cells transformed with p415-GFP, and cells of the *CUP1-GFP* reporter strain. Figure 3D shows that amplification ensued across all samples, indicating that phi29 is capable of amplifying linear DNA. However, further quantitative analysis shown in Figure 3E revealed a difference in the abundance of *GFP* from the plasmid versus the linear reporter. The ΔCt was negative for only the p415-GFP samples, regardless of ExoV treatment, which demonstrates a strong preference for circular DNA by phi29. The calculated fold-change of *GFP* abundance between ExoV-treated and -untreated samples additionally revealed a synergistic effect between ExoV treatment and phi29 RCA, optimal for enriching circular DNA.

Circular DNA Enrichment Assay Enables Detection of eccDNA

To observe the production of *CUP1* eccDNA in yeast induced with copper, the optimized assay was performed in wildtype yeast with p415-GFP. Figure 4A and 4B quantitatively show the enrichment of circular DNA in the samples over each of the previously described steps. As expected, *ACT* was initially less abundant than both *GFP*

and *CUP1*, and its abundance decreased with each circular DNA enrichment step. While *CUP1* also seemed to decrease in abundance with each enrichment step, the dramatic amplification of *GFP* shown in Figure 4A accounts for these increasing ΔCt values. Despite the decrease in *CUP1* abundance relative to *GFP* following RCA, a significant cycle difference was revealed between the copper-induced and un-induced samples suggesting the presence of *CUP1* eccDNA. Furthermore, the fold-change in Figure 4C was observed to increase with each successive step, both confirming the utility of the described enrichment assay as well as providing further evidence for *CUP1* eccDNA.

Surprisingly, whole genome sequencing of the enriched samples did not yield any sequence reads from the *CUP1* locus in the copper-induced sample, but rather a few low-confidence reads in the un-induced sample (Figure 4D). It is interesting that no significant *CUP1* eccDNA was detected while other loci known to exist on circular DNA were highly represented (Supplementary Figure 4). The S288C *GAL1* promoter and *LEU2* gene, both also present on p415-GFP, showed high sequencing coverage in both samples, as did the rDNA locus on Chr. XII which is known to form ERCs (Supplementary Figure 4; Prada-Luengo et al., 2020). Furthermore, each sample revealed significant sequencing alignments to various loci that are both previously described and undescribed in the literature as capable of forming eccDNA. Of these, eccDNA from the *ENA1/2/5* tandem repeat locus was previously described in yeast to contain an ARS, and Figure 4E reveals the presence of this eccDNA in the copper-induced sample (Møller et al., 2016). Interestingly, no sequencing reads from the un-induced sample mapped to the *ENA1/2/5* locus. The ENA proteins are highly similar ion transport ATPases, thus the

cellular stress conferred by copper induction may have selected for a copy number increase of these genes to remediate toxicity in the cell.

The un-induced sample contained many sequencing reads which mapped to telomeres, such as that of Chr. I (Supplementary Figure 4). Telomeres were previously shown to commonly form eccDNA as a result of internal loops which form at points of DNA damage (Mazzucco et al., 2020). Sequences from another known circular DNA, the mitochondrial genome, were significantly represented only in the copper-induced sample, (Supplementary Figure 4). These results confirm the ability of the circular DNA enrichment protocol to reveal populations of eccDNA, however the discrepancy in detection of *CUP1* eccDNA between qPCR and sequencing suggests deficiencies in the ability of the assay to specifically identify a target eccDNA population. Other notable loci not previously reported to form eccDNA but detected in the copper-induced sample were the cell wall stress transducer *WSC3* and the DNA Ligase IV gene, which has been implicated in eccDNA formation in mice (Supplementary Figure 4; Cohen et al., 2006). Finally, sequences from the hexose transporter genes *HXT5* and *HXT8* were well-represented in the un-induced sample, which supports previous reports of other paralogous hexose transporter genes, namely *HXT6/HXT7*, forming eccDNA in yeast (Supplementary Figure 4; Møller et al., 2016; Prada-Luengo et al., 2020).

Inconclusive Detection of eccDNA from *CUP1-GFP* Locus

Application of the eccDNA enrichment and detection qPCR assay to the *CUP1-GFP* reporter locus yielded unanticipated results which conflict with the basic assumptions made by the experiment. First, the construction of the chromosomal *GFP* reporter by integration between tandem *CUP1* repeats implies that, at a maximum, *CUP1*

should be twice as abundant as *GFP* in the cell. Both trials of the experiment consistently showed a greater abundance of *GFP* than *CUP1* relative to the housekeeping gene *mCherry*. Additionally, cells cultured in histidine dropout media were assumed to contain the *GFP* cassette either chromosomally or on eccDNA. However, in both trials *GFP* was detected slightly earlier in non-selective conditions than under selective pressure to retain the cassette. This result was more prominent in Trial 2, wherein a significant increase in *GFP* as a result of galactose induction was observed primarily in non-selective culture. One explanation that may account for these results is a potential bias of phi29 for plasmids like p415 or 2μ over other eccDNA in the sample. Sequencing of the samples shown in Figure 4 demonstrated that only 2% of reads aligned to the S288C genome, while around 30% of reads from each sample mapped to either plasmid. Phi29 has previously been shown to exhibit template-dependent biases when performing RCA on small DNA circles, thus a similar preferential mechanism may cause an underrepresentation of some eccDNA in the final enriched sample (Joffroy et al., 2018).

An interesting difference between the two trials was the induction condition necessary to form *CUP1* eccDNA at the reporter locus. While Trial 1 demonstrated a requirement for both copper and galactose, double-induction instead inhibited *CUP1* eccDNA formation in Trial 2. Although induction and circular DNA enrichment were performed following the same protocol in both trials, the sub-nanogram quantity of DNA transferred to phi29 RCA reactions following linear DNA digestion could result in random variability between trials. In the future, the enrichment protocol should be optimized for larger starting quantities of eccDNA by extending the exonuclease digestion step, or by the physical separation of circular DNA through the use of magnetic

beads (Mouakkad-Montoya et al., 2021). A final and likely possibility explaining the unexpectedly lesser abundance of *CUP1* than *GFP* in this experiment is the potential inhibition of *CUP1* eccDNA formation due to the large, intervening *GFP* cassette. Because *CUP1* eccDNA likely forms by a homology-mediated mechanism, separation of the tandem repeats by ~6 kb may disrupt its formation (Hull et al., 2019). To investigate this possibility further the cassette should be adjusted to remove unnecessary sequences from the original cloning vector, such as the *URA3* buffer and *AmpR* gene.

CHAPTER 5

CONCLUSION

Summary

The purpose of this study was to develop an approach for investigating the formation of eccDNA in yeast, both by manipulating an endogenous eccDNA-forming locus and detecting suspected eccDNA in a sample. A galactose inducible *GFP* reporter was constructed within the *CUP1* repeat array, and initial flow cytometry of cells induced with both copper and galactose showed an apparent accumulation of the reporter in cells under selective pressure. To investigate whether this was the result of eccDNA formation at the locus, a circular DNA enrichment and detection assay was designed. Upon its initial application to wildtype yeast the qPCR assay revealed a distinct population of *CUP1* eccDNA in copper-induced cells, however *CUP1* was significantly underrepresented in the aligned sequence reads of these samples. Despite this, other known eccDNAs were identified in the sequence reads from either sample, demonstrating the utility of the circular DNA enrichment protocol.

Application of the enriched eccDNA qPCR assay to the *CUP1-GFP* reporter strain yielded inconclusive results that varied significantly between two separate induction trials. Additionally, the unexpectedly greater abundance of *GFP* than *CUP1* relative to the housekeeping circular DNA suggests a flaw in the initial design or assumptions of the experiment. To rectify these, future efforts should be made towards optimizing the reporter cassette to make it more amenable to eccDNA formation. In addition to reducing the size of the cassette itself to restore homology-mediated eccDNA formation between repeats of *CUP1*, inclusion of other elements on the cassette such as

an ARS may be considered to ensure retainment of eccDNA from the locus. While the circular DNA enrichment protocol was sufficient for identifying various populations of eccDNA by sequencing, more optimization at the point of RCA should be done to ensure greater accuracy in representing all eccDNA populations in the final enriched sample.

Future Perspectives

Elucidating the dynamics which govern eccDNA formation in cells under selective pressure will be necessary for determining the mechanism of eccDNA-mediated tumor evolution and aggression (Robert and Crasta, 2022). At present, only two studies have been published which use differing fluorescent reporter approaches to track eccDNA in the cell and observe segregation behaviors (Yi et al., 2022; Lange et al., 2021). Optimizing the *CUP1-GFP* reporter described in this study presents an opportunity to develop the first fluorescent reporter that instead measures the rate of eccDNA formation at a locus under selective pressure. Gaining a better understanding of these formation events may provide a target for therapeutic intervention at the point of inhibiting *de novo* eccDNA formation. Alternative approaches proposed for therapeutically targeting ecDNA in tumors include inhibiting chromosomal reintegration and interfering with intracellular ecDNA hubs (van Been et al., 2022). In addition to measuring formation events, the fluorescent reporter described here may also be optimized for visualizing such alternative eccDNA functions. In one approach called CIRSPR-C, targeted DNA breaks were induced at a reporter locus such that formation of circles were indicated by reporter activation (Møller et al., 2018). This technique could be modified to instead report on eccDNA reintegration, and furthermore, can be broadly applied to investigate the limits of eccDNA formation and load in the cell.

eccDNA is relevant not only to cancer but has also become implicated in other conditions such as neurodegeneration and osteoarthritis (Ain et al., 2020; Xiang et al., Preprint). Furthermore, novel roles for eccDNA in the maintenance of healthy cells are also being proposed. In one example, eccDNA was shown to be produced by apoptotic cells and stimulate an innate immune response (Wang et al., 2021). In another, eccDNA in healthy human cells were shown to produce regulatory microRNAs that are capable of repressing endogenous genes (Paulsen et al., 2019). Despite our increasing knowledge of potential eccDNA functions, we still lack a basic comprehension of eccDNA formation mechanisms and the dynamics which govern eccDNA formation in cells under selective pressure. Resolving these questions will be necessary for explaining various phenomena reported in the literature, such as eccDNA accumulation in aged yeast cells or the spread of drug resistance on ecDNA in tumors.

CHAPTER 6

SUPPLEMENTAL

Supplementary Methods

CUP1 *locus validation*: The Chr. VIII *CUP1* locus of *S. cerevisiae* 834 was assayed by PCR to confirm the presence of a tandem repeat. PCR reactions were performed as previously described with the substitution of HF buffer for GC buffer, and the CUP1a and CUP1b primer pairs shown in Supplementary Table 1 were used. PCR products were stained with 1X Purple Gel Loading Dye (NEB, B7024) and separated by gel electrophoresis using 1% agarose and 1X TAE at 100 V for an hour.

Phi29 Primer Control: RCA reactions were composed as previously described, instead using Exo-resistant Random Hexamer and the various qPCR forward primers shown in Supplementary Table 1. One ng of genomic DNA extracted from un-induced 834 p415-GFP cells was added to each reaction without prior exonuclease treatment. The universal primer was used at a final concentration of 25 μM, while target-specific primers were added at a final concentration of 10 μM. Samples were analyzed by gel electrophoresis and qPCR as previously described.

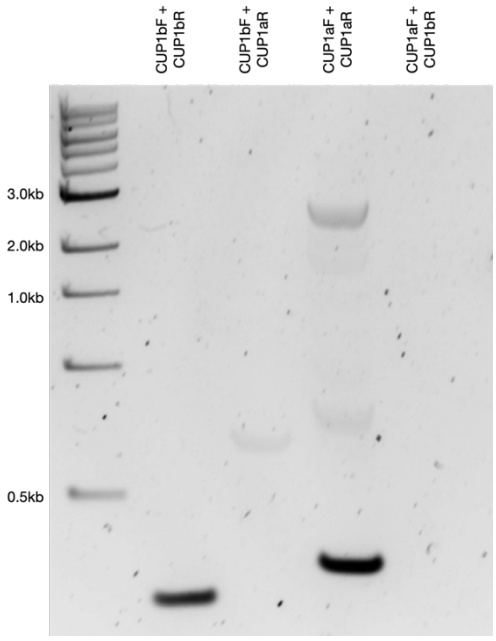
Supplementary Table 1

PCR Primers Used in this Study

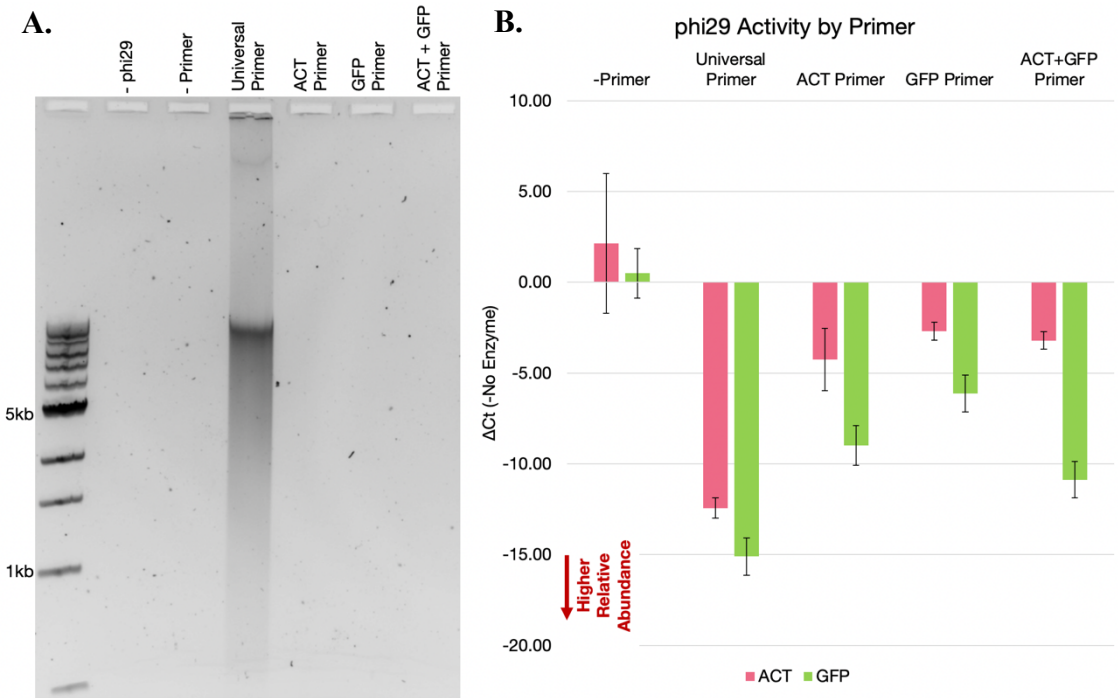
Supplementary Table 1. Lists the primer name, use in this study, sequence, annealing temperature, and predicted amplicon size for all primer pairs used. For studying the *CUP1* repeat locus in Supplementary Figure 1, multiple primer combinations were used and are listed accordingly.

PRIMER	**USE**	**SEQUENCE**	**T_a (°C)**	**POTENTIAL AMPLICONS (bp)**
CUP1aF	*CUP1* locus structure	GAAAGTGACGGGGATAACAGC	aR: 60 bR:61	aR: 288, 2286 bR: 4154
CUP1aR	*CUP1* locus structure	GGAGAAGGAAGTTTAATCGAC	aF and bF: 60	aF: 288, 2286 bF: 371
CUP1bF	*CUP1* locus structure	GGCAAGTAGAAAGGAACACC	bR: 61 aR: 60	bR: 24, 2239 aR: 371
CUP1bR	*CUP1* locus structure	GCTGTCTGTTAAGCTTCCAG	aF and bF: 61	bF: 241, 2239 aF: 4154
CUP1a_SpeI	*CUP1-GFP* cloning	ATGAATACTAGTTTACTCACGCCCATAGTCAAGGT	66	188
CUP1a_EcoRI	*CUP1-GFP* cloning	ATCATGAATTCAACCTATCTCCGATACCTGCCTC	66	188
CUP1b_SacI	*CUP1-GFP* cloning	ATCATGAGCTCTTCAGCAGCGGGTACCATGAAT	66	231
CUP1b_SpeI	*CUP1-GFP* cloning	ATGAATACTAGTCGACGATTCTACTCAGTTTGAGTACAC	66	231
mCherry_BamHI	p415-mCherry cloning	ATAGTGGATCCATGGTGAGCAAGGGCGAGGA	61	730
mCherry_XhoI	p415-mCherry cloning	ACTATCTCGAGTTACTTGTACAGCTCGTCCA	61	730
qACTF	qPCR	CTGCCGGTATTGACCAAACT	60	144
qACTR	qPCR	CGGTGATTTCCTTTTGCATT	60	144
qCUPF	qPCR	CTGCCATGGACAAAGAGGAT	60	154
qCUPR	qPCR	ACCTTGACTATGGGCGTGAG	60	154
qGFPF	qPCR	TTGGTGATGGTCCAGTCTTG	60	132
qGFPR	qPCR	CATGGGTAATACCAGCAGCA	60	132

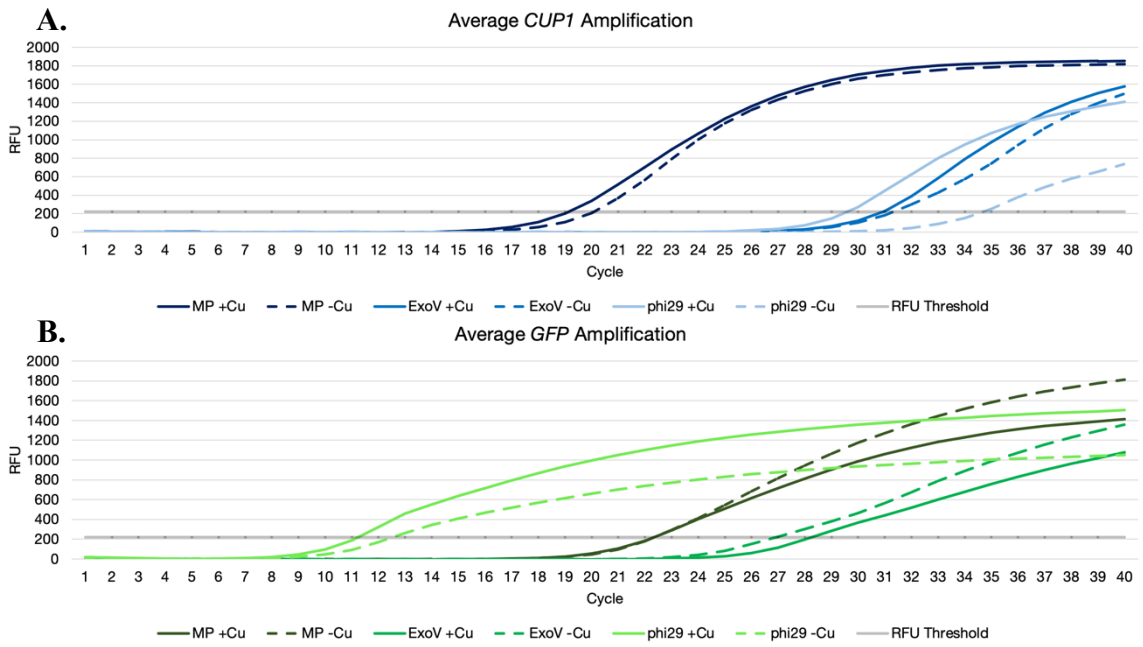
Supplementary Figures



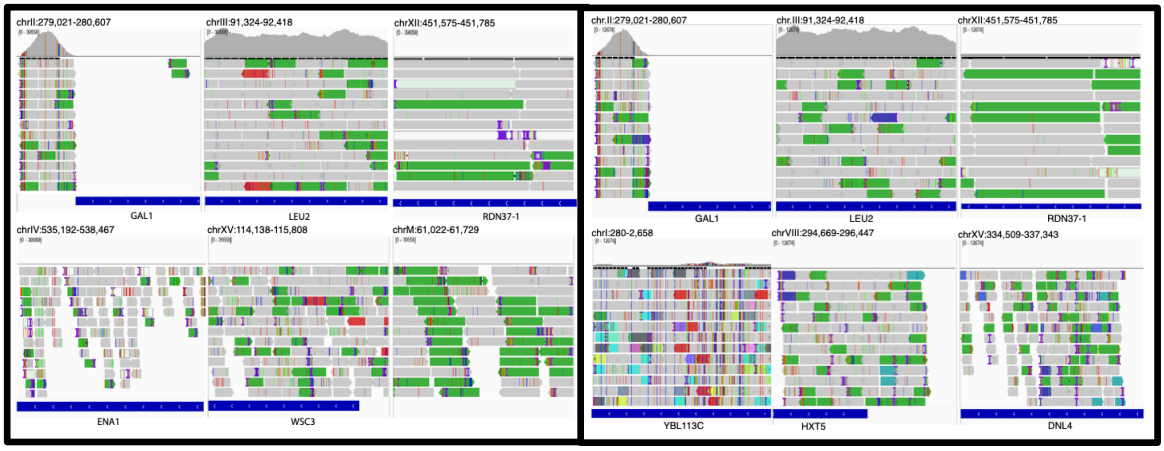
Supplementary Figure 1. PCR amplicons from the *CUP1* repeat locus of *S. cerevisiae* 834. Different combinations of the *CUP1* primers shown in Supplementary Table 1 were used to confirm the presence of at least two *CUP1* repeats in tandem. The 2286 bp *CUP1A* repeat was achieved and is present in Lane 3, while other combinations yielded smaller, unintended amplicons of 241 bp and 288 bp. The >4 kb double-repeat was consistently not achieved by PCR amplification, indicated by its empty lane.



Supplementary Figure 2. phi29 primer control. **A.** Gel image illustrating the activity of phi29 on 834 p415-GFP whole genomic DNA in reactions using various primers. From left to right: no enzyme, no primer, random hexamer, *ACT*-specific primer, *GFP*-specific primer, and both *ACT*- and *GFP*-specific primers. **B.** Samples from A were used in a qPCR assay measuring the relative abundance of linear *ACT* and circular *GFP*. The Ct values for each sample were normalized to the no-enzyme control to give the plotted ΔCt.



Supplementary Figure 3. Enriched circular DNA samples from copper-induced 834 p415-GFP cells. Average *CUP1* **(A)** and *GFP* **(B)** amplification curves from the experiment described in Figure 4.



Supplementary Figure 4. Sequencing coverage and reads from copper-induced (left) and un-induced (right) enriched samples, which mapped to various loci in the S288C reference genome. The top row displays the control loci GAL1 and LEU2 which were present in the sample also on the circular plasmid p415-mCherry. RDN37-1 represents an rDNA locus that is commonly circularized in yeast. The bottom row shows various loci with mapped reads that were unique to each sample. For the copper-induced sample, some of these were the ATPase *ENA1*, the cellular stress transducer *WSC3*, and non-coding mtDNA. For the un-induced sample, some examples were telomeric sequences like *YBL113C*, the hexose transporter *HXT5*, and the DNA ligase IV gene.

REFERENCES

Ain, Q., Schmeer, C., Wengerodt, D., Witte, O.W., and Kretz, A. (2020). Extrachromsomal circular DNA: current knowledge and implications for CNS aging and neurodegeneration. Int. J. Mol. Sci. *21*, 2477.

Brewer, B.J., Payen, C., Di Rienzi, S.C., Higgins, M.M., Ong, G., Dunham, M.J., and Raghuraman, M.K. (2015). Origin-dependent inverted-repeat amplification: tests of a model for inverted DNA amplification. PLoS Genet. *11*, e1005699.

Cohen, S. and Mechali, M. (2001). A novel cell-free system reveals a mechanism of circular DNA formation from tandem repeats. Nucleic Acids Res. *29*, 2542-2548.

Cohen, Z., Bacharach, E., and Lavi, S. (2006). Mouse major satellite DNA is prone to eccDNA formation via DNA Ligase IV-dependent pathway. Oncogene *25*, 4515-4524.

Fogel, S. and Welch, J.W. (1982). Tandem gene amplification mediates copper resistance in yeast. Proc. Natl. Acad. Sci. *79*, 5342-5346.

Henriksen, R.A., Jenjaroenpun, P., Sjøstrøm, I.B., Jensen, K.R., Prada-Luengo, I., Wongsurawat, T., Nookaew, I., and Regenberg, B. (2022). Circular DNA in the human germline and its association with recombination. Mol. Cell *82*, 209-217.

Hotta, Y. and Bassel, A. (1965). Molecular size and circularity of DNA in cells of mammals and higher plants. Proc. Natl. Acad. Sci. USA *53*, 356-362.

Hull, R.M. and Houseley, J. (2020). The adaptive potential of circular DNA accumulation in ageing cells. Curr. Genet. *66*, 889-894.

Hull, R.M., King, M., Pizza, G., Krueger, F., Vergara, X., and Houseley, J. (2019). Transcription-induced formation of extrachromosomal DNA during yeast ageing. PLoS Biol. *17*, e3000471.

Joffroy, B., Uca, Y.O., Prešern, D., Doye, J.P.K., and Schmidt, T.L. (2018). Rolling circle amplification shows a sinusoidal template length-dependent amplification bias. Nucleic Acids Res. *46,* 538-545.

Kim, H., Nguyen, N., Turner, K., Wu, S., Gujar, A.D., Luebeck, J., Liu, J., Deshpande, V., Rajkumar, U., Namburi, S., et al. (2020). Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. Nat. Genet. *52*, 891-897.

Koche, R.P., Rodriguez-Fos, E., Helmsauer, K., Burkert, M., MacArthur, I.C., Maag, J., Chamorro, R., Munoz-Perez, N., Puiggròs, M., Garcia, H.D., et al. (2019).

Extrachromosomal circular DNA drives oncogenic genome remodeling in neuroblastoma. Nat. Genet. *52*, 29-34.

Koo, D., Molin, W.T., Saski, C.A., Jiang, J., Putta, K., Jugulam, M., Friebe, B., and Gill, B.S. (2018). Extrachromosomal circular DNA-based amplification and transmission of herbicide resistance in crop weed Amaranthus palmeri. Proc. Natl. Acad. Sci. USA *115*, 3332-3337.

Lange, J.T., Chen, C.Y., Pichugin, Y., Xie, L., Tang, J., Hung, K.L., Yost, K.E., Shi, Q., Erb, M.L., Rajkumar, U., et al. (2021). Principles of ecDNA random inheritance drive rapid genome change and therapy resistance in human cancers. BioRxiv, Preprint.

Langmead, B., Salzburg, S.L., (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9,* 357-359.

Liao, Z., Jiang, W., Ye, L., Li, T., Yu, X., and Liu, L. (2020). Classification of extrachromosomal circular DNA with a focus on the role of extrachromosomal DNA (ecDNA) in tumor heterogeneity and progression. Biochem. Biophys. Acta Rev. Cancer *1874*, 188392.

Mazzucco, G., Huda, A., Galli, M., Piccini, D., Giannattasio, M., Pessina, F., and Doksani, Y. (2020). Telomere damage induces internal loops that generate telomeric circles. Nat. Commun. *11* 5297.

Molin, W.T., Yaguchi, A., Blenner, M., and Saski, C.A. (2020). Autonomous replication sequences from the Amaranthus palmeri eccDNA replicon enable replication in yeast. BMC Res. Notes *13*, 330.

Molin, W.T., Yaguchi, A., Blenner, M., and Saski, C.A. (2020). The eccDNA replicon: a heritable, extranuclear vehicle that enables gene amplification and glyphosate resistance in Amaranthus palmeri. Plant Cell *32*, 2132-2140.

Møller, H.D., Bojsen, R.K., Tachibana, C., Parsons, L., Botstein, D., and Regenberg, B. (2016). Genome-wide purification of extrachromosomal circular DNA from eukaryotic cells. J. Vis. Exp. *110*, e54239.

Møller, H.D., Mohiyuddin, M., Prada-Luengo, I., Sailani, M.R., Halling, J.F., Plomgaard, P., Maretty, L., Hansen, A.J., Snyder, M.P., Pilegaard, H., et al. (2018). Circular DNA elements of chromosomal origin are common in healthy human somatic tissue. Nat. Commun. *9*, 1069.

Møller, H.D., Parsons, L., Jørgensen, T.S., Botstein, D., and Regenberg, B. (2015). Extrachromosomal circular DNA is common in yeast. Proc. Natl. Acad. Sci. USA *112*, 3114-3112.

Møller, H.D., Ramos-Madrigal, J., Prada-Luengo, I., Gilbert, M.T.P., and Regenberg, B. (2019). Near-random distribution of chromosome-derived circular DNA in the condensed genome of pigeons and the larger, more repeat-rich human genome. Genome Biol. Evol. *12*, 3762-3777.

Mouakkad-Montoya, L., Murata, M.M., Sulovari, A., Suzuki, R., Osia, B., Malkova, A., Makoto, K., Giuliano, A.E., Eichler, E.E., and Tanaka, H. (2021). Quantitative assessment reveals the dominance of duplicated sequences in germline-derived extrachromosomal circular DNA. Proc. Natl. Acad. Sci. USA *118*, e2102842118.

Nathanson, D.A., Gini, B., Mottahdeh, J., Visnyei, K., Koga, T., Gomez, G., Eskin, A., Hwang, K., Wang, J., Masui, K., et al. (2014). Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. Science *343*, 72-76.

Paulsen, T., Malapti, P., Shibata, Y., Wilson, B., Eki, R., Benamar, M., Abbas, T., and Dutta, A. (2021). MicroDNA levels are dependent on MMEJ, repressed by c-NHEJ pathway, and stimulated by DNA damage. Nucleic Acids Res. *49*, 11787-11799.

Paulsen, T., Shibata, Y., Kumar, P., Dillon, L., and Dutta, A. (2019). Small extrachromosomal circular DNAs, microDNA, produce short regulatory RNAs that suppress gene expression independent of canonical promoters. Nucleic Acids Res. *47*, 4586-4596.

Prada-Luengo, I., Møller, H.D., Henriksen, R.A., Gao, Q., Larsen, C.E., Alizadeh, S., Maretty, L., Houseley, J., and Regenberg, B. (2020). Replicative aging is associated with loss of genetic heterogeneity from extrachromosomal circular DNA in Saccharomyces cerevisiae. Nucleic Acids Res. *48*, 7883-7898.

Robert, M. And Crasta, K. (2022). Breaking the vicious circle: extrachromosomal circular DNA as an emerging player in tumor evolution. Semin. Cell Dev. Biol. *123*, 140-150.

Shoshani, O., Brunner, S.F., Yaeger, R., Ly, P., Nechemia-Arbely, Y., Kim, D.H., Fang, R., Castillon, G.A., Yu, M., Li, J.S.Z., et al. (2021). Chromothripsis drives the evolution of gene amplification in cancer. Nature *591*, 137-141.

Stanfield, S.W. and Lengyel, J.A. (1979). Small circular DNA of Drosophila melanogaster: chromosomal homology and kinetic complexity. Proc. Natl. Acad. Sci. USA *76*, 6142-6146.

Storlazzi, C.T., Fioretos, T., Surace, C., Lonoce, A., Mastrorilli, A., Strömbeck, B., D'Addabbo, P., Iacovelli, F., Minervini, C., Aventin, A., et al. (2006). MYC-containing double minutes in hematologic malignancies: evidence in favor of the episome model and exclusion of MYC as the target gene. Hum. Mol. Genet. *15*, 933-942.

Sunnerhagen, P., Sjöberg, R., Karlsson, A., Lundh, L., and Bjursell, G. (1986). Molecular cloning and characterization of small poly disperse circular DNA from mouse 3T6 cells. Nucleic Acids Res. *14*, 7823-7838.

van Leen, E., Brückner, L., and Henssen, A.G. (2022). The genomic and spatial mobility of extrachromosomal DNA and its implications for cancer therapy. Nat. Genet. *54*, 107-114.

Wang, Y., Wang, M., Djekidel, M.N., Chen, H., Liu, D., Alt., F.W., and Zhang, F. (2021). eccDNAs are apoptotic products with high innate immunostimulatory activity. Nature *599,* 308-314.

Xiang, W., Zhang, T., Li, S., Gong, Y., Luo, X., Yuan, J., Wu, Y., Yan, X., Xiong, Y., Lian, J., et al. (2022). Cir-DNA sequencing revealed the landscape of extrachromosomal circular DNA in articular cartilage and the potential roles in osteoarthritis. BioRxiv, Preprint.

Yi, E., Gujar, A.D., Guthrie, M., Kim, H., Zhao, D., Johnson, K.C., Amin, S.B., Costa, M.L., Yu, Q., Das, S., et al. (2022). Live-cell imaging shows uneven segregation of extrachromosomal DNA elements and transcriptionally active extrachromosomal DNA hubs in cancer. Cancer Discov. *12*, 468-483.