Causation and Consequences of Polyploidy in Cyanobacteria, *Synechococcus* sp. PCC 7002

Jillian Frank

Honors Thesis

Department of Molecular, Cellular, and Developmental Biology

University of Colorado, Boulder

5 April 2022

Thesis Advisor:

Dr. Kristin Moore, Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder

Thesis Committee:

Dr. Jeffrey Cameron, Department of Chemistry and Biochemistry, University of Colorado, Boulder, Renewable and Sustainable Energy Institute, University of Colorado, Boulder, National Bioenergy Center, National Renewable Energy Laboratory, Golden, Colorado

Dr. Joel Kralj, Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder

Dr. Andrea Feldman, Program of Writing and Rhetoric, University of Colorado, Boulder

Table of Contents

Abstract

The number of chromosome copies within an organism can directly impact gene expression, phenotype, and offspring viability. The presence of multiple copies of chromosomes, known as polyploidy, may confer additional consequences that remain unknown. It is important to understand the driving mechanisms behind chromosome replication so as to maintain the integrity of gene expression and inheritance of those genes and to avoid the retention of deleterious mutations. Currently, most model prokaryotic organisms are monoploid, therefore the effects of multiple chromosome copies in prokaryotes remains understudied as does the chromosome replication mechanisms for polyploid microbes. The cyanobacterium, *Synechococcus* sp. PCC 7002 is a strong candidate as a model organism particularly for use in synthetic biology and biofuel production based on its photosynthetic abilities. In this study, through CRISPR-interference gene knockdown of possible chromosome replication initiation factors as well as conferring heterogeneity of antibiotic resistance at a single, competing locus, we aim to elucidate both possible mechanisms and consequences of polyploidy in a potential model prokaryotic organism. We find that two conserved genes, *priA* and *recG* have no major phenotypic effect on ploidy level, and likely are not key regulators in chromosome replication. We also find that in the absence of selective pressure, there is random gene elimination to a homogenous allelic state, suggesting that heterogenetic polyploidy may not be a naturally occurring phenomenon or benefit in polyploid prokaryotes. Further investigation is required to confirm the definitive functions and significance of *priA* and *recG* in the chromosome replication pathway, as well as if gene elimination always occurs in a heterozygotic polyploid cyanobacteria in the absence of selection.

Acknowledgements

First and foremost, a huge thank you is garnered for Dr. Kristin Moore, Instructor of Cell Biology Laboratory in the Molecular, Cellular, and Developmental Biology (MCDB) department at the University of Colorado, Boulder and my thesis advisor for this work. Without Dr. Moore's time, expert knowledge, ingenuity and passion for the life sciences this work would not have been brought to fruition. It was a privilege to work under her guidance and mentorship this last year navigating the reality of research. The research opportunity she provided drastically enhanced my academic experience during otherwise unprecedented times, and the knowledge and hands-on skills I now take with me will undoubtedly prove both invaluable and influential to my future in the life sciences, whether as a veterinarian or MCDB graduate student. It is with true appreciation that I thank and acknowledge Dr. Kristin Moore's vital role in my undergraduate experience, and I contribute a significant portion of my success in this field to her mentorship. Future thesis students of the Moore Lab should consider themselves lucky.

I also thank Dr. Jeffrey Cameron, Institute Fellow and Assistant Professor in the Biochemistry Department at the University of Colorado, Boulder for the provision of critical resources such as cyanobacteria specimens, templates, and primers also without which, this work would not have been possible.

I would also like to acknowledge Dr. Andrea Feldman, Senior Professor in the Program of Writing and Rhetoric. Her class, Honor's Writing, initially sparked my interest in writing a thesis, and without having taken that class, I cannot say I would have sought this opportunity.

Additionally, two entities at the University of Colorado Boulder were huge assets in data collection and analysis: BioCore and the MCDB department. BioCore provided free access to the qPCR machine, a vital tool in measuring allelic ratios and obtaining useful data. The MCDB department provided all the remaining tools, space and funding required to execute this research. I am incredibly grateful to have been provided the opportunity and resources to contribute research to the scientific community made possible by these departments and individuals.

Literature Review

Prokaryotic Polyploidy

Polyploidy is a genetic quality of an organism in which it possesses multiple complete (mostly identical) copies of its chromosome(s). The condition has been observed in many plants, fish, and amphibians where its effects and consequences have been thoroughly studied. Evidence from genome analyses revealed that many eukaryotic organisms have arisen from polyploid ancestry, suggesting that most eukaryotes have experienced polyploidy (Comai, 2005). Where research is lacking is in the study of polyploid prokaryotes, in particular cyanobacteria, and the possible consequences it may confer to them both in metabolism and DNA regulation. Studying the causation and consequences of polyploidy in cyanobacteria may offer deeper insight into their evolution, both past and future. These insights could reveal novel mechanisms of chromosome regulation and whether polyploidy could provide benefits in using it in industrial applications. With this insight, horizons may broaden in terms of the available, manipulatable, and predictable model prokaryotic microbes for use in industrial and synthetic biology.

The degree of ploidy in cyanobacteria is flexible and depends on cell culture density and growth conditions (Watanabe, 2020). Thus there is an ideal cell phase in which to study ploidy. Four phases of the cyanobacterium growth cycle influences chromosome count: lag, log, and linear, and stationary phase. Lag phase consists of stationary and reproductively dormant cells adjusting to their environment. The duration of lag phase depends on how long it takes bacteria to acclimate to their environment. After lag phase cells enter log phase or “logarithmic” phase which is characterized by rapid or exponential cell doubling. Log phase is followed by linear phase in which both cell growth and polyploidy are less extreme than in log phase due to nutrient depletion and light limitation caused by cell-cell shading (Watanabe, 2020). Thus, the ideal timing for studying ploidy in cyanobacteria is log phase when cell reproduction and chromosome replication are highest.

To address gaps in knowledge surrounding the mechanisms that regulate cyanobacteria polyploidy, our study focused on the cyanobacterium *Synechococcus* sp. PCC 7002 (hereafter PCC 7002). PCC 7002 has been an increasing subject of interest

in biotechnical applications because of its euryhaline (salinity-tolerant) properties, high-light irradiation tolerance and rapid growth (Ludwig & Bryant, 2012). Additionally, as a marine strain, it does not require freshwater, offering an extreme advantage in cost-effective biofuel production (Ruffing et al., 2016). It is also polyploid, possessing an average of 2-5 circular chromosomes per individual cell (dependent on cell phase) with a genome size of 3.0 Mbp (Cassier-Chauvat et al., 2016). PCC 7002 is a good candidate for studying gene expression because it is naturally transformable and its genome has been fully sequenced, therefore, its predicted genes and their functions are available.

DNA Replication in Prokaryotes

Accurate and consistent replication and segregation of genomic DNA are fundamental processes of not only polyploid organisms, but across all kingdoms of life to ensure life continues and heritable traits are passed faithfully through generations (Moore et al., 2019). Most studies surrounding polyploidy in cyanobacteria are based on chromosome replication mechanisms conserved in monoploid model bacterial species such as *E. coli*. These studies focus on the importance of *dnaA* and *oriC* in chromosome replication initiation as well as strict regulation of replication timing and synchrony. In *E. coli*, all replication initiation occurs in sync at a fixed cell size that is independent of cell growth. This mechanism is tightly coupled with cell division to ensure appropriate chromosome distribution to daughter cells (Chen et al., 2012). However, these mechanisms do not seem to be as important in polyploid species, given that the presence of multiple chromosomes increases the statistical likelihood that any cell division event will result in daughter cells receiving at least one chromosome copy (Chen et al., 2012). Previous studies on the cyanobacterium, *Synechococcus* sp. PCC 7942 by Chen et al., (2012) found chromosome duplication is asynchronous and correlates with cell length, not cell division.

Similar findings for a related cyanobacterium, PCC 7002 were observed by Moore et. al. (2019). The relationship between chromosome number to cell size ratio was analyzed via manipulation of environmental growth conditions and supported the Chen study in correlating chromosome number to cell length. The study also examined

the effects of chromosomal mis-segregation and depletion through CRISPR interference of genes responsible for cell division and chromosome replication: *misD, ftsZ*, *dnaX* and *dnaA*. Using CRISPRi these genes were targeted for knockdown and the effects were examined via single-cell time lapse microscopy. *dnaX* encodes for a component of the DNA polymerase holoenzyme, the molecular complex responsible for synthesizing the nascent ssDNA replicate from the parent strand during DNA replication (Blinkova et al., 1993). Through the inhibition of chromosome replication via *dnaX* knockdown, chromosome count within individual cells were diluted over time. They found that decreasing chromosome number and/or complete chromosome depletion via *dnaX* knockdown still yielded functional daughter cells in following generations. Extended chromosome depletion did eventually lead to cell division inhibition and eventual growth arrest which was likely due to limited chromosome availability for gene expression required for cell viability (Moore et al., 2019). The immediate generation of chromosome-depleted daughter cells could have been utilizing the gene products that were efficiently expressed by the initial high levels of ploidy and redundant DNA. In conjunction with revealing expected phenotype on ploidy levels when replication machinery are targeted for gene knockdown, this study also proves that CRISPR-interference is a suitable option for manipulating chromosome number in cyanobacteria.

In Monoploid bacteria with circular chromosomes, replication is initiated at a unique origin site (*oriC*) and proceeds bidirectionally via DnaA-dependence to the termination site (*ter*) (Ohbayashi et al., 2016). DnaA is an initiation factor that binds *oriC*, initiating DNA unwinding and drawing replication machinery to the DNA. It was believed to be a conserved critical function across all prokaryotes until Ohbayashi et al., (2016) revealed a varying degree of dependency on DnaA chromosome replication initiation. The study traced the phylogeny of *dnaA* in cyanobacteria, noting strains and species in which the encoding sequence of *dnaA* is entirely absent. Additionally, three model strains of cyanobacteria were selected for gene knockout of *dnaA*: *Synechococcus* sp. PCC 7942, *Synechocystis* sp. PCC 6803 and *Anabaena* sp. PCC 7120 (hereafter PCC 7942, PCC 6803 and PCC 7120). PCC 7942 was found to be DnaA-dependent, but adaptable to *dnaA* deletion by plasmid integration, with the

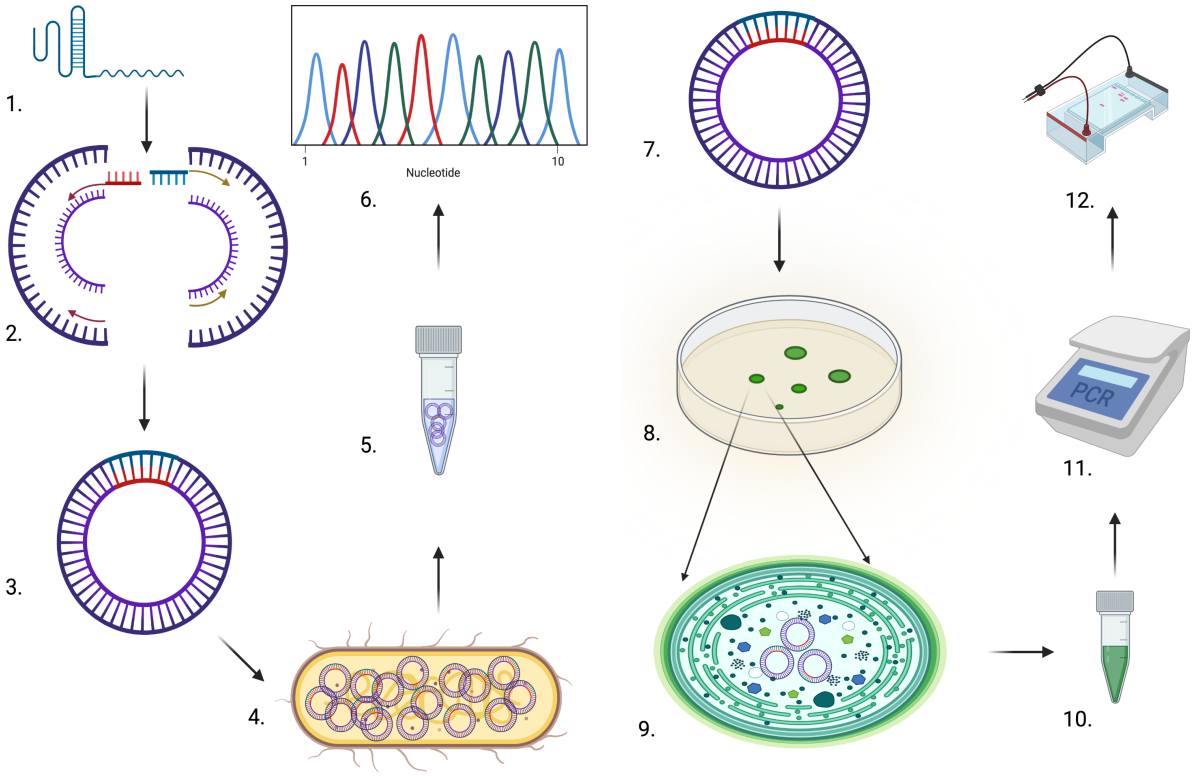
long-term effects of *dnaA* deletion proving advantageous in culturing conditions. PCC 6803 and PCC 7120 were found to be DnaA-independent and sequence analysis suggested the presence of multiple replication origins (Ohbayashi et al., 2016). Moore et al. (2019) presented similar findings on the significance of *dnaA* in chromosome replication in PCC 7002 through gene knockdown and knockout studies, finding that DnaA was not required for chromosome replication in conditions studied (Moore et al., 2019). The appearance of DnaA-independence is evolutionarily independent in each cell lineage (Ohbayashi et al., 2020). However, the prevalence of DnaA-independent chromosome replication draws into question what other mechanisms of chromosome replication initiation exist in these prokaryotes, and more specifically how they affect polyploidy and chromosome number.

To address alternative mechanisms (dnaA-independent) chromosome replication initiation, two alternative genes from the PCC 7002 genome suspected to play a role in chromosome replication were selected based on inferred function through homologs and sequence conservation between PCC 7002 and *E. coli*. The genes selected were: *recG* and *priA*. *recG* encodes for ATP-dependent DNA helicase. This gene is highly conserved across cyanobacteria and other genuses of bacteria, suggesting importance to the DNA regulation pathway (Cassier-Chauvat et al., 2016). Its known function includes a catalyst for chromosome unwinding at very specific dsDNA repair sites. Specifically, as a helicase and translocase in remodeling and directing DNA synthesis at Holliday junctions and replication forks during homologous recombination (Azeroglu et al., 2016). Its primary function remains unclear, but may act as a similar catalyst in primary chromosome replication.

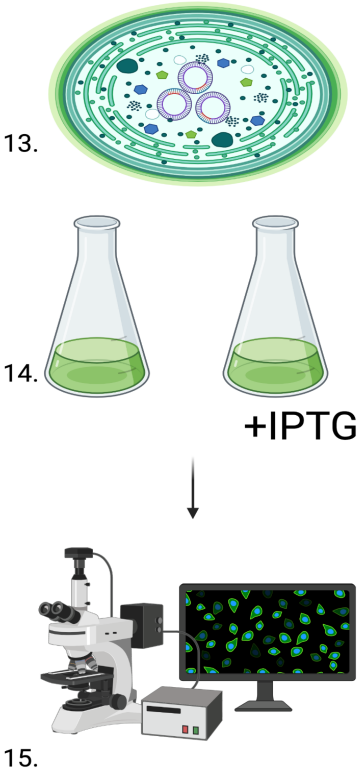
priA also encodes a helicase, but not a primary helicase. The protein complex responsible for replicating chromosomes is called the replisome. During replication the replisome will inevitably encounter and collide with different transcription complexes, biomolecules, DNA damage or aberrant secondary DNA structures. All of which can either stall or completely eject the replisome from the DNA and effectively arrest chromosome replication entirely. To counter the lethal effects of DNA synthesis arrest, cells have developed replication restart mechanisms with complexes known as primosomes. *priA* acts as a critical component of the primosome and serves to

recognize and remodel abandoned replication forks for replisome reloading (Duckworth et al., 2021). It initiates replisome restart by recognizing specific structures of stalled/abandoned replication forks, allowing it to function independently of sequence, thus allowing it to recognize replisome arrest anywhere along the chromosome. In studies performed on *E. coli*, the deletion of *priA* has resulted in cells extremely sensitive to DNA damage (Duckworth et al., 2021).

Both *priA* and *recG* have proven functionally significant in the genomic stability of *E. coli* (Duckworth et al., 2021) (Azeroglu et al., 2016). Based on homologous sequence conservation between both *E. coli* and PCC 7002, it is inferred that these genes may function similarly in PCC 7002. This makes *priA* and *recG* excellent gene candidates for targeting dnaA-independent chromosome replication initiation in PCC 7002. To determine whether either recG or priA are primary drivers of chromosome replication in a polyploid bacterium, we used pre-constructed PCC 7002 cells, scJC0164 (courtesy of the Cameron Lab) that allowed for chromosome visualization using a 256X TetO array with TetR-GFP fusion protein and were genetically modified with the insertion of CRISPR interference (CRISPRi) machinery inducible by isopropyl β-d-1-thiogalactopyranoside (IPTG). Methods for scJC0164 construction are described by Moore et al., (2019). To elucidate the possible mechanisms of polyploidy single-guide RNAs (sgRNA) targeted to *priA* and *recG* were inserted into scJC0164. *priA* and *recG* sgRNAs were successfully inserted into template plasmid first via Gibson Assembly (Gibson et al., 2009), amplified via chemically competent *E. coli*, sequence-verified via Quintara Sanger Sequencing and then transformed into scJC0164. CRISPRi was induced to knockdown *priA* and *recG* through the addition of IPTG. These strains were visualized via fluorescent microscopy and compared to cells lacking sgRNAs. A full experimental schematic can be found in **Figure 1**.



scJCC0164:sgRNA



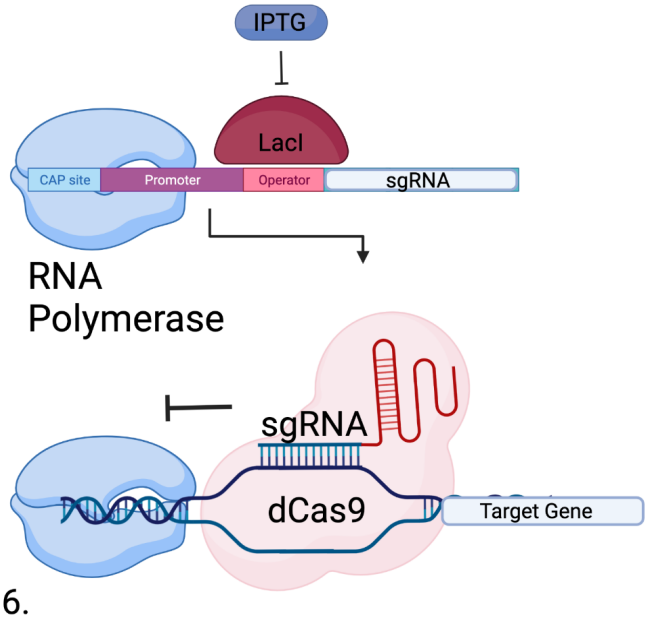


Figure 1. Full experimental schematic of sgRNA transformation and CRISPRi induction via IPTG. In chronological order: (**1.1**) sgRNAs were designed via CRISPR sgRNA Design by Synthego that were targeted to two genes suspected to partake in chromosome replication initiation in PCC 7002: *recG* and *priA*. **1.2.** Insertion of sgRNAs into template plasmid KAMc0205 via 2-piece overhang Phusion PCR. **1.3**. 2-piece plasmid Phusion products were fused together via Gibson reaction and template plasmid was eliminated via DpnI digest. **1.4.** sgRNA plasmids were transformed and amplified in chemically competent *E. coli* and then extracted (**1.5**) and Sanger sequenced through Quintara (**1.6**) Plasmids were then transformed into scJC0164 in **1.7** and **1.8**. Transformed cells (**1.9**) were used as template DNA in colony PCR (**1.10** and **1.11**). Colony PCR products were visualized via gel electrophoresis and UV imaging (**1.12**). Validated *ΔscJC0164:sgRNA* cells (**1.13**) were grown in 2 separate 25mL flasks of A+ media. 1mM of IPTG was added to one flask to induce CRISPRi (**1.14**). The cells were visualized via fluorescent microscopy as to visualize knockdown of *priA* and *recG* on ploidy phenotype (**1.15**). **1.16.** IPTG acts as a inhibitor to *lacI*, a *lacO* repressor. In the absence of a repressor, RNA polymerase binds and transcribes the lac operon of which the CRISPRi cassette is attached. dCas9 is expressed and is guided to target sequences via sgRNA where it binds the target sequence, blocking RNA polymerase and inhibiting transcription of the target gene.

Physiological Consequences of Prokaryotic Polyploidy

Hypothetically, because of the risk of accumulation of deleterious mutations, polyploidy is an unfavorable genotypic quality that should have been evolutionarily selected against; however, advantages have been proposed as well. The possible advantages of polyploidy in prokaryotes include double-strand DNA damage repair through homologous recombination, higher rate of protein synthesis in unfavorable conditions, phenotypic repression of deleterious mutations, and the use of redundant genomic DNA as a phosphate storage polymer. Possible harmful effects of polyploidy include the accumulation of deleterious mutations (Markov & Kaznacheev, 2016). However, since DNA is naturally subjected to DNA damage via ultraviolet or other radiation, exogenous chemicals and/or free radicals or other endogenous metabolic processes that release waste products, increasing ploidy level also increases mutation level. The accumulation of deleterious mutations as a result of polyploidy can then result in compromised offspring viability. This is because chromosome segregation is an apparently stochastic process that leaves polyploid prokaryotes susceptible to a phenomenon known as Muller's Ratchet. Segregation load is the resulting chromosome content daughter cells end up with after cell division. Since segregation is stochastic, there is no mechanism that ensures daughter cells receive functional copies of wild type

chromosomes, instead receiving a segregation load of the accumulation of mutated chromosomes that result in non-viable offspring (Markov & Kaznacheev, 2016). Non-viable offspring will eventually lead to species extinction.

A possible advantage of polyploidy that remains unexplored in cyanobacteria is heterogeneity, where the cell possesses multiple copies of homologous chromosomes with variation in alleles at the same locus. Previous work in the polyploid haloarchaeon *Haloferax volcanii* indicated that gene conversion may occur in this organism resulting in a loss of heterozygosity in nonselective environments (Lange et al, 2011). The studies in *Haloferax volcanii* utilized a heterozygous strain with modification at a single amino acid locus that could possess either wild-type *leuB* or *ΔleuB:trpA*. The cells were then grown under varying selective pressures: in the presence of tryptophan to select for *leuB*, in leucine to select for *ΔleuB:trpA* and in the absence of selection and then allelic ratios were determined using quantitative Polymerase Chain Reaction (qPCR). Lange et al. found that in the presence of tryptophan there was complete loss of *ΔleuB:trpA* genomes, in the presence of leucine, all *leuB* genomes were lost, and in the absence of selection, the genes converted back to their *leuB*-containing genomes (Lange et al., 2011) suggesting that heterozygotic polyploidy may not naturally occur unless under constant selective pressure for both alleles.

To elucidate the possible consequences of heterozygotic polyploidy in PCC 7002, we conferred competing antibiotic resistant genes, kanamycin-resistance (KmR) and gentamicin-resistance (GmR) at a single neutral allelic locus (*glpk*) in wild type PCC 7002 and studied the evolution of allele retention over multiple generations. The cells were grown in a variety of selectable conditions and DNA was extracted every 2-4 weeks. Allelic ratios were then determined and compared via qPCR, similarly to the Lange et al. experimental design. The selective pressures of this design included both gentamicin and kanamycin to select for both alleles, only gentamicin to select for GmR, only kanamycin to select for KmR, and in the absence of selection. A full experimental schematic can be found in **Figure 2.**

Summary

Ultimately, our study contains two objectives: 1. To determine if the two genes, *priA* and *recG*, suspected to partake in chromosome replication initiation are also part of the mechanism responsible for maintaining polyploidy in the PCC 7002 genome and 2. To determine if heterogeneity is manipulatable, and therefore a possible beneficial consequence of polyploidy in PCC 7002. We hypothesize that if *priA* and *recG* play major roles in chromosome replication and the stability of ploidy level in PCC 7002, then in repressing the expression of these genes, a phenotypic decrease in ploidy level will be visible. We also hypothesize that if heterogeneity is a possible benefit of polyploidy, in conferring a heterozygotic cell and placing it in an environment void of selective pressure, the ratio of competing alleles will remain relatively equal and constant across generations.

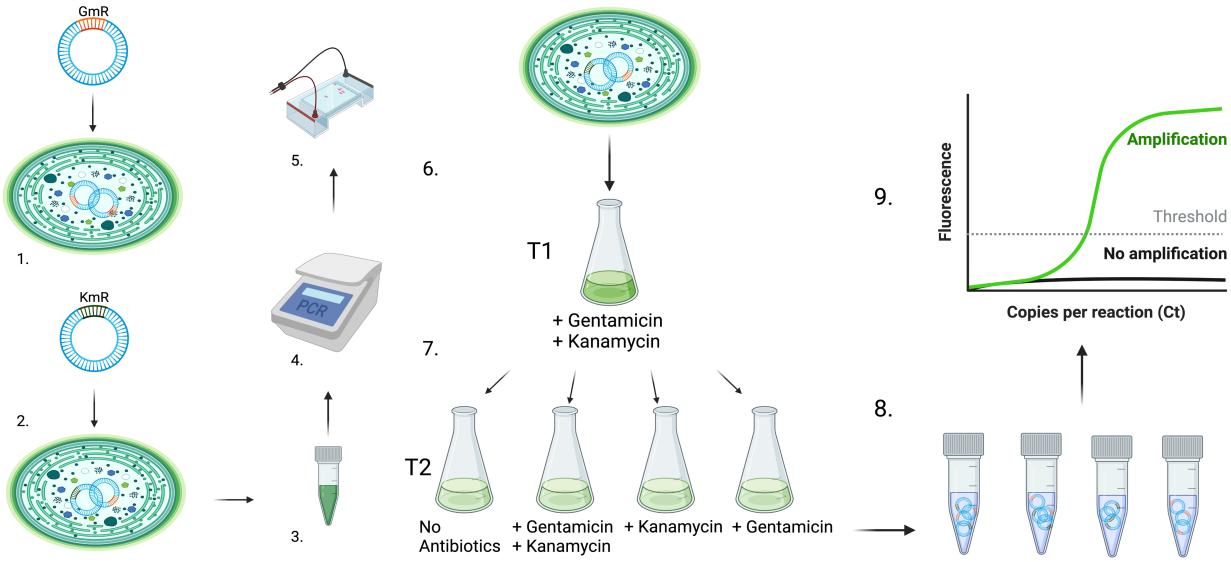


Figure 2. Full experimental schematic of polyploid heterogeneity and competing alleles under different selective pressures in PCC 7002. In chronological order: **2.1.** Pre-designed GmR plasmids, JCC_ES_014 were transformed into WT PCC 7002 *glpK* neutral site. **2.2**. Pre-designed KmR plasmids, JCC_ES_015 were then transformed into *ΔglpK:GmR*. Transformation was validated via PCR and gel electrophoresis using whole colony swabs as template (**2.3-2.5**). A validated colony containing *ΔglpK:GmR+ΔglpK:KmR* (**2.6**) was selected to seed an initial culture, denoted T1, in normalized conditions of a gentamicin concentration of 15µg/mL and 30µg/mL of kanamycin to select for cells containing both alleles. **2.7.** T1 culture was then used to seed four different conditions: no antibiotics, gentamicin at 5µg/mL with kanamycin at 30µg/mL, gentamicin at 15µg/mL, and kanamycin at 90µg/mL at an initial $OD_{750\ nm}$ of between 0.07-1.00. Cells were split every 2-5 days for 25 weeks. **2.8.** DNA was extracted

from each condition every 2-4 weeks starting at culture T5. **2.9.** Allelic ratios were determined from DNA passages using qPCR set against a standard curve.

Materials and Methods

sgRNA and Primer Design

CRISPR sgRNA Design by Synthego was used to design primers containing sgRNAs for two target genes in PCC 7002: *priA*, and *recG*. sgRNA sequences (**Table 1**) were selected under the following parameters: ~20 base pairs in length near the end of the gene sequence, 50% G/C content, no nucleotide repeats greater than 3, higher on-target score and fewer off-targets. On-target score predicts the ability of the sgRNA to knockdown the gene based on nucleotide composition, and the off-target score predicts the likelihood of the sgRNA binding an alternative sequence. sgRNA scoring parameters were determined by Doench et al., 2016. sgRNA sequences in **Table 1** were appended to primer sequences noted in **Table S1**. All primers and templates designed and used for studying polyploidy mechanisms are described in supplemental figure **Table S1**.

sgRNA/CRISPR-interference Plasmid Construction

In order to amplify the sgRNA sequences to transform into PCC 7002, pJSF003 and pJSF005 were inserted into plasmid template KAMc0205 (courtesy of the Jeff Cameron Lab) via two-piece fragment overlap Gibson Assembly (Gibson et al., 2009) using Phusion PCR products. KAMc0205 construction methods can be found in Moore et al, (2019) who used CRISPRi methodology from Markley et al., (2015). JSFO_005 and JSFO_009 were paired with KAM0167 reverse primers. JSFO_006 and JSFO_010 were paired with KAM0166 forward primers in a Phusion PCR master mix. KAM0166/KAM0167 primers were courtesy of Cameron Lab. PCR thermocycler settings were: 98°C for 2 min (initial denaturation), 98°C for 10 sec, 60°C for 15 sec, 72°C for 45 sec repeated for 35 cycles. Final extension time: 72°C for 5 min, and 12°C until removed. All PCR products in this work were validated via gel electrophoresis and UV imaging. The two-piece Phusion fragments were fused into circular plasmids via Gibson cloning. A 1:1 ratio of each fragment was used in a 1.3X Gibson Mix at a

thermocycler setting of 50°C for 20 min. DpnI digest reaction was used to eliminate template plasmid. The digest reactions incubated at 37°C for 15 min. DpnI was deactivated at 80°C for 5 min. Samples were stored at -20°C.

Plasmid Amplification

sgRNA plasmids, pJSF003 and pJSF005 were transformed and amplified in chemically competent *E. coli* prior to transformation into PCC 7002 cells pre-modified with CRISPRi machinery, scJC0164 (courtesy of Cameron Lab). 0.5ng of each plasmid was transformed into 50µL of *E. coli*. Tranformants were selected for on TSA plates with 50µg/mL ampicillin. Colonies incubated at 37°C for 18 hours. Correct plasmid transformation was validated via colony PCR and gel electrophoresis. Individual colonies were suspended in 10µL of Nuclease-free water, 1µL of which was used as templates in 1X GoTaq master mix PCRs with Pet_81F (forward) and KAMo0100 (reverse) primers. Thermocycler settings were 95°C for 5 min (initial denaturation), 95°C for 30 sec, 55°C for 30 sec, 72°C for 4 min repeated for 35 cycles. Final extension was 72°C for 5 min and 12°C until removed. Individual validated transformed *E. coli* colonies were cultured in 3mL liquid LB media with 50µg/mL ampicillin for 18 hours at 37°C and 220 rpm for plasmid extraction.

Gel Electrophoresis

Gels were made with between 0.7-1.5% gel agarose contingent on expected PCR product size in 1X TAE buffer and 1X Sybr Safe. GeneRuler 1 kilobase (kb) and 100bp ladders were used contingent on the longest expected PCR product length. All gels were run at 120V for 20-35 min. See supplemental figure **Table S3** for agarose percent used for noted PCR product sizes

Plasmid Extraction and Sequencing

pJSF003 and pJSF005 were extracted from *E. coli* grown in liquid LB culture using Zymo Research Zyppy™ Miniprep kits. Plasmids were extracted from 3mL of transformed *E. coli* cultures. Zymo manufacturer instructions were followed. Plasmid

quality and concentration were determined via NanoDrop ND-1000 Spectrophotometer. pJSF003 and pJSF005 sequences were validated via Quintara Sanger Sequencing.

ΔscJC0164:sgRNA Transformation

Validated pJSF003 and pJSF005 sequences were transformed into scJC0164. scJC0164 cells were grown from a plate streak in 25mL A+ media ($Na_2EDTA*2H2O$, KCl, $CaCl_2$, $NaNO_3$, KH_2PO_4, Tris HCl, A+ Trace Minerals, B12, NaCl, $MgSO_4$, milliQ H_2O) with 25µg/mL spectinomycin, 30µg/mL kanamycin, 15µg/mL gentamicin, and 0.5µg/mL anhydrotetracycline (aTC) for 25 hours at 100% red light -2 brightness intensities, 25% white light -1 brightness intensity. All lights used for these experiments were the AIBOO RGBW RGB + White Color Changing Christmas Xmas Under Cabinet LED Lights Kit RF Remote Puck Lights for Kitchen Counter Shelf Furniture Ambiance Lighting (RGBW, 4 Lights, 12W) (purchased from amazon.com). Lights were placed under flasks in a shaking incubator at 230 rpm at 37°C. 300µL of this culture was incubated separately with 1000ng of each pJSF003 and pJSF005 for 42 hours at the same light and temperature conditions listed above. The solutions were plated on A+ plates with the same antibiotic conditions as described above.

scJC0164:sgRNA CRISPRi Induction

25mL liquid cultures each of *ΔscJC0164:pJSF003* and *ΔscJC0164:pJSF005* were grown for 3 days in the conditions and antibiotic concentrations described above. These cultures were used to seed 25mL cultures with the same initial $OD_{750\ nm}$ of ~0.09. CRISPRi was induced in one of each liquid culture via the addition of 1mM of IPTG. Cultures incubated in the settings described above for 72 hours.

Microscopy

Cells were visualized and photographed via widefield Olympus fluorescent microscopy using cellSens Dimensions imaging software on the 100X oil objective lens. Excitation/emission settings were TRITC for ~20 millisec and FITC for ~450 millisec.

Kanamycin-Resistant Plasmid Amplification

Pre-designed plasmids with KmR insertion, JCC_ES_015 courtesy of Jeff Cameron Lab were amplified and extracted in chemically competent *E. coli* in the methods described above for plasmid amplification and extraction.

Δglpk:GmR+ΔglpK:KmR Transformation

WT PCC 7002 (courtesy of Cameron Lab) were grown in 25mL of A+ media. Cultures incubated for 48 hours at 37°C, 220 rpm and 50% white light. WT PCC 7002 cells were diluted in a 1:1 ratio of A+ media and mixed with 1000ng of pre-designed plasmid with GmR insertion, JCC_ES_014 (plasmid courtesy of Cameron Lab). The mixture incubated in 25% white light at 37°C for 4 hrs. To select for transformants with the *GmR* gene inserted into the *glpK* neutral site, cells were plated on A+ plates with 30µg/mL gentamicin and incubated at 37°C in 25% white light for 5 days. 12 individual colonies were selected and patch plated on A+ plates. A colony was selected and grown on an A+ plate with 15µg/mL gentamicin at 37°C, 50% white light for 48 hours. A streak from this plate seeded a 25mL liquid *Δglpk:GmR* culture in A+ media of which was transformed with 1000ng of JCC_ES_015 using the same methods described for *Δglpk:GmR* transformation. Transformation of both antibiotic resistant plasmids into the *glpk* neutral site of PCC 7002 genome was validated via PCR and gel electrophoresis. JCCo0388 and JCCo0389 primers targeted the parameter of *glpK*, EBJo0178 and EBJo0179 targeted *glpK* internal as a negative control, NCH_9F and JCCo0389 targeted the GmR cassette, and EBJo0182 and JCCo0389 targeted the KmR cassette.

Δglpk:GmR+ΔglpK:KmR Cell Passaging

Initial *Δglpk:GmR+Δglpk:KmR* liquid cell culture - termed T1 were maintained in a gentamicin concentration of 15µg/mL and 30µg/mL of kanamycin. Cells incubated for 6 days at 37°C, 25% white light, 100% red light and 230 rpm. T1 cells seeded 4 separate culture conditions: no antibiotics, gentamicin at 5µg/mL with kanamycin at 30µg/mL, gentamicin at 15µg/mL, and kanamycin at 90µg/mL at an initial $OD_{750\ nm}$ of between 0.07-1.00. The four conditions were continuously passaged in liquid A+ media every 2-5 days for 25 weeks at 37°C, 25% white light, 100% red light and 230 rpm.

DNA Extraction

Genomic DNA was extracted every 2-4 weeks from each culture condition starting at T5 (5 weeks after initial culturing). ~15mL of each culture was centrifuged for 15 min at 4000 rpm. Pellets were resuspended in 200µL of sterile water and DNA extraction was performed using Zymo Research *Quick*-DNA™ Fungal/Bacterial Miniprep Kits. Extraction followed manufacturer instructions. Cells were vortexed in bead beaters for 30 min and eluted with 40µL of DNA Elution Buffer. Sample quality was validated via NanoDrop ND-1000 Spectrophotometer and frozen for storage.

Primer Design for qPCR

Primers were designed to anneal to the KmR and GmR cassettes using Primer3Plus. Primer parameters included the following: length of 20 nucleotides, Tm of 55°C, GC content of 50%, likelihood of homodimer formation anywhere in the primer: < 6, at the 3' end of the primer: < 5, chance of primer binding somewhere other than expected: < 50, Tm of hairpin structure in a single primer: < 40, chance of mis-priming at 3' end: low, and an ideal PCR product length of 200-1000 base pairs (bp). Homology of the designed primer sets were compared to the rest of the PCC 7002 genome using BLAST. All primers and templates designed and used for studying polyploidy consequences are described in supplemental figure **Table S2**. All primer working stocks were 10µM.

Quantitative Polymerase Chain Reaction (qPCR)

Quantitative analysis of *GmR:KmR* allelic ratios was determined using qPCR with an absolute value standard curve. Standards were created for KmR, GmR and WT qPCR products using 50µL PCR gel extracts of KmR primers, GmR primers and WT primers with T1 template DNA. DNA was extracted from gel electrophoresis using Zymo Research Zymoclean™ Gel DNA Recovery Kit and then cleaned and purified with Zymo Research DNA Clean & Concentrator-5™. Manufacturer instructions were followed. qPCR was run with ProMega GoTaq® Master Mix at 95°C for 30 sec, 55°C for 30 sec

and 72°C for 30sec for 40 cycles using an Applied Biosystems 7500 Fast Real-Time PCR System.

Results

Determining the impact of PriA and RecG on Ploidy Levels of PCC 7002

Construction of sgRNA plasmid vectors, pJSF003 and pJSF005

In order to utilize the pre-designed scJC0164 PCC 7002 cells from the Cameron Lab that already contain the inducible CRISPR-interference machinery and a green-fluorescent protein (GFP) targeted to chromosome sites enabling visualization of individual chromosomes (Moore et. al. 2019), single-guide RNAs needed to be constructed specific to *priA* and *recG* and successfully integrated into a plasmid vector prior to transformation into scJC0164. sgRNA sequences for *priA* and *recG* were designed via Synthego CRISPR design and their 5'-3' sequences can be found in **Table 1.** sgRNAs targeting *priA* were designed with on-target scores of 0.630 and 0.616, respectively, with no off-target scores and cloned into pJSF003 and pJSF004. sgRNAs targeting *recG* were designed with on-target scores of 0.690 and 0.686, respectively, and no off-targets and cloned into pJSF005 and pJSF006. All sgRNAs were within the first few hundred nucleotides of the 5' end of their target genes and all sgRNAs were 20bp in length. All four sgRNAs were appended to primers complementary to the first 20 nucleotides of the edges of the insertion site in the template plasmid, KAMc0205. Primer sequences with sgRNA attachment can be found in **Supplemental Figure Table S1**.

Plasmid Name	5'-3' Sequence	Target Gene
pJSF003	AGTGCCGCTAGAATTGGCGG	*priA*
pJSF004	AGAATTGGCGGTGGTGCCAG	*priA*
pJSF005	TTATCCATCGAAGCAGAACG	*recG*
pJSF006	CGAGGCTTTCAGGATCTCCA	*recG*

Table 1. 5'-3' single-guide RNA sequences complementary to genes *priA* and *recG* designed through Synthego. Two sgRNAs were designed per target gene suspected to

function during chromosome replication in PCC 7002. sgRNA sequences were 20 nucleotides in length with ~50% G/C content and minimal repeats and were selected based on on and off target scores.

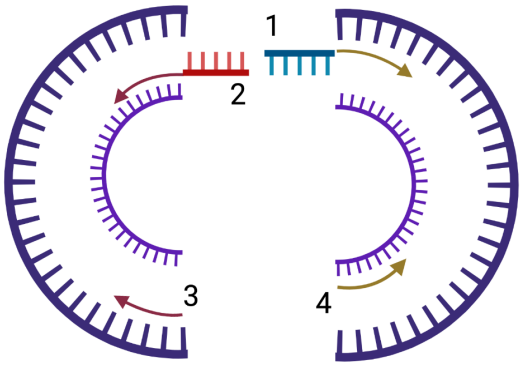
All sgRNAs were inserted into vector plasmid, KAMc0205 via two-piece Phusion PCR. **Figure 3A** depicts a schematic of the two-piece Phusion PCR primer annealing. Primers 1 and 2 represent primers JSFO_005 (F) and JSFO_006 (R) for amplification of KAMc0205 with pJSF003 insertion. Primers 1 and 2 also represent primers JSFO_009 (F) and JSFO_010 (R) for the amplification of KAMc0205 with pJSF005 insertion. Primers 3 and 4 represent primers KAMo0167 (F) and KAMo0166 (R) that were paired with corresponding JSFO forward and reverse primers in order to form the vector plasmid in two pieces. Expected Phusion PCR product lengths were 3805bp and 3019bp for a complete plasmid length of 6824. The Phusion products were visualized on a 1.0% agarose gel to confirm PCR lengths prior to Gibson fusion depicted in **Figure 3B.** All products were the expected corresponding band lengths of ~3805bp for one fragment and ~3019bp for the second fragment, indicating that PCR fragments could be used with Gibson Assembly to fuse the fragments into one complete plasmid, and DpnI digest to remove any template plasmids (see materials and methods).

The Gibson products were then inserted into chemically competent *E. coli* for amplification. Prior to sequencing, the transformed *E. coli* were run through colony PCR and visualized via gel electrophoresis to verify transformation of the vector plasmid. **Figure 3C** depicts a schematic for primer annealing to *E. coli* transformed with either pJSF003 or pJSF005 for colony PCR. Primers annealed to the perimeter of the sgRNA cassette, resulting in an expected product of ~2000bp. The PCR products were run on a 1% agarose gel and visualized via UV imaging seen in **Figure 3D.** The top gel presented bands in lanes 3 and 7 consistent with the expected band length of ~2000bp. These lanes contained template colonies of pJSF003, validating their transformation into the *E. coli*. The bottom gel contained an expected band length of ~2000bp in lane 3 with template colony pJSF005, validating transformation of pJSF005 into that colony. The aberrant band appearance was secondary to gel error. All other lanes containing bands at ~1200bp were *E.coli* lacking the correct pJSF003/pJSF005 cloning products.

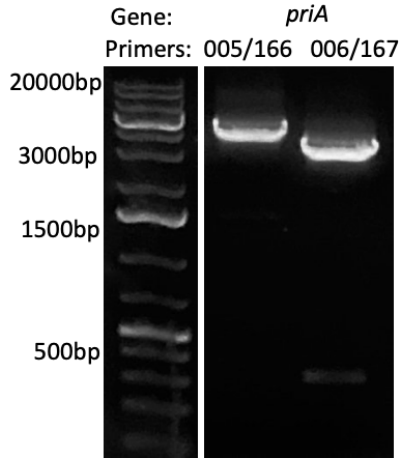
pJSF004 and pJSF006 did not yield validated transformation into *E. coli*, therefore only pJSF003 and pJSF005 were utilized going forward.

All validated *E.coli* colonies from **Figure 3D** were grown up for plasmid amplification and extraction. Extracted plasmids were analyzed for DNA quality via spectrometer. Plasmids with acceptable quality were Sanger sequenced to confirm correct sgRNA sequence insertion . Resulting chromatograms (**Figure 3E**) for pJSF003 and pJSF005 revealed validated sgRNA sequences with no mutations. Not depicted is a single point mutation in the pJSF005 *lacI* sequence changing a single asparagine to an isoleucine.

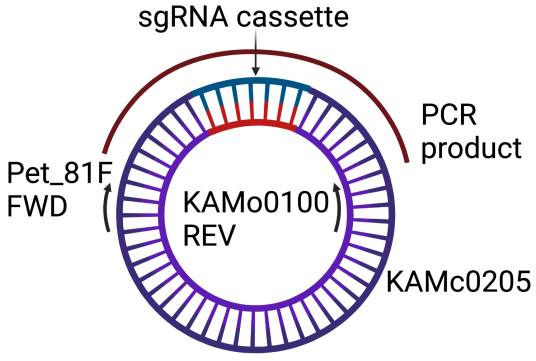
A.

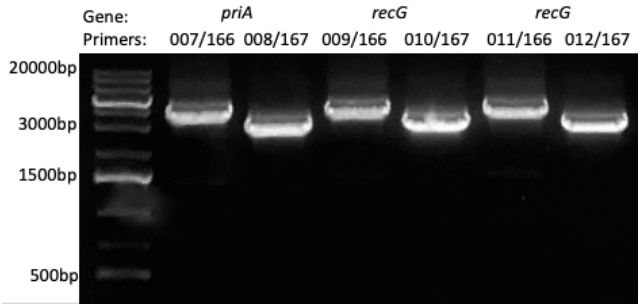


B.



C.





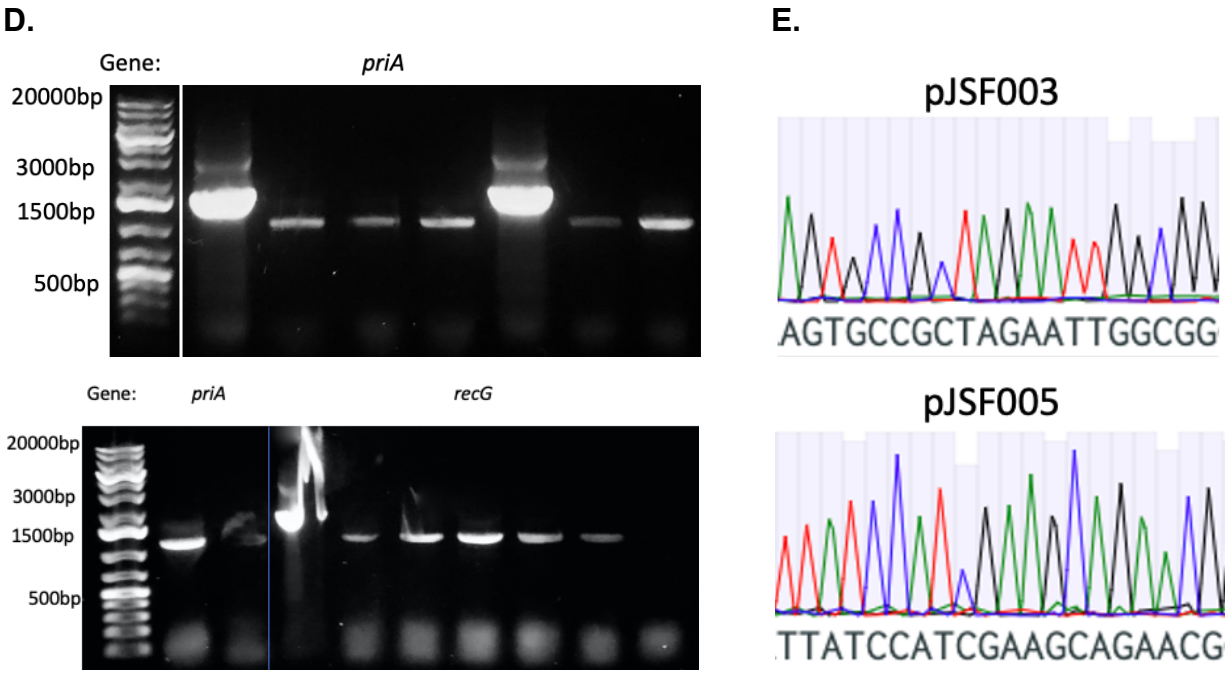


Figure 3. sgRNA plasmid design and construction. A. A schematic of two-piece Phusion PCR. Primer 1 represents JSFO reverse primers with sgRNAs appended (red fragment). Primer 2 represents JSFO forward primers with sgRNAs appended (blue fragment). JSFO forward primers were paired with KAMo0166 reverse primers (4) and JSFO reverse primers were paired with KAMo0167 (3) to form a two-piece vector plasmid with sgRNA sequences inserted. **B.** Depicts a 1% agarose gel run of the two-piece Phusion products. All products were at the expected length of ~3800 for JSFO forward and KAMo reverse primers, and ~3000bp for JSFO reverse and KAMo forward primers. **C.** A schematic of the primer annealing for colony PCR for pJSF003 (labeled *priA*) and pJSF005 (labeled *recG*). Pet_81F forward primer and KAMo0100 reverse primer annealed along the perimeter of the sgRNA cassette insertion site with an expected PCR product length of ~2000bp. **D.** A 1% agarose gel to visualize the *E.coli* colony PCR products. Lanes 3 and 7 in the top gel contained a bright band at the expected band length of ~2000bp, validating the cloning of pJSF003. Lane 3 in the bottom gel contained a band at ~2000bp, validating the cloning of pJSF005. **E.** Chromatograms of final pJSF003 and pJSF005 reveal validated sequences of the sgRNAs plasmids.

Insertion of pJSF003 and pJSF005 into scJC0164

To insert sgRNAs into the scJC0164 genome, pJSF003 and pJSF005 extracted from sequenced-validated *ΔE.coli* were transformed into scJC0164 cells as described in materials and methods. Transformation was validated using PCR and gel electrophoresis. **Figure 4A** depicts the primer annealing schematic for validation of *ΔscJC0164:pJSF003/pJSF005*. KAMo0252 (F) and KAMo061 (R) targeted the

scJC0164 genome on the perimeter of the sgRNA cassette. Expected PCR product length was 3686bp with sgRNA cassette integration. **Figure 4B** reveals the final gel result of the PCR described in Figure 4A. Template was *ΔscJC0164:pJSF005* colonies for lanes 1-3 and *ΔscJC0164:pJSF003* colonies for lanes 4-6. Lane 7 used a scJC0164 colony void of sgRNA insertion for template as a control. The presence of consistent bright bands at ~3600bp in lanes 1-6 and the presence of a shorter band at ~800bp in lane 7 validated the segregation of pJSF003 and pJSF005 into the scJC0164 genome in the transformed colonies and the absence of the cassette in the non-transformed scJC0164 colony.

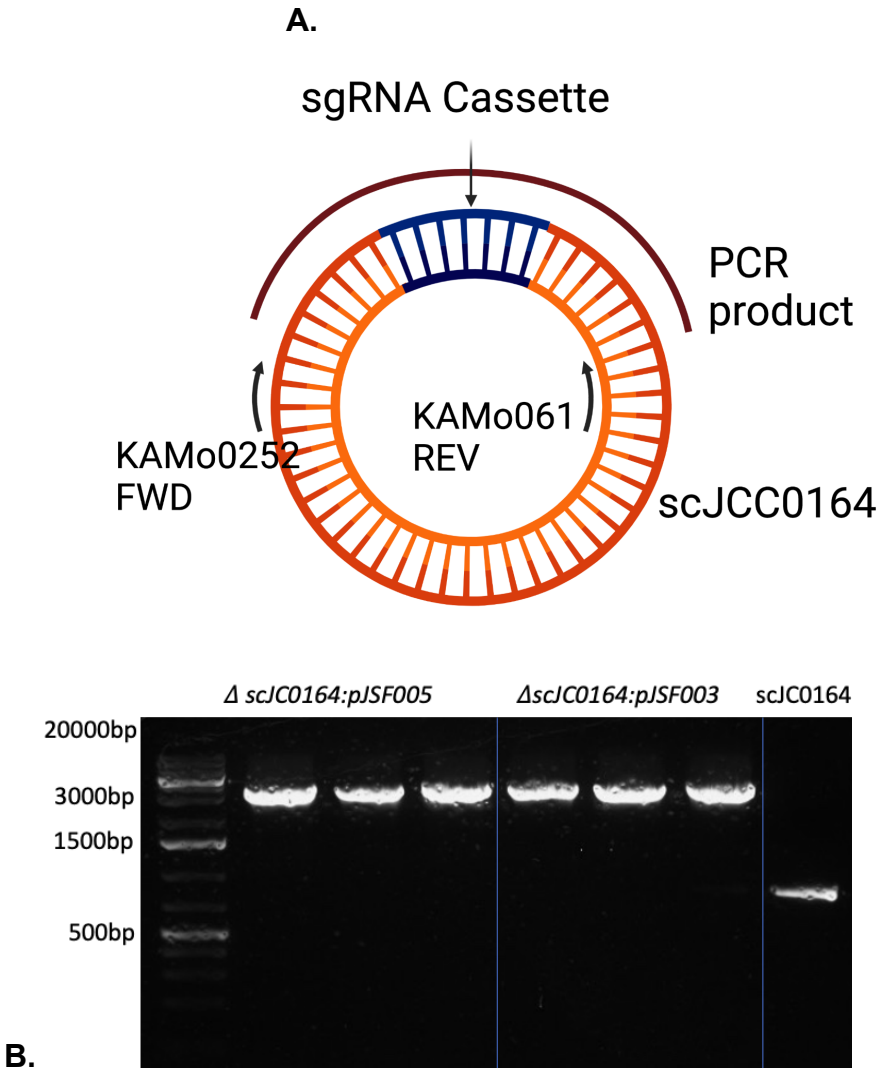


Figure 4. *ΔscJC0164:pJSF003/pJSF005* Transformation Analysis A. Schematic of primers used in PCR for segregation analysis of pJSF003 and pJSF005 into the scJC0164 genome. Primers KAMo0252 (F) and KAMo0061 (R) annealed to the external portion of the sgRNA

cassette producing a PCR product of 3686bp. **B.** A 1.2% agarose gel of colony PCR of *ΔscJC0164:pJSF003/pJSF005*. Lanes 1-6 present bands at ~3600bp. Lane 7 presents a band at ~800bp as a positive control, indicating the absence of the sgRNA cassette.

Knockdown of priA and recG reveal no phenotypic effect on ploidy

Following the validation of *ΔscJC0164:pJSF003/pJSF005*, the cells were induced to knockdown *priA* and *recG* expression by addition of IPTG. *ΔscJC0164:pJSF003/pJSF005* were grown in two liquid A+ medias in the presence of antibiotics to maintain the sgRNA/CRISPRi cassette (see materials and methods). The cells were passaged once to ensure a consistent initial $OD_{750\ nm}$ for all cultures to serve as a control to determine if knockdown of *priA* or *recG* compromised cell viability and density. Monolayers of cells were then visualized via fluorescent microscopy immediately after removal from incubation. **Figure 5** depicts the images obtained of *ΔscJC0164:pJSF003/pJSF005* in the presence and absence of IPTG. **Figure 5.A.i** depicts scJC0164 cells without pJSF003 or pJSF005 insertion. **Figure 5.A.ii** is the addition of IPTG to scJC0164 cells void of pJSF003/pJSF005 insertion. Overall cell quality as determined by level of chlorosis was normal as was ploidy level, with the average chromosome count subjectively measured to be between 2-5 chromosomes per cell.

Three replicates of different *ΔscJC0164:pJSF005* cultures in the absence of IPTG are shown in **Figure 5B**. Cell ploidy and quality appeared similar to scJC0164 cells. *ΔscJC0164:pJSF005* in the presence of IPTG (**Figure 5C**) and assumed knockdown of *recG* revealed a single generation (**Figure 5Ci**) with compromised ploidy level. However, replicates **(Figures 5Cii** and **5Ciii**) revealed decreased fluorescence and increased chlorosis, but failed to reveal consistent decreased ploidy level.

The process was repeated for *ΔscJC0164:pJSF003*, with three replicates in the absence of IPTG (**Figure 5Di-iii**) again revealing average cell ploidy level and viability that were similar to scJC0164. The second replicate (Figure 5Dii) had decreased fluorescence but normal ploidy level. In the presence of IPTG (**Figure 5E**) *ΔscJC0164:pJSF003* cell viability and fluorescence overall seemed mild to moderately compromised, but ploidy level remained similar to -IPTG levels. .

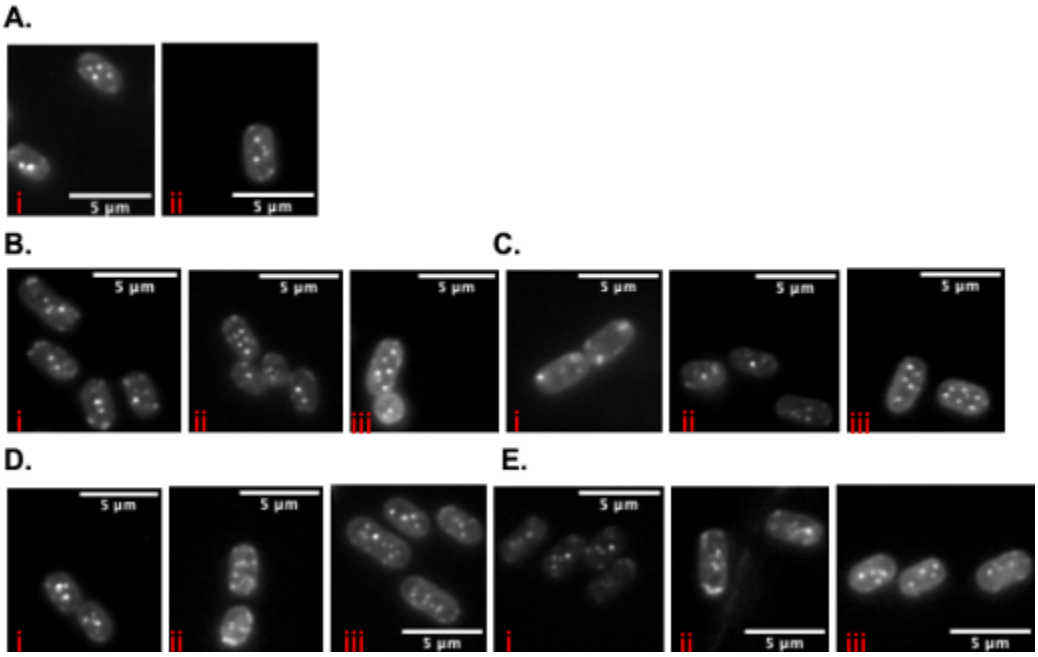


Figure 5. Fluorescent Microscopy Images of *ΔscJC0164:pJSF003/pJSF005* +/- IPTG. A.i. scJC0164 cells without sgRNA with normal cell viability and ploidy level. **A.ii.** scJC0164 cells void of sgRNA + IPTG with normal cell viability and ploidy level **B.** Replicates of *ΔscJC0164:pJSF005* cells in the absence of IPTG with average ploidy level **C.** Replicates of *ΔscJC0164:pJSF005* cells + IPTG with a single generation with compromised ploidy (**Ci**). **D.** Replicates of *ΔscJC0164:pJSF003* in the absence of IPTG with cells possessing average ploidy level. **E.** Replicates of *ΔscJC0164:pJSF003* cells + IPTG with normal ploidy level, but compromised fluorescence and quality.

Determining the Impact of Selective Environments on Allelic Ratios in PCC 7002

Insertion of GmR and KmR into PCC 7002

In order to generate PCC 7002 cells with competing alleles at a single chromosomal locus, WT PCC 7002 cells were first transformed with gentamicin-resistance utilizing the pre-constructed Cameron Lab plasmid, JCC_ES_014 (GmR). Validation of homologous recombination of the *GmR* gene into the *glpK* neutral site within the PCC 7002 genome was performed via PCR using two primer sets targeting the internal and immediate external sites of the *glpK* site. Primer set JCCo0388/JCCo0389 annealed to the perimeter of the *glpK* insertion site denoted *glpK* external (**Fig. 6Ai**). These primers produce a 1658bp PCR product from wild type cells

and a 1476bp product if homologous recombination occurred. EBJo0179/EBJo0179 targeted an internal *glpK* sequence (**Fig. 6Aii**). WT expected band length was 482bp. No bands were expected in *ΔglpK:GmR* colonies. **Figure 6A** depicts the gel result using colony swabs of *ΔglpK:GmR* as template for the aforementioned primer sets. All *glpK* external products presented bands at the expected length of ~1500bp. Colonies 8, 15, and 17 in the second gel image presented unexpected faint bands at ~500bp, indicating incomplete segregation in those colonies.

Colony 10 of *ΔglpK:GmR* presented no bands with *glpK* internal primers (**Fig. 6Aii lane 2**) and presented an expected band length of ~1500bp with *glpK* external primers (**Fig. 6Ai)** indicating complete segregation of *ΔglpK:GmR*, therefore it was selected for transformation with the kanamycin-resistant plasmid, JCC_ES_014 (KmR), which also integrates into the *glpK* neutral site. To ensure that both the *GmR* and *KmR* genes were both retained, transformed cells were selected on an A+ plate containing both gentamicin and kanamycin. To validate that transformed cells contained both the *GmR* and *KmR* genes into the *glpK* site, two additional primer sets specific to internal sequences of the KmR and GmR cassettes were also used to validate the segregation of *ΔglpK:GmR+ΔglpK:KmR* cells. Colony swabs of *ΔglpK:GmR+ΔglpK:KmR* were used as template in the PCR reaction. KmR internal primers were EBJp0082 (F) and JCCo0389 (R) with an expected PCR product length of 344bp. GmR internal primers were NCH_9F (F) and JCCo0389 (R) with an expected PCR product length of 97bp. **Figure 6B** reveals validated segregated *ΔglpK:GmR+ΔglpK:KmR* colonies with all bands at expected lengths except for lane 1 in **Figure 6Bi**, with an unexpected complete absence of a band. Colony 12 was selected for culturing based on strongest band luminosity for GmR internal primers (**Fig. 6Bii**) and validation of complete *ΔglpK:GmR+ΔglpK:KmR* segregation.

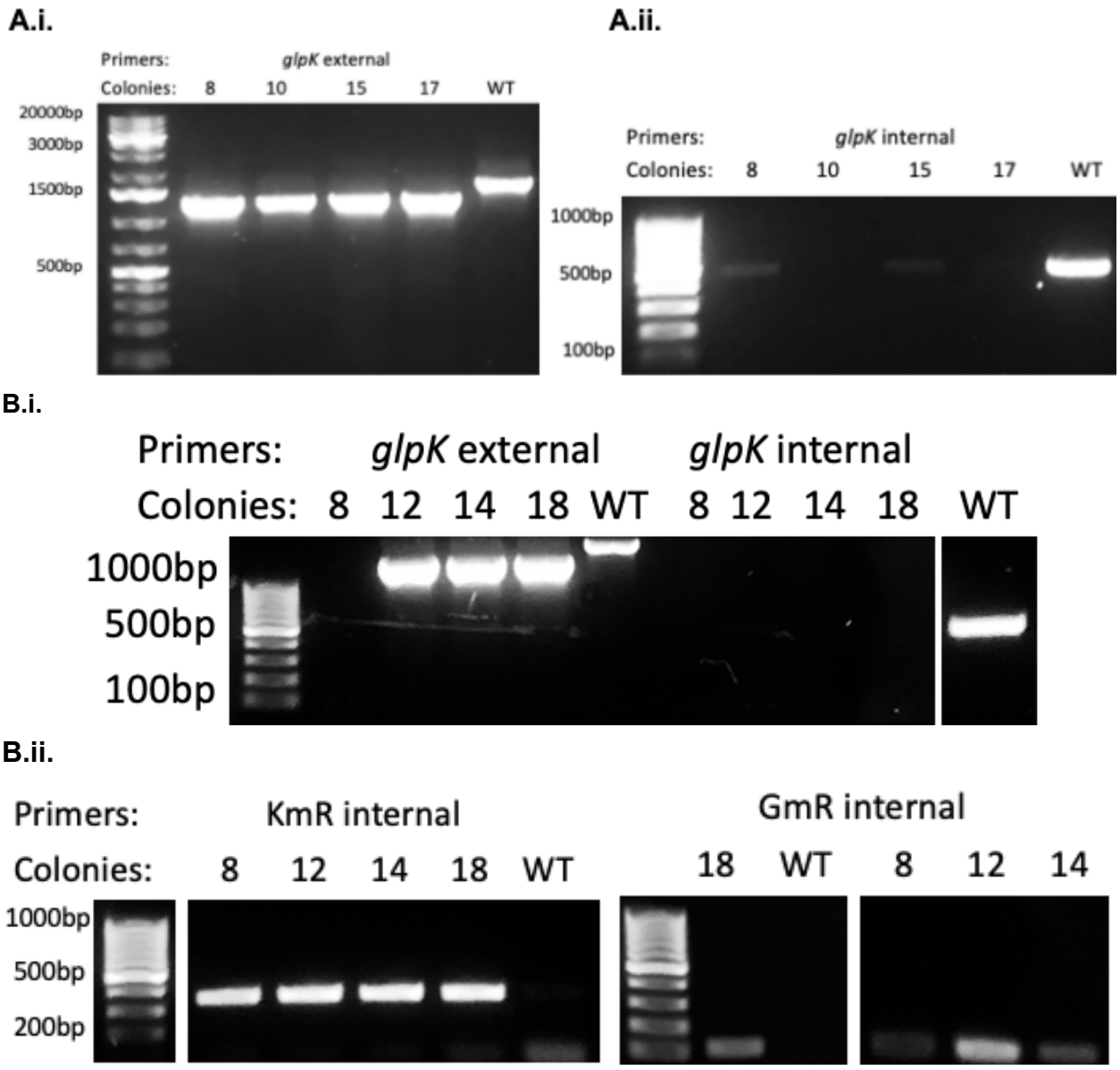


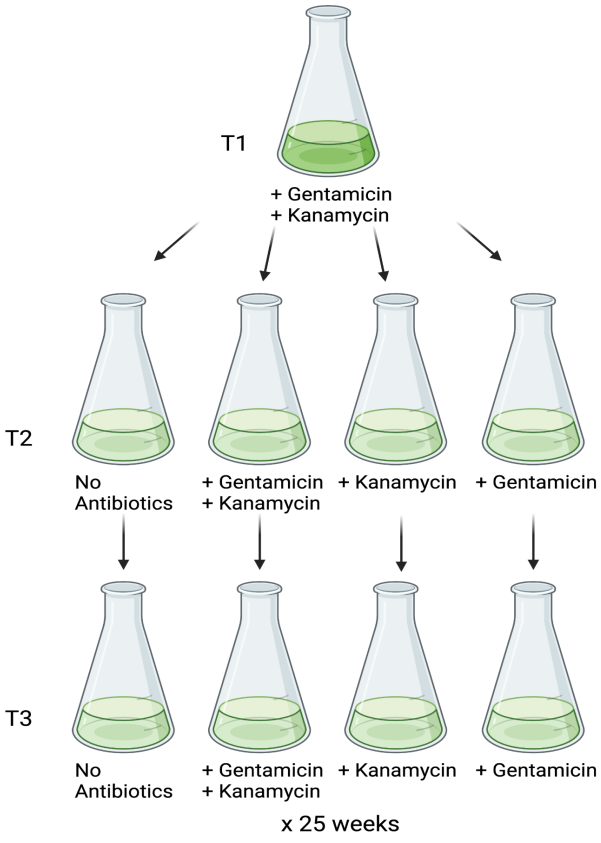
Figure 6. *ΔglpK:GmR* and *ΔglpK:GmR+ ΔglpK:KmR* Segregation Validation A. A 1.2% agarose gel of *ΔglpK:GmR* colony PCR products. All colonies with *glpK* external primers presented bands at ~1500bp, as expected. Colonies 8, 15, and 17 presented bands at ~500bp with *glpK* internal primers, suggesting persistent wild-type *glpK* sequences in those colonies and incomplete *ΔglpK:GmR* segregation. **B.** A 1.2% agarose gel of *ΔglpK:GmR+ΔglpK:KmR* colony PCR products. Expected bands for colonies 12, 14, and 18 were at ~1500bp for *glpK* external. Colony 8 presented no band. All colonies presented no bands for *glpK* internal. All colonies produced bands at expected lengths for KmR internal primers at ~344bp. A faint band at ~344bp in the WT lane indicates probable contamination. All colonies presented expected band lengths at ~100bp for GmR internal primers.

Passaging ΔglpK:GmR+ΔglpK:KmR

After validation of complete *ΔglpK:GmR+ΔglpK:KmR* segregation, a single culture swab was used to seed a culture grown in both kanamycin and gentamicin, denoted T1. T1 cultures were then used to seed four separate environmental conditions: both gentamicin and kanamycin, only gentamicin, only kanamycin, and the absence of antibiotics. Cells were passaged every 2-5 days for 25 weeks, starting at an initial $OD_{750\ nm}$ of between 0.7-1.00 each time, depicted in **Figure 7A.**

Figure 7B reveals the growth trends of the subcultures used in quantitative analysis in each selective condition over the course of 42 passages (~25 weeks). Conditions were kept constant with the same antibiotic concentrations, lighting and temperature maintained in each subculture. Cell density was measured for each subculture immediately after removal from incubation via a spectrophotometer set to $OD_{750\ nm}$. Fluctuations in growth density varied, but overall trend between each antibiotic condition remained relatively consistent. DNA was extracted for qPCR allelic ratio analysis every 2–4 weeks. T1 DNA was extracted as a control.

A.



B.

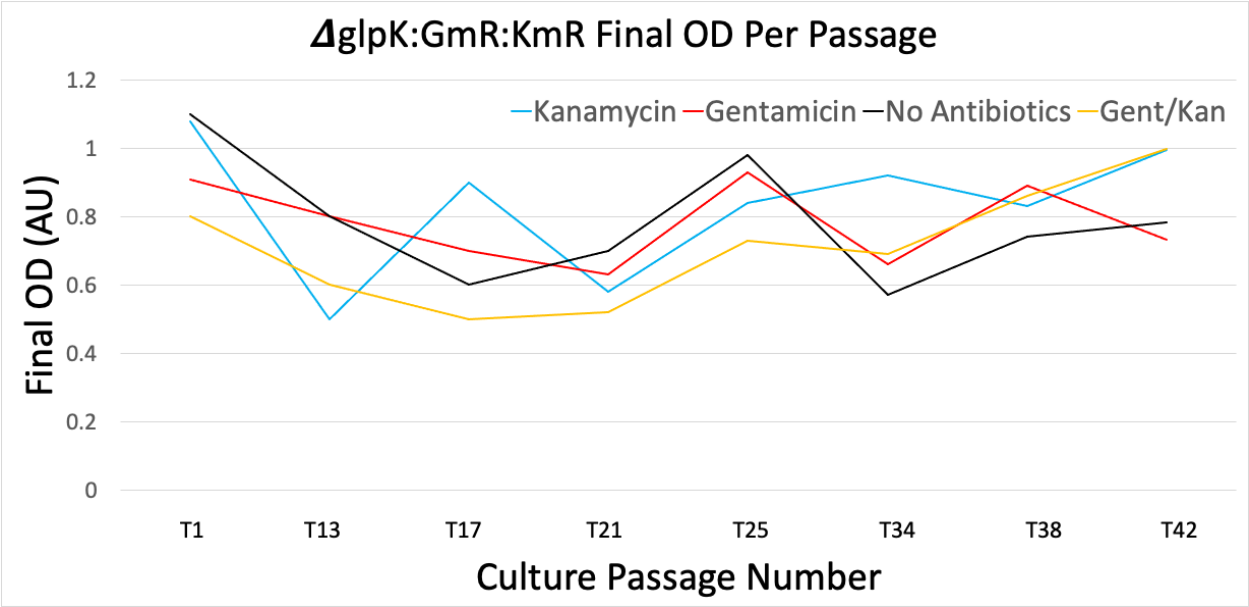


Figure 7. Passaging schematic and $OD_{750\ nm}$ density evolution of *ΔglpK:GmR+ΔglpK:KmR* T1-T42. **7A.** Depicts a schematic of the culture passaging of *ΔglpK:GmR+ΔglpK:KmR*. T1 culture was normalized in A+ media with gentamicin at 15µg/mL and kanamycin at 30µg/mL before seeding 4 separate conditions of no antibiotics, gentamicin at 5µg/mL and kanamycin at 30µg/mL, only gentamicin at 15µg/mL, and only kanamycin at 90µg/mL. Cells were passaged every 2-5 days with the prior culture seeding the new passage at an initial $OD_{750\ nm}$ of between 0.07-1.00. **7B.** Reveals the final OD of each passage (T1-T42) after 2-5 days incubation at 37°C, 25% white light, 100% red light and 230 rpm.

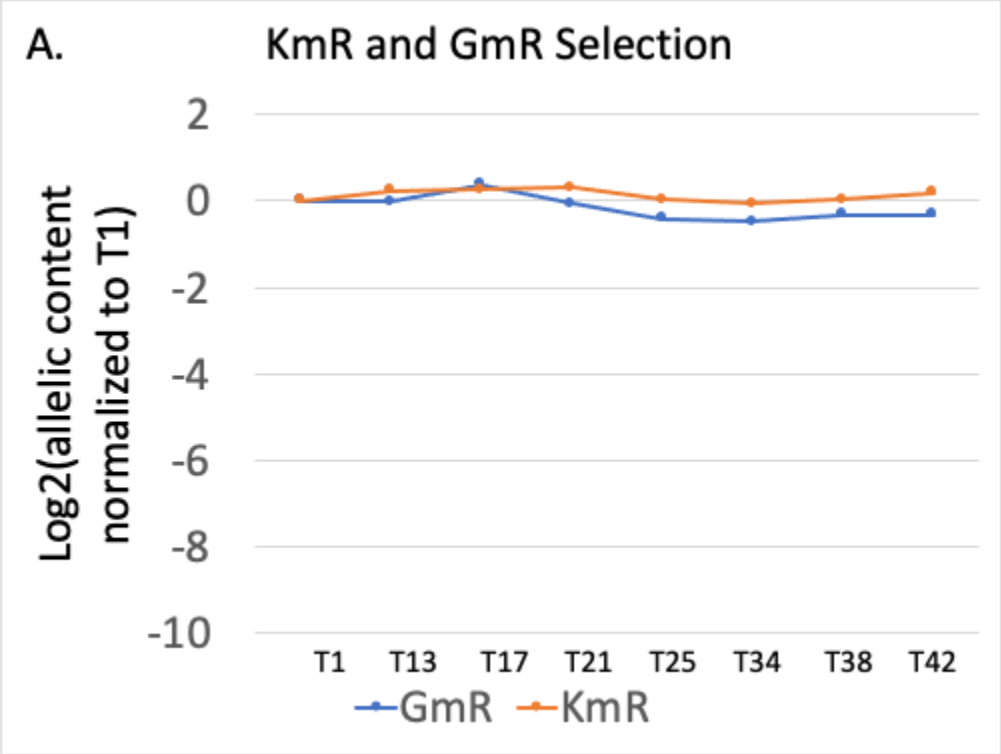
Evolution of KmR:GmR allelic ratios in 4 selectable conditions

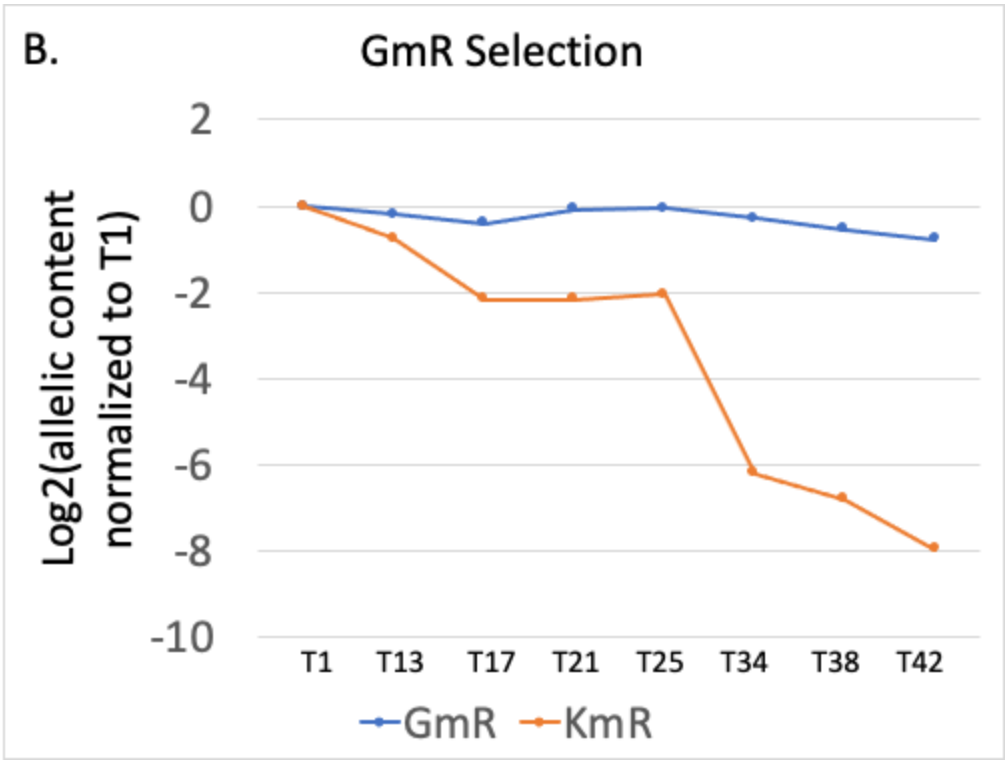
GmR:KmR allelic ratios were determined from DNA extracts from passages T1, T13, T17, T21, T25, T34, T38 and T42 through qPCR set to standard curves. The allelic ratios of GmR and KmR from each condition derived from qPCR analysis were normalized to chromosomal content within each sample, and converted to log(2) values and then graphed as a function of time (passage). Allelic content has been normalized to T1 for all conditions. **Figure 8** depicts the evolution of GmR:KmR allelic ratios under the selective pressure for both alleles (grown in kanamycin and gentamicin), only KmR (grown in kanamycin only), only GmR (grown in gentamicin only), and no selection (the absence of either antibiotic). In **Figure 8A** under the selective pressure of both gentamicin and kanamycin for both alleles, both GmR and KmR allele quantity remained approximately equal and constant, as expected. **Figure 8B** also revealed

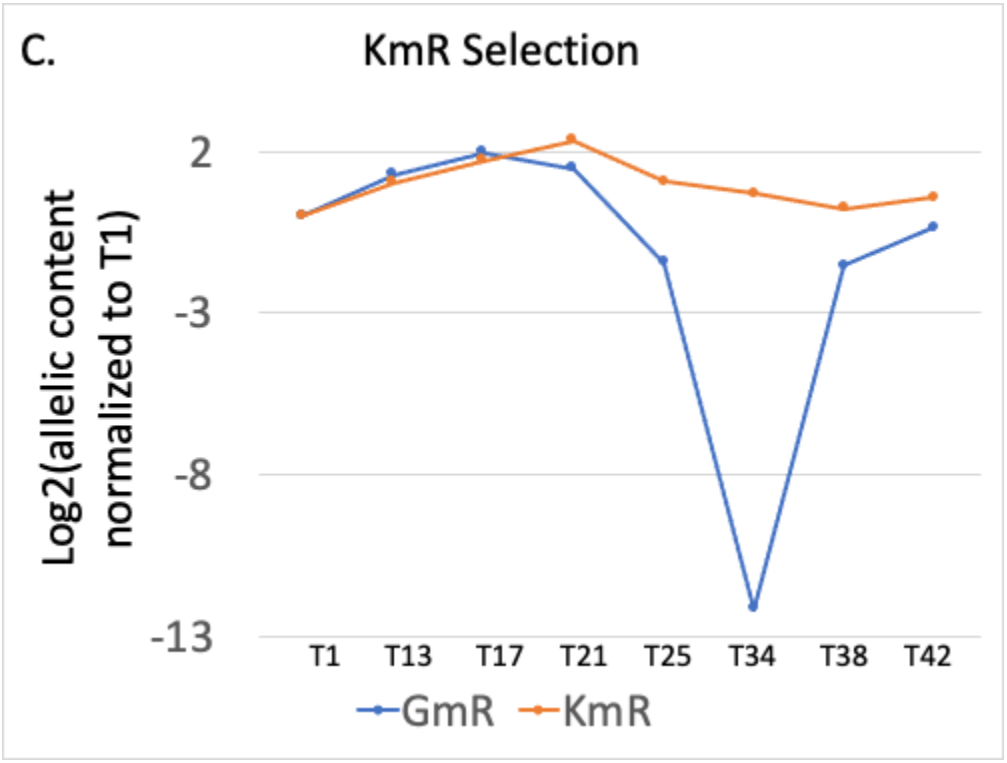
expected results under selection for GmR as seen in a steady decline in KmR allele content over time.

Figures 8C and **8D** yielded unexpected results. Under selective pressure for KmR, there is minimal change in GmR allele frequency and approximately equal KmR:GmR ratio for T13, T17 and T21. T25 indicates a decline in GmR allele frequency, and is exacerbated by a drastic decrease in GmR frequency in T34. However, this trend seems to immediately reverse in T38 and T42 with GmR allele frequency increasing in spite of selective pressure for KmR.

Figure 8D depicts KmR:GmR allelic ratios in the absence of selection. KmR frequency remains relatively constant throughout all passages. For the first five passages, GmR frequency steadily declines. At T38 and T42, however, GmR allele frequency begins to starkly increase, approaching a comparable value to KmR in the final passage, T42.




B.
GmR Selection
Log2(allelic content normalized to T1)
2
0
-2
-4
-6
-8
-10
T1
T13
T17
T21
T25
T34
T38
T42
GmR
KmR


C.
KmR Selection
Log2(allelic content normalized to T1)
2
-3
-8
-13
T1
T13
T17
T21
T25
T34
T38
T42
GmR
KmR

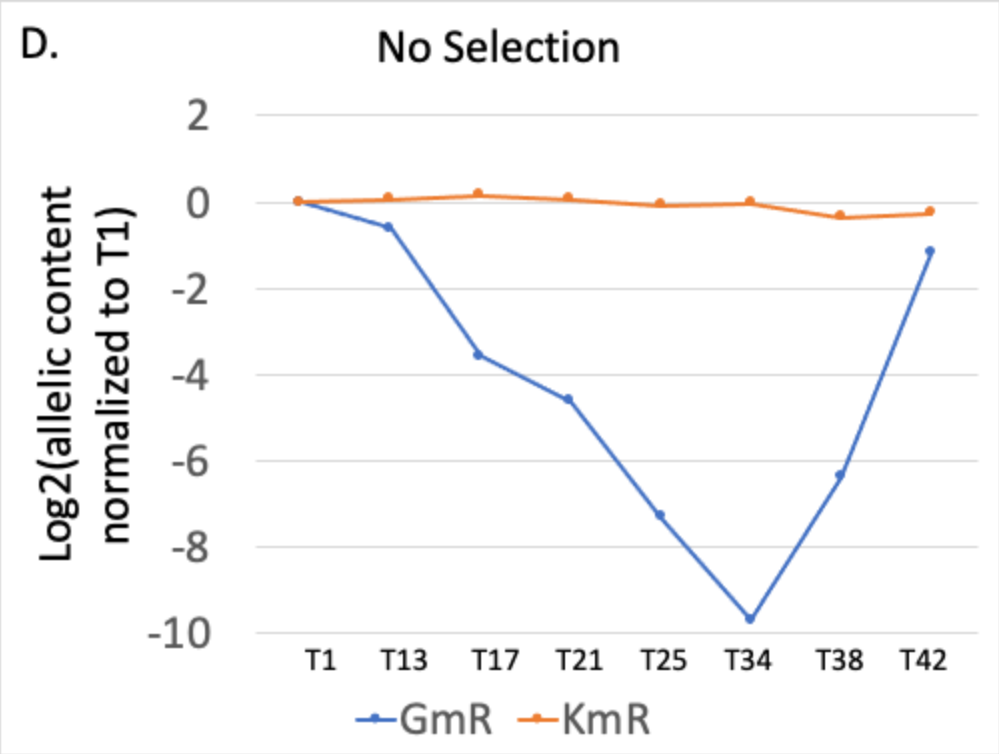


Figure 8. log(2) values of qPCR KmR:GmR allelic ratios in four separate conditions. Each graph represents GmR:KmR allelic ratios as a function of time under different selective pressures. **A.** KmR and GmR selective pressure maintain approximately equal and constant GmR:KmR frequencies. **B.** Under GmR selection, KmR allele frequency gradually decreases over time. **C.** KmR selection initially maintains equal GmR:KmR frequencies until T25 and T34 when GmR frequency drastically decreases. This trend immediately reverses in the following generations, T38 and T42 as GmR frequency increases. **D.** In the absence of selection, KmR frequency remains constant. GmR frequency gradually decreases in T13, T17, T21, T25 and T34 before a stark increase in frequency at T38, with the trend continuing to increase into T42.

Discussion

Lack of a clear role for priA or recG in chromosome replication in PCC 7002

In order for an organism to be utilized as a biological chassis for industrial and basic research, its mechanisms of DNA replication as well as chromosome stability and integrity must be thoroughly understood. The process of DNA transformation is the first step in designing a genetically modified organism (Ruffing et al., 2016), and DNA transformation is entirely dependent on the organism's replication mechanisms. Although PCC 7002 remains a strong candidate as a biological chassis, particularly for use in synthetic biology, both the consequences and mechanisms of its polyploidy

genome remain poorly understood. Here, we used CRISPR interference and conferred heterogeneity to determine possible genes regulating the DNA replication pathway and possible benefits of polyploidy in PCC 7002. Our results suggest that neither *priA* nor *recG*, alone play major roles in DNA replication initiation and that selective pressure or lack thereof plays an important role in regulating allelic ratios in a heterozygous polyploid bacterium. .

To inducibly knockdown *priA* and *recG in vivo*, a vector was required to guide the CRISPRi machinery to the genes to inhibit their transcription. This was achieved through the design of sgRNA-integrated plasmids. Two sgRNAs for each gene were selected to increase the likelihood of successful sgRNA plasmid design and then successful CRISPR knockdown of *priA* and *recG*. pJSF003 and pJSF005 were eventually the utilized guide RNAs as they were successfully transformed and amplified by *E. coli*, sequenced-validated (**Figures 3D and 3E**) and successfully transformed into scJC0164 PCC 7002 cells (**Figure 4B**).

There was not an obvious phenotypic effect on ploidy level after CRISPRi induction of *recG* knockdown (**Figure 5C**) compared to uninduced cells (**Figure 5B**) and -sgRNA cells (**Figure 5A**), this does not dismiss *recG*'s hypothesized function as a key component of the replisome and DNA damage repair pathway (Azeroglu et al., 2016). The cells harboring pJSF005, regardless of the addition of IPTG, had a tendency to appear more chlorotic in comparison to -sgRNA cells. There also appeared to be a higher frequency of dead cells in the *ΔscJC0164:pJSF005* colonies when imaged under fluorescent microscopy. These findings allude to some importance of *recG* in cell viability, if not direct association with chromosome replication initiation.

Similar results were observed for knockdown of *priA* (**Figure 5E**). There was no obvious phenotypic effect on cell length or ploidy level in comparison to uninduced cells (**Figure 5D**) or -sgRNA cells (**Figure 5A**). However, similarly to the knockdown of *recG*, cells harboring pJSF003, regardless of IPTG induction, overall appeared more chlorotic with higher frequency of dead cells imaged during microscopy. The overall compromised viability of both *ΔscJC0164:pJSF005* and *ΔscJC0164:pJSF003*, regardless of the addition of IPTG could be attributed to the presence of the CRISPRi machinery and leaky expression of its components, implying some level of knockdown

in the -IPTG samples. These results are in contrast to knockdown of a replisome-associated gene (*dnaX*), which resulted in obvious phenotypic differences compared to cells without knockdown, including cell elongation due to delayed division and decreased ploidy level (Moore et al., 2019).

priA is hypothesized to function as a stalled replisome rescue helicase (Duckworth et al., 2021), *recG* is a key player in DNA damage repair, and due to its photosynthetic properties, PCC 7002 is highly subjected to DNA damage (Cassier-Chauvat et al., 2016), therefore repressing the expression of these genes could explain the compromised viability regardless of phenotypic ploidy changes.

Environmental conditions appear to impact allelic ratios in PCC 7002

In addressing heterogeneity as a possible benefit of polyploidy in PCC 7002, we first conferred competing alleles at a single locus. This was achieved through the transformation of pre-designed GmR and KmR-containing plasmids into wild-type PCC 7002. Transformation was validated through colony PCR (**Figure 6**). All colonies were validated to have segregated barring colony 8, of which was missing a band in the *glpK* external lane (**Figure 6Bi**). This would indicate that this colony lacked a portion of the *glpK* cassette if bands were also missing from its KmR internal and GmR internal lanes, as these three PCRs used JCCo0389 as the reverse primer. The presence of bands at the expected lengths in colony 8 KmR internal and GmR internal (**Figure 6Bii**) indicates that the missing band in Figure 6.B.i was likely an experimental error. Colony 12 had expected band lengths in all PCRs and had the brightest band for GmR internal, thus it was selected as *ΔglpK:GmR+ΔglpK:KmR* culture to be normalized for seeding four conditions of varying selective pressures (**Figure 7A**).

The cells were passaged in each condition every 2-5 days for 25 weeks. The growth conditions and antibiotic concentrations were kept constant throughout the study. Fluctuations in $OD_{750\ nm}$ (**Figure 7B**) could be due to variation in incubation duration (2-5 days) as well flask location within the incubator. Throughout this experiment there was a correlation between where flasks were incubated, their final $OD_{750\ nm}$, and chlorotic appearance, indicating that some places in the incubator were more optimal either due to uneven heat distribution or light reflection. To mitigate the

impacts of differential growth conditions within the incubator, flasks were rotated to different positions throughout the 25 week experimental time course. There did not appear to be a direct correlation between culture location and extracted DNA concentration yields and purity noted in **Supplemental Figure Table S3**.

The DNA samples selected for quantitative PCR analysis were T13, T17, T21, T25, T34, T38, and T42 (**Supplemental Figure Table S3**). These samples were selected based on their DNA quality determined by a nanodrop spectrophotometer. Later samples were selected to account for the time component of evolution under selective pressure. Samples extracted from earlier passages were less likely to yield significant differences in KmR:GmR allelic ratios.

Examining the evolution of allelic ratios the the GmR selectable condition (cells grown only with gentamicin) and the GmR/KmR selection (cells grown with both kanamycin and gentamicin) yielded expected results. In the presence of both gentamicin and kanamycin, the ratio of both KmR and GmR alleles (**Figure 8A**) remained constant and equal due to the requirement for both alleles to survive in the presence of both antibiotics. The selection for GmR (**Figure 8B**) also yielded expected results in that, in the presence of gentamicin, the frequency of the KmR allele gradually decreased with each passage.

Surprisingly, however, in the presence of kanamycin, the GmR allele frequency was inconsistent (**Figure 8C**). The GmR:KmR ratios initially remained relatively constant and equal in the presence of kanamycin which could be attributed to kanamycin as a weak selective agent. However, at T25 there is a relatively dramatic decrease in GmR frequency, followed by a severe decrease at T34. The GmR frequency trend suddenly reverses at T38 where it continues to increase in frequency in the last two passages. It is unlikely that this is an accurate representation of the actual GmR ratio for the T34. The nanodrop qualities of the T34 DNA extraction were a 260/280 ratio of 1.83, a 260/230 ratio of 0.46 and a concentration of 165.5ng/μL.T34's 260/230 ratio of 0.46 indicates that the sample may include impurities that could have influenced qPCR results. Therefore, it is more likely that the actual T34 GmR frequency was closer to the T25 and T38 values, indicating a brief mild decline in GmR frequency before it began to trend up again in T42.

We were especially interested in how allelic ratios would be impacted by the absence of selection, as previous studies had indicated that complete gene conversion occurred between competing alleles in a polyploid haloarchaea species (Lange et al. 2011). However, the direction of the gene conversion coincided with an unsymmetrical aspect of the experimental design. The conversion to one allele required the synthesis of 53 bp while the conversion in the opposite direction required the synthesis of 958bp (Lange et al., 2011). The resulting conversion favored the smaller allele. Here, we eliminated this variable by conferring heterogenetic alleles of similar sizes. Surprisingly, however, there did still appear to be a trend towards the conversion to homogeneity in the absence of selection. **Figure 8D** reveals KmR allele frequency remaining constant throughout the study, while GmR frequency significantly decreases until T34, where it sharply reverts to increasing frequency. Gene conversion (converting one homologous gene to another through homologous recombination) is a possible explanation for the decrease in GmR frequency. However, if gene conversion were the sole mechanism, we also would have expected to see an increase in KmR frequency. KmR frequency remained relatively constant. Gene elimination of GmR, therefore, is more likely the explanation for the initial decrease in GmR frequency. It is also possible that gene conversion did revert and increased GmR frequency at later passages under these conditions. However, given the unexpected results of the same generations in the KmR selection, it is possible that the increase of GmR frequency in the final two generations in **Figure 8D** is a result of experimental error and/or DNA contamination. Further experiments and a replicate are garnered to definitely determine the correct occurrence.

Future Directions

A substantial amount of information remains unknown surrounding the mechanisms behind polyploidy in PCC 7002. Although this study offers some insight into the significance of *recG* and *priA* in this process, there are many additional experiments required to further elucidate these results. First, determining ideal growth times and incubator locations to ensure visualization of the cells is always during log phase when ploidy conditions are ideal. Second, cells that are induced to knockdown genes should also undergo RNA extraction and cDNA synthesis. Visualizing cDNA

through PCR with *priA* and *recG* specific primers and gel electrophoresis will confirm knockdown of the genes is actually occurring. These experiments were attempted but DNA contamination led to inconclusive results. In conjunction with the growth optimization and RNA extraction, the cells should then undergo quantitative chromosome analysis to determine finite ploidy level. It is possible there are changes in ploidy level that are subtle but quantitatively significant. Subtle phenotypic ploidy changes could also be better visualized through time-lapse microscopy. This study examined single cell generations of *priA* and *recG* knockdown. Time-lapse visualization would reveal whether effects of gene knockdown of *priA* or *recG* are not immediately obvious, as well as possible long term effects on offspring viability.

Additionally, evidence suggests that *recG* and *priA* function together in double-stranded DNA damage repair, specifically homologous recombination. RecG proposedly remodels the 3' end of stalled replication forks to permit PriA binding, where PriA functions as a helicase on the parental strand (Azeroglu et al., 2016). Therefore, *recG* and *priA* could simultaneously be knocked down to determine if PCC 7002 DNA replication is *recG-priA*-dependent. Alternative sgRNA sequences targeted to *priA* and *recG* could be explored to account for variability in knockdown efficacy between individual sgRNAs based on binding strength and on and off target scores. Alternative genes suspected to function in chromosome replication could also be targeted for knockdown.

To further elucidate the evolution of heterogeneity in polyploidy PCC 7002, further experiments are also required. First, the experiment described here should be repeated in a reversed transformation order. PCC 7002 cells should be transformed with KmR and then GmR. This could reveal whether or not the gene conversion/elimination process favors a certain allele or is stochastic. The study may also be repeated with increased antibiotic concentrations or a kanamycin-alternative and an antibiotic that is more comparable to gentamicin's strength as a selective agent. Additionally, the study should be repeated for a longer period of time. The 25 week duration of this study revealed changes in allelic ratios over time, but with a possible reversal in allele content towards the end of the study. The experiment should be repeated for a longer duration to determine if the change was experimental error or an actual fluctuation in allele

frequencies. Lastly, like the Lange study on *Haloferax volcanii*, an allele already present in the genome should be measured against a novel allele competing for its locus. This study examined two alleles novel to PCC 7002. It would be interesting to reveal whether the cell would favor gene retention of a wild type allele versus a novel allele, revealing whether or not allele retention is stochastic in the absence of selection.

It would also be interesting to examine the knockdown of genes suspected to function in homologous recombination (possibly *priA* and *recG*) in *ΔglpK:GmR+ΔglpK:KmR*. Homologous recombination is one major mechanism in gene conversion (Lange et al., 2011), therefore in knocking down a key driver in gene conversion, it may reveal whether gene conversion is a mechanism at play, and if so, if PCC 7002 would maintain a heterogenetic state if the homologous recombination pathway is compromised.

Ultimately, with the ever-increasing demand for sustainable and renewable modes of energy and chemical production, the demand for biological agents capable of executing these duties also increases. Yet, options for these agents remain limited. Cyanobacteria, particularly PCC 7002, are strong candidates to fulfill these demands, but the insufficient understanding of their genomic regulation hinders their use. In order to genetically modify PCC 7002, the genome regulation pathways and consequences of its polyploidy nature must be well defined. This study offers steps in revealing the significance of *priA* and *recG* on ploidy level in PCC 7002, as well as insight into the cells' evolution of allele retention in absence of selective pressure. It bolsters the efforts in advancing PCC 7002 as a viable chassis for use in synthetic and biological applications.

References

Azeroglu, B., Mawer, J. S. P., Cockram, C. A., White, M. A., Hasan, A. M. M., Filatenkova, M., & Leach, D. R. F. (2016). RecG Directs DNA Synthesis during Double-Strand Break Repair. *PLOS Genetics*, *12*(2), e1005799. https://doi.org/10.1371/journal.pgen.1005799

Blinkova, A., Hervas, C., Stukenberg, P. T., Onrust, R., O'Donnell, M. E., & Walker, J. R. (1993). The Escherichia coli DNA polymerase III holoenzyme contains both products of the dnaX gene, tau and gamma, but only tau is essential. *Journal of Bacteriology*, *175*(18), 6018–6027. https://doi.org/10.1128/jb.175.18.6018-6027.1993

Cassier-Chauvat, C., Veaudor, T., & Chauvat, F. (2016). Comparative Genomics of DNA Recombination and Repair in Cyanobacteria: Biotechnological Implications. *Frontiers in Microbiology*, *7*. https://www.frontiersin.org/article/10.3389/fmicb.2016.01809

Chen, A. H., Afonso, B., Silver, P. A., & Savage, D. F. (2012). Spatial and Temporal Organization of Chromosome Duplication and Segregation in the Cyanobacterium Synechococcus elongatus PCC 7942. *PLOS ONE*, *7*(10), e47837. https://doi.org/10.1371/journal.pone.0047837

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nature Reviews Genetics*, *6*(11), 836–846. https://doi.org/10.1038/nrg1711

Doench, J. G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E. W., Donovan, K. F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., Virgin, H. W., Listgarten, J., & Root, D. E. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature Biotechnology*, *34*(2), 184–191. https://doi.org/10.1038/nbt.3437

Duckworth, A. T., Windgassen, T. A., & Keck, J. L. (2021). Examination of the roles of a conserved motif in the PriA helicase in structure-specific DNA unwinding and processivity. *PLOS ONE*, *16*(7), e0255409. https://doi.org/10.1371/journal.pone.0255409

Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009).

Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods*, *6*(5), 343–345. https://doi.org/10.1038/nmeth.1318

Lange, C., Zerulla, K., Breuert, S., & Soppa, J. (2011). Gene conversion results in the equalization of genome copies in the polyploid haloarchaeon Haloferax volcanii. *Molecular Microbiology*, *80*(3), 666–677. https://doi.org/10.1111/j.1365-2958.2011.07600.x

Ludwig, M., & Bryant, D. (2012). Synechococcus sp. Strain PCC 7002 Transcriptome: Acclimation to Temperature, Salinity, Oxidative Stress, and Mixotrophic Growth Conditions. *Frontiers in Microbiology*, *3*. https://www.frontiersin.org/article/10.3389/fmicb.2012.00354

Markley, A. L., Begemann, M. B., Clarke, R. E., Gordon, G. C., & Pfleger, B. F. (2015). Synthetic Biology Toolbox for Controlling Gene Expression in the Cyanobacterium Synechococcus sp. Strain PCC 7002. *ACS Synthetic Biology*, *4*(5), 595–603. https://doi.org/10.1021/sb500260k

Markov, A. V., & Kaznacheev, I. S. (2016). Evolutionary consequences of polyploidy in prokaryotes and the origin of mitosis and meiosis. *Biology Direct*, *11*(1), 28. https://doi.org/10.1186/s13062-016-0131-8

Moore, K. A., Tay, J. W., & Cameron, J. C. (2019). *Multi-generational Analysis and Manipulation of Chromosomes in a Polyploid Cyanobacterium* (p. 661256). bioRxiv. https://doi.org/10.1101/661256

Ohbayashi, R., Hirooka, S., Onuma, R., Kanesaki, Y., Hirose, Y., Kobayashi, Y., Fujiwara, T., Furusawa, C., & Miyagishima, S. (2020). Evolutionary Changes in DnaA-Dependent Chromosomal Replication in Cyanobacteria. *Frontiers in Microbiology*, *11*. https://www.frontiersin.org/article/10.3389/fmicb.2020.00786

Ohbayashi, R., Watanabe, S., Ehira, S., Kanesaki, Y., Chibazakura, T., & Yoshikawa, H. (2016). Diversification of DnaA dependency for DNA replication in cyanobacterial evolution. *The*

ISME Journal, *10*(5), 1113–1121. https://doi.org/10.1038/ismej.2015.194

Ruffing, A. M., Jensen, T. J., & Strickland, L. M. (2016). Genetic tools for advancement of Synechococcus sp. PCC 7002 as a cyanobacterial chassis. *Microbial Cell Factories*, *15*, 190. https://doi.org/10.1186/s12934-016-0584-6

Watanabe, S. (2020). Cyanobacterial multi-copy chromosomes and their replication. *Bioscience, Biotechnology, and Biochemistry*, *84*(7), 1309–1321. https://doi.org/10.1080/09168451.2020.1736983

Schematics created using Bio Render

Supplemental Figures

Oligo Name	Sequence 5'-3'	Target	Use
JSFO_005	*AGTGCCGCTAGAATTGGCGG***GTTTTAGAGC TAGAAATAGC**	F	*priA* sgRNA 2-piece Phusion insertion
JSFO_006	*CCGCCAATTCTAGCGGCACT***TGTGTGAAAT TGTTATCCGC**	R	*priA* sgRNA 2-piece Phusion insertion
JSFO_009	*TTATCCATCGAAGCAGAACG***GTTTTAGAGCT AGAAATAGC**	F	*recG* sgRNA 2-piece Phusion insertion
JSFO_010	*CGTTCTGCTTCGATGGATAA***TGTGTGAAATT GTTATCCGC**	R	*recG* sgRNA 2-piece Phusion insertion
KAMo0166	AACTGGATCTCAACAGCGGT	R	sgRNA 2-piece Phusion insertion
KAMo0167	ACCGCTGTTGAGATCCAGTT	F	sgRNA 2-piece Phusion insertion
KAM0100	CAGACCGCTTCTGCGTTCTG	R	sgRNA segregation validation
Pet_81F	CTTCAGCTTACAGTGACGCCTGTC	F	sgRNA segregation validation
JSFo_023	CGACAGCCAAATTATCATCG	F	*priA* expression +/- IPTG
JSFo_024	CCTGAACGTAATGGTAGGAG	R	*priA* expression +/- IPTG
JSFo_025	GATCACTGAAATTCTTGCGG	F	*recG* expression +/- IPTG
JSFo_026	CAGTAAACCCACGTTAAAGC	R	*recG* expression +/- IPTG

Table S1. All 5’-3’ primer sequences used in the experiments examining mechanisms of polyploidy via CRISPRi.
F denotes ”forward” primers, R denotes “reverse” primers. Italicized portion of sequences in JSFO_005-009 denote sgRNA sequence, bold portion denotes appended primer. Pet_81F and all KAM primers were courtesy of Cameron Lab.

Oligo Name	Sequence 5'-3'	Target	Use
JSFO.013	AACTTGCTCCGTAGTAAGAC	F	GmR cassette qPCR
JSFO.014	ACGTAGATCACATAAGCACC	R	GmR cassette qPCR
JSFO.017	TATGAGCCATATTCAACGGG	F	KmR cassette qPCR
JSFO.018	TAACCATGCATCATCAGGAG	R	KmR cassette qPCR
NCH_9F	CGGCCTTGGCCTGCATAT	F	GmR cassette segregation validation
JCCo0388	CAATGGCGAAGGTTTTCTGT	F	*glpK* upstream external
JCCo0389	GGGAGATGCTGTAGGCAAGA	R	*glpK* downstream external
EBJp0082	GATCTTGCCATCCTATG	F	KmR cassette segregation validation
EBJo0178	CTTGGGCAAGATCATTC	F	*glpK* internal
EBJo0179	GATCACAGTAATGCCAG	R	*glpK* internal
KAMo0149	GCAACTGACCCCGACTACAT	F	WT sequence control
KAMo0150	CAGACGTTGGGTGAGGAAAT	R	WT sequence control

Table S2. All 5’-3’ primer sequences used in the experiments examining consequences of polyploidy via conferred heterogeneity at a single competing locus.
F denotes ”forward” primers, R denotes “reverse” primers. All primers aside from JSFO were courtesy of Cameron Lab.

Generation	Condition	260/280	260/230	Concentration ng/µL		Generation	Condition	260/280	260/230	Concentration ng/µL
T1	BOTH	1.89	2.31	101		T21	NONE	1.87	1.2	160.6
T5	GM	1.84	0.63	33.1		T21	BOTH	2.14	2.43	50.9
T5	KM	1.77	0.71	42		T25	GM	1.87	0.92	167.4
T5	NONE	1.79	0.99	53.9		T25	KM	1.85	1.32	197.9
T5	BOTH	1.73	1.12	57.7		T25	NONE	1.84	1.56	179.6
T9	GM	1.8	1.2	62.8		T25	BOTH	1.86	1.08	136.5
T9	KM	1.8	0.92	58		T34	GM	1.85	0.67	115.6
T9	NONE	1.72	0.56	70.3		T34	KM	1.83	0.46	165.5
T9	BOTH	1.7	0.53	51.3		T34	NONE	1.85	1.34	270.8
T13	GM	1.78	0.65	59.4		T34	BOTH	1.84	1.42	148.9
T13	KM	1.84	1.04	146.3		T38	GM	1.84	1.47	100.1
T13	NONE	1.82	1.03	134.7		T38	KM	1.86	0.96	136
T13	BOTH	1.77	0.55	117.7		T38	NONE	1.88	1.65	163.3
T17	GM	1.82	0.86	80.7		T38	BOTH	1.84	1.26	94.5
T17	KM	1.86	1.69	135.6		T42	GM	2.29	2.36	43
T17	NONE	1.78	1.07	98.7		T42	KM	1.8	0.87	95.8
T17	BOTH	1.78	0.81	47.7		T42	NONE	1.82	1.4	76.5
T21	GM	1.93	2.06	80		T42	BOTH	1.85	0.97	140.7
T21	KM	1.85	1.36	183.8						

Table S3. Quality and concentrations of *ΔglpK:GmR+ΔKmR* DNA extracts determined by Nanodrop spectrophotometer. Ideal 260/280 ratios were >1.8 and >0.5 for 260/230 ratios. Ideal DNA concentrations were >60ng/µL for samples used in qPCR analysis.

DNA Size (bp)	% Agarose
800-12000	0.7
500-10000	1
400-7000	1.5
200-3000	1.7

Table S3. Percent gel agarose to 1X TAE used to visualize PCR products via gel electrophoresis. Most PCR products in this work were between 200bp and 3000bp. Percent agarose used to visualize products was determined based on the longest expected PCR product length.