Evolution of a minimal cell

Moger-Reischer RZ[1], Glass JI[2], Wise KS[2], Sun L[2,3], Bittencourt D[2,4], Lynch M[5], Lennon JT[1,*]

[1] Department of Biology, Indiana University, Bloomington, IN 47405, USA

[2] J. Craig Venter Institute, La Jolla, CA 92037, USA

[3] Sorrento Therapeutics, Inc., San Diego, CA 92121 USA

[4] Embrapa Genetic Resources and Biotechnology, Brasília, 70770-917, Brazil

[5] Arizona State University, Tempe, AZ 85287, USA

* Corresponding author: lennonj@indiana.edu

ABSTRACT

Possessing only essential genes, a minimal cell can reveal mechanisms and processes that are critical for the persistence and stability of life. Here, we report on how a synthetically constructed minimal cell contends with the forces of evolution compared to a non-minimized cell from which it was derived. Genome streamlining was costly, but 80% of fitness was regained in 2000 generations. Although selection acted upon divergent sets of mutations, the rates of adaption in the minimal and non-minimal cell were equivalent. The only apparent constraint of minimization involved epistatic interactions that inhibited the evolution of cell size. Together, our findings demonstrate the power of natural selection to rapidly optimize fitness in the simplest autonomous organism, with implications for the evolution of cellular complexity.

INTRODUCTION

The complexity of an organism is reflected in the number of genes it possesses, a quantity that varies by orders of magnitude across the tree of life. While some obligately endosymbiotic bacteria have fewer than 200 protein-coding genes, many plant and animal genomes contain more than 20,000 genes (Lynch 2007). In principle, the simplest organism would have no functional redundancies and possess only the minimum number of genes essential for life. Any mutation in such an organism could lethally disrupt one or more cellular functions, placing constraints on evolution, as reflected in the fact that essential proteins change more slowly than those encoded by dispensable genes (Graur and Li 2000; Hahn and Kern 2005). Furthermore, organisms with streamlined genomes have fewer targets upon which positive selection can act, thus limiting opportunities for adaptation. Understanding how organisms with simplified genomes overcome these evolutionary challenges has important implications for long-standing problems in biology, including the treatment of clinical pathogens, the persistence of host-associated endosymbionts, the refinement of engineered microorganisms, and the origin of life itself (Leprince et al. 2012; Moran and Bennett 2014; Glass et al. 2017).

The cell is the simplest independent functional unit of life. Yet even unicellular model organisms that are touted for their tractability are complex, possessing thousands of genes and proteins, many of which remain uncharacterized even after decades of in-depth interrogation. The quest for the simplest organism has been aided by advances in synthetic biology, which involves the redesign or novel construction of biological parts and modules (Benner and Sismour 2005; Glass et al. 2017). Promoted as a way to address practical challenges for society (Cameron et al. 2014; DeLisi 2019), synthetic biology also provides a platform for developing powerful simplest-

case models through streamlining, whereby nonessential chromosomal sequences are removed from an organism's genome (Leprince et al. 2012; Hutchison et al. 2016; Glass et al. 2017; Richardson et al. 2017; Lachance et al. 2019). Guided by such strategies, a "minimal cell" was constructed with a genome containing only the smallest set of genes required for autonomous cellular life (Hutchison et al. 2016; Breuer et al. 2019). While these efforts succeeded in experimentally identifying the genetic requirements for basic cellular processes such as metabolism and cell division, it remains unclear how a minimal cell should respond to the forces of evolution, particularly given the limited raw materials upon which natural selection can operate as well as the uncharacterized input of new mutations.

To gain insight into the dynamics and outcomes of evolution in a minimal cell, we conducted experiments with synthetic strains of *Mycoplasma mycoides* (Hutchison et al. 2016; Breuer et al. 2019). The minimal cell, *M. mycoides* JCVI-syn3B, was derived from a non-minimal strain, *M. mycoides* JCVI-syn1.0, by reducing the chromosome from 901 to 493 genes, resulting in the smallest genome of any known free-living organism (Hutchison et al. 2016; Breuer et al. 2019). With these two strains, we first investigated whether genome minimization, which included the removal of two DNA replication genes and eight DNA repair genes, altered the rate and spectrum of new mutations in the minimal cell relative the non-minimal organism. Second, with knowledge of the mutational input, we evaluated whether genome minimization altered the rate and mechanisms of evolution in response to natural selection, as measured using whole-genome sequencing, estimates of population fitness, and phenotypic changes in cell size.

RESULTS AND DISCUSSION

Highest recorded mutation rate is robust to genome minimization

Through serial bottlenecking under relaxed selection, we conducted mutation-accumulation experiments with populations of *Mycoplasma mycoides* (see Methods). The number of mutations per nucleotide per generation for the non-minimal cell ($3.13 \pm 0.12 \times 10^{-8}$, mean ± SEM) was indistinguishable from that of the minimal cell ($3.25 \pm 0.16 \times 10^{-8}$) ($t_{140} = 0.43$, $P = 0.667$, Fig. 1a). These mutation rates, which are the highest recorded for any cellular organism, are consistent with other reports where organisms with smaller genomes have higher mutation rates (Kuo and Ochman 2009; Sung et al. 2012; Lynch et al. 2016; Long et al. 2018). For example, *Mesoplasma florum* also has an exceptionally high mutation rate ($1.16 \pm 0.07 \times 10^{-8}$) (Sung et al. 2012), which may reflect the fact that these closely related bacteria both lack mismatch repair, a system for recognizing and fixing errors that are made during DNA replication (Carvalho et al. 2005). Our data are also consistent with predictions from the drift-barrier hypothesis. This theory posits that mutation rates evolve downward until the selective advantage of another incremental decrease in the mutation rate is small enough to be effectively neutral and outweighed by genetic drift (Sung et al. 2012; Lynch et al. 2016). In other words, populations with lower effective population size (N_e) experience stronger drift, and hence evolve higher mutation rates (Lynch et al. 2016). Notably, the natural history (obligate pathogen) and genomic features (small genome size and low GC content) of *M. mycoides* and *Me. florum* suggest these mollicute bacteria most likely have low N_e (Moran 2003; Kuo and Ochman 2009; Hershberg 2015; Long et al. 2018).

Small effect of genome minimization on mutational spectrum

Although the mutation rate was unaffected by minimization, the types of mutations that arise in a

94 population can still influence evolution. Therefore, we compared the spectra of mutations in the minimal and non-minimal *M. mycoides.* Overall, the composition of mutation types (insertions, 96 deletions, and single-nucleotide mutations) was unaffected by genome minimization ($\chi^2_2 = 4.16$, $P = 0.125$) (Fig. 1b). Both strains exhibited bias toward deletions over insertions (non-minimal: 98 $\chi^2_1 = 21.3$, $P = 4.0 \times 10^{-6}$; minimal: $\chi^2_1 = 68.4$, $P < 2.2 \times 10^{-16}$), consistent with findings reported for other bacteria (Kuo and Ochman 2009; Hershberg 2015; Sung et al. 2016). However, the com-100 position of single-nucleotide mutations (SNMs), which constituted the largest category of mutations (88%), differed between the minimal and non-minimal cell (Monte Carlo $\chi^2 = 69.9$, $P = 1.0$ 102 $\times 10^{-4}$). Mutations from a G or C nucleotide to an A or T nucleotide occurred at a rate ~30-fold higher in the non-minimal cell ($\chi^2_1 = 3736$, $P < 2.2 \times 10^{-16}$) and ~100-fold higher in the minimal 104 cell ($\chi^2_1 = 1444$, $P < 2.2 \times 10^{-16}$) compared to mutations from A or T to G or C (Fig. 1c). As a consequence, *M. mycoides* exhibits more A:T bias than any known organism (Hershberg 2015; 106 Long et al. 2018).

108 There are a few explanations for the extreme A:T mutational bias in the minimal cell. First, *M. mycoides* naturally lacks Dut, a dUTPase protein that draws down the intracellular concentration of dUTP, thereby avoiding its misincorporation into DNA (Tye and Lehman 1977). Because 110 adenine would pair with a misincorporated uracil, the lack of Dut should lead to an increased 112 frequency of A:T-biased mutations in the spectrum of cells without Dut (Sedwick et al. 1986). However, since both cells lack Dut, this mechanism is insufficient to explain the discrepancy in 114 A:T bias between the non-minimal and minimal cells, respectively ($\chi^2_1 = 21.8$, $P = 3.08 \times 10^{-6}$). Instead, the difference in A:T bias is more likely due to the removal of *ung* during minimization

(Hutchison et al. 2016). This nonessential DNA repair gene encodes for the uracil-DNA glycosylase enzyme Ung, which excises ectopic uracils from DNA via hydrolysis. The process results in a free uracil and an abasic site in the DNA strand (Lindahl et al. 1977), which allows the correct nucleotide to be inserted using base complementarity. Without Ung, the uracil would remain in place and be paired with A during DNA replication, even if the original nucleotide was not T. One common cause for the presence of uracil in DNA is the deamination of cytosine (Lindahl 1974). Thus, a cell lacking *ung* would be enriched in C-to-T mutations when C is replaced by U via deamination and U is replaced by T in subsequent DNA replication via base-pairing with A. In the minimal cell, we observed an excess of these C:G-to-T:A mutations relative to the non-minimal bacterium ($\chi^2_1 = 89.8$, $P = 1.60 \times 10^{-20}$), consistent with this aspect of the mutation spectrum being related to the minimal cell lacking *ung* (Fig. 1c).

Rapid recovery of fitness in the minimal cell

We next considered how populations of the minimal and non-minimal cell would respond to natural selection. With measured mutation rates of ~3×10^{-8} per nucleotide per generation and population sizes in excess of 10^7 individuals, we estimated that a new mutation would hit every nucleotide in the genome >250 times during 2000 generations of experimental evolution. Thus, given their equal mutation rates per site, populations of either cell type should not be limited by the availability of genetic variation to fuel adaptation, suggesting that any differences in the ways the two strains adapt would be driven by alterations in genome architecture created by minimization.

To study natural selection, we passaged replicate populations of *M. mycoides* for ~2000 generations (see Methods), a period during which rapid adaptation is often observed (Vasi et al.

1994; Gifford et al. 2011). At the end of the experiment, we measured the fitness (competitive growth rates) of all ancestors and evolved populations (see Methods). For the ancestral strains, we determined that genome minimization led to a 50% reduction in fitness (Fig. 2). Despite this major initial cost, the minimal cell rapidly regained fitness during the course of evolution. Approximately 80% of the fitness lost to minimization was recovered during 300 days of serial passaging (Fig. 2). This suggests that a streamlined *M. mycoides* genome is not inherently crippled and can perform almost as well as the non-minimized cell following readaptation.

Contrary to our initial predictions, adaptation was not constrained by genome minimization (Fig. 2). The rate of fitness gain was the same for replicate populations of the non-minimal and minimal cells ($t_6 = 0.40$, $P = 0.704$, Fig. 2). This interpretation was bolstered by results from population-genomic sequencing (see Methods). The relative ratio of nonsynonymous to synonymous fixed SNMs (d_N/d_S) did not differ between the two cell types ($t_6 = 1.18$, $P = 0.859$, Fig. S1), consistent with the interpretation that the rates of molecular evolution were comparable even though all of the genes in the minimal cell are critical for fitness (Graur and Li 2000; Hahn and Kern 2005).

Divergent mechanisms of adaptation

Using a combination of statistical simulation and reverse genetics, we identified mutations that likely contributed to the observed patterns of adaptation. First, we analyzed the gene-by-population matrix for non-synonymous mutations that arose in essential genes during the natural selection experiment (see Methods). The two cell types acquired mutations in different sets of essential genes (PERMANOVA, $F_7 = 4.12$, $P = 0.029$, Fig. 3), suggesting that, despite adapting at similar

rates, the populations evolved via divergent routes. To explore this hypothesis, we looked for genes that acquired a higher number of non-synonymous mutations than expected under assumptions of neutrality (see Methods). We identified 16 genes in the non-minimal genome and 14 in the minimal genome that were potential targets of positive selection (Table 1). Second, we used reverse genetics to experimentally verify that one of the common types of mutation observed in replicate populations of both strains was in fact beneficial (Table S1). Using CRISPR editing, we recreated *ftsZ* C-terminal nonsense mutations by inserting an *ftsZ* E315* mutation into the ancestral genomes of the minimized and non-minimized strains (see Methods). Competition assays with the constructs revealed that this putatively adaptive mutation conferred a 25% fitness increase in the non-minimal cell and a 14% fitness increase in the minimal cell (Fig. S2).

Comparative analysis of genes under positive selection provided insight into the functional consequences of adaptation in the minimal cell. We hypothesized that mutations in genes related to membrane transport would be critical for adaptation because the minimal cell relies on the import and export of metabolites and other biomolecules for metabolism (Antczak et al. 2019; Breuer et al. 2019). However, mutations in membrane transport functions were enriched to a similar degree in both cell types over 2000 generations of evolution (Fisher's Exact Test $P = 0.934$). Instead, we detected a signal of enrichment for mutations in biosynthetic genes for the minimal cell (Fisher's Exact Test $P = 0.090$), including those involved in lipid metabolism. Specifically, *fakA* and *clsA* (Table 1) are considered essential for synthesizing cardiolipin and other lipids from free fatty acids (Breuer et al. 2019), which are important for the construction of cell membranes and the regulation of cell division. The gene *lgt* is also critical for membrane construction, encoding the protein that transfers diacylglyceryl moieties to anchor surface lipoproteins in the lipid bilayer

(Breuer et al. 2019). Thus, metabolic innovations involving lipid synthesis and distribution may be more important for the minimal cell than enhanced acquisition of metabolites that are already present in the growth medium.

Genome minimization constrains evolution of cell size

The size of single-celled organisms is variable and often linked to fitness in complex ways (Gounand et al. 2016; Mei et al. 2009; Chien et al. 2012). In resource-rich environments, cell size tends to be positively correlated with growth rate, one of the most important components of fitness (Vasi et al. 1994; Mongold and Lenski 1996; Mei et al. 2009; Chien et al. 2012; Hill et al. 2013). For example, in the first 2000 generations of a classic long-term evolution experiment, cell volume and fitness concomitantly increased by 50% and 30%, respectively (Vasi et al. 1994). While an increase in size can accommodate more macromolecules needed for growth and division, it also decreases a cell's surface-to-volume ratio which increases molecular diffusion times. Given these opposing pressures, we evaluated how cell size changed in response to genome streamlining over the course of evolution. Prior to initiating experiments, we observed that the spherically shaped non-minimal cell was 575 ± 28 nm (mean ± SEM) and the minimal cell was 422 ± 23 nm (see Methods), comparable to measures of cell diameter previously reported (Hutchison et al. 2016; Breuer et al. 2019). Following 2000 generation of evolution, the non-minimal cell exhibited a 73% increase in diameter (993 ± 53 nm, $t_3 = 7.84$, $P = 0.004$) and a corresponding five-fold increase in volume compared to its ancestor (Fig. 4; Table S2). In contrast, the size of the minimal cell did not change during evolution ($t_3 = 0.85$, $P = 0.459$).

While cell size is a complex multigenic trait, previous studies have implicated *ftsZ* with

changes in morphology of the minimal cell (Pelletier et al. 2021). It encodes for FtsZ, a protein that localizes to the midcell and determines the site of membrane constriction during cell division. Prevalent among diverse lineages of bacteria and archaea (McQuillen and Xiao 2020; Liao et al. 2021), *ftsZ* is nevertheless nonessential in *M. mycoides*. However, cells lacking *ftsZ* exhibit aberrant cell division and morphology (Hutchison et al. 2016; Breuer et al. 2019; Pelletier et al. 2021). Thus, along with 18 other nonessential genes, *ftsZ* was retained in JCVI-syn3B to aid in culture maintenance and stable growth (Breuer et al. 2019; Pelletier et al. 2021). In our study, *ftsZ* was consistently mutated over 2000 generations of evolution and was identified as a target of positive selection in both the minimal and non-minimal cell (Table 1, Fig. S2). Introduction of this termination codon could eliminate the C-terminal region of FtsZ, known to interact with membrane-associated products that recruit FtsZ (Cohan et al. 2020), The early stop codon could also create a transcriptional polar effect (Graffeuil et al. 2020) that reduces expression of two adjacent downstream genes within a likely polycistronic operon: MMSYN1_0521, an ortholog of cell division protein *sepF* and MMSYN1_0520, encoding a hydrolase of unknown functions. Irrespective of mechanism, we demonstrated that mutations in *ftsZ* had a non-additive effect that contributed to the evolutionary divergence of cell size. Using scanning electron microscopy, we documented that the *ftsZ* E315* nonsense mutation in the ancestral non-minimal cell led to a 25% increase in cell diameter ($t_{118} = 5.56$, $P = 8.71 \times 10^{-8}$) and a corresponding two-fold increase in cell volume. In contrast, the same *ftsZ* nonsense mutation in the ancestral minimal cell led to a 19% decrease in the cell diameter ($t_{122} = 3.29$, $P = 0.001$, Fig. S4), which reduced cell volume by half. Thus, the *ftsZ* E315* mutation recapitulated nearly 60% of the evolved divergence in cell size, indicating that FtsZ plays a central role in cell size in *M. mycoides*.

Although changes in *ftsZ* had opposing effects on the size of the minimal and non-minimal cell, mutations in this gene were adaptive for both strains (Fig. S2, Table 1). Smaller cell size may be advantageous for the minimal cell. Owing to its slower growth rate, JCVI-syn3B would have spent more time in an exponential growth phase in our evolution experiment. Some studies have shown that smaller cell size is selected for when bacteria evolve under such conditions (Gounand et al. 2016) even though other studies report a positive association between metabolic rate and body mass across species (DeLong et al. 2010). Meanwhile, the faster-growing non-minimal cell should experience bouts of feast-or-famine over thousands of generations in a serial batch environment. Under such conditions, repeated transitions between exponential and stationary growth phases has been shown to select for an increase in cell size (Vasi et al. 1994; Gounand et al. 2016). The observed patterns may also reflect constraints imposed by genome streamlining on the ability of the minimal cell to evolve an adaptive increase in cell size (Mei et al. 2009; DeLong et al. 2010; Chien et al. 2012; Westfall and Levin 2017). With more than 50% of its membrane transport proteins removed, the minimal cell may be unable to sequester the resources needed for constructing and maintaining a larger cell (Chien et al. 2012; Hutchison et al. 2016; Westfall and Levin 2017). However, one might expect that these biosynthetic constraints would be reflected in a reduced rate of adaptation, which was not the case (Fig. 2). In addition, cell size could evolve as a fitness-neutral byproduct of selection on other traits, such as DNA replication rate (Amir 2017). For example, the two strains could have evolved different size trajectories despite similar selection pressures, due to epistatic effects of genome minimization such as those demonstrated using the *ftsZ* E315* mutants (Table S3, Fig. S4). In any case, our findings highlight that cell size, a fundamental feature of biological complexity in multicellular and single-celled organisms alike, evolves in a way that is highly dependent on genomic context.

Outlook

We documented genes, proteins, and traits that are critical for evolutionary performance in the synthetically constructed *M. mycoides*, a bacterium with the smallest genome of any self-replicating form of life. In its ancestral state, this working approximation of a minimal cell had low fitness, fewer than 500 protein-coding genes, and the highest mutation rate ever recorded in a cellular organism. A single gene disruption could have proven lethal. Instead, the force of natural selection outweighed any deleterious consequences of mutation and drift that could have led populations to extinction. Yet, even in one of the simplest organisms, evolution was still complex. Rapid adaptation involved epistatic interactions and selection on divergent targets, 25% of which encoded for proteins of unknown function. Evolution studies like the ones described here can improve gene characterization and the mapping of regulatory networks, which will inform subsequent rounds of rational engineering and streamlining. In addition, minimal cells can be leveraged to address long-standing problems in biology. Derived from suites of diverse ancestors, they may serve as new models for investigating innovations of early life (Kacar et al. 2017), while providing bottom-up insight into the development of multicellularity (Márquez-Zacarías et al. 2021), the mechanisms of coevolution (Hillesland and Stahl 2010), and the assembly of multi-species consortia and their emergent functions (Goldford et al. 2018).

METHODS

Strains and growth conditions

We maintained synthetic *Mycoplasma mycoides* JCVI-syn1.0 and synthetic *M. mycoides* JCVI-syn3B in SP4 medium with KnockOut™ Serum Replacement (Gibco) substituted for fetal bovine

serum (Table S4). Cultures were grown in a dark, static growth chamber at 37 °C. The non-minimal JCVI-syn1.0 strain has been described in detail elsewhere (2010). The minimal JVCI-syn3B is identical to the strain synthesized in previous studies (2016) with the following exceptions: JVCI-syn3B possesses a second rRNA operon copy, lacks a gene (MMSYN1_0531) coding for an efflux protein, and has 19 genes that were added back into the minimal genome to render the cell easier to use (Breuer et al. 2019; Pelletier et al. 2021) (Table S5). The strain also contains a landing pad system (*Cre* recombinase and *LoxP*) facilitating genetic manipulation. For competition experiments used to quantify relative fitness, we used a JCVI-syn1.0 strain that expresses mCherry, which allowed us to distinguish it in mixed culture from other strains via flow cytometry (see below).

Mutation accumulation experiment

Overview. Mutation accumulation (MA) experiments are designed to reduce the influence of natural selection through repeated bottlenecks of evolving populations (Lynch et al. 2016). When used with microbial populations, this is typically achieved by transferring single colonies, which have undergone single-cell bottlenecks. Prior to initiating MA experiments, we acclimated JCVI-syn1.0 and JCVI-syn3B to laboratory conditions by maintaining populations in SP4 liquid medium. We took a clone of each acclimated strain to begin the MA experiment. We propagated replicate MA lineages (JCVI-syn1.0: 87 replicates; JCVI-syn3B: 57 replicates) of each strain for 20 to 36 weekly transfers.

Number of generations. In order to compare rates of mutation across replicates, we normalized all rates as per-generation values. To calculate the number of generations per transfer in the MA, we

grew cells on SP4 agar for one week and diluted a sample of 7th day colonies into 1 mL of phosphate-buffered saline (pH = 7.4). Cells were fixed with 20 µL of 25% glutaraldehyde and stained with 2X SYBR Green, and then counted with a NovoCyte flow cytometer (ACEA Biosciences). We used the dilutions to calculate the number of cells in the original colony, from which we inferred the number of generations ($\log_2(N)$, where N is the number of cells in the undiluted colony) that must have occurred to reach a colony of that size (Dillon et al. 2017), assuming each colony is formed by a single progenitor cell. Because growth rate and other fitness components can decrease during an MA experiment (Behringer and Hall 2016), we also measured the number of cells per colony during and at the end of the MA, averaging across time points to estimate the total number of generations. We then used the number of generations per transfer to estimate the effective population size (N_e) via the harmonic mean method (Behringer and Hall 2016). Specifically, N_e was approximated as the harmonic mean of the series (2^0, 2^1, 2^2, …, 2^f), where f is equal to the number of generations per transfer inferred from the previous step.

Whole-genome sequencing and sequence analysis. We performed DNA extractions from evolved MA lines using a DNeasy UltraClean Microbial Kit (Qiagen) following the manufacturer's instructions, with the additional step of adding 50 µL of 50 mg/mL lysozyme to improve cell lysis. Genomic DNA was sequenced using Illumina MiSeq sequencing to a depth of ≥35X coverage. Library preparation and DNA sequencing were conducted by the Indiana University Bloomington Center for Genomics and Bioinformatics. Whole-genome sequencing reads were quality controlled using cutadapt (Martin 2011) to trim low-quality base pairs and remove residual adapter sequences. We used breseq with default parameters (Barrick et al. 2014; Deatherage and Barrick 2014) to call mutations using the trimmed reads. We only considered fixed mutations for the MA

lines. We checked for mutations that had arisen in experimental ancestor strains prior to evolution. Ancestral mutations were removed from analysis of all evolved lines derived from that strain using gdtools (Barrick et al. 2014; Deatherage and Barrick 2014). We used the sequence data to check for contamination or cross-contamination in evolved lines.

Statistical analyses. To compare mutation rate and spectrum between strains, we used two-sample *t*-tests for numerical response variables and two-sample χ^2 tests with continuity correction for comparing proportions. For comparing proportions to theoretical expectations within a strain, we used one-sample χ^2 tests with continuity correction.

Adaptive evolution

Overview. We conducted experiments that allowed bacteria to achieve large population sizes to increase the efficacy of natural selection. This involved serial passaging of cells in liquid cultures with limited bottlenecking at each transfer. For example, in our experiment, the minimum population size was 2-4 × 10^7 for both JCVI-syn1.0 and JCVI-syn3B. We passaged replicate 3 mL liquid cultures of each strain (n = 4 per strain) in 13 mm glass culture tubes *via* 1% v/v serial transfer each day for 300 days in a dark, static incubator held at 37 °C. We calculated the number of generations per day as the $\log_2$ of the dilution factor, i.e., $\log_2(101)$, the number of binary fissions needed to regenerate the original population size after the 1% v/v transfer (Lenski and Travisano 1994). Based on this, we estimated that the *M. mycoides* strains were maintained for 1,997 generations, which based on other experiments, is long enough for the majority of adaptation to occur (Lenski and Travisano 1994; Lang et al. 2011).

346 *Measurement of fitness.* We measured relative fitness by competing ancestral and evolved strains against a third-party *M. mycoides* JCVI-syn1.0 strain labeled with mCherry (Wiser and Lenski 348 2015). Each strain was grown in liquid medium to log phase, and then the labeled and unlabeled strains were simultaneously diluted into a mixed culture in fresh medium. The abundances of the 350 labeled and unlabeled strains in the mixed culture were calculated by immediately (t_0) sampling the culture, fixing the cells with 20 µL of cold 25% glutaraldehyde incubated at 4 °C for 20 min, 352 and staining with 2X SYBR Green. After 24 hr of growth (t_f), the mixed culture was sampled again in the same manner. We quantified the abundance of each strain using either our Novocyte flow 354 cytometer (ACEA Biosciences) or an LSR II flow cytometer (BD Biosciences) at Indiana University's Flow Cytometry Core Facility. The abundance of the strain of interest was calculated from 356 the observed number of SYBR Green-only positive cells while the abundance of the third-party strain was calculated from the number of cells positive for both SYBR Green and mCherry. From 358 these cell counts, we calculated relative fitness as the change in the relative abundance of the strain of interest during the 24 hr period of competitive growth versus the third-party strain. Specifically, 360 the relative fitness versus the mCherry third-party strain W_C is

$$W_C = \frac{\ln(\frac{N_f}{N_0})}{\ln(\frac{N_{Cf}}{N_{C0}})}$$

362 where N_0 represents the initial abundance of the focal strain, N_f the abundance of the focal strain after 24 hr, and N_{Cf} and N_{C0} are final and initial abundances of the mCherry third-party strain, 364 respectively (Wiser and Lenski 2015). We standardized fitness values to be relative to the original *Mm* JCVI-syn1.0 ancestor strain. In other words, we represent the fitness as W: $W = \frac{W_C}{W_{JCVI-syn1.0}}$, 366 where $W_{JCVI-syn1.0}$ is the value of W_C for *M. mycoides* JCVI-syn1.0. Axenic control cultures were always run contemporaneously in order account for a proportion of mCherry-positive cells that

failed to detectably express mCherry.

Whole-genome sequencing and sequence analysis. DNA extraction, sequencing, and bioinformatics followed the same methods as for the MA with a few exceptions. Specifically, each replicate population was sequenced to a depth of ≥ 100X coverage, and polymorphic mutations were included in our analyses. In addition, we calculated d_N/d_S for evolved lines using the Jukes-Cantor-corrected equation (Yang 2006) as a possible indicator of molecular constraint among the lines. We counted the number of synonymous and nonsynonymous sites using the gdtools module of breseq (Deatherage and Barrick 2014). The observed numbers of synonymous and nonsynonymous substitutions were obtained directly from breseq outputs. Synonymous and nonsynonymous polymorphisms were included in the observed count with probability proportionate to their frequency in mapped reads. We added a pseudocount of 1 synonymous substitution for all calculations (Shpak et al. 2015) because two of the lines had 0 synonymous substitutions.

To identify mutations possibly contributing to adaptation, we looked for genes which had mutations in two or more lines. Mutations in the same gene, arising and increasing in frequency in independent lineages, suggests that mutation's rise could be driven by positive selection (Johnson et al. 2021). To test this hypothesis, we statistically assessed whether multiply mutated genes had acquired more mutations than would be expected by chance under the assumption that the mutations were neutral (Johnson et al. 2021). To do this, we recorded all of the polymorphic and fixed mutations that were called within genes. Synonymous mutations were excluded. We then simulated placing these mutations at random across all genes. The probability of any given gene receiving any given mutation was proportional to the gene's length and relativized to the gene's GC

content using the known mutation rates of G:C nucleotides and A:T nucleotides from the mutation accumulation experiment. We repeated this random placement of mutations 100,000 times. In each simulation, we counted the number of mutations received by each gene, with each fixed mutation increasing the count by 1 and each polymorphism counted at a value equal to its frequency. For each multiply-mutated gene from the real adaptation experiment, we measured the proportion of the 100,000 simulations in which the gene received at least as many mutations as were truly observed and called this proportion the *P* value. To correct for the multiple tests (Non-minimal cell: 32; Minimal: 22), we report a Benjamini-Hochberg corrected value P_{adj} (Benjamini and Hochberg 1995; Benjamini et al. 2009) for which the false discovery rate $\alpha = 0.05$ for the reported significant genes (Table 1).

Generation of ftsZ *E315* mutant cells.* This process required mutating the bacterial genomes while they were yeast centromeric plasmids (YCPs) followed by genome transplantation of the mutated genomes. The YCPs were mutated using rounds of CRISPR/Cas9 and yeast homologous recombination that is a modification of a method we used previously to mutate *M. mycoides* strains (Kannan et al. 2016).

In the first CRISPR/Cas9 step, the molecule to be mutated was cleaved and the donor DNA comprising sequences from the two flanking genes was recombined with the cut JCVI-syn1.0 or JCVI-syn3B YCP, removing parts of genes of the flanking genes and all of the target gene. The donor DNA had 40 basepair overlaps to both genes flanking the target gene and had a 22 bp *Mycoplasma gallisepticum* 161 CRISPR/Cas9 target sequence with a protospacer adjacent motif (PAM) (5'-GTATAAATACATCCAGGAGtgg-3') that had no homology elsewhere in JCVI-

414 syn1.0 or JCVI-syn3B. The *M. gallisepticum* sequence put a new PAM in the genome, employed in the second round of CRISPR/Cas9.

416

The second round of CRISPR/CAS9 cut the JCVI-syn1.0 or JCVI-syn3B YCP at the new 418 *M. gallisepticum* PAM. The cut YCP was then re-circularized using a donor DNA containing the desired point mutation. The mutagenized regions of the YCPs were PCR amplified and the muta- 420 tion was confirmed by Sanger sequencing. Correctly mutagenized JCVI-syn1.0 or JCVI-syn3B YCPs were then transplanted into *Mycoplasma capricolum* recipient cells as reported previously 422 (Lartigue et al. 2007, 2009; Hutchison et al. 2016). The mutagenized regions of the transplants were PCR amplified and sequenced to confirm the presence of the desired mutations.

424

Microscopy and image analysis. To compare changes in cell size of evolved populations, we cen- 426 trifuged stationary-phase cultures and resuspended the pellet in 1 mL of phosphate-buffered saline (pH = 7.4). The resuspended cells were fixed by adding 20 µL of cold 25% glutaraldehyde and 428 incubating at 4 °C for 20 min. For microscopic observation, fixed cells were concentrated 50X via centrifugation and resuspension. The centrifugation steps were performed at room temperature for 430 4 min at 2000 *g*.

432 For phase contrast microscopy, we pipetted 5 µL of concentrated culture onto an agarose pad (1% agarose in phosphate-buffered saline) and observed cells in the Plan Apo 100X Ph 3 434 objective lens of a Nikon Eclipse 80i phase contrast microscope. Image data were analyzed using MicrobeJ (Ducret et al. 2016). Cell outlines were drawn using default parameters. Area (number 436 of pixels) within the outline was use as a metric of the area of one face of the spherical cells. Pixel

values were converted to μm^2 using the conversion factor 1 pixel length = 0.0645 μm. Cell diameter was inferred from this area using the assumption of spherical geometry.

Scanning electron microscopy (SEM) was performed at the Indiana University Bloomington Electron Microscopy Center. Fixed cells in PBS were pelleted and resuspended in 100 mM sodium cacodylate buffer (pH = 7.2) with 2 mM calcium chloride and 2% sucrose. We coated 12 mm diameter glass coverslips with 0.1% poly-L lysine for 5 min, after which coverslips were washed with a few drops of double distilled water. Resuspended cells were added to the coverslip surface and allowed to adhere. After 5 min the coverslips were washed twice with 100 mM sodium cacodylate buffer (pH = 7.2) with 2 mM calcium chloride and 2% sucrose. Next, 300 μL of 2% osmium tetroxide in 100 mM sodium cacodylate buffer (pH = 7.2) with 2% sucrose was added to the surface of the coverslips while on ice. After 30 minutes the coverslips were washed with double distilled water. Coverslips were placed in a CPD coverslip holder (Electron Microscopy Sciences, catalog number 70193-01). Samples were dehydrated in a graded ethanol series (30%, 50%, 70%, 90%, 95%) while on ice. At room temperature, coverslips were rinsed three times with 100% ethanol. Each dehydration step lasted 2 min. Critical point drying was done using a Tousimis Samdri 790 critical-point dryer. The dried coverslips were placed on aluminum SEM stubs and sputter coated, using a Safematic CCU-010 with SP-010 Sputter Head, with 45 nm of gold/palladium (80%/20%). We viewed the samples with a FEI Teneo scanning electron microscope at 2.0 kV, 25 pA probe current, and 3.0 mm working distance. The T2 detector was used. We analyzed the SEM image data using ImageJ (Schneider et al. 2012). We used the “straight” and “measure” features combined with image scale metadata to measure the vertical diameters of imaged cells that met the following criteria: cells must be round; cells must not have apparent holes or punctures; cells

must be completely within the field of view; cells must have an unambiguous perimeter; there must be no suggestion that a cell is currently or has recently undergone binary fission; cells must ≥ 0.1 μm across. Each image was processed counterclockwise starting from east. The samples were processed in a randomized order.

Statistical analyses. We compared relative fitness, d_N/d_S, and cell size between evolved populations using two-sample *t*-tests, and to ancestral values using one-sample *t*-tests. We also used two-sample *t*-tests to compare the relative fitness and cell size of *ftsZ* E315* mutants to the respective ancestral values.

We compared the composition of genes acquiring mutations among the evolved replicate lines by first constructing a gene-by-population matrix. Herein, each row represented an evolved line and each column represented a gene that had acquired at least one mutation among all of the lines. Each cell of the matrix was filled with the sum value of mutations occurring in that gene in that line, where fixed mutations were valued at 1 and polymorphisms were valued equal to the allele frequency. Only essential genes, shared between JCVI-syn1.0 and JCVI-syn3B, were considered. We used a permutational multivariate analysis of variance (PERMANOVA) on the Bray-Curtis distances generated from the gene-by-population matrix to test for the significance of cell type (minimal vs. non-minimal) on the composition of mutations using the adonis function in the R package vegan (Oksanen et al. 2020). For visualization, the Bray-Curtis distances were decomposed into two dimensions using principal coordinate analysis via the cmdscale function.

ACKNOWLEDGEMENTS

We acknowledge financial support from the US National Science Foundation (DEB-1442246 and 1934554 JTL; MCB-1818344 and 1840320 JIG, LS, and KSW), US Army Research Office Grant (W911NF-14-1-0411 JTL and ML), the National Aeronautics and Space Administration (80NSSC20K0618 JTL), the National Institutes of Health (R35-GM122566-01 and 2017-202 ML), and the Brazilian Agricultural Research Corporation (DMCB). We also acknowledge B. Stein for assistance with scanning electron microscopy (NIH 1S10OD023501-01); J. French, E. Snider, K. McKenzie, B. Lehmkuhl for assistance in the laboratory; and M. Behringer for discussion regarding sequence analyses. Data and code are available at https://github.com/LennonLab/MinimalCell.

LITERATURE CITED

Amir, A. 2017. Is cell size a spandrel? eLife 6:e22186.

Antczak, M., M. Michaelis, and M. N. Wass. 2019. Environmental conditions shape the nature of a minimal bacterial genome. Nat. Commun. 10:3100.

Barrick, J. E., G. Colburn, D. E. Deatherage, C. C. Traverse, M. D. Strand, J. J. Borges, D. B. Knoester, A. Reba, and A. G. Meyer. 2014. Identifying structural variation in haploid microbial genomes from short-read resequencing data using breseq. BMC Genomics 15:1039.

Behringer, M. G., and D. W. Hall. 2016. Genome-wide estimates of mutation rates and spectrum in *Schizosaccharomyces pombe* indicate CpG sites are highly mutagenic despite the absence of DNA methylation. G3 Bethesda Md 6:149–160.

Benjamini, Y., R. Heller, and D. Yekutieli. 2009. Selective inference in complex research. Philos. Trans. R. Soc. Math. Phys. Eng. Sci. 367:4255–4271.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. 57:289–300.

Benner, S. A., and A. M. Sismour. 2005. Synthetic biology. Nat. Rev. Genet. 6:533–543.

Breuer, M., T. M. Earnest, C. Merryman, K. S. Wise, L. Sun, M. R. Lynott, C. A. Hutchison, H. O. Smith, J. D. Lapek, D. J. Gonzalez, V. de Crécy-Lagard, D. Haas, A. D. Hanson, P. Labhsetwar, J. I. Glass, and Z. Luthey-Schulten. 2019. Essential metabolism for a minimal cell. eLife 8:e36842.

Cameron, D. E., C. J. Bashor, and J. J. Collins. 2014. A brief history of synthetic biology. Nat. Rev. Microbiol. 12:381–390.

Carvalho, F. M., M. M. Fonseca, S. Batistuzzo De Medeiros, K. C. Scortecci, C. A. G. Blaha, and L. F. Agnez-Lima. 2005. DNA repair in reduced genome: the Mycoplasma model. Gene 360:111–119.

Chien, A.-C., N. S. Hill, and P. A. Levin. 2012. Cell size control in bacteria. Curr. Biol. 22:R340–R349.

Cohan, M. C., A. M. P. Eddelbuettel, P. A. Levin, and R. V. Pappu. 2020. Dissecting the functional contributions of the intrinsically disordered C-terminal tail of *Bacillus subtilis* FtsZ. J. Mol. Biol. 432:3205–3221.

Deatherage, D. E., and J. E. Barrick. 2014. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. Methods Mol. Biol. Clifton NJ 1151:165–188.

DeLisi, C. 2019. The role of synthetic biology in climate change mitigation. Biol. Direct 14:14.

DeLong, J. P., J. G. Okie, M. E. Moses, R. M. Sibly, and J. H. Brown. 2010. Shifts in metabolic scaling, production, and efficiency across major evolutionary transitions of life. Proc. Natl. Acad. Sci. 107:12941–12945.

Dillon, M. M., W. Sung, R. Sebra, M. Lynch, and V. S. Cooper. 2017. Genome-wide biases in the rate and molecular spectrum of spontaneous mutations in *Vibrio cholerae* and *Vibrio fischeri*. Mol. Biol. Evol. 34:93–109.

Ducret, A., E. M. Quardokus, and Y. V. Brun. 2016. MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. Nat. Microbiol. 1:16077.

Gibson, D. G., J. I. Glass, C. Lartigue, V. N. Noskov, R.-Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z.-Q. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison, H. O. Smith, and J. C. Venter. 2010. Creation of a bacterial cell controlled by a chemically synthesized genome. Science 329:52–56.

Gifford, D. R., S. E. Schoustra, and R. Kassen. 2011. The length of adaptive walks is insensitive to starting fitness in *Aspergillus nidulans*. Evol. Int. J. Org. Evol. 65:3070–3078.

Glass, J. I., C. Merryman, K. S. Wise, C. A. Hutchison, and H. O. Smith. 2017. Minimal cells—real and imagined. Cold Spring Harb. Perspect. Biol. 9:a023861.

Goldford, J. E., N. Lu, D. Bajić, S. Estrela, M. Tikhonov, A. Sanchez-Gorostiaga, D. Segrè, P. Mehta, and A. Sanchez. 2018. Emergent simplicity in microbial community assembly. Science 361:469–474.

Gounand, I., T. Daufresne, D. Gravel, C. Bouvier, T. Bouvier, M. Combe, C. Gougat-Barbera, F. Poly, C. Torres-Barceló, and N. Mouquet. 2016. Size evolution in microorganisms masks trade-offs predicted by the growth rate hypothesis. Proc. R. Soc. B Biol. Sci. 283:20162272.

542 Graffeuil, A., B. E. Uhlin, and D. A. Cisneros. 2020. "Polar mutagenesis of bacterial transcriptional units using Cas12a." Microbiology.

544 Graur, D., and W.-H. Li. 2000. Fundamentals of molecular evolution. 2nd ed. Sinauer Associates, Sunderland, Mass.

546 Hahn, M. W., and A. D. Kern. 2005. Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks. Mol. Biol. Evol. 22:803–806.

548 Hershberg, R. 2015. Mutation—The engine of evolution: studying mutation and its role in the evolution of bacteria. Cold Spring Harb. Perspect. Biol. 7:a018077.

550 Hill, N. S., P. J. Buske, Y. Shi, and P. A. Levin. 2013. A moonlighting enzyme links *Escherichia coli* cell size with central metabolism. PLoS Genet. 9:e1003663.

552 Hillesland, K. L., and D. A. Stahl. 2010. Rapid evolution of stability and productivity at the origin of a microbial mutualism. Proc. Natl. Acad. Sci. 107:2124–2129.

554 Hutchison, C. A., R.-Y. Chuang, V. N. Noskov, N. Assad-Garcia, T. J. Deerinck, M. H. Ellisman, J. Gill, K. Kannan, B. J. Karas, L. Ma, J. F. Pelletier, Z.-Q. Qi, R. A. Richter, E. A. Strychalski, L. Sun, Y.
556 Suzuki, B. Tsvetanova, K. S. Wise, H. O. Smith, J. I. Glass, C. Merryman, D. G. Gibson, and J. C. Venter. 2016. Design and synthesis of a minimal bacterial genome. Science 351:aad6253.

558 Johnson, M. S., S. Gopalakrishnan, J. Goyal, M. E. Dillingham, C. W. Bakerlee, P. T. Humphrey, T. Jagdish, E. R. Jerison, K. Kosheleva, K. R. Lawrence, J. Min, A. Moulana, A. M. Phillips, J. C. Piper,
560 R. Purkanti, A. Rego-Costa, M. J. McDonald, A. N. Nguyen Ba, and M. M. Desai. 2021. Phenotypic and molecular evolution across 10,000 generations in laboratory budding yeast populations.
562 eLife 10:e63910.

Kacar, B., X. Ge, S. Sanyal, and E. A. Gaucher. 2017. Experimental evolution of *Escherichia coli* harboring
564 an ancient translation protein. J. Mol. Evol. 84:69–84.

Kannan, K., B. Tsvetanova, R.-Y. Chuang, V. N. Noskov, N. Assad-Garcia, L. Ma, C. A. Hutchison III, H.
566 O. Smith, J. I. Glass, C. Merryman, J. C. Venter, and D. G. Gibson. 2016. One step engineering of the small-subunit ribosomal RNA using CRISPR/Cas9. Sci. Rep. 6.

568 Kuo, C.-H., and H. Ochman. 2009. Deletional bias across the three domains of life. Genome Biol. Evol. 1:145–152.

570 Lachance, J.-C., S. Rodrigue, and B. O. Palsson. 2019. Minimal cells, maximal knowledge. eLife 8:e45379.

Lang, Gregory. I., D. Botstein, and M. M. Desai. 2011. Genetic variation and the fate of beneficial mutations

572 in asexual populations. Genetics 188:647–661.

Lartigue, C., J. I. Glass, N. Alperovich, R. Pieper, P. P. Parmar, C. A. Hutchison, H. O. Smith, and J. C.

574 Venter. 2007. Genome transplantation in bacteria: changing one species to another. Science 317:632–638.

576 Lartigue, C., S. Vashee, M. A. Algire, R.-Y. Chuang, G. A. Benders, L. Ma, V. N. Noskov, E. A. Denisova, D. G. Gibson, N. Assad-Garcia, N. Alperovich, D. W. Thomas, C. Merryman, C. A. Hutchison, H.

578 O. Smith, J. C. Venter, and J. I. Glass. 2009. Creating bacterial strains from genomes that have been cloned and engineered in yeast. Science 325:1693–1696.

580 Lenski, R. E., and M. Travisano. 1994. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. Proc. Natl. Acad. Sci. U. S. A. 91:6808–6814.

582 Leprince, A., M. W. van Passel, and V. A. M. dos Santos. 2012. Streamlining genomes: toward the generation of simplified and stabilized microbial systems. Curr. Opin. Biotechnol. 23:651–658.

584 Liao, Y., S. Ithurbide, C. Evenhuis, J. Löwe, and I. G. Duggin. 2021. Cell division in the archaeon *Haloferax volcanii* relies on two FtsZ proteins with distinct functions in division ring assembly and con-

586 striction. Nat. Microbiol. 6:594–605.

Lindahl, T. 1974. An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing

588 deaminated cytosine residues. Proc. Natl. Acad. Sci. U. S. A. 71:3649–3653.

Lindahl, T., S. Ljungquist, W. Siegert, B. Nyberg, and B. Sperens. 1977. DNA N-glycosidases: properties

590 of uracil-DNA glycosidase from *Escherichia coli*. J. Biol. Chem. 252:3286–3294.

Long, H., W. Sung, S. Kucukyildirim, E. Williams, S. F. Miller, W. Guo, C. Patterson, C. Gregory, C.

592 Strauss, C. Stone, C. Berne, D. Kysela, W. R. Shoemaker, M. E. Muscarella, H. Luo, J. T. Lennon,

Y. V. Brun, and M. Lynch. 2018. Evolutionary determinants of genome-wide nucleotide composition. Nat. Ecol. Evol. 2:237–240.

Lynch, M. 2007. The origins of genome architecture. Sinauer Associates, Sunderland, Mass.

Lynch, M., M. S. Ackerman, J.-F. Gout, H. Long, W. Sung, W. K. Thomas, and P. L. Foster. 2016. Genetic drift, selection and the evolution of the mutation rate. Nat. Rev. Genet. 17:704–714.

Márquez-Zacarías, P., P. L. Conlin, K. Tong, J. T. Pentz, and W. C. Ratcliff. 2021. Why have aggregative multicellular organisms stayed simple? Curr. Genet., doi: 10.1007/s00294-021-01193-0.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10.

McQuillen, R., and J. Xiao. 2020. Insights into the structure, function, and dynamics of the bacterial cytokinetic FtsZ-ring. Annu. Rev. Biophys. 49:309–341.

Mei, Z.-P., Z. V. Finkel, and A. J. Irwin. 2009. Light and nutrient availability affect the size-scaling of growth in phytoplankton. J. Theor. Biol. 259:582–588.

Mongold, J. A., and R. E. Lenski. 1996. Experimental rejection of a nonadaptive explanation for increased cell size in *Escherichia coli*. J. Bacteriol. 178:5333–5334.

Moran, N. A. 2003. Tracing the evolution of gene loss in obligate bacterial symbionts. Curr. Opin. Microbiol. 6:512–518.

Moran, N. A., and G. M. Bennett. 2014. The tiniest tiny genomes. Annu. Rev. Microbiol. 68:195–215.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, and H. Wagner. 2020. R package 'vegan': community ecology package.

Pelletier, J. F., L. Sun, K. S. Wise, N. Assad-Garcia, B. J. Karas, T. J. Deerinck, M. H. Ellisman, A. Mershin, N. Gershenfeld, R.-Y. Chuang, J. I. Glass, and E. A. Strychalski. 2021. Genetic requirements for cell division in a genomically minimal cell. Cell S0092867421002932.

Richardson, S. M., L. A. Mitchell, G. Stracquadanio, K. Yang, J. S. Dymond, J. E. DiCarlo, D. Lee, C. L. V. Huang, S. Chandrasegaran, Y. Cai, J. D. Boeke, and J. S. Bader. 2017. Design of a synthetic yeast genome. Science 355:1040–1044.

Schneider, C. A., W. S. Rasband, and K. W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9:671–675.

Sedwick, W. D., O. E. Brown, and B. W. Glickman. 1986. Deoxyuridine misincorporation causes site-specific mutational lesions in the lacI gene of *Escherichia coli*. Mutat. Res. 162:7–20.

Shpak, M., M. M. Goldberg, and M. C. Cowperthwaite. 2015. Rapid and convergent evolution in the Glioblastoma multiforme genome. Genomics 105:159–167.

Sung, W., M. S. Ackerman, M. M. Dillon, T. G. Platt, C. Fuqua, V. S. Cooper, and M. Lynch. 2016. Evolution of the insertion-deletion mutation rate across the tree of life. G3 Genes Genomes Genet. 6:2583–2591.

Sung, W., M. S. Ackerman, S. F. Miller, T. G. Doak, and M. Lynch. 2012. Drift-barrier hypothesis and mutation-rate evolution. Proc. Natl. Acad. Sci. 109:18488–18492.

Tye, B.-K., and I. R. Lehman. 1977. Excision repair of uracil incorporated in DNA as a result of a defect in dUTPase. J. Mol. Biol. 117:293–306.

Vasi, F., M. Travisano, and R. E. Lenski. 1994. Long-term experimental evolution in *Escherichia coli*. II. Changes in life-history traits during adaptation to a seasonal environment. Am. Nat. 144:432–456.

Westfall, C. S., and P. A. Levin. 2017. Bacterial cell size: multifactorial and multifaceted. Annu. Rev. Microbiol. 71:499–517.

Wiser, M. J., and R. E. Lenski. 2015. A comparison of methods to measure fitness in *Escherichia coli*. PLoS ONE 10:e0126210.

Yang, Z. 2006. Computational molecular evolution. Oxford University Press, Oxford.

Table 1. Mutations in genes that were identified as being under positive selection during 2000 generations of experimental evolution. Statistical simulations were performed to find genes acquiring more mutations than expected to occur by chance, indicative of positive selection increasing the allelic proportion of such mutations. P_{adj} corresponds to significance following Benjamin-Hochberg correction to account for multiple comparisons. Genes are assigned to categories based on the secondary functional classifications (Breuer et al. 2019). Note that "Central metabolism" corresponds to "Central carbon metabolism" in the original source (Breuer et al. 2019). While most nonessential genes are absent from the minimal cell (*), a few nonessential genes were retained in *M. mycoides* JCVI-syn3B to facilitate cultivation and robust growth (†). Uncharacterized genes did not fall into defined category (N/A). For mutations that were found in both strains, "non" refers to non-minimal and "min" refers to minimal.

Non-minimal cell only

Gene	Annotation	Category	P_{adj}
lpdA	Dihydrolipoyl dehydrogenase	Central metabolism	<0.0001
pyk	Pyruvate kinase	Central metabolism	0.016
dnaA_1	Chromosomal replication initiator protein	DNA maintenance	0.0005
ecfT	ECF transporter T component	Membrane transport	<0.0001
Locus 0339*	Na^+ ABC transporter, ATP-binding component	Membrane transport	0.016
Locus 0030	Uncharacterized ABC transporter ATP-binding	Membrane transport	0.016
rpoA	DNA-directed RNA polymerase, alpha subunit	Transcription	0.0090
Locus 0187*	Transcriptional regulator, GntR family	Transcription	0.0098
tnpA_1*	IS1296 transposase protein A	DNA maintenance	0.0015
tnpB_1*	IS1296 transposase protein B	DNA maintenance	0.0023
Locus 0892*	Putative membrane protein	N/A	0.016

Locus 0898*	Conserved hypothetical protein	N/A	0.029
Locus 0471*	Hypothetical protein	N/A	0.0003
Locus 0751*	Putative viral A-type inclusion protein	N/A	0.016

Minimal cell only

Gene	Annotation	Category	P_{adj}
atpD	F0F1 ATP synthase subunit beta	Central metabolism	<0.0001
pyrG	CTP synthase	Central metabolism	0.0002
dnaN	DNA polymerase III subunit beta	DNA maintenance	<0.0001
lgt	Diacylglyceryl transferase	Biosynthesis	0.0038
clsA	Cardiolipin synthase	Biosynthesis	0.015
fakA	Fatty acid kinase subunit A	Biosynthesis	0.0025
amiF	Oligopeptide ABC transporter ATP-binding	Membrane transport	0.0036
rpoC	DNA-directed RNA polymerase, beta subunit	Transcription	0.0024
rpsC	30S ribosomal protein S3	Translation	0.021
Locus 0373	Uncharacterized protein	N/A	0.015
Locus 0691	Uncharacterized protein	N/A	0.027
Locus 0430	Uncharacterized DNA-binding protein	N/A	0.0012

Minimal and non-minimal cell

Gene	Annotation	Category	P_{adj}
ftsZ †	Cell division protein ftsZ	Cell growth	non <0.0001 min <0.0001
tetM	Tetracycline resistance ribosomal protection	Drug resistance	non = 0.029 min = 0.0079

FIGURE CAPTIONS

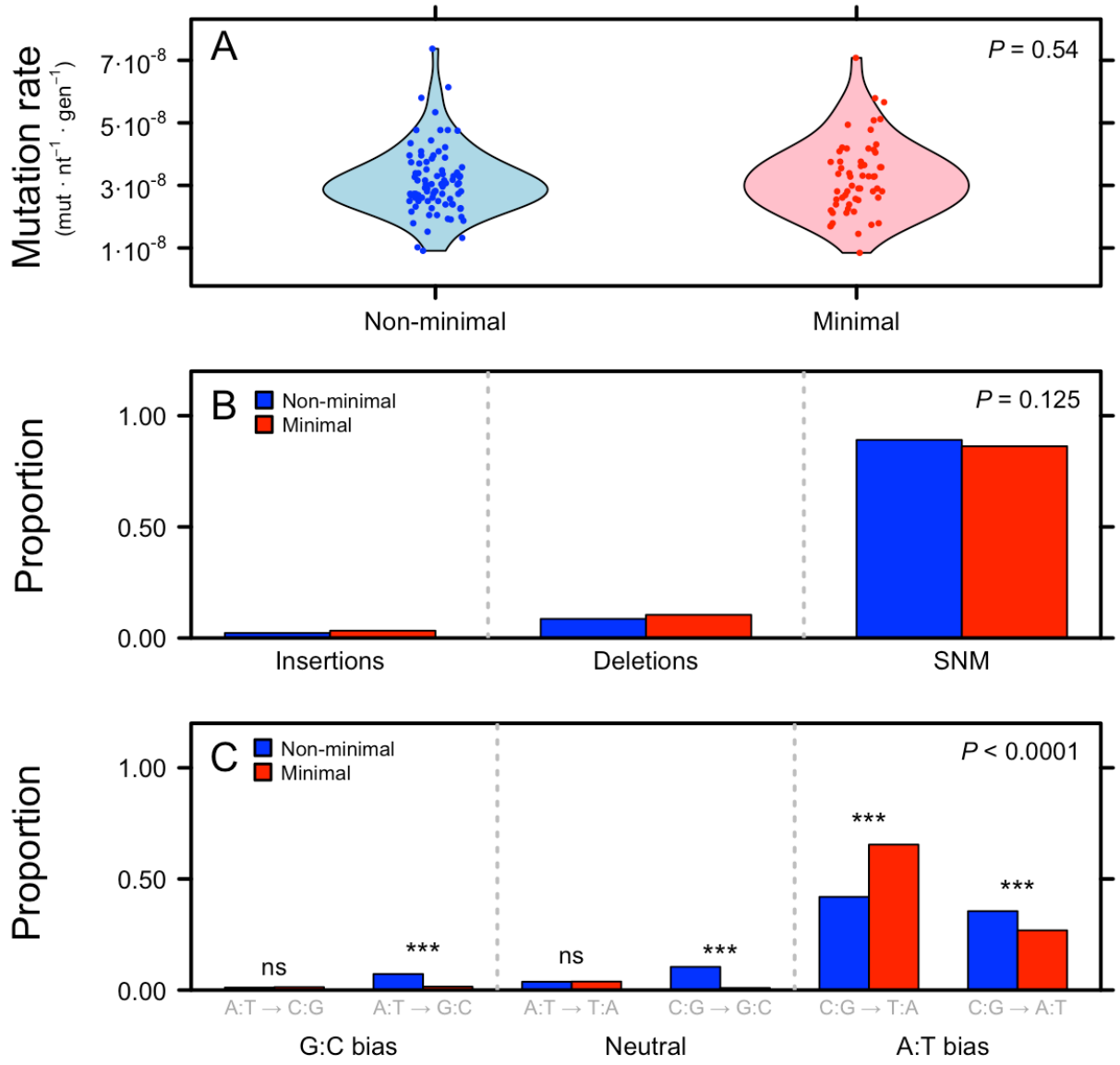
Fig. 1. Mutation rate and spectrum of the minimal and non-minimal cell estimated from mutation accumulation experiments. (**A**) Although *M. mycoides* has the highest recorded mutation rate (base substitutions and indels), it was not affected by genome minimization. (**B)** The proportions of insertions, deletions, and single nucleotide mutations (SNMs) were also the same for the minimal and non-minimal cell. (**C**) Among SNMs, which accounted for 88% of all mutations, the minimal cell exhibited a stronger AT bias in its mutation spectrum than the non-minimal cell, particularly in the C:G-to-A:T category. In this panel, ns = not significant and *** = $P < 0.001$.

Fig. 2. Genome minimization resulted in a 50% reduction in relative fitness, but 80% of this cost was regained over 2000 generations of evolution. Despite removal of nearly half of its genome, the minimal cell adapted at the same rate as the non-minimal cell. Data represent mean ± SEM. Because the experiment was initiated with a single clone, error bars for the ancestral time point were calculated from experimental replicates, while error bars for evolved lines were calculated from replicate populations.

Fig. 3. The non-minimal cell and minimal cell acquired adaptive mutations in different sets of shared essential genes ($P = 0.029$) over 2000 generations of evolution (Table 1) as depicted in ordination from a principal coordinates analysis (PCoA) created with a population x mutation matrix using Bray-Curtis distance metric. Dashed lines represent 95% confidence ellipses around replicate populations (solid symbols).

Fig. 4. Effect of genome minimization on the evolution of cell size. (**A**) The size of non-minimal cells increased by 73% over 2000 generations of evolution, while the size of minimal cell remained the same. Data represent means ± SEM based on images obtained via phase contrast microscopy. Scanning electron micrographs obtained from evolved replicate populations of the (**B**) non-minimal and (**C**) minimal cell.

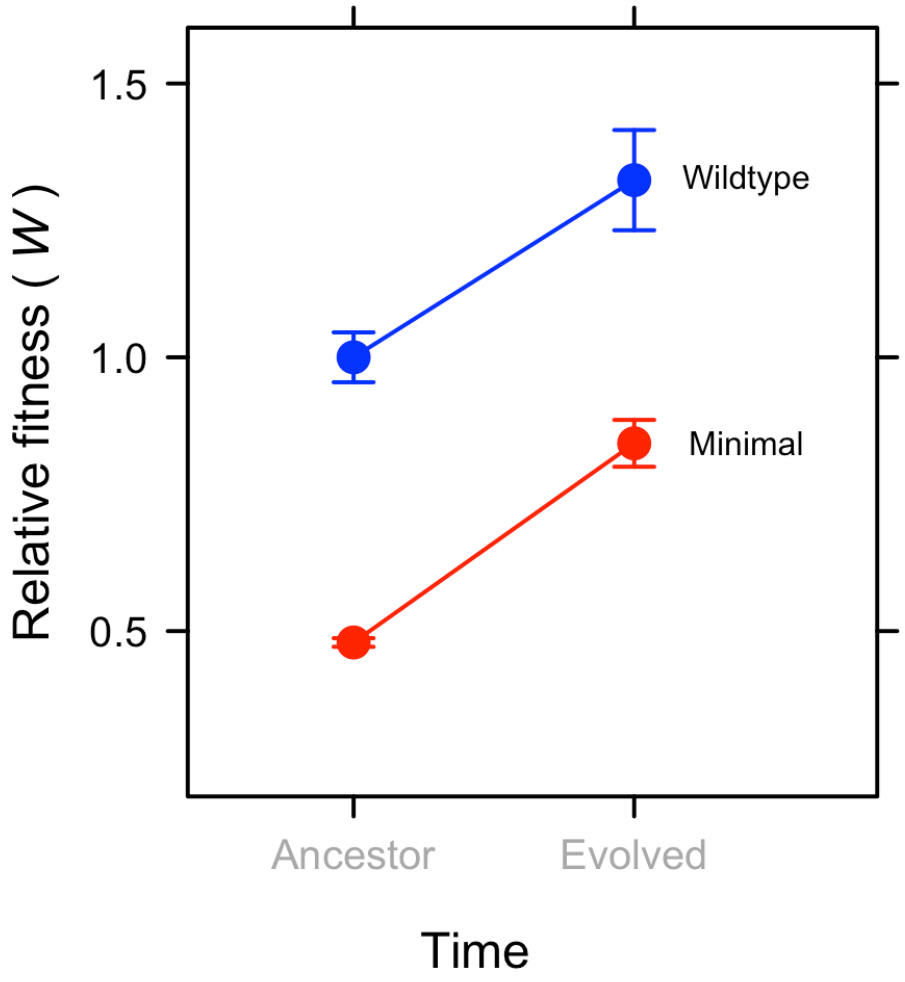
682 **Fig. 1**



684 Fig. 1. Mutation rate and spectrum of the minimal and non-minimal cell estimated from mutation accumulation experiments. (**A**) Although *M. mycoides* has the highest recorded mutation rate (base substitutions and indels), it was not
686 affected by genome minimization. (**B**) The proportions of insertions, deletions, and single nucleotide mutations (SNMs) were also the same for the minimal and non-minimal cell. (**C**) Among SMNs, which accounted for 88% of all
688 mutations, the minimal cell exhibited a stronger AT bias in its mutation spectrum than the non-minimal cell, particularly in the C:G-to-A:T category. In this panel, ns = not significant and *** = $P < 0.001$.

690

Fig. 2



692

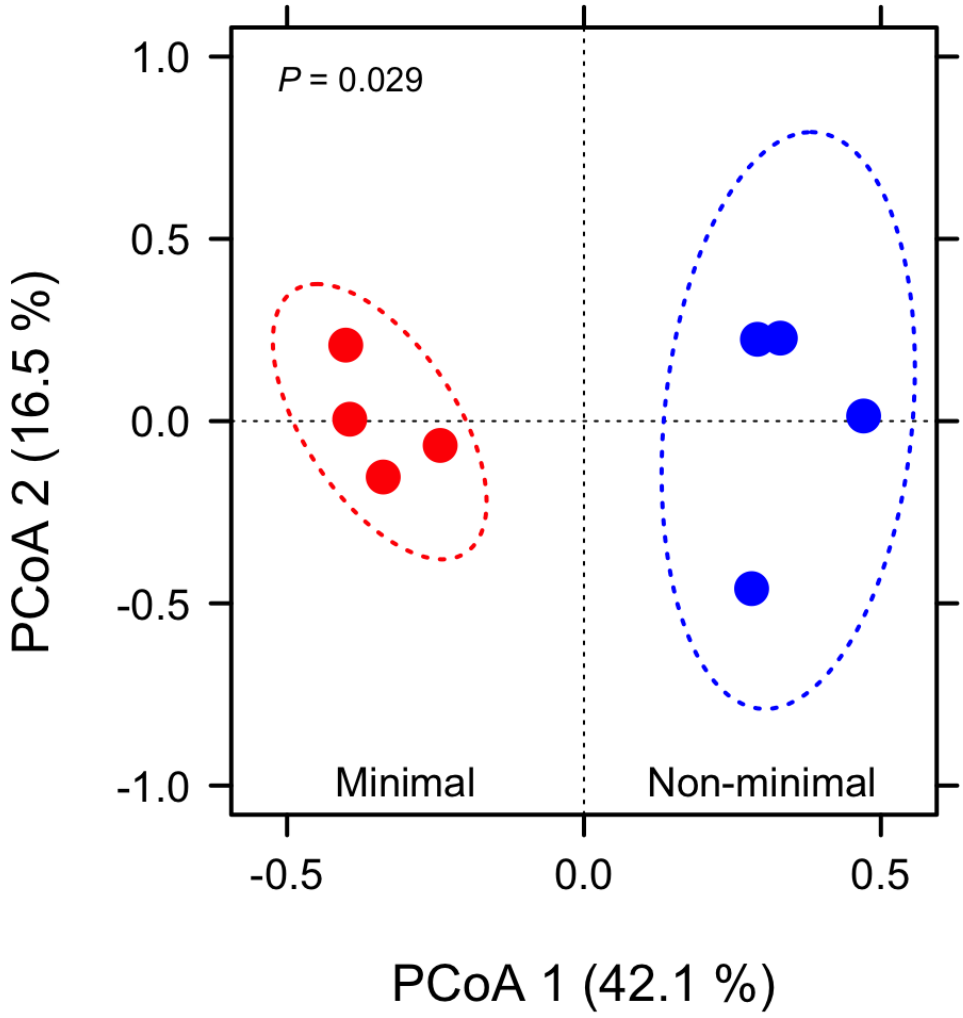
694 Fig. 2. Genome minimization resulted in a 50% reduction in relative fitness, but 80% of this cost was regained over
2000 generations of evolution. Despite removal of nearly half of its genome, the minimal cell adapted at the same
696 rate as the non-minimal cell. Data represent mean ± SEM. Because the experiment was initiated with a single clone,
error bars for the ancestral time point were calculated from experimental replicates, while error bars for evolved lines
698 were calculated from replicate populations.

700

702

704

706 **Fig. 3**



708

710 **Fig. 3.** The non-minimal cell and minimal cell acquired adaptive mutations in different sets of shared essential genes (*P* = 0.029) over 2000 generations of evolution (Table 1) as depicted in ordination from a principal coordinates analy-
712 sis (PCoA) created with a population x mutation matrix using Bray-Curtis distance metric. Dashed lines represent 95% confidence ellipses around replicate populations (solid symbols).

714

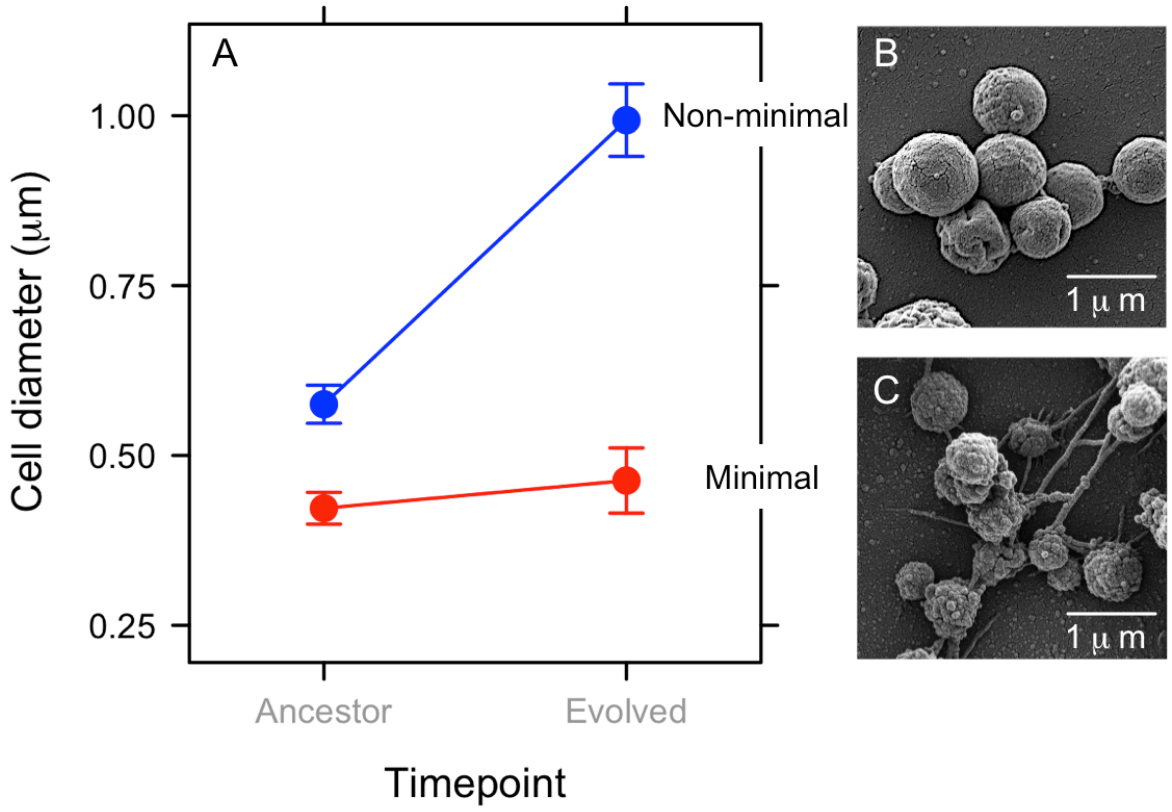
716

718

720

722 **Fig. 4**

724



726

728 **Fig. 4.** Effect of genome minimization on the evolution of cell size. (**A**) The size of non-minimal cells increased by 73% over 2000 generations of evolution, while the size of minimal cell remained the same. Data represent means ± SEM based on images obtained via phase contrast microscopy. Scanning electron micrographs obtained from 730 evolved replicate populations of the (**B**) non-minimal and (**C**) minimal cell.