

1 **Experimental evolution of Escherichia coli**

2 **harboring an ancient translation protein**

3 Betül Kaçar^{1,2, *}, Xueliang Ge³, Suparna Sanyal³ and Eric A. Gaucher^{4, 5}

4

5 1. NASA Astrobiology Institute, Mountain View, CA 94035

6 2. Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street,
7 Cambridge, MA 02138

8 3. Department of Cell and Molecular Biology, Uppsala University, Box-596, BMC,
9 75124, Uppsala, Sweden

10 4. School of Biology, Georgia Institute of Technology, 950 Atlantic Drive, Atlanta,
11 GA 30332

12 5. Petit H. Parker Institute for Bioengineering and Bioscience, Georgia Institute of
13 Technology, Atlanta, GA 30332

14

15

16 * Corresponding author

17 E-mail: kacar@fas.harvard.edu

18

19

20

21

22

23

24

1 Abstract

2 The ability to design synthetic genes and engineer biological systems at the genome
3 scale opens new means by which to characterize phenotypic states and the responses
4 of biological systems to perturbations. One emerging method involves inserting artificial
5 genes into bacterial genomes, and examining how the genome and its new genes adapt
6 to each other. Here we report the development and implementation of a modified
7 approach to this method, in which phylogenetically inferred genes are inserted into a
8 microbial genome, and laboratory evolution is then used to examine the adaptive
9 potential of the resulting hybrid genome. Specifically, we engineered an approximately
10 700-million-year old inferred ancestral variant of *tufB*, an essential gene encoding
11 Elongation Factor Tu, and inserted it in a modern *Escherichia coli* genome in place of
12 the native *tufB* gene. While the ancient homolog was not lethal to the cell, it did cause
13 a two-fold decrease in organismal fitness, mainly due to reduced protein dosage. We
14 subsequently evolved replicate hybrid bacterial populations for 2,000 generations in the
15 laboratory, and examined the adaptive response via fitness assays, whole-genome
16 sequencing, proteomics and biochemical assays. Hybrid lineages exhibit a general
17 adaptive strategy in which the fitness cost of the ancient gene was ameliorated in part
18 by up-regulation of protein production. We expect that this ancient-modern
19 recombinant method may pave the way for the synthesis of organisms that exhibit
20 ancient phenotypes, and that laboratory evolution of these organisms may prove useful
21 in elucidating insights into historical adaptive processes.

22
23
24
25
26

Background

Understanding historical evolutionary pathways is crucial to understanding how life became the way it is today across millions of years of environmental and ecosystem change (Gould 1989). One of the most difficult aspects of characterizing these historical pathways is the limited amount of knowledge available about how ancient organisms behaved and changed through time. Fossils provide useful morphological and anatomical details, but only traces of information about sub-organismal level processes and states can be inferred from fossilized specimens alone (Pagel 1999). Ancestral sequence reconstruction permits phylogenetics-based sequence inferences of ancestral genes at interior nodes of a tree using likelihood or Bayesian statistics, and offer an opportunity to determine the selectively advantageous amino acid replacements responsible for changes in protein behavior associated with adaptive events for particular molecular systems (Benner 1995; Chang et al. 2002a; Huelsenbeck and Bollback 2001; Liberles 2007; Pauling and Zuckerkandl 1963; Thornton 2004; Ugalde et al. 2004). Going backwards in time at the protein level and studying the biochemical properties of reconstructed ancient proteins in the laboratory may also improve our ability to engineer proteins for specific tasks (Benner 2007; Chang et al. 2002b; Merkl and Sterner 2016; Ogawa and Shirai 2014; Pal et al. 2014). However, mathematical sequence reconstructions and *in vitro* characterization alone may not necessarily provide the salient details of why the protein evolved along a particular evolutionary pathway (Bar-Rogovsky et al. 2015; Copley 2012; Dean and Thornton 2007; Kacar 2016; Zhu et al. 2005), nor does it automatically indicate how the protein's evolution might be mapped onto the phenotypic evolution of the whole organism, making *a priori* predictions that connect inferred genotype to ancestral phenotype a challenge. Incorporating a functional perspective into the study of ancient proteins may shed light

1 onto the biochemical mechanisms of enzyme evolution, and may also provide clues to
2 the development of ancient adaptive pathways (Dean and Thornton 2007; Harms and
3 Thornton 2013; Kacar and Gaucher 2012; Kacar and Gaucher 2013; Lunzer et al. 2005;
4 Zhu et al. 2005).

5

6 Advances in synthetic biology, bioengineering and genomics now allow us to detect
7 changes in genotype and to connect these changes with various perturbations
8 triggered at multiple levels within the cellular environment such as engineering protein-
9 protein interaction networks (Isalan et al. 2008; Wang et al. 2009; Wright et al. 2013),
10 metabolic pathways (Gallagher et al. 2014; Nyerges et al. 2016; Xu et al. 2012), and
11 constructing genetic circuits (Esvelt and Wang 2013; Nandagopal and Elowitz 2011).
12 Further, current methods of evolutionary genome engineering rely upon modifying the
13 genetic content of microbes to execute desired tasks (Andersson et al. 2015; Cambray
14 et al. 2011; Dalchau et al. 2012; Feher et al. 2012; Sandoval et al. 2012; Storz et al.
15 2015), reducing the genome size to study efficacy of certain proteins and protein
16 families (Annaluru et al. 2015; Deutschbauer et al. 2014; Kolisnychenko et al. 2002; Lee
17 et al. 2005) and heterologously replacing genomic components with homologs obtained
18 from other organisms (Acevedo-Rocha et al. 2013; Agashe et al. 2013; Andersson and
19 Hughes 2009; Pena et al. 2010; Urbanczyk et al. 2012) as well as engineering
20 endogenous promoters with variants obtained across taxa (Nevoigt et al. 2006;
21 Peisajovich et al. 2010). Observing the adaptive pathways of a modified organism can
22 reveal a variety of biochemical pathways adjusting regulatory networks (Dragosits and
23 Mattanovich 2013; Kacar and Gaucher 2012; Michener et al. 2014; Quan et al. 2012),
24 yielding invaluable information on the evolution and prediction of protein function

1 (Counago et al. 2006; Harms and Thornton 2014; Johnsen and Levin 2010; Lind et al.
2 2010).

3

4 An evolutionary bioengineering approach has been proposed to characterize the
5 adaptation of an ancient protein to a modern genome on time scales of laboratory
6 observation; an adaptive laboratory evolution method we previously coined as paleo-
7 experimental evolution (Kacar and Gaucher 2012) (**Figure 1**). It remains to be seen
8 whether it is possible to elucidate and discern ancient adaptive steps from adjustments
9 taken by a modern cell to a maladapted gene. We can begin with a top-level query: if
10 integrated into a modern genome, and given a chance to interact with descendant
11 protein partners in the cell, would the observed mutations alter the ancestral
12 component's protein structure and function in a manner comparable to the historical
13 pathway inferred for the organism, or will it develop an alternative adaptive pathway? In
14 order to address this, the paleo-experimental evolution system requires an organism
15 with a short generation time and a protein under strong selective constraints in the
16 modern host but whose ancestral genotype and phenotype, if genomically integrated,
17 would cause the modern host to be less fit than a modern population hosting the
18 modern form of the protein. The modern organism hosting the ancient protein would
19 need to be viable, but maladapted. *Escherichia coli* (*E. coli*) and an essential protein
20 family of the bacterial translation machinery, Elongation Factor-Tu (EF-Tu), are ideal for
21 this type of experiment. *E. coli* is an organism that grows quickly in the laboratory, the
22 genetics of the organism are well-known, easy to manipulate, utilizes a range of energy
23 sources and can be stored frozen and later can be re-animated to test ancestral versus
24 evolved populations (Blount 2015). Elongation Factor Tu (bacteria)/elongation factor 1A
25 (archaea and eukaryota) is a GTPase family member involved in the protein translation

system (Kavaliauskas et al. 2012). EF-Tu forms a complex with GTP that in turn favors the binding of an aminoacyl-tRNA complex (Agirrezabala and Frank 2009). This ternary complex binds to mRNA-programmed ribosomes thereby delivering aminoacyl-tRNA to the ribosomal A site (Czworkowski and Moore 1996). The biochemistry of EF-Tu has been studied for over three decades giving rise to a clear understanding of the functional aspects of the protein (Negrutskaa and El'skaya 1998). Previously, *in vitro* peptide synthesis assays demonstrated that ancestral EF-Tus can participate in a translation system in which all other components necessary for translation besides EF-Tu are provided from modern *E. coli* (Zhou et al. 2012) indicating the possibility that co-evolution between EF-Tu and aa-tRNAs/ribosome/nucleotide-exchange-factors in *E. coli* since the divergence of the ancestral and modern EF-Tu forms has not prevented the ancestral EF-Tu from interacting with the modern *E. coli* translation components. We hypothesized that the diminished capacity of the ancestral EF-Tu is sufficient enough to provide a strong selective constraint for the bacteria containing the ancient gene to acquire beneficial mutations.

Our method builds upon heterologous gene replacement by engineering a synthetic ancient EF-Tu into the genome of *E. coli* bacteria. The engineered EF-Tu represents that of an ancestral γ -proteobacteria that is inferred to be approximately 700 million years old and has 21 (out of 392) amino acid differences with the modern EF-Tu protein (Gaucher et al. 2008). Our experimental system exploits the unique scenario in which *E. coli* bacteria have a paralogous copy of the EF-Tu gene *tufA*, in the form of *tufB*, that frequently recombines with the original copy (Abdulkarim and Hughes 1996). Each of the EF-Tu genes has its own specific expression machinery, and EF-Tu produced through *tufB* accounts for one-third of the cellular EF-Tu as that produced by the *tufA*

gene in bacteria (Van Delft et al. 1987; van der Meide et al. 1983; Zengel and Lindahl 1982). Through recombination mediate-engineering (recombineering), the *tufA* gene was deleted from the bacterial genome and the *tufB* copy of a laboratory strain of *E. coli* was replaced with an ancient EF-Tu variant under the control of the endogenous *tufB* promoter, followed by adaptive laboratory evolution of the ancient-modern hybrid populations in replicate lineages through daily propagation of bacterial cultures (**Figure 1**) (Bell 2016; Dragosits and Mattanovich 2013; Elena and Lenski 2003). Evolved populations were sampled for whole genome sequencing, followed by identification of the total number of genomic changes in each population relative to the founding strain and assaying the change in adaptive response through fitness assays. We further investigated whether *in vivo* analyses into the functionality of ancestral components can be used to discern effects arising from the substituted gene when screened from adaptive responses taken by the host cell to the sub-adapted genetic component. Altogether this work provides the first demonstration of an artificial ancient essential gene variant inside a bacterial genome and provides insights into the principles of using experimental evolution for exploring adaptation of artificial genes in modern organisms.

RESULTS

Replacement of modern EF-Tu with ancient EF-Tu is detrimental to *E. coli* fitness

Complete replacement of endogenous EF-Tu protein requires disruption of both native *tufA* and *tufB* genes, and insertion of the inferred ancient gene (**Supplementary Figure 1**) (Schnell et al. 2003). We first disrupted the native *tufA* gene. This intermediate *tufA*⁻ *tufB*⁺ construct displays a fitness of 0.89 ($P < 0.001$) relative to the parent clone. Subsequent replacement of *tufB* with the reconstructed ancient *tuf* gene produced a

1 further fitness decline, to 0.77 ($P < 0.001$) relative to the parent clone (**Figure 2**). This
 2 dramatic fitness detriment of complete EF-Tu replacement suggests that the ancient
 3 gene is compatible with the modern *E. coli* genome, though unfit. Co-evolution between
 4 EF-Tu and aa-tRNAs/ribosome/nucleotide-exchange-factors in *E. coli* since the ancestral
 5 state for which the ancient *tuf* gene was inferred has thus not prevented the inferred
 6 ancestral EF-Tu from interacting with the modern *E. coli* translation system in a viable
 7 manner.

8 **Experimental evolution allows bacteria to restore fitness**

9 To examine the co-adaptation between *E. coli* and the ancient EF-Tu, we conducted
 10 evolution experiments with both the ancient-modern hybrid and the *tufA*⁻ *tufB*⁺
 11 construct. Six replicate populations were generated for each of the two modified
 12 genomes by selecting identical clones that were verified to be free of any plasmid
 13 vectors that could mediate genetic exchange. The twelve populations were then
 14 evolved for 2000 generations under a daily 100-fold serial transfer regime in DM25
 15 minimal glucose medium, at 37°C and 150 rpm. Under these conditions, each
 16 population grew $\log_2(100) = 6.64$ generations per day before reaching stationary phase.
 17 The daily maximum population size for each population is approximately 2.5×10^8 cells.
 18 Fitness assays were conducted every 500 generations, in which evolved populations
 19 were competed against the ancestral clone. The ancient-modern hybrid populations
 20 displayed a mean fitness of 1.06 ($P < 0.043$) at generation 500 (**Figure 2**). This increased
 21 to 1.12 ($P < 0.005$) at generation 1000, and 1.15 ($P < 0.009$) and 1.16 ($P < 0.001$) at
 22 generations 1500 and 2000, respectively. The *tufA*⁻ *tufB*⁺ populations also exhibited
 23 fitness increases. Mean fitness relative to the ancestor is 1.097 ($P < 0.001$) at generation
 24 500, 1.16 ($P < 0.004$) at generation 1000, and 1.15 at both 1500 ($P < 0.001$) and 2000

1 (P<0.004), respectively (**Figure 2**).

2 ***Promoter level* mutations upregulate ancient EF-Tu expression and restore**
 3 **bacterial fitness**

4 To identify the genetic bases of the observed fitness increases, whole genomes of
 5 clones were periodically isolated and sequenced from all six evolved populations during
 6 the evolution experiment. Mutations generally accumulated in similar genes across all
 7 experimental construct populations (**Table 1**). However, five out of the six ancient-
 8 modern hybrid lineages, but none of the other controls, evolved mutations in the
 9 *thrT/tufB* promoter region, with four variant alleles being observed (Lee et al. 1981). The
 10 majority of these *thrT/tufB* promoter region mutations accumulated early in the
 11 experiment, and rose to high frequency, if not fixation, by 2000 generations across all
 12 five populations in which they occurred (**Figure 3**). Such cis-regulatory mutations have
 13 been shown to be a common means of adaptation (Hoekstra and Coyne 2007; Jacob
 14 and Monod 1961; Lynch and Wagner 2008). Others have observed the accumulation of
 15 mutations in orthologous genes engineered into microbes during laboratory evolution
 16 experiments. By contrast, we observed no mutations in the ancient or modern EF-Tu
 17 gene-coding region in any of the evolved lineages, suggesting that compensatory
 18 amino acid replacements may have only occurred at other sites in the genome.

19
 20 We performed whole-cell shotgun proteomic analysis on five of the evolved hybrid
 21 populations with EF-Tu promoter mutations to examine the impact of these mutations
 22 on EF-Tu protein levels. The assayed time points were those for each population at
 23 which the mutations had reached over 90% frequency in the population. For

1 comparison, we also assayed unevolved ancient-modern hybrid bacteria, the wild-type
2 parent *E. coli* strain, and an unevolved *tufA*⁻ *tufB*⁺ construct. Deletion of the *tufB* copy
3 and the subsequent insertion of the ancient reconstructed gene into *E. coli* causes EF-
4 Tu protein levels to drop by approximately ~66% relative to that observed in the wild-
5 type. The evolved hybrid populations with *tufB* promoter mutations all show significant
6 increases in EF-Tu levels (**Figure 3, Figure 4a**). We also assessed the effect of these
7 promoter mutations on EF-Tu expression level *in vitro* by examining their effect on a
8 plasmid-borne fluorescent reporter. The mutant promoters increase expression
9 between 1.5 to 20 fold (**Supplementary Figure 2**). Interestingly, promoter mutations
10 that rose to high frequency later during the experiment had lower relative effects on
11 ancient EF-Tu protein expression than those that did so earlier during laboratory
12 evolution. To test whether increased ancient EF-Tu levels would correlate with
13 increased fitness, the unevolved ancient-modern cells, as well as *E. coli* harboring a
14 single wild-type *tufB* gene was transformed with pASK plasmids expressing ancient EF-
15 Tu proteins (Materials and Methods). Overexpression of ancient EF-Tu protein in *E. coli*
16 isogenic strain decreases the doubling time from 35 min to 26 min. Similarly,
17 overexpression of ancient EF-Tu protein in ancient-modern hybrid ancestor decreases
18 the doubling time from 52 min to 29 min (**Figure 4b**). This observation is in agreement
19 with the previous studies demonstrating the correlation between the cellular
20 concentration of EF-Tu and organismal fitness (Brandis et al. 2016; Tubulekas and
21 Hughes 1993). Taken together, these results indicate that each experimental population
22 exhibited parallel patterns of response such as upregulation of EF-Tu, as well as more
23 idiosyncratic means of compensating for altered EF-Tu expression and activity.

24

1 **Replacement of the endogenous EF-Tu with the ancient counterpart abolishes** 2 **previously existing protein level interactions**

3 The single ancient-modern hybrid population in which no *tufB* promoter mutations
4 occurred did accumulate a number of mutations particular to this population, including
5 one in *nusA* gene (**Table 1**). The *nusA* gene is a translation regulator and its protein
6 product is thought to exhibit chaperone activity with direct interaction to ribosomal
7 proteins (Shazand et al. 1993). NusA may affect the efficiency of the translation
8 machinery in a manner similar to EF-Tu and other ribosomal proteins with chaperone
9 activity (Caldas et al. 2000; Caldas et al. 1998). The mutation region is located at the C-
10 terminal domain of the protein, and the complete deletion may be potentially
11 detrimental for the protein's chaperone activity (**Supplementary Figure 4, 5**). To
12 examine the biochemical effects of the *nusA* mutation, specifically the 27bp deletion
13 (*nusA*Δ9), the mutant protein was cloned in an expression vector and subsequently
14 purified. Changes in interaction of the mutant NusA protein with EF-Tu, ancient or
15 modern was examined by measuring protein-protein binding via isothermal titration
16 calorimetry. While the wild-type EF-Tu bound NusA with a robust binding constant (Kd)
17 of $14.6 \pm 5.2 \mu\text{M}$, the ancient EF-Tu binds only weakly to the native NusA protein
18 (**Figure 5**). Moreover, dipeptide formation assays detected no NusA-EF-Tu interaction
19 in the ribosome, and the interaction between EF-Tu and NusA had no observable effect
20 on dipeptide formation in the ribosome (**Supplementary Figure 5**).

21 On the other hand, *nusA*Δ9 and ancient EF-Tu exhibit a Kd of $680 \pm 66 \mu\text{M}$, suggesting
22 virtually no interaction (**Supplementary Figure 5**). The loss of interaction might be due
23 to the lack of interaction between EF-Tu in the ancestral context in which the

1 reconstructed ancient EF-Tu existed. To test this hypothesis, ancient *nusA* gene
 2 representing the γ -proteobacterial ancestor was phylogenetically reconstructed,
 3 synthesized, expressed, purified, and examined the ancient NusA protein's capacity to
 4 bind to the ancient EF-Tu protein (**Supplementary Figure 6**). No detectable interaction
 5 between the two ancient proteins was observed. Deletion of the mutant *nusA* gene from
 6 the evolved and the ancestral allele has no observable fitness effect. However, deletion
 7 of *nusA* did cause a statistically significant fitness drop of ~8% in the *tufA⁻ tufB⁺*
 8 background (**Figure 5**). The *nusA* mutation is therefore neutral in evolved hybrid
 9 background and the mutation is independent of the ancient-EF-Tu mechanism.

10 Discussion

11 By combining a unique set of tools drawn from synthetic biology, evolutionary biology
 12 and genomics we experimentally evolved and then analyzed the adaptive properties of
 13 a single-celled organism containing a reconstructed ancestral gene inserted in its
 14 genome. A majority of the evolved lineages accumulated mutations in the promoter
 15 region of the ancestral *tuf* gene and these lead to increased expression of the ancient
 16 EF-Tu protein. It is possible that these promoter mutations constitute 'low hanging fruit'
 17 of compensatory genetic changes, particularly for highly-conserved essential proteins,
 18 and that structural mutations in the ancient *tuf* gene might have been observed had
 19 evolution been allowed to continue. Understanding the lack of direct accumulation of
 20 mutations on the ancient EF-Tu requires a full accounting of the fitness effects of all
 21 potentially contributing mutations. Considering the important role of EF-Tu in the
 22 translational machinery, mutations accumulating directly on the EF-Tu gene can cause
 23 cell lethality and thus may not be readily adaptive (Goldman et al. 2010; Kacar and
 24 Gaucher 2013; Pereira-Leal et al. 2006). On the other hand, there likely are beneficial

1 mutations that can occur, but they don't confer the advantage that others do under the
2 same conditions, resulting in tuf mutations always being out-competed early on.
3 Increasing the cellular EF-Tu protein level may be the first emergency response of the
4 organism to cope with a drastic alteration introduced by a maladapted protein central
5 to translation machinery (Bridgham et al. 2009; Gong and Bloom 2014; Kryazhimskiy et
6 al. 2014; Kvitek and Sherlock 2011; Lunzer et al. 2010).

7

8 Engineering native genomes with ancient genes has been considered a challenging
9 experimental approach due to the possibility of functional incompatibility of the
10 ancestral genes in modern organisms (Hobbs et al. 2015). Moreover, altering essential
11 genes carries the risk of drastic effects on cellular epistatic networks (Coulomb et al.
12 2005; Drummond et al. 2005; Zotenko et al. 2008). Even a single mutation in the
13 translation machinery can interfere with protein expression, and thereby drastically
14 impact an organism's viability (Ito et al. 1998; Lind and Andersson 2013). Indeed, if
15 phenotypically altered by subsequent functional divergence and promiscuity over time
16 (Copley 2003), ancestral sequences could be maladapted to the host cell to the extent
17 that a functional organism is all but precluded. However, this experimental limitation
18 does not apply to reconstructed ancestral genes alone. It has been suggested that as
19 the number of nodes connecting a protein within its protein-protein interaction network
20 increases, the capacity to replace that protein with another homolog decreases despite
21 the presumed functional equivalence between the endogenous gene and the homolog
22 (Jain et al. 1999). While a careful assessment of candidate ancestral protein properties
23 prior to integration is helpful, in most cases, studying gene-triggered genomic
24 perturbations through integration of ancestral genes offers a valuable and

1 complementary alternative to existing methodologies that use orthologous proteins (Pal
2 et al. 2014).

3

4 This work was originally conceived as a way to reconstruct the historical pathways by
5 which ancestral genes evolved into modern ones by replacing the modern gene with a
6 reconstructed ancestral one in a modern organism, and then performing experimental
7 evolution with the hybrid. No mutation was detected within the ancestral gene. Instead,
8 all likely adaptive mutations that compensated for the fitness detriment caused by the
9 substitution of the suboptimal ancestral gene occurred at other sites, including within
10 the promoter region of the ancestral gene. This result does not, however, rule out the
11 possibility that useful information on historical gene evolution might be gleaned with
12 this method. The most consistent drivers of historic mutational change may be
13 macroscopic variables (i.e., atmospheric composition, nutrient availability, ecological
14 partitioning, or long-term climate fluctuations) that cannot be readily incorporated into
15 laboratory-scale synthetic evolution experiments. In just one example relevant to this
16 experimental setup, the EF-Tu protein phenotype is tightly coupled to the optimal
17 growth temperature of its host organism (Gromiha et al. 1999), but bacterial clades are
18 not thought to have gone through any significant temperature-dependent evolutionary
19 bottlenecks over the 700 million years of evolution that has occurred between the ages
20 of the modern and ancestral homologous sequences (Blattler and Higgins 2014; Knauth
21 2004). This is one possible interpretation for the observed lack of mutations on EF-Tu
22 itself. Future applications of this method should focus on linking substituted component
23 behavior with a demonstrable organismal phenotype that can be independently
24 retraced over the inferred age of the component (Kacar et al. 2016).

25

1 **Conclusions**

2 Engineering bacterial genomes with phylogenetically reconstructed genes complements
3 the current technique of genome level alterations of gene and gene clusters with
4 currently existing homologs, and provides insights into molecular mechanisms of
5 adaptation by providing access to the historical states of currently existing proteins.
6 However, these methods are also severely constrained by limited existing knowledge of
7 how laboratory evolution setups impact evolutionary trajectories, especially for
8 synthetic biology applications, and the connection between environment and protein
9 phenotype. This knowledge is critical for discerning the change in behavior due to the
10 ancestral state of the protein from the change in systems-level behavior attributable to
11 its intrinsic response to a suboptimal cellular component. The synthetic system
12 described here may enable the development of ancient-modern hybrid model systems
13 that will provide new insights related to the role of evolutionary history and the “tape” of
14 evolution as well as the degree of coupling between protein-level biochemical attributes
15 and macroscale evolutionary trajectories and biogeochemical cycles.

1 **Acknowledgements**

2 We gratefully acknowledge Zachary Blount, Donald Burke and Seth Childers for their
3 comments on this manuscript. Neerja Hajela for providing the *E. coli* REL606 strain,
4 members of the Richard Lenski laboratory for providing feedback during the earlier
5 phase of this work, Deepak Unni for his assistance during the Breseq analyses,
6 undergraduates Lily Tran (Georgia Tech) and Jennifer Zhang (Georgia Tech) for their
7 assistance in the laboratory, and Brian Hammer (Georgia Tech) for providing the
8 pBBRlux plasmid.

9

10 **METHODS**

11 **Bacterial Strains and Culture Conditions**

12 All experiments were done at 37 °C unless stated otherwise. Luria-Bertani (LB) broth
13 was used as the base medium for liquid cultures and agar plates. Experimental
14 evolution and competition assays were carried out in Davis-Mingioli minimal medium
15 (DM) supplemented with 25 mg/L glucose (Davis 1950). Tetrazolium Arabinose (TA)
16 plates were used as the base for competition experiment plating (Reference). When
17 required, LB and DM media were supplemented with kanamycin, chloramphenicol and
18 tetracycline antibiotics. All dilutions were carried out in 0.1% sterile saline. LB and DM
19 cultures were incubated on a rotary shaker 200 rpm and 150 rpm, respectively. The
20 REL606 parent strain was kindly donated by Richard Lenski. Genes coding for ancestral
21 EF-Tu were codon optimized for expression in *E. coli*, chemically synthesized by DNA
22 2.0, and cloned into a pET15b plasmid as reported previously (Gaucher et al. 2008).

23

24

25

1 **Recombineering**

2 *Construction of the ancient-modern hybrid strain*

3

4 Integration of the ancient EF-Tu gene (*AnEF*) into the chromosome of *E. coli* strain
5 REL606 was carried out via the λ -red homology recombineering approach as described
6 by Datsenko and Wanner (Datsenko and Wanner 2000). First, linear DNA containing
7 homology sequences of upstream and downstream of *tufA* gene was amplified by PCR,
8 via (5' GTGGTTGCGAAAATCATCGCTAGAAATTCGGGGATCCGTCGACC 3' and 5'
9 TGTAATTAGCCCAGAACTTTAGCAACTGTAGGCTGGAGCTGCTTCG 3'), and pKD13
10 plasmid as template, and then transferred in REL606 cells through electrophoration,
11 together with the temperature sensitive pKD46 plasmid. Recombinants were isolated
12 from LB agar plates containing 50 μ g/ml kanamycin at 37 °C, grown in in liquid LB
13 medium containing 50 μ g/ml Kanamycin and their genomic DNA was isolated using
14 Promega Wizard Genomic DNA Purification Kit. Confirmation PCR was performed using
15 genomic DNA isolated from colonies as a template, with primers aligning to the
16 chromosome outside of the recombination site (5'
17 CAGGCCGTAATTGAAGCCCGTGGTAAATAAGCC 3' and 5'
18 GAATAATTTATTCGTTCTGACAGTACGAATAAG 3'). Once the successful replacement
19 of *tufA* gene with the kanamycin marker was confirmed via Sanger sequencing, the
20 strain was transformed with linear DNA containing homology sequences of upstream
21 and downstream *tufB* flanked in between the AnEF DNA construct soed to a
22 chloramphenicol marker originally amplified from the A007 loxP-Cm-loxP plasmid
23 (Gene Bridges GmbH) via Gibson Assembly. The transformants were selected on LB
24 plates containing 25 μ g/ml Chloramphenicol and 50 μ g/ml Kanamycin at 37 °C and
25 the correct insert was screened with primers aligning to the chromosome outside

1 recombination site using Fwd primer 5' TCCGTGTCTTAGAGGGACAATCGATG 3' and
2 Rev primer 5' GCAATTAGCTCAGAACTTTTGCTAC 3'. Once confirmed, both the
3 Kanamycin and the Chloramphenicol markers were removed using pCP20 and 706-Cre
4 plasmids (Gene Bridges GmbH), respectively, followed by the confirmation of the
5 deletions by genomic PCR analysis. Plasmids pKD46 and pCP20 were cured by
6 growing the cultures at 42 °C, the final $\Delta tufA$, $\Delta tufB$:AnEF construct was moved into a
7 fresh ancestral strain via bacteriophage P1 transduction. Freezer stocks of the REL606
8 $\Delta tufA$, $\Delta tufB$:AnEF were prepared by mixing 50% sterile glycerol and overnight liquid
9 cultures originated from a single colony, in 1:2 ratios. All stocks were stored in -80 °C.
10 Isogenic Ara⁺ variants of the REL606 $\Delta tufA$, $\Delta tufB$:AnEF were obtained through gene-
11 goring protocol (Herring et al. 2003) (plasmid pJEB12 is kindly donated by Jeff
12 Barrick).

13

14 *Deletion of nusA gene from the chromosome*

15 The nusA gene from the chromosome of REL606, ancestral REL606 $\Delta tufA$, $\Delta tufB$:AnEF
16 and evolved REL606 $\Delta tufA$, $\Delta tufB$:AnEF strain from lineage Rip2 were replaced with a
17 FRT-kan-FRT fragment in the presence of pKD46 helper plasmid as described by
18 Datsenko and Wanner using primers 5' TCCTGCGTGAAGATATGCTG 3' and 5'
19 TCACTTCTTCGCCGATTCT 3'. PCR amplification of the recombination region and
20 sanger sequencing of this amplified region confirmed the correct replacement and the
21 removal of the selection cassette. The cassette was then removed from the
22 chromosome via pCP20, followed by the curation of pKD46 and pCP20 plasmids at 42
23 °C.

24

1 **Growth assays**

2 Saturated overnight cultures were preconditioned by dilution into sterile saline by a
3 1:100, then again by 1:100 into DM25 medium, followed by an overnight growth.
4 Preconditioned cultures were diluted 1:100 into the assay medium and 100 μ L
5 transferred into a 96-well microplate. OD readings were taken at 420 nm every 15
6 minutes with continuous shaking between readings for 24 hours.

8 **Experimental Evolution**

9 Experimental evolution was carried out using a serial transfer regime in DM25 medium
10 for 2000 generations (~6.6 generations per day) as described previously (Elena and
11 Lenski 2003). Relative fitness change was measured by competing evolved strains or
12 populations against the ancestral genotype, REL606 or REL607, every 500 generations
13 using a standard competition assay protocol. Relative fitness was defined as the ratio
14 of the Malthusian parameter of one competitor to the other. The Malthusian parameter
15 was calculated as follows: $m = (cdx \cdot f^x) / cd_0$, where cd_0 = count of the competitor on
16 day 0, and cdx = count of the competitor on day x , f = growth of the population over
17 time $(x-0)$. In our competitions, $f = 100$ because our transfers involve 100-fold dilution
18 and subsequent outgrowth.

20 **Whole genome sequencing**

21 Sequencing libraries of clones of interest were prepared by isolating 3 mg of genomic
22 DNA from bacteria grown in 10 mL LB overnight, fragmented and tagged the isolated
23 DNA with specific Illumina adapters using Nextera DNA sample preparation kit. The
24 product was purified using the Zymo DNA Clean and Concentrator Kit, dual-indexed

1 the libraries with TruSeq Dual Indexed Sequencing primer sets and ensured the
2 products were pure using an Agilent 2100 BioAnalyzer. Sets of compatible barcodes
3 (11plex) were combined into a single lane in an Illumina HiSeq 2500 Rapid Run flow cell
4 (v1) after QC. Sequencing was in a paired end 2 x 100 basepair format (PE100) using
5 TruSeq Rapid SBS reagents. Mutations were identified using the Breseq (0.23) pipeline
6 (Deatherage and Barrick 2014).

7

8 *Fitness measurement of the ancestral strain in the presence of over-expressed EF-Tu*

9 Ancient EF-Tu was cloned into a pASK-IBA43 (IBA Life Sciences) vector inducible
10 under a tetracycline promoter using the following primers Forward 5'
11 GTTGAATTCATGTCTAAAGAAAAGTTTGAACGTAC 3' and Reverse 5'
12 CGGGATCCTCAAGCGATGATTTTCGCAACCAC 3', between the Xho and Nde sites
13 leading to plasmid pASK-IBA43-AnEF. Ligation was confirmed using Forward primer 5'
14 GAGTTATTTTACCACTCCCT 3' and Reverse primer 5' CGCAGTAGCGGTAAACG 3'.
15 The plasmid was transferred to REL606 $\Delta tufA$, $\Delta tufB:AnEF$ cells via electrophoration
16 and selected on LB agar plate with Chloramphenicol. Five representative colonies were
17 picked, preconditioned in LB media containing 250 μ M anhydrous tetracycline for 24
18 hours, followed by a 1:100 dilution into DM media containing glucose. Over-expression
19 of the EF-Tu protein was confirmed through SDS-PAGE gel analysis in comparison to
20 ancestral cells that harbored no plasmid and non-induced plasmid. A REL607 strain
21 was acclimated to the competition environment by separate grown under the same
22 environmental conditions as REL606 $\Delta tufA$, $\Delta tufB:AnEF$ harboring pASK-IBA43-AnEF.
23 The competitors were then mixed in equal stoichiometric ratios by diluting into fresh
24 DM medium with glucose containing 250 μ M anhydrous tetracycline. Samples were
25 plated on tetrazolium arabinose agar plate every 4 hours during the 24 hour

1 competition. The competitions were carried out two times to ensure the precision of
2 fitness estimates.

3

4 Whole genome sequencing was completed for 2000 generations for eight lineages
5 harboring ancient EF-Tu, as well as the wild-type strains. To prepare the sequencing
6 library, we isolated 3 mg of genomic DNA from bacteria grown in 10 mL LB overnight,
7 fragmented and tagged the isolated DNA with specific Illumina adapters using Nextera
8 DNA sample preparation kit. We purified the product using Zymo DNA Clean and
9 Concentrator Kit, dual-indexed the libraries with TruSeq Dual Indexed Sequencing
10 primer sets and ensured the products were pure using a Agilent 2100 BioAnalyzer. We
11 combined the sets of compatible barcodes (11plex) into a single lane on Illumina HiSeq
12 2500 Rapid Run flow cell (v1) after QC. Sequencing was in a paired end 2 x 100bp
13 format (PE100) using TruSeq Rapid SBS reagents. The Breseq (0.23) software was used
14 for the generation and the analysis of the mutations (Deatherage and Barrick 2014).

15 **Fitness measurement of the ancestral strain in the presence of over-expressed** 16 **EF-Tu**

17 Ancient EF-Tu was cloned into a pASK-IBA43 (IBA Life Sciences) vector inducible
18 under a tetracycline promoter using the primers (Forward) 5'
19 GTTGAATTCATGTCTAAAGAAAAGTTTGAACGTAC 3' and (Reverse) 5'
20 CGGGATCCTCAAGCGATGATTTTCGCAACCAC 3', between the *Xho* and *Nde* sites.
21 Ligation was confirmed using Forward primer 5' GAGTTATTTTACCACTCCCT 3' and
22 Reverse primer 5' CGCAGTAGCGGTAAACG 3'. The plasmid was transformed into
23 REL606 $\Delta tufA$, $\Delta tufB$:AncientEFTu cells via electroporation, and transformants selected

on LB agar plate with chloramphenicol. Five representative colonies were picked, preconditioned in LB media containing 250 μ M anhydrous tetracycline for 24 hours, followed by a 1:100 dilution into DM media containing glucose. Over-expression of the EF-Tu protein was confirmed through SDS-PAGE gel analysis in comparison to ancestral cells that harbored no plasmid, or non-induced plasmid. A REL607 strain was acclimated to the competition environment by separate growth under the same environmental conditions as REL606 $\Delta tufA$, $\Delta tufB:Rip$ harboring pASK-IBA43 with the ancient EF-Tu gene. The competitors were then mixed in 50:50 ratios by volume by diluting each into fresh DM25 supplemented with 250 μ M anhydrous tetracycline. Samples were plated on tetrazolium arabinose agar plate every 4 hours during the 24-hour competition. The competitions were carried out twice to increase the precision of fitness estimates.

Luciferase Assay

tufB and pBBRlux plasmid cloning

The wild-type and mutant (evolved) promoter region of *tufB* gene (EF-Tu protein) was cloned into the pBBRlux plasmid as adapted from Lenz et. al (2004) (kindly provided by Prof. Brian Hammer, Georgia Tech). Phusion High-Fidelity DNA polymerase, dNTPs, restriction enzymes (high fidelity), and T4 ligases were all obtained from New England Biolabs. DNA purification materials were purchased from QIAGEN. Promoters were amplified using PCR primers 5'-CAGAATGAAAATCAGGTAGCCGAGTTCCAG-3' and 5'-TAGTGATTGCAGCGGTCAGCGTTGTTTTAC-3' and resulted in a 403 bp product from REL606 *E. coli* in the 4155251-4155654 region of the genome. Restriction sites were subsequently added to the ends of the *tufB* promoter with the following primers: Forward 5'-GATACTAGTCAGAATGAAAATCAGGTAGCCGAGTTCCAG-3' and Reverse

1 5'-TATGGATCCTAGTGATTGCAGCGGTCAGCGTTGTTTTAC-3' (underlying restriction
2 sites correspond to *SpeI* and *BamHI*, respectively). The EF-Tu promoter was cloned
3 upstream of the luciferase operon in the pBBRlux plasmid in order to drive
4 transcription. pBBRlux provides chloramphenicol (CMP) resistance.

5

6 *Scintillation Counts*

7 Four experimental constructs: +86/-29 (G+86A), +54/-61 (G+54T), +87/-28 (A+87C),
8 +94/-21 (19bp duplication, +96); and two control constructs: P (no promoter), Patuf
9 (wild-type, or unevolved ancestor, *tufB* promoter) were transformed into chemically
10 competent *E. coli* (REL606) cells and incubated at 37 °C for 24 hours on
11 chloramphenicol (CMP) agar plates. A single colony was cultured in LB media
12 containing CMP at 37°C for 24 hours. A 100µL aliquot of the overnight culture was
13 diluted one thousand-fold prior to being transferred into a 50 mL Erlenmeyer flask
14 containing 9.9 mL of DM25 media. Cells were grown for ~8.25 hours, or ~5 doublings
15 as monitored by plating (this represents the end of log growth since these cultures
16 reach stationary phase after ~6.6 generations in DM25) and then pelleted. The
17 supernatant was aspirated until 100 µL of media remained, and the pellet was then
18 resuspended in the remaining 100 µL. Scintillation counting was used to quantify the
19 amount of light signal generated by the luciferase pathway. For all six constructs, three
20 readings per sample were averaged for each of the two replicates assayed.

21

22

23

24

25

1 **Bacterial Enumeration**

2 For each construct, a 10µL aliquot was serially diluted 50 thousand-fold and 50 µL was
3 plated on agar petri dishes containing CMP. Extrapolation was utilized to determine the
4 total amount of cells in each scintillation assay. Three plates per flask were averaged.

6 *Luciferase Assay Statistical Analysis*

7 The luciferase expression per cell was normalized by:

$$\frac{\text{total scintillation counts}}{\text{total number of cells}}$$

8
9 Luciferase expression for each construct was subtracted by the amount of luciferase
10 signal from P to eliminate any leaky expression from the pBBRlux vector without
11 promoter and presented as fold-change relative to the amount of luciferase signal from
12 Patuf. A one-way ANOVA with $\alpha = 0.05$ and a post-hoc Tukey HSD Test were
13 performed against Patuf to determine significant differences.

15 **Protein biochemistry**

16 *Cloning, expression and purification of modern EF-Tu and ancient EF-Tu proteins*

17 Both of the EF-Tu genes were ligated into pET15b plasmid between BamH1/EcoR1
18 sites, containing an N-term His-Tag with Ampicilin resistance. For expression, the
19 plasmids were transferred in a BL21(DE3) strain, the cells were grown in LB media until
20 OD₆₀₀ reading reached 0.6-0.8 and then induced with 1 mM Imidazole for 4 hours. The
21 cells were lysed using Bugbuster protein extraction reagent (EMD Millipore) containing
22 benzonase. For purification of the His-tagged protein from the supernatant, the cleared
23 lysate was transferred into nitrilotriacetic acid (Ni-NTA) resin gravity-flow columns
24 (Qiagen, Hilden, Germany) at 4°C that was pre-equilibrated with lysis buffer (50 mM

1 NaH₂PO₄, 300 mM NaCl, 10 mM imidazole, pH 8). The Ni-NTA gravity-flow column was
2 washed by two times wash with lysis buffer containing 20 mM imidazole. His-tagged
3 protein was eluted using elution buffer (50 mM NaH₂PO₄, 300 mM NaCl, 200 mM
4 imidazole, pH 8).

5

6 *Cloning, expression and purification of NusA proteins*

7 Both the wild type *nusA* and the evolved *nusA* genes were amplified from their host
8 bacterial genome using Forward primer 5'-GTGAAGGTGTCGACGCTGCGTGCGCT-3'
9 and Reverse primer 5'-AGCGCACGCAGCGTCGACACCTTCAC-3'. The amplified DNA
10 was purified through gel extraction, removed from salt and then cloned into a pET15b
11 vector (Novagen) using 5' GGCGACATATGAACAAAGAAATTTTGGC 3' and 5'
12 GGAGCTCGAGTTACGCTTCGTCACCGA 3' primers in between BamH1 and XhoI sites.
13 The plasmids were transferred into a BL21(DE3) strain for expression and induced by
14 IPTG. Cells were broken by French Press in Buffer A (20 mM Tris-HCl at pH 7.5, 50 mM
15 MgCl₂, 200 mM NaCl, 5% glycerol). 25 μM GDP was added in Buffer A for EF-Tu
16 purification. After centrifugation for 30 min at 16,000 rpm (F21-8x50 rotor, Thermo), the
17 supernatant was applied to Ni-NTA column and elute with gradient Buffer B (Buffer A
18 supplied with 500 mM imidazole). To prepare EF-Tu and NusA for ITC experiment, the
19 proteins were dialyzed in Buffer C (20 mM Tris-HCl at pH 7.5, 50 mM MgCl₂, 100 mM
20 KCl) for 16 hours at 4 °C.

21

22 *ITC Analysis*

23 The ITC data was measured on a Microcal ITC200 System (GE Healthcare). The syringe
24 was loaded with 42 μL of 0.6-1 mM NusA and the sample cell was filled with 10uM EF-
25 Tu. NusA was titrated (2.5 μL for each) into EF-Tu with 120 s intervals and the first

1 injection was 0.25 μ L. The stirring speed was set at 1000 rpm. Blank experiment was
2 measured by titrating NusA into Buffer C (20 mM Tris-HCl at pH 7.5, 50 mM $MgCl_2$, 100
3 mM KCl).

4

5 **LC-MS/MS Analysis**

6 *Sample preparation*

7 Whole cell lysate was generated from each ancestral and evolved strains using Bug
8 Buster reagent (EMD Millipore) and following manufacturer's instructions. Total protein
9 was quantified via BCA assay using Pierce BCA protein assay kit (Thermo Fisher
10 Scientific). 30 mg of whole cell lysate were submitted to the Proteomics and
11 Metabolomics Facility at Colorado State University. Samples were processed for in-
12 solution trypsin digestion as previously described (Schauer et al. 2013). Briefly, protein
13 was precipitated out of solution in the presence of 4 volumes of 100% -20°C acetone
14 and then resolubilized in 8 M urea, 0.2% ProteaseMAXtm surfactant trypsin enhancer
15 (Promega, Madison, WI). Samples were reduced and alkylated with 5mM dithiothreitol
16 and 5 mM iodoacetamide. Trypsin (MS Grade, Thermo Pierce, San Jose, CA) was
17 added at an enzyme to substrate ratio of 1:50 and incubated at 37 °C for 3-hours.
18 Trypsin was deactivated with the addition of 5% trifluoroacetic acid and desalted using
19 C18 OMIX tips (Agilent Technologies, Santa Clara, CA) using manufacturer's
20 instructions. Peptide eluate was dried in a vacuum evaporator and resuspended in 3%
21 acetonitrile/0.1% formic acid at a concentration of approximately 1 μ g/ μ l. Relative
22 Quantitation of EF-Tu proteins were carried out using spectral counting approach.
23 Approximately 2 μ g of tryptic digest for each sample was injected using an EASY
24 nanoLC-II system (Thermo Scientific, San Jose, CA). Peptides were purified and
25 concentrated using an on-line enrichment column (EASY-Column, 100 μ m ID x 2cm

ReproSil-Pur C18). Subsequent chromatographic separation was performed on a reverse phase nanospray column EASY-Column, 3 m m, 75 m m ID x 100mm ReproSil-Pur C18) using a 180 minute linear gradient from 10%-55% buffer B (100% ACN, 0.1% formic acid) at a flow rate of 400 nanoliters/min. Peptides were eluted directly into the mass spectrometer (Thermo Scientific Orbitrap Velos). The instrument was operated in Orbitrap-LTQ mode where precursor measurements were acquired in the Orbitrap (60,000 resolution) and MS/MS spectra (top 20) were acquired in the LTQ ion trap with normalized collision energy of 35%. Mass spectra were collected over a m/z range of 400-2000 Da using a dynamic exclusion limit of 2 MS/MS spectra of a given peptide mass for 30 s (exclusion duration of 90 s). Compound lists of the resulting spectra were generated using Xcalibur 2.2 software (Thermo Scientific) with a S/N threshold of 1.5 and 1 scan/group.

Data Analysis – Spectral Counting:

Database searching Tandem mass spectra were extracted, charge state deconvoluted and deisotoped by ProteoWizard version 3.0. All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.3.02). Mascot was set up to search the Uniprot_e_coli_custom_reverse database (Updated August 2014, 8750 entries) (Elias and Gygi 2010) assuming the digestion enzyme trypsin, allowing up to 3 missed cleavages. Mascot was searched with a fragment ion mass tolerance of 0.80 Da and a parent ion tolerance of 20 PPM. Oxidation of methionine M (+15.99) and carbamidomethyl of cysteine C (+57) were specified in Mascot as variable modifications.

1 *Criteria for protein identification*

2 Scaffold (version Scaffold_4.3.4, Proteome Software Inc., Portland, OR) was used to
3 validate MS/MS based peptide and protein identifications. Peptide identifications were
4 accepted if they could be established at greater than 69.0% probability to achieve an
5 FDR less than 0.1% by the Scaffold Local FDR algorithm. Protein identifications were
6 accepted if they could be established at greater than 99.0% probability to achieve an
7 FDR less than 1.0% and contained at least 2 identified peptides (Kall et al. 2008; Keller
8 et al. 2002). Protein probabilities were assigned by the Protein Prophet algorithm
9 (Nesvizhskii et al. 2003). Proteins that contained similar peptides and could not be
10 differentiated based on MS/MS analysis alone were grouped to satisfy the principles of
11 parsimony.

13 *Quantitative analysis*

14 Binary comparisons were created in separate Scaffold files comparing wild type *E. coli*
15 REL606 and unevolved ancestor harboring the ancient protein and the evolved lineages
16 tested (biological replicates n=3) to Strain/Treatment Group (each n=3). Biological
17 samples were organized into Categories based on strain type. Each Category had 3
18 biological replicates. Normalization of spectral counts was not applied based on these
19 criteria: An equal amount of sample from each replicate was loaded into the mass
20 spectrometer and there was no deviation in processing and the number of spectra
21 between samples is closely similar (% CV < 5% between biological replicates). Spectral
22 counting uses the sum of the MS/MS spectra assigned to each protein as a measure of
23 abundance (Paoletti and Washburn 2006). A T-Test was performed on Total Spectral
24 Counts for each MS sample using the embedded algorithm in Scaffold v 4.3.4. Proteins
25 with P-values less than 0.05 are excluded in calculation of fold changes compared to *E.*

1 *coli* REL606.

2

References

- 1
2 Abdulkarim F, Hughes D (1996) Homologous recombination between the tuf genes of
3 *Salmonella typhimurium*. *J Mol Biol* 260:506

- 4 Acevedo-Rocha CG, Fang G, Schmidt M, Ussery DW, Danchin A (2013) From essential
5 to persistent genes: a functional approach to constructing synthetic life. *Trends*
6 *Genet* 29:273

- 7 Agashe D, Martinez-Gomez NC, Drummond DA, Marx CJ (2013) Good codons, bad
8 transcript: large reductions in gene expression and fitness arising from
9 synonymous mutations in a key enzyme. *Mol Biol Evol* 30:549

- 10 Agirrezabala X, Frank J (2009) Elongation in translation as a dynamic interaction among
11 the ribosome, tRNA, and elongation factors EF-G and EF-Tu. *Q Rev Biophys*
12 42:159

- 13 Andersson DI, Hughes D (2009) Gene amplification and adaptive evolution in bacteria.
14 *Annu Rev Genet* 43:167

- 15 Andersson DI, Jerlstrom-Hultqvist J, Nasvall J (2015) Evolution of new functions de
16 novo and from preexisting genes. *Cold Spring Harb Perspect Biol* 7

- 17 Annaluru N, Ramalingam S, Chandrasegaran S (2015) Rewriting the blueprint of life by
18 synthetic genomics and genome engineering. *Genome Biol* 16:125

- 19 Bar-Rogovsky H, Stern A, Penn O, Kobl I, Pupko T, Tawfik DS (2015) Assessing the
20 prediction fidelity of ancestral reconstruction by a library approach. *Protein Eng*
21 *Des Sel*

- 22 Bell G (2016) Experimental macroevolution. *Proc Biol Sci* 283

- 23 Benner S (1995) Reconstructing Ancient Forms Of Life. *Journal Of Cellular*
24 *Biochemistry*:200

- 25 Benner S (2007) The early days of paleogenetics: connecting molecules to the planet
26 Oxford University Press

- 1 Blattler CL, Higgins JA (2014) Calcium isotopes in evaporites record variations in
2 Phanerozoic seawater SO₄ and Ca. *Geology* 42:711
- 3 Blount ZD (2015) The unexhausted potential of *E. coli*. *Elife* 4
- 4 Brandis G, Bergman JM, Hughes D (2016) Autoregulation of the *tufB* operon in
5 *Salmonella*. *Mol Microbiol* 100:1004
- 6 Bridgham JT, Ortlund EA, Thornton JW (2009) An epistatic ratchet constrains the
7 direction of glucocorticoid receptor evolution. *Nature* 461:515
- 8 Caldas T, Laalami S, Richarme G (2000) Chaperone properties of bacterial elongation
9 factor EF-G and initiation factor IF2. *J Biol Chem* 275:855
- 10 Caldas TD, El Yaagoubi A, Richarme G (1998) Chaperone properties of bacterial
11 elongation factor EF-Tu. *J Biol Chem* 273:11478
- 12 Cambray G, Mutalik VK, Arkin AP (2011) Toward rational design of bacterial genomes.
13 *Curr Opin Microbiol* 14:624
- 14 Chang BS, Jonsson K, Kazmi MA, Donoghue MJ, Sakmar TP (2002a) Recreating a
15 functional ancestral archosaur visual pigment. *Mol Biol Evol* 19:1483
- 16 Chang BSW, Kazmi MA, Sakmar TP (2002b) Synthetic gene technology: Applications to
17 ancestral gene reconstruction and structure-function studies of receptors G
18 Protein Pathways, Pt A, Receptors, pp. 274-294
- 19 Conrad TM, Joyce AR, Applebee MK, Barrett CL, Xie B, Gao Y, Palsson BØ. (2009)
20 Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing
21 short-term laboratory evolution in lactate minimal media reveals flexible
22 selection of adaptive mutations. *Genome Biol.* 10:10
- 23 Copley SD (2003) Enzymes with extra talents: moonlighting functions and catalytic
24 promiscuity. *Curr Opin Chem Biol* 7:265
- 25 Copley SD (2012) Toward a systems biology perspective on enzyme evolution. *J Biol*
26 *Chem* 287:3

- 1 Coulomb S, Bauer M, Bernard D, Marsolier-Kergoat MC (2005) Gene essentiality and
2 the topology of protein interaction networks. *Proc Biol Sci* 272:1721
- 3 Counago R, Chen S, Shamoo Y (2006) In vivo molecular evolution reveals biophysical
4 origins of organismal fitness. *Molecular Cell* 22:441
- 5 Czworkowski J, Moore PB (1996) The elongation phase of protein synthesis. *Prog*
6 *Nucleic Acid Res Mol Biol* 54:293
- 7 Dalchau N, Smith MJ, Martin S, Brown JR, Emmott S, Phillips A (2012) Towards the
8 rational design of synthetic cells with prescribed population dynamics. *J R Soc*
9 *Interface* 9:2883
- 10 Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in
11 *Escherichia coli* K-12 using PCR products. *Proceedings of the National*
12 *Academy of Sciences of the United States of America* 97:6640
- 13 Dean AM, Thornton JW (2007) Mechanistic approaches to the study of evolution: the
14 functional synthesis. *Nature Reviews Genetics* 8:675
- 15 Deatherage DE, Barrick JE (2014) Identification of mutations in laboratory-evolved
16 microbes from next-generation sequencing data using breseq. *Methods Mol*
17 *Biol* 1151:165
- 18 Deutschbauer A, Price MN, Wetmore KM, Tarjan DR, Xu Z, Shao W, Leon D, Arkin AP,
19 Skerker JM (2014) Towards an informative mutant phenotype for every bacterial
20 gene. *J Bacteriol* 196:3643
- 21 Dillon MM, Rouillard NP, Van Dam B, Gallet R, Cooper VS (2013) Diverse phenotypic
22 and genetic responses to short-term selection in evolving *Escherichia coli*
23 populations. *Evolution* 70:3
- 24 Dragosits M, Mattanovich D (2013) Adaptive laboratory evolution -- principles and
25 applications for biotechnology. *Microb Cell Fact* 12:64
- 26 Drummond DA, Bloom JD, Adami C, Wilke CO, Arnold FH (2005) Why highly expressed
27 proteins evolve slowly. *Proc Natl Acad Sci U S A* 102:14338
- 28 Elena SF, Lenski RE (2003) Evolution experiments with microorganisms: The dynamics
29 and genetic bases of adaptation. *Nature Reviews Genetics* 4:457

- 1 Elias JE, Gygi SP (2010) Target-decoy search strategy for mass spectrometry-based
2 proteomics. *Methods Mol Biol* 604:55
- 3 Esvelt KM, Wang HH (2013) Genome-scale engineering for systems and synthetic
4 biology. *Mol Syst Biol* 9:641
- 5 Feher T, Burland V, Posfai G (2012) In the fast lane: large-scale bacterial genome
6 engineering. *J Biotechnol* 160:72
- 7 Gallagher RR, Li Z, Lewis AO, Isaacs FJ (2014) Rapid editing and evolution of bacterial
8 genomes using libraries of synthetic DNA. *Nat Protoc* 9:2301
- 9 Gaucher EA, Govindarajan S, Ganesh OK (2008) Palaeotemperature trend for
10 Precambrian life inferred from resurrected proteins. *Nature* 451:704
- 11 Goldman AD, Samudrala R, Baross JA (2010) The evolution and functional repertoire of
12 translation proteins following the origin of life. *Biol Direct* 5:15
- 13 Gong LI, Bloom JD (2014) Epistatically interacting substitutions are enriched during
14 adaptive protein evolution. *PLoS Genet* 10:e1004328
- 15 Gould SJ (1989) *Wonderful Life: The Burgess Shale and the Nature of History*. W. W.
16 Norton & Company, New York
- 17 Gromiha MM, Oobatake M, Sarai A (1999) Important amino acid properties for
18 enhanced thermostability from mesophilic to thermophilic proteins. *Biophysical*
19 *Chemistry* 82:51
- 20 Harms MJ, Thornton JW (2013) Evolutionary biochemistry: revealing the historical and
21 physical causes of protein properties. *Nat Rev Genet* 14:559
- 22 Harms MJ, Thornton JW (2014) Historical contingency and its biophysical basis in
23 glucocorticoid receptor evolution. *Nature* 512:203
- 24 Herring CD, Glasner JD, Blattner FR (2003) Gene replacement without selection:
25 regulated suppression of amber mutations in *Escherichia coli*. *Gene* 311:153
- 26 Herron MD, Doebeli M (2013) Parallel evolutionary dynamics of adaptive diversification
27 in *Escherichia coli*. *PLoS Biol*. 11:12

- 1 Hobbs JK, Prentice EJ, Groussin M, Arcus VL (2015) Reconstructed Ancestral Enzymes
2 Impose a Fitness Cost upon Modern Bacteria Despite Exhibiting Favourable
3 Biochemical Properties. *J Mol Evol*
- 4 Hoekstra HE, Coyne JA (2007) The locus of evolution: evo devo and the genetics of
5 adaptation. *Evolution* 61:995
- 6 Huelsenbeck JP, Bollback JP (2001) Empirical and hierarchical Bayesian estimation of
7 ancestral states. *Syst Biol* 50:351
- 8 Isalan M, Lemerle C, Michalodimitrakis K, Horn C, Beltrao P, Raineri E, Garriga-Canut
9 M, Serrano L (2008) Evolvability and hierarchy in rewired bacterial gene
10 networks. *Nature* 452:840
- 11 Ito K, Uno M, Nakamura Y (1998) Single amino acid substitution in prokaryote
12 polypeptide release factor 2 permits it to terminate translation at all three stop
13 codons. *Proc Natl Acad Sci U S A* 95:8165
- 14 Jacob F, Monod J (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J*
15 *Mol Biol* 3:318
- 16 Jain R, Rivera MC, Lake JA (1999) Horizontal gene transfer among genomes: the
17 complexity hypothesis. *Proc Natl Acad Sci U S A* 96:3801
- 18 Johnsen PJ, Levin BR (2010) Adjusting to alien genes. *Mol Microbiol* 75:1061
- 19 Kacar B (2016) Rolling the Dice Twice: Evolving Reconstructed Ancient Proteins in
20 Extant Organisms. In: Ramsey G, Pence CH (eds) *Chance in Evolution* The
21 University of Chicago Press, United States of America, pp. 264-276
- 22 Kacar B, Adam ZR, Hanson-Smith V, Boekelheide N (2016) Constraining the Great
23 Oxidation Event within the Rubisco phylogenetic tree *The Geological Society of*
24 *America*, p 289
- 25 Kacar B, Gaucher EA (2012) Towards the Recapitulation of Ancient History in the
26 Laboratory: Combining Synthetic Biology with Experimental Evolution.
27 *Proceedings of the Artificial Life* 13:11
- 28 Kacar B, Gaucher EA (2013) Experimental evolution of protein-protein interaction
29 networks. *Biochem J* 453:311

- 1 Kall L, Storey JD, MacCoss MJ, Noble WS (2008) Assigning significance to peptides
2 identified by tandem mass spectrometry using decoy databases. *J Proteome*
3 *Res* 7:29
- 4 Kavaliauskas D, Nissen P, Knudsen CR (2012) The busiest of all ribosomal assistants:
5 elongation factor Tu. *Biochemistry* 51:2642
- 6 Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to
7 estimate the accuracy of peptide identifications made by MS/MS and database
8 search. *Anal Chem* 74:5383
- 9 Knauth PL (2004) Temperature and salinity history of the Precambrian ocean:
10 implications for the course of microbial evolution. *Palaeogeography,*
11 *Palaeoclimatology, Palaeoecology* 219:53
- 12 Kolisnychenko V, Plunkett G, 3rd, Herring CD, Feher T, Posfai J, Blattner FR, Posfai G
13 (2002) Engineering a reduced *Escherichia coli* genome. *Genome Res* 12:640
- 14 Kryazhimskiy S, Rice DP, Jerison ER, Desai MM (2014) Microbial evolution. Global
15 epistasis makes adaptation predictable despite sequence-level stochasticity.
16 *Science* 344:1519
- 17 Kvitek DJ, Sherlock G (2011) Reciprocal sign epistasis between frequently
18 experimentally evolved adaptive mutations causes a rugged fitness landscape.
19 *PLoS Genet* 7:e1002056
- 20 Lee JS, An G, Friesen JD, Fill NP (1981) Location of the *tufB* promoter of *E. coli*:
21 cotranscription of *tufB* with four transfer RNA genes. *Cell* 25:251
- 22 Lee SY, Lee DY, Kim TY (2005) Systems biotechnology for strain improvement. *Trends*
23 *Biotechnol* 23:349
- 24 Liberles DA (2007) Using phylogeny to understand genomic evolution Parsimony,
25 Phylogeny, and Genomics. Oxford University Press
- 26 Lind PA, Andersson DI (2013) Fitness costs of synonymous mutations in the *rpsT* gene
27 can be compensated by restoring mRNA base pairing. *PLoS One* 8:e63373

- 1 Lind PA, Tobin C, Berg OG, Kurland CG, Andersson DI (2010) Compensatory gene
2 amplification restores fitness after inter-species gene replacements. *Mol*
3 *Microbiol* 75:1078
- 4 Lunzer M, Golding GB, Dean AM (2010) Pervasive cryptic epistasis in molecular
5 evolution. *PLoS Genet* 6:e1001162
- 6 Lunzer M, Milter SP, Felsheim R, Dean AM (2005) The biochemical architecture of an
7 ancient adaptive landscape. *Science* 310:499
- 8 Lynch VJ, Wagner GP (2008) Resurrecting the role of transcription factor change in
9 developmental evolution. *Evolution* 62:2131
- 10 Maddamsetti R, Lenski RE, Barrick JE (2016) Adaptation, Clonal Interference, and
11 Frequency-Dependent Interactions in a Long-Term Evolution Experiment with
12 *Escherichia coli*. *Genetics* 200: 2
- 13 Merkl R, Sterner R (2016) Reconstruction of ancestral enzymes. *Perspectives in*
14 *Science*
- 15 Michener JK, Camargo Neves AA, Vuilleumier S, Bringel F, Marx CJ (2014) Effective use
16 of a horizontally-transferred pathway for dichloromethane catabolism requires
17 post-transfer refinement. *Elife* 3
- 18 Nandagopal N, Elowitz MB (2011) Synthetic biology: integrated gene circuits. *Science*
19 333:1244
- 20 Negrutskii BS, El'skaya AV (1998) Eukaryotic translation elongation factor 1 alpha:
21 structure, expression, functions, and possible role in aminoacyl-tRNA
22 channeling. *Prog Nucleic Acid Res Mol Biol* 60:47
- 23 Nesvizhskii AI, Keller A, Kolker E, Aebersold R (2003) A statistical model for identifying
24 proteins by tandem mass spectrometry. *Anal Chem* 75:4646
- 25 Nevoigt E, Kohnke J, Fischer CR, Alper H, Stahl U, Stephanopoulos G (2006)
26 Engineering of promoter replacement cassettes for fine-tuning of gene
27 expression in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 72:5266
- 28 Nyerges A, Csorgo B, Nagy I, Balint B, Bihari P, Lazar V, Apjok G, Umenhoffer K, Bogos
29 B, Posfai G, Pal C (2016) A highly precise and portable genome engineering

- 1 method allows comparison of mutational effects across bacterial species. Proc
2 Natl Acad Sci U S A 113:2502
- 3 Ogawa T, Shirai T (2014) Tracing ancestral specificity of lectins: ancestral sequence
4 reconstruction method as a new approach in protein engineering. Methods Mol
5 Biol 1200:539
- 6 Pagel M (1999) Inferring the historical patterns of biological evolution. Nature 401:877
- 7 Pal C, Papp B, Posfai G (2014) The dawn of evolutionary genome engineering. Nat Rev
8 Genet 15:504
- 9 Paoletti AC, Washburn MP (2006) Quantitation in proteomic experiments utilizing mass
10 spectrometry. Biotechnol Genet Eng Rev 22:1
- 11 Pauling L, Zuckerkandl E (1963) Chemical Paleogenetics Molecular Restoration Studies
12 of Extinct Forms of Life. Acta Chemica Scandinavica 17:S9
- 13 Peisajovich SG, Garbarino JE, Wei P, Lim WA (2010) Rapid diversification of cell
14 signaling phenotypes by modular domain recombination. Science 328:368
- 15 Pena MI, Van Itallie E, Bennett MR, Shamoo Y (2010) Evolution of a single gene
16 highlights the complexity underlying molecular descriptions of fitness. Chaos
17 20:026107
- 18 Pereira-Leal JB, Levy ED, Teichmann SA (2006) The origins and evolution of functional
19 modules: lessons from protein complexes. Philos Trans R Soc Lond B Biol Sci
20 361:507
- 21 Phillips KN, Castillo G, Wünsche A, Cooper TF (2016) Adaptation of Escherichia coli to
22 glucose promotes evolvability in lactose Evolution 70:2
- 23 Quan S, Ray JC, Kwota Z, Duong T, Balazsi G, Cooper TF, Monds RD (2012) Adaptive
24 evolution of the lactose utilization network in experimentally evolved populations
25 of Escherichia coli. PLoS Genet 8:e1002444
- 26 Sandoval NR, Kim JY, Glebes TY, Reeder PJ, Aucoin HR, Warner JR, Gill RT (2012)
27 Strategy for directing combinatorial genome engineering in Escherichia coli.
28 Proc Natl Acad Sci U S A 109:10540

- 1 Schauer KL, Freund DM, Prenni JE, Curthoys NP (2013) Proteomic profiling and
2 pathway analysis of the response of rat renal proximal convoluted tubules to
3 metabolic acidosis. *Am J Physiol Renal Physiol* 305:F628

- 4 Schnell R, Abdulkarim F, Kalman M, Isaksson LA (2003) Functional EF-Tu with large C-
5 terminal extensions in an E-coli strain with a precise deletion of both
6 chromosomal tuf genes. *Febs Letters* 538:139

- 7 Shazand K, Tucker J, Grunberg-Manago M, Rabinowitz JC, Leighton T (1993) Similar
8 organization of the nusA-infB operon in *Bacillus subtilis* and *Escherichia coli*. *J*
9 *Bacteriol* 175:2880

- 10 Storz JF, Bridgham JT, Kelly SA, Garland T, Jr. (2015) Genetic approaches in
11 comparative and evolutionary physiology. *Am J Physiol Regul Integr Comp*
12 *Physiol* 309:R197

- 13 Thornton JW (2004) Resurrecting ancient genes: experimental analysis of extinct
14 molecules. *Nat Rev Genet* 5:366

- 15 Tubulekas I, Hughes D (1993) Growth and translation elongation rate are sensitive to the
16 concentration of EF-Tu. *Mol Microbiol* 8:761

- 17 Ugalde JA, Chang BS, Matz MV (2004) Evolution of coral pigments recreated. *Science*
18 305:1433

- 19 Urbanczyk H, Furukawa T, Yamamoto Y, Dunlap PV (2012) Natural replacement of
20 vertically inherited lux-rib genes of *Photobacterium aquimaris* by horizontally
21 acquired homologues. *Environ Microbiol Rep* 4:412

- 22 Van Delft JH, Schmidt DS, Bosch L (1987) The tRNA-tufB operon transcription
23 termination and processing upstream from tufB. *J Mol Biol* 197:647

- 24 van der Meide PH, Vijgenboom E, Talens A, Bosch L (1983) The role of EF-Tu in the
25 expression of tufA and tufB genes. *Eur J Biochem* 130:397

- 26 Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM (2009)
27 Programming cells by multiplex genome engineering and accelerated evolution.
28 *Nature* 460:894

1 Wright PC, Jaffe S, Noirel J, Zou X (2013) Opportunities for protein interaction network-
2 guided cellular engineering. IUBMB Life 65:17

3 Xu P, Vansiri A, Bhan N, Koffas MA (2012) ePathBrick: a synthetic biology platform for
4 engineering metabolic pathways in E. coli. ACS Synth Biol 1:256

5 Zengel JM, Lindahl L (1982) A secondary promoter for elongation factor Tu synthesis in
6 the str ribosomal protein operon of Escherichia coli. Mol Gen Genet 185:487

7 Zhou Y, Asahara H, Gaucher EA, Chong S (2012) Reconstitution of translation from
8 Thermus thermophilus reveals a minimal set of components sufficient for protein
9 synthesis at high temperatures and functional conservation of modern and
10 ancient translation components. Nucleic Acids Research 40:7932

11 Zhu GP, Golding GB, Dean AM (2005) The selective cause of an ancient adaptation.
12 Science 307:1279

13 Zotenko E, Mestre J, O'Leary DP, Przytycka TM (2008) Why do hubs in the yeast
14 protein interaction network tend to be essential: reexamining the connection
15 between the network topology and essentiality. PLoS Comput Biol 4:e1000140
16
17

18

19

20

21

22

23

24

25

26

27

28

1 **Table Captions**

2

3 **Table 1**

4 Parallel mutations in genes for six, initially identical, independently evolved populations
5 harboring an ancient EF-Tu.

6 Top part represents the genes that accumulated mutations in at least three
7 populations containing the ancient EF-Tu gene and occupied the population by
8 minimum 20% across generation 500 to 2000 are shown for a total of 7
9 populations evolved in parallel. thrT/tuf represents the intergenic region between
10 ancient EF-Tu gene and thrT gene. Bottom three are the mutated genes that are
11 specific only to the single lineage that did not accumulate a mutation in the
12 thrT/tuf region (referred as Rip 2). Mutations were detected in genes shown with
13 asterisk are the genes in which mutations were detected in at least one
14 population containing the native EF-Tu gene. Prior studies that report mutations
15 in genes reported here include (Barrick et al. 2009, Maddamsetti et al. 2015,
16 Dillon et al. 2016, Conrad 2009, Herron 2013 and Philips et al. 2014)

17

1 **Figure Captions**

2 **Figure I**

3 Overall experimental scheme

4 A) Phylogenetic reconstruction allows inferences of the ancestral protein sequence.

5 Replacement of the endogenous EF-Tu gene (orange) with the reconstructed

6 ancient EF-Tu (black) allele is indicated by the black-orange hybrid. B) The

7 hybrid population was subjected to adaptive laboratory evolution through daily

8 propagation and transfer of cultures in minimal media C) Relative fitness of

9 evolved populations were assessed via competitive fitness assays as described

10 by Elena 2005, followed by the identification of mutations through whole-

11 genome sequencing and the characterization of the mutants via in-vivo and in-

12 vitro assays.

13

14

1 **Figure II**

2 Fitness values of *E. coli* populations relative to the ancestral strain during adaptive
3 evolution

4 Replacement of the endogenous EF-Tu gene with the reconstructed ancient EF-
5 Tu allele significantly reduces the fitness of the ancient-modern hybrid relative to
6 the original strain (blue dashed line). Hybrid population mean fitness rapidly
7 improved during experimental evolution in minimal glucose medium (blue line).
8 *E. coli* Δ tufA represents the bacteria that contain a single tufB gene Fitness
9 change in *E. coli* Δ tufA is shown in orange. Error bars show 95 % confidence
10 interval among six replicate populations for each system. Red line represents
11 the average total number of genomic changes relative to the ancestor in each
12 sampled hybrid lineage. Mutations are total number of genomic changes relative
13 to the ancestor in each sampled lineage that reach to over 20% fixation.

14

15

1 **Figure III**

2 Analysis of the mutations accumulated in the cis-regulatory region thrT/tufB

3 A) The thrT/tufB promoter region in which five of six evolved hybrid populations
 4 were found to have accumulated mutations. B) The allelic frequency of the
 5 mutations located in ancient EF-Tu gene's promoter region per generation per
 6 population during laboratory evolution.

7

8

1 **Figure IV**

2 A) Relative abundance of ancient EF-Tu protein among evolved hybrid strains
3 using the peak area quantification from MS proteomics data. Error bars obtained
4 using Anova/t-test.

5 B) Growth rates of an isogenic strain of *E. coli* REL606 lacking the *tufA* gene, as
6 well as the unevolved ancient-modern hybrid *E. coli* was evaluated in the
7 presence of with Anhydrotetracycline (ATC) inducer. Strains were induced with
8 500 mg/mL ATC in rich growth media for 3-4 hours to achieve proper induction.
9 Cells from these fresh induced cultures were inoculated in 96-well plates and
10 grown at 37°C with a starting OD₆₀₀ of ~0.06 under respective ATC
11 condition. Doubling times were determined by fitting the exponential growth
12 curves with an exponential function.

13

14

15

16

1 **Figure V**

2 Fitness change after deletion of nusA gene from the ancestral and evolved
3 bacterial genome. (Left) Bacterial constructs with nusA knockouts are
4 constructed and competed against the native E. coli bacteria for fitness
5 measurement. (Right) The interaction between the native EF-Tu, ancient EF-Tu
6 and nusA variants are measured via in vitro Isothermal Calorimetry binding
7 assays.

8

9

10

1 **Supplementary Material Captions**

2 **Supplementary Figure I**

3 Recombineering outline for the allelic replacement of the *tufA* with the ancient
4 EF-Tu gene

5 **Supplementary Figure II**

6 A) Effect of mutations on the expression of reporter gene for the EF-Tu promoter
7 region measured using luciferase reporter assay. B) Relative whole cell
8 concentration of EF-Tu (percent) based on MALDI TOF MS detection

9 **Supplemental Figure III**

10 The replacement of the modern EF-Tu with the ancient EF-Tu caused *E. coli* to
11 be maladapted. As the ancient-modern hybrid populations evolve, the doubling
12 time reduces from ~70 minutes to ~45 minutes. Each red line represents the
13 mean doubling time calculated via three representative clones for each lineage
14 (n=7) in minimal glucose media (DM25).

15 **Supplemental Figure IV**

16 ITC profiles for the titration of NusA to EF-Tu. NusA protein that was purified
17 from plasmids pLT20 (wild-type nusA) (a and c), pLT21 (mutant nusA with 27bp
18 deletion in C-terminal, nusA Δ CTD27) (b and d) was injected to 10 μ M wild type
19 (a and b) and ancient EF-Tu (c and d), respectively. The upper panel of each
20 figure represents the raw plots of enthalpy for each injection (μ cal/s) against
21 time (min). The corresponding bottom panels show integrated heats (closed
22 squares) in each injection against mole ratio. The data points were fitted to a
23 one-site model, suggesting that native EF-Tu interacts with nusA with a K_d of
24 $14.6 \pm 5.2 \mu$ M.

Supplemental Figure V

A) SDS PAGE analysis of the ribosomal elongation complex shows that NusA cannot bind to 70S ribosome. Initiation complex (containing 70S ribosome, initiation factors Kd (μ M) Wild type NusA NusA Δ 9 Wild type EF-Tu 14.6 ± 5.2 153.6 ± 5.4 Ancient EF-Tu 30.1 ± 2.6 680 ± 66 (IF1, IF2 and IF3), XR7-ML mRNA and fMet-tRNA^{fMet}) was mixed with elongation mix (containing EF-Tu GTPase mutant H84A, EF-Ts, Leu tRNA synthetase, Leucine) in the absence and presence of NusA. After incubating the reactions for 10s the mixes were loaded directly on 37% sucrose cushion (100 ml). The samples were centrifuged at 80,000 rpm for 2 hours at 4°C. Ribosomal pellets were loaded in the SDS-PAGE. The gel was stained with Coomassie blue and only the upper part of the gel showing bands corresponding to 30S protein S1, EF-Tu H84A and NusA is displayed. EF-Tu H84A was chosen to show EF-Tu retention in the ribosome since this mutant EF-Tu is defective in release.

The lanes contained following samples:

1. 70S initiation complex (IC), 2. Elongation mix (EM), 3. EF-Tu H84A protein,
4. 70S IC + EM + NusA, 5. NusA protein alone,
6. 70S IC + EM + NusAD27CTD, 7. NusA Δ 9 alone,
8. 70S IC + EM, 9. 70S ribosome

The absence of NusA bands in lanes 4 and 6 suggest the lack of binding of NusA to the 70S ribosome.

1 Supplementary Figure VI

2 NusA protein structure highlighting the two alleles observed during the adaptive
3 laboratory evolution. Nus A protein is made up of a dual repeated acidic domain
4 of 70 residues. These modules are referred to as acidic repeats (AR) 1 and 2.
5 Mutations are located in the C-Terminal, away from the its N-protein binding
6 site.

8 Supplementary Figure VII

9 Sequence alignment of the reconstructed ancestral and present-day NusA
10 protein.

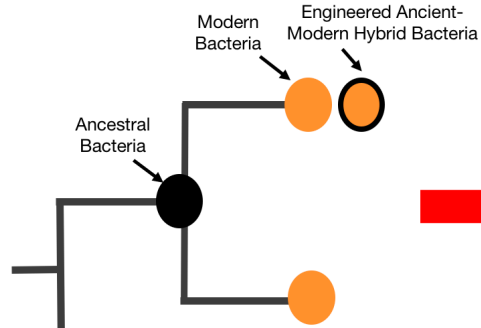
11
12
13
14
15
16
17
18

1 Figures

2

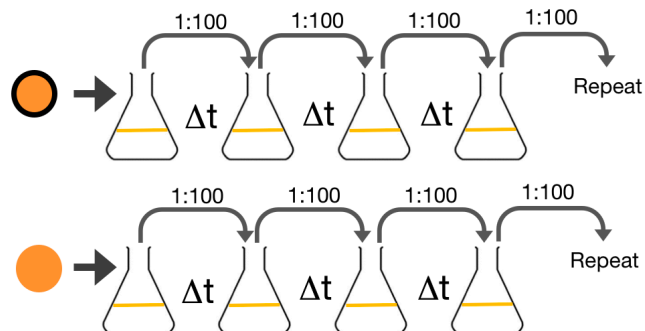
3 Figure 1

a) Phylogenetic and In-vivo resurrection

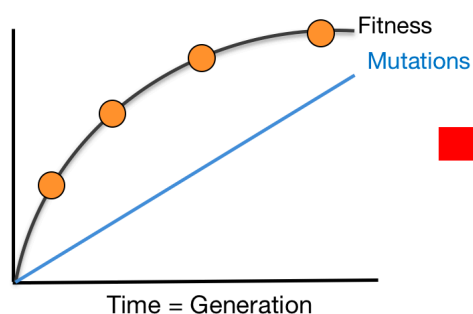


b) Laboratory Evolution

Selection pressure: Environment & Synthetic Ancient Gene



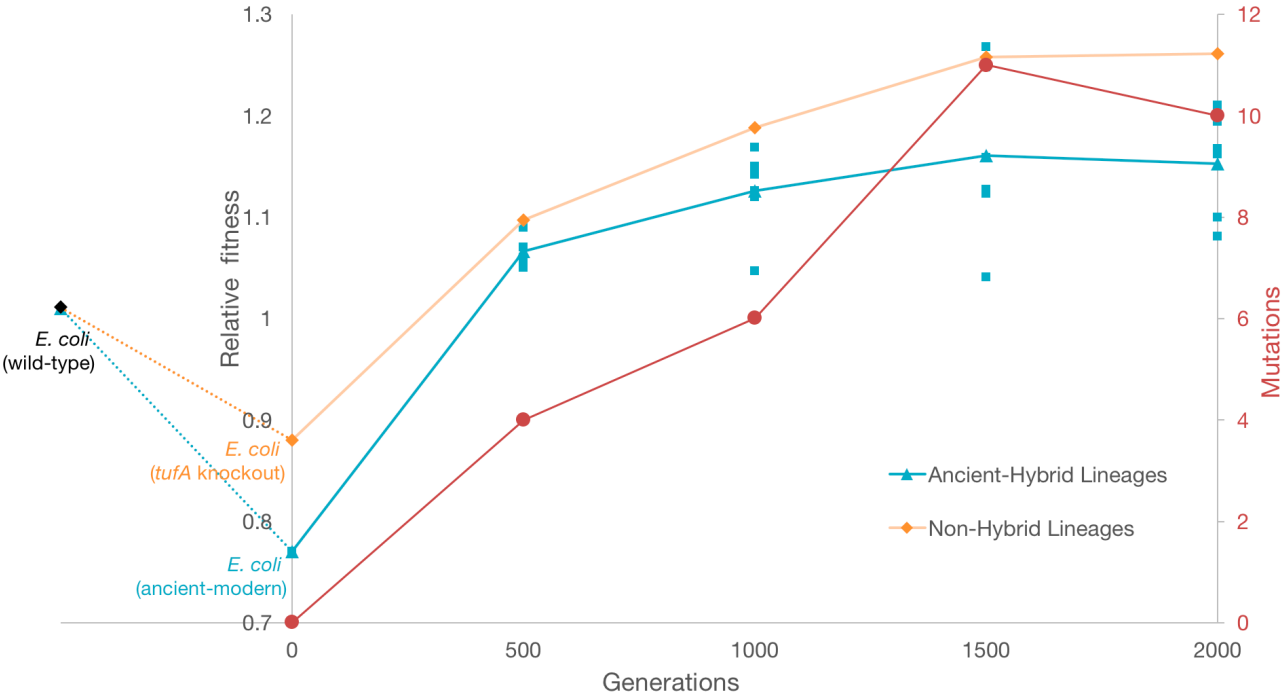
c) Characterization



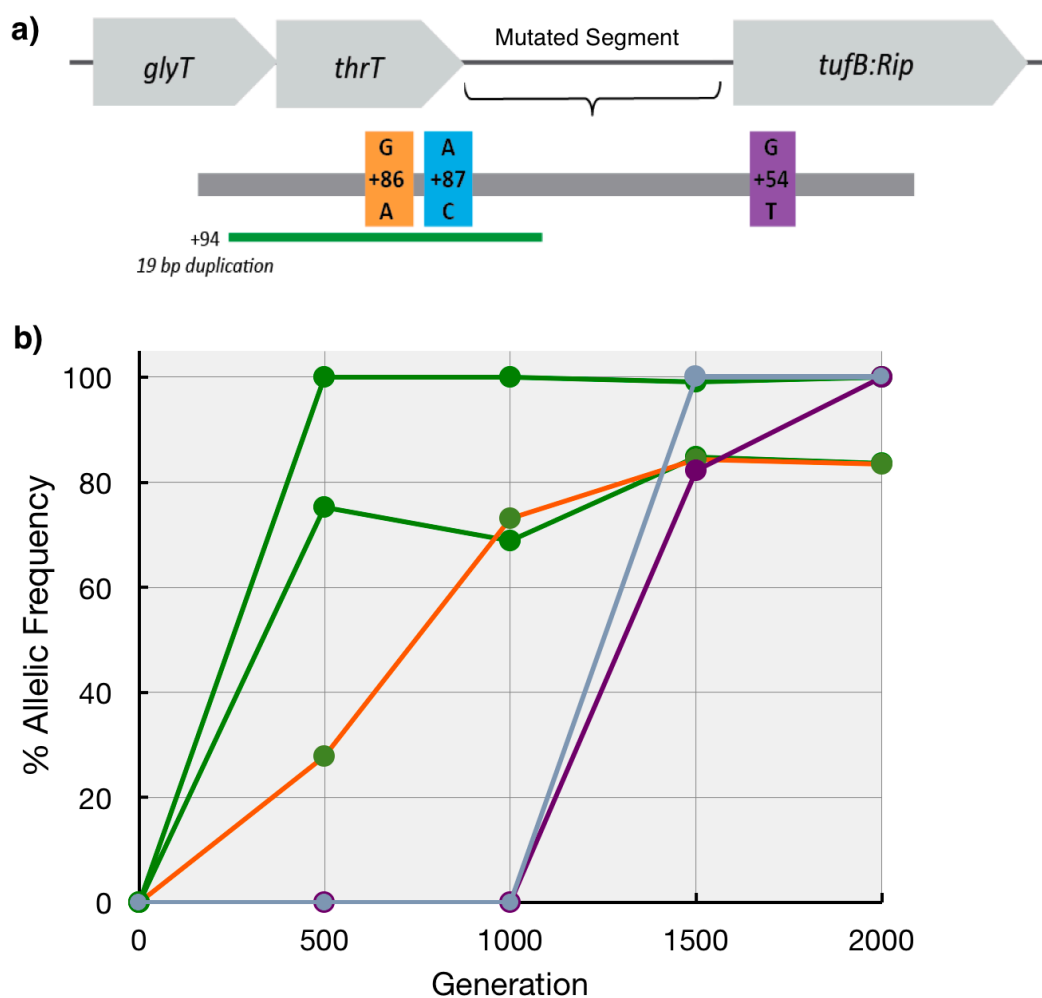
Identify mutations:
Biochemical, genetic and system level characterization

4
5

Figure II

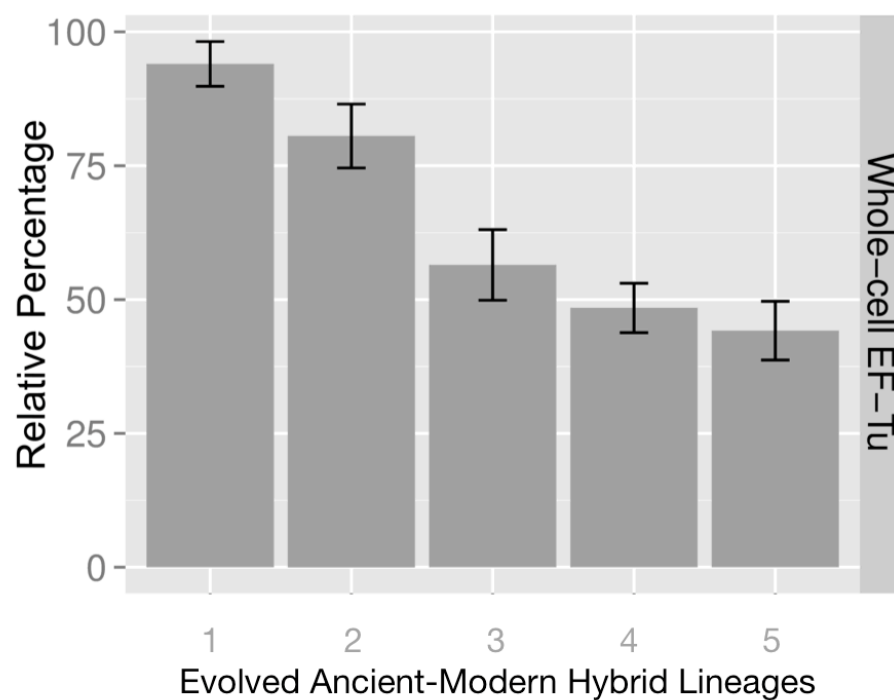


1 **Figure III**



2
3
4
5
6
7

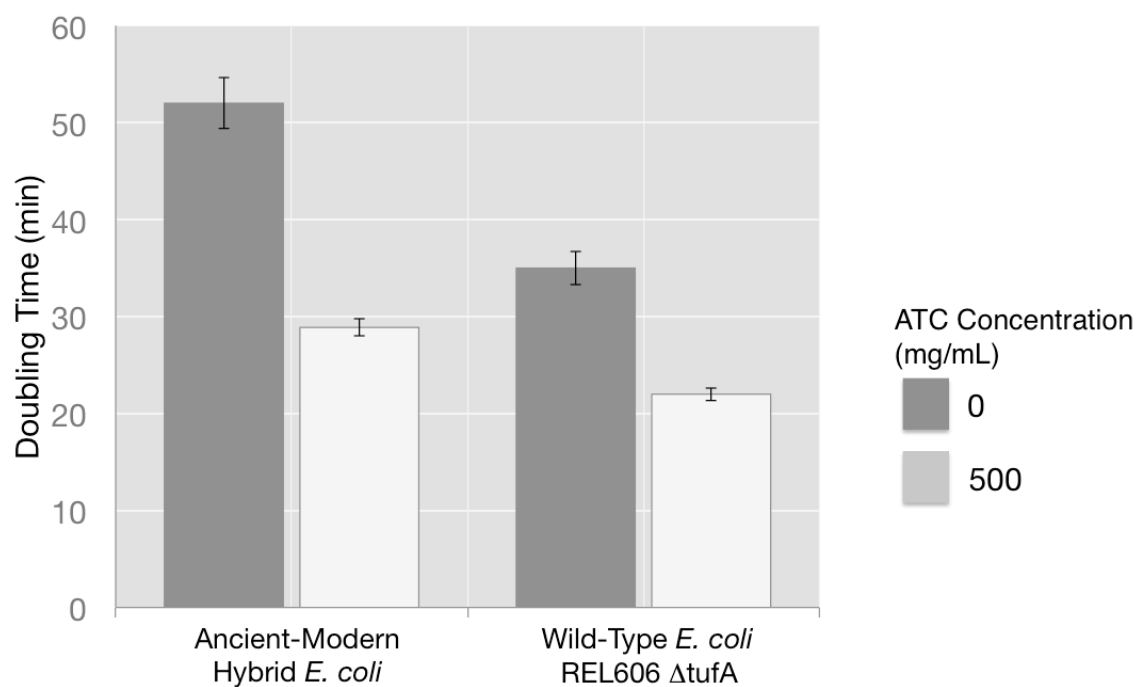
1 Figure IVa



2

3

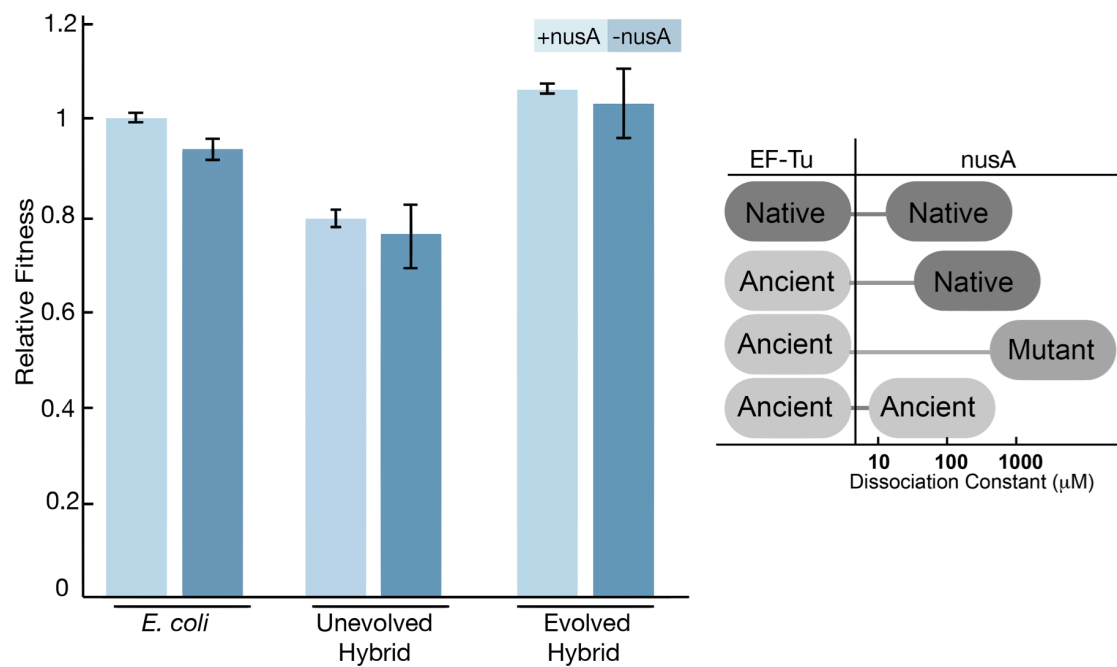
1 Figure IVb



2

3

1 Figure V



2

3

4

5

1 Table I

| Gene | Function | Number of lineages that accumulated a non-synonymous mutation over 20% frequency | | | | Present in non-hybrid lines? | Previously reported for LTEE? |
|----------|--|--|------|------|------|------------------------------|-------------------------------|
| | | Generation | | | | | |
| | | 500 | 1000 | 1500 | 2000 | | |
| mreB | cell wall structural complex MreBCD, actin-like component MreB | 2 | 2 | 3 | 3 | Yes | Yes |
| mrdB | cell wall shape-determining protein | 2 | 2 | 2 | 2 | Yes | Yes |
| thrT/tuf | tRNA-Thr/protein chain elongation factor EF-Tu | 4 | 4 | 5 | 5 | No | No |
| fadA | acetyl-CoA acetyltransferase | 3 | 3 | 3 | 3 | No | Yes |
| ftsZ | cell division protein FtsZ | 1 | 1 | 2 | 5 | No | Yes |
| iclR | DNA-binding transcriptional repressor | - | 1 | 4 | 3 | No | Yes |
| accC | acetyl-CoA carboxylase | - | 1 | 4 | 5 | Yes | Yes |
| pykF | pyruvate kinase | - | 1 | 3 | 5 | Yes | Yes |
| topA | DNA topo isomerase | 1 | 1 | 1 | 1 | Yes | Yes |
| nusA | transcription elongation factor NusA | 1 | 1 | 1 | 1 | No | Yes |
| infB | translation initiation factor IF-2 | 1 | 1 | 1 | - | No | Yes |
| hupA | HU, DNA-binding transcriptional regulator, alpha subunit | 1 | 1 | 1 | 1 | No | Yes |

2

3