The birth of a bacterial tRNA gene

Gökçe B. Ayan[1], Hye Jin Park[1,2,3], Jenna Gallie[1]*

[1]Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany
[2]Asia Pacific Center for Theoretical Physics, Pohang, 37673, Korea
[3]Department of Physics, POSTECH, Pohang, 37673, Korea
*To whom correspondence should be addressed: gallie@evolbio.mpg.de

ABSTRACT

While the major function of transfer RNA is conserved across the tree of life, organisms differ in the tRNA genes that they carry. Here, we provide insight into the mechanisms by which bacterial tRNA gene sets evolve. We use mathematical modelling to demonstrate that while rapid translation generally promotes the elimination of non-essential tRNA types, there are conditions under which retention is predicted. We test our predictions by eliminating non-essential tRNA-Ser(CGA) from *Pseudomonas fluorescens* SBW25, generating a conditional growth defect that is rapidly compensated during an evolution experiment. Compensation occurs through large-scale, tandem duplication events encompassing the gene encoding tRNA-Ser(UGA). Each genetic duplication is accompanied by a two-fold increase in mature tRNA-Ser(UGA). We revisit our models to demonstrate how eliminating tRNA-Ser(CGA) and subsequently increasing tRNA-Ser(UGA) can decrease and then elevate translation speed. Together, our results reveal spontaneous gene duplication as a novel mechanism of bacterial tRNA gene set evolution.

INTRODUCTION

Translation requires 61 types of mRNA codons to be rapidly and accurately decoded into 20 standard amino acids. Decoding is performed by adapter molecules, called transfer RNAs (tRNAs), which carry an amino acid and an anticodon at opposing ends. Each 3-bp anticodon base pairs with, and decodes, specific 3-bp codons. However, anticodon-codon matching is not straightforward; some anticodons can recognize more than one codon (coding for the same amino acid). This allows organisms to use all 61 codons, yet carry fewer than 61 corresponding anticodons (Ikemura, 1981; Ikemura, 1985).

The recognition of multiple codons by a single anticodon is mechanistically possible due to relaxed base pairing rules between the first anticodon and third codon position. Here, some non-traditional, “wobble” base pairings are tolerated in addition to standard Watson and Crick base pairings (Crick, 1966; for an example see **Fig. S1**). Prominent examples are guanine-uracil (G-U) base pairs, and inosine (formed by post-transcriptional modification of an anticodon-first-position adenosine) pairing with uracil, adenine or cytosine (Crick, 1966; Machnicka et al., 2016). Using a combination of the G-U and inosine wobble rules, the number of tRNA types required to translate all 61 codons is reduced to 32 (31 elongators and one initiator; Nelson and Cox, 2013).

Generally speaking, bacteria exploit the G-U and inosine wobble rules to reduce the types of tRNAs that they carry (Chan and Lowe, 2016; Marck and Grosjean, 2002). Typically, bacterial tRNA gene sets consist of a core set of around 33 tRNA types, that together can translate all 61 codons, plus a smaller, more variable set of non-essential tRNA types (Marck and Grosjean, 2002; Wald and Margalit, 2014). These non-essential types are retained even though the codons that they translate can, at least theoretically, be translated by a core tRNA type. In addition to differing in the numbers and types of non-essential tRNAs present, bacterial tRNA gene sets also differ substantially in the total number of tRNA genes they carry. Many tRNA types are coded for by multiple, and often identical, gene copies (Chan and Lowe, 2016; Marck and Grosjean, 2002).

The evolution of variation in bacterial tRNA gene sets is influenced by several factors (reviewed in Rak et al., 2018). One factor is translational speed; a comparative study of tRNA gene sets across 102 bacterial species demonstrates that shorter division times (and concurrently faster translation rates) correlate with fewer tRNA types encoded by higher numbers of tRNA genes (Rocha, 2004). Other factors that co-vary with tRNA gene content are codon use (Bulmer, 1987; Higgs and Ran, 2008; Rudorf, 2019), GC content (Wald and Margalit, 2014), and genome size (Rak et al., 2018).

The molecular mechanisms by which tRNA gene sets evolve remain poorly understood. Thus far, most evidence for tRNA gene set evolution is indirect. Phylogenetic analyses indicate horizontal gene transfer events in the history of a number of bacterial tRNA genes (McDonald et al., 2015; Wald and Margalit, 2014). There is also computational evidence for anticodon switching – the process by which point mutation(s) alter the anticodon and thus the codons recognized by the tRNA – in several bacterial tRNA genes (Wald and Margalit, 2014). An anticodon switch has also been reported in laboratory populations of the yeast *Saccharomyces cerevisiae*; following deletion of the single-copy gene encoding tRNA-Arg(CCU), one of eleven copies of the gene encoding tRNA-Arg(UCU) acquired a point mutation, switching the anticodon to that of the deleted tRNA (Yona et al., 2013).

In this work, we provide insight into the mechanisms by which bacterial tRNA gene sets evolve. Firstly, we use mathematical modelling to demonstrate conditions that promote the elimination or retention of non-essential tRNA types. Secondly, we use a combination of genetic engineering, experimental evolution, and deep sequencing of mature tRNA pools in the bacterium *Pseudomonas fluorescens* SBW25 to reveal spontaneous gene duplication as a novel mechanism of tRNA gene set evolution in bacteria. Finally, we combine our experimental data and modelling to provide a molecular explanation of how the changes observed during experimental evolution may elevate translation speed.

RESULTS

Rapid translation favours fewer tRNA types

In order to investigate conditions under which non-essential tRNA types are lost from tRNA gene sets, mathematical modelling was employed. First, a simple theoretical system was devised to examine how selection for translational speed affects the presence of non-essential tRNA types (**Fig. 1A**). The

model system consists of an mRNA molecule with two sequential codons (*AB*), plus a tRNA pool with two tRNA types (α and β). *Alpha* can translate *A* and *B*, while β can translate only *B*. (*i.e.*, α is an essential tRNA type, β is non-essential).

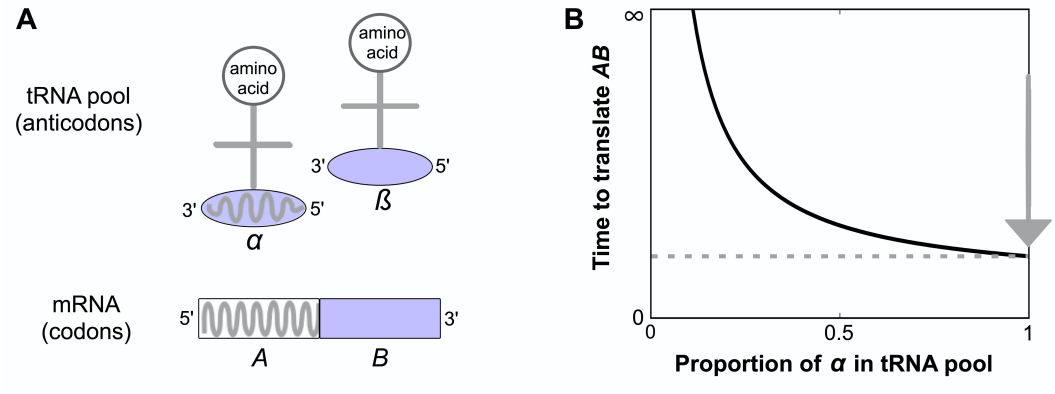

Figure 1. A model system in which rapid translation favours elimination of non-essential tRNA types. (**A**) The system consists of two consecutive mRNA codons (*A*, *B*) that code for the same amino acid, and a pool of equal proportions of two tRNA types (α, β) that carry the same amino acid. *Alpha* (α) can translate *A* (by cognate pairing) and *B* (by wobble pairing), while β can translate only *B* (by cognate pairing). The model assumes that all codon-anticodon pairs that result in translation do so at equal rates. (**B**) Graph showing the time taken to translate *AB* as a function of the relative proportion of α and β in the anticodon pool. Translation is fastest if β is eliminated (grey arrow).

The time taken to translate *AB* can be approximated by the sum of the time required to translate *A* and *B*. The time needed to translate a codon is dependent on the rate of translation by each matching anticodon. Accordingly, the time taken to translate *A* depends only on the rate of translation by α ($\lambda_{\alpha A}$), while the time taken to translate *B* depends on the rate of translation by α ($\lambda_{\alpha B}$) and β ($\lambda_{\beta B}$). Overall, the time taken to translate *AB* (τ) can be estimated by the inverse sum of the rate of translation of *A* and *B* (**Equation 1**). Assuming that each functional anticodon-codon pair translates at an equal rate (*i.e.*, $\lambda_{\alpha A}=\lambda_{\alpha B}=\lambda_{\beta B}=1$), translation of *A* takes twice as long as translation of *B*.

$$\tau = \frac{1}{\lambda_{\alpha A}} + \frac{1}{\lambda_{\alpha B} + \lambda_{\beta B}} \qquad \textbf{Equation 1}$$

In a biological system, the rate of translation by a functional anticodon-codon pair is determined by a number of factors, one of which is the proportion of the anticodon in the mature tRNA pool (Varenne et al., 1984). Under the assumption that the relative proportion of α and β is the only factor influencing translation rates, **Equation 1** can be re-written as **Equation 2**.

$$\tau = \frac{1}{p_\alpha} + \frac{1}{p_\alpha + p_\beta} \qquad \textbf{Equation 2}$$

When the proportion of α and β are equal (*i.e.*, $p_\alpha = p_\beta = 0.5$), then A again takes twice as long to translate as B (**Equation 3**). However, as the proportion of α increases relative to β, the time taken to translate A decreases, and translation of AB is fastest when β is eliminated (**Fig. 1B**).

$$\tau = 2 + 1$$ **Equation 3**

Thus far, the model makes a number of assumptions, including that (i) rapid translation is the driving factor in tRNA gene set evolution, (ii) codons occur at equal frequencies, and (iii) the rate of translation by an anticodon-codon pair is determined solely by the proportion of the anticodon in the tRNA pool. Violation of any of these assumptions may lead to a different prediction. For example, the third assumption would be violated by the existence of translation rates that vary independently of anticodon proportion (*e.g.*, through different base pairing efficiencies, or different tRNA turnover rates). Indeed, varying translation rates independently of anticodon proportions leads to the prediction of retention in cases where translation by a wobble tRNA type is significantly less efficient than that achieved by the cognate nonessential tRNA type (**Text S1**). Specifically, the extended model predicts retention of anticodon β when the translation efficiency of αB is less than 62 % of that of βB.

The models in this section illustrate an important point: unnecessary options reduce translational speed. The simple version of the model predicts that, in cases where rapid translation is important, tRNA gene sets will be streamlined by the elimination of functionally redundant tRNA types. The extended model demonstrates that there are nevertheless conditions under which selection for rapid translation promotes the retention of non-essential tRNA types. These results are in agreement with computational analyses showing that bacteria with faster growth rates tend to encode fewer tRNA types, but still retain some non-essential tRNA types (Rocha, 2004).

Six non-essential tRNA types are predicted in the *P. fluorescens* SBW25 tRNA gene set

Pseudomonas fluorescens SBW25 is a fast-growing bacterium that is often used as an evolutionary model system in the laboratory. The *P. fluorescens* SBW25 genome is predicted to contain 67 tRNA genes of 39 types (GtRNAdb; Chan and Lowe, 2016). According to the G-U and inosine wobble rules (Crick, 1966), six of the 39 types are non-essential; the cognate codon recognized by each can, at least theoretically, be translated by another tRNA type through wobble base pairing.

In the remainder of this work, we delve into the evolutionary and biological relevance of one of the theoretically non-essential, retained tRNA types in SBW25: tRNA-Ser(CGA). This tRNA is encoded by the single copy gene *serCGA*, and translates codon 5'-UCG-3'. According to the G-U wobble rule, codon 5'-UCG-3' can also be translated by tRNA-Ser(UGA), the essential tRNA type encoded by the single copy gene *serTGA* (see **Fig. S1**).

Deletion of *serCGA* limits rapid growth

The entire 90-bp, single copy *serCGA* gene was removed from *P. fluorescens* SBW25 by two-step allelic exchange (**Text S2**). This process was performed two independent times, yielding biological replicate strains Δ*serCGA*-1 and Δ*serCGA*-2. A third round of allelic exchange gave an engineering control strain, SBW25-eWT. Successful deletion of *serCGA* demonstrates that it is indeed a non-essential gene (see also **Fig. 5A** for evidence that *serCGA* is expressed in SBW25). Hence, retention of *serCGA* contravenes the prediction of the simple version of our model: that non-essential tRNA types will be eliminated to achieve rapid translation (**Fig. 1B**).

Next, we tested whether the phenotype of the *serCGA* deletion mutant is consistent with the extended model, which hypothesizes that the retained, non-essential tRNA type contributes uniquely to translational speed (**Text S1**). Accordingly, loss of the non-essential tRNA type is expected to decrease the maximum rate of translation. To test this expectation, the growth of the *serCGA* deletion mutant was examined on two types of media: (i) KB, a full medium that promotes rapid growth, and (ii) M9, a minimal medium that supports slower growth. If the prediction holds, the deletion mutant is expected to show a more dramatic reduction under conditions that support faster growth (*i.e.*, on KB).

Deletion of *serCGA* results in an immediately obvious growth defect on KB medium: Δ*serCGA*-1 produces visibly smaller colonies on KB agar (**Fig. 2A**), and shows a reduction in maximum growth rate in liquid KB (two sample *t*-test $p=1.18\text{x}10^{-6}$; **Fig. 2B**, **2C**). Further, both independent deletion mutants lose when grown in direct, 1:1 competition with the neutrally-marked SBW25-*lacZ* in liquid KB (one sample *t*-tests $p<0.0001$; **Fig. 2D**). In M9 the growth defect is much less pronounced; SBW25 and Δ*serCGA*-1 colonies are similar sizes on M9 agar (**Fig. 2E**), and no negative effects were detected in liquid M9 (**Fig. 2F, 2G**). However, direct 1:1 competitions between Δ*serCGA* and SBW25-*lacZ* indicate a slight negative effect of *serCGA* deletion (one sample *t*-tests $p=0.06882$ and 0.00908; **Fig. 2H**).

The results in this section demonstrate two points: (i) that *serCGA* encodes a non-essential, yet retained, tRNA in *P. fluorescens* SBW25, and (ii) that deletion of *serCGA* results in a decrease in maximum growth rate and fitness that is most pronounced under conditions that support rapid growth. These results are consistent with the predictions of the extended model, and thus with the hypothesis that tRNA-Ser(CGA) is retained because it translates its cognate codon at a significantly faster rate than that achieved by the wobble tRNA, tRNA-Ser(UGA).

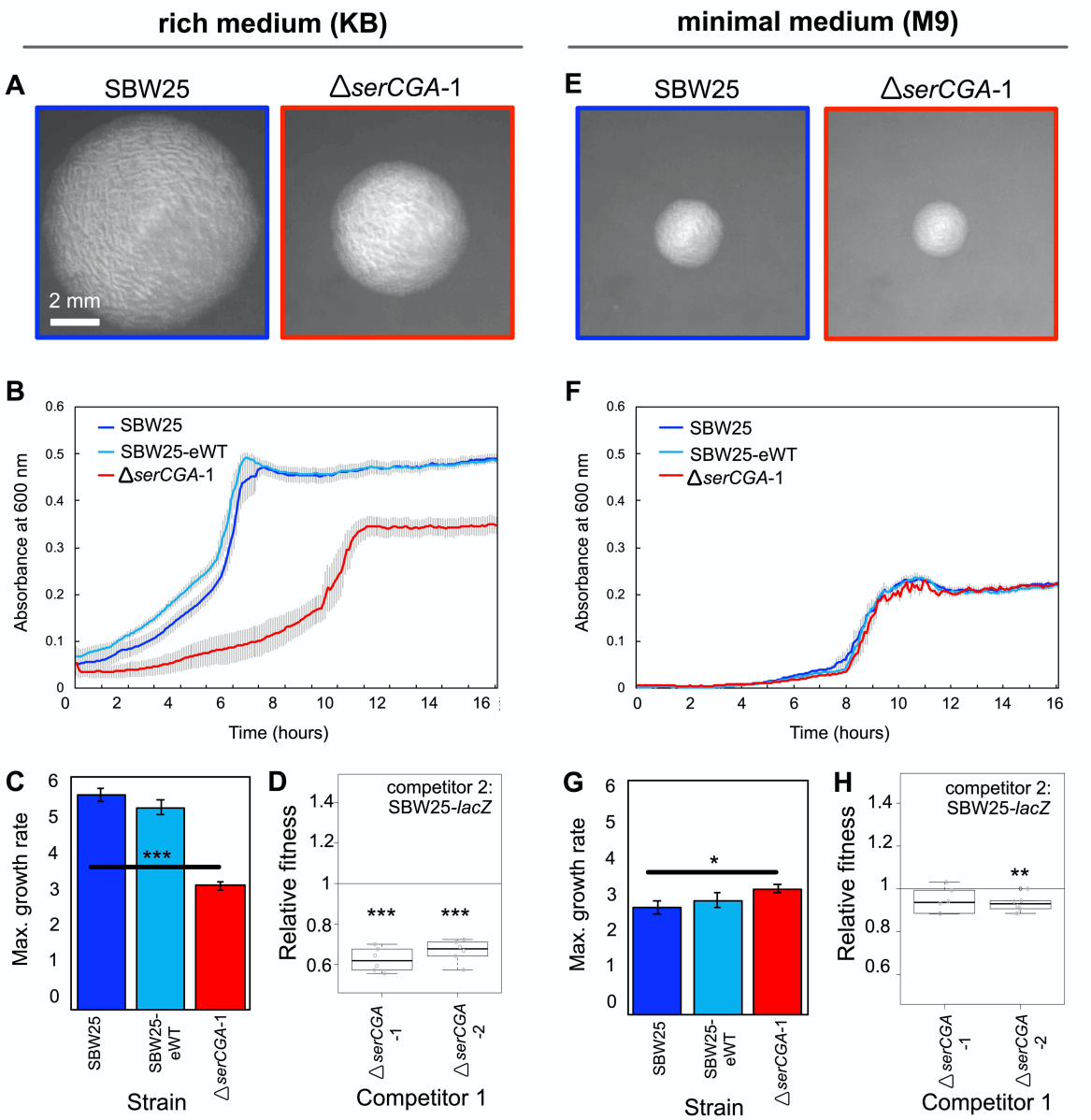

Figure 2. Deletion of *serCGA* leads to a growth defect in a rich medium. (**A**) 45-hour colonies on KB agar. (**B**) Growth (absorbance at 600 nm) in KB. Lines=mean of 5-7 replicates, error bars=one standard error. (**C**) Maximum growth rate (mOD min^{-1}) in KB. Bars=mean of six replicates, error bars=one standard error. (**D**) Relative fitness values from direct 1:1 competitions between competitor 1 (Δ*serCGA*-1 or Δ*serCGA*-2) and the neutrally marked wild type strain, SBW25-*lacZ* (six replicates *per* competition), in KB. Relative fitness of 1=no difference, <1=SBW25-*lacZ* wins, >1=competitor 1 wins. (**E**) 45-hour colonies on M9 agar, at same magnification and time as in panel A. (**F**) Growth in M9. Lines=mean of seven replicates, error bars=one standard error. (**G**) Maximum growth rate (mOD min^{-1}) in M9. Bars=mean of six replicates, error bars=one standard error. (**H**) Relative fitness values from direct 1:1 competitions between competitor 1 (Δ*serCGA*-1 or Δ*serCGA*-2) and SBW25-*lacZ* (six replicates *per* competition), in M9. Two (panel C, G) and one (panel D, H) sample *t*-tests ***$p<0.001$, **$p<0.01$, *$p<0.05$.

The growth defect is repeatedly and rapidly compensated during experimental evolution

If tRNA-Ser(CGA) is retained because it contributes uniquely to rapid translation of 5'-UCG-3', then logic dictates that maximum translational speed and growth of the *serCGA* deletion mutant can only be restored by regaining tRNA-Ser(CGA). Theoretically, this could be achieved through the acquisition of a mutation in the anticodon of another tRNA (anticodon switching). However, there are no obvious candidates for anticodon switching; only three other genes encode serine-carrying tRNAs, and each one is a single-copy gene encoding an essential tRNA (**Table S1**; GtRNAdb). This means that, according to the extended model, there is no immediately obvious evolutionary route by which the *serCGA* deletion mutant is expected to regain its previous growth profile.

In order investigate whether the growth defect exhibited by the *serCGA* deletion mutant can be genetically compensated, a serial transfer evolution experiment was performed. This experiment consisted of eight Lines: W1-W4 were control Lines each founded by a wild type strain (SBW25 or SBW25-eWT), while M1-M4 were founded by the deletion mutant (Δ*serCGA*-1 or Δ*serCGA*-2). Lines were started from single colonies, and grown in liquid KB for 15 days. Every 24 hours, 1 % of each population was transferred into fresh medium, and population samples frozen.

After 13 days (~90 generations), Lines M1-M4 showed visibly improved growth. Plating of Day 13 populations on KB agar revealed that many colonies from Lines M1-M4 were larger than those of the founding *serCGA* deletion mutant (**Fig. 3A**). Notably, Line M2 showed two phenotypically distinct types of large colonies: a phenotypically standard type, and an opaque type. The opaque type closely resembles the previously reported switcher phenotype (Beaumont et al., 2009; Gallie et al., 2019, 2015; Remigi et al., 2019). No obvious change was observed in the size of colonies derived from Day 13 of any of the four wild type Lines (**Fig. 3A**).

Five colonies were isolated from the Day 13 mutant lineages for further analysis. These included one standard-looking, large colony from each mutant Line (referred to hereafter as M1-L, M2-L, M3-L, M4-L), and a second large, opaque colony from Line M2 (M2-Lop). Two representative colonies were isolated from Day 13 of the wild type Lines (W1-L, W3-L). A growth curve of each of the seven isolates in liquid KB showed varying degrees of improved growth in the isolates from the mutant Lines (**Fig. 3B**), and each of the five mutant isolates outcompetes Δ*serCGA* in direct, 1:1 competition (one sample *t*-tests $p<0.01$; **Fig. 3C**). No fitness difference was detected when M1-L and M4-L were directly competed with SBW25-*lacZ* in liquid KB (one sample *t*-tests $p>0.05$; **Fig. 3C**).

The results in this section demonstrate that, despite the lack of obviously available evolutionary routes, the growth defect caused by the deletion of *serCGA* was repeatedly and rapidly restored to near wild-type levels by experimental evolution.

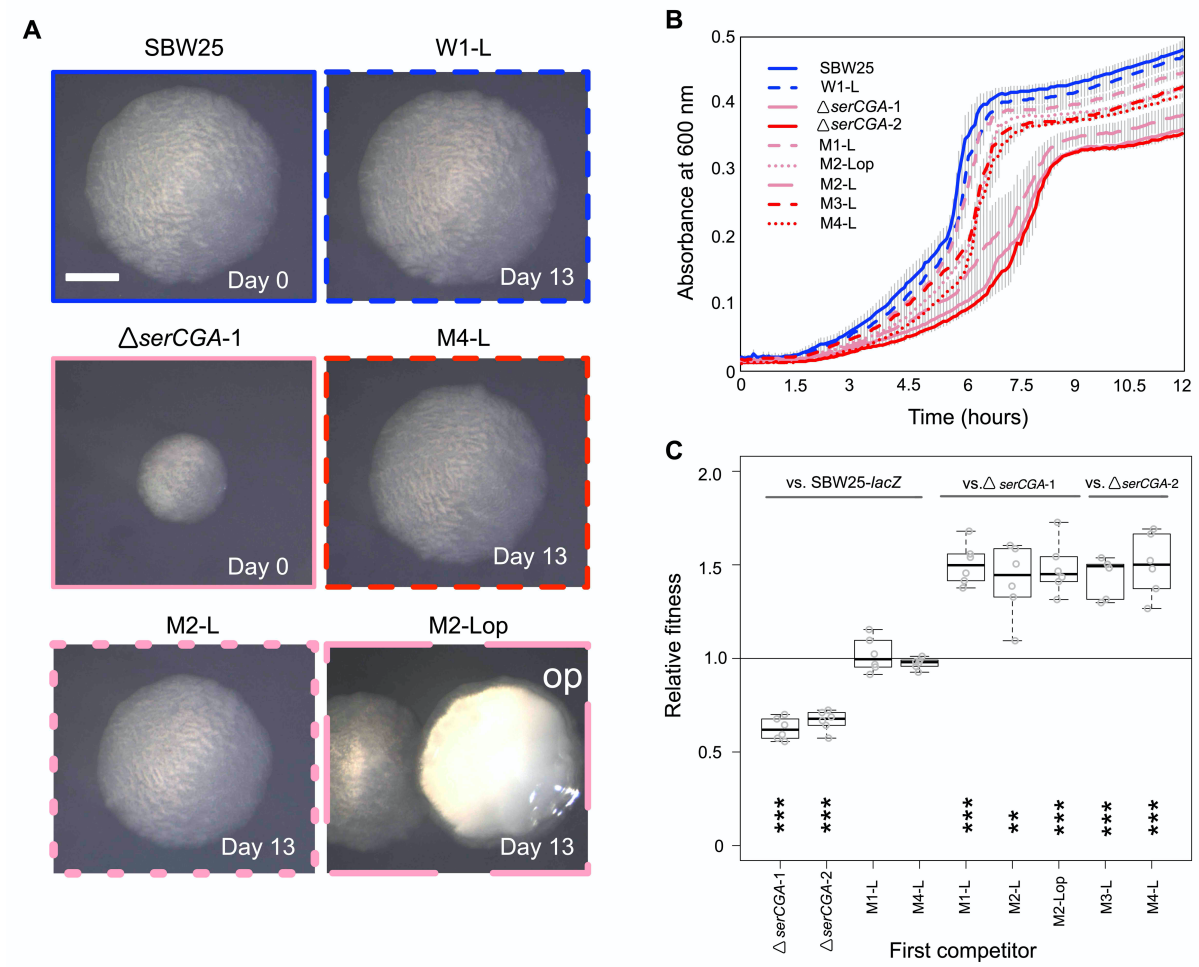

Figure 3. Loss of *serCGA* is repeatedly and rapidly compensated by experimental evolution. (**A**) Colony morphology of founder (Day 0) and evolved (Day 13) isolates on KB agar (30 h, 28˚C). Lineage M2 yielded two large colony morphotypes: standard (left), and opaque ("op", right). Image borders match lines in panel B. Scale bar=3 mm. (**B**) 12-hour growth curves in liquid KB for founder (Day 0, solid lines) and evolved (Day 13, dotted lines) isolates. Lines=mean of six replicates, error bars=1 standard error). (**C**). Box plots of the relative fitness of competitor 1 (*x*-axis) and competitor 2 (horizontal bars at top). Direct, 1:1 competitions were performed in liquid KB for 24 hours (28˚C, shaking). Relative fitness >1 means competitor 1 wins, <1 means competitor 2 wins. The first two competitions are also presented in **Fig. 2G**. One-sample *t*-tests ***p<0.001, **p<0.01, *p<0.05.

The genetic basis of compensation is large duplications spanning *serTGA*

To determine the genetics of Δ*serCGA* compensation, whole genome sequencing was performed for seven strains (W1-L, W3-L, M1-L, M2-L, M2-Lop, M3-L, M4-L). Each of the isolates from mutant linages was shown to contain a large, direct, tandem duplication at around 4.16 Mb in the SBW25 chromosome (**Fig. 4A, 4B**). No evidence was found of any such duplications in either of the isolates from wild type Lines (**Fig. S2**). A full list of candidate mutations is provided in **Table S2.**

A combination of bioinformatic analyses, PCR, and Sanger sequencing was used to determine the precise region of duplication in four out of five isolates (**Text S3**). The duplications range in size from 45 kb (in M1-L) to 290 kb (in M4-L), and occur between 4.05 Mb – 4.34 Mb of the chromosome

(**Table 1**, **Fig. 4A**). The precise location of the duplication in the fifth isolate (M3-L) could not be determined due to highly repetitive flanking DNA.

Strain	Junction side 1		Junction side 2		Size
	Base	Region	Base	Region	(bp)
M1-L	4,164,966	*murB*	4,119,923	intergenic repeat	45,043
M2-L	4,310,940 – 4,311,029	intergenic repeat	4,119,235 – 4,119,352	intergenic repeat	191,833
M2-Lop	4,224,306	*nuoL*	4,042,455	*pflu3649*	182,877
M3-L	~4,310,800	intergenic repeat	~4,119,100	intergenic repeat	~192,000
M4-L	4,339,314	intergenic	4,048,979	*pflu3655*	290,335

Table 1. Duplication junctions in five isolates from the mutant Lines reveal duplication fragments of 45-290 kb. Base positions refer to the SBW25 genome sequence (Silby et al., 2009). For details of the genes contained within each duplication segment see **Table S3**.

Next, we sought to identify the region(s) within the duplications responsible for the observed gain in fitness. Closer examination revealed a ~45 kb segment that is duplicated in all five strains (4,119,923 – 4,164,966). This segment is predicted to contain 45 genes, including 44 protein-coding genes, and one tRNA gene: *serTGA* (**Fig. 4A**, **Table S3**). Manual inspection of the *serTGA* sequences in each strain revealed no evidence of anticodon switching events (**Table S4**), meaning that each duplication strain contains a straightforward additional copy of *serTGA*. In addition, strain M4-L carries additional copies of four other tRNA genes (*argTCT*, *hisGTG, leuTAA, hisGTG*; **Table S3**). Notably, this experiment has revealed three different tRNA gene sets that each provide a similar level of fitness in KB: SBW25 encodes 67 tRNA genes of 39 tRNA types, four of five compensated strains carry 67 tRNA genes of 38 types, and the fifth compensated strain has 71 tRNA genes of 38 types.

Given that the duplicated *serTGA* gene encodes tRNA-Ser(UGA), the tRNA type that can theoretically perform the function of tRNA-Ser(CGA), it is a logical candidate for the cause of compensation. However, there are also 44 protein-coding genes in the shared duplication segment, any of which could underpin the compensatory effect. We therefore tested whether a plasmid-based increase in *serTGA* expression can compensate for *serCGA* loss. The *serCGA* and *serTGA* genes were amplified from SBW25, and ligated into the expression vector pSXn (giving pSXn-CGA, pSXn-TGA; **Text S2**). Each plasmid construct was inserted into SBW25, Δ*serCGA*-1, and Δ*serCGA*-2. Growth analyses demonstrate that expression of either *serCGA* or *serTGA* improves the growth of the *serCGA* deletion mutants in rich medium (**Fig. 4C**). In particular, maximum growth rate increases (**Fig. 4E**; two sample *t*-tests $p<0.05$). Contrastingly, expression of neither tRNA improved growth of SBW25 (**Fig. 4D**), with *serCGA* expression actually leading to a decrease in SBW25 maximum growth rate (**Fig. 4E;** one-sided two sample *t*-test $p=0.000158$). These results demonstrate that increased expression of *serTGA* can compensate, at least in part, for *serCGA* loss.

The *serTGA* duplications were isolated from all lineages within 90 generations, demonstrating that large-scale duplications occur repeatedly and rapidly. In order to investigate the rapidity with which the duplications arose, a PCR was performed to identify the first time point at which the emergent M1-L duplication junction could be amplified from Line M1 (see **Fig. 4B**). The Line M1 population samples that were frozen daily during the evolution experiment were revived, and used as templates for the PCR. The M1-L duplication junction was first visible by PCR on Day 3, and grew stronger as the experiment progressed (**Fig. 4F**). Similar PCRs for the remaining duplication strains detected the presence of duplications junctions on Day 2 (M2-Lop), Day 4 (M3-L) and Day 5 (M2-L, M4-L) (**Fig. S3**). These results demonstrate that although the duplication strains were isolated on Day 13, the duplication events occurred much earlier (within 5 days, ~35 generations).

Together, the results of this section show that the growth defect caused by deletion of the *serCGA* is repeatedly and rapidly compensated by large-scale, direct, tandem duplication events containing *serTGA*.

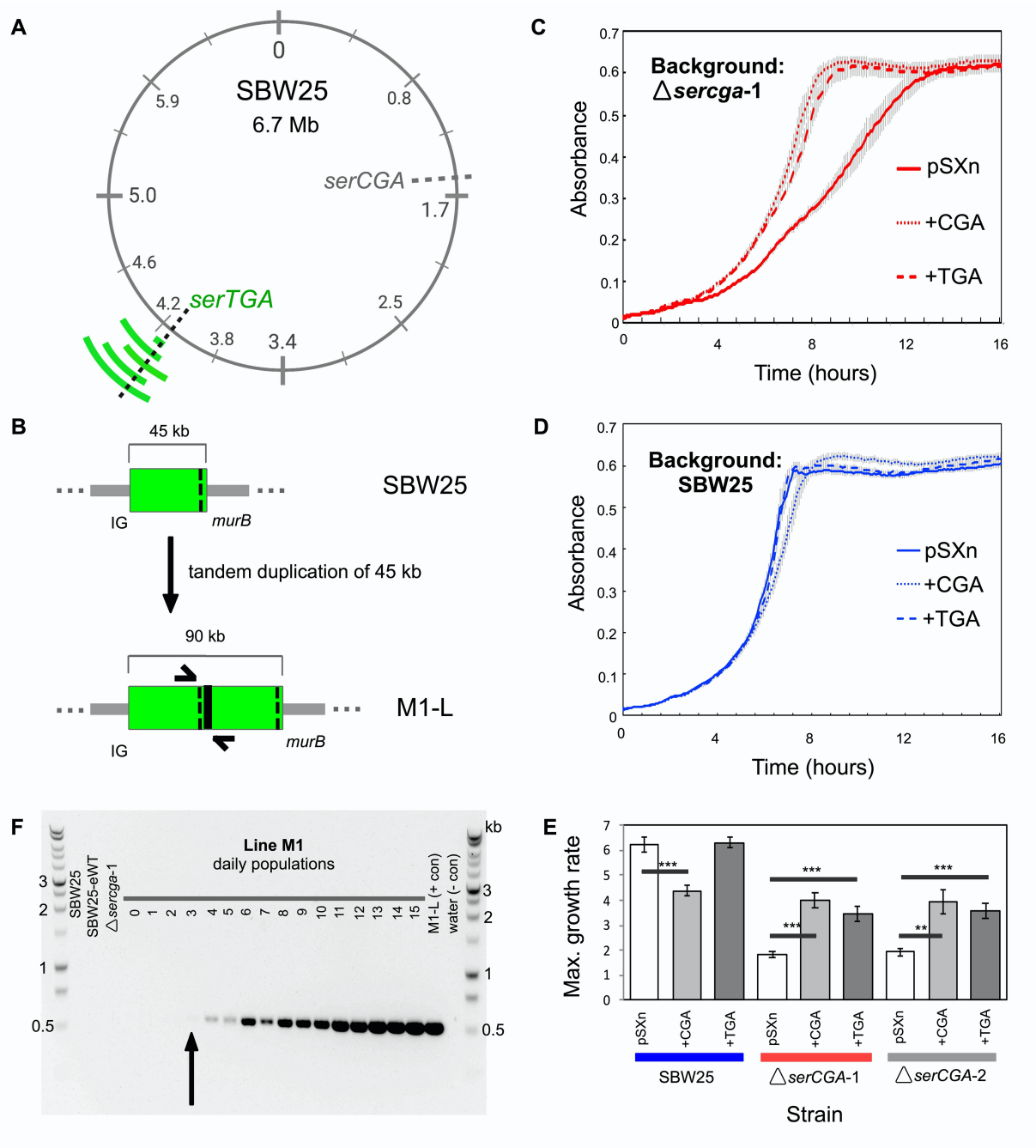

Figure 4. Direct, tandem duplications spanning *serTGA* compensate for the loss of *serCGA*. (A) Five isolates from the mutant Lines have unique, large, tandem duplications between 4.05 and 4.34 Mb of the SBW25 chromosome (green arcs; moving outwards: M1-L, M2-L, M2-Lop, M3-L, M4-L). The duplications contain a shared 45 kb region with *serTGA* (dotted black line). (**B**) Cartoon depiction of the duplication event in M1-L, resulting in two copies of the 45 kb fragment (green) and an emergent junction (thick black line). The junction can be amplified using primers ~250 bp to either side (black arrows). IG=intergenic, black dotted line=*serTGA*. (**C, D**) 12-hour growth curves in LB+Gm (20 µg ml^{-1}) for Δ*serCGA*-1 (panel C) and SBW25 (panel D) expressing *serCGA* or *serTGA* from the pSXn plasmid. Lines=mean of six replicates, error bars=1 standard error. (**E**) Maximum growth speed (mOD min^{-1}) of SBW25, Δ*serCGA*-1, Δ*serCGA*-2 carrying empty pSXn, pSXn-*serCGA* ("+CGA") and pSXn-*serTGA* ("+TGA"). Bars=mean of six replicates, error bars=1 standard error. Two sample *t*-tests ***$p<0.001$, **$p<0.01$, *$p<0.05$. (**F**) The M1-L junction was first amplified from Line M1 populations on Day 3 (black arrow). For the history of other junctions, see **Fig. S3**.

Duplication events increase the proportion of tRNA-Ser(UGA) in the mature tRNA pool

Next, we sought to quantify the effect of *serCGA* deletion and subsequent *serTGA* duplication on the mature tRNA pool. YAMAT-seq – an established method of deep-sequencing tRNA pools in human cells (Shigematsu et al., 2017) – was adapted for SBW25. Briefly, YAMAT-seq involves (i) ligation of hybrid RNA/DNA adapters to the universally conserved 5'-CCA-3' ends of mature tRNAs, (ii) reverse transcription and amplification of adapter-tRNA complexes, and (iii) high throughput sequencing.

YAMAT-seq was performed on three replicates of each of nine strains: wild type (SBW25), the two independent *serCGA* deletion mutants (Δ*serCGA*-1, Δ*serCGA*-2), and six isolates from Day 13 of the evolution experiment (M1-L, M2-L, M2-Lop, M3-L, M4-L, W1-L). Sequencing reads for each sample were aligned to a reference set of 42 unique tRNA sequences in the SBW25 genome (see **Text S4**); numbers and proportions of reads aligning to each reference sequence are provided in **Table S5**. Overall, 41 of the 42 reference tRNAs were detected. Cys-GCA-2, the secondary structure of which is predicted to deviate significantly from the conserved cloverleaf structure of tRNAs (Chan and Lowe, 2016), was not detected in any sample. This strongly indicates that Cys-GCA-2 does not contribute to translation in SBW25. Of the remaining 41 tRNAs, three gave consistently low reads (<0.01 % of reads *per* sample in **Table S5**: Glu-UUC, Ile2-CAU, Phe-GAA). The low read numbers may reflect low proportions of these tRNAs in the cell, or they may be the result of post-transcriptional modifications impeding the sequencing process (Machnicka et al., 2016; Shigematsu et al., 2017). In either case, as these three tRNA types were not central to our study, they were excluded from downstream analyses. This left a total of 38 reference tRNAs *per* sample .

DESeq2 was used to compare normalized expression levels of 38 tRNAs in pairs of strains (Love et al., 2014; **Table S6**). Firstly, the effect of deleting *serCGA* was investigated by comparing tRNA sequences from each of the two independent *serCGA* deletion strains with those in SBW25 (**Fig. 5A**; **Table S6**). The absence of tRNA-Ser(CGA) from both deletion mutants demonstrates that (i) tRNA-Ser(CGA) is encoded solely by the deleted *serCGA* gene, and (ii) tRNA-Ser(CGA) is indeed a non-essential tRNA type in SBW25. The *serCGA* deletion mutants also showed lower expression of tRNA-Thr(CGU) (DESeq2 adjusted $p<0.001$), a result that may reflect the close metabolic relationship between threonine and serine tRNAs (Liu et al., 2015; Sawers, 1998). No significant differences were detected between the independent deletion mutants (DESeq2 adjusted $p>0.1$).

Next, the effect of evolution on the *serCGA* deletion mutant was investigated by pairwise comparisons between each Day 13 isolate and its corresponding ancestor (**Fig. 5B**; **Table S6**). Importantly, no differences in tRNA pools were detected in the wild type control Line (W1-L versus SBW25; DESeq2 adjusted $p>0.1$). Contrastingly, a single consistent, statistically significant difference was observed across the five mutant lineage isolates: the expression of tRNA-Ser(UGA) was 2.06 to 2.60-fold higher than in the Δ*serCGA* ancestor (DESeq2 adjusted $p<0.0001$). A number of other tRNA types co-vary with the rise in tRNA-Ser(UGA) (**Fig. 5B**; **Table S6**). While none of these differences

are statistically significant in all strains, they are consistently in the same direction (increase or decrease). Thus, while the main effect of *serCGA* deletion and *serTGA* duplication has been the loss of tRNA-Ser(CGA) and duplication of tRNA-Ser(UGA), other more subtle effects are likely. Finally, pairwise comparisons between the duplication strains reveal some tRNA pool differences between the different duplication strains (**Table S6**; **Fig. S4**). However, no differences were detected between strain M4-L and any of the other duplication strains (DESeq2 adjusted p>0.3), indicating that the four additional tRNA genes duplicated in M4-L are not expressed. It is possible that the duplication junction in M4-L – which lies 112-bp upstream of the duplicated *argTCT-hisGTG-leuTAG-hisGTG* tRNA genes – truncates a promoter, leading to little or no expression of the duplicated copies. This result highlights that tRNA gene copy number does not always correlate with mature tRNA levels.

The results in this section demonstrate that deletion of *serCGA* eliminates tRNA-Ser(CGA) from the mature tRNA pool of SBW25. The increase in *serTGA* copy number caused by the compensatory large-scale duplication events is accompanied by an increase in tRNA-Ser(UGA).

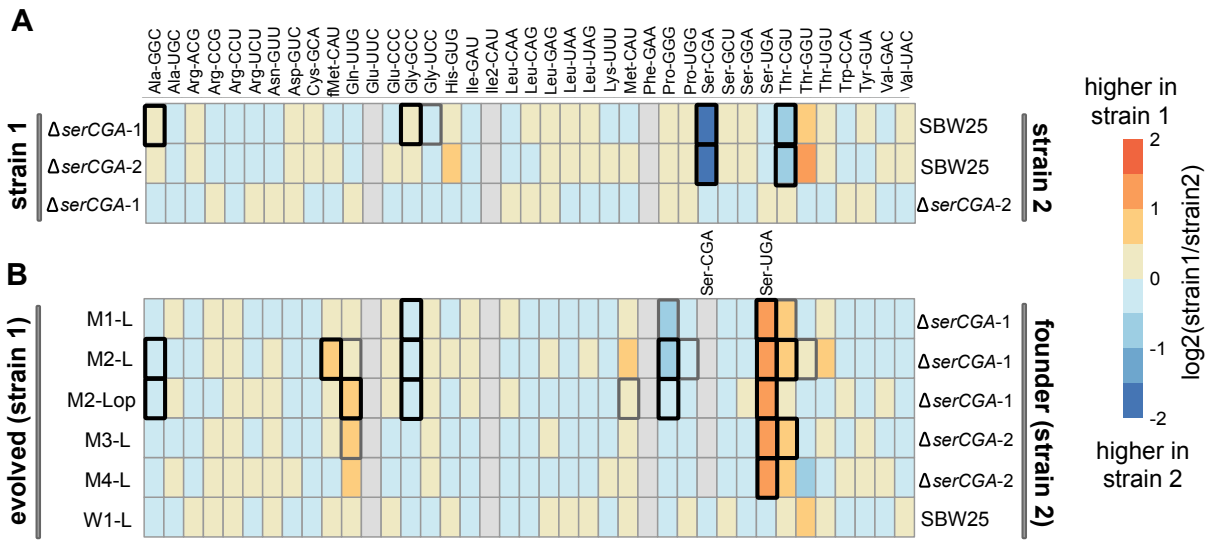

Figure 5. Heatmaps showing differences in the expression of tRNA types between strains. The log2-fold.change(strain1/strain2) difference in expression for 36 tRNA types (with 38 different primary sequences) was determined for pairs of strains using DESeq2 (Love et al., 2014). (**A**) tRNA expression levels in the *serCGA* deletion mutants compared with SBW25, demonstrating a consistently lower expression of tRNA-Ser(CGA) and tRNA-Thr(CGU) upon deletion of *serCGA* (with no statistically significant differences detected between the two deletion mutants; row 3). (**B**) tRNA-Ser(UGA) is higher in the mature tRNA pool in each of the five *serTGA* duplication isolates compared with the deletion mutant (with no significant differences in expression detected in the wild type control Line, row 6). Some tRNA types were removed from the DESeq2 analysis (filled grey boxes): Glu-UUC, Ile2-CAU, Phe-GAA, and – in some comparisons – Ser-CGA, consistently gave low read numbers for which DESeq2 results are not meaningful. Box borders represent statistical significance: thin grey=adjusted p>0.01, thick grey=0.01>adjusted p >0.001, black=adjusted p<0.001.

The model revisited: a higher tRNA-Ser(UGA) proportion leads to faster translation

To better understand how an increase in tRNA-Ser(UGA) can compensate for loss of translational speed in the *serCGA* deletion mutant, we turn our attention back to our models. The serine translation system from our experiments can be presented within the framework of the original model system (see **Fig. 1A**). That is, 5'-UCA-3' is codon *A*, 5'-UCG-3' is codon *B*, tRNA-Ser(UGA) is the wobble anticodon α, and tRNA-Ser(CGA) is the redundant β anticodon (**Fig. 6A**). Clearly, the mature tRNA pools of real organisms are far more complex, with many tRNA types present in varying proportions. Our YAMAT-seq data shows that the SBW25 tRNA pool is composed of 39 tRNA types in proportions varying from 0.14 to $1.7x10^{-6}$ (**Table S5**). The proportion of tRNA-Ser(CGA) is around 2.5-fold higher than that of tRNA-Ser(UGA) (0.015 and 0.0059, respectively; **Fig. 6B**). The dominant effect of *serCGA* deletion is elimination of tRNA-Ser(CGA), with the proportions of the other types, including tRNA-Ser(UGA) remaining relatively stable (**Fig. 6C**). Subsequent *serTGA* duplication then generates a two-fold increase in the proportion of tRNA-Ser(UGA) (**Fig. 6D**).

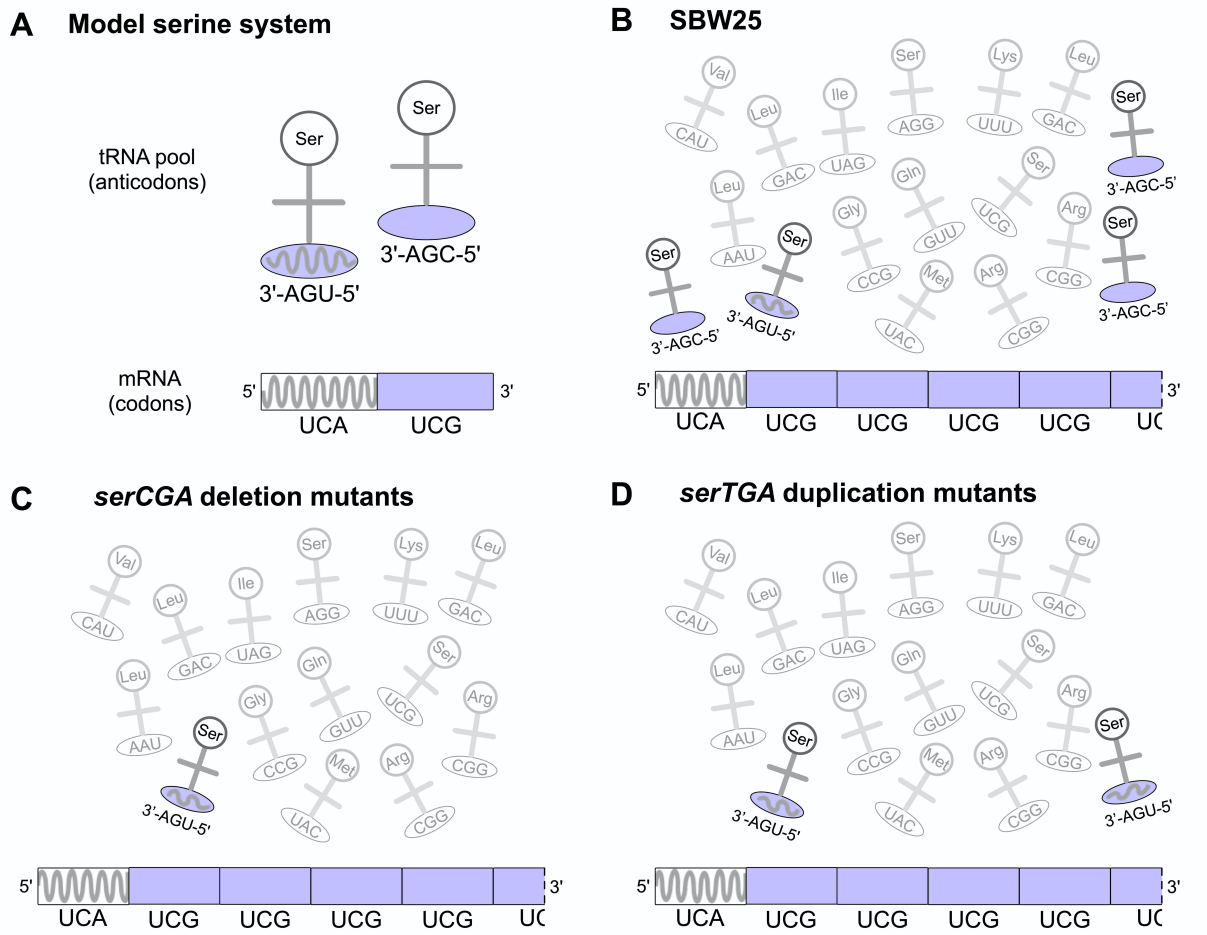

Figure 6. Illustration of the serine translation system in SBW25-derived strains. The system consists of two codons (5'-UCA-3', 5'-UCG-3') and two tRNA types (tRNA-Ser(UGA), tRNA-Ser(CGA)). As in the original model (see **Fig. 1A**), codon 5'-UCA-3' is translated by tRNA-Ser(UGA), and codon 5'-UCG-3' is translated by either tRNA type (panel **A**). The mature tRNA pools in SBW25 (panel **B**), the *serCGA* deletion mutants (panel **C**), and the *serTGA* duplication strains (panel **D**) are more complex, containing 38 or 39 tRNA types in variable proportions (**Table S5**). These three tRNA pools differ predominantly and consistently in their tRNA-Ser(UGA)

and tRNA-Ser(CGA) proportions, with the other 36 or 37 tRNA types remaining relatively stable (represented by the unchanging, light coloured background tRNAs in panels B-D (not to scale)). The mRNA in panels B-D consists of one 5'-UCA-3' and 4.5 5'-UCG-3' codons, to reflect the relative genome-wide use of these codons in SBW25.

Next we consider a defined mRNA molecule, which consists entirely of 5'-UCA-3' and 5'-UCG-3' codons. The proportion of each codon reflects its usage level in the SBW25 genome, with 5'-UCG-3' occurring 4.5 times more frequently than 5'-UCA-3' (Chan and Lowe, 2016; **Fig. 6B-D**). Using the framework outlined in **Fig. 6A**, the relative time taken to translate this defined mRNA can be estimated for each of SBW25, the *serCGA* deletion mutant, and the *serTGA* duplication strains. To this end, **Equation 4** – which is **Equation 2** modified to account for codon bias – was formulated. This equation calculates τ, the time taken to translate the mRNA, as a function of the inverse sum of the proportion (p) in the tRNA pool of tRNA types that can translate each codon.

$$\tau = 1\left(\frac{1}{p_{UGA}}\right) + 4.5\left(\frac{1}{p_{UGA} + p_{CGA}}\right) \qquad \textbf{Equation 4}$$

Substitution of the YAMAT-seq tRNA proportions obtained for strains SBW25, Δ*serCGA*-1, and M1-L into **Equation 4** gives translation times of 390, 1108 and 411 respectively (see also **Text S5**). Hence, this model makes two predictions: (i) that tRNA-Ser(CGA) elimination will slow translation of the UCA-UCG transcriptome dramatically (nearly 3 fold), due to a reduction in the chance of sampling a tRNA type that can translate codon 5'-UCG-3', and (ii) that increasing tRNA-Ser(UGA) is predicted to almost completely compensate for the lack of *serCGA*, by increasing the chance of sampling a tRNA that can translate 5'-UCG-3'. These predictions sit well with our experimental observations: (i) deletion of *serCGA* slows growth under rapid conditions, and (ii) *serTGA* duplication leads to (near) restoration of growth rate.

DISCUSSION

Given that tRNAs perform the same function – translating 61 codons into 20 amino acids – in all organisms, tRNA gene sets show a considerable degree of variability across the tree of life (Chan and Lowe, 2016; Marck and Grosjean, 2002). The evolutionary and molecular mechanisms by which different tRNA gene sets emerge have been of consistent, long-standing interest (periodically reviewed in Ibba and Söll, 2000; Ikemura, 1985; Rak et al., 2018). In recent years, a number of theoretical and computational studies have focused on various aspects of tRNA gene set evolution (*e.g.*, Lei and Burton, 2020; Rudorf, 2019; Schwark et al., 2020). In particular, computational analyses of large numbers of bacterial genomes have provided evidence of historical tRNA gene set evolution by two mechanisms: (i) the acquisition of tRNA genes through horizontal gene transfer (McDonald *et al*, 2015; Wald and Margalit, 2014), and (ii) anticodon switching (Wald and Margalit, 2014). The observation of an anticodon switch event has also been reported in laboratory populations of *S.*

cerevisiae (Yona et al., 2013). In this work, we have observed a third mechanism of tRNA gene set evolution: spontaneous tRNA gene duplication.

The *P. fluorescens* SBW25 tRNA gene set

During the evolution experiment, the growth defect caused by the elimination of tRNA-Ser(CGA) from *P. fluorescens* SBW25 was repeatedly compensated by spontaneous, large-scale, tandem duplication events that serve to double the copy number of the tRNA gene, *serTGA* (**Fig. 4A**, **Fig. S2**). Duplication of *serTGA* is accompanied by a two-fold increase of tRNA-Ser(UGA) in the mature tRNA pool (**Fig. 5B**). Presumably, tRNA-Ser(UGA) can perform the function of tRNA-Ser(CGA) – translation of codon 5'-UCG-3' – through wobble base pairing, according to the G-U wobble rule (Crick, 1966; **Fig. S1**). As a result, elevating the proportion of tRNA-Ser(UGA) in the mature tRNA pool is expected to speed up translation and growth by increasing the chance of sampling the required tRNA type at 5'-UCG-3' codons (**Fig. 6**; Varenne et al., 1984). In short, high-level expression of streamlined tRNA sets promotes rapid translation.

While the observation that increasing the proportion of tRNA-Ser(UGA) can compensate for tRNA-Ser(CGA) loss can be explained on the molecular level, it raises the question of why *P. fluorescens* SBW25 retains a copy of *serCGA*. One possible explanation is that tRNA-Ser(CGA) provides a unique contribution to the speed of translation that is too small to easily detect in the laboratory, but that matters over evolutionary time scales. A second possibility is that tRNA-Ser(CGA) contributes to rapid translation under different environmental conditions; temperature, acidity and ion concentration all alter the stability of RNA base pairings (Nikolova and Al-Hashimi, 2010; Serra et al., 2002). A third explanation could be that tRNA-Ser(CGA) contributes uniquely to an aspect of 5'-UCG-3' translation besides rapidity (*e.g.*, accuracy). Any one, or combination of, these options may contribute to the retention of *serCGA* in *P. fluorescens* SBW25.

The possibility that restoration of *serCGA* may further increase the fitness of the *serTGA* duplication strains could be investigated by continuing the evolution experiment past Day 13. This is because, in addition to increasing fitness, the duplication of *serTGA* has provided a possible route by which *serCGA* could be regained, and the original *P. fluorescens* SBW25 tRNA gene set restored. Specifically, one of the two *serTGA* copies could conceivably acquire a T→C transition at the first anticodon position, changing the gene from *serTGA* to *serCGA* (*i.e.*, an anticodon switch event). In order for such a mutation to spread through the population, it must encode a functional tRNA. This requires the new, hypothetical tRNA-Ser(CGA) to be recognized by seryl-tRNA ligase (SerS), the enzyme that adds serine to all types of serine-carrying tRNAs in the cell (for a review of tRNA ligases function, see Ibba and Söll, 2000). Recognition of serine-carrying tRNAs by SerS depends not on the tRNA sequence or the anticodon, but rather on the characteristic three-dimensional shape of serine-carrying tRNAs (Lenhard et al., 1999). Hence, even though *serTGA* and the original *serCGA* encode tRNAs with very different sequences (see **Fig. S1A**), it is feasible that the new, hypothetical *serCGA*

could form a functional tRNA (see **Fig. S1B**). Notably, none of the Day 13 mutant Line isolates shows evidence of anticodon switch events in either copy of *serTGA* (see **Table S4**); whether anticodon switching occurs across a longer evolutionary time scale remains to be seen.

The evolutionary fate of the large-scale, tandem duplications

Regardless of possible anticodon switch events, the fate of the large-scale, tandem duplications obtained during the evolution experiment remains a point of interest. Each duplication contains large amounts of excess DNA (~99.9 % of each fragment, amounting to between 45,950 and 290,200 bp). This surplus DNA encodes large numbers of genes (between 44 and 279), at least some of which show evidence of expression (*e.g.*, the CAP genes underpinning the colony opacity in M2-Lop; **Fig. 3A**). This means that the large-scale duplications affect cellular replication and expression profiles. Given this, one might expect further evolution to result in loss or shrinkage of the duplication fragments.

Similar duplications are a well-documented adaptive solution to various phenotypic challenges in phage, bacteria and yeast, and are known to typically be highly unstable (reviewed in Anderson and Roth, 1977; Reams and Roth, 2015). That is, such duplications are gained and lost at high rates: measurement of spontaneous duplication rates across the *Salmonella typhimurium* chromosome demonstrate that some loci are duplicated in as many as one in every 300 cells, and that these duplications are lost by 1-30% of cells in an overnight culture (in the absence of selection) (Anderson and Roth, 1981; Sonti and Roth, 1989). Due to their combined ease of gain and loss, it has been suggested that duplications may serve as fleeting solutions to transient phenotypic challenges (Sonti and Roth, 1989). However, while it is probable that the duplications observed in this experiment will be lost at high-frequency, spread of such revertants through the population is only expected if the extra *serTGA* copy is no longer required (*e.g.*, through acquisition of another compensatory mutation outside the duplication fragment, or an environmental change). A second possibility is that further evolution results in partial loss of the duplication fragments. This could occur through unequal recombination events between sets of direct repeats present to one side of *serTGA*. Such repeats may even be imperfect and/or short (*e.g.*, Lin et al., 1984; Shyamala et al., 1990). This process could conceivably reduce surplus DNA without loss of the second *serTGA* copy.

Conclusion

In this work, we have demonstrated an unexpectedly high degree and rate of flexibility in a bacterial tRNA gene set, with regards to both the tRNA types present and gene copy number. The elimination of one tRNA type from *P. fluorescens* SBW25 was readily counteracted by large-scale duplication events that increased the copy number of a gene encoding a second, compensatory tRNA type. Together with the sequence conservation of bacterial multi-copy tRNA genes, the ease and repeatability with which spontaneous duplications arose suggests that gene duplication may be an important source of new tRNA genes in bacteria.

MATERIALS AND METHODS

Strains growth conditions. A list of the strains used, and constriction details are in the Extended Materials and Methods (**Text S2**). Unless otherwise stated, *P. fluorescens* SBW25 cultures were grown in King's Medium B (KB; King et al., 1954) for ~16 hours at 28˚C with shaking. *E. coli* strains were grown in Lysogeny broth (LB) for 16-18 hours at 37˚C with shaking.

Growth curves. Strains were streaked from glycerol stocks on KB, M9, or LB+Gm (20 μg ml^{-1}) plates. After 48 hours incubation, six colonies *per* strain were grown in 200 μl of liquid KB, M9 or LB+Gm (20 μg ml^{-1}) in a 96-well plate. Two microliters of each culture were transferred to a fresh 198 μl of medium in a new 96-well plate, sealed with a breathable rayon film (VWR), and grown at 28˚C in a BioTek Epoch 2 plate reader. Absorbance at 600 nm (OD_{600}) of each well was measured at 5-minute intervals, with five seconds of 3 mm orbital shaking before each read. Medium control wells were used to standardize other wells. The mean absorbance, and standard error, of the replicates at every time point were used to draw the growth curves in **Figures 2, 3** and **4**. Maximum growth speed (V_{max}) and lag time were calculated using a sliding window of nine (LB+Gm) or 12 (KB/M9) data points (Gen5™ software from BioTek).

Fitness assays. Six single colonies of each competitor were grown independently in shaken KB. Eight competition tubes were inoculated with ~$5x10^6$ cells of each competitor, and incubated at 28˚C (shaking, 24 hours). Competitor frequencies were determined by plating on KB agar or LB+X-gal (60 μg ml^{-1}) agar at 0 and 72 hours. Competing genotypes were readily distinguished by their distinctive morphologies (on KB agar) or colour (neutrally-marked SBW25-*lacZ* forms blue colonies on LB+X-gal; Zhang and Rainey, 2007). Relative fitness was expressed as the ratio of Malthusian parameters (Lenski, 1991). Deviation of relative fitness from one was determined by one sample *t*-tests.

Evolution experiment. SBW25 (wild type), SBW25-eWT (engineering control), and the two independent tRNA-Ser(CGA) deletion mutants Δ*serCGA*-1 and Δ*serCGA*-2 were streaked from glycerol stocks onto KB agar and grown at 28˚C for 48 hours. Two colonies from every strain were used to found one Line each. This gave four independent wild type Lines (W1-W4), and four mutant Lines (M1-M4), plus a medium control. Each colony was inoculated into 4 ml KB in a 13 ml plastic tube, and incubated overnight at 28˚C (shaking). Each grown culture ("Day 0") was vortexed for 1 minute, and 100 μl used to inoculate 10 ml KB in a 50 ml Falcon tube (28˚C, shaking, 24 hours). Every 24 hours thereafter, 1 % of each culture was transferred to a fresh 10 ml KB in a 50 ml Falcon tube, and a sample of the population frozen at -80˚C. The experiment was continued until Day 15. Every few days, populations were dilution plated on KB agar to check for changes in colony size.

Genome sequencing. Seven isolates were purified and stored from Day 13 of the evolution experiment (W1-L, W3-L, M1-L, M2-L, M2-Lop, M3-L, M4-L). Genomic DNA was isolated from 0.5 ml overnight culture of each using a DNeasy Blood & Tissue Kit (Qiagen). DNA quality was checked by agarose gel electrophoresis. Whole genome sequencing was performed by the sequencing facility at the Max Planck Institute for Evolutionary Biology (Ploen, Germany). Paired-end, 150 bp reads were generated with a NextSeq 550 Output v2.5 kit (Illumina). Raw reads are available at NCBI sequence read archive (SRA accession number: PRJNA558233). A minimum of 4.5 million raw reads *per* strain were aligned to the SBW25 genome sequence (NCBI reference sequence NC_012660.1; Silby et al., 2009) using breseq (Barrick et al., 2014; Deatherage et al., 2015; Deatherage and Barrick, 2014) and Geneious (v11.1.4). A minimum mean coverage of 94.7 reads *per* base pair was obtained. A full list of mutation predictions is in **Table S2**.

Identification of duplication junctions. The duplication junctions in M1-L, M2-L, M2-Lop, M3-L and M4-L were identified using a combination of analysis of whole genome sequencing, and laboratory-based techniques (see **Text S3**). Briefly, the raw reads obtained from whole genome sequencing of each isolate were aligned to the SBW25 genome sequence (Silby et al., 2009) using breseq and Geneious (v11.1.4). Coverage analyses were performed in Geneious, and coverage plots generated in R (v3.6.0) (**Fig. S2**). Manual inspection of the Geneious alignment in coverage shift regions led to predicted junctions in all isolates except M3-L. Each predicted junction was checked by alignment to raw reads and previously unused sequences (Geneious). Finally, junction sequences were confirmed by PCR and Sanger sequencing.

Historical junction PCRs. Glycerol stock scrapings of the frozen daily populations from each of the four mutant Lines (Day 0 – Day 15) were grown in liquid KB. Washed cells were used as PCR templates, alongside positive and negative controls (see **Text S2**). The PCR products from each Line were run on a 1 % agarose gel against a 1 kb DNA ladder (Promega) at 100 volts for 90 minutes. Gels were stained with SYBR Safe (Life Technologies), and photographed under UV illumination. In order to better detect faint PCR products in the earlier days of the evolution experiment, the colours in each photograph were inverted using Preview (v11.0) (**Fig. 4F** and **Fig. S3**).

YAMAT-seq. Originally described for human cell lines, YAMAT-seq (Shigematsu et al., 2017) is adapted in this work for *P. fluorescens* SBW25. Three independent replicates of nine strains (*i.e.*, 27 samples) were grown to mid-exponential phase in 250 ml flasks containing 20 ml KB. Total RNA was isolated from 1.5 ml aliquots (TRIzol Max Bacterial RNA isolation kit; Life Technologies). For each sample, 5 μg of total RNA was subjected to tRNA deacylation treatment, and hybrid DNA/RNA Y-shaped adapters (Eurofins) were ligated to exposed 5'-CCA-3' tRNA ends using T4 RNA ligase 2 (New England Biolabs). Ligation products were reverse transcribed into cDNA (SuperScript III

reverse transcriptase; Thermo Fisher), and amplified by eleven rounds of PCR (Phusion; Thermo Fisher). One of 27 sample-specific indices (Illumina; **Table S5**) was added to each of the 27 reactions. The quality and quantity of each PCR product was checked using an Agilent DNA 7500 kit on a Bioanalyzer (Agilent), and samples combined in equimolar amounts into a single tube. The mixture was run on a 5 % polyacrylaminde gel (Bio-Rad Laboratories), and the group of bands ranging between 180 bp and 250 bp excised. DNA was extracted in deionized water overnight, and agarose removed by centrifugation through filter paper. The final product was sequenced at the Max Planck Institute for Evolutionary Biology (Plön, Germany). Single-end, 150 bp reads were generated with a NextSeq 550 Output v2.5 kit (Illumina). The raw YAMAT-seq data is available at NCBI Gene Expression Omnibus (GSE144791).

Analysis of YAMAT-seq data. The YAMAT-seq data was sorted into 27 samples by extracting exact matches to each unique, 6 bp long Illumina index. The resulting 27 raw read files, each containing at least 637,037 reads, were analysed using Geneious (version 11.1.4). First, reads were trimmed to 80-151 bp (99.99 % read retention). The trimmed reads were assembled to a set of 42 reference tRNA sequences from SBW25 (GtRNAdb, Chan & Lowe, 2016; **Text S4**). Assembly parameters: up to 10 % mismatches or gaps of <3 bp *per* read, up to 5 ambiguities *per* read, no assignment of reads that map equally well to more than one reference sequence, and no iterations. The numbers of raw reads assembled to each of the 42 reference tRNAs were converted to proportions of the total tRNA pool for each of the 27 samples. Mean proportions were calculated using three replicates of the nine strains (**Table S5**). DESeq2 (Love et al., 2014) was used in R (version 3.6.0) to detect tRNA expression differences between pairs of strains (**Table S6**). Unaligned reads for each strain were saved, *de novo* assembled into contigs, and manually examined. No notable contigs were found.

Statistical tests. One sample *t*-tests were performed on competition data (**Fig. 2D, 2H, 3C**). Two sample *t*-tests or Mann-Whitney-Wilcoxon tests were performed to detect differences in maximum growth speed (V_{max}) in growth curves (**Fig. 2B, 2F**, **4F**). DESeq2 adjusted *p*-values were used to detect differences in tRNA expression during YAMAT-seq (**Fig. 5**, **Fig. S4**, **Table S6)**. Analyses were performed in R (v3.6.0). Significance levels: *$0.05<p<0.001$, **$0.01<p<0.001$, ***$p<0.001$.

ACKNOWLEDGEMENTS

The work was supported by the Max Planck Society (all authors). The authors thank Gunda Dechow-Seligmann for technical assistance, and Dr Sven Künzel for assistance with troubleshooting the YAMAT-seq protocol. We thank Dr Frederic Bertels for discussions during development of the YAMAT-seq analysis, and the gift of the pSXn plasmid. We thank Dr Frederic Bertels, Dr Anuradha Mukherjee, Devika Bhave and Bilal Haider and helpful comments on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors hereby declare that there are no competing financial interests.

AUTHOR CONTRIBUTIONS

JG conceived the research. HJP and JG developed the mathematical model. GBA and JG performed the lab work and data analyses. JG wrote the manuscript, and all authors commented.

REFERENCES

Anderson P, Roth J. 1981. Spontaneous tandem genetic duplications in *Salmonella typhimurium* arise by unequal recombination between rRNA (*rrn*) cistrons. *Proceedings of the National Academy of Sciences of the United States of America* **78**:3113–3117. doi:10.1073/pnas.78.5.3113

Anderson RP, Roth JR. 1977. Tandem genetic duplications in phage and bacteria. *Annual Review of Microbiology* **31**:473–505. doi:10.1146/annurev.mi.31.100177.002353

Barrick JE, Colburn G, Deatherage DE, Traverse CC, Strand MD, Borges JJ, Knoester DB, Reba A, Meyer AG. 2014. Identifying structural variation in haploid microbial genomes from short-read resequencing data using *breseq*. *BMC Genomics* **15**:1039. doi:10.1186/1471-2164-15-1039

Beaumont HJE, Gallie J, Kost C, Ferguson GC, Rainey PB. 2009. Experimental evolution of bet hedging. *Nature* **462**:90–93. doi:10.1038/nature08504

Bulmer M. 1987. Coevolution of codon usage and transfer RNA abundance. *Nature*. doi:10.1038/325728a0

Chan PP, Lowe TM. 2016. GtRNAdb 2.0: An expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Research* **44**:D184-189. doi:10.1093/nar/gkv1309

Crick FHC. 1966. Codon-anticodon pairing: The wobble hypothesis. *Journal of Molecular Biology*. **19**:548-555. doi:10.1016/S0022-2836(66)80022-0

Deatherage DE, Barrick JE. 2014. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods in Molecular Biology* **1151**:165–188. doi:10.1007/978-1-4939-0554-6_12

Deatherage DE, Traverse CC, Wolf LN, Barrick JE. 2015. Detecting rare structural variation in evolving microbial populations from new sequence junctions using breseq. *Frontiers in Genetics* **5**:468. doi:10.3389/fgene.2014.00468

Gallie J, Bertels F, Remigi P, Ferguson GC, Nestmann S, Rainey PB. 2019. Repeated phenotypic evolution by different genetic routes in *Pseudomonas fluorescens* SBW25. *Molecular Biology and Evolution* **36**:1071–1085. doi:10.1093/molbev/msz040

Gallie J, Libby E, Bertels F, Remigi P, Jendresen CB, Ferguson GC, Desprat N, Buffing MF, Sauer U, Beaumont HJE, Martinussen J, Kilstrup M, Rainey PB. 2015. Bistability in a metabolic network underpins the *de novo* evolution of colony switching in *Pseudomonas fluorescen*s. *PLoS Biology* **13**:e1002109. doi:10.1371/journal.pbio.1002109.

Higgs PG, Ran W. 2008. Coevolution of codon usage and tRNA genes leads to alternative stable states of biased codon usage. *Molecular Biology and Evolution*. **25**:2279-2291. doi:10.1093/molbev/msn173

Ibba M, Söll D. 2000. Aminoacyl tRNA synthesis. *Annual Review of Biochemistry* **69**:617-650. doi:10.1146/annurev.biochem.69.1.617.

Ikemura T. 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. *Journal of Molecular Biology*. **151**:389-409. doi:10.1016/0022-2836(81)90003-6

Ikemura T. 1985. Codon usage and tRNA content in unicellular and multicellular organisms. *Molecular Biology and Evolution.* **2**:13-34.

King EO, Ward MK, Raney DE. 1954. Two simple media for the demonstration of pyocyanin and fluorescin. The *Journal of Laboratory and Clinical Medicine* **44**:301–307. doi:10.5555/uri:pii:002221435490222X

Lei L, Burton ZF. 2020. Evolution of life on Earth: tRNA, aminoacyl-tRNA synthetases and the genetic code. *Life*. **10**:21. doi:10.3390/life10030021

Lenhard B, Orellana O, Ibba M, Weygand-Durašević I. 1999. tRNA recognition and evolution of determinants in seryl-tRNA synthesis. *Nucleic Acids Research* **27**:721–729. doi:10.1093/nar/27.3.721

Lenski RE. 1991. Quantifying fitness and gene stability in microorganisms. *Biotechnology* **15**:173-192. doi:10.1016/b978-0-409-90199-3.50015-2

Lin RJ, Capage M, Hill CW. 1984. A repetitive DNA sequence, *rhs*, responsible for duplications within the *Escherichia coli* K-12 chromosome. *Journal of Molecular Biology* **177**:1–18. doi:10.1016/0022-2836(84)90054-8

Liu Z, Vargas-Rodriguez O, Goto Y, Novoa EM, Ribas De Pouplana L, Suga H, Musier-Forsyth K. 2015. Homologous trans-editing factors with broad tRNA specificity prevent mistranslation caused by serine/threonine misactivation. *Proceedings of the National Academy of Sciences of the United States of America* **112**:6027–6032. doi:10.1073/pnas.1423664112

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**:550. doi:10.1186/s13059-014-0550-8

Machnicka MA, Dunin-Horkawicz S, de Crécy-Lagard V, Bujnicki JM. 2016. tRNAmodpred: A computational method for predicting posttranscriptional modifications in tRNAs. *Methods* **107**:34–41. doi:10.1016/j.ymeth.2016.03.013

Marck C, Grosjean H. 2002. tRNomics: Analysis of tRNA genes from 50 genomes of eukarya, archaea, and bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA* **8**:1189-1232. doi:10.1017/S1355838202022021

McDonald MJ, Chou CH, Swamy KBS, Huang H Da, Leu JY. 2015. The evolutionary dynamics of tRNA-gene copy number and codon-use in *E. coli*. *BMC Evolutionary Biology* **15**:163. doi:10.1186/s12862-015-0441-y

Nelson DL, Cox MM. 2013. Protein metabolism: wobble allows some tRNAs to recognize more than one codon. In: Lehninger principles of biochemistry. New York: W.H. Freeman & Co.

Niolova EN, Al-Hashimi HM. 2010. Thermodynamics of RNA melting, one base pair at a time. *RNA* **16**:1687-1691. doi:10.1261/rna/2235010.

Rak R, Dahan O, Pilpel Y. 2018. Repertoires of tRNAs: The couplers of genomics and proteomics. *Annual Review of Cell and Developmental Biology* **34**:239–264. doi:10.1146/annurev-cellbio-100617-062754

Reams AB, Roth JR. 2015. Mechanisms of gene duplication and amplification. *Cold Spring Harbor Perspectives in Biology* **7**:a016592. doi:10.1101/cshperspect.a016592

Remigi P, Ferguson GC, McConnell E, De Monte S, Rogers DW, Rainey PB. 2019. Ribosome provisioning

activates a bistable switch coupled to fast exit from stationary phase. *Molecular Biology and Evolution* **36**:1056–1070. doi:10.1093/molbev/msz041

Rocha EPC. 2004. Codon usage bias from tRNA's point of view: Redundancy, specialization, and efficient decoding for translation optimization. *Genome Research* **14**:2279-2286. doi:10.1101/gr.2896904

Rudorf, S. 2019. Efficiency of protein synthesis inhibition depends on tRNA and codon compositions. *PLoS Computational Biology* **15**:e1006979. doi:10.1371/journal.pcbi.1006979

Serra MJ, Baird JD, Dale T, Fey BL, Retatagos K, Westhof E. 2002. Effects of magnesium ions on the stabilization of RNA oligomers of defined structures. *RNA* **8**:307-323. doi:10.1017/s1355838202024226.

Sawers G. 1998. The anaerobic degradation of L-serine and L-threonine in enterobacteria: Networks of pathways and regulatory signals. *Archives of Microbiology* **171**:1–5. doi:10.1007/s002030050670

Schwark DG, Schmitt AM, Biddle W, Fisk JD. 2020. The influence of competing tRNA abundance on translation: quantifying the efficinecy of sense codon reassignment at rarely used codons. *ChemBioChem*. Accepted author manuscript. doi:10.1002/cbic.202000052

Shigematsu M, Honda S, Loher P, Telonis AG, Rigoutsos I, Kirino Y. 2017. YAMAT-seq: An efficient method for high-throughput sequencing of mature transfer RNAs. *Nucleic Acids Research* **45**:e70. doi:10.1093/nar/gkx005

Shyamala V, Schneider E, Ames GF. 1990. Tandem chromosomal duplications: role of REP sequences in the recombination event at the join-point. The *EMBO Journal* **9**:939–946. doi:10.1002/j.1460-2075.1990.tb08192.x

Silby MW, Cerdeño-Tárraga AM, Vernikos GS, Giddens SR, Jackson RW, Preston GM, Zhang XX, Moon CD, Gehrig SM, Godfrey SAC, Knight CG, Malone JG, Robinson Z, Spiers AJ, Harris S, Challis GL, Yaxley AM, Harris D, Seeger K, Murphy L, Rutter S, Squares R, Quail MA, Saunders E, Mavromatis K, Brettin TS, Bentley SD, Hothersall J, Stephens E, Thomas CM, Parkhill J, Levy SB, Rainey PB, Thomson NR. 2009. Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens.* *Genome Biology* **10**:R51. doi:10.1186/gb-2009-10-5-r51

Sonti R V., Roth JR. 1989. Role of gene duplications in the adaptation of *Salmonella typhimurium* to growth on limiting carbon sources. *Genetics* **123**:19–28

Varenne S, Buc J, Lloubes R, Lazdunski C. 1984. Translation is a non-uniform process: Effect of tRNA availability on the rate of elongation of nascent polypeptide chains. *Journal of Molecular Biology* **180**:549–576. doi:10.1016/0022-2836(84)90027-5

Wald N, Margalit H. 2014. Auxiliary tRNAs: Large-scale analysis of tRNA genes reveals patterns of tRNA repertoire dynamics. *Nucleic Acids Research* **42**:6552-6566. doi:10.1093/nar/gku245

Yona AH, Bloom-Ackermann Z, Frumkin I, Hanson-Smith V, Charpak-Amikam Y, Feng Q, Boeke JD, Dahan O, Pilpel Y. 2013. tRNA genes rapidly change in evolution to meet novel translational demands. *eLife* **2**:e01339. doi:10.7554/eLife.01339

Zhang XX, Rainey PB. 2007. Construction and validation of a neutrally-marked strain of *Pseudomonas fluorescens* SBW25. *Journal of Microbiological Methods* **71**:78–81. doi:10.1016/j.mimet.2007.07.001