The birth of a bacterial tRNA gene

Gökçe B. Ayan, Hye Jin Park, Jenna Gallie*

Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, 24306 Plön, Germany
*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 types and numbers of tRNA genes that they carry. The evolutionary mechanisms behind the emergence of different tRNA gene sets remain largely obscure. Here, we report the rapid and repeated evolution of a tRNA gene set in laboratory populations of the bacterium *Pseudomonas fluorescens* SBW25. Deletion of the non-essential, single-copy tRNA gene *serCGA* from SBW25 results in a sub-optimal tRNA gene set. Compensation occurs within 35 generations *via* large (45-290 kb), direct, tandem duplications in the chromosome. Each duplication contains a *serTGA* gene, and is accompanied by a two-fold increase in tRNA-Ser(UGA) in the mature tRNA pool. This work demonstrates that the composition of tRNA gene sets – and mature tRNA pools – can readily evolve by duplication of existing tRNA genes, a phenomenon that could explain the presence of multiple identical tRNA gene copies within genomes.

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, 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).

The recognition of multiple codons by one 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**). A prominent example of wobble base pairing is the tolerance of G:U base pairs. Using only the G:U wobble rule, the number of tRNA types required to translate the entire set of 61 codons is reduced to 33 (32 elongators and one initiator). A second commonly utilized wobble rule involves the post-transcriptional modification of an anticodon first-position adenosine to hypoxanthine ("I"). Hypoxanthine base pairs with U, A or C, allowing the modified anticodon to translate three codons (Crick, 1966; Machnicka et al., 2016).

The G:U and hypoxanthine wobble rules are commonly exploited by bacteria, reducing the number of tRNA types they encode (Chan and Lowe, 2016; Marck and Grosjean, 2002). Bacterial tRNA gene sets typically carry a highly similar core set of 32 to 33 tRNA types that can – according to a combination of the G:U and hypoxanthine wobble rules – translate all 61 codons. In addition, bacteria contain a more variable set of non-essential tRNA types (Wald and Margalit, 2014). Bacterial tRNA gene sets also differ in the total number of tRNA genes, with many tRNA types being coded for by multiple, and usually identical, gene copies (Chan and Lowe, 2016; Marck and Grosjean, 2002).

The evolution of variation in bacterial tRNA gene sets is influenced by a number of factors, one of which is translational demand. A comparative study of tRNA gene sets across 102 bacterial species demonstrates that shorter division times correlate with higher numbers of tRNA genes of fewer tRNA types (Rocha, 2004). Presumably, this tRNA gene set promotes translational efficiency by reducing the amount of time spent sampling different tRNAs during translation (the limiting step of elongation; Varenne et al., 1984). Other factors that co-vary with tRNA gene content are codon use (Bulmer, 1987; Higgs and Ran, 2008; Ikemura, 1981), GC content (Wald and Margalit, 2014), and genome size (Rak *et al*., 2018).

The molecular mechanisms by which bacterial tRNA gene sets evolve remain poorly understood. Logically, new tRNA genes could be acquired from external sources (by horizontal gene transfer), or from existing genomic DNA (by spontaneous duplication events, and/or anticodon mutations in multi-copy tRNA genes). However, thus far, most evidence for tRNA gene set evolution is indirect. For example, 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). Empirical evidence of the emergence of a tRNA type was recently reported in laboratory populations of *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 that switched the anticodon to that of the deleted tRNA (Yona et al., 2013).

Here we report the evolution of a bacterial tRNA gene set through spontaneous, large-scale duplication events. Initially, we construct a sub-optimal tRNA gene set in *Pseudomonas fluorescens* SBW25 by deleting a single-copy, non-essential tRNA gene (*serCGA*). Next, we perform a serial transfer experiment to investigate the evolutionary mechanisms by which compensation occurs. We determine the effect of evolution on the tRNA gene sets (using whole genome sequencing) and mature tRNA pools (using an adaptation of YAMAT-seq; Shigematsu et al., 2017). We find that the tRNA gene set of *P. fluorescens* SBW25 is unexpectedly flexible; the loss of one tRNA type (*serCGA*) can be readily compensated by spontaneous duplication events that increase the copy number of a second, functionally related, tRNA type (*serTGA*).

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).

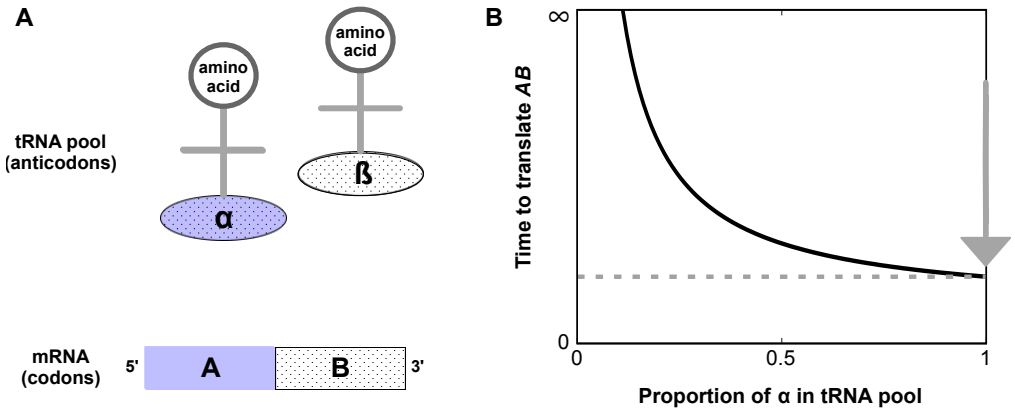

Figure 1. A model system in which rapid translation favours elimination of non-essential tRNA types. (**A**) The system consists of two consecutive mRNA codons (*A*, *B*), and a pool of equal proportions of two tRNA types (α, β). *Alpha* (α) can translate *A* (*via* cognate pairing) and *B* (*via* wobble pairing), while β can translate only *B* (*via* cognate pairing). The model assumes that all codon-anticodon pairs that can translate do so at equal rates. (**B**) Graph showing 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). See also **Text S1** for more details.

The speed at which *AB* is translated can be estimated by τ, the inverse sum of the time taken to translate *A* and *B*. The speed at which each codon is translated in the model is affected by two factors: the rate (λ) of translation by each matching anticodon, and the proportion of matching anticodons in the tRNA pool (p; **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 speed is determined solely by the proportion of each anticodon in the tRNA pool (p; **Equation 2**). Notably, tRNA proportions are also an important factor biologically; tRNA sampling is the rate-limiting step in the elongation phase of protein synthesis (Varenne et al., 1984). When there are equal proportions of α and β in the pool (*i.e.*, $p_\alpha = p_\beta$), *A* is translated at half the speed of *B* (**Equation 3**). As the proportion of α increases in the pool, the time taken to translate *AB* decreases, and translation fastest when tRNA type β is eliminated (**Fig. 1B**).

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

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

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

The model illustrates an important point: unnecessary options reduce translational speed. It predicts that, in cases where rapid translation is important, non-essential tRNA types will be eliminated. This result is in agreement with previous work showing that bacteria with faster growth rates – and accordingly, faster translation rates – tend to carry fewer tRNA types (Rocha, 2004).

The model can be extended to investigate the effect of other parameters on translation rate and tRNA retention. For example, varying translation rates of each functional anticodon-codon pair (*i.e.*, $\lambda_{\alpha A} \neq \lambda_{\alpha B} \neq \lambda_{\beta B}$), leads to the prediction that non-essential tRNA types are retained in cases where wobble pairing is significantly less efficient than the cognate pairing (*i.e.*, when $\lambda_{\alpha B}$ is less than $\lambda_{\beta B}$; **Text S1**).

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

The predictions of the model were tested using the fast-growing bacterium *Pseudomonas fluorescens* SBW25, which contains 67 tRNA genes of 39 types (GtRNAdb; Chan and Lowe, 2016). According to the G:U and hypoxanthine wobble rules (see introduction and Crick, 1966), six of the 39 types are non-essential (**Table S1**). The cognate codon recognized by each of these non-essential types can, at least theoretically, be translated through wobble pairing by another tRNA type present in the genome. According to our model, the retention of the six non-essential tRNA types indicates that they contribute positively to translation rate.

A retained, non-essential tRNA type was chosen for further investigation: tRNA-Ser(CGA), encoded by the single copy gene *serCGA.* tRNA-Ser(CGA) translates codon 5'-UCG-3', which can also be translated by the essential tRNA type, tRNA-Ser(UGA) (encoded by the single copy gene *serTGA* in SBW25; **Fig. S1**).

Deletion of *serCGA* limits rapid growth

The entire 90-bp, single copy *serCGA* gene was removed from the *P. fluorescens* SBW25 chromosome by allelic exchange (**Text S2**). This process was performed two independent times, giving rise to biological replicate strains Δ*serCGA*-1 and Δ*serCGA*-2. A third round of the exchange process yielded an engineering control strain, SBW25-eWT. This strain has passed through all steps of the engineering process, but retains the wild type *serCGA* gene.

Deletion of *serCGA* results in an immediately obvious growth defect: Δ*serCGA* produces smaller colonies on KB agar (**Fig. 2A**), and shows a reduction in maximum growth rate in liquid KB (**Fig. 2B**, **2C**). The deletion mutant loses in 1:1 competition with SBW25-*lacZ* (SBW25 carrying a neutral marker) in liquid KB (**Fig. 2D**). The observed growth defect was less pronounced in minimal

medium (M9, which supports slower growth and, presumably, requires slower translation); SBW25 and Δ*serCGA* colonies are similar on M9 agar (**Fig. 2E**), and no negative effects were detected on growth in liquid M9 (**Fig. 2F, 2G**). However, a direct competition between Δ*serCGA* and SBW25-*lacZ* in liquid M9 shows some degree of negative effect of *serCGA* deletion (**Fig. 2H**).

The results in this section are in line with the predictions of the model. Firstly, *serCGA* can be deleted, demonstrating that tRNA-Ser(CGA) is non-essential. Secondly, the observation that the reduction in growth caused by deletion of *serCGA* is most pronounced under conditions that support rapid growth, is consistent with tRNA-Ser(CGA) contributing to rapid translation.

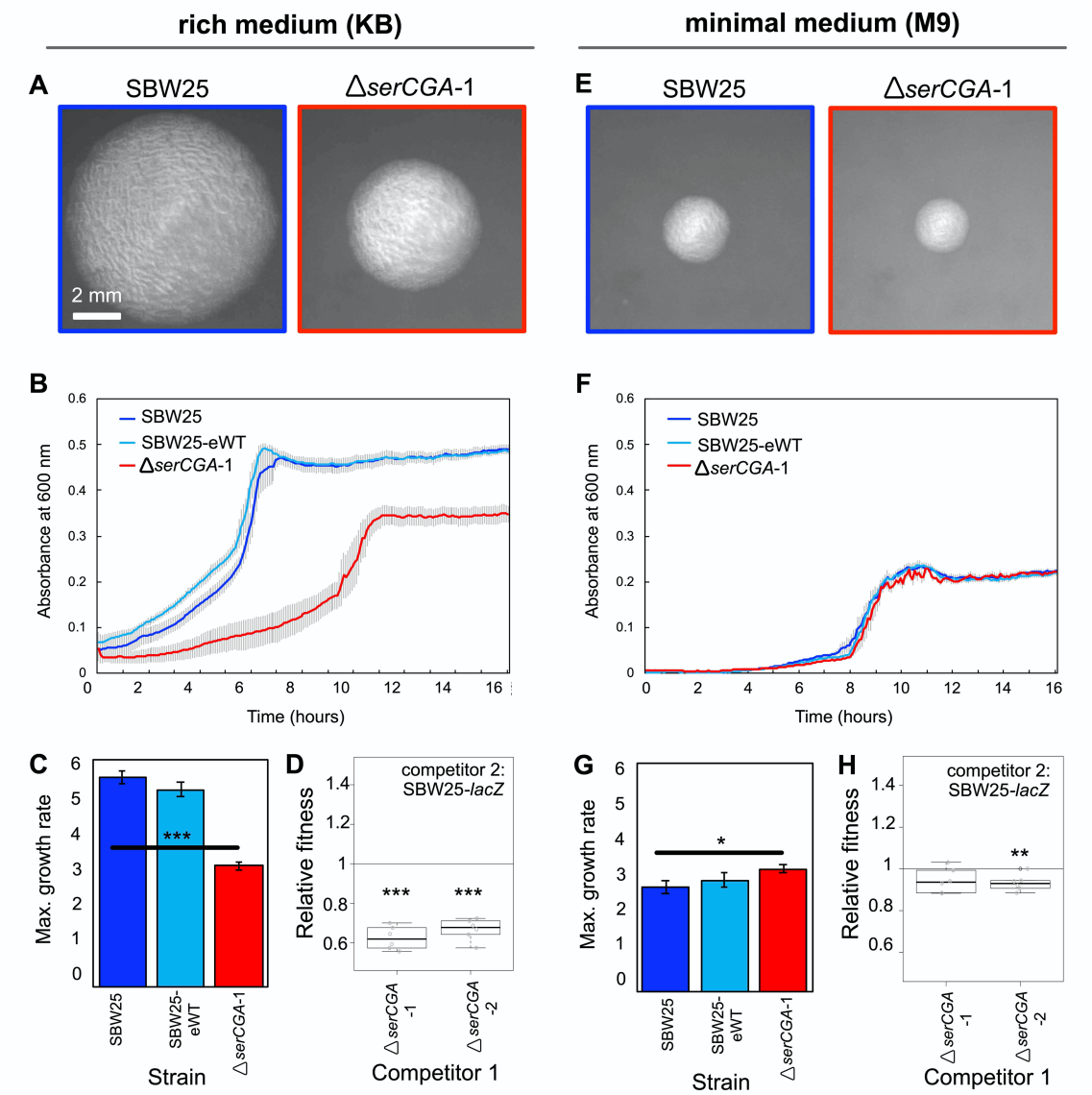

Figure 2. Deletion of *serCGA* leads to a growth defect in rich medium. (**A**) 45-hour colonies on KB agar. (**B**) Growth (absorbance at 600 nm) in KB. Lines=mean of five, six or seven replicates, error bars=one standard error. (**C**) Maximum growth rate (mOD min^{-1}) in KB. Bars=mean of six replicates, error bars=one standard error. (**D**) Relative fitness values from 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 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 1:1 competitions between competitor 1 (Δ*serCGA*-1 or Δ*serCGA*-2) and SBW25-*lacZ* (six replicates *per* competition), in M9. Two sample *t*-tests ***$p<0.001$, **$p<0.01$, *$p<0.05$.

The growth defect is repeatedly and rapidly compensated during experimental evolution

Next, a serial transfer evolution experiment was performed to investigate whether the growth defect observed on deleting *serCGA* can be compensated genetically. The evolution experiment consisted of eight Lines: W1-W4 were each founded by a wild type strain (SBW25 or SBW25-eWT), while M1-M4 were founded by the growth-impaired *serCGA* deletion mutant (Δ*serCGA*-1 or Δ*serCGA*-2). A medium control Line was also included. 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 (*i.e.*, had a turbidity approaching that of the wild type Lines after eight hours' incubation). This observation was corroborated by plating Day 13 populations on KB agar; Lines M1-M4 gave rise to colonies that were larger than those of the founding tRNA deletion strain (**Fig. 3A**). Notably, Line M2 carries two phenotypically distinct types of large colonies: a standard type, and an opaque type. On closer inspection, the opaque morphotype closely resembles the previously reported switcher phenotype (Beaumont et al., 2009; Gallie et al., 2019, 2015; Remigi et al., 2019). That is, Line M2 opaque colonies consist of cells that produce large capsules, and, when streaked onto KB agar, repeatedly produce a mixture of standard and opaque morphotypes. There was no obvious change in the morphology of colonies in any of the four wild type Lines (**Fig. 3A**).

Five colonies were isolated from Day 13 of the evolved mutant Lines for further analysis. These include one large colony from each mutant Line (M1-L, M2-L, M3-L, M4-L), and a second large, opaque colony from Line M2 (M2-Lop). Two representative colonies from the wild type Lines (W1-L, W3-L) were also isolated. A growth curve for each of the seven isolates in liquid KB showed varying degrees of improved growth in the isolates from the mutant Lines (**Fig. 3B**). Additionally, each of the five mutant isolates outcompetes the tRNA deletion strain (founder) in direct competitions (**Fig. 3C**; one sample *t*-tests $p<0.01$).

The results in this section demonstrate that the growth defect caused by the deletion of *serCGA* can be repeatedly and rapidly compensated by evolution.

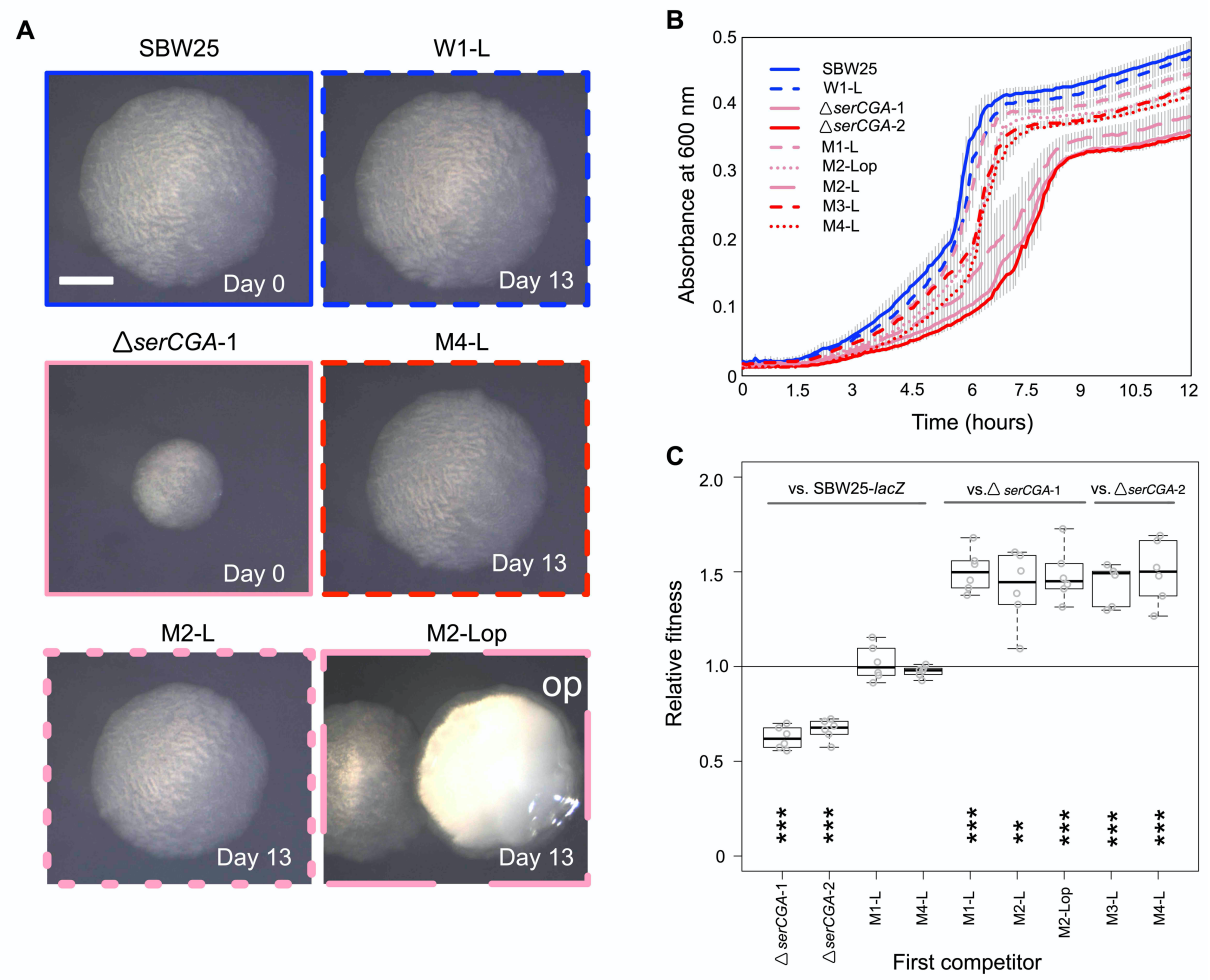

Figure 3. Loss of *serCGA* is repeatedly and rapidly compensated by experimental evolution. (**A**) Colony morphology of founder (Day 0) and evolved (Day 13) isolates on KB agar (30 h, 28˚C). Line M2 shows two distinct large colony morphotypes: standard (left), and opaque ("op", right). Picture borders match the lines in panel B. 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). 1:1 competition assays were performed in liquid KB for 24 hours (28˚C, shaking). Relative fitness >1 means competitor 1 wins, <1 means competitor 2 wins. The results 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*

Whole genome sequencing was used to determine the genetic basis of improvement in the five strains isolated from the mutant Lines of the evolution experiment. Genomic DNA was extracted from seven strains: W1-L, W3-L, M1-L, M2-L, M2-Lop, M3-L, M4-L, and used for whole genome sequencing. Analyses revealed that each of the isolates from the mutant Lines contains a large, tandem duplication at around 4.16 Mb in the SBW25 chromosome (**Table 1**, **Fig. 4A**). No evidence was found of any such duplication events in either of the isolates from the wild type Lines (**Fig. S2**). A full list of the mutations identified is provided in **Table S2.**

Strain	Junction side 1		Junction side 2		Size
	Base	Region	Base	Region	(bp)
M1-L	4,119,923	intergenic repeat	4,164,966	*murB*	45,043
M2-L	4,119,235 – 4,119,352	intergenic repeat	4,310,940 – 4,311,029	intergenic repeat	191,833
M2-Lop	4,042,455	*pflu3649*	4,224,306	*nuoL*	182,877
M3-L	~4,119,100	intergenic repeat	~4,310,800	intergenic repeat	~192,000
M4-L	4,048,979	*pflu3655*	4,339,314	intergenic	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). The duplicated region in M3-L is not precisely defined due to highly repetitive sequences flanking the duplication. For details of the genes contained within each duplication segment see **Table S3**.

A combination of bioinformatic analyses of the whole genome sequencing data, 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 SBW25 chromosome (**Table 1**, **Fig. 4A**). The fifth isolate (M3-L) contains a duplication very similar – but not identical – to that in M2-L. The precise location of the M3-L duplication could not be determined due to the presence of highly repetitive flanking DNA.

Direct tandem duplication events generate a novel duplication junction, the presence of which can be shown by PCR (**Fig. 4B**). This technique was used to screen for the presence of the M1-L duplication on each day of the 15-day evolution experiment. A sample of the population frozen each day was revived and used as template for the PCR. The emergent PCR product was first visible on Day 3, and grew stronger as the experiment progressed (**Fig. 4C**). In similar PCRs for the remaining junctions, emergent products were first detected 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 in the experiment (within 5 days, ~35 generations).

The two phenotypically distinct isolates from Line M2 (see **Fig. 3A**) are also genetically distinct; M2-L carries a 192 kb duplication, while M2-Lop contains a 183 kb duplication (with ~104 kb overlap). Notably, neither duplication encompasses the other, indicating that M2-L and M2-Lop arose independently (as opposed to a single duplication event followed by shrinking). Further, there is PCR evidence for a third duplication junction in two other large colony isolates from Line M2, meaning that there are at least three distinct duplications in the Day 13 Line M2 population (see **Fig. S3**). Overall, there are at least five distinct duplication segments across the experimental Lines, demonstrating an unexpected degree of flexibility in the regions from which duplications originate.

Each of the genome-sequenced duplications is unique and, as such, contains a different set of duplicated genes (**Table S3**). The M2-Lop duplication is the only one of the five genome-sequenced duplication events that encompasses the entire CAP locus (*pflu3655-pflu3678*). Upregulation of

pflu3655 has previously been shown to underpin the emergence of opaque colony bistability (Gallie et al., 2019; Remigi et al., 2019). Together with the opaque colony phenotype (see **Fig. 3A**), duplication of the CAP operon in M2-Lop strongly indicates a gene dosage effect of the duplications; at least some of the duplicated protein-coding genes lead to changes in expression and, ultimately, phenotype.

There is a 45 kb segment that is duplicated in all five duplication strains (4,119,923 – 4,164,966). This shared segment contains 45 genes, including one tRNA gene: *serTGA* (**Fig. 4A**). The *serTGA* gene encodes the essential tRNA, tRNA-Ser(UGA), that is thought to functionally compensate for the deleted tRNA (*i.e.*, can theoretically translate codon 5'-UCG-3'). Close inspection of the genome sequencing data shows no evidence of any anticodon mutations in either copy of *serTGA* (**Table S4**), meaning that each of the five mutant isolates contains two identical copies of *serTGA* (as compared with a single copy in the ancestral strains). M4-L is the only isolate in which the duplication segment contains any other tRNA genes; four tRNA genes – *argTCT*, *hisGTG, leuTAA, hisGTG* – occur in close proximity at the 3' end of the M4-L duplication segment.

While it seems logical that the duplication of *serTGA* underpins the observed gain in fitness, there are 44 other genes in the shared duplication segment, any one or combination of which could contribute to the phenotype. Therefore, we sought to test whether an increase in tRNA-Ser(UGA) can compensate for loss of tRNA-Ser(CGA). The *serCGA* and *serTGA* genes were amplified from SBW25 by PCR, and each ligated into the expression vector pSXn (**Text S2**). Along with the empty vector (pSXn), each constructed plasmid (pSXn-CGA, pSXn-TGA) was inserted into SBW25, Δ*serCGA*-1 and Δ*serCGA*-2. Growth analyses of the resulting strains demonstrated that expression of either *serCGA* or *serTGA* improves the growth of the Δ*serCGA* strains in rich medium (**Fig. 4D**). In particular, maximum growth rate increases (**Fig. 4F**; two sample *t*-tests $p<0.05$). Contrastingly, expression of neither tRNA gene improved the growth of SBW25 (**Fig. 4E**), with *serCGA* expression actually leading to a decrease in the SBW25 maximum growth rate (**Fig. 4F;** one-sided two sample *t*-test $p=0.000158$).

Together, the results of this section show that the growth defect caused by deletion of the *serCGA* can be repeatedly and rapidly compensated by large duplication events containing *serTGA*.

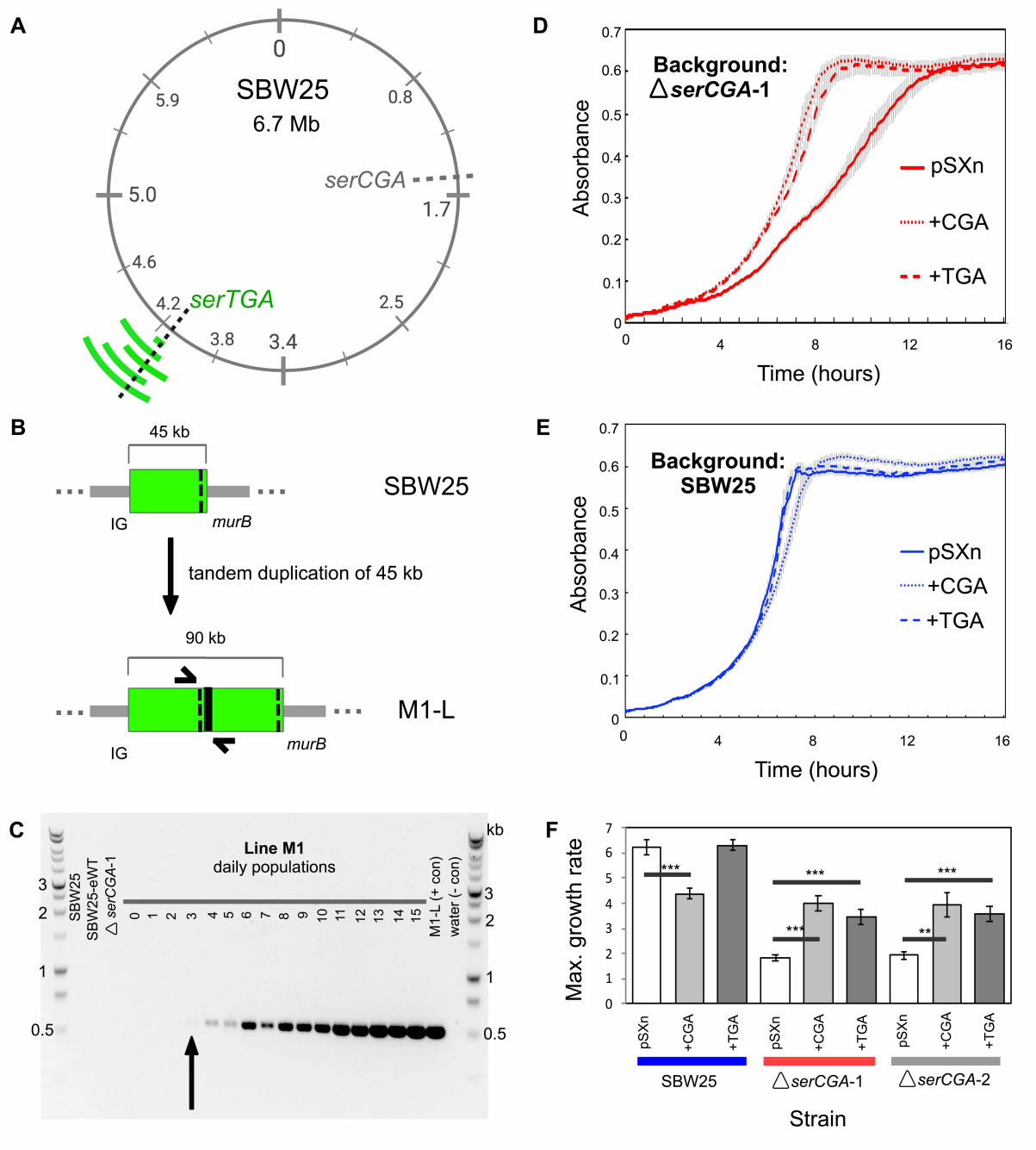

Figure 4. Direct, tandem duplications spanning *serTGA* compensate for the loss of *serCGA*. (A) Five isolates from the mutant Lines have unique, large, direct, tandem duplications between 4.05 and 4.34 Mb of the SBW25 chromosome (green arcs; moving outwards: M1-L, M2-L, M2-Lop, M3-L, M4-L). The duplications 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**) 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**. (**D, E**) 12-hour growth curves in LB+Gm (20 µg ml^{-1}) for Δ*serCGA*-1 (D) and SBW25 (E) expressing *serCGA* or *serTGA* from the pSXn plasmid. Lines=mean of six replicates, error bars=1 standard error. (**F**) 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 ***$p<0.001$, **$p<0.01$, *$p<0.05$.

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

Work to this point shows that the fitness decrease resulting from *serCGA* deletion can be compensated by duplication of a 45 kb chromosomal segment that includes *serTGA*. Next, the effect of the deletion and subsequent evolution on the mature tRNA pool was examined. To this end, YAMAT-seq – an established method of deep-sequencing mature tRNA pools in human cells (Shigematsu et al., 2017) – was adapted for use in SBW25. Briefly, YAMAT-seq involves (i) ligation of hybrid RNA/DNA adapters to the universally conserved 5'-CCA-3' end 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 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**); umbers 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 predicted secondary structure of which deviates significantly from the conserved cloverleaf structure (GtRNAdb; Chan and Lowe, 2016) – was not detected in any sample, indicating that this sequence does not contribute to the mature tRNA pool. Of the remaining 41 tRNAs, three consistently gave very low read numbers (<0.01 % of reads *per* sample in **Table S5**: Glu-UUC, Ile2-CAU, Phe-GAA). This may be the result of post-transcriptional modifications impeding the sequencing process (Machnicka et al., 2016; Shigematsu et al., 2017). These three tRNAs were excluded from the downstream analyses, leaving a total of 38 reference tRNAs *per* sample. Notably, these 38 reference tRNAs include all six retained non-essential tRNA types listed in **Table S1**, and their six functionally corresponding tRNA types.

In order to detect changes in mature tRNA pools, DESeq2 was used to compare normalized expression levels of 38 tRNAs in pairs of strains (Love et al., 2014; **Table S6**). First, the effect of deleting *serCGA* was investigated by comparing tRNA sequences from each of the two independent deletion strains with those in SBW25 (**Fig. 5A**; **Table S6**). As expected, tRNA-Ser(CGA) was absent from the pool of both deletion mutants. The deletion mutants also showed lower expression of tRNA-Thr(CGU) (67 % and 59 % of SBW25, respectively; DESeq2 adjusted $p<0.001$). This may reflect the close metabolic relationship between threonine and serine amino acids and tRNAs (Liu et al., 2015; Sawers, 1998). No significant differences in expression were detected between the two independent deletion mutants (DESeq2 adjusted $p>0.1$).

Next, the effect of subsequent evolution was investigated; pairwise comparisons were performed between each of the six Day 13 isolates and its 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 deletion mutant ancestor (DESeq2 adjusted $p<0.0001$). This demonstrates that duplication of *serTGA* is accompanied by doubling of tRNA-Ser(UGA) in the mature tRNA pool.

A number of other tRNA types co-vary with the rise in tRNA-Ser(UGA) (**Fig. 5B**; **Table S6**). Namely, the Day 13 mutant Line isolates show consistently higher levels of five tRNA types than the deletion mutant, and lower expression of ten others. While none of these differences are statistically significant in all strains, they are consistently in the same direction. 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), there exist a plethora of more subtle effects on the mature tRNA pool.

While the five duplication strains all contain two copies of *serTGA*, each strain differs in regards to the content of the remainder of the duplicated region (**Fig. 4A**, **Table S3**). To investigate whether these differences affect the mature tRNA pool, all possible pairwise comparisons were made between M1-L, M2-L, M2-Lop, M3-L and M4-L (**Table S6**, **Fig. S4**). The strain with the lowest number of differences is M4-L; no differences were detected in any tRNAs between M4-L and any of the other duplication strains (DESeq2 adjusted $p>0.3$). The duplication fragment in M4-L is the largest of the five, engulfing all other duplication fragments (except a small region of M2-Lop). M2-Lop shows the highest number of differences to the other strains (DESeq2 adjusted $p<0.01$). The M2-Lop duplication fragment contains eight genes at the 5' end that are not duplicated in any other isolate. This includes *pflu3655*, the upregulation of which is known to lead to a plethora of metabolic and growth effects (Gallie et al., 2019, 2015; Remigi et al., 2019). It is therefore not especially surprising that M2-Lop demonstrates a relatively high number of differences in the tRNA pool compared with the other isolates. No differences were detected between M2-L and M3-L (DESeq2 adjusted $p>0.3$), the two strains carrying the most similar duplication fragment.

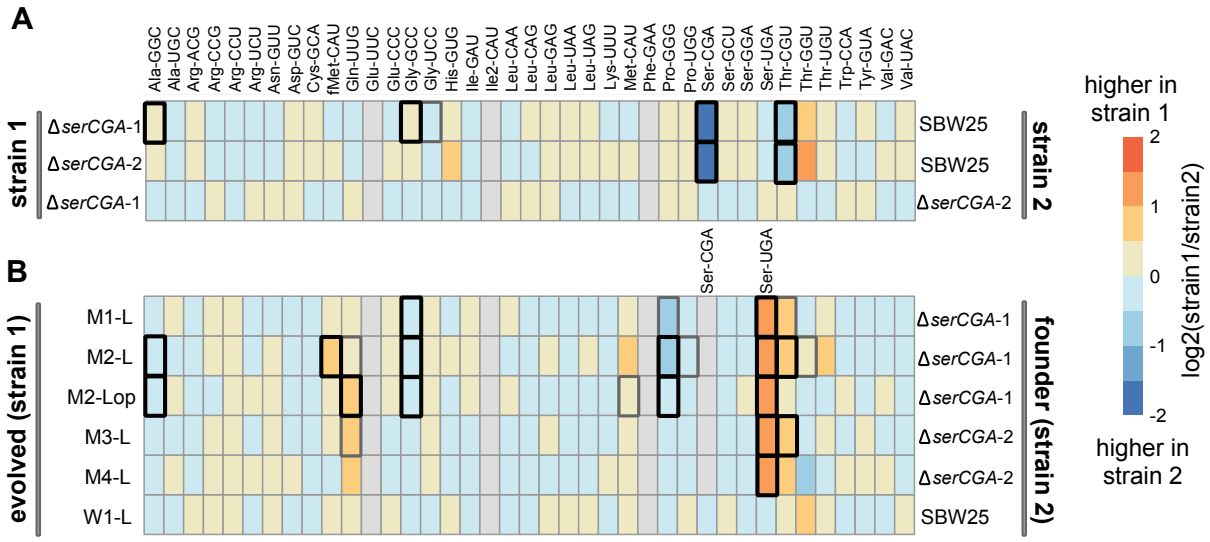

Figure 5. Heatmaps showing differences in the expression of tRNA types between strains. The log2-fold.change(strain1/strain2) difference in expression for 36 tRNA types (with 38 different primary sequences) was determined for pairs of strains using DESeq2 (Love et al., 2014). (**A**) tRNA expression levels in two independent *serTGA* 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 very 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. See also **Table S5, Table S6** and **Fig. S4**.

DISCUSSION

This work reports the emergence of a new bacterial tRNA gene in direct response to selection acting on a sub-optimal tRNA gene set; the deleterious effect of removing *serCGA* from *P. fluorescens* SBW25 is readily compensated through duplications containing *serTGA*. Importantly, increased expression of *serTGA* can compensate for *serCGA* loss (**Fig. 4D**, **4E**), and each duplication leads to a two-fold increase in tRNA-Ser(UGA) in the mature tRNA pool (**Fig. 5B**). Presumably, tRNA-Ser(UGA) can perform the function of tRNA-Ser(CGA) – translating codon 5'-UCG-3' – through wobble base pairing. Elevating tRNA-Ser(UGA) levels is expected to speed up translation and growth by increasing the chance of sampling the required tRNA type at 5'-UCG-3' codons.

The mechanism behind the duplication of *serTGA* is large-scale, direct, tandem duplications in the SBW25 chromosome. Similar duplications are a well-documented adaptive solution to various phenotypic challenges in phage, bacteria and yeast (reviewed in Anderson and Roth, 1977; Reams and Roth, 2015). Such duplications occur spontaneously, and at high rates. For example, spontaneous duplication rates measured at 38 loci across the *Salmonella typhimurium* chromosome vary from $1x10^{-4}$ to $3x10^{-2}$, meaning that some loci are duplicated in one in every 300 cells (in the absence of selection; Anderson and Roth, 1981). Regions with the highest duplication rates contain DNA that is repeated elsewhere in the chromosome, from which duplications can arise by unequal recombination (reviewed in Reams and Roth, 2015). Examples include *rrn* cistrons (Anderson and Roth, 1981), *rhs* genes (Lin et al., 1984), and short repeat sequences such as REPs (Shyamala et al., 1990).

In the case of the duplications isolated in this work, two (M2-L and M3-L) have endpoints that are obviously repetitive: both have one endpoint in a 1.4 kb region of repetitive DNA between *pflu3722* and *pflu3723*, and the other in a similar repetitive region between *pflu3908* and *pflu3909*. Presumably, these duplications are the result of unequal recombination between direct repeats at the two endpoints. The mechanism of duplication is less clear in the remaining strains; in each case, either one (M2-L) or both (M2-Lop, M4-L) endpoints occur in genomic regions in which direct repeats are not readily identifiable. The considerable variation observed in duplication endpoints likely reflects the abundance of repetitive sequences in the genome; SBW25 contains numerous classes of repetitive elements, covering approximately 12 % of the genome (Bertels and Rainey, 2011; Silby et al., 2009). Many of these types of repeats could feasibly potentiate duplication by unequal recombination.

Large duplications are highly unstable; large duplications in the *S. typhimurium* chromosome are reported to be lost in 1-30 % of cells in an overnight culture (in the absence of selection; Sonti and Roth, 1989). It has been suggested that, due to their combined ease of gain and loss, duplications may serve as fleeting solutions to transient phenotypic challenges (Sonti and Roth, 1989). However, in the case of the duplication mutants in this work, loss of the duplication is unlikely to be straightforward due to continued selection for higher levels of tRNA-Ser(UGA) (in rich media). One possibility is partial loss of duplication fragments through unequal recombination events occurring to one side of the *serTGA* gene. This seems probable, given that the largest duplication is 290,335 bp long (M4-L), of which only around 300 bp (~0.1 %) is required for compensation.

The evolutionary fate of the two copies of *serTGA* is also a consideration. As for any duplicated gene, diploidy enables a degree of evolutionary flexibility; mutations may be tolerated in one copy of the gene because the second gene copy remains intact. In the case of tRNAs, the presence of multiple gene copies sets the stage for anticodon switching, whereby mutations in the 3-bp anticodon change the tRNA type encoded by the gene. Evidence for several historic anticodon switching events was provided by a comprehensive computational study of 319 bacterial genomes (Wald and Margalit, 2014). Further, an example of anticodon switching has been observed in laboratory *S. cerevisiae* populations (Yona *et al*., 2013). In this study, one of eleven copies of *argTCT* switched, by one point mutation, to encode *argCCT*.

An anticodon switch event may be a possibility in the case of the evolved mutants in this study. Specifically, one of the two *serTGA* copies could conceivably acquire a T→C transition at the first anticodon position. This would change the gene from *serTGA* to *serCGA*, restoring the original tRNA gene set of *P. fluorescens* SBW25. In order for such a mutation to rise in frequency in the population it must firstly encode a functional tRNA, which requires the new, hypothetical tRNA-Ser(CGA) to be recognized by seryl-tRNA ligase (SerS). The SerS enzyme covalently bonds the amino acid serine to all types of serine tRNAs in the cell (for a review of tRNA ligases, see Ibba and Söll, 2000). Recognition of serine tRNAs by SerS depends not on the tRNA sequence or the anticodon, but rather on the characteristic three-dimensional shape of serine tRNAs (Lenhard et al., 1999). Thus, even though *serTGA* and the original *serCGA* encode tRNAs with very different sequences (see **Fig. S1A**), it is feasible that the new *serCGA* – which would encode the tRNA-Ser(UGA) backbone, but with a CGA anticodon – may form a functional tRNA (see **Fig. S1B**). However, for an anticodon switching mutation to spread through the population in our evolution experiment, the new, hypothetical *serCGA* must not only encode a functional tRNA, but must provide a selective advantage over carrying two copies of *serTGA*. Whether such an anticodon switch actually occurs in our mutants remains to be seen; however, none of the Day 13 mutant Lineage isolates show any evidence of anticodon switch events in either *serTGA* copy (see **Table S4**).

A considerable degree of variability in the SBW25 tRNA gene set has been demonstrated in this work. At least three different tRNA gene sets provide similar growth rates under the conditions

tested; SBW25 encodes 66 tRNA genes of 39 tRNA types, four of the five compensated strains carry 66 tRNA genes of 38 types (M1-L, M2-L, M2-Lop, M3-L), and the fifth (M4-L) carries 70 tRNA genes of 38 types. Such a high – and rapid – degree of tRNA gene set variability points to under-appreciated fluidity in tRNA pools at the population level; even in the absence of selection, a significant proportion of a population can be expected to vary in tRNA gene copy number. Further, the fact that M4-L contains four additional tRNA genes (the *argTCT-hisGTG-leuTAG-hisGTG* gene quad), and yet shows no differences in the mature tRNA pool compared with the other isolates from the mutant Lines (see **Fig. S4**) illustrates an important point: the contribution of individual tRNA to the mature tRNA pool is not necessarily straightforward, and thus tRNA gene copy number is not a reliable proxy for the relative level of a tRNA in the mature pool. The lack of reliable correlation between tRNA gene copy number and level in the mature tRNA pool results from a plethora of regulatory processes that affect the mature tRNA pool (reviewed in Rak *et al*., 2018). In the case of M4-L, it is probable that the duplication junction – which lies 112 bp upstream of the duplicated *argTCT-hisGTG-leuTAG-hisGTG* tRNA gene quad – truncates a promoter, leading to little or no expression of the duplicated copies.

A notable feature of tRNA gene sets is that tRNA types are often encoded by multiple, identical gene copies, dispersed through the genome (Chan and Lowe, 2016; Marck and Grosjean, 2002). For example, 14 of the 39 tRNA types in SBW25 are encoded by multicopy genes, and only two of these (*asnGTT* and *fmetCAT*) show any diversity in primary sequence among gene copies (see **Text S4**). It is possible that the high degree of tRNA gene primary sequence conservation reflects, at least in part, new tRNA gene copies arising from within-genome duplication events.

In this work, we have shown a bacterial tRNA gene set to have an unexpectedly high degree and rate of flexibility. An engineered, sub-optimal tRNA gene set was readily compensated by large scale duplication events that increased the copy number of a remaining tRNA gene. The ease and repeatability with which strains carrying duplication events were isolated suggests that duplications may be an important source of new tRNA genes in bacteria.

MATERIALS AND METHODS

Strains growth conditions. A list of the strains used, and the associated constriction details (including a list of oligonucleotides) are available 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. The relevant 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 each grown in 200 μl of liquid

KB, M9 or LB+Gm (20 μg ml^{-1}) in wells of a 96-well plate. Two microliters of each 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 9 (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 (the 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 lineage 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 on 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). Quality of DNA 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). 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. A full description is provided in **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 (Barrick et al., 2014; Deatherage et al., 2015; Deatherage and Barrick, 2014) 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 (in which the precise duplication junction could not be identified due to highly repetitive DNA sequence). Each predicted junction sequence was checked by alignments to raw reads and previously unused sequences (Geneious). Finally, the junction sequences were confirmed by PCR amplification and Sanger sequencing. The presence of the imprecisely identified M3-L junction was also confirmed by PCR (Sanger sequencing of the PCR product was unsuccessful due to repetitive DNA).

Historical junction PCRs. Duplication junction PCR primers were also used to identify the earliest time point at which each junction could be detected in its lineage history. Glycerol stock scrapings of the frozen daily populations from each of the four mutant Lines (Day 0 – Day 15) were grown in overnight KB cultures. Cells were washed and used as PCR templates, alongside appropriate positive and negative controls (see **Text S2** for full details). 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. 4C** 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).

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 were: 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 the expression of tRNA types during YAMAT-seq (**Fig. 5**, **Fig. S4**, **Table S6)**. Analyses were performed in R version 3.6.0. Levels of statistical significance: *$0.05<p<0.001$, **$0.01<p<0.001$, ***$p<0.001$.

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.

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

Bertels F, Rainey PB. 2011. Within-genome evolution of REPINs: A new family of miniature mobile DNA in bacteria. *PLoS Genetics* **7**:e1002132. doi:10.1371/journal.pgen.1002132

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

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

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

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

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

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