Abstract

Design and Construction of a Recoded Organism Containing 62 Codons

Daniel Wood Moonan
2018

In *Escherichia coli*, 64 codons code for 20 natural amino acids and translation termination. This redundancy raises several questions. First, are all 64 codons and the translational factors that recognize them essential? Second, can this redundancy be exploited to create organisms with novel phenotypes? Both questions can be addressed through the creation of a genomically recoded organism (GRO) where all instances of a redundant codon are mutated to a synonymous codon and the translational factor that recognizes the mutated codon is deleted or engineered. Such recoded organisms exhibit several novel phenotypes that spring from the removal of redundancy in the genetic code.

Recoding requires the insertion of a large number of site-specific mutations, which generally relies upon the synthesis of DNA containing the mutations. Integration of a DNA fragment containing multiple desired mutations may save both DNA synthesis cost and effort compared to separate DNA fragments. For a few sparse mutations, recognition of such opportunities is simple. However, an algorithm for the optimum design of DNA fragments is required when the number and density of mutations increases. Here are described two algorithms for the optimum design of genome wide mutations. The first finds the minimum number of fragments to cover all mutations given a maximum DNA fragment size. This minimizes the number of DNA fragments but ignores cost. The second algorithm casts the problem as a weighted Set Cover

optimization and minimizes both DNA synthesis costs and the number of DNA fragments. For UGA recoding, this provides a design with DNA synthesis costs only 33% of those for whole genome synthesis, while maintaining the same number of DNA fragments. Thus, I have developed an algorithm which provides a general method to optimize DNA synthesis strategies for the insertion of genome-wide mutations.

Additional free codon(s) would enhance the properties of the ΔUAG recoded organism. The UGA codon – the second rarest codon in *E. coli* – is an attractive choice for a second free codon since only 1,178 total UGA codons exist in the MG1655 genome. Given the difficulty to recode all 1178 UGA codons, a pilot experiment to validate UGA to UAA recoding was undertaken in three steps. First, I computationally analyzed all of the UGA codons present in the genome and designed mutations to recode them to UAA while minimizing unintended polar affects. Second, I removed 109 native UGA codons, including those at the end of all essential genes, in the C321.ΔA.2 strain already lacking the UAG codon. Finally, the fitness of the resulting strain was characterized by determining doubling times and measuring growth on minimal media. This establishes an important result and a clear path toward an E. coli strain with only one stop codon (UAA).

Since the peptide release factors provide recognition of stop codons, removal of UGA codons from the genome does not affect the translational meaning of the UGA codon. In order to obtain an *E. coli* strain with two free codons (ΔUAG, ΔUGA), the release factors must be altered. Two main experiments were conducted to establish compatibility at the translational level for the UGA recoding efforts. In the first, I sought

to remove native recognition of the UGA codon by removing RF2 in an essential UGA recoded strain (C321.2.109.F) with RF1 reinserted in order to act as the sole release factor for the remaining UAA codons. In the second translational experiment, we sought to engineer a mutant release factor that terminates at only UAA codons - an essential step towards an organism with a free UGA and UAG codon. The resulting mutant release factor had specificity towards only UAA codons which was validated using both *in vitro* and in *vivo assays*. This mutant release factor provides the translation component necessary for an *E. coli* strain with two free codons (ΔUAG, ΔUGA).

Of the remaining 1069 UGA codons, 995 have been removed across twelve strains. Of the remaining 74 codons, 29 of them have previously been removed in a reduced genome *E. coli* strain. This provides strong evidence that no individual UGA codon is essential for the growth of *E. coli* and removes the last major barrier to an *E. coli* strain lacking both UAG and UGA. These twelve strains have been consolidated into ten strains using conjugative assembly genome engineering (CAGE). These data together show the viability of achieving a single strain lacking both UGA and UAG codons. This strain is expected to have several interesting phenotypes, including increased phage resistance and the ability to incorporate two distinct nonstandard amino acids into proteins. Together, these phenotypes should establish the utility of an *E. coli* strain with two free codons (ΔUAG, ΔUGA).

Design and Construction of a Recoded Organism Containing 62 Codons

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

By
Daniel Wood Moonan

Dissertation Director: Farren Isaacs
Susan Baserga, Jason Crawford, Matt Simon

December 2018

© 2019 by Daniel Moonan

All rights reserved

Table of Contents

Abstract ..... **Page i**

List of Figures ..... **Page ix**

List of Tables ..... **Page xi**

Dedication ..... **Page xii**

Acknowledgements ..... **Page xiii**

Chapter 1 ..... **Page 1**

Introduction: Genetic code allows for removal of redundant codons

Chapter 2 ..... **Page 11**

Design and optimization of recoding

Chapter 3 ..... **Page 23**

Genomic recoding of the Escherichia coli essential genes terminating in a UGA codon

Chapter 4 ..... **Page 42**

Engineering a release factor with activity at only UAA codons

Chapter 5 ..... **Page 60**

Progress toward genomic recoding of all UGA codons in Escherichia coli

Chapter 6 ..... **Page 79**

Future Directions

Appendix A ..... **Page 88**

Chapter 3 supplemental tables and plasmid sequences

Appendix B ..... **Page 158**

Chapter 4 supplemental data

Appendix C ..... **Page 160**

Chapter 5 supplemental data

References ..... **Page 167**

List of Figures

Figure 1.1. Codon frequency in MG1655 of the 20 rarest codons ……Page 9

Figure 1.2. Characteristics of UGA codons ……Page 10

Figure 2.1. Possible DNA fragment designs to cover the AGA codons in a segment of the *E.* coli genome ……Page 19

Figure 2.2. DNA fragment designs to recode UGA and AGG provide differing DNA synthesis costs for an equivalent number of DNA fragments ……Page 20

Figure 2.3. Weighted set cover algorithm provides optimal solution to recoding at low to moderate codon densities ……Page 21

Figure 3.1. Design of UGA genomic recoding ……Page 35

Figure 3.2. All possible mutations caused by UGA to UAA recoding ……Page 36

Figure 3.3. Allele frequency of conversion for a subset of converted essential UGA codons ……Page 37

Figure 3.4. Allele frequency of conversion is not explained by several variables ..Page 38

Figure 3.5. Outline of the CAGE process and the necessary selectable markers …Page 39

Figure 3.6. Genotypic and phenotypic characterization of UGA recoded strains ..Page 40

Figure 3.7. Parameters for design of mutagenic ssDNA oligos ……Page 41

Figure 4.1. Engineering of mutant release factor 1 ……Page 57

Figure 5.1. The distribution of UGA codons in the regions of the genome deleted in the MDS42 strain……Page 71

Figure 5.2 Two options for marker placement in CAGE process ……Page 72

Figure 5.3. Initial characterization of recoding progress ……Page 73

Figure 5.4. Full UGA recoding progress ……Page 74

Figure 5.5. Violin plot of the aggregate doubling times of picked colonies over time for strain G ……Page 75

Figure 5.6. Comparison of properties of genes with converted UGA codons versus genes with non-converted UGA codons ……Page 76

Figure 5.7. Off-target mutations present in the full UGA recoding strains ……Page 77

Figure 6.1. Future phage infectivity assay ……Page 86

Figure 6.2. The effect or recoding and release factor engineering on the incorporation of non-standard amino acids at UAG and UGA codons ……..Page 87

List of Tables

Table 2.1. The costs associated with DNA synthesis used in model ……..Page 22

Table 4.1. Mutations in *prfB* knockout escape colonies ……Page 58

Table 5.1. Division of the MAGE oligo pools amongst the 12 strains ……..Page 78

Dedication

For my family and friends who helped me throughout my dissertation

Acknowledgements

Thank you to all of those in the lab who have helped me during my PhD. Jesse Tordoff, Spencer Katz, Felix Radford, Jaymin Patel, Paul Muir, and Jack Moen have all done work that directly contributed to my projects. Eddie Barbieri, Ryan Gallagher, Alexis Rovner, Natalie Ma, Adrian Haimovich and Mira Amiram all helped guide me during the process. Thank you to Koen Vanderschuren and Amy Mulholland for keeping the lab running. A special thank you to Dan Shapiro for putting up with me for nearly four years.

Chapter 1

Introduction: Genetic code allows for removal of redundant codons

The central dogma of biology holds that information stored in DNA is transcribed to RNA before being translated into proteins[1]. The information contained within DNA is stored with 4 nucleotides, while the eventual proteins consist of 20 amino acids. Clearly each nucleotide can't correspond to an individual amino acid since 16 amino acids would lack a nucleotide. Instead, groupings of three nucleotides produce 64 unique codons. Since this exceeds the minimum of 21 necessary to code for the 20 amino acids and a signal for stop, each amino acid besides tryptophan and methionine is signified by between two and six codons[2].

The meaning of each codon is determined by its corresponding translational factors. For sense codons (those that code for amino acids), their meaning is derived from a tRNA and aminoacyl tRNA synthetase[3]. Each codon is matched to at least one tRNA, which in turn is matched to a synthetase that charges it with the appropriate amino acid. If the tRNA is charged with a different amino acid, then its cognate codon will be decoded with this altered amino acid[4]. For the stop signal, its meaning is derived from the peptide release factor proteins that break the connection between the final amino acid and its tRNA[5-8].

This redundancy raises several questions. At any given position where a particular amino acid is needed in a protein why would one codon be favored over another? Whatever the cause for codon biases, does it render each of the 64 codons essential? Since many codons are recognized by more than one translational factor, are multiple translation factors for the same codon essential?

The first question has been extensively studied for sense codons[9-12], with one prominent hypothesis being that codon choice alters translational rate[13-15]. The causes[14,16,17] and impacts[16,18] of this altered translational rate have also been investigated. The impact of stop codon choice has not been as well studied. Genome-wide stop codon preference appears to correlate with GC content[19,20]. Differential readthrough could explain some biases. UGA readthrough has been reported as 10-fold more common than UAG readthrough and 100-fold more common than UAA readthrough[21-23]. Readthrough of stop codons in E. coli has been reported to have biological function[24]. Evidence of readthrough at specific stop codons has also been observed in genome wide ribosome profiling experiments[25]. The choice of stop codon could alter the presence of readthrough or the identity of the suppressing amino acid, both of which could impact biological function.

Do these differential functions of stop codons render any of them essential? This can rigorously be solved by removing all instances of a codon and replacing it with a synonymous codon—a process termed recoding. In a groundbreaking study, all UAG codons in *E.coli* were recoded to the synonymous UAA stop codon[26,27]. This conclusively proved that the UAG codon is not essential in *E.coli.* In addition, with the removal of the UAG codon, the translation factor RF1 was rendered redundant. Thus, this experiment also showed that the translational factors that recognize codons are not essential in those codons absence.

Not only did the resulting strain answer these biological questions, it also had several useful and interesting phenotypes. Foremost, it exhibited an increased ability to

suppress UAG codons with exogenously supplied non-standard amino acids (NSAAs)[26,28]. Second, the strain exhibits increased resistance to phage infection[29,30]. Finally, with some additional manipulations, the strain can provide novel solutions for biocontainment[31,32].

I aimed to expand upon these advances by recoding all instances of a second codon to create a genomically recoded organism comprised of 62 codons. This was tackled in two major steps:

1. Genome Recoding: remove all instances of a codon in addition to the already recoded UAG codon to create 62 standard codons and 2 free codons.

2. Translation Factor Engineering: engineer translational factor such that 2 codons are unrecognized by native factors.

Recoding of other codons will provide insights into the differential function of those additional codons. Recoding will also allow for investigation of the function of the translational factors that decode the recoded codons. In addition, engineering of the appropriate translational factors will allow for the incorporation of more than one NSAA. Previously the incorporation of more than one NSAA was used to incorporate FRET pairs[33,34] or a fluorophore along with a biorthogonal tag[35]. The incorporation of an acetyl-lysine along with an amino acid for cross-linking was used to determine the change in protein interactions upon post-translational modification[36].

Several factors could be considered in recoding the next additional codon(s) to recode. First, the nature of the translational factors that recognize the codon to be

recoded. Many codons are recognized by more than one translational factor[37,38]. If the recoded codon is to become a free codon, then multiple translational factors would have to be removed or engineered[39,40]. The second major concern is the frequency of the codon to be recoded. Recoding a more common codon increases the likelihood of fitness effects or lethality caused by recoding. In addition, unless the entire genome is synthesized, recoding a more common codon increases the requisite work.

To investigate possible additional codons to recode, we compiled the frequency of the 20 rarest codons in the MG1655 strain of E. *coli* (**Fig. 1**). The frequency of these codons ranges from 321 for UAG to roughly 15,000 for CTC. All of them are sense codons with the exception of UAG and UGA, the two rarest codons. Several of these codons have already been targeted for recoding. 247 TCG and TCA serine codons across 15 genes were recoded to AGT, AGC or either AGT/AGC in a pilot experiment. This amounts to about 1% of TCG and TCA codons[41]. However, they were unable to successfully recode all TCG and TCA codons to AGT.

To test the viability of sense codon recoding in essential genes, 13 rare codons (UAG, AGA, AGG, CUU, CUC, CCC,ACC, AUA, GUC, GCC, UCC, CGG, UGA) in 45 essential genes were recoded to synonymous codons, testing each gene individually[42]. In total, this amounted to 405 total codons, only 4 of which were UGA codons. While all 405 codons were successfully recoded individually, several required multiple recoding designs. The logic behind the success of some designs and the failure of others was not apparent.

The AGG codon has been targeted for recoding multiple times. In one effort, 29 AGG codons were recoded to CGN, while 2 others were recoded to AGA or the lysine codon AAG[40]. This effort failed to genomically recode an additional 11 AGG codons. In another pilot experiment, all AGG codons in essential genes were recoded in a single strain [43]. Again, initial designs for recoding did not successfully recode all of the targeted codons. Subsequent analysis suggested that the targeted codons were important for RNA structure, which recoding had disrupted.

Despite the challenges encountered in these initial sense codon recoding projects, recoding of 7 codons (including 6 sense codons) was attempted by synthesizing the entire *E.coli* genome[44]. The partial success of this effort provided important insights into the mechanisms of recoding. However, the project also suffered from the following flaws that limit the interpretation of the data and likely make complete success of the project more challenging. The genome was divided into 87 50kb segments, each of which was synthesized as a recoded version then tested individually. Thus, interactions between one 50kb segment and any of the other 86 segments was not investigated. This assumes that recoding sense codons most likely leads to dominant lethality instead of synthetic lethality. Even within the individual 50kb segments, many problems were encountered. First, validation of recoding was done by inserting an episomal copy of a 50kb segment of the recoded genome into strains that continued to have a native genomic copy of the 50kb segment. The genomic copy was then subsequently deleted for only 63% of the 50kb segments. While the plasmid is “low copy-number”, the gene dosage in these instances is still likely higher than a genomic copy. This suggests that

37% of the segments harbor a dominant lethal allele. Of these 50kb segments, only about 13% successfully replaced the native segment in the genome.

Due to technical limitations in DNA synthesis, additional mutations had to be inserted in order to reduce GC content and remove repetitive sequences. While the authors report off-target mutations, these correspond only to unintended errors, not the non-recoding mutation they inserted to make DNA synthesis more feasible. In any recoding effort, mutations other than those required for recoding can confound the analysis of the effects of the recoding mutations.

Given these challenges in recoding sense codons, further recoding of nonsense codons appears comparatively attractive. Since the UGA codon is the second rarest codon, I chose this codon to investigate further.

Beyond the codons themselves, the translational factors that recognize those codons must also be considered. Since the peptide release factors provide recognition of stop codons, simply removing all instances of UGA from the genome will not alter the translational meaning if an instance of UGA is reintroduced. In order to obtain an *E. coli* strain with two free codons (ΔUAG, ΔUGA), the release factors that recognize these codons must be altered. Since release at UAA remains essential, and both RF1 and RF2 terminate at UAA, either could form the basis for an engineered release factor specific to only UAA.

Given the difficulties in recoding sense codons, we searched for a subset of UGA codons that could be used to test the viability of recoding. Recoding this subset first might illuminate potential challenges in recoding all UGA codons. UGA codons at the end of essential genes were attractive for a few reasons. First, since there are 71 essential genes that terminate in UGA, they form a tractable set of UGA codons[45]. Second, these genes are essential and are generally highly expressed (**Fig 2**)[46]. If recoding disrupts expression, these genes are most likely to exhibit a phenotype. Finally, recoding of the essential UGA codons might allow for the investigation of the release factor (RF2) that recognizes UGA. Recoding of the essential UAG codons allowed for the deletion of the translational factor that recognizes UAG (RF1), so we hypothesized that recoding the essential UGA codons might allow a similar deletion of the translational factor that recognizes UGA (RF2). If RF2 could be successfully deleted in this strain, then UAG, UGA and UAA release activity could all be easily assayed *in vivo*.

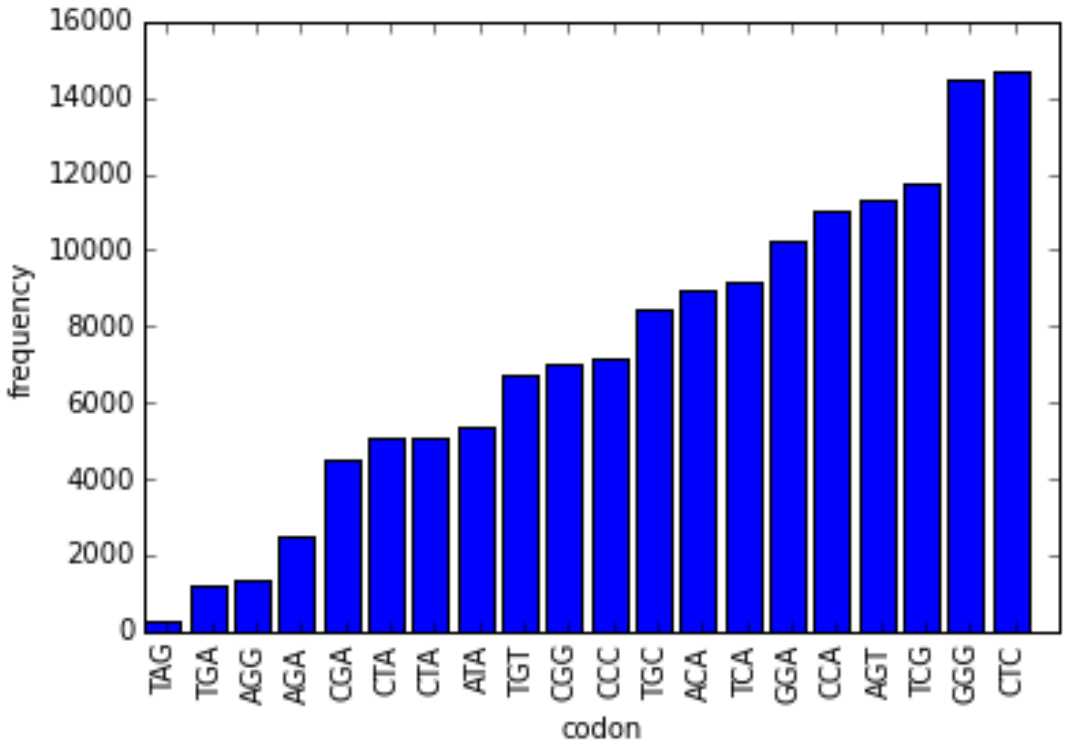


Figure 1. Codon frequency in MG1655 of the 20 rarest codons. The first two codons TAG and TGA are stop codons. The remainder of codons are sense codons. The final stop codon (UAA) is excluded.

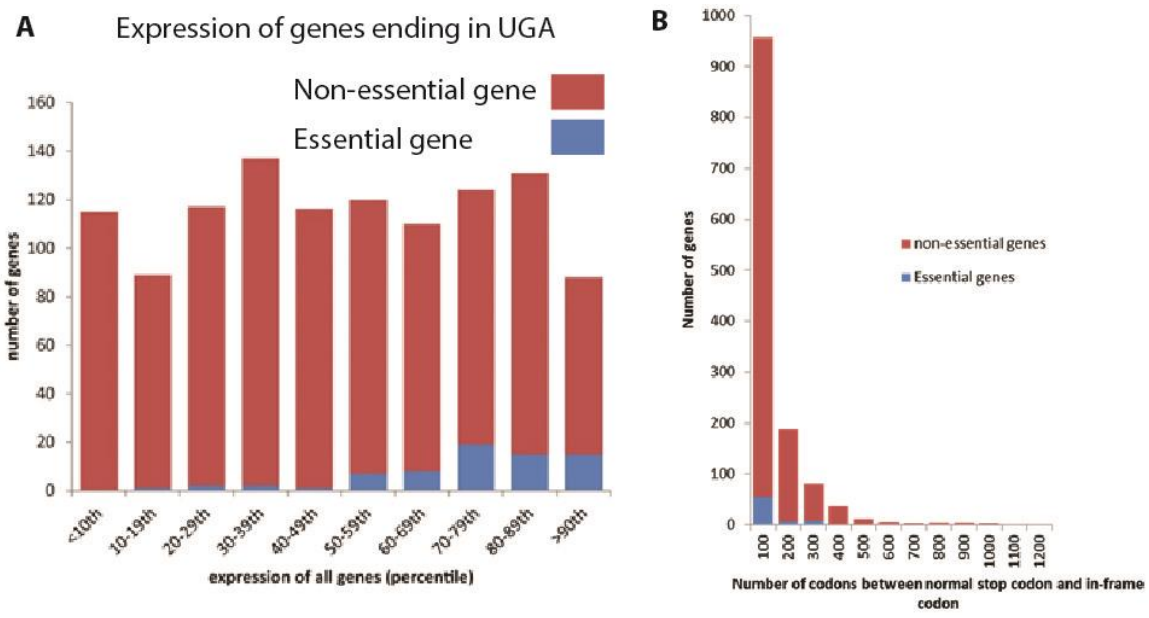


Figure 2. Characteristics of UGA codons a) the expression level of all genes terminating in TGA. The percentiles were calculated from the distribution of expression of all genes. b) Suppression of a native TGA will cause translation to continue past the normal stop codon. If an in-frame TAA exists nearby, then translation will be terminated at this codon instead, with the addition of several C-terminal amino acids. This length was calculated for every TGA terminating gene and the distribution plotted

Chapter 2

Design and optimization of recoding

Optimum design of genome wide mutations

Projects in synthetic biology increasingly call for the introduction of many site-specific mutations across the genome[47]. Two methods have been primarily utilized for the introduction of these site-specific genome-wide mutations: genome synthesis and targeted recombination[27,48]. Whole genome synthesis allows for the facile introduction of many mutations[49], but at significant cost – the first bacterial genome cost $40 million[50]. Recombination based techniques require significantly less expenditure on DNA synthesis[26], but need many manipulations and thus a significant investment of time and labor.

Ideally a strategy could be devised that uses large scale DNA synthesis where the desired mutations are most dense, while targeted recombination creates mutations where the desired mutations are sparse. Such an approach could minimize both the time invested and DNA synthesis costs. If only a few mutations are desired, the decision of which mutations to create using targeted recombination and which to create using large scale DNA synthesis could be made manually. As the number of mutations to insert increases, the complexity of that choice increases (**Fig. 1**). Because of this ballooning complexity, an optimal solution must be determined computationally.

We conceived two computational methods to apply these this idea. The first deemed clustering is the minimization of the number of DNA fragments of a maximum size required to cover all desired mutations. While reducing the portion of the genome synthesized compared to whole genome synthesis, this method ignores any associated

costs. The second casts the problem as a weighted set cover problem[51]. This allows costs (weights) to be associated with each potential DNA fragment. Therefore, the cost associated with inserting all required mutations can be minimized. For genome-wide recoding of UAG codons to UAA, this method provided a set of under 400 fragments that would insert all mutations at a DNA synthesis cost roughly 70% less than that predicted for whole genome synthesis. This demonstrates that the introduction of many mutations into the genome may be made more efficient by better design, rather than the development of new experimental techniques.

Methods

In order to create genome-wide mutations, a set of DNA fragments must be designed that will introduce all of the desired mutations. These desired mutations are represented by an integer **m** denoting the genomic position of each mutation. This is intuitive in the case of single nucleotide polymorphisms. For the mutation of all three of the nucleotides in a codon, a single position provides an approximation. The exact representation of larger insertions or deletions is still possible, but requires special considerations. The set of all desired mutations **m** is denoted as **U**

$$U = \{m_1, m_2, \ldots, m_n\}$$

Subsets **s** represent the individual DNA fragments that can be selected from **U**. These subsets consist of contiguous desired mutations. A set **S** of these subsets represents a tentative solution to the DNA fragment design problem. As long as the constraint that all desired mutations **m** are contained within **S** at least once, then the solution **S** is valid.

Both whole genome synthesis and the synthesis of individual DNA fragments for every mutation are valid solutions.

$$S = \{s_1, s_2, \dots, s_i\}$$

$$s = \{m_j, \dots, m_{j+k}\}$$

Finding an optimum solution requires a maximization (or minimization) function and further constraints. In the method we have deemed clustering, our additional constraint is that all DNA fragments must be less than a given size. The function to be minimized is the number of DNA fragments required to cover all mutations in the set U. An optimum solution to this problem can be found. First, the mutations in U are sorted from least to most. The first mutation m_1 is added to the first set s_1. The next mutation is added to this set as long the distance between it and m_1 is less than the maximum DNA fragment size. When the next mutation exceeds this maximum DNA fragment size, it is added instead to the next set s_2. This process is iterated until all the mutations in U are covered.

An alternative approach optimizes for the lowest cost to introduce all the desired mutations. This can be approached as a weighted set cover problem[51]. A cost (or weight) corresponding to the DNA synthesis cost (**Table 1**) can be associated with each DNA fragment **s** in the prospective solution **S**. The sum of the costs of each DNA

fragment represents the total cost for this solution. The challenge is to find the solution S with the minimum total cost. However, if only the DNA synthesis costs are considered, then the optimum solution will always have the least amount of DNA synthesized. Because of the trivial nature of this solution, we add an additional per DNA fragment cost. This represents a premium paid in order to reduce the total number of DNA fragments. A single fragment would pay this cost only once, while multiple fragments would each incur this cost.

In order to find the solution with the lowest total cost, all possible DNA fragments s must be elucidated. Starting from a fragment containing just the first mutation m_1, additional DNA fragments are created that contain all mutations m_1 to m_{1+i} until m_{1+i} creates a DNA fragment greater than the maximum length. This process repeats for every desired mutation. A cost can then be associated with each DNA fragment as a function of fragment length.

An optimum solution to this problem would require the evaluation of every possible solution. Since each fragment can be either included or excluded, this requires the evaluation of 2^n possible solutions, where n is the number of possible DNA fragments[52]. This proves tractable in simple cases, but rapidly becomes impossible. Thus, an approximation technique is necessary. A greedy algorithm provides the current best approximation to the weighted set cover problem[53]. Instead of trying to find a combination of fragments that is optimal, the greedy algorithm attempts to find only a single optimal fragment, and then repeat this until all mutations are covered. This DNA fragment is chosen by picking the fragment with the highest cost efficiency **h** which

represents the ratio between the number of mutations covered by this mutation but not yet covered by previously chosen DNA fragments and the cost **c** of choosing this ratio

$$h = \frac{m_{uncovered}}{c}$$

Results

The main drivers of cost in DNA synthesis are the total amount of DNA synthesized (which can be represented as the percent of the genome synthesized) and the length of individual fragments synthesized. Longer fragments have higher per base costs than shorter fragments (**Table 1**). The cheapest option therefore is to have many short oligos that synthesize as little of the genome as possible. This strategy is employed when MAGE is utilized for creating genome-wide mutations[26,40,43]. As shown in **Figure 2**, when the number of DNA fragments approaches the number of mutations to insert (~1200 fragments for both UGA and AGG), both the weighted set cover and clustering algorithms return this trivial solution that is compatible with using only MAGE to insert mutations. The major disadvantage of this approach is that unless the desired mutations are densely clustered, short oligos won't span more than one mutation. Every mutation will require its own DNA fragment and each fragment will require additional time for insertion and screening.

To achieve more insightful solutions, the cost per fragment was increased to provide solutions with a lower number of DNA fragments. If the *E. coli* genome was synthesized as 10kb fragments, it would require roughly 450 fragments at a DNA synthesis cost slightly greater than a million dollars. Using an equivalent number of DNA

fragments, the weighted set cover and clustering algorithms provided solutions to UGA and AGA recoding in *E. coli* that decreased the DNA synthesis costs to 350 and 600 thousand dollars respectively. As shown in the inset to **Figure 2**, much of this savings as compared to whole genome synthesis is due to the synthesis of a significantly lower portion of the genome.

Since the clustering algorithm never optimizes for costs it provides a more expensive solution than weighted set cover as the number of fragments increases. This divergence accelerates as the number of fragments decreases, causing the clustering algorithm to provide an answer twice as expensive when 350 fragments are synthesized. Interestingly, however, the fraction of the genome synthesized is only slightly higher than with weighted set cover. This likely stems from the fact that the cost per base for DNA synthesis jaggedly increases as the length of a fragment increases. Thus, even though both methods create designs with similar total DNA synthesis, the weighted set cover favors fragments that fall inside of the length cutoffs that increase the cost per base. This clearly significantly impacts the DNA synthesis costs.

To determine if these results applied to a wide range of mutations and organisms, these methods were applied to the reference genomes of 100 prokaryotes for UAG, UGA, AGG and AGA codons (**Fig. 3**). A ratio of the DNA synthesis costs for whole genome synthesis in 10kb fragments versus the DNA synthesis costs calculated from weighted set cover was taken. This provides a measure of how much optimization of design can reduce the cost of recoding where lower numbers indicate greater improvement. The plot clearly indicates that the mutation density determines the

degree of optimization afforded by the weighted set cover algorithm. As the density of mutations increases (assuming they are roughly evenly distributed) the impact of optimization lessens. This fits with intuition, since closely packed mutations will favor larger DNA fragments, and enough large DNA fragments will be functionally very similar to whole genome synthesis.

Discussion

These results clearly show value in optimizing the design of DNA fragments for the genome-wide insertion of many mutations. This optimization proves most valuable with a moderate density of mutations. Too little density and the mutations are too far apart to practically span with a single DNA fragment. Too dense and optimization returns an answer very similar to whole genome synthesis. Several current recoding projects fall within this region of attractive density[40,43]. In addition, the ambitious project of synthesizing the human genome has been proposed[54]. While few specifics exist, such a project would require incredible expenditure on DNA synthesis. Given this fact, it is likely that at least some portion of the project would benefit from a technique to reduce DNA synthesis costs through optimization. The technique we have presented provides a method to lower DNA synthesis costs when inserting many genome-wide mutations.

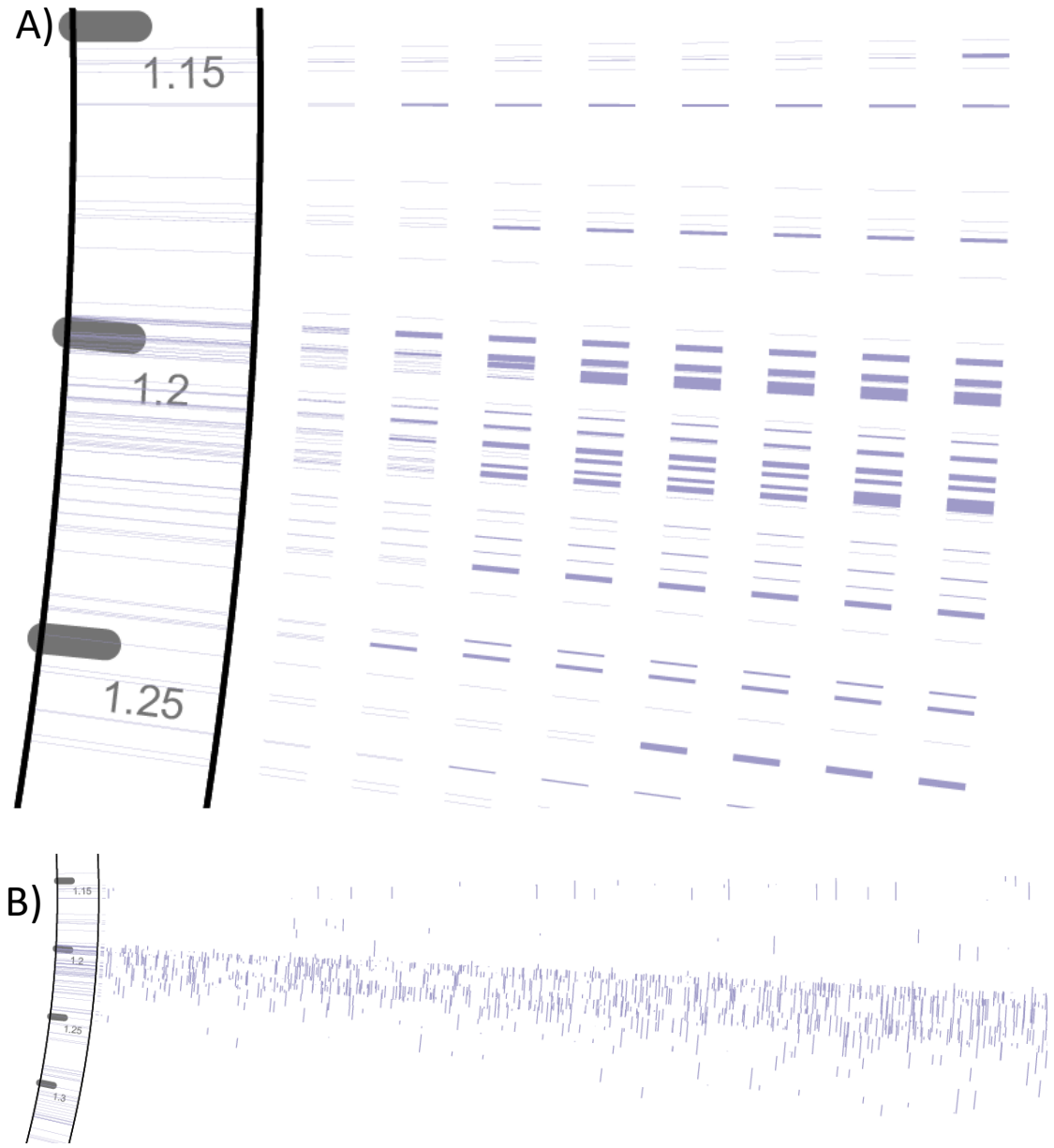


Figure 1. Possible DNA fragment designs to cover the AGA codons in a segment of the *E.* coli genome. A) The leftmost black lines represent the genome, with each purple band within those lines representing an individual codon. To the right of those lines are a sample of eight possible DNA fragment designs to cover all of those codons. Each design is represented by a column of DNA fragments represented by lines of varying width. The width corresponds to the length of the DNA fragment that would be synthesized. B) the same as A, but showing a greater number of the possible designs over a larger segment of the genome

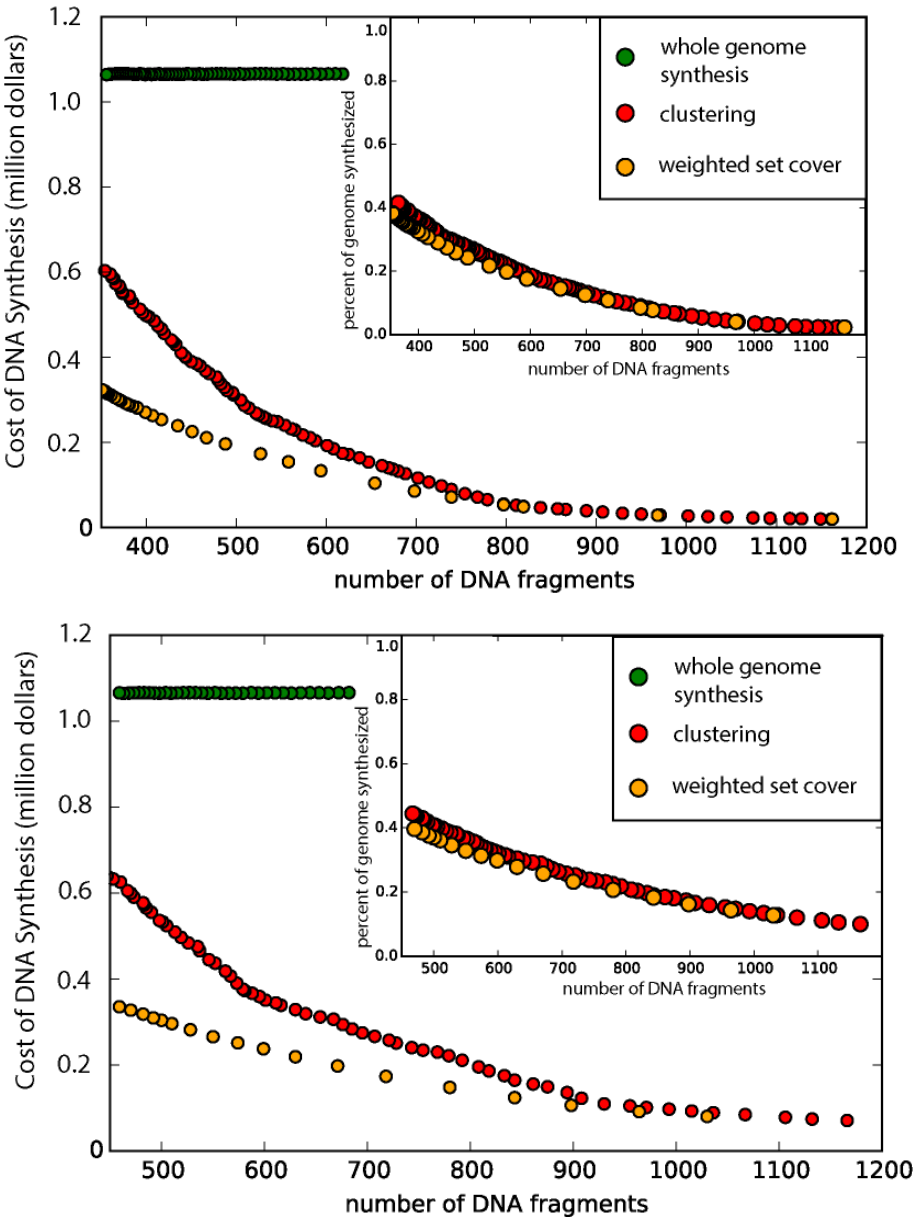


Figure 2. **DNA fragment designs to recode UGA(top) and AGG(bottom)provide differing DNA synthesis costs for an equivalent number of DNA fragments**. Red and yellow dots represent the DNA synthesis cost of an optimal solution using either clustering or weighted set cover at the given number of DNA fragments. Green dots represent the cost of whole genome synthesis using 10kb long fragments. The inset shows the fraction of the genome synthesized for each solution shown in the larger image.

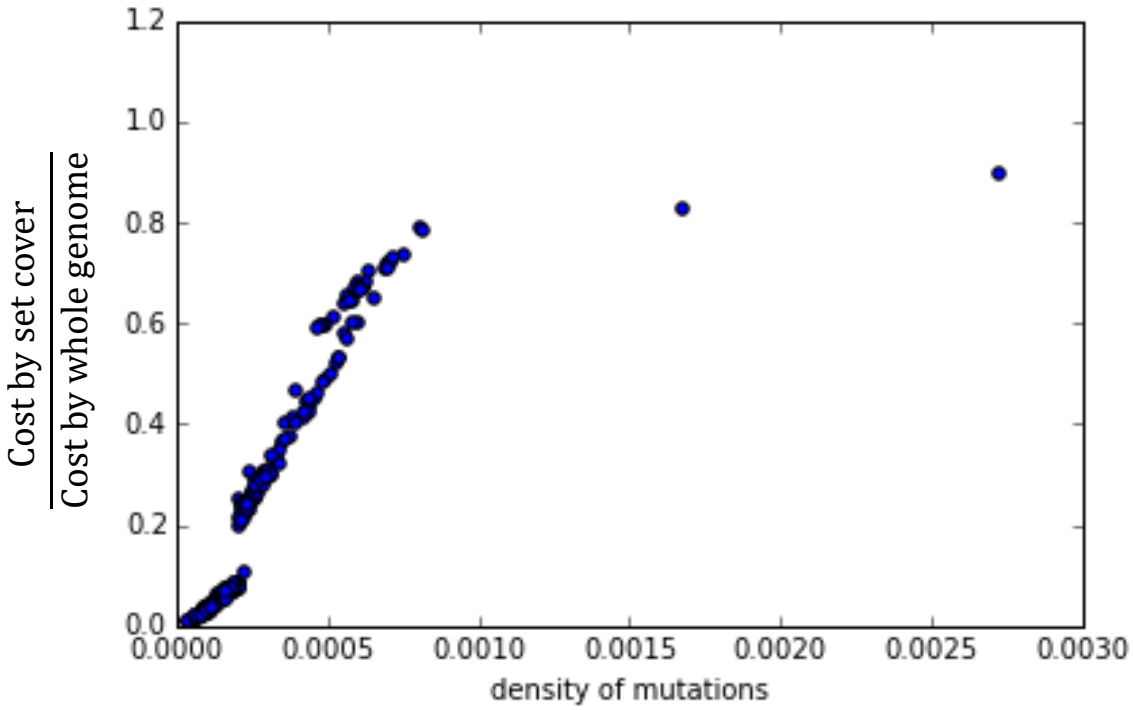


Figure 3. Weighted set cover algorithm provides optimal solution to recoding at low to moderate codon densities. The density of mutations equals the number of mutations to make divided by the total nucleotides present in the genome. The cost by whole genome synthesis is calculated using 10kb fragments.

Size ≤	Cost	Discounted Cost	Vendor[55,56]	Type of recoding
90	$24.30	$17	IDT	MAGE
500	$89	$62	IDT	DNA synthesis/recombination
750	$129	$90	IDT	DNA synthesis/recombination
1000	$149	$104	IDT	DNA synthesis/recombination
1250	$209	$146	IDT	DNA synthesis/recombination
1500	$249	$174	IDT	DNA synthesis/recombination
1750	$289	$202	IDT	DNA synthesis/recombination
2000	$329		IDT	DNA synthesis/recombination
3000	0.23 per base		Genscript	DNA synthesis/recombination
10000	0.23 per base		Genscript	DNA synthesis/recombination

Table 1. The costs associated with DNA synthesis used in model. Discounted costs were obtained by assuming a 30% discount for a bulk DNA order. DNA fragments ≤ 90nt would be inserted using single stranded recombineering, while larger fragments would be inserted using double stranded recombination.

Chapter 3

Genomic recoding of the *E. coli* essential genes terminating in a UGA codon

*This chapter is adapted from the manuscript for: DW Moonan, JR Patel, J Moen, Muir P, J Tordoff, J Rinehart, FJ Isaacs, Genomic recoding of the Escherichia coli essential genes terminating in a UGA codon, submitted to Nature Chemical Biology, 2018,

The genetic code of organisms, with few exceptions, contains 64 codons that degenerately encode for 20 natural amino acids and translation termination. The redundancy of the genetic code provides an opportunity to alter the meaning of redundant codons by first removing all instances of the codon from the genome and then removing the translation factor(s) that recognize the codon. Such a strategy previously allowed the recoding of all 321 UAG codons to the synonymous stop codon UAA in the C321 strain of *Escherichia coli* and subsequent deletion of Release Factor 1 (RF1) - the sole factor that recognizes UAG[26,27]. Early efforts motivated by this work to recode sense codons are underway[41,43,44,57]. These genomically recoded organisms (GROs) possess several advantageous phenotypes, which include increased phage resistance[29], novel solutions for biocontainment[31,32], and more efficient incorporation of nonstandard amino acids into proteins[28].

Additional free codon(s) would enhance the properties of the ΔUAG recoded organism in three important ways. First, a second free codon would facilitate multi-site incorporation of two distinct nsAAs in proteins at high yield and purity[28]. Second, a doubly modified genetic code would likely further obstruct the horizontal transfer of genetic elements (*e.g.,* resistance against multiple phage). Third, a GRO with two open codons would strengthen efforts in biocontainment by establishing the possibility of engineering synthetic auxotrophs with dependencies on two distinct nsAAs.

This work establishes two key advances toward the design and construction of an *E. coli* strain with two free codons. First, we computationally analyzed all of the UGA codons present in the genome and designed mutations to recode them to UAA while

minimizing unintended polar affects. Second, we have removed 109 native UGA codons, including those at the end of all essential genes, in the C321.ΔA.2 already lacking the UAG codon[26].

The UGA codon – the second rarest codon in *E. coli* – is an attractive choice for a second free codon since only 71 essential genes terminate in UGA and only 1,178 total UGA codons exist in the MG1655 genome (**Fig. 1a, Supplementary Table 1**). Since almost all UGA codons occur at the end of a gene, they often overlap a neighboring gene. This overlap can fall in one of three categories (**Fig. 1b**). In the first instance, the A of UGA and AUG is shared, and thus mutating G to A in the UGA codon does not alter the AUG codon. In the second instance, both the U and G are shared between the AUG and UGA codons. Therefore, recoding UGA to UAA disrupts the start codon, which could preclude translation of the overlapped gene. To avoid this problem, the insertion of three bases – AAU – preserves both the new UAA codon and the AUG start codon, albeit with the ribosome binding sequence site (RBS) shifted three nucleotides further from the start codon (**Fig. 1b, Supplementary Table 1**). Since the spacing of the RBS from the start codon can impact translation initiation rate[58], we used the RBS calculator to predict translation initiation rates for both the wild type and recoded sequences (**Fig. 1c**)[59,60], and found little predicted deviation in translation initiation upon recoding. The final category of overlap encompasses all instances where two genes overlap extensively so that recoding inevitably introduces a mutation in the overlapped gene (**Fig. 2**). In the four essential genes with this type of overlap, recoding *mukF* causes a

synonymous mutation in the overlapped gene, while recoding of *rnpA*, *mukF*, and *murC* causes non-synonymous mutations.

The 71 essential genes terminating in UGA as well as 44 proximal UGA codons in non-essential genes were converted to UAA across two C321.ΔA.2 (*ΔUAG*, *ΔprfA*) strains using multiplex automated genome engineering (MAGE) (**Fig 3**, **Supplementary Table 2**)[61,62]. Four non-essential genes that terminate in UGA in six regions of the genome known to be expendable were removed with the *tolC* selectable marker (**Supplementary Table 3**)[63,64]. In total 109, or ~9% of all UGA codons were removed across two strains (C321.ΔA.2.39 and C321.ΔA.2.64) (**Supplementary Table 4**). The efficiency of recoding varied between locations (**Fig. 3**), although neither genomic location, conversion of an UGA overlapping with another gene, nor the Gibbs free energy of folding for the mutagenic oligo appeared to predict this efficiency (**Fig. 4**). Previous work on removing AGG codons in essential genes encountered difficulty in recoding 13 AGG codons, necessitating redesign of the recoding strategy[43]. This was attributed to the initial recoding causing a disruption to either mRNA folding or RBS accessibility. Despite recoding of UGA in 20 essential genes causing an increase in spacing between the RBS and start codon of an overlapped gene, all mutations were incorporated successfully as initially designed without an observable change in phenotype.

The gene *dnaX* provides a particularly interesting case of a UGA codon extensively overlapping with another gene. The *dnaX* gene codes for both the full length τ and the truncated γ subunits of DNA polymerase III. The non-essential γ

subunit is produced after a programmed frameshift partway through *dnaX* leads to a UGA codon[65]. Recoding of this UGA codon to UAA causes a non-synonymous mutation in the essential τ subunit. We attempted to convert this UGA to UAA in the EcnR2 strain[62], but failed to isolate a mutant (**Fig. 1d, Supplementary Table 5**). If the UGA is mutated to a sense codon instead, a UAA codon a further 11 amino acids downstream could terminate translation. Pursuing this strategy, we designed three synthetic oligonucleotides with six positions of degeneracy stretching from two positions before the UGA codon to the position after the UGA codon. Three rounds of MAGE with these oligonucleotides provided seven clones which successfully recoded the UGA codon (**Fig. 1d**).

To unite all of the essential gene mutations in C321.ΔA.2.39 and C321.ΔA.2.64 into a single strain, conjugative assembly genome engineering (CAGE) was utilized (**Fig. 5**)[27,30]. The C321.ΔA.2.64 strain maintained the doubling time of the ancestor C321.ΔA.2, while the C321.ΔA.2.39 strain developed an auxotrophy and higher doubling time in rich media during recoding (**Fig. 6a**). The resulting conjugated strain (C321.ΔA.2.104) inherited both an auxotrophy from the C321.ΔA.2.39 recipient strain and increased doubling time compared to the C321.ΔA.2.64 donor strain (**Fig. 6a**) due to off-target mutations (**Fig. 6b**). We performed whole genome sequencing to confirm all of the UGA to UAA conversions in essential genes and to determine the cause of the auxotrophy (**Supplementary Table 6**). Eighteen essential or conditionally essential (growth on minimal media) genes in the conjugated strain contained off-target non-synonymous mutations (**Supplementary Table 7**). Oligonucleotides correcting these

mutations were used in five rounds of MAGE and the resulting population was plated onto minimal media plates to select against the auxotrophy. The colony with the fastest doubling time was sequenced. Of the 18 essential genes with mutations in the conjugated strain (C321.ΔA.2.104), eight of these were reverted in the resulting C321.ΔA.2.109.F prototrophic strain (**Fig. 6a, Supplementary Table 7**). Together, these results show that CAGE, followed by targeted editing by MAGE, can be rapidly used to select for strains with enhanced properties (*i.e.,* prototrophs with quicker doubling times).

To further gauge the impact of recoding, we investigated recoding at the protein level by introducing an opal suppressor (*glyT*) to incorporate glycine residues at UGA codons[66]. Using MS/MS, we detected 25 peptides in which a UGA codon was efficiently suppressed in MG1655 containing *glyT* (**Fig. 6c**). Five of these peptides were from essential genes. In C321.ΔA.2.109.F with *glyT*, 15 peptides with UGA suppression events were detected, none of which were from essential genes (**Supplementary Table 8**). Notably, these results demonstrate that essential recoding has removed opal suppressor activity at the recoded UGA codons.

In this study, we have removed 109 native UGA codons (9% of the total), including those at the end of all essential genes, in the C321.ΔA.2 GRO strain of *E. coli*. This result suggests that the remaining UGA codons are amenable to recoding. The second task of removing the translation factor that recognizes UGA has progressed in two important directions. First, we have demonstrated the importance of removing all UGA codons by showing that RF2 remains essential in a UGA essential recoded strain, a

distinct result from similar studies involving the UAG codon[26,67]. Finally, through rational engineering, we constructed a mutant RF1 unable to recognize UAG codons. Taken together, our work on recoding and translational factor engineering establish important results and a clear path toward an *E. coli* strain with two free codons (ΔUAG, ΔUGA) possessing a mutant release factor with UAA as the single stop codon.

Methods

Design of MAGE Oligonucleotides

Genome coordinates and essentiality information for all annotates genes in E. coli MG1655 (based on genbank reference sequence U00096.3) were acquired from ecogene.org[68]. For all genes terminating in "TGA" codons, 90mer MAGE oligonucleotides were designed computationally. These oligonucleotides 1) targeted the lagging replication strand, 2) contained >15bp flanking the TGA->TAA mutation position, and 3) were optimized to reduce secondary structure (**Figure 7**). For the latter, in brief, all possible candidate 90mers were analyzed for secondary structure with the ViennaRNA package[69]. The 90mer that minimized Mean Folding Energy was selected. Oligonucleotides were synthesized by IDT with 2 5' phosphorothioate bonds.

Design of MASC PCR primers

Genes were grouped into sets of up to 11 for multiplex allele-specific PCR analysis. For each gene, two forward PCR primers were designed, one with a wild-type "G" base at the 3' end; and one with a mutant "A" base at the 5' end. Reverse primers were designed to create 100bp, 150bp, 200bp, 250bp, 300bp, 400bp, 500bp, 600bp, 700bp,

800bp, and 1000bp amplicons within the gene set. Primers were trimmed until reaching a Tm of 60C using Kun's Oligonucleotide Tm Calculator with default parameters[70,71].

Genomic Recoding

The starting strain was a standard C321.ΔA strain with a zeocin cassette replacing a chloramphenicol cassette at the *mutS* locus[26]. Three mutations were made to this strain. The *ybcG-tolQRA* locus was duplicated between the *ycgV* and *ychF* genes as previously described to increase the fidelity of colicin selections[63]. The previously reported mutations to *dnaG* and *exoX* to improve MAGE efficiency were also present in the starting strain[72]. The oligonucleotides were designed as previously described (**Supplementary Table 4**)[61,62], ordered from IDT and were grouped into pools ranging from 7-10 oligonucleotides in size. These pools were used for ~10 rounds of MAGE before the populations were plated. MAGE was performed as previously described[61,62]. From the plated population, 48 colonies were picked and screened using MASC PCR. The colony containing the largest number of conversions was used for subsequent rounds of MAGE. In cases in which multiple colonies had an equal number of conversions, the colony which converted a UGA codon that had a lower conversion frequency was selected. This process was repeated until all UGA codons in that pool were converted. Allele frequencies were calculated by dividing the number of colonies that had a targeted UGA codon converted by the total number of colonies screened (**Supplementary Table 4**)

MASC PCR

Performed using the protocol as previously described[61,62] Pools consisted of 7-10 sites that corresponded to the oligonucleotides used for MAGE (**Supplementary Table 4**). Reverse primers were designed for each pool to produce a ladder from 100bp-850bp in size. This allowed for the resolution of each site.

Assembly of the final strain

Conjugative Assembly Genome Engineering (CAGE) was performed using the protocol as previously described[27,30]. Deviating from the previous protocol, the donor strain (C321.ΔA.2.64) had two positive markers, the *tolC* marker and the gentamicin marker in addition to the oriTKan cassette (**Fig. 5**). The donor strain also contained a modified RK24 plasmid in which all genes that end in a UAG codon were recoded to UAA[29]. The recipient strain (C321.ΔA.2.39) lacked tolC, but had a spectinomycin marker. Upon conjugation, the desired genotype was selected by growing the cells on LB agar with SDS, gentamicin, and spectinomycin. The concentrations were as previously reported. From the plated populations, 48 of the resulting colonies were screened using MASC PCR to confirm successful conjugation. Five of the UGA codons in essential genes were not present in the conjugated strain (C321.ΔA.2.104) (*dxr*, *ispU*, *fmt*, *glmU*, *lptF*). Four of these were not converted in the ancestor strains, while one was overwritten because it was at the junction between the spectinomycin and gentamicin cassettes. These 5 UGA codons were subsequently converted to UAA using MAGE. The resulting strain contained the three selectable markers, all of which were subsequently deleted to yield C321.ΔA.2.109. *tolC* was first deleted utilizing the colicin selection protocol as previously described[30,63]. Once the strain lacked *tolC*, the *tolC* cassette was used to

remove spectinomycin via lambda mediated recombineering and subsequent selection on sodium dodecyl sulfate (SDS). This *tolC* cassette was subsequently deleted using a colicin selection. The same process was used to delete the gentamicin cassette. The resulting strain has resistance only to zeocin and carbenicillin. To select for a strain with improved doubling times and no auxotrophy, MAGE was performed five times with oligonucleotides that reverted the mutations listed in **Supplementary Table 7** to WT and the resulting population was plated onto M9 minimal media supplemented with biotin (since the strain lacks the *bioAB* locus). The chosen colony was C321.ΔA.2.109.F. The strain C321.2.109.F was created by first inserting a *tolC* cassette into the *prfA* native locus using recombineering and selecting for incorporation on an SDS containing LB plate. Then the *prfA* gene was inserted into its native locus by recombineering. Insertion was selected for using a colicin selection as referenced above. The C321.ΔB.2.109.F was created by first transforming the pBAD *prfB* plasmid, then deleting the native prfB locus by recombineering in a *tolC* selectable marker.

Whole Genome Sequencing

Whole genome DNA was isolated using the Qiagen DNeasy Blood and Tissue isolation kit. 100bp paired-end sequencing data was collected with a Hiseq 4000. Libraries were prepared by the Yale Center for Genome Analysis (YCGA). Paired end sequencing reads were first preprocessed using Trimmomatic with the parameters "LEADING:3 TRAILING:3 SLIDINGWINDOW:2:30", which removes low quality bases within the first or

last three bases of each read and truncates reads where the quality score drops below 30 over a two base pair sliding window[73]. Reads were then aligned to the C321.ΔA genome using BWA mem[74]. The AddOrReplaceReadGroups and MarkDuplicates Picard Tools were then used to modify the SAM files for GATK variant calling[75]. Variant calling was performed using GATK's HaplotypeCaller with ploidy 1 and stand_call_conf 30. The variants were filtered using GATK VariantFiltration with the filterExpression "QD < 2.0 || MQ < 40.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0" for SNPs and "QD < 2.0 || ReadPosRankSum < -20.0" for indels[76]. The potential impact of the filtered variants was then estimated using SnpEff with the following tags: -canon -no-downstream -no-upstream -noLof[77]. The full results are in **Supplementary Table 6**. This table includes both off-target and target UGA recoding events.

Mutation of *dnaX*

Five Rounds of MAGE were performed in EcnR2 with a mixture of oligos. To attempt to convert the UGA codon to UAA, a mixture of dnaXTAA1, dnaXTAA2, dnaXTAA3 was used oligo. This was done to prevent any bias caused by a particular oligo design. To randomize the UGA codon the dnamix1 and dnamix2 oligonucleotides were utilized. Populations were plated and 48 colonies picked and screened using MASC PCR and the dnaXTGAwt and dnaXTGArev for each experiment. Colonies that produced no band had their dnaX locus sequenced via Sanger sequencing. Doubling times for all 48 colonies were obtained.

Doubling Times

Growth profiles were obtained in a BioTek Synergy HT at 34°C with fast linear shaking and reads every 10 minutes for 16 hours. Doubling times were obtained with a computational script that first takes the natural log of the growth[27,32]. It then takes a linear regression of the growth profile across a moving window and reports the steepest slope.

Mass spectrometry

The observed peptides are listed in **Supplementary Table 8**. Samples were prepped, digested, and analyzed as previously outlined elsewhere[24].

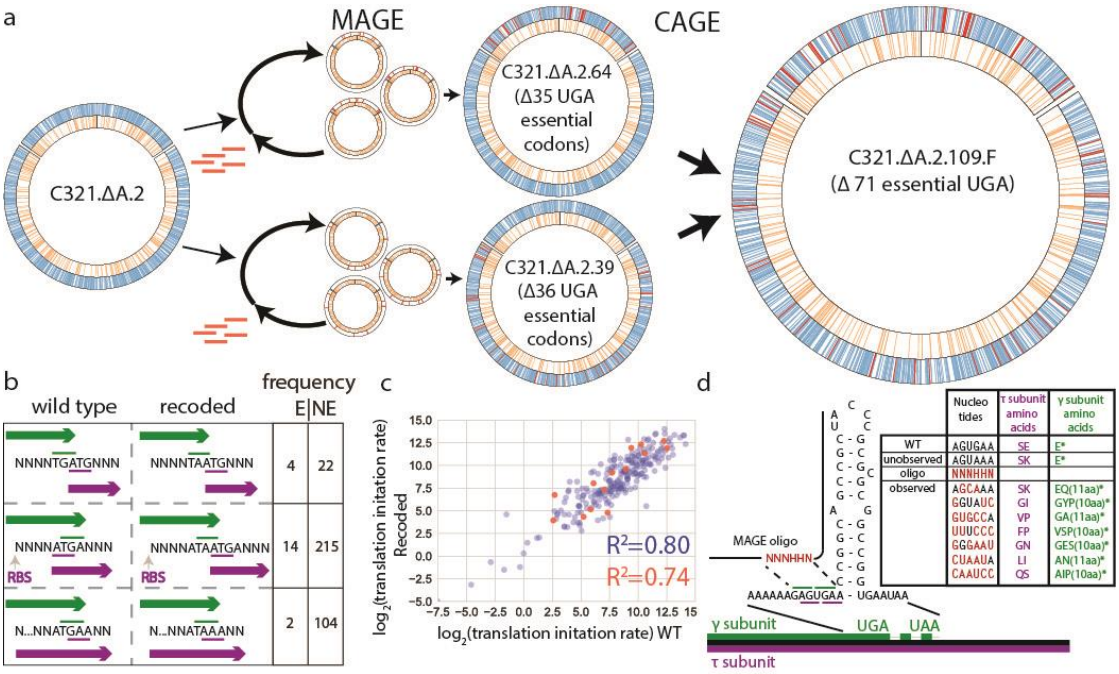


Figure 1. Design of UGA genomic recoding. A) Overview of the locations of the recoding and strain assembly process. Leftmost genome shows the ancestor C321.ΔA.2 strain. The two middle genomes show the strains resulting from MAGE to recode UGA to UAA. The rightmost genome shows the final UGA essential recoded strain. Previously recoded UAG codons are shown gold in the inner circle. Non-essential UGA codons are shown in blue while essential UGA codons are shown in red in the outer circle. B) Types and frequencies of UGA codon overlaps with other genes. Green genes represent the gene containing a UGA codon, while the purple gene is the overlapped gene. E stands for essential genes, NE stands for Non-Essential genes. C) Predicted translation initiation rate for wild type genes vs the recoded designs shown in C for all genes ending in UGA. Essential genes are shown in orange while non-essential genes are shown in blue. The diagonal corresponds to recoding events that do not alter the expected translation rate. R^2 is the Pearson correlation coefficient. D) *dnaX* subunit overlap caused by a programmed frameshift. The positions randomized by the MAGE oligonucleotide are shown in red. The reading frame of the γ subunit is shown by green lines over 3 nucleotides, while the τ subunit reading frame is shown by lines under 3 nucleotides. The table shows both observed and not observed dnaX sequences and the corresponding amino acid sequences for both subunits.

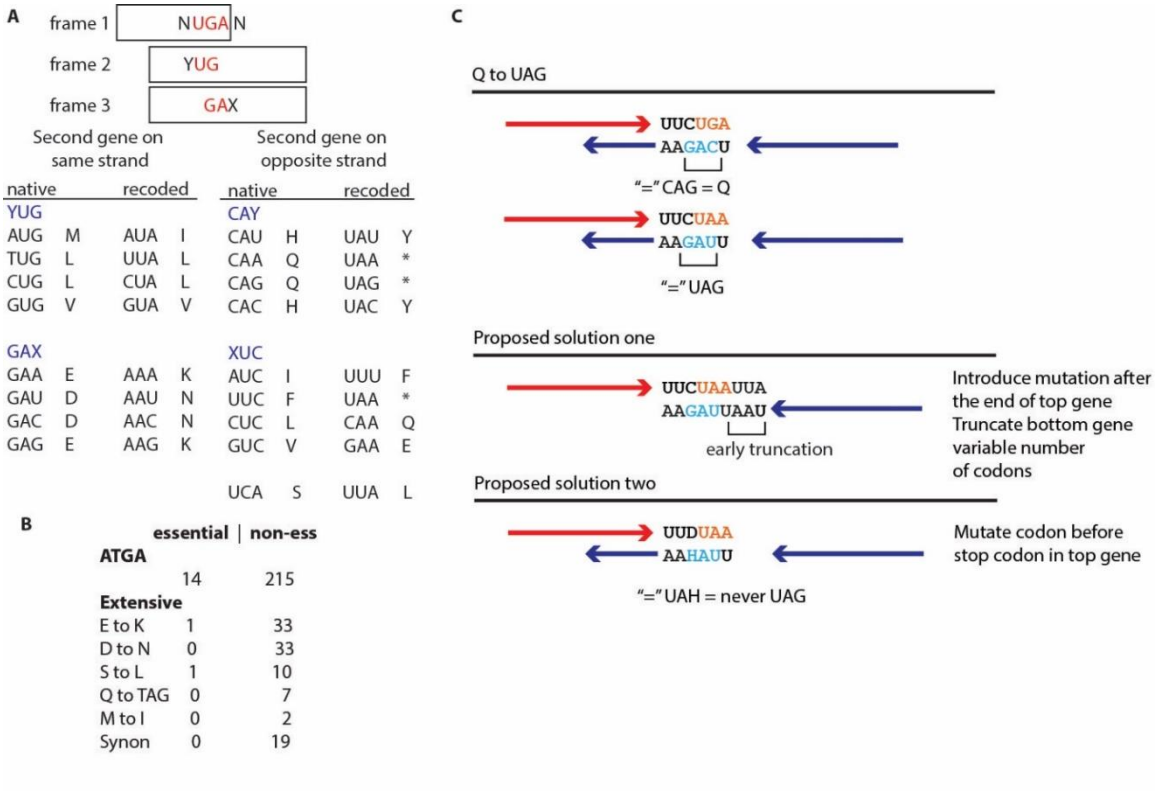


Figure 2. All possible mutations caused by UGA to UAA recoding A) All possible mutations when a second gene extensively overlaps the UGA codon recoded to UAA. B) The number of such mutations caused by recoding in essential genes and in all non-essential genes. C) The mutation of a CAG codon in the second overlapping gene leads to insertion of a UAG codon which is an undesirable outcome. There are three proposed solutions that cause varying amounts of disruption to gene 1 and gene 2 to avoid introduction of the UAG codon. In proposed solution one, the UGA containing gene has a UUA insertion placed after the recoded UGA codon. In the overlapped gene this is a UAA codon which truncates the gene before the UAG codon is reached. This prevents introducing the UAG codon into the overlapped gene. In the second scenario, a mutation to the codon preceding the recoded UGA codon in the top gene prevents the UAG codon being formed in the overlapped gene

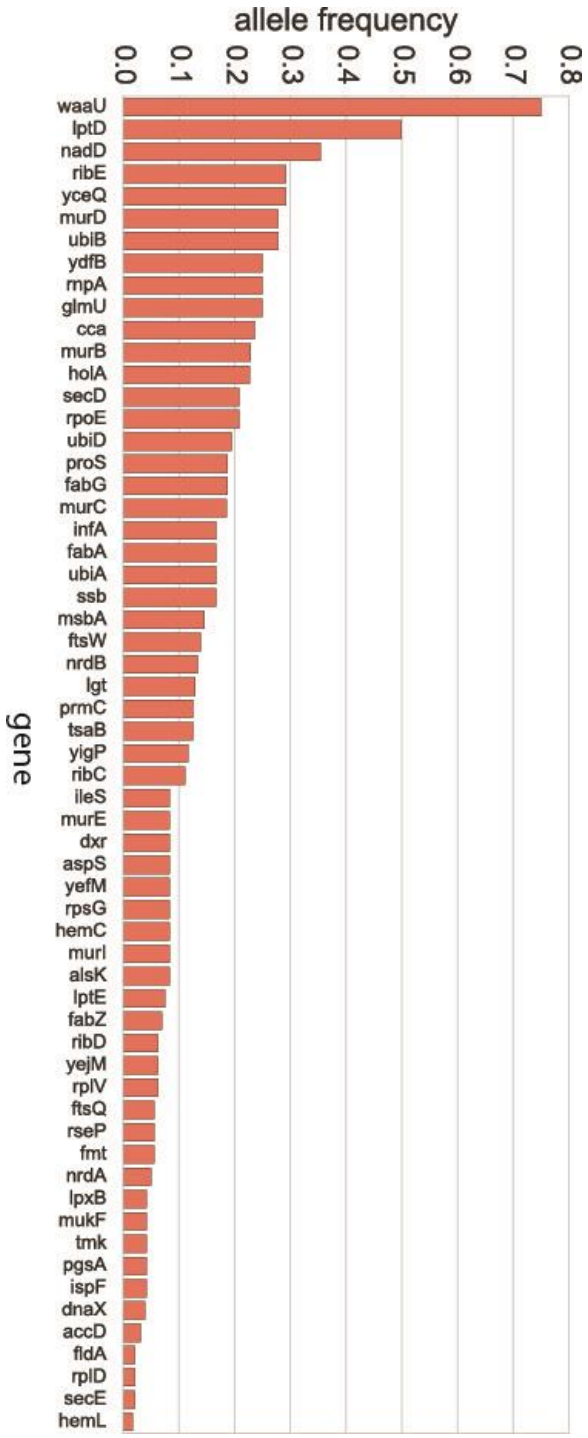


Figure 3. Allele frequency of conversion for a subset of converted essential UGA codons. About 14% of the essential genes were converted and fixed in the population before screening. Therefore, they moved from no conversions, to all colonies showing the UGA converted to UGA. The conversion is shown for 10 rounds of MAGE on average, although the exact number of rounds of MAGE before these measurements varies.

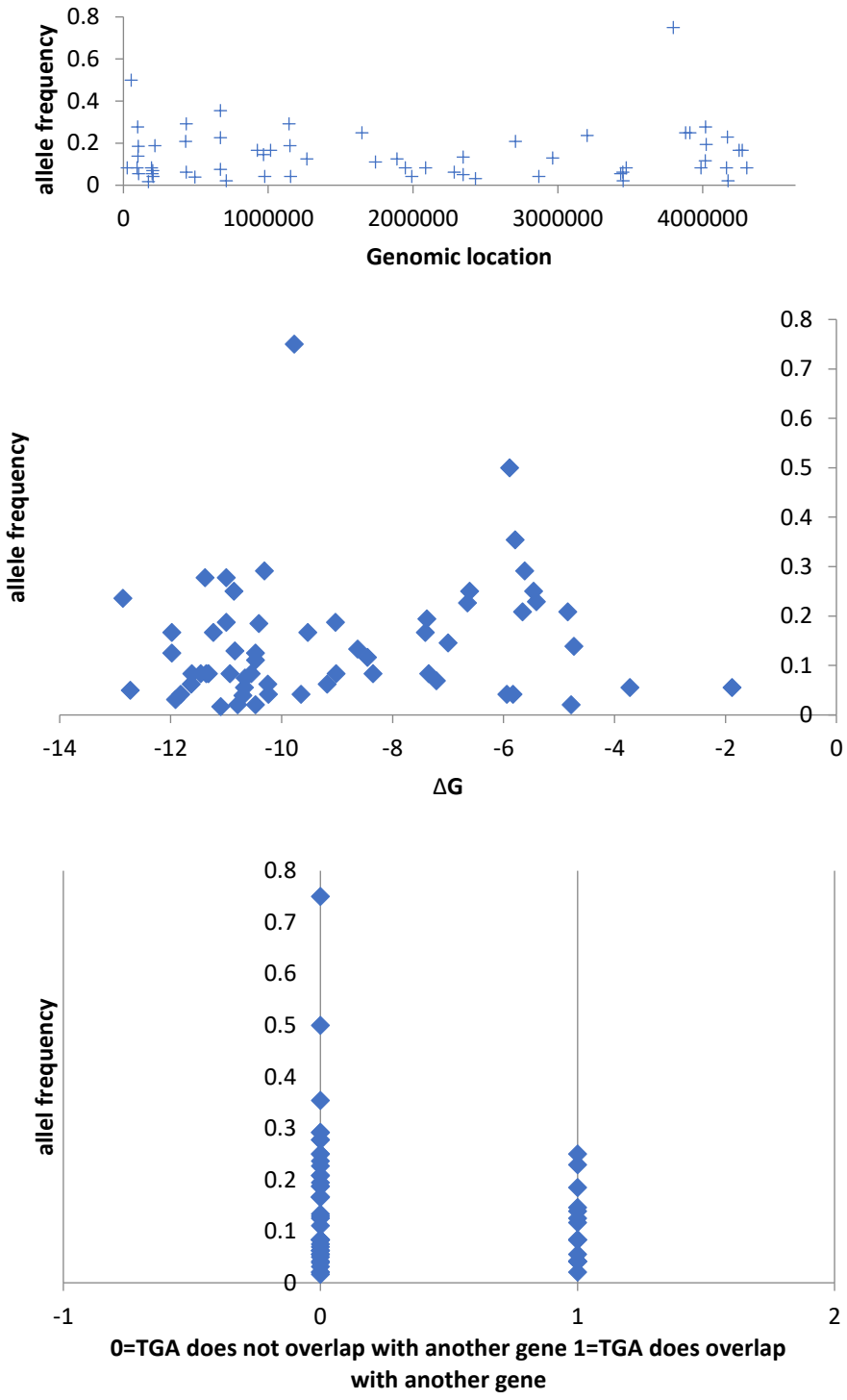


Figure 4. Allele frequency of conversion is not explained by several variables. A) allele frequency against genomic location. B) allele frequency vs Gibbs free energy of folding for the mutagenic oligo. C) allele frequency for oligos targeting UGS codons overlapping other genes and for UGA codons not overlapping other genes.

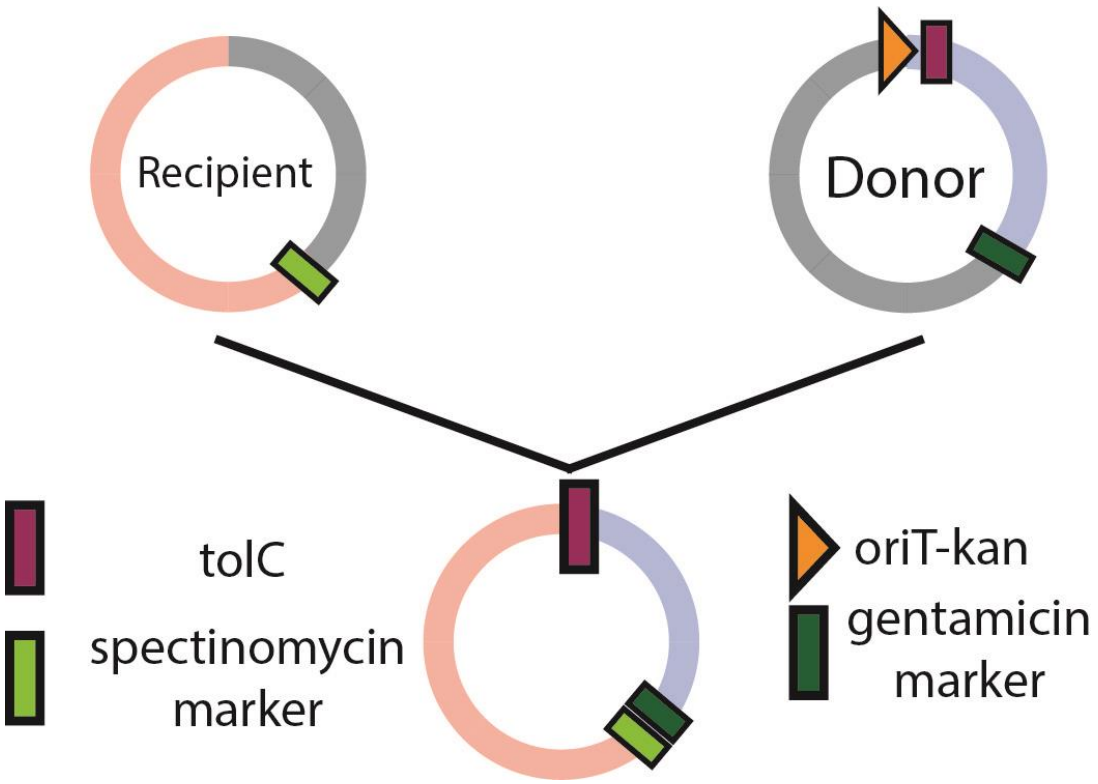


Figure 5. Outline of the CAGE process and the necessary selectable markers. The recipient strain(C321.ΔA.2.39) contains a single spectinomycin selectable marker which specifies the maximum extent of the genome that is transferred. The donor strain(C321.ΔA.2.64) contains a *oriTKan* and *tolC* marker that specify the beginning of the region to be transferred. The gentamicin marker in the donor strain that specifies the minimum extent of transfer.

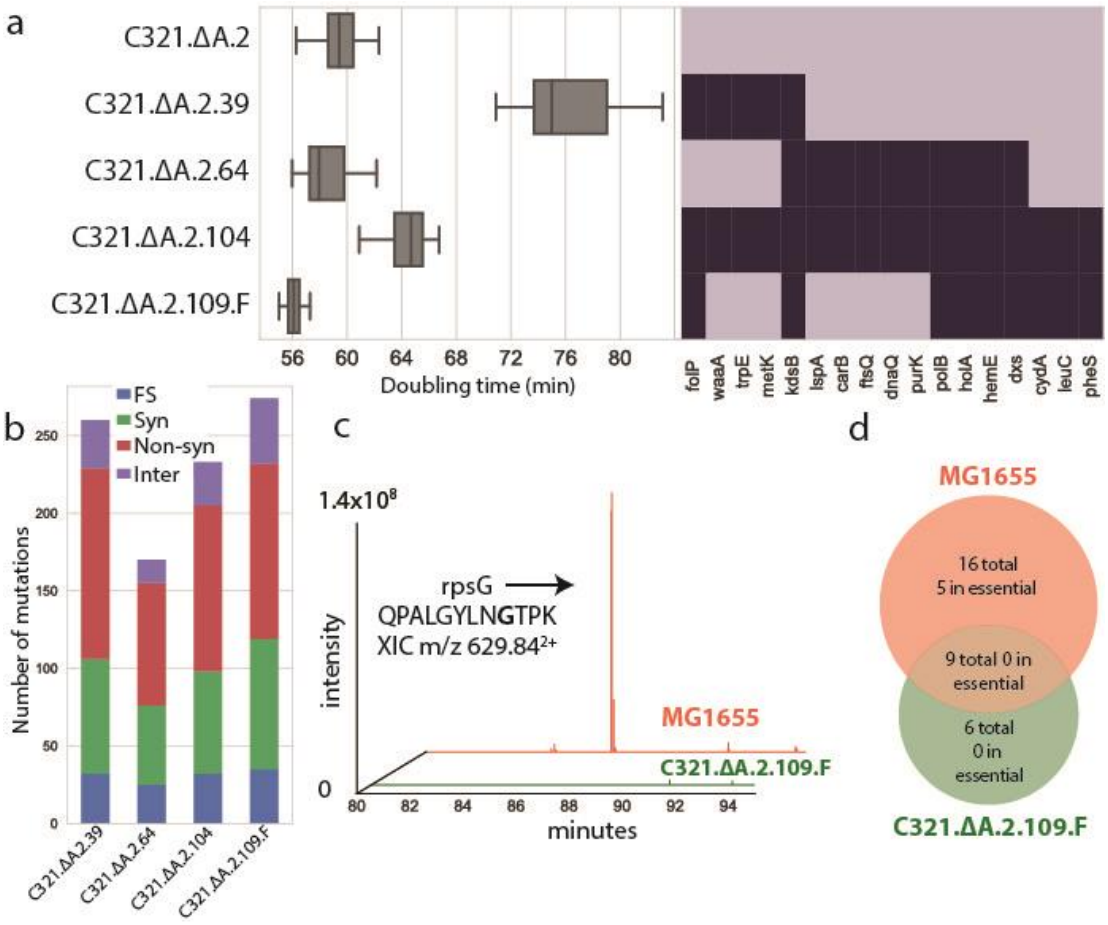


Figure 6. Genotypic and phenotypic characterization of UGA recoded strains. A) Growth of partially and fully UGA essential gene recoded strains in rich media and the off-target mutations in essential genes of each strain. Dark shading indicates the mutation is present, while light shading indicates it is absent. B) Off-target mutations observed by whole genome sequencing. Frameshift (FS), synonymous (Syn) and non-synonymous (Non-syn) mutations occur in coding regions of genes, while intergenic (inter) means the mutations falls between coding regions. c) Extracted ion chromatograph of a peptide showing glycine suppression of the UGA codon of the essential *rpsG* gene by the *glyT* suppressor in MG1655 (red) but not the UGA essential recoded strain C321.ΔA.2.109.F (green). d) Venn diagram providing a summary by strain of the peptides with a UGA suppression event observed by mass spectrometry.

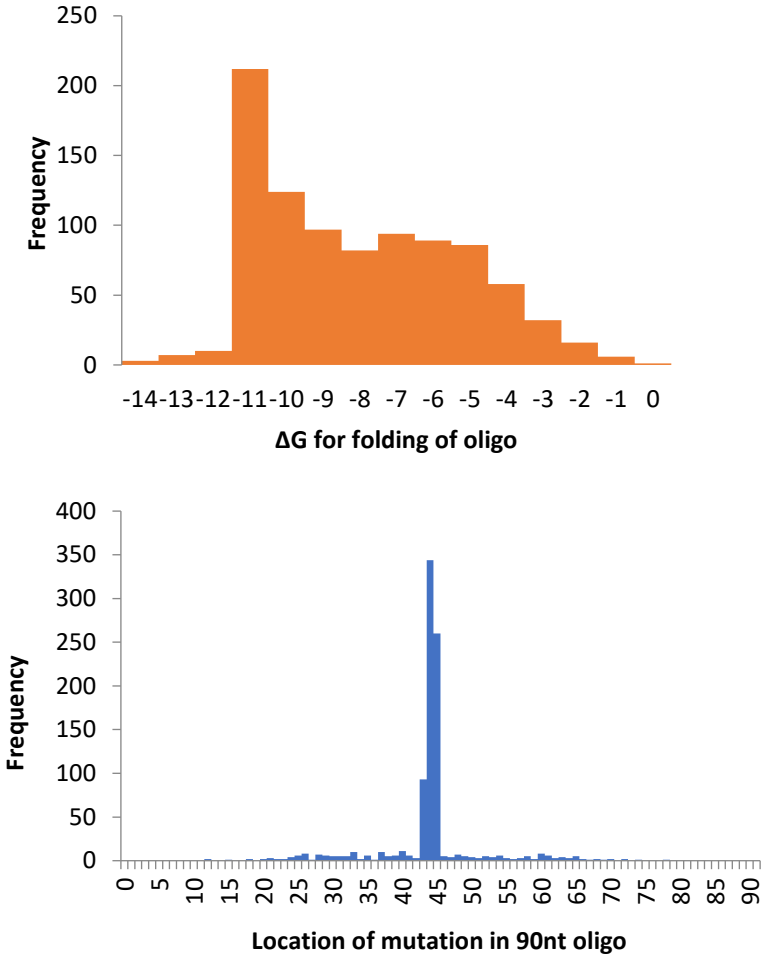


Figure 7. Parameters for design of mutagenic ssDNA oligos. a) Distribution of the free energy of folding for all ssDNA oligos. The more negative the value, the more likely the oligo is to form inhibitory secondary structure. b) Distribution of the position within the 90nt oligo of the mutation. At least 12nt of homology is required on both 5' and 3' end of the mutation for efficient incorporation.

Chapter 4

Engineering a release factor with activity at only UAA codons

*This chapter has segments adapted from the manuscript for: DW Moonan, JR Patel, J Moen, Muir P, J Tordoff, J Rinehart, FJ Isaacs, Genomic recoding of the Escherichia coli essential genes terminating in a UGA codon, submitted to Nature Chemical Biology, 2018,

Since the peptide release factors provide recognition of stop codons, removal of UGA codons from the genome does not affect the translational meaning of the UGA codon. In order to obtain an *E. coli* strain with two free codons (ΔUAG, ΔUGA), the release factors that recognize these codons must be altered. First, a mutant release factor that recognizes only the UAA codon was rationally designed. Second, this release factor was assayed *in vitro* and *in vivo*. Finally, we investigated the feasibility of this mutant release factor displacing the native release factors in a doubly recoded strain.

The first step in rationally designing a mutant release factor that recognizes only the UAA codon was to investigate the specificities of natural release factors. In most known prokaryotes, RF1 terminates translation at UAG and UAA, and RF2 terminates translation at UGA and UAA[78,79]. If there exists a known natural release factor that terminates at only UAA, it might be possible to import that release factor into *E. coli*. This release factor would likely have to be prokaryotic in origin due to a major split in release factors exists between eukaryotic and prokaryotic release factors. While having the same peptide release functionality, they have distinct primary sequence and tertiary structure[80], making it unlikely that a eukaryotic release factor ported into a bacteria would function properly.

If directly utilizing such a protein in *E. coli* was impractical, the cause of the altered specificity could be investigated. Finally, if no known UAA specific release factor exists, release factors with different specificities could be investigated. The differences between these release factors with altered specificities might give clues as to which

residues or domains must stay invariant to maintain function and which residues or domains are likely to alter specificity.

While I was unable to locate any release factor specific to only UAA, there are natural systems that demonstrate release factors with altered discrimination of stop codons. While most eukaryotes have a single release factor eRF1 that terminates at UAA, UGA and UAG, some ciliates possess an eRF1 that terminate at only UGA codons[81,82].

An alternative to searching for organisms where the specificities of their release factors has been determined is to investigate organisms whose genetic code varies such that UAA/UAG/UGA likely code for a sense codon[2,83,84]. The RFs of these organisms might have altered specificities to accommodate UAA/UAG/UGA coding for an amino acid. Alternatively, since prokaryotes have release factors with overlapping specificities, one of the release factors could simply be deleted.

Certain bacteria are known to use UGA for Trp or Gly [84,85]. This occurs in both *Mycoplasma* and in uncultured bacteria from the gut[2,85]. In many *Mycoplasma*, this recoding is accompanied by the deletion of RF2 instead of a reassignment of specificity[86-88]. Unfortunately, the uncultured bacteria were discovered from sequencing of the gut microbiota which did not allow for the assembly of a complete genome. The raw reads do not contain a protein with homology to RF2. Thus, the change in specificity likely stems from a deletion of RF2, although it is possible that an RF2 with altered specificity exists in the genome but was not sequenced.

While prokaryotes do not have much variation in the genetic code, mitochondrial genomes do[89]. Since these organelles are derived from prokaryotes, the RF's used by the mitochondrial translation machinery shows significant homology to prokaryotic RF[89]. In fact, the recognition domain of mitochondrial RF can be swapped with the prokaryotic domain and still maintain function[90,91]. Sadly, in many cases where the stop codons vary from the canonical code, no protein with homology to either RF is known. This is because only the mitochondria genome has been sequenced in these organisms. The mitochondrial RFs are often encoded in the nuclear genome, which in these organisms has not been sequenced.

Since no known release factors specific to only UAA were found, we chose instead to rationally engineer a release factor. The rational design of mutant proteins requires an extensive knowledge of the native protein. In prokaryotes RF1 terminates translation at UAG and UAA, and RF2 terminates translation at UGA and UAA, although mutations are known that expand the specificity of RF2[78,79]. Thus, either RF1 or RF2 could serve as the starting point for engineering of a release factor with specificity exclusive for UAA. Given the availability of crystal structures[92,93] showing RF1 bound to UAG and UAA codons, we attempted to rationally design a variant that no longer recognizes the UAG codon. RF1 and RF2 have a distinct domain for recognizing a stop codon (recognition domain), a distinct catalysis domain and a few other domains through which recognition of a codon is transmitted to the catalytic domain[7,78,94]. While mutations in these auxiliary domains have been shown to affect stop codon specificity, there are no clear patterns arising from these mutational studies[79]. Thus, we

hypothesized that the best way to create the desired mutant RF is through mutations to the recognition domain.

When RF1 is bound to a UAG codon, the Q185 residue of the recognition domain of RF1 acts as a hydrogen bond donor to the carbonyl group of the final nucleotide of the UAG codon. When RF1 is bound to UAA, this carbonyl is replaced with an amino group. Thus, no hydrogen bond is formed, and the Q185 residue flips orientation (**Fig. 1c**). We hypothesized that a Q185E mutation would remove the hydrogen bond to the UAG codon isosterically, thus limiting recognition of UAG codons.

Rational design does not always produce proteins with the intended function. Therefore, any mutants must be carefully assayed for activity at the three stop codons (UAA, UAG, UGA). We utilized two assays, both of which derive from the idea that without a release factor to recognize a stop codon, the ribosome should stall. Thus, a fluorescent reporter ending in a stop codon not recognized by a release factor should lead to ribosomal stalling and low fluorescence production. If instead the reporter ends in a stop codon recognized by a release factor, normal protein production should result in high fluorescence.

To test the release factor activity *in vivo* an mCherry construct terminating in UAG was transformed into C321.ΔA.Δlambda with either wildtype or Q185E RF1. Wildtype RF1 terminates translation at UAG and produces significant mCherry signal (**Fig. 1d**). In contrast, if Q185E RF1 indeed fails to terminate translation at UAG codons, ribosomal stalling at UAG and subsequent tmRNA dependent degradation should lead to

minimal mCherry signal[95]. As expected, the mCherry fluorescence in a strain with Q185E RF1 is comparable to a strain lacking RF1, but is eight times less than the fluorescence in a strain with native RF1. Thus, these data indicate that Q185E RF1 does not terminate at UAG codons and establishes a promising mutant release factor for a GRO with UAA as the single stop codon.

While this assay proves Q185E RF1 no longer terminates at UAG codons, the Q185E mutation might simply destroy function at both UAG and UAA instead of changing specificity to only UAA. Release factor activity at only UAA can't be assayed *in vivo* because of the essentiality of *prfB* in wildtype cells[45]. While *prfB* is present, any termination at UGA or UAA codons in an assay is likely due to the *prfB* gene. If a prokaryotic release factor specific to only UGA existed, then a mutant release factor's ability to complement a *prfB* knockout could potentially give insights into the mutant release factor's ability to terminate at UGA codons. However, no such prokaryotic UGA specific release factor is known.

To assay mutant release factor activity at UAA codons, an *in vitro* assay was utilized (**Fig. 1e**). A purified transcription translation kit was utilized to allow the introduction of a single release factor[96]. Just like the *in vivo* assay, the *in vitro* assay relies on ribosomal stalling at unrecognized stop codons. A fluorescent reporter ending in a stop codon not recognized by a release factor should lead to ribosomal stalling which should inhibit production of more fluorescent protein. Even the GFP on the stalled ribosomes should not lead to fluorescence because key residues should stay sequestered in the ribosome exit channel[97]. If instead the reporter ends in a stop codon

recognized by a release factor, normal protein production should result in high fluorescence.

As expected from the *in vivo* assay, the Q185E RF1 mutant does not release at UAG codons. This is demonstrated by release at UAG codons for Q185E equaling the background release of UGA codons by RF1 (RF1 does not recognize UGA). More importantly, Q185E maintains activity at UAA, albeit at a slightly reduced level compared to wildtype RF1. The physiological importance of this difference will eventually be assayed in a strain capable of deleting RF2. Together, the *in vitro* and in *vivo assays* show that the Q185E RF1 mutant release factor has specificity towards only UAA codons.

The final step in creating a mutant release factor that provides the translation component necessary for an *E. coli* strain with two free codons (ΔUAG, ΔUGA) is proving it is capable of solely supporting growth. To test this, *prfB* must first be deleted. Normally this is not possible due to the need to terminate translation at both UGA and UAA codons. However, the C321.2.109.F has removed the UGA codons in essential genes. This might attenuate the need to terminate at UGA codons in this strain. The successful removal of RF1 in several genetic backgrounds that still had UAG codons provided inspiration[26,67,98,99]

To test this hypothesis, we sought to remove native recognition of the UGA codon by removing RF2 in an essential UGA recoded strain (C321.2.109.F) with RF1 reinserted in order to act as the sole release factor for the remaining UAA codons. Since

an amber suppressor was required in the removal of RF1 in a UAG essential gene recoded background, we hypothesized that *glyT* may help alleviate ribosomal stalling at the remaining 1,069 UGA codons[26,67]. Attempts to replace the native *prfB* gene with *tolC* or a kanamycin resistance marker were unsuccessful in the presence or absence of the *glyT*. However, the native *prfB* gene was easily deleted when it was complemented with a plasmid-borne *prfB* gene under the arabinose inducible pBAD promoter (**Fig. 1a**). This strain failed to grow in the absence of the inducer, even in the presence of the *glyT* suppressor, although escape colonies were isolated (**Fig. 1b**). Two such colonies were sequenced and both showed mutations in the *araC* gene—one of which was previously characterized[100] - that allowed for constitutive expression of the *prfB* gene (**Table 1**). We hypothesize that, in the absence of RF2, ribosomal stalling at the remaining 1,069 UGA codons represents too great a burden on the translational machinery for the partially recoded strain to survive and that further recoding is necessary to delete RF2.

The continued essentiality of the *prfB* gene in the UGA essential recoded strain highlights differences between RF1 and RF2. Since eukaryotes have only a single release factor for all stop codons, the split in release activity between two proteins in prokaryotes might indicate the importance of these differences. Previous studies have been able to isolate RF2 mutants that terminate at all three stop codons and complement *prfA* and *prfB* knockouts[78,79,101]. Those same studies failed to produce an RF1 mutant that complemented deletion of *prfB*, despite *in vitro* assays indicating some RF1 mutants terminate at all three stop codons. Thus, RF2 may have additional functionalities beyond translation termination that render it more essential than RF1.

Two additional activities beyond translation termination have been ascribed to RF2 but not RF1. First, after ribosomal stalling on a non-stop or truncated mRNA, RF2 can pair with the ribosomal rescue factor *arfA* to terminate translation at any codon[102]. Deletion of *arfA* does not impair growth in rich media, and thus RF2's pairing with *arfA* is unlikely to be responsible for the essentiality of *prfB*. Second, RF2 appears to exhibit translational proofreading activity *in vitro* by terminating translation after the incorporation of an incorrect amino acid[103,104]. In addition, domain swapping between RF1 and RF2 obtained a chimera that could successfully complement both RF1 and RF2 deletions[94]. This technique may allow for the localization of the domain(s) responsible for the additional functions of RF2.

If this additional function is not responsible for *prfB*'s continued essentiality, RF1's contribution to termination at UAA codons may not sufficiently complement RF2's. However, the K12 strain already has an impaired RF2 caused by a A246T mutation, shifting more burden onto RF1[101,105,106]. This mutation causes a 3 to 4-fold decrease in RF2 activity but it does not impair fitness on rich media, suggesting that RF1 activity at UAA is not limiting[90,107]. Further evidence is provided by growth of K12 or BL21 strains in a release factor targeted antibiotic[108]. In the K12 strain with impaired RF2 function, antibiotic resistance mutations appear in the RF1 gene, suggesting that the major source of lethality stems from impairing RF1 function. In BL21 however, the resistance mutations appear in RF2, suggesting it plays a more central role when it has greater activity. This suggests RF1 already complements an impairment of RF2 activity in the K12 strain.

Instead, we hypothesize that ribosomal stalling at remaining 1069 UGA codons represents too great a burden on the translational machinery for the cell to survive in the absence of RF2. It remains unclear what level of recoding would sufficiently alleviate this burden, or if there are specific non-essential genes that require recoding.

Taken together, this translational factor engineering establishes important results and a clear path toward a mutant release factor with UAA as the single stop codon.

The Q185E mutant could be altering codon specificity by altering binding to the stop codon or changing the rate of release once the codon is bound. Both steps could consist of several fundamental steps. The kinetics of release for RF1 at 12 codons and RF2 at 13 codons have been determined previously[109]. Release at UGA had a k_{cat} roughly 50 times slower than at UAA. The K_m of release was also roughly 2000-fold higher. If K_m is used as a proxy for K_d, this would indicate that binding to the stop codon is the predominant discriminatory step. Later studies investigated both the K_d and kinetics of binding to stop and sense codons by RF1[110]. While only UAA, CAA, UUU and CUC were investigated, the binding constant to UAA was roughly a 1000-fold lower than at the sense codons, due mostly to differences in the off rate.

Mutants to the recognition domain of RF1 have shown an effect in binding to stop codons. Specifically, mutation of H197 - involved in binding to the second codon – to alanine causes a 1000-fold reduction in binding affinity, due almost entirely to an increase in the off rate[111]. Interestingly, they observed a biphasic binding profile,

suggesting an initial rapid binding to the codon followed by a slow conformational change to a conformation capable of release. The rate of this change was roughly equivalent to the rate of release, suggesting that this conformational change may be the rate limiting step of release. This conformational change is supported by recent structural data showing RF1 in a pre-release complex[112-114]. A RF1 mutant lacking the switch loop residues 302-304 was captured in this pre-release complex by X-ray crystallography. The increase in specificity against sense codons shown by this mutant is presumably due to a decreased rate of the post-binding rearrangement. This slowed rearrangement also likely explains why this mutant could be captured in the pre-release conformation.

Mutation of Q185 to alanine caused a 20-fold increase in the dissociation constant at UAA[115], but did not alter the biphasic nature of binding nor show a decrease in catalysis under saturating conditions. This suggests that the Q185A mutation alters initial binding but does not impact the subsequent conformational rearrangement or catalysis. Given the similarities between the Q185A and Q185A mutations, I hypothesize that the Q185E mutation also alters initial binding but does not impact the subsequent conformational rearrangement or catalysis. The Q185A mutation likely abolishes favorable interactions with both adenine and guanine at the third position, whereas our Q185E could conceivably maintain a favorable interaction with adenine, leading to an alteration of specificity. This is supported by our results that mutation of the Q185 residue does not destroy catalytic activity but does alter specificity.

In a separate study, the effects of altering the bases of the stop codons were investigated[116]. Altering the final base to purine removed all exocyclic functional groups from the base. This has the effect of removing the hydrogen bond between the Q185 reside and the final base. This altered base increases the dissociation constant 30-fold in comparison to the constant for UAA, but does not alter the rate of catalysis under saturating conditions. This agrees with the Q185A experiment. In addition, replacement of the final adenine with inosine only reduced activity of RF1 to about 70% of wildtype. This indicates that amino group of guanine is not very important for the binding of RF1. Thus, our focus the focus on the differential amino/carbonyl group of A/G is further validated by this result.

Methods

Measurement of Release Activity at TAG codon *in vivo*

C321.ΔA.Δlambda was utilized because C321.ΔA normally has an ampicillin resistance cassette next to the lambda cassette, but the reporter plasmid also has an ampicillin resistance cassette. C321.ΔA.Δlambda was created previously using lambda mediated recombineering by replacing the lambda cassette with the native biotin synthesis genes, then selecting on minimal media that lacked supplemental biotin. In doing so the genomic *tetR* copy that is present in C321.ΔA was also removed. Therefore, Tet based promoters are constitutive in this strain. The resulting strain was transformed with the mCherry-TAG plasmid (mCherry terminating in a TAG codon driven by a pLtet promoter, modified from[24]) and successful transformants were selected on carbenicillin. These were then transformed with a pBAD *prfA* plasmid (WT or Q185E *prfA* driven by pBAD promoter) and plated onto carb and kanamycin to select for transformants. The strains were grown in LB media supplemented with carb and kan in a 96 well plate at 34°C for 16 hours with an individual well volume of 145µL. A time course of mCherry fluorescence was taken, then at 16 hours the cells were twice washed in PBS before resuspending in PBS and taking a final fluorescence measurement. This measurement after washing did not appreciably differ from mCherry levels measured in LB. Fluorescent measurements were taken in BioTek Synergy HT at 34°C with a 590/20 filter for excitation and 645/40 filter for emission.

Measurement of Release Activity at TAG codon *in vitro*

PURExpress® Δ RF123 from New England Biolabs was utilized as an *in vitro* transcription, translation kit. Manufacturer's instructions were used with the following additions and exceptions. Optional RNase inhibitor was utilized as per the manufacturer's instructions. To the reaction was added 65ng of linear Release Factor template and 30ng of linear sfGFP template (**Sequences, Appendix C**). Release factor is initially produced from this template due to low levels of release, as detailed by the manufacturer and in previous reports[96]. No purified Release Factor (1,2,3) was added to the reaction. Reactions were run in a 384 well plate at a final volume of 10µL for 3 hours. Plates were sealed with a clear, impermeable film to prevent evaporation of the reaction. Fluorescent measurements were taken in a BioTek Synergy HT.

Deletion of RF2 in C321.ΔB.2.109.F

The C321.ΔB.2.109.F was created by first transforming the pBAD *prfB* plasmid, then deleting the native prfB locus by recombineering in a *tolC* selectable marker. The ability of the cells to grow in the absence of arabinose was tested with several conditions. LB plates were created that contained no arabinose, no arabinose with gentamicin, arabinose, arabinose with gentamicin. For each strain to be tested a ten-fold dilution series from 1x to 10,000x was created and spotted at 10µL onto the appropriate plate.

Sequencing of escape mutants from C321.ΔB.2.109.F

To obtain escape colonies, 50µL of the appropriate strain was plated onto plates lacking arabinose. A colony was then picked into LB lacking arabinose. The genomic DNA was

then isolated as described previously in this thesis. Whole genome sequencing was then performed, as described previously

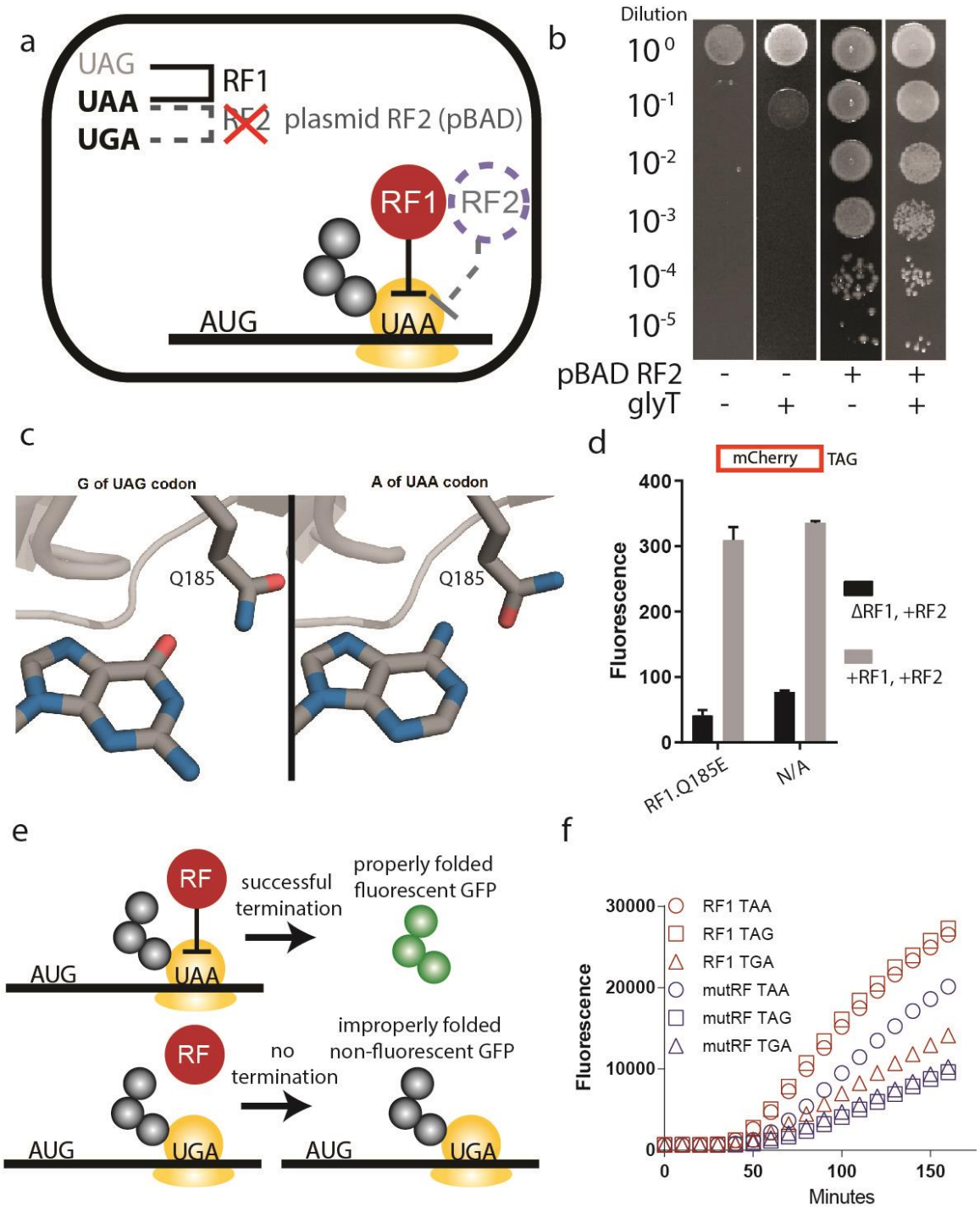


Figure 1. Engineering of mutant release factor 1. A) Schematic of the strain C321.ΔB.2.109.F used to test the feasibility of deletion of *prfB*. The genomic copy of *prfB* was complemented by a high copy number plasmid with *prfB* under an arabinose inducible promoter. B) The strain shown in A with or without a *glyT* was plated with or without the inducible promoter arabinose. C) Crystal structures of RF1 bound to UAG codon on the left and UAA codon on the right. Shown is the final nucleotide in the

codon and its interaction with the Q185 residue of RF1. D) *In vivo* assay of release factor activity at UAG codon tested in two genetic backgrounds (+RF1, +RF2 and ΔRF1, +RF2) with and without a mutant release factor (RF1.Q185E). Termination at UAG of mCherry leads to fluorescence, while lack of release leads to degradation. E) schematic of the *in vitro* release assay. A GFP transcript ending in one of the 3 stop codons is assayed against a release factor. If the release factor terminates translation at the stop codon, fluorescent GFP is produced. If the release factor does not recognize the stop codon, no release occurs, precluding proper maturation of GFP. F) Results from *in vitro* release assay for both WT and Q185E RF1(mutRF).

Table 1. Mutations in *prfB* knockout escape colonies

gene	mutation	escape mutant 1	escape mutant 2	Previously observed
araC	P11S	present	absent	doi: 10.1128/JB.01529-08.
araC	G141R	absent	present	

Chapter 5

Progress toward genomic recoding of all UGA codons in Escherichia coli

The inability to remove RF2 in a strain with only the essential UGA codons recoded reinforced the importance of recoding all of the UGA codons to UAA. Using the strain lacking essential UGA codons as a starting point (C321.ΔA.2.109.F), we began recoding or deleting the remaining 1069 UGA codons. Because of the ambitious nature of the project, a detailed plan was necessary.

Given the low cost of DNA synthesis associated with MAGE, this technique was chosen to convert the majority of the remaining 1069 UGA codons. The oligonucleotides to accomplish this were designed using the same procedure detailed earlier. Of the 1069 remaining codons, 75 UGA codons were targeted for deletion across 15 regions. The regions earmarked for deletion have all previously been deleted in a single strain without impacting the doubling time of the resulting MDS42 strain[64]. Not all of the deletions used in the MDS42 strain were targeted for deletion- only those that contain a high number of UGA codons **(Fig. 1**). In addition, one deletion (MD43) that contained 18 UGA codons was excluded, since this deletion targets a significant number of non-prophage genes of unknown function, and the deletion was not utilized in a number of follow up studies[117,118]. These regions will be deleted by inserting the *tolC* selectable marker[63,119] using dsDNA recombineering[120] before subsequently deleting the *tolC* using a colicin selection.

As with the essentials recoding effort, the MAGE oligos were grouped into pools. However, for full recoding, these pools consisted of 12 oligonucleotides each. These were screened using Multiplex Allele Specific (MASC) PCR primers that corresponded to each UGA codon. Each UGA codon would produce a PCR product of a defined size if it

was converted to UAA. For a pool of 12 oligos, the corresponding MASC pool would have a ladder ranging from 100bp to 1200bp with each amplicon corresponding to a specific UGA codon. However, because of the close proximity of many stop codons the primers that are supposed to pair with each other to form a defined amplicon could pair with adjacent pairs to create amplicons of unexpected sizes. This would ruin the interpretation of the experiment. To overcome this barrier, the initial MASC pools for most oligonucleotide pools did not screen every locus.

The full pool of oligonucleotides was used in MAGE. After approximately ten rounds the population was plated and individual colonies were screened with the limited MASC pool. This information, paired with phenotypic information on doubling time and growth in M9 minimal media, was used to determine the colony to carry forward to further rounds of MAGE. After eight months of continuous MAGE, the resulting strains were sequenced with whole genome sequencing. This confirmed MASC data, as well as informed us as to which UGA codons that were targeted for MAGE but not screened by MASC had been surreptitiously converted. From this point, new MASC PCR pools that corresponded to the previously unscreened (but not converted) UGA codons were created. Because only a few codons in each oligonucleotide pool were unscreened, these new MASC PCR pools spanned several oligonucleotide pools. The total number of UGA codons screened by these new primers amounted to 178.

The linear nature of recoding via MAGE would make recoding UGA in a single strain a decade-long undertaking. Therefore, we parallelized recoding across eight strains initially, then eventually twelve to quicken the pace of recoding. The strains

were initially split to maintain a roughly equal number of UGA codons in each strain. They were also split to ensure at least 10kb of space between the last UGA codon of one strain and the first UGA codon of the neighboring strain (**Table 1**). Both of these requirements were relaxed as the strains were expanded to twelve.

While recoding across multiple strains gives insight into the feasibility of recoding individual UGA codons, it doesn't investigate if all UGA codons can be removed simultaneously. In order to answer this question, we designed the CAGE process that will be used to stich the constituent strains into a single strain. A key step in the CAGE process is the placement of four markers: two positive selectable markers, a positive/negative selectable marker and the origin of transfer for the RK2 plasmid. The precise locations of these markers determine how the twelve constituent strains are reformed into a single strain. In the donor strain, the origin of transfer determines the point at which transfer begins, while the positive marker determines the minimum region of the genome transferred. In the recipient strain, the other positive marker delineates the maximum amount of the genome transferred, since if the donor portion of the genome was transferred past this point, the marker would be removed and the selection would fail.

Transferring the correct amount of the genome is a critical step. If the markers are placed such that too little of the genome is transferred, then UGA codons that had been converted in the donor strain will not be transferred to the recipient strain, necessitating the repeated conversion of those codons to UAA. If, the markers are placed so that too much of the genome is transferred, then some of the converted UGA

codons in the recipient strain will be overwritten with the wildtype UGA codons. This will again necessitate redundant work. The gaps shown in **Table 1** are those in which the markers must be placed. Otherwise some number of recoded UGA codons will be reverted back to wildtype.

Choosing the donor and recipient strains will drastically impact the speed of the CAGE process. There are two basic choices to conjugate neighboring regions (**Fig. 2**). The first has transfer from the donor strain directed towards the recipient strain. Because of the placement of the selectable markers, only a precise region will be transferred. While this will result in the desired strain, the selectable markers will be left in the junction between the two regions. Both of those markers must be subsequently removed, before markers can be placed for another round of conjugation. In the alternative arrangement, the donor strain has transfer directed away from the recipient region of the genome. Thus, in addition to the desired region of transfer a variable amount of the donor genome will be transferred. Since deleterious or favorable mutations could exist in the region not explicitly selected for, this technique may allow for the selection of conjugants with a favorable phenotype. More importantly, this arrangement results in the selectable markers remaining on the outer edges of the recipient and donor regions. These markers can then be converted directly to the markers necessary for the next round of CAGE.

Results

With the design determined, recoding commenced. Initial progress showed fairly consistent conversion of UGA codons across all eight strains, even after 16 rounds of MAGE (**Fig. 2**). In addition, the strains seemed to maintain fairly consistent doubling times, despite the proclivity of the strains to acquire deleterious mutations due to the *mutS* negative genotype. After 14 months when approximately 45% of the UGA codons had been converted, we sequenced the 8 strains with WGS to determine which unscreened UGA codons had been converted. Unfortunately, WGS showed that several strains had been cross-contaminated during the recoding process. Because both strain E and G had UGA codons converted in the region targeted in strain F, we concluded that strain F had contaminated both E and G. This happened partway through recoding, such that the conversions believed to be in a single strain were in fact spread across two strains. This split became the basis for E' and G'. H' and D' were also split at this time.

In these 12 strains, all but 75 of the 1178 UGA codons have been converted or deleted (**Fig. 3**). While most of the 1103 converted UGA codons have been recoded to UAA, 55 were deleted across 11 deletions. Even after ~40 rounds of screening, the strains have maintained a doubling time very close to that of their ancestor (**Fig. 4, Supplemental Fig. 1-6**). With the large majority of recoding across the 12 strains completed, the focus turned towards stitching the strains back together using CAGE. Although this might result in the final conjugated strain containing a few UGA codons, there is precedence with UAG recoding that suggests a fair number of unconverted codons can be tolerated without altering fitness[48]. In addition, the remaining UGA

codons could be converted via MAGE in the final strain, or they could be converted by MAGE during the CAGE process.

For the full CAGE process, markers will have to be placed 46 times at 24 different locations throughout the genome (since markers in subsequent rounds will displace those already inserted). Markers have been placed at all but 4 of those locations. In total, 23 marker placements have been achieved. Marker placement has been hindered by a decrease in recombineering efficiency. This is likely caused by a non-synonymous mutation in the *gam* gene involved in recombineering that occurred in the ancestor C321.ΔA.2.109.F strain. In addition, strain B which has failed to place either of two markers after many attempts across tens of locations has a frameshift mutation in the *exo* gene of the red operon. Reversion of both of these mutations to wildtype will hopefully reverse this phenotype.

Despite the difficulties encountered placing markers, enough markers were placed to allow several conjugations. The D and D', E and E', and H and H' conjugations have all been performed successfully (**Fig. 4**). In addition, the resulting EE' strain was successfully conjugated to the F strain. The resulting strain had roughly a third of all UGA codons recoded or deleted and a doubling time comparable to the ancestor strain. Given the success of all attempted conjugations and the likely identification of the cause for low marker recombineering, the main barriers to conjugation into a single strain have been overcome.

A potential barrier could be that the remaining unconverted UGA codons have some property that makes them resistant to recoding. We investigated several properties to determine if any properties correlated with the remaining UGA codons. Recoding UGA codons that overlap with another gene can impact the overlapped gene. This impact might increase with the number of nucleotides of overlap. A single nucleotide of overlap results in no change in the overlapped gene, but recoding UGA codons with three nucleotides of overlap likely impacts the translation initiation of the overlapped gene and recoding of UGA codons that overlap by more than three can lead to a non-synonymous mutation. However, the unconverted genes don't appear to be enriched in UGA codons with extensive overlap (**Fig. 5a**).

Changes in the overlapped genes might not be particularly deleterious if those overlapped genes are non-essential. If those genes are essential, alterations in translation initiation rate or non-synonymous mutations might have more of an impact. However, the unconverted UGA codons are not enriched for those that overlap with essential genes (**Fig. 5c**).

Changing UGA to a synonymous UAA could alter readthrough potential of the stop codon. To result in productive protein, this readthrough must eventually reach another stop codon. In addition, tandem codons have been ascribed biological importance[121,122]. If one of these stop codons is altered, it should impact the termination efficiency[123]. Therefore, the unconverted UGA codons may display an altered distribution of distance to the next stop codon. However, the distribution of

lengths to the next stop codon does not seem to vary between unconverted and converted UGA codons.

Since specific stop codons are enriched in highly expressed genes[124,125], the expression level of a gene and its stop codon may be important. In addition, alteration of a highly expressed gene could have more of an impact than expression of a moderately or lowly expressed gene. As shown in Figure 4d, the distribution of counts per cell for proteins with converted UGA codons vs proteins with unconverted UGA codons does show some difference. It is possible therefore, that UGA codons in highly expressed genes are more difficult to recode to UAA than typical UGA codons. But given the few UGA codons with proteomics data remaining, conversion of one or two can have profound effects on the distribution, casting some doubt on this conclusion. Regardless, because many highly expressed UGA codons have been converted to UAA, this insight does not provide reason to doubt the viability of recoding the remaining UGA codons.

UGA codons don't overlap with just other proteins. They overlap with various non-coding elements as well. UGA codons in *yaiC*, *cynX*, *arcB* and *tnaC* all overlap with annotated terminators[68]. Of these, only the UGA codon in *yaiC* has not been converted. There are no overlaps between UGA codons and any annotated riboswitches[68] and the boundaries of transcription factor binding sites are too ambiguous to effectively determine if they overlap with any UGA codons[126].

The final potential biological impediment to recoding UGA to UAA is selenocysteine incorporation[127]. Of the three selenocysteine incorporating proteins in *E. coli* (*fdhF*, *fdoG*, *fdnG*) only *fdhF* has had its UGA codon converted[128]. Conversion of UGA to UAA in these proteins results in a truncated protein. However, deletion of the selenocysteine incorporation machinery only appears to alter fitness in anaerobic conditions[45]. Thus, while it may be possible to convert the remaining selenocysteine UGA codons to UAA, we will pursue a strategy to convert these to UGC (cysteine) instead.

While MAGE has successfully created a thousand intended mutations, there have also been a number of unintended mutations (**Fig. 7**). These mutations do not appear to be enriched in any particular lineage, nor do they seem to be enriched for a particular type of mutation. Several interesting mutations have appeared in the strains. Two in particular that effect translation activity are a mutation to the RBS and start codon of *prmC* and a T246A mutation in *prfB* (which codes for RF2). The *prmC* mutation appeared in a common ancestor to all of the strains, while the *prfB* mutation appeared in strains E-H.

PrmC encodes a protein that methylates the release factors[129]. Methylation of the release factors significantly increases termination efficiency[107]. The *prmC* mutations are predicted to decrease its translation initiation 60-fold[60]. Thus, this mutation could lead to reduced methylation of RF2 and thus less active RF2. Since RF2 is the only release factor present in these strains, this could severely impact fitness. However, this mutation is present in all of the strains without any increase in doubling times indicating

either a lack of severity from the mutation, or some type of compensation. Because RF2 levels are autoregulated by RF2 termination activity, it is possible that more RF2 is produced than would be in the absence of the *prmC* mutations[130].

Interestingly, the T246A mutation in *prfB* that appeared in strains E through H has been shown previously to complement a *prmC* knockout[105]. The T246A mutation has also appeared when the C321 strain (UAG recoded) has been passaged over many generations[131]. This mutation increases the release efficiency of both release factors[107,132]. This continued mutation of genes relevant to translation termination likely points to some unobserved perturbations caused by the deletion of *prfA* in the C321 strain.

The recoding of over a thousand UGA codons to UAA represent an impressive achievement. While not all UGA codons have been recoded and those that have been recoded have not been assembled into a single strain, the path towards such an achievement is clear.

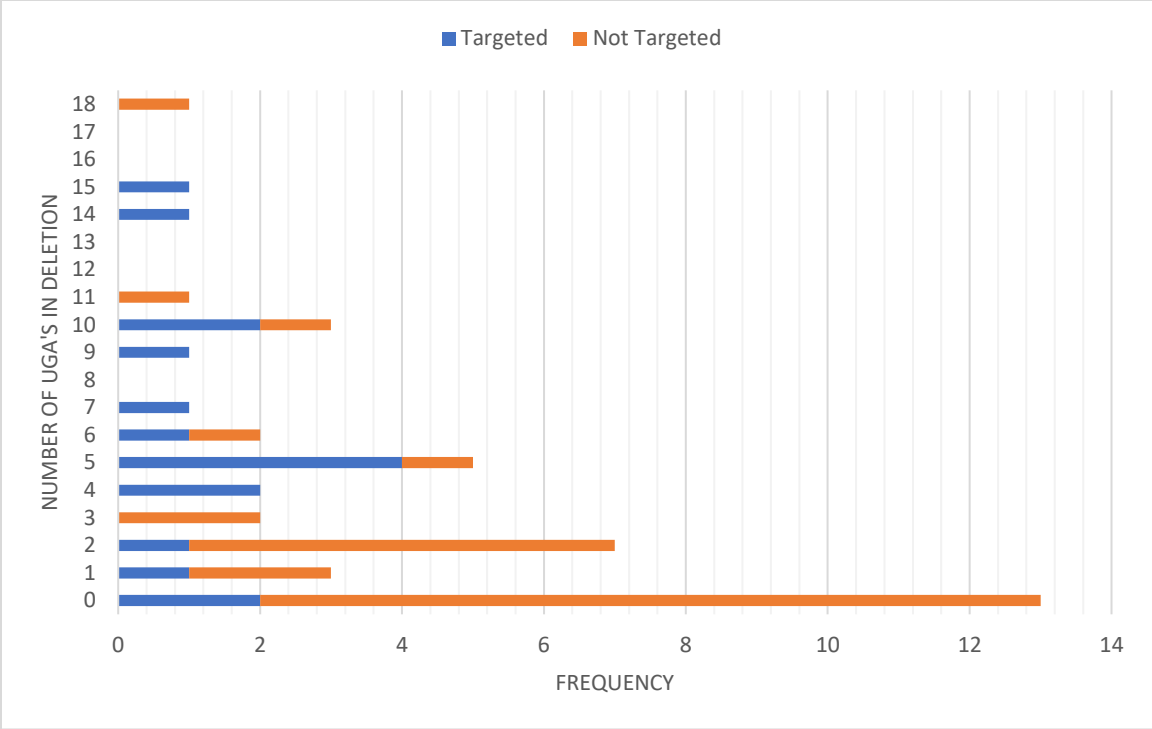


Figure 1. The distribution of UGA codons in the regions of the genome deleted in the MDS42 strain. Not targeted means that the deletion was not utilized in the full recoding of UGA, while those in blue were targeted for deletion

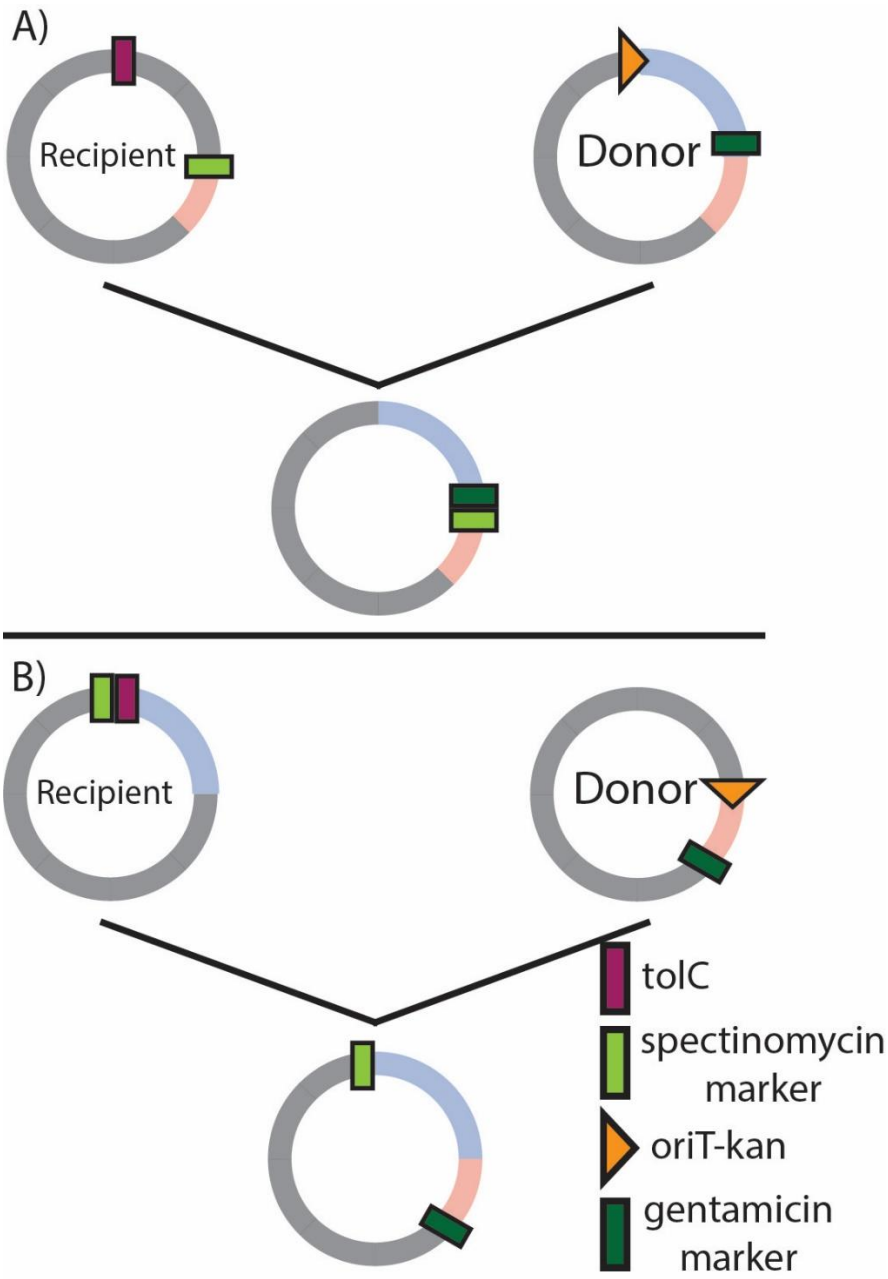


Figure 2. Two options for marker placement in CAGE process. A) Conjugation of the donor and recipient strains leaves both positive markers in the junction between the donor and recipient genomic regions. B) Conjugation of the donor and recipient strains leaves both positive markers on the edges of the donor and recipient genomic regions.

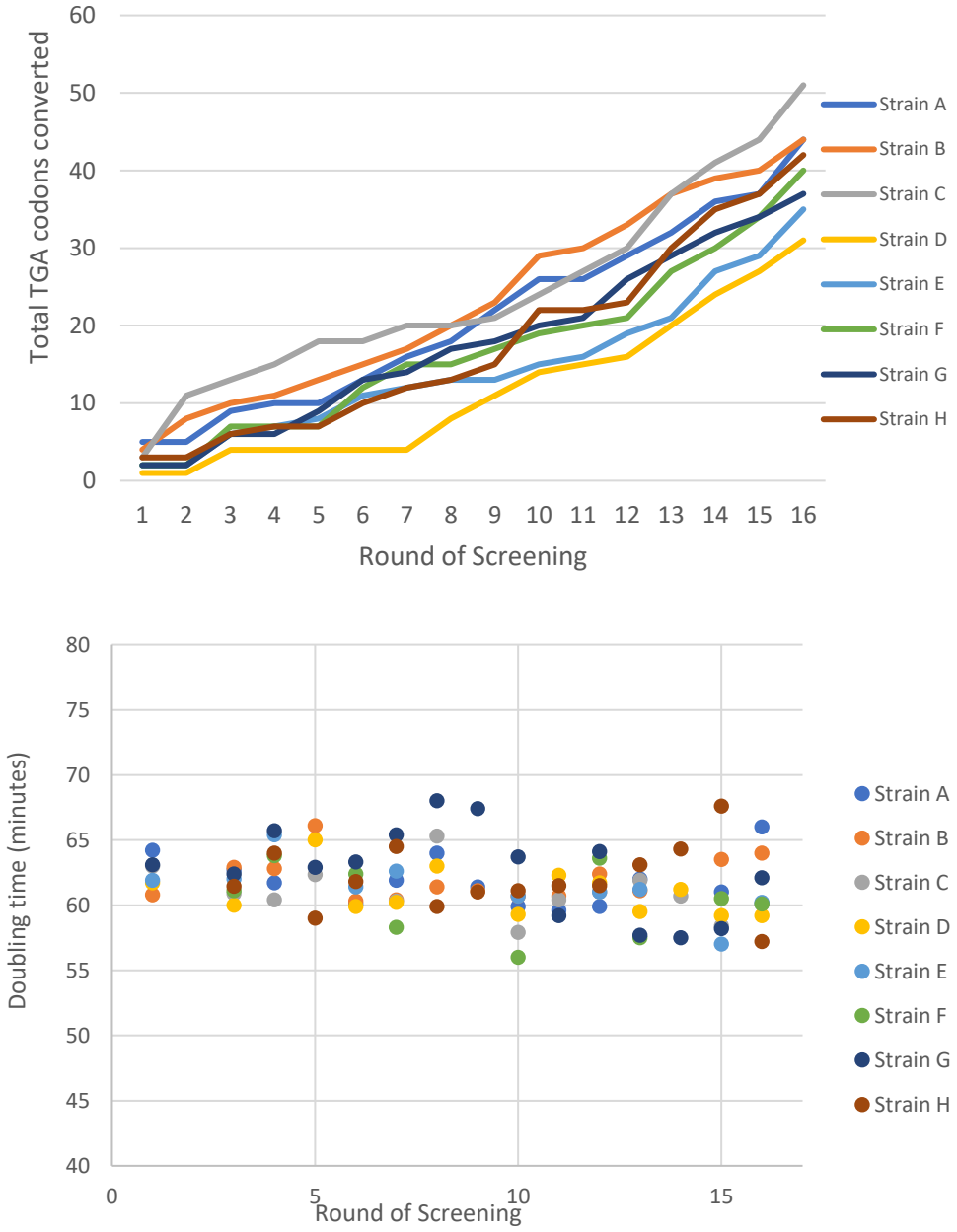
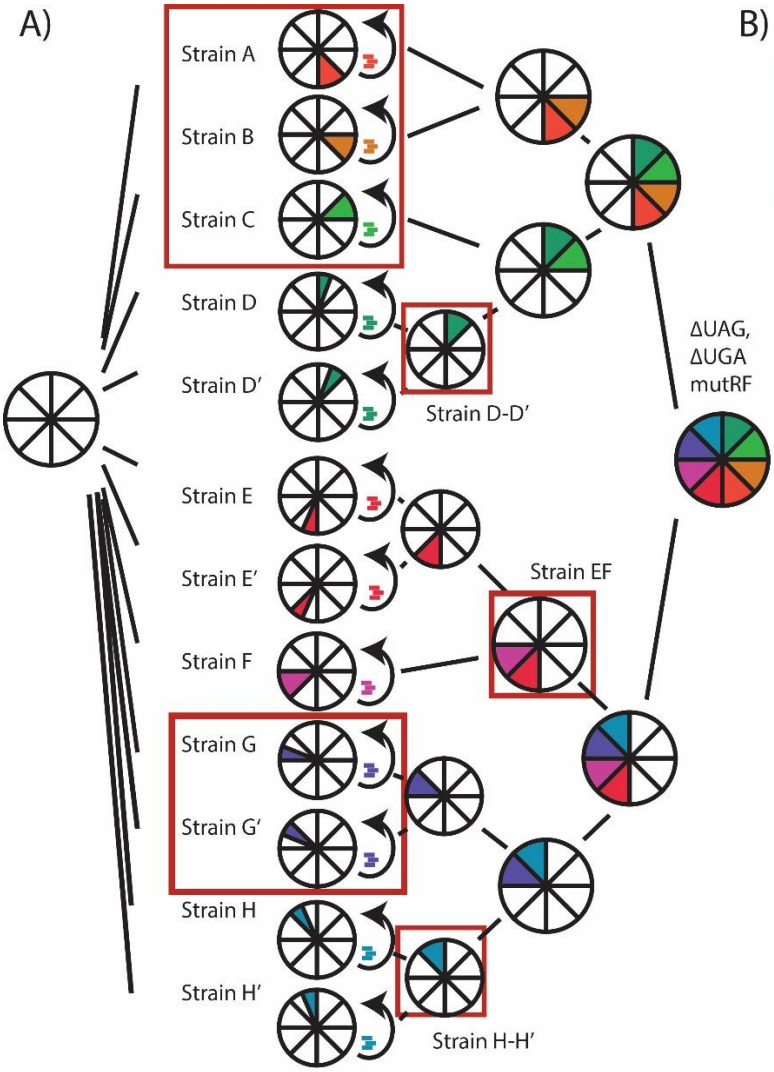


Figure 3. Initial characterization or recoding progress. The top chart shows the rate of recoding across the eight initial strains. The bottom shows the doubling times of the eight strains



B)

Strain	Number of codons remaining
A	4
B	5
C	3
D-D'	2
E-F	14
G	9
G'	17
H-H'	1

Figure 4. Full UGA recoding progress. A) shows the portion of the genome with non-essential UGA codons recoded in each strain. It also shows the conjugation pathway from 12 strains to a single strain. The strains in boxed in red are the most progressed strains. B) The current UGA codons remaining in each strain. This only includes the 55 codons remaining for conversion via MAGE, not the 20 remaining for deletion.

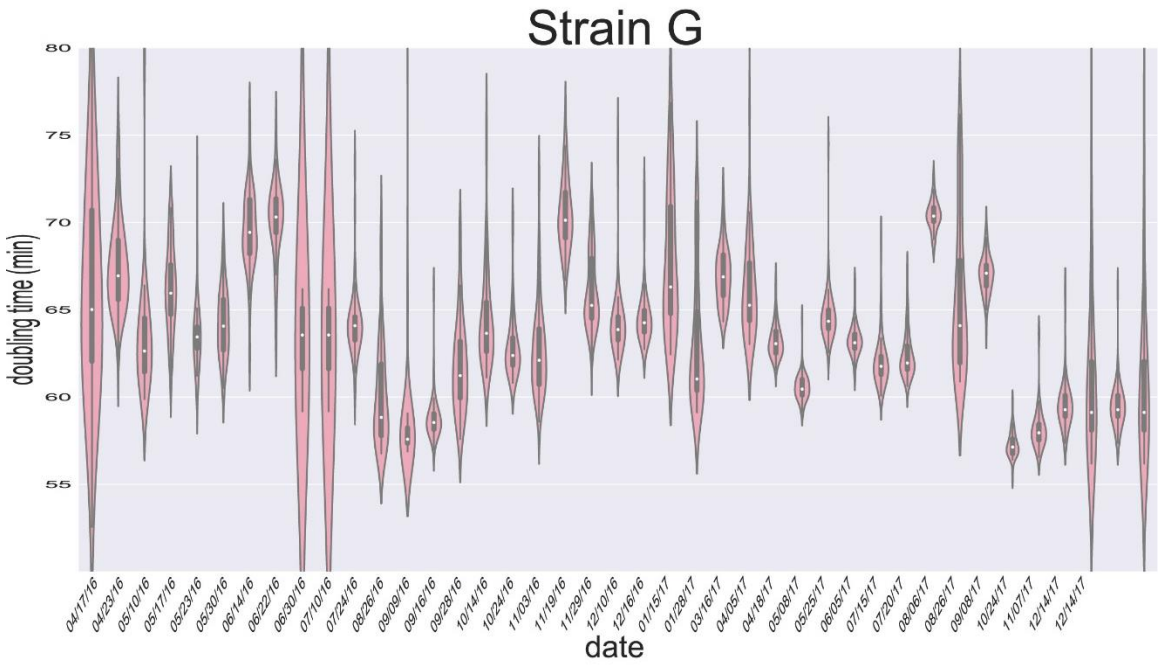


Figure 5. Violin plot of the aggregate doubling times of picked colonies over time for strain G. Each violin corresponds to a single experiment after several rounds of MAGE where 24-48 colonies were picked and their doubling time was measured on a plate reader. The violin represents the density of colonies at each doubling time. The thick lines within the violin represent the interquartile range and the white dot represents the median

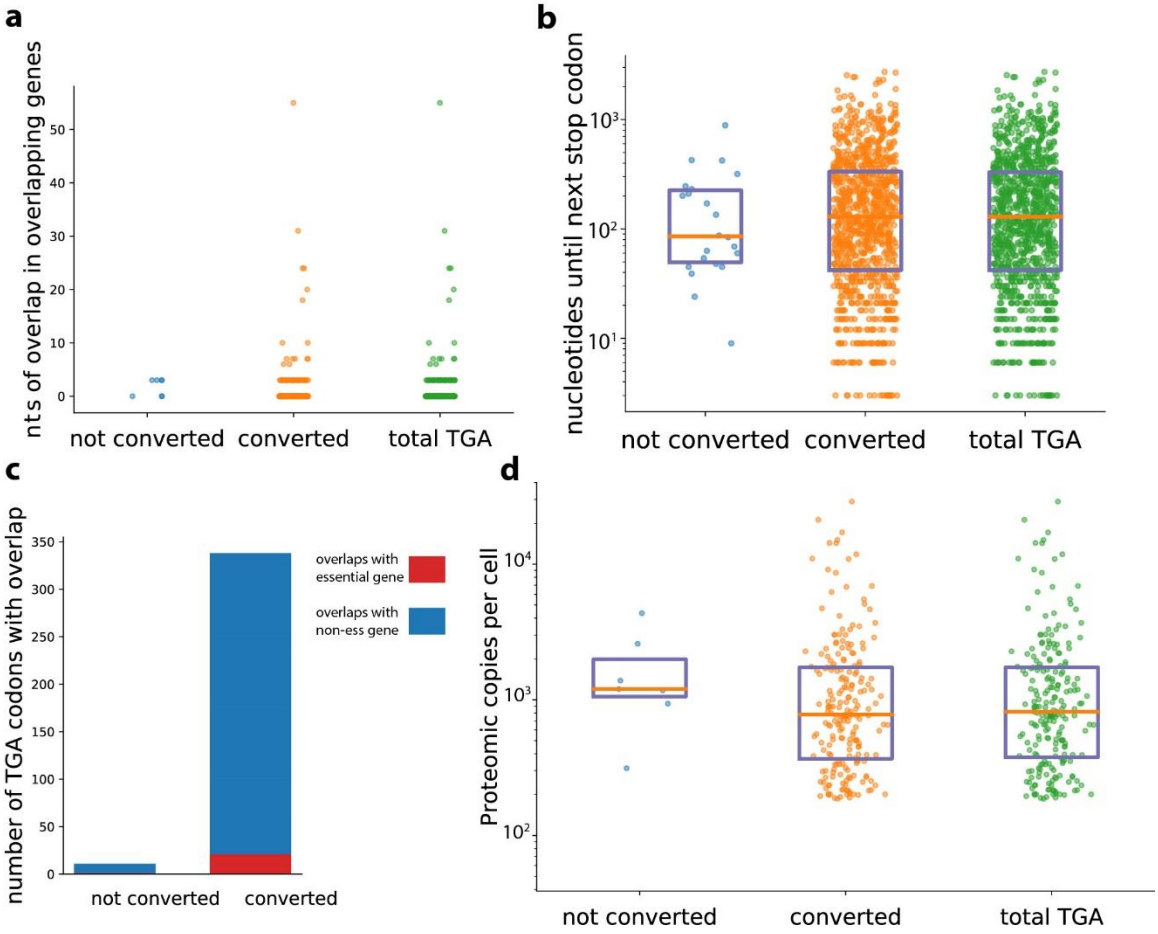


Figure 6. Comparison of properties of genes with converted UGA codons versus genes with non-converted UGA codons. A) The number of nucleotides of overlap in instances where the UGA codon overlaps with another gene. Each point represents a gene. B) The nucleotides after the UGA codon until the next UAA codon, which would be the only valid stop codon in the *E. coli* strain with two free codons. Each point represents a gene. The box represents the 75th-25th percentiles and the red line represent the median. C) The number of genes with overlap where the overlapped gene is either essential(red) or non-essential(blue). D) The number of protein copies observed by mass spectrometry in MG1655 grown in LB media[133] for genes ending in UGA. The points represent individual genes, while the boxes represent the interquartile range and the red line represents the media. Not all genes were observed in the proteomics data set

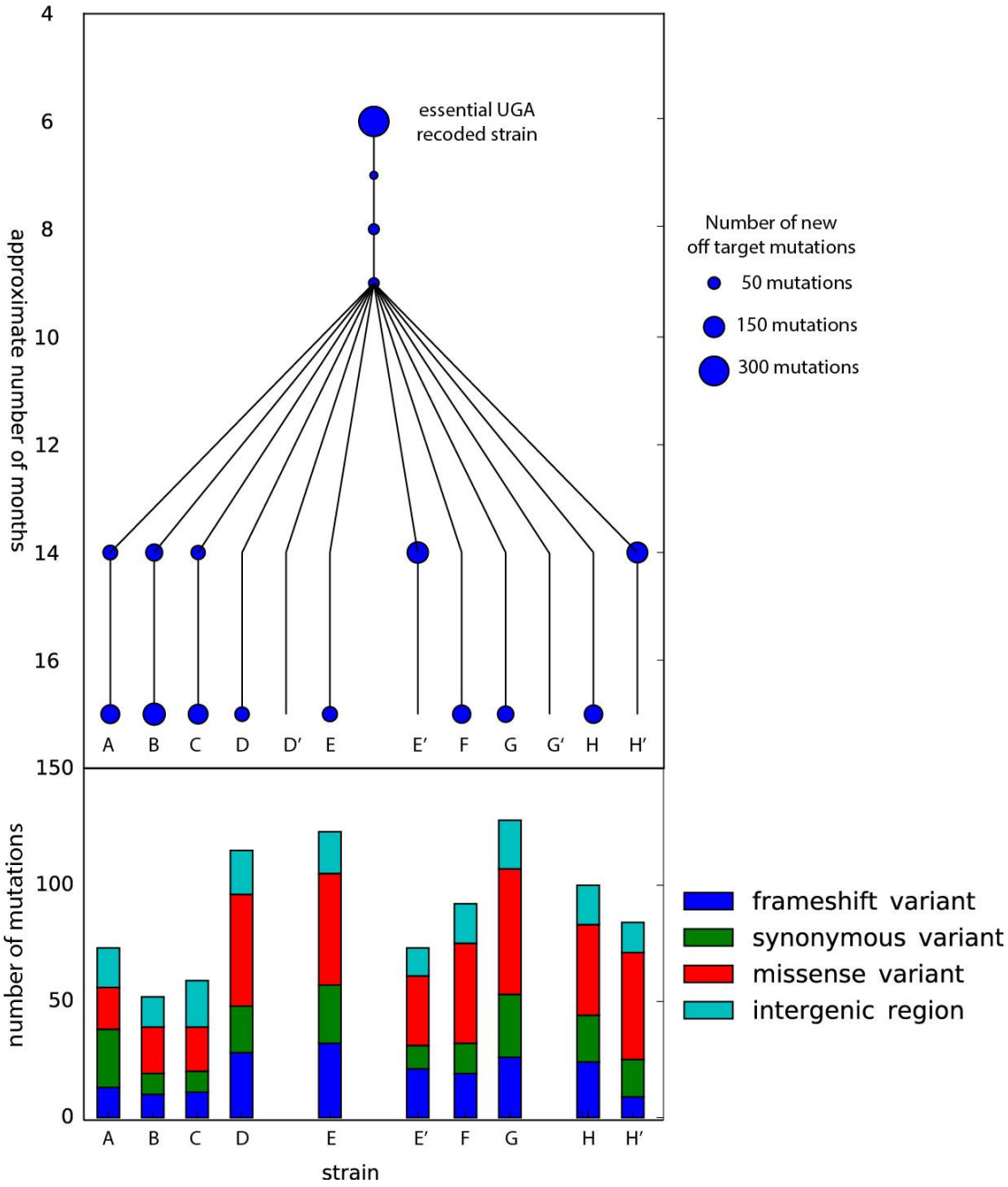


Figure 7. Off-target mutations present in the full UGA recoding strains. The top panel shows the lineage of the 12 strains. The number of new off-target mutations present in each strain is represented by the area of the dot. In the bottom panel, the specific type of mutations observed in each of these strains is shown.

Strain	MAGE pools	# of pools	Left coordinate	Right coordinate	Gap to strain before	Gap to strain after
A	83-86, 1-7	11	4489228	519640	10kb	12kb
B	8-18	11	532154	1123340	12kb	56kb
C	19-29	11	1179579	1701291	56kb	8kb
D	30-36	7	1709544	2030407	8kb	10kb
D '	37-40	4	2040046	2182534	10kb	21kb
E	41-44	4	2203703	2407541	21kb	2kb
E '	45-50	6	2409460	2705819	2kb	10kb
F	51-61	11	2715512	3258145	10kb	10kb
G'	62,63 71,72	4	3266412	3856423	10kb	10kb
G	64-70	7	3410819	3746600	62kb	7kb
H old	73-77	5	3866080	4225531	10kb	4kb
H new	78-82	5	4229246	4473334	4kb	10kb

Table 1. Division of the MAGE oligo pools amongst the 12 strains.

Chapter 6

Future directions

The next important achievement for this project is assembling all of the recoded UGA codons into a single strain. While recoding 100% of UGA codons to UAA still represents a worthy goal, a strain containing 95% of all UGA codons should allow for the deletion of RF2, and the testing of the effects of recoding on incorporation of NSAAs at UGA codons.

The constructs for the deletion of RF2 have already been validated (Chapter 4). If deletion of RF2 fails, this means either one or several of the unconverted UGA codons remain essential, or RF2 remains essential. To further investigate this possibility, the *prfA* and *prfB* genes should be shuffled to create a library of chimeric release factors[134]. This library should be used to complement a temperature sensitive RF2[78]. The domains of RF2 that remain in the chimeric release factors capable of complement RF2 should indicate those that continue to be essential. If RF2 domains outside of the recognition domain are conserved, this likely indicates that termination at UGA is not responsible for the continued essentiality of *prfB*. The domains conserved in these shuffling experiments could likely be introduced into the Q185E RF1 to produce a chimera capable of terminating at only UAA and complementing *prfB* deletion. If instead the conserved domains from the complementation assay include the recognition domain of RF2, this may indicate termination at UGA continues to be essential for some reason.

To test this scenario, we should attempt to engineer RF2 to remove UGA termination activity. Luckily, we have already validated an assay for *in vivo* UAG termination activity which easily could be converted to testing UGA termination activity. Since this assay uses fluorescence as an output, the assay could utilize FACS to rapidly

interrogate many variants, increasing the odds of success. If this successfully delivers a RF2 variant with impaired release at UGA as expected, then clearly the continued essentiality of RF2 is dues to a function outside of termination at UGA codons. Otherwise, the two combined assays provide strong evidence of the continued essentiality of termination at UGA codons.

Phage resistance experiments

The proposed phage resistance experiments are similar to those previously performed (**Fig. 1**)[26,29]. Specifically, they would measure the number of plaques after infection with one of three types of phage. These experiments serve two main purposes. First, they will provide further evidence that recoding insulates genomically recoded organisms from greedy genetic elements. Second, they will provide an elegant test of the codon specificity of the engineered release factor.

The current hypothesis as to why the C321.ΔA strain (lacking UAG and RF1) exhibits increased resistance to certain phage is that in the absence of RF1, UAG no longer codes for anything, which leads to ribosomal stalling. This leads to tagging with *ssrA* and subsequent degradation of both the protein and mRNA[95]. The reduced expression of the gene impedes productive infection. This hypothesis makes two hypotheses testable by our planned experiment. First, deletion of RF2 should lead to a similar reduction in viability of phage. Second, this reduction of viability should vary with the number of UGA codons in the phage genome. Phage lacking UGA codons

should be unaffected by an RF2 deletion, while those lacking UAG should be unaffected by an RF1 deletion.

All phage containing a UAG or UGA codon should be affected by a strain containing only a release factor that terminates at only UAA. Thus, this experiment provides a simple test of the specificity of the mutant release factor. If the mutant release factor supports plaques of the mu phage, this implies that it terminates at least in part at UGA codons. If it supports plaques of the MS2 phage, it has activity towards UAG codons. If neither grow on the strain with only the mutant release factor, then this implies activity at both of these codons below the level needed to support a phage infection.

Effects of recoding, and release factor manipulation on incorporation of a single nsAA

Recoding and translational factor engineering purportedly provide a strain with significantly higher expression of proteins containing non-standard amino acids. However, previous investigations into the effect of recoding and manipulation of translation factors (i.e. recoding of UAG and deletion of RF1) have focused mostly on the difference between the progenitor strain and the strain with both changes[26,28]. Thus, it is essential to test the effect of recoding then both release factor engineering on the incorporation of nsAA's (**Fig. 2**).

First, the effect of recoding UAG and deleting RF1 on suppression at UAG codons will be investigated. This provides an internal frame of reference to compare the effects of UGA recoding against. Second, the effects of UGA recoding on UAG suppression will

be investigated. While we do not expect UGA recoding to significantly alter UAG suppression, off-target mutations could potentially alter this phenotype. By comparing the results of this experiment against the results of the same experiment with the C321 strain, the effect of these off-target mutations can be estimated. This again provides a framing for interpreting UGA suppression experiments. The final experiment with UAG suppression is with a UGA recoded strain with a mutant release factor. If the mutant release factor behaves as hoped, this experiment should look similar to the experiment with RF1 deleted.

The UGA suppression experiments will be identical to the UAG experiments, with the caveat that UGA suppression systems are less well characterized. Therefore, the choice of orthogonal translation system may significantly impact the experiment. Two systems have been identified. The most convenient system utilizes a pEvol system[135] synthetase with the anticodon mutated to recognize UGA instead of UAG[36]. The alternative approach requires removal of the native tryptophan synthetase and tRNA and complementation with the yeast tryptophan tRNA and synthetase[136]. The *E. coli* tryptophan OTS can then be mutated to decode UGA and reintroduced.

Investigation of the impact of off-target mutations

The presence of many off-target mutations from both the UAG and UGA recoding efforts makes interpretation of the effects of recoding more difficult. Any change in phenotype in a UAG and UGA recoded strain that still maintained both RF1 and RF2 could be ascribed to the effects of off-target mutations. And any changes in

phenotype due to removing RF1 or RF2 could be exaggerated or ameliorated by those same off-target mutations. The nature of many of these mutations makes them unlikely to have an effect on observed phenotypes. Others, however, are known to have caused phenotypes including lowered doubling times and an inability to grow on minimal media. Therefore, future efforts should work to investigate the impact of the mutations, or endeavor to remove them.

These mutations could be removed by recreating a UAG and UGA recoded strain using genome synthesis. Much of the value of recoding using MAGE is the ability to convert codons in a sequential manner. Since the work is linear, defects in the design for recoding only impede the introduction of individual mutations. In contrast, recoding via synthesis of the genome can lead to the introduction of many recoding events in parallel. If any of these mutations prove lethal, the entire construct may fail. This makes determining the root design failure difficult. However, this work determined a successful design for UGA recoding. Therefore, whole genome synthesis using this design would be considerably less risky.

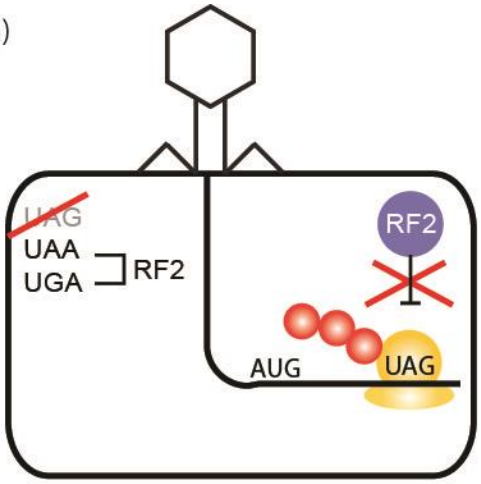
While many mutations likely negatively impact phenotypes, some mutations might have enabled further recoding. Excluding these from a design for whole genome synthesis would therefore lead to failure of that design. Therefore, it is still important to attempt to determine the impact of off-target mutations in the UAG and UGA recoded strain.

A classical approach to resolving the impact of different alleles is to recombine them into various backgrounds. Given the large number of off-target mutations,

isolating, recreating or removing each mutation would prove impractical. Luckily, the process of CAGE provides a series of strains with varying combinations of the off-target mutations. The phenotype of interest could be measured in each one of the precursor strains and the presence or absence of off-target mutations determined by the whole genome sequencing data. A regression analysis with the presence or absence of the off-target mutations as the inputs and the phenotype of interest as the output would provide information on the impact of each mutation[137].

This approach presents several problems. First, the large number of mutations compared to the number of samples. Hundreds of off-target mutations exist in the strains, while only roughly thirty strains will have genome data. Thus, even a linear regression would overfit the data. Several solutions exist. The number of inputs could be trimmed with a heuristic such as looking at only non-synonymous mutations in essential genes. A LASSO regression could reduce the number of inputs in an unbiased nature[138]. Even with these approaches, the data will still suffer from significant collinearity since mutations in a transferred segment of the genome will always appear with the other mutations from the transferred segment. This collinearity could negatively impact the regression's reflection of reality.

a)



b)

	UAG	UGA	UAA
M13	1	3	6
Mu	0	23	42
MS2	2	0	2

c)

Recoding	RF1	RF2	M13	Mu	MS2
ΔUAG	WT	WT			
ΔUAG	none	WT			
ΔUAG, ΔUGA	WT	WT			
ΔUAG, ΔUGA	none	WT			
ΔUAG, ΔUGA	WT	none			
ΔUAG, ΔUGA	Q185E RF1	none			

Figure 1. Future phage infectivity assay. A) demonstrates the biological explanation for phage resistance in genomically recoded strains lacking translational factors. B) The number of each stop codon present in each of the three representative phage chosen for the assay. C) The expected results from plaque assay. Dark blue represents no plaques expected. Light green represents plaques expected.

reporter construct	**plasmid RF1**	UAG termination						
mTAG	Q185E							
myTAG	Q185E							
mTAGy	Q185E							
mTAGy	WT							
mTAGy	none							
reporter construct	**plasmid RF2**	UGA termination						
mTGA	none							
mTGAy	none							
mTGAy	WT							
	RF1	WT	KO	Q185E	KO	WT	WT	Q185E
	RF2	WT	WT	WT	WT	WT	KO	KO
	recoding	C31	C31	C31	DM1179	DM1179	DM1179	DM1179

Figure 2. The effect or recoding and release factor engineering on the incorporation of non-standard amino acids at UAG and UGA codons. C31 signifies the strain that lacks UAG codon. DM1179 signifies the UGA recoded strain.

Appendix A

Chapter 3 supplemental tables and plasmid sequences

Supplemental Table 1. Genes containing a UGA. 0 signifies not essential while 1 signifies essential. Numbering is from MG1655

Gene	Location	Essential	overlapping with other gene	Overlapped Gene	Is overlapped gene essential?	Bases of overlap
rzoQ	1637104	0	1	rzpQ	0	447
rzoD	577732	0	1	rzpD	0	402
mokB	1489986	0	1	hokB	0	109
hofP	3518637	0	1	hofQ	0	88
btuB	4163506	0	1	murI	1	55
cho	1822426	0	1	ves	0	40
ybhQ	824263	0	1	ybhR	0	38
rnpA	3882875	1	1	yidD	0	36
dpiB	653116	0	1	dpiA	0	31
yebS	1915565	0	1	yebT	0	31
aceK	4218355	0	1	arpA	0	31
holD	4606239	0	1	rimI	0	31
yafJ	245094	0	1	yafK	0	29
tfaP	1207768	0	1	tfaE	0	28
smf	3430458	0	1	smg	0	28
nnr	4393636	0	1	tsaE	1	28
torR	1056485	0	1	torT	0	27
nrdI	2799397	0	1	nrdE	0	27
mpaA	1387894	0	1	ycjG	0	25
eutP	2573000	0	1	eutQ	0	25
mdtL	3890813	0	1	yidZ	0	25
yrdB	3427788	0	1	yrdA	0	24
yrdA	3427812	0	1	yrdB	0	24
pmrR	4330228	0	1	basS	0	24

cutA	4363041	0	1	dsbD	0	24
ycbC	971845	0	1	ycbJ	0	23
ycbJ	971868	0	1	ycbC	0	23
xylG	3731765	0	1	xylH	0	22
rffH	3972512	0	1	wecD	0	22
yhbO	3297514	0	1	yhbP	0	20
sdhE	3038826	0	1	cptA	0	19
mukF	974864	1	1	mukE	1	19
glpR	3557870	0	1	rtcR	0	18
rtcR	3557888	0	1	glpR	0	18
mobA	4039438	0	1	mobB	0	18
queG	4390951	0	1	yjeV	0	17
thiP	72911	0	1	thiQ	0	16
ybjC	890423	0	1	nfsA	0	16
hofN	3519455	0	1	hofO	0	16
ygdB	2960771	0	1	ppdC	0	15
iaaA	866756	0	1	gsiA	0	13
mdtJ	1671160	0	1	mdtI	0	13
ypeA	2549735	0	1	yfeZ	0	13
yhbQ	3298290	0	1	yhbS	0	13
phnN	4313548	0	1	phnO	0	13
vsr	2028472	0	1	yedA	0	11
cyoB	447874	0	1	cyoC	0	10
ybgK	744398	0	1	ybgL	0	10
ymcE	1051300	0	1	gnsA	0	10
ydcU	1512796	0	1	ydcV	0	10
glpA	2352297	0	1	glpB	0	10
cysW	2538826	0	1	cysA	0	10

eutJ	2567523	0	1	eutG	0	10
ygaZ	2808376	0	1	ygaH	0	10
hybA	3142176	0	1	hybB	0	10
agaC	3283510	0	1	agaD	0	10
yhdE	3395807	0	1	rng	0	10
yhdT	3405639	0	1	panF	0	10
frvB	4088938	0	1	frvX	0	10
yfdR	2473542	0	1	yfdS	0	9
yjjP	4600111	0	1	yjjB	0	9
ydcK	1497493	0	1	rimL	0	8
hcaT	2664729	0	1	csiE	0	8
csiE	2664737	0	1	hcaT	0	8
yaeP	213925	0	1	rof	0	7
metN	221614	0	1	metI	0	7
ninE	572314	0	1	ybcO	0	7
uof	709862	0	1	fur	0	7
ybhG	828197	0	1	ybhF	0	7
ddpD	1556055	0	1	ddpF	0	7
fliG	2013899	0	1	fliH	0	7
yehM	2205996	0	1	yehP	0	7
yehY	2215422	0	1	yehX	0	7
hyfI	2609486	0	1	hyfJ	0	7
rnlA	2765013	0	1	rnlB	0	7
proV	2804039	0	1	proW	0	7
srlR	2827842	0	1	gutQ	0	7
hycH	2841058	0	1	hycI	0	7
casA	2880652	0	1	casB	0	7
ptrA	2954018	0	1	recB	0	7

aas	2971877	0	1	lplT	0	7
yqeI	2987333	0	1	yqeJ	0	7
hybE	3138333	0	1	hybF	0	7
hybD	3138814	0	1	hybE	0	7
yrbN	3305875	0	1	deaD	0	7
waaG	3803966	0	1	waaP	0	7
thiE	4191592	0	1	thiF	0	7
frdA	4378533	0	1	frdB	0	7
mutL	4397282	0	1	miaA	0	7
hypT	4555401	0	1	iraD	0	7
murC	102240	1	1	ddlB	0	7
ybgJ	743472	0	1	ybgK	0	6
galK	788054	0	1	galM	0	6
lsrA	1601049	0	1	lsrC	0	6
arnF	2371300	0	1	pmrD	0	6
yheT	3482246	0	1	yheU	0	6
fixC	45466	0	1	fixX	0	3
pdxA	52427	0	1	rsmA	0	3
rsmH	91035	0	1	ftsL	1	3
ampD	119284	0	1	ampE	0	3
yadG	143705	0	1	yadH	0	3
pcnB	157729	0	1	folK	1	3
fhuD	171465	0	1	fhuB	0	3
yaeQ	214836	0	1	arfB	0	3
rcsF	219591	0	1	tsaA	0	3
paoB	300155	0	1	paoC	0	3
paoA	301108	0	1	paoB	0	3
insE1	314814	0	1	insF1	0	3

mhpD	372148	0	1	mhpF	0	3
mhpF	373095	0	1	mhpE	0	3
tauB	386198	0	1	tauC	0	3
tauC	387022	0	1	tauD	0	3
insE2	391826	0	1	insF2	0	3
sbcD	414974	0	1	sbcC	0	3
qmcA	514080	0	1	ybbJ	0	3
ybbA	519643	0	1	ybbP	0	3
ylbF	549665	0	1	ybcF	0	3
purE	551814	0	1	purK	0	3
insE3	566364	0	1	insF3	0	3
ylcH	571692	0	1	ybcN	0	3
ybcO	572597	0	1	rusA	0	3
rusA	572956	0	1	ylcG	0	3
rrrD	577333	0	1	rzpD	0	3
ybdZ	613383	0	1	entF	0	3
fepG	619419	0	1	fepC	0	3
fepD	620408	0	1	fepG	0	3
citD	649710	0	1	citE	0	3
ybeR	676641	0	1	djlB	0	3
ybeU	679438	0	1	djlC	0	3
ybeZ	691561	0	1	ybeY	0	3
speF	717485	0	1	potE	0	3
kdpD	720953	0	1	kdpE	0	3
ybgA	738733	0	1	phr	0	3
ybgP	748202	0	1	ybgO	0	3
ybgC	774379	0	1	tolQ	0	3
bioB	809607	0	1	bioF	0	3

ybhP	822959	0	1	clsB	0	3
glnP	845683	0	1	glnQ	0	3
mntR	852873	0	1	ybiR	0	3
potH	896310	0	1	potI	0	3
ybjQ	904139	0	1	amiD	0	3
macA	919573	0	1	macB	0	3
ycaR	970078	0	1	kdsB	1	3
smtA	973545	0	1	mukF	1	3
ssuC	993264	0	1	ssuB	0	3
ssuA	995208	0	1	ssuD	0	3
pqiB	1014122	0	1	ymbA	0	3
hyaA	1032480	0	1	hyaB	0	3
hyaC	1034996	0	1	hyaD	0	3
hyaD	1035580	0	1	hyaE	0	3
hyaE	1035975	0	1	hyaF	0	3
gfcB	1047911	0	1	gfcC	0	3
torA	1061025	0	1	torD	0	3
insE4	1094361	0	1	insF4	0	3
ymfA	1180487	0	1	ycfZ	0	3
potC	1182049	0	1	potD	0	3
potB	1182840	0	1	potC	0	3
ymfL	1203048	0	1	ymfM	0	3
narG	1282830	0	1	narH	0	3
narH	1284365	0	1	narJ	0	3
oppD	1303791	0	1	oppF	0	3
yciV	1322125	0	1	yciO	0	3
yciK	1326378	0	1	btuR	0	3
sapA	1353491	0	1	sapB	0	3

ycjS	1375911	0	1	ycjT	0	3
ycjT	1378175	0	1	ycjU	0	3
ycjX	1383538	0	1	ycjF	0	3
ydbH	1443714	0	1	ynbE	0	3
paaE	1455524	0	1	paaF	0	3
tehA	1499589	0	1	tehB	0	3
mnaT	1516352	0	1	ydcZ	0	3
narW	1534638	0	1	narV	0	3
narZ	1536874	0	1	narY	0	3
ddpC	1557038	0	1	ddpD	0	3
ddpB	1557931	0	1	ddpC	0	3
rrrQ	1637548	0	1	rzpQ	0	3
ydfR	1638078	0	1	rrrQ	0	3
dicB	1647821	0	1	ydfD	0	3
uidA	1692284	0	1	uidB	0	3
ydhX	1748369	0	1	ydhU	0	3
sufD	1758544	0	1	sufS	0	3
ydiJ	1763653	0	1	menI	0	3
ydiS	1780708	0	1	ydiT	0	3
astD	1826280	0	1	astB	0	3
astA	1827755	0	1	astD	0	3
astC	1828786	0	1	astA	0	3
znuC	1941441	0	1	znuB	0	3
yecD	1949422	0	1	yecE	0	3
cmoA	1951469	0	1	cmoB	0	3
motA	1974276	0	1	motB	0	3
uvrY	1992727	0	1	uvrC	0	3
yecS	1995835	0	1	yecC	0	3

yedE	2007506	0	1	yedF	0	3
fliJ	2016413	0	1	fliK	0	3
fliM	2019115	0	1	fliN	0	3
cobU	2063243	0	1	cobS	0	3
yeeR	2074335	0	1	yeeS	0	3
hisD	2090425	0	1	hisC	0	3
rfbC	2107605	0	1	wzxB	0	3
wcaK	2115148	0	1	wcaL	0	3
wcaH	2123767	0	1	wcaI	0	3
wcaB	2130091	0	1	wcaC	0	3
yegL	2150493	0	1	yegK	0	3
mdtD	2160903	0	1	baeS	0	3
baeS	2162303	0	1	baeR	0	3
insE5	2168559	0	1	insF5	0	3
yegT	2178120	0	1	yegU	0	3
yegU	2179121	0	1	yegV	0	3
thiM	2182535	0	1	thiD	0	3
yehU	2210981	0	1	yehT	0	3
yohJ	2229044	0	1	yohK	0	3
dsbE	2290429	0	1	ccmH	0	3
ccmF	2290983	0	1	dsbE	0	3
ccmE	2292923	0	1	ccmF	0	3
ccmD	2293399	0	1	ccmE	0	3
ccmC	2293605	0	1	ccmD	0	3
ccmA	2295043	0	1	ccmB	0	3
napH	2296737	0	1	napB	0	3
napD	2300772	0	1	napA	0	3
atoS	2319891	0	1	atoC	0	3

atoA	2322781	0	1	atoE	0	3
yfaA	2332978	0	1	yfaT	0	3
glpB	2353546	0	1	glpC	0	3
arnA	2368043	0	1	arnD	0	3
arnT	2370582	0	1	arnE	0	3
menC	2373022	0	1	menE	0	3
menD	2375611	0	1	menH	0	3
nuoK	2393065	0	1	nuoL	0	3
nuoJ	2393364	0	1	nuoK	0	3
nuoG	2395461	0	1	nuoH	0	3
nuoE	2399574	0	1	nuoF	0	3
hisQ	2423252	0	1	hisM	0	3
yfcQ	2448609	0	1	yfcP	0	3
yfcR	2449094	0	1	yfcQ	0	3
gtrA	2466239	0	1	gtrB	0	3
gtrB	2467156	0	1	gtrS	0	3
yffO	2560549	0	1	yffP	0	3
eutH	2564903	0	1	eutA	0	3
eutT	2571524	0	1	eutD	0	3
eutQ	2572324	0	1	eutT	0	3
hyfH	2608731	0	1	hyfI	0	3
hcaE	2668415	0	1	hcaF	0	3
hcaC	2669250	0	1	hcaB	0	3
rnc	2701405	0	1	era	1	3
rseB	2705820	0	1	rseC	0	3
rluD	2733053	0	1	yfiH	0	3
yfjL	2761559	0	1	yfjK	0	3
nrdH	2798990	0	1	nrdI	0	3

yqaB	2814959	0	1	yqaA	0	3
srlA	2824417	0	1	srlE	0	3
norV	2831937	0	1	norW	0	3
hycG	2841465	0	1	hycH	0	3
hycB	2847260	0	1	hycC	0	3
hypD	2851279	0	1	hypE	0	3
ygbJ	2860360	0	1	ygbK	0	3
ygcR	2893798	0	1	ygcQ	0	3
ppdB	2961175	0	1	ygdB	0	3
xdhB	3001514	0	1	xdhC	0	3
ygfM	3019341	0	1	xdhD	0	3
pepP	3051537	0	1	ubiH	0	3
yggD	3072708	0	1	yggC	0	3
yggP	3074201	0	1	yggF	0	3
yggT	3094408	0	1	yggU	0	3
hybB	3141008	0	1	hybC	0	3
qseB	3168509	0	1	qseC	0	3
ttdA	3205396	0	1	ttdB	0	3
ebgA	3223747	0	1	ebgC	0	3
higB	3232163	0	1	higA	0	3
kdsC	3340861	0	1	lptC	1	3
yhbJ	3345991	0	1	npr	0	3
elbB	3347828	0	1	mtgA	0	3
nanK	3367497	0	1	yhcH	0	3
nanE	3368369	0	1	nanK	0	3
aroE	3428042	0	1	yrdB	0	3
gspE	3457842	0	1	gspF	0	3
gspH	3459999	0	1	gspI	0	3

gspL	3463107	0	1	gspM	0	3
nirB	3494576	0	1	nirD	0	3
frlC	3501192	0	1	frlD	0	3
yhfW	3506609	0	1	php	0	3
ompR	3533887	0	1	envZ	0	3
glgB	3569339	0	1	glgX	0	3
ugpC	3586133	0	1	ugpQ	0	3
ugpA	3588047	0	1	ugpE	0	3
livM	3592226	0	1	livG	0	3
livH	3593500	0	1	livM	0	3
nikB	3614208	0	1	nikC	0	3
nikD	3615802	0	1	nikE	0	3
yhiI	3627558	0	1	rbbA	0	3
bcsE	3696052	0	1	bcsF	0	3
bcsF	3696240	0	1	bcsG	0	3
dppD	3700888	0	1	dppF	0	3
tag	3711678	0	1	yiaC	0	3
insJ	3719224	0	1	insK	0	3
lyxK	3746603	0	1	yiaQ	0	3
selA	3757881	0	1	selB	0	3
lldR	3777853	0	1	lldD	0	3
waaQ	3805087	0	1	waaG	0	3
yidG	3853631	0	1	yidF	0	3
yidK	3856424	0	1	yidJ	0	3
dgoA	3871018	0	1	dgoD	0	3
dgoR	3872494	0	1	dgoK	0	3
wecB	3969286	0	1	wecC	0	3
wzxE	3975551	0	1	wecF	0	3

wecF	3976627	0	1	wzyE	1	3
dapF	3993609	0	1	yigA	0	3
yigA	3994313	0	1	xerC	0	3
yihY	4075041	0	1	dtd	0	3
dtd	4075475	0	1	yiiD	0	3
fdoI	4079248	0	1	fdhE	0	3
fdoH	4079880	0	1	fdoI	0	3
rhaB	4094002	0	1	rhaA	0	3
cpxR	4102995	0	1	cpxA	0	3
thiG	4189888	0	1	thiH	0	3
zraS	4201346	0	1	zraR	0	3
yjbF	4234923	0	1	yjbG	0	3
yjcH	4282922	0	1	actP	0	3
nrfB	4287834	0	1	nrfC	0	3
nrfC	4288502	0	1	nrfD	0	3
nrfF	4291569	0	1	nrfG	0	3
mdtO	4299050	0	1	mdtP	0	3
phnL	4315238	0	1	phnM	0	3
phnJ	4316784	0	1	phnK	0	3
phnG	4319267	0	1	phnH	0	3
crfC	4327386	0	1	yjcZ	0	3
eptA	4331970	0	1	basR	0	3
dcuS	4348054	0	1	dcuR	0	3
yjeN	4383730	0	1	yjeO	0	3
tamA	4442138	0	1	tamB	0	3
fecC	4510434	0	1	fecD	0	3
fecB	4511429	0	1	fecC	0	3
fecI	4515737	0	1	fecR	0	3

sgcX	4528553	0	1	sgcB	0	3
yjiH	4558020	0	1	yjiG	0	3
hsdM	4579482	0	1	hsdS	0	3
yjjU	4611507	0	1	yjjV	0	3
murE	94653	1	1	murF	1	3
ftsW	99647	1	1	murG	1	3
ftsQ	103985	1	1	ftsA	1	3
lpxB	204496	1	1	rnhB	0	3
msbA	967592	1	1	lpxK	1	3
tmk	1154988	1	1	holB	1	3
prmC	1266150	1	1	ychQ	0	3
yefM	2087486	1	1	yoeB	0	3
ispF	2869323	1	1	truD	0	3
rplD	3449703	1	1	rplW	1	3
hemC	3987848	1	1	hemD	1	3
yigP	4018252	1	1	ubiB	1	3
murB	4171108	1	1	birA	1	3
thrL	255	0	0	none	0	0
thrA	2799	0	0	none	0	0
yaaW	10643	0	0	none	0	0
nhaA	18655	0	0	none	0	0
yaaY	21399	0	0	none	0	0
caiE	34781	0	0	none	0	0
fixB	44129	0	0	none	0	0
fixX	45750	0	0	none	0	0
djlA	58179	0	0	none	0	0
polB	63429	0	0	none	0	0
sgrR	75644	0	0	none	0	0
sgrT	77519	0	0	none	0	0

setA	78799	0	0	none	0	0
leuA	81958	0	0	none	0	0
ilvI	87354	0	0	none	0	0
ilvH	87848	0	0	none	0	0
yacG	111649	0	0	none	0	0
panD	146314	0	0	none	0	0
yadC	149715	0	0	none	0	0
gluQ	159186	0	0	none	0	0
fhuC	170575	0	1	fhuD	0	0
clcA	176528	0	0	none	0	0
yadS	177001	0	0	none	0	0
glnD	185978	0	0	none	0	0
rnhB	205089	0	0	none	0	0
rof	213678	0	0	none	0	0
yaeF	216179	0	0	none	0	0
gmhB	223408	0	0	none	0	0
mltD	232597	0	0	none	0	0
gloB	234027	0	0	none	0	0
yafV	239419	0	0	none	0	0
dinB	251953	0	0	none	0	0
yafO	252699	0	0	none	0	0
crl	258230	0	0	none	0	0
ykfI	262552	0	0	none	0	0
ykfH	263250	0	0	none	0	0
ykfG	263480	0	0	none	0	0
yafX	263972	0	0	none	0	0
ykfF	264528	0	0	none	0	0
ykfB	264844	0	0	none	0	0

yafZ	266408	0	0	none	0	0
ykfA	267321	0	0	none	0	0
perR	268513	0	0	none	0	0
insI1	270978	0	0	none	0	0
mmuM	276871	0	0	none	0	0
yagE	282410	0	0	none	0	0
yagI	287628	0	0	none	0	0
argF	288525	0	0	none	0	0
yagN	294363	0	0	none	0	0
intF	294920	0	0	none	0	0
paoD	296994	0	0	none	0	0
ykgO	311598	0	0	none	0	0
ykgM	311738	0	0	none	0	0
ykgB	316950	0	0	none	0	0
ykgD	320305	0	0	none	0	0
ykgE	321551	0	0	none	0	0
ykgG	323677	0	0	none	0	0
betA	324801	0	0	none	0	0
betT	330720	0	0	none	0	0
yahA	332683	0	0	none	0	0
yahF	337549	0	1	yahG	0	0
yahG	338967	0	0	none	0	0
yahJ	341731	0	0	none	0	0
yahK	343157	0	0	none	0	0
yahN	344890	0	0	none	0	0
prpR	346081	0	0	none	0	0
codA	356678	0	0	none	0	0
cynS	359183	0	0	none	0	0

cynX	360370	0	0	none	0	0
lacI	365652	0	0	none	0	0
mhpR	366811	0	0	none	0	0
mhpA	369499	0	0	none	0	0
mhpB	370445	0	0	none	0	0
mhpT	375894	0	0	none	0	0
frmB	376759	0	0	none	0	0
frmA	377686	0	0	none	0	0
yaiO	379293	0	0	none	0	0
ampH	394354	0	0	none	0	0
yaiW	398190	0	0	none	0	0
psiF	402825	0	0	none	0	0
yaiC	404042	0	0	none	0	0
proC	404059	0	0	none	0	0
aroL	406153	0	0	none	0	0
yaiA	406394	0	0	none	0	0
yajD	430176	0	0	none	0	0
tsx	430353	0	0	none	0	0
nusB	434780	0	0	none	0	0
clpP	456524	0	0	none	0	0
ybaE	464836	0	0	none	0	0
mdlB	471641	0	0	none	0	0
ylaB	476291	0	0	none	0	0
acrB	480478	0	0	none	0	0
ybaM	489334	0	0	none	0	0
ybaN	490483	0	0	none	0	0
ybaB	493629	0	1	recR	0	0
htpG	496218	0	0	none	0	0

fsr	502700	0	0	none	0	0
glsA	511797	0	0	none	0	0
cueR	513624	0	0	none	0	0
ybbM	516586	0	0	none	0	0
ybbN	516649	0	0	none	0	0
ybbO	517564	0	0	none	0	0
ybbP	522054	0	0	none	0	0
allA	532157	0	0	none	0	0
glxR	536688	0	0	none	0	0
ybcJ	555884	0	0	none	0	0
sfmH	562542	0	0	none	0	0
ylcG	573093	0	0	none	0	0
ybcW	579309	0	0	none	0	0
ylcI	579474	0	0	none	0	0
appY	583653	0	0	none	0	0
envY	585370	0	0	none	0	0
cusC	596196	0	0	none	0	0
cusB	597925	0	0	none	0	0
fepA	609477	0	0	none	0	0
fes	613162	0	0	none	0	0
entE	626903	0	0	none	0	0
entH	628936	0	0	none	0	0
ybdH	631612	0	0	none	0	0
ybdO	635939	0	0	none	0	0
ybdR	642549	0	0	none	0	0
citC	650021	0	0	none	0	0
dpiA	653765	0	0	none	0	0
pagP	656340	0	0	none	0	0

tatE	658373	0	0	none	0	0
ybeF	659648	0	0	none	0	0
rlmH	667471	0	0	none	0	0
cobC	668519	0	0	none	0	0
djlB	678065	0	0	none	0	0
djlC	680886	0	0	none	0	0
umpH	698797	0	0	none	0	0
chiP	708963	0	0	none	0	0
ybfK	720063	0	0	none	0	0
kdpE	720279	0	0	none	0	0
kdpB	724211	0	0	none	0	0
kdpA	726282	0	0	none	0	0
kdpF	727955	0	0	none	0	0
ybfA	728563	0	0	none	0	0
ybgO	747144	0	0	none	0	0
sdhD	755130	0	1	sdhA	0	0
mngB	769834	0	0	none	0	0
cydX	773532	0	1	ybgE	0	0
tolB	778255	0	0	none	0	0
modF	791539	0	0	none	0	0
acrZ	794145	0	0	none	0	0
modC	796835	0	0	none	0	0
ybhH	801034	0	0	none	0	0
ybhN	820765	0	0	none	0	0
clsB	821721	0	0	none	0	0
ybiX	837753	0	0	none	0	0
fiu	838472	0	0	none	0	0
ybiU	857019	0	0	none	0	0

ybiV	858436	0	0	none	0	0
bssR	877854	0	0	none	0	0
ybjG	882015	0	0	none	0	0
ybjJ	885354	0	0	none	0	0
grxA	889719	0	0	none	0	0
ybjN	892656	0	0	none	0	0
potI	897152	0	0	none	0	0
clpS	922456	0	0	none	0	0
aat	925951	0	0	none	0	0
dmsB	943254	0	0	none	0	0
ycaP	956677	0	0	none	0	0
aroA	959318	0	0	none	0	0
pepN	992457	0	0	none	0	0
ycbU	1002654	0	0	none	0	0
ycbF	1003880	0	0	none	0	0
rlmL	1009175	0	0	none	0	0
uup	1011094	0	0	none	0	0
pqiA	1012477	0	0	none	0	0
rmf	1015105	0	0	none	0	0
ycbZ	1015762	0	0	none	0	0
yccT	1026334	0	0	none	0	0
gfcC	1047168	0	0	none	0	0
gnsA	1051463	0	0	none	0	0
torS	1052657	0	0	none	0	0
cbpM	1061773	0	0	none	0	0
rutC	1070996	0	0	none	0	0
rutA	1072086	0	0	none	0	0
ycdU	1096052	0	0	none	0	0

ycdZ	1100010	0	0	none	0	0
csgG	1100074	0	0	none	0	0
ymdB	1105576	0	0	none	0	0
yceK	1113029	0	0	none	0	0
mdtG	1113487	0	0	none	0	0
mdtH	1123341	0	0	none	0	0
flgN	1128637	0	0	none	0	0
flgM	1129058	0	0	none	0	0
flgI	1136594	0	1	flgJ	0	0
flgJ	1137535	0	0	none	0	0
ycfH	1156797	0	0	none	0	0
fhuE	1158585	0	0	none	0	0
cobB	1179582	0	0	none	0	0
phoP	1188999	0	0	none	0	0
hflD	1191213	0	0	none	0	0
mnmA	1191890	0	0	none	0	0
ymfE	1196756	0	0	none	0	0
lit	1198811	0	0	none	0	0
intE	1198902	0	0	none	0	0
ycgG	1218074	0	0	none	0	0
ymgJ	1222672	0	0	none	0	0
minC	1224608	0	0	none	0	0
ycgJ	1226191	0	0	none	0	0
ycgN	1228499	0	0	none	0	0
hlyE	1228706	0	0	none	0	0
umuD	1230409	0	1	umuC	0	0
nhaB	1232399	0	0	none	0	0
ycgR	1243016	0	0	none	0	0

dhaR	1252208	0	0	none	0	0
chaA	1269972	0	0	none	0	0
chaC	1272425	0	0	none	0	0
narL	1274402	0	0	none	0	0
tpr	1286310	0	0	none	0	0
rssA	1289373	0	0	none	0	0
rssB	1290478	0	0	none	0	0
oppB	1301843	0	0	none	0	0
yciI	1308593	0	0	none	0	0
yciF	1313294	0	0	none	0	0
yciG	1313880	0	0	none	0	0
trpB	1315246	0	0	none	0	0
trpE	1319408	0	0	none	0	0
trpL	1321062	0	0	none	0	0
pyrF	1340682	0	1	yciH	0	0
ycjD	1349431	0	0	none	0	0
sapD	1350660	0	0	none	0	0
puuD	1359908	0	0	none	0	0
puuC	1362254	0	0	none	0	0
pspC	1367408	0	0	none	0	0
pspD	1367638	0	0	none	0	0
ompG	1380876	0	0	none	0	0
mppA	1392864	0	0	none	0	0
fnr	1396798	0	0	none	0	0
abgA	1401279	0	0	none	0	0
smrA	1404566	0	0	none	0	0
ydaQ	1411261	0	0	none	0	0
racC	1415512	0	0	none	0	0

kilR	1416032	0	0	none	0	0
ydaG	1417346	0	0	none	0	0
ydaS	1418685	0	0	none	0	0
uspF	1433209	0	0	none	0	0
ynbE	1443896	0	0	none	0	0
tynA	1447100	0	0	none	0	0
paaB	1453179	0	0	none	0	0
paaD	1454446	0	0	none	0	0
paaI	1458917	0	1	paaJ	0	0
paaJ	1460122	0	0	none	0	0
paaK	1461462	0	0	none	0	0
insI3	1468533	0	0	none	0	0
ydbD	1475474	0	0	none	0	0
ynbB	1477146	0	0	none	0	0
ynbC	1478919	0	0	none	0	0
cybB	1489456	0	0	none	0	0
ydcA	1489874	0	0	none	0	0
trg	1492134	0	0	none	0	0
ydcI	1492172	0	0	none	0	0
tehB	1500179	0	0	none	0	0
insQ	1502889	0	0	none	0	0
ydcO	1502929	0	0	none	0	0
yncJ	1506858	0	0	none	0	0
ydcT	1511854	0	0	none	0	0
yncL	1515123	0	0	none	0	0
ydcY	1515905	0	0	none	0	0
yncG	1524888	0	0	none	0	0
yddE	1532989	0	0	none	0	0

fdnG	1548472	0	0	none	0	0
dosP	1561358	0	0	none	0	0
gadB	1568669	0	0	none	0	0
ydeO	1580950	0	0	none	0	0
safA	1581786	0	0	none	0	0
hipA	1588878	0	0	none	0	0
lsrB	1604097	0	0	none	0	0
lsrG	1605313	0	0	none	0	0
yneE	1606132	0	0	none	0	0
yneF	1608931	0	0	none	0	0
yneG	1609990	0	0	none	0	0
sad	1611339	0	0	none	0	0
yneJ	1613709	0	0	none	0	0
ydeE	1620543	0	0	none	0	0
ynfO	1635013	0	0	none	0	0
gnsB	1635633	0	0	none	0	0
ynfN	1635978	0	0	none	0	0
cspI	1636479	0	0	none	0	0
cspF	1640091	0	0	none	0	0
relE	1643370	0	0	none	0	0
relB	1643657	0	0	none	0	0
dicC	1645644	0	0	none	0	0
ydfA	1646687	0	0	none	0	0
ynfL	1666723	0	0	none	0	0
ynfM	1668976	0	0	none	0	0
ydgU	1669884	0	0	none	0	0
mdtI	1670844	0	0	none	0	0
tqsA	1672971	0	0	none	0	0

rstB	1682207	0	0	none	0	0
uidR	1694486	0	0	none	0	0
ydgJ	1701292	0	0	none	0	0
rsxE	1709547	0	1	nth	0	0
nth	1710182	0	0	none	0	0
pdxY	1713050	0	0	none	0	0
pdxH	1715375	0	0	none	0	0
mliC	1716090	0	0	none	0	0
anmK	1716517	0	0	none	0	0
slyA	1718414	0	0	none	0	0
ydhL	1723705	0	0	none	0	0
ynhF	1735480	0	0	none	0	0
ydhC	1739146	0	0	none	0	0
ydhR	1745029	0	0	none	0	0
ydhT	1746771	0	0	none	0	0
ydhW	1749101	0	0	none	0	0
ydhY	1751875	0	0	none	0	0
ydhZ	1752956	0	0	none	0	0
ydiH	1762958	0	0	none	0	0
menI	1763246	0	0	none	0	0
ydiB	1772679	0	0	none	0	0
ydiQ	1778405	0	0	none	0	0
aroH	1787505	0	0	none	0	0
btuD	1790833	0	0	none	0	0
pheM	1797250	0	0	none	0	0
yniD	1803296	0	0	none	0	0
yniA	1806680	0	0	none	0	0
katE	1814152	0	0	none	0	0

ydjZ	1833539	0	1	ynjA	0	0
ydjA	1846149	0	0	none	0	0
msrB	1860040	0	0	none	0	0
yeaE	1862806	0	0	none	0	0
mipA	1863750	0	0	none	0	0
yeaJ	1871555	0	0	none	0	0
yeaO	1875280	0	0	none	0	0
yeaP	1876764	0	0	none	0	0
yeaR	1877613	0	0	none	0	0
leuE	1878145	0	0	none	0	0
dmlR	1878910	0	0	none	0	0
yeaV	1882657	0	0	none	0	0
fadD	1886085	0	0	none	0	0
yoaC	1892456	0	0	none	0	0
nudL	1894772	0	0	none	0	0
cspC	1905250	0	0	none	0	0
mgrB	1906647	0	0	none	0	0
kdgR	1907332	0	0	none	0	0
proQ	1912860	0	0	none	0	0
rsmF	1919686	0	0	none	0	0
yebV	1920040	0	0	none	0	0
yebZ	1921743	0	0	none	0	0
yebG	1928481	0	0	none	0	0
ruvA	1943389	0	0	none	0	0
yebC	1945435	0	0	none	0	0
nudB	1946204	0	0	none	0	0
yecE	1950237	0	0	none	0	0
torY	1955056	0	0	none	0	0

cutC	1956544	0	0	none	0	0
flhE	1960604	0	0	none	0	0
cheZ	1964417	0	0	none	0	0
cheY	1965072	0	0	none	0	0
tap	1967407	0	0	none	0	0
tar	1969054	0	0	none	0	0
cheA	1971384	0	0	none	0	0
motB	1973353	0	0	none	0	0
flhD	1975871	0	0	none	0	0
araH	1980578	0	0	none	0	0
araG	1981579	0	0	none	0	0
ftnB	1985452	0	0	none	0	0
yecH	1987275	0	0	none	0	0
tyrP	1988916	0	0	none	0	0
uvrC	1990898	0	0	none	0	0
sdiA	1994134	0	0	none	0	0
fliS	2003737	0	1	fliT	0	0
fliT	2004102	0	0	none	0	0
yedK	2008513	0	0	none	0	0
yedL	2009103	0	0	none	0	0
fliO	2019893	0	1	fliP	0	0
yodD	2023237	0	0	none	0	0
yodC	2026212	0	0	none	0	0
yedJ	2030408	0	0	none	0	0
zinT	2040049	0	0	none	0	0
yeeJ	2050038	0	0	none	0	0
shiA	2052983	0	0	none	0	0
flu	2072682	0	0	none	0	0

yeeT	2075062	0	0	none	0	0
cbtA	2075967	0	0	none	0	0
dacD	2079405	0	0	none	0	0
yeeD	2082250	0	0	none	0	0
yoeB	2087235	0	0	none	0	0
hisG	2089115	0	0	none	0	0
hisC	2091492	0	1	hisB	0	0
hisH	2093149	0	1	hisA	0	0
hisI	2095249	0	0	none	0	0
wbbK	2101415	0	0	none	0	0
wbbH	2104084	0	0	none	0	0
wzxB	2106361	0	0	none	0	0
wzxC	2116704	0	0	none	0	0
wcaJ	2118184	0	0	none	0	0
wcaE	2126928	0	0	none	0	0
yegH	2137509	0	0	none	0	0
yegE	2144607	0	0	none	0	0
alkA	2144716	0	0	none	0	0
yegI	2147063	0	0	none	0	0
yegK	2149735	0	0	none	0	0
mdtA	2153287	0	1	mdtB	0	0
yegQ	2165053	0	0	none	0	0
yegR	2166013	0	0	none	0	0
gatD	2169857	0	0	none	0	0
gatB	2172304	0	0	none	0	0
fbaB	2175534	0	0	none	0	0
thiD	2181738	0	0	none	0	0
yehL	2203706	0	0	none	0	0

yehP	2207125	0	0	none	0	0
mglC	2234765	0	0	none	0	0
folE	2241006	0	0	none	0	0
yeiG	2242768	0	0	none	0	0
cirA	2242800	0	0	none	0	0
nfo	2249719	0	0	none	0	0
nupX	2250917	0	0	none	0	0
fruK	2259449	0	0	none	0	0
yejF	2275911	0	0	none	0	0
bcr	2276592	0	0	none	0	0
atoE	2324100	0	0	none	0	0
yfaP	2325389	0	0	none	0	0
ubiG	2338311	0	0	none	0	0
yfaE	2346790	0	0	none	0	0
rhmA	2356064	0	0	none	0	0
yfaY	2360453	0	0	none	0	0
yfaZ	2361755	0	0	none	0	0
arnD	2368930	0	1	arnT	0	0
menH	2374856	0	0	none	0	0
yfbK	2382017	0	0	none	0	0
yfbM	2385459	0	0	none	0	0
nuoL	2391227	0	0	none	0	0
yfbS	2407542	0	0	none	0	0
yfbT	2409461	0	0	none	0	0
yfbU	2410122	0	0	none	0	0
ackA	2412694	0	0	none	0	0
yfcD	2416656	0	0	none	0	0
yfcF	2417863	0	0	none	0	0

yfcH	2420623	0	0	none	0	0
hisM	2422539	0	0	none	0	0
argT	2425031	0	0	none	0	0
purF	2426743	0	0	none	0	0
epmC	2442225	0	0	none	0	0
smrB	2447179	0	0	none	0	0
yfcS	2449606	0	0	none	0	0
sixA	2454349	0	0	none	0	0
fadL	2460668	0	0	none	0	0
intS	2465724	0	0	none	0	0
yfdK	2469099	0	0	none	0	0
yfdQ	2472878	0	0	none	0	0
emrY	2478660	0	0	none	0	0
frc	2490026	0	0	none	0	0
ypdA	2498390	0	0	none	0	0
fryA	2500012	0	0	none	0	0
ypdE	2502532	0	0	none	0	0
xapA	2522067	0	0	none	0	0
yfeN	2523913	0	0	none	0	0
cysA	2537739	0	0	none	0	0
cysP	2540534	0	0	none	0	0
ucpA	2541854	0	0	none	0	0
murR	2542774	0	0	none	0	0
murP	2546119	0	0	none	0	0
eutR	2552152	0	0	none	0	0
eutB	2555340	0	0	none	0	0
intZ	2558088	0	0	none	0	0
yffM	2559635	0	0	none	0	0

yffP	2561139	0	0	none	0	0
yffR	2562394	0	0	none	0	0
eutA	2563503	0	0	none	0	0
eutD	2570511	0	0	none	0	0
tktB	2579661	0	0	none	0	0
nudK	2580925	0	0	none	0	0
aegA	2581568	0	0	none	0	0
dapE	2590756	0	0	none	0	0
tmcA	2591866	0	0	none	0	0
purC	2594927	0	0	none	0	0
bcp	2598970	0	0	none	0	0
hyfJ	2609892	0	0	none	0	0
focB	2612804	0	0	none	0	0
yfgF	2624717	0	0	none	0	0
guaA	2628980	0	0	none	0	0
guaB	2630626	0	0	none	0	0
yfgM	2636685	0	0	none	0	0
rlmN	2641151	0	0	none	0	0
iscA	2657585	0	0	none	0	0
hcaF	2668930	0	1	hcaC	0	0
yphE	2674872	0	0	none	0	0
yphH	2682078	0	0	none	0	0
glrR	2685491	0	0	none	0	0
yfhH	2697629	0	0	none	0	0
recO	2699763	0	0	none	0	0
rseC	2705344	0	0	none	0	0
yfiF	2715513	0	0	none	0	0
pka	2720635	0	0	none	0	0

pheL	2735668	0	0	none	0	0
pheA	2736927	0	0	none	0	0
alpA	2756878	0	0	none	0	0
yfjI	2758416	0	0	none	0	0
yfjK	2759373	0	0	none	0	0
rnlB	2765377	0	0	none	0	0
yfjP	2766595	0	0	none	0	0
yfjT	2769637	0	0	none	0	0
yfjY	2774890	0	0	none	0	0
ypjJ	2775099	0	0	none	0	0
ypjA	2776168	0	0	none	0	0
gabP	2793675	0	0	none	0	0
ygaP	2796066	0	0	none	0	0
alaE	2797635	0	0	none	0	0
nrdF	2802483	0	0	none	0	0
gshA	2812905	0	0	none	0	0
recX	2820161	0	0	none	0	0
pncC	2821871	0	0	none	0	0
srlD	2826538	0	0	none	0	0
ascF	2839003	0	0	none	0	0
hycF	2842232	0	0	none	0	0
hycD	2844511	0	0	none	0	0
hycC	2845437	0	0	none	0	0
hycA	2847996	0	0	none	0	0
hypA	2849019	0	0	none	0	0
ygbN	2864487	0	0	none	0	0
cysC	2871409	0	0	none	0	0
iap	2875640	0	0	none	0	0

ygbF	2876591	0	0	none	0	0
ygbT	2876877	0	0	none	0	0
casE	2877810	0	0	none	0	0
casC	2879073	0	0	none	0	0
queD	2890601	0	0	none	0	0
ygcP	2892793	0	0	none	0	0
queE	2902769	0	0	none	0	0
gudX	2917428	0	0	none	0	0
yqcA	2920557	0	0	none	0	0
fucP	2933573	0	0	none	0	0
fucK	2936908	0	0	none	0	0
fucR	2938121	0	0	none	0	0
mltA	2944103	0	0	none	0	0
amiC	2945779	0	0	none	0	0
recC	2957082	0	0	none	0	0
ygdQ	2969155	0	0	none	0	0
araE	2978786	0	0	none	0	0
yqeH	2986190	0	0	none	0	0
yqeK	2987957	0	0	none	0	0
ygeR	2997158	0	0	none	0	0
xdhA	3000625	0	0	none	0	0
ygfS	3026546	0	0	none	0	0
ygfT	3027034	0	0	none	0	0
prfB	3033206	0	0	none	0	0
xerD	3036869	0	0	none	0	0
fldB	3038398	0	0	none	0	0
gcvT	3047595	0	0	none	0	0
ubiH	3050362	0	0	none	0	0

zapA	3053963	0	0	none	0	0
rpiA	3056688	0	0	none	0	0
yggC	3071998	0	0	none	0	0
yggF	3073239	0	0	none	0	0
cmtA	3075493	0	0	none	0	0
yqgB	3083942	0	0	none	0	0
yqgE	3091522	0	1	yqgF	1	0
yggR	3092122	0	0	none	0	0
yqgA	3108282	0	0	none	0	0
glcE	3123492	0	0	none	0	0
glcD	3124544	0	0	none	0	0
yghS	3131266	0	0	none	0	0
hybG	3137738	0	0	none	0	0
yghW	3144472	0	0	none	0	0
gpr	3146959	0	0	none	0	0
yqhA	3146999	0	0	none	0	0
yghB	3152244	0	0	none	0	0
yqhG	3156598	0	0	none	0	0
yqhH	3156906	0	0	none	0	0
ygiS	3164133	0	0	none	0	0
ygiV	3166771	0	0	none	0	0
yqiA	3173447	0	0	none	0	0
cpdA	3174028	0	0	none	0	0
tolC	3177618	0	0	none	0	0
ygiB	3178437	0	0	none	0	0
glgS	3189761	0	0	none	0	0
glnE	3194823	0	0	none	0	0
plsY	3203333	0	0	none	0	0

dnaG	3210874	0	0	none	0	0
ygjP	3234485	0	0	none	0	0
exuR	3245450	0	0	none	0	0
yhaO	3254701	0	0	none	0	0
tdcG	3256307	0	0	none	0	0
tdcE	3258146	0	0	none	0	0
yhaB	3266415	0	0	none	0	0
garK	3268647	0	0	none	0	0
garD	3274875	0	0	none	0	0
yhaV	3275823	0	0	none	0	0
agaR	3275878	0	0	none	0	0
agaV	3278712	0	0	none	0	0
yraQ	3295120	0	0	none	0	0
yhbU	3300502	0	0	none	0	0
yhbX	3318010	0	0	none	0	0
yhbE	3329792	0	0	none	0	0
sfsB	3333209	0	0	none	0	0
ibaG	3334571	0	0	none	0	0
mlaE	3336488	0	0	none	0	0
lptB	3342691	0	0	none	0	0
rpoN	3344172	0	0	none	0	0
arcB	3348711	0	0	none	0	0
yhcD	3363210	0	0	none	0	0
yhcF	3365664	0	0	none	0	0
nanA	3370705	0	0	none	0	0
aaeX	3387156	0	0	none	0	0
yhdP	3390480	0	0	none	0	0
yhdU	3410822	0	0	none	0	0

yhdY	3420450	0	0	none	0	0
rsmB	3434518	0	0	none	0	0
zntR	3436727	0	0	none	0	0
rpmJ	3440640	0	0	none	0	0
gspA	3451951	0	0	none	0	0
gspD	3456351	0	0	none	0	0
bfr	3464271	0	0	none	0	0
tusB	3472700	0	0	none	0	0
tusC	3472995	0	0	none	0	0
tusD	3473354	0	0	none	0	0
slyX	3475880	0	0	none	0	0
kefG	3478629	0	0	none	0	0
yheS	3481224	0	1	yheT	0	0
yhfA	3483436	0	0	none	0	0
pabA	3488288	0	0	none	0	0
yhfY	3509082	0	0	none	0	0
hofM	3519994	0	0	none	0	0
mrcA	3523445	0	0	none	0	0
yrfG	3527359	0	0	none	0	0
yhgE	3528737	0	0	none	0	0
gntT	3545897	0	0	none	0	0
rtcA	3553855	0	0	none	0	0
glgX	3567369	0	0	none	0	0
yhhZ	3581064	0	0	none	0	0
yrhB	3583066	0	0	none	0	0
livK	3594474	0	0	none	0	0
yhhN	3604400	0	0	none	0	0
tusA	3606774	0	0	none	0	0

yhhT	3610937	0	0	none	0	0
acpT	3611579	0	0	none	0	0
nikR	3617012	0	0	none	0	0
rbbA	3624826	0	0	none	0	0
yhiJ	3628991	0	0	none	0	0
yhiN	3634231	0	0	none	0	0
prlC	3641163	0	0	none	0	0
dinQ	3645728	0	0	none	0	0
dctR	3653236	0	0	none	0	0
gadW	3661913	0	0	none	0	0
gadA	3664203	0	0	none	0	0
yhjB	3669315	0	0	none	0	0
yhjC	3671336	0	0	none	0	0
yhjG	3674313	0	0	none	0	0
bcsA	3690641	0	0	none	0	0
yiaB	3725430	0	0	none	0	0
xylH	3732924	0	0	none	0	0
ysaA	3739132	0	0	none	0	0
yiaK	3741754	0	0	none	0	0
aldB	3752996	0	0	none	0	0
yibF	3759370	0	0	none	0	0
yibH	3768266	0	0	none	0	0
secB	3781684	0	0	none	0	0
kbl	3789378	0	0	none	0	0
waaF	3793998	0	0	none	0	0
yicL	3837194	0	0	none	0	0
nepI	3838572	0	0	none	0	0
yicN	3839973	0	0	none	0	0

uhpC	3845328	0	0	none	0	0
uhpA	3848159	0	0	none	0	0
yidF	3853137	0	0	none	0	0
yidJ	3854934	0	0	none	0	0
yidQ	3866083	0	0	none	0	0
tnaC	3886532	0	0	none	0	0
yieE	3892653	0	0	none	0	0
bglG	3903754	0	0	none	0	0
pstB	3905616	0	0	none	0	0
pstA	3906572	0	0	none	0	0
asnC	3924568	0	0	none	0	0
rbsD	3931793	0	0	none	0	0
rbsK	3936246	0	0	none	0	0
yieP	3938658	0	0	none	0	0
ilvM	3950487	0	0	none	0	0
ilvY	3954950	0	0	none	0	0
wecC	3970545	0	1	rffG	0	0
wecD	3973164	0	0	none	0	0
wecE	3974299	0	0	none	0	0
wecG	3978719	0	0	none	0	0
cyaA	3991722	0	0	none	0	0
yigF	4000442	0	0	none	0	0
yigI	4002253	0	0	none	0	0
pldA	4003754	0	0	none	0	0
rhtC	4006400	0	0	none	0	0
rhtB	4006462	0	0	none	0	0
ubiE	4017633	0	0	none	0	0
fre	4025251	0	0	none	0	0

pepQ	4030515	0	1	yigZ	0	0
trkH	4032619	0	0	none	0	0
yshB	4051670	0	0	none	0	0
glnG	4051892	0	0	none	0	0
ompL	4061626	0	0	none	0	0
yiiE	4077532	0	0	none	0	0
fdoG	4080795	0	0	none	0	0
frvR	4086130	0	0	none	0	0
sbp	4107846	0	0	none	0	0
yiiQ	4109638	0	0	none	0	0
yiiS	4111289	0	0	none	0	0
glpX	4112592	0	0	none	0	0
zapB	4116783	0	0	none	0	0
rraA	4116868	0	0	none	0	0
yijE	4135036	0	0	none	0	0
thiH	4188758	0	0	none	0	0
thiS	4190660	0	0	none	0	0
thiC	4192227	0	0	none	0	0
rsd	4194355	0	0	none	0	0
nudC	4195699	0	0	none	0	0
yjaH	4199284	0	0	none	0	0
iclR	4220827	0	0	none	0	0
metH	4225534	0	0	none	0	0
rluF	4229249	0	0	none	0	0
dinF	4257144	0	0	none	0	0
aphA	4268150	0	0	none	0	0
yjbR	4269037	0	0	none	0	0
yjcB	4272783	0	0	none	0	0

yjcF	4279806	0	0	none	0	0
fdhF	4295242	0	0	none	0	0
rpiB	4311822	0	0	none	0	0
yjdP	4312220	0	0	none	0	0
phnP	4312367	0	0	none	0	0
phnM	4314105	0	0	none	0	0
phnK	4316029	0	0	none	0	0
phnH	4318686	0	0	none	0	0
phnF	4319720	0	0	none	0	0
phnC	4322400	0	0	none	0	0
yjdM	4324422	0	0	none	0	0
adiY	4335191	0	0	none	0	0
melB	4342813	0	0	none	0	0
yjdJ	4350380	0	0	none	0	0
yjdC	4360756	0	0	none	0	0
dsbD	4361368	0	0	none	0	0
yjeI	4371185	0	0	none	0	0
yjeO	4384041	0	0	none	0	0
rsgA	4388480	0	0	none	0	0
miaA	4398225	0	0	none	0	0
yjeT	4402606	0	0	none	0	0
nsrR	4404638	0	0	none	0	0
rnr	4407118	0	0	none	0	0
yjfN	4414040	0	0	none	0	0
ytfE	4429344	0	0	none	0	0
ytfF	4430114	0	0	none	0	0
ytfQ	4448941	0	0	none	0	0
nrdG	4457923	0	0	none	0	0

treR	4464322	0	0	none	0	0
bdcR	4472740	0	0	none	0	0
tabA	4473337	0	0	none	0	0
idnT	4489229	0	0	none	0	0
idnK	4493209	0	0	none	0	0
yjgB	4493213	0	0	none	0	0
yjhB	4503298	0	0	none	0	0
insI4	4505489	0	0	none	0	0
fecA	4512376	0	0	none	0	0
yjhG	4520150	0	0	none	0	0
yjhH	4522128	0	0	none	0	0
yjhI	4523038	0	0	none	0	0
sgcR	4524129	0	0	none	0	0
yjhQ	4531262	0	0	none	0	0
nanS	4534637	0	0	none	0	0
fimE	4540656	0	0	none	0	0
fimG	4546811	0	0	none	0	0
uxuR	4553372	0	0	none	0	0
yjiJ	4559520	0	0	none	0	0
yjiL	4561945	0	0	none	0	0
mdtM	4565310	0	0	none	0	0
yjiS	4569938	0	0	none	0	0
hsdS	4578091	0	0	none	0	0
hsdR	4581272	0	0	none	0	0
mrr	4585886	0	0	none	0	0
yjjN	4595035	0	0	none	0	0
yjjA	4597718	0	0	none	0	0
fhuF	4602898	0	0	none	0	0

yjjZ	4604063	0	0	none	0	0
yjjG	4607346	0	0	none	0	0
yjjW	4612703	0	0	none	0	0
deoB	4618849	0	0	none	0	0
ytjB	4622168	0	0	none	0	0
slt	4630693	0	0	none	0	0
trpR	4631109	0	0	none	0	0
ileS	25207	1	1	lspA	1	0
lptD	54755	1	0	none	0	0
murD	98403	1	1	ftsW	1	0
hemL	173602	1	0	none	0	0
frr	193429	1	0	none	0	0
dxr	194717	1	0	none	0	0
ispU	195664	1	0	none	0	0
rseP	197898	1	0	none	0	0
fabZ	202556	1	0	none	0	0
proS	217057	1	0	none	0	0
secD	428718	1	0	none	0	0
ribD	433782	1	0	none	0	0
ribE	434341	1	0	none	0	0
dnaX	493247	1	0	none	0	0
nadD	669154	1	0	none	0	0
holA	669797	1	0	none	0	0
lptE	670828	1	0	none	0	0
fldA	710158	1	0	none	0	0
infA	925448	1	0	none	0	0
fabA	1015175	1	0	none	0	0
yceQ	1144045	1	0	none	0	0

fabG	1150627	1	0	none	0	0
ydfB	1646817	1	0	none	0	0
ribC	1740625	1	0	none	0	0
pheT	1793581	1	0	none	0	0
tsaB	1888596	1	0	none	0	0
aspS	1946774	1	0	none	0	0
pgsA	1990293	1	0	none	0	0
yejM	2284158	1	0	none	0	0
nrdA	2345172	1	0	none	0	0
nrdB	2346536	1	1	yfaE	0	0
accD	2431034	1	0	none	0	0
ligA	2526183	1	0	none	0	0
zipA	2528269	1	0	none	0	0
rpoE	2707459	1	0	none	0	0
lgt	2963184	1	0	none	0	0
ribB	3181835	1	0	none	0	0
cca	3201151	1	0	none	0	0
fmt	3433183	1	0	none	0	0
rplV	3447923	1	0	none	0	0
rpsG	3471564	1	0	none	0	0
waaU	3796262	1	0	none	0	0
waaA	3807840	1	0	none	0	0
gmk	3820074	1	0	none	0	0
yidC	3884745	1	0	none	0	0
glmU	3911853	1	0	none	0	0
ubiB	4019889	1	0	none	0	0
ubiD	4024504	1	0	none	0	0
ftsN	4120403	1	0	none	0	0

murI	4164308	1	0	none	0	0
secE	4175764	1	0	none	0	0
ubiA	4251911	1	0	none	0	0
ssb	4272684	1	0	none	0	0
alsK	4304893	1	0	none	0	0
lptF	4485341	1	1	lptG	1	0

Supplemental Table 2. Strain genotypes.

Ancestor Reference	genotype	plasmid	UAG codons	UGA essential codons	Total UGA codons recoded	Total UGA codons deleted	Total UGA codons	genomic prfA	genomic prfB
doi:10.1038/nature08187	MG1655 ΔmutS::cat.Δ(ybhB-bioAB)::[λcI857 N(cro-ea59)::tetR-bla])	none	321	71	0	0	1178	present	present
doi: 10.1126/science.1241459	MG1655 dnaG.Q576A.exoX.K28TAA.ΔtolC.ΔprfA.ΔmutS::zeo.Δ(ybhB-bioAB)::[λcI857 N(cro-ea59)::tetR-bla]).12B.tolQRA	none	0	71	0	0	1178	deleted	present
This Study	MG1655 dnaG.Q576A.exoX.K28TAA.ΔprfA.ΔmutS::zeo.Δ(ybhB-bioAB)::[λcI857 N(cro-ea59)::tetR-bla]).12B.tolQRA.Conj-spec	none	0	36	37	2	1139	deleted	present
This Study	MG1655 dnaG.Q576A.exoX.K28TAA.ΔprfA.ΔmutS::zeo.Δ(ybhB-bioAB)::[λcI857 N(cro-ea59)::tetR-bla]).12B.tolQRA.Conj-tolC.Conj-gent	none	0	35	63	1	1114	deleted	present
This Study	MG1655 dnaG.Q576A.exoX.K28TAA.ΔprfA.ΔmutS::zeo.Δ(ybhB-bioAB)::[λcI857 N(cro-ea59)::tetR-bla]).12B.tolQRA.Conj-tolC.Conj-gent.Conj-spec	none	0	5	101	3	1074	deleted	present
This Study	MG1655 dnaG.Q576A.exoX.K28TAA.ΔprfA.ΔmutS::zeo.Δ(ybhB-bioAB)::[λcI857 N(cro-ea59)::tetR-bla]).12B.tolQRA	none	0	0	106	3	1069	deleted	present
This Study	MG1655 ΔprfA.ΔmutS::zeo	none	0	71	0	0	1178	deleted	present
This Study	MG1655 dnaG.Q576A.exoX.K28TAA.ΔtolC.ΔmutS::zeo.Δ(ybhB-	none	0	0	106	3	1069	present	present

	bioAB)::[λcI857 N(cro-ea59)::tetR-bla]).12B.tolQRA								
This Study	MG1655 dnaG.Q576A.exoX.K28TAA.ΔtolC.ΔprfB::tolC.ΔmutS::zeo.Δ(ybhB-bioAB)::[λcI857 N(cro-ea59)::tetR-bla]).12B.tolQRA	pBAD prfB	0	0	106	3	1069	present	deleted

*12B refers to the locus between *ycgV* and *ychF*

Supplemental Table 3. Deletions. Numbering is from MG1655

Blattner deletion name [1,2]	Position of Deletion	Position of Deletion	Gene containing UGA codon
MD13	15388	20562	nhaA
MD31	728588	738185	no TGA codons deleted
MD33	3617014	3623701	no TGA codons deleted
MD23	2284423	2288202	no TGA codons deleted
MD24	3360184	3365664	yhcD, yhcF
Citations			
1	doi: 10.1101/gr.217202		
2	doi: 10.1126/science.1126439		

Supplemental Table 4. Recoding

Gene Name	Sense	Replicore	LP	RP	Essential	Overlapping	Extended Overlapping	Overlapped gene	Overlapped gene essential?	Bases of overlap	Deleted or Converted	Strain in which it was converted	allele frequency
ileS	+	1	252047	252207	1	0	0	lspA	1	0	Converted	C321.ΔA.2.64	8%
fixB	+	1	441126	441129	0	0	0	none	0	0	Converted	C321.ΔA.2.64	2%
fixC	+	1	454463	454466	0	1	0	fixX	0	3	Converted	C321.ΔA.2.64	8%
fixX	+	1	457447	457500	0	0	0	none	0	0	Converted	C321.ΔA.2.64	2%
pdxA	-	1	524426	524429	0	1	0	rsmA	0	3	Converted	C321.ΔA.2.64	10%
lptD	-	1	547554	547557	1	0	0	none	0	0	Converted	C321.ΔA.2.64	50%
djlA	+	1	581176	581179	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
polB	-	1	634428	634431	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
thiP	-	1	729	729	0	1	1	thiQ	0	16	Converted	C321.ΔA.2.64	Converted before

			10	13									screening
murE	+	1	946503	946533	1	1	0	murF	1	3	Converted	C321.ΔA.2.64	8%
murD	+	1	984000	984003	1	0	0	ftsW	1	0	Converted	C321.ΔA.2.64	28%
ftsW	+	1	996444	996447	1	1	0	murG	1	3	Converted	C321.ΔA.2.64	14%
murC	+	1	1022237	1022240	1	1	1	ddlB	0	7	Converted	C321.ΔA.2.64	19%
ftsQ	+	1	1039822	1039855	1	1	0	ftsA	1	3	Converted	C321.ΔA.2.64	6%
hemL	-	1	1736001	1736004	1	0	0	none	0	0	Converted	C321.ΔA.2.64	2%
glnD	-	1	1859777	1859800	0	0	0	none	0	0	Converted	C321.ΔA.2.64	40%
frr	+	1	1934266	1934299	1	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
rseP	+	1	1978	1978	1	0	0	none	0	0	Converted	C321.ΔA.2.64	6%

			95	98									
fabZ	+	1	2025533	2025556	1	0	0	none	0	0	Converted	C321.ΔA.2.64	7%
lpxB	+	1	2044493	2044496	1	1	0	rnhB	0	3	Converted	C321.ΔA.2.64	4%
yaeP	-	1	21392 24	21392 27	0	1	1	rof	0	7	Converted	C321.ΔA.2.64	40%
yaeQ	+	1	2148333	2148336	0	1	0	arfB	0	3	Converted	C321.ΔA.2.64	25%
proS	-	1	2170056	2170059	1	0	0	none	0	0	Converted	C321.ΔA.2.64	19%
gmhB	+	1	2234005	2234008	0	0	0	none	0	0	Converted	C321.ΔA.2.64	4%
psiF	+	1	4028222	4028225	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
yaiA	+	1	4063391	4063394	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
sbcD	-	1	41	41	0	1	0	sbcC	0	3	Converted	C321.ΔA.2.64	Converted

			4973	4976									before screening
secD	+	1	4287115	4287118	1	0	0	none	0	0	Converted	C321.ΔA.2.64	21%
ribD	+	1	4333779	4333782	1	0	0	none	0	0	Converted	C321.ΔA.2.64	6%
ribE	+	1	4343338	4343341	1	0	0	none	0	0	Converted	C321.ΔA.2.64	29%
dnaX	+	1	4932444	4932447	1	0	0	none	0	0	Converted	C321.ΔA.2.64	4%
nadD	-	1	6691153	6691156	1	0	0	none	0	0	Converted	C321.ΔA.2.64	35%
holA	-	1	6697796	6697799	1	0	0	none	0	0	Converted	C321.ΔA.2.64	23%
lptE	-	1	6708827	6708830	1	0	0	none	0	0	Converted	C321.ΔA.2.64	8%
fldA	-	1	7101157	7101160	1	0	0	none	0	0	Converted	C321.ΔA.2.64	2%

ybgA	+	1	738730	738733	0	1	0	phr	0	3	Converted	C321.ΔA.2.64	Converted before screening
ybgJ	+	1	7434469	7434472	0	1	1	ybgK	0	6	Converted	C321.ΔA.2.64	17%
ybgK	+	1	7443955	7443988	0	1	1	ybgL	0	10	Converted	C321.ΔA.2.64	17%
ybgO	-	1	7471443	7471446	0	0	0	none	0	0	Converted	C321.ΔA.2.64	3%
sdhD	+	1	7551127	7551130	0	0	0	sdhA	0	0	Converted	C321.ΔA.2.64	Converted before screening
cydX	+	1	7735229	7735532	0	0	0	ybgE	0	0	Converted	C321.ΔA.2.64	14%
ybgC	+	1	7743776	7743779	0	1	0	tolQ	0	3	Converted	C321.ΔA.2.64	14%
galK	-	1	7880053	7880056	0	1	1	galM	0	6	Converted	C321.ΔA.2.64	17%
tnaC	+	2	38886	38886	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before

			529	532									screening
yieE	+	2	3892650	3892653	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
bglG	-	2	3903753	3903756	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
pstB	-	2	3905615	3905618	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
pstA	-	2	3906571	3906574	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
rbsD	+	1	3931790	3931793	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
rbsK	+	1	3936243	3936246	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
yieP	-	1	3938657	3938660	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening

ilvM	+	1	3950484	3950487	0	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
hemC	-	1	39878477	39878500	1	1	0	hemD	1	3	Converted	C321.ΔA.2.64	8%
yigP	+	1	40182249	40182252	1	1	0	ubiB	1	3	Converted	C321.ΔA.2.64	12%
ubiB	+	1	40198486	40198489	1	0	0	none	0	0	Converted	C321.ΔA.2.64	28%
ubiD	+	1	40245401	40245404	1	0	0	none	0	0	Converted	C321.ΔA.2.64	19%
ftsN	-	1	41204002	41204005	1	0	0	none	0	0	Converted	C321.ΔA.2.64	Converted before screening
murI	+	1	41643005	41643008	1	0	0	none	0	0	Converted	C321.ΔA.2.64	8%
murB	+	1	417	417	1	1	0	birA	1	3	Converted	C321.ΔA.2.64	23%

			1105	1108									
secE	+	1	4175761	4175764	1	0	0	none	0	0	Converted	C321.ΔA.2.64	2%
ubiA	+	1	42519081	42519111	1	0	0	none	0	0	Converted	C321.ΔA.2.64	17%
ssb	+	1	42726811	42726844	1	0	0	none	0	0	Converted	C321.ΔA.2.64	17%
alsK	-	1	43048922	43048955	1	0	0	none	0	0	Converted	C321.ΔA.2.64	8%
dxr	+	1	19471144	19471177	1	0	0	none	0	0	Converted	C321.ΔA.2.109.F	8%
ispU	+	1	19566161	19566644	1	0	0	none	0	0	Converted	C321.ΔA.2.109.F	Converted before screening
fmt	+	2	34331180	34331183	1	0	0	none	0	0	Converted	C321.ΔA.2.109.F	6%

glmU	-	2	3911852	3911855	1	0	0	none	0	0	Converted	C321.ΔA.2.109.F	25%
lptF	+	1	44853381	44853411	1	0	0	lptG	1	0	Converted	C321.ΔA.2.109.F	Converted before screening
infA	-	1	9254447	9254500	1	0	0	none	0	0	Converted	C321.ΔA.2.39	17%
msbA	+	1	9675889	9675922	1	1	0	lpxK	1	3	Converted	C321.ΔA.2.39	15%
mukF	+	1	97486161	97486464	1	1	1	mukE	1	20	Converted	C321.ΔA.2.39	Converted before screening
fabA	-	1	10151774	10151777	1	0	0	none	0	0	Converted	C321.ΔA.2.39	17%
yceQ	+	1	11440042	11440045	1	0	0	none	0	0	Converted	C321.ΔA.2.39	29%
fabG	+	1	115006	115006	1	0	0	none	0	0	Converted	C321.ΔA.2.39	19%

			24	27									
tmk	+	1	1154985	1154988	1	1	0	holB	1	3	Converted	C321.ΔA.2.39	4%
prmC	+	1	12661470	12661500	1	1	0	ychQ	0	3	Converted	C321.ΔA.2.39	13%
ydfB	+	2	16468144	16468177	1	0	0	none	0	0	Converted	C321.ΔA.2.39	25%
ribC	-	2	17406244	17406277	1	0	0	none	0	0	Converted	C321.ΔA.2.39	11%
pheT	-	2	17935800	17935833	1	0	0	none	0	0	Converted	C321.ΔA.2.39	Converted before screening
tsaB	-	2	18885955	18885988	1	0	0	none	0	0	Converted	C321.ΔA.2.39	13%
aspS	-	2	19467733	19467766	1	0	0	none	0	0	Converted	C321.ΔA.2.39	8%

pgsA	-	2	1990292	1990295	1	0	0	none	0	0	Converted	C321.ΔA.2.39	4%
yefM	-	2	2087485	2087488	1	1	0	yoeB	0	3	Converted	C321.ΔA.2.39	8%
yohJ	+	2	22290441	22290444	0	1	0	yohK	0	3	Converted	C321.ΔA.2.39	Converted before screening
folE	-	2	22410005	22410008	0	0	0	none	0	0	Converted	C321.ΔA.2.39	Converted before screening
yejF	+	2	22759908	22759111	0	0	0	none	0	0	Converted	C321.ΔA.2.39	Converted before screening
yejM	+	2	22841555	22841558	1	0	0	none	0	0	Converted	C321.ΔA.2.39	6%
nrdA	+	2	23451669	23451672	1	0	0	none	0	0	Converted	C321.ΔA.2.39	5%
nrdB	+	2	234	234	1	0	0	yfaE	0	0	Converted	C321.ΔA.2.39	13%

			6533	6536									
accD	-	2	2431033	2431036	1	0	0	none	0	0	Converted	C321.ΔA.2.39	3%
ligA	-	2	2526182	2526185	1	0	0	none	0	0	Converted	C321.ΔA.2.39	Converted before screening
zipA	-	2	2528268	2528271	1	0	0	none	0	0	Converted	C321.ΔA.2.39	Converted before screening
rpoE	-	2	2707458	2707461	1	0	0	none	0	0	Converted	C321.ΔA.2.39	21%
ispF	-	2	2869322	2869325	1	1	0	truD	0	3	Converted	C321.ΔA.2.39	4%
lgt	-	2	2963183	2963186	1	0	0	none	0	0	Converted	C321.ΔA.2.39	13%
ribB	-	2	3181818	3181818	1	0	0	none	0	0	Converted	C321.ΔA.2.39	Converted before screening

			34	37									
cca	+	2	3201148	3201151	1	0	0	none	0	0	Converted	C321.ΔA.2.39	24%
rplV	-	2	34479222	34479225	1	0	0	none	0	0	Converted	C321.ΔA.2.39	6%
rplD	-	2	34497002	34497005	1	1	0	rplW	1	3	Converted	C321.ΔA.2.39	2%
rpsG	-	2	34715563	34715566	1	0	0	none	0	0	Converted	C321.ΔA.2.39	8%
waaU	-	2	37962611	37962644	1	0	0	none	0	0	Converted	C321.ΔA.2.39	75%
waaA	+	2	38078337	38078400	1	0	0	none	0	0	Converted	C321.ΔA.2.39	Converted before screening
gmk	+	2	38200071	38200074	1	0	0	none	0	0	Converted	C321.ΔA.2.39	Converted before screening

rnpA	+	2	3882872	3882875	1	1	1	yidD	0	36	Converted	C321.ΔA.2.39	25%
yidC	+	2	3884742	3884745	1	0	0	none	0	0	Converted	C321.ΔA.2.39	Converted before screening
nhaA	+	1	174889	186555	0	0	0	none	0	0	Deleted	C321.ΔA.2.64	
yhcD	+	2	3362807	3365188	0	0	0	none	0	0	Deleted	C321.ΔA.2.39	
yhcF	+	2	3366926	3367642	0	0	0	none	0	0	Deleted	C321.ΔA.2.39	

Supplemental Table 5. Oligonucleotides used in dnaX mutation

name	sequence	purpose
dnaXmix1	C*C*GCGCGCGGGKAGCGGCTGCCGGNDDNNNCTTTTTTgctttggttgctccctgcacgcgctgcaactgctggcgcgccg	dnaX UGA randomize
dnaXmix2	G*T*GACCGAAGCCAGTCTTTCCAGCGCAGCGTTATTtACCGGCCGCGCGCGGGTAGCGGCTGCCGGNDDNNNCTTTTTTGCTTTGGTTGCT	dnaX UGA randomize
dnaXTAA1	C*C*GCGCGCGGGTAGCGGCTGCCGGTTtACTCTTTTTTgctttggttgctccctgcacgcgctgcaactgctggcgcgccg	dnaX UGA->UAA
dnaXTAA2	C*G*ATCGGTGACCGAAGCCAGTCTTTCCAGCGCAGCGTTATTCACCGGCCGCGCGCGGtTAGCGGCTGCCGGTCcaCTCTTTTTTGCTTTG	dnaX UGA->UAA
dnaXTAA3	G*T*GACCGAAGCCAGTCTTTCCAGCGCAGCGTTATTtACCGGCCGCGCGCGGGTAGCGGCTGCCGGTCcaCTCTTTTTTGCTTTGGTTGCT	dnaX UGA->UAA
dnaXTGAwt	ggagcaaccaaagcAAAAAAGAGTg	MASC PCR of dnaX UGA
dnaXTGArev	GGGATTATCATCTTCAACGATAGTCAGT	MASC PCR of dnaX UGA

Supplemental Table 6. SNVs and indels

Aligned to CP006698.1

strain	position	ref	allele	annotation	gene name	transcript biotype	mutation	amino acid mutation
C321.ΔA.2.104	5323	G	A	synonymous_variant	yaaX	protein_coding	c.90G>A	p.Pro30Pro
C321.ΔA.2.104	25207	G	A	splice_region_variant&stop_retained_variant	ileS	protein_coding	c.2816G>A	p.Ter939Ter
C321.ΔA.2.104	25295	A	G	missense_variant	lspA	protein_coding	c.88A>G	p.Ile30Val
C321.ΔA.2.104	30896	G	T	missense_variant	carB	protein_coding	c.79G>T	p.Asp27Tyr
C321.ΔA.2.104	44129	G	A	splice_region_variant&stop_retained_variant	fixB	protein_coding	c.941G>A	p.Ter314Ter
C321.ΔA.2.104	45750	G	A	splice_region_variant&stop_retained_variant	fixX	protein_coding	c.287G>A	p.Ter96Ter
C321.ΔA.2.104	54757	C	T	splice_region_variant&stop_retained_variant	lptD	protein_coding	c.2354G>A	p.Ter785Ter
C321.ΔA.2.104	58179	G	A	splice_region_variant&stop_retained_variant	djlA	protein_coding	c.815G>A	p.Ter272Ter
C321.ΔA.2.104	63431	C	T	splice_region_variant&stop_retain	polB	protein_coding	c.2351G>A	p.Ter784Ter

				ed_variant				
C321.ΔA.2.104	64263	C	T	missense_variant	polB	protein_coding	c.1519G>A	p.Ala507Thr
C321.ΔA.2.104	72743	C	T	missense_variant	thiQ	protein_coding	c.16G>A	p.Asp6Asn
C321.ΔA.2.104	72743	C	T	splice_region_variant&stop_retained_variant	thiP	protein_coding	c.1610G>A	p.Ter537Ter
C321.ΔA.2.104	79908	T	C	missense_variant	leuC	protein_coding	c.788A>G	p.Lys263Arg
C321.ΔA.2.104	85664	C	T	synonymous_variant	ilvI	protein_coding	c.204C>T	p.Val68Val
C321.ΔA.2.104	88188	T	C	synonymous_variant	cra	protein_coding	c.330T>C	p.Ile110Ile
C321.ΔA.2.104	98233	G	A	splice_region_variant&stop_retained_variant	murD	protein_coding	c.1316G>A	p.Ter439Ter
C321.ΔA.2.104	102070	G	A	missense_variant	ddlB	protein_coding	c.7G>A	p.Asp3Asn
C321.ΔA.2.104	102070	G	A	splice_region_variant&stop_retained_variant	murC	protein_coding	c.1475G>A	p.Ter492Ter
C321.ΔA.2.104	102995	G	A	missense_variant	ftsQ	protein_coding	c.10G>A	p.Ala4Thr
C321.ΔA.2.104	116886	T	C	synonymous_variant	hofB	protein_coding	c.45A>G	p.Gly15Gly
C321.ΔA.2.104	133742	T	A	missense_variant	acnB	protein_coding	c.2297T>A	p.Ile766Asn
C321.ΔA.2.104	156361	A	G	synonymous_variant	yadN	protein_coding	c.354T>C	p.Ser118Ser

C321.ΔA.2.104	160814	C	T	synonymous_variant	sfsA	protein_coding	c.504G>A	p.Ala168Ala
C321.ΔA.2.104	171836	A	G	missense_variant	fhuB	protein_coding	c.544A>G	p.Thr182Ala
C321.ΔA.2.104	173434	C	T	splice_region_variant&stop_retained_variant	hemL	protein_coding	c.1280G>A	p.Ter427Ter
C321.ΔA.2.104	185809	C	T	splice_region_variant&stop_retained_variant	glnD	protein_coding	c.2672G>A	p.Ter891Ter
C321.ΔA.2.104	193258	G	A	splice_region_variant&stop_retained_variant	frr	protein_coding	c.557G>A	p.Ter186Ter
C321.ΔA.2.104	197727	G	A	splice_region_variant&stop_retained_variant	rseP	protein_coding	c.1352G>A	p.Ter451Ter
C321.ΔA.2.104	202385	G	A	splice_region_variant&stop_retained_variant	fabZ	protein_coding	c.455G>A	p.Ter152Ter
C321.ΔA.2.104	213756	C	T	missense_variant	rof	protein_coding	c.7G>A	p.Asp3Asn
C321.ΔA.2.104	213756	C	T	splice_region_variant&stop_retained_variant	yaeP	protein_coding	c.200G>A	p.Ter67Ter
C321.ΔA.2.104	216888	C	T	splice_region_variant&stop_retained_variant	proS	protein_coding	c.1718G>A	p.Ter573Ter

C321.ΔA.2.104	221945	A	G	synonymous_variant	metN	protein_coding	c.531T>C	p.Arg177Arg
C321.ΔA.2.104	223237	G	A	splice_region_variant&stop_retained_variant	gmhB	protein_coding	c.575G>A	p.Ter192Ter
C321.ΔA.2.104	223835	C	T	non_coding_transcript_exon_variant	rrsH	rRNA	n.235C>T	
C321.ΔA.2.104	236026	A	G	missense_variant	dnaQ	protein_coding	c.130A>G	p.Thr44Ala
C321.ΔA.2.104	253368	T	C	missense_variant	ykfJ	protein_coding	c.61T>C	p.Trp21Arg
C321.ΔA.2.104	257670	A	G	missense_variant	crl	protein_coding	c.13A>G	p.Ser5Gly
C321.ΔA.2.104	264888	T	C	missense_variant	ykfB	protein_coding	c.253A>G	p.Thr85Ala
C321.ΔA.2.104	295866	C	T	missense_variant	intF	protein_coding	c.283G>A	p.Asp95Asn
C321.ΔA.2.104	296893	A	G	missense_variant	paoD	protein_coding	c.886T>C	p.Ser296Pro
C321.ΔA.2.104	299130	A	G	missense_variant	paoC	protein_coding	c.857T>C	p.Leu286Pro
C321.ΔA.2.104	311939	G	A	intergenic_region	ykgM-ykgR		n.311939G>A	
C321.ΔA.2.104	336301	T	C	missense_variant	yahF	protein_coding	c.473T>C	p.Leu158Pro
C321.ΔA.2.104	348768	T	C	intergenic_region	prpB-prpC		n.348768T>C	
C321.ΔA.2.104	362284	A	G	intragenic_variant	lacZ		n.362284A>G	
C321.ΔA.2.104	369824	A	G	missense_variant	mhpB	protein_coding	c.497A>G	p.Asn166Ser
C321.ΔA.2.104	402649	G	A	splice_region_variant&stop_retain	psiF	protein_coding	c.320G>A	p.Ter107Ter

				ed_variant				
C321.ΔA.2.104	406218	G	A	splice_region_variant&stop_retained_variant	yaiA	protein_coding	c.191G>A	p.Ter64Ter
C321.ΔA.2.104	411506	C	T	missense_variant	araJ	protein_coding	c.25G>A	p.Ala9Thr
C321.ΔA.2.104	413600	C	T	synonymous_variant	sbcC	protein_coding	c.1203G>A	p.Ala401Ala
C321.ΔA.2.104	428542	G	A	splice_region_variant&stop_retained_variant	secD	protein_coding	c.1847G>A	p.Ter616Ter
C321.ΔA.2.104	433606	G	A	splice_region_variant&stop_retained_variant	ribD	protein_coding	c.1103G>A	p.Ter368Ter
C321.ΔA.2.104	434165	G	A	splice_region_variant&stop_retained_variant	ribE	protein_coding	c.470G>A	p.Ter157Ter
C321.ΔA.2.104	438536	C	T	missense_variant	dxs	protein_coding	c.691G>A	p.Glu231Lys
C321.ΔA.2.104	443297	T	C	missense_variant	panE	protein_coding	c.268A>G	p.Thr90Ala
C321.ΔA.2.104	473260	C	T	synonymous_variant	amtB	protein_coding	c.1246C>T	p.Leu416Leu
C321.ΔA.2.104	474984	C	T	missense_variant	ybaY	protein_coding	c.557C>T	p.Ala186Val
C321.ΔA.2.104	486909	T	C	missense_variant	mscK	protein_coding	c.1325T>C	p.Ile442Thr
C321.ΔA.2.104	493071	G	A	splice_region_variant&stop_retain	dnaX	protein_coding	c.1931G>A	p.Ter644Ter

				ed_variant				
C321.ΔA.2.104	495240	C	T	stop_gained	htpG	protein_coding	c.1072C>T	p.Gln358*
C321.ΔA.2.104	499472	A	G	missense_variant	gsk	protein_coding	c.299A>G	p.His100Arg
C321.ΔA.2.104	500836	C	T	stop_gained	ybaL	protein_coding	c.1452G>A	p.Trp484*
C321.ΔA.2.104	503766	G	A	intergenic_region	fsr-ushA		n.503766G>A	
C321.ΔA.2.104	550961	A	G	missense_variant	purK	protein_coding	c.683T>C	p.Phe228Ser
C321.ΔA.2.104	611931	A	G	synonymous_variant	fes	protein_coding	c.69A>G	p.Val23Val
C321.ΔA.2.104	623295	C	T	synonymous_variant	fepB	protein_coding	c.264G>A	p.Gln88Gln
C321.ΔA.2.104	668980	C	T	splice_region_variant&stop_retained_variant	nadD	protein_coding	c.641G>A	p.Ter214Ter
C321.ΔA.2.104	669623	C	T	splice_region_variant&stop_retained_variant	holA	protein_coding	c.1031G>A	p.Ter344Ter
C321.ΔA.2.104	669665	A	G	missense_variant	holA	protein_coding	c.989T>C	p.Leu330Pro
C321.ΔA.2.104	670654	C	T	splice_region_variant&stop_retained_variant	lptE	protein_coding	c.581G>A	p.Ter194Ter
C321.ΔA.2.104	679212	A	G	missense_variant	ybeU	protein_coding	c.658A>G	p.Thr220Ala
C321.ΔA.2.104	696926	C	A	synonymous_variant	asnB	protein_coding	c.1299G>T	p.Leu433Leu
C321.ΔA.2.104	709983	C	T	splice_region_variant&stop_retain	fldA	protein_coding	c.530G>A	p.Ter177Ter

				ed_variant				
C321.ΔA.2.104	722801	A	G	missense_variant	kdpD	protein_coding	c.662T>C	p.Leu221Pro
C321.ΔA.2.104	723403	C	T	synonymous_variant	kdpD	protein_coding	c.60G>A	p.Pro20Pro
C321.ΔA.2.104	740043	T	C	intergenic_region	phr-dtpD		n.740043T>C	
C321.ΔA.2.104	744222	G	A	missense_variant	ybgL	protein_coding	c.10G>A	p.Asp4Asn
C321.ΔA.2.104	744222	G	A	splice_region_variant&stop_retained_variant	ybgK	protein_coding	c.932G>A	p.Ter311Ter
C321.ΔA.2.104	754954	G	A	splice_region_variant&stop_retained_variant	sdhD	protein_coding	c.347G>A	p.Ter116Ter
C321.ΔA.2.104	757053	A	G	missense_variant	sdhB	protein_coding	c.317A>G	p.Asp106Gly
C321.ΔA.2.104	771998	G	A	missense_variant	cydA	protein_coding	c.1493G>A	p.Arg498His
C321.ΔA.2.104	774870	G	A	missense_variant	tolQ	protein_coding	c.670G>A	p.Val224Ile
C321.ΔA.2.104	775990	G	A	missense_variant	tolA	protein_coding	c.601G>A	p.Glu201Lys
C321.ΔA.2.104	778079	G	A	splice_region_variant&stop_retained_variant	tolB	protein_coding	c.1292G>A	p.Ter431Ter
C321.ΔA.2.104	808519	C	T	missense_variant	N840_4502	protein_coding	c.181G>A	p.Asp61Asn
C321.ΔA.2.104	808519	C	T	synonymous_variant	N840_4503	protein_coding	c.9G>A	p.Glu3Glu
C321.ΔA.2.104	816521	C	A	missense_variant	bla	protein_coding	c.109C>A	p.Gln37Lys

C321.ΔA.2.104	816550	A	G	synonymous_variant	bla	protein_coding	c.138A>G	p.Glu46Glu
C321.ΔA.2.104	816640	C	T	synonymous_variant	bla	protein_coding	c.228C>T	p.Gly76Gly
C321.ΔA.2.104	816656	A	G	missense_variant	bla	protein_coding	c.244A>G	p.Ile82Val
C321.ΔA.2.104	816886	T	C	synonymous_variant	bla	protein_coding	c.474T>C	p.Thr158Thr
C321.ΔA.2.104	816957	T	C	missense_variant	bla	protein_coding	c.545T>C	p.Val182Ala
C321.ΔA.2.104	817129	G	A	synonymous_variant	bla	protein_coding	c.717G>A	p.Gly239Gly
C321.ΔA.2.104	819327	C	T	stop_gained	bioC	protein_coding	c.700C>T	p.Gln234*
C321.ΔA.2.104	823721	A	G	synonymous_variant	ybhK	protein_coding	c.33T>C	p.Arg11Arg
C321.ΔA.2.104	848553	C	T	missense_variant	fiu	protein_coding	c.85G>A	p.Ala29Thr
C321.ΔA.2.104	916196	G	A	missense_variant	ltaE	protein_coding	c.205C>T	p.His69Tyr
C321.ΔA.2.104	933332	C	T	splice_region_variant&stop_retained_variant	infA	protein_coding	c.218G>A	p.Ter73Ter
C321.ΔA.2.104	940119	C	T	synonymous_variant	lrp	protein_coding	c.420C>T	p.Gly140Gly
C321.ΔA.2.104	978135	C	T	missense_variant	kdsB	protein_coding	c.179C>T	p.Ala60Val
C321.ΔA.2.104	982745	G	A	missense_variant	mukE	protein_coding	c.19G>A	p.Glu7Lys
C321.ΔA.2.104	982745	G	A	splice_region_variant&stop_retained_variant	mukF	protein_coding	c.1322G>A	p.Ter441Ter

C321.ΔA.2.104	983284	C	T	synonymous_variant	mukE	protein_coding	c.558C>T	p.Arg186Arg
C321.ΔA.2.104	1E+06	C	T	splice_region_variant&stop_retained_variant	fabA	protein_coding	c.518G>A	p.Ter173Ter
C321.ΔA.2.104	1E+06	G	A	synonymous_variant	ycbZ	protein_coding	c.915C>T	p.Gly305Gly
C321.ΔA.2.104	1E+06	C	T	missense_variant	agp	protein_coding	c.170C>T	p.Thr57Met
C321.ΔA.2.104	1E+06	C	T	missense_variant	rutA	protein_coding	c.1138G>A	p.Glu380Lys
C321.ΔA.2.104	1E+06	G	A	splice_region_variant&stop_retained_variant	yceQ	protein_coding	c.320G>A	p.Ter107Ter
C321.ΔA.2.104	1E+06	G	A	splice_region_variant&stop_retained_variant	fabG	protein_coding	c.734G>A	p.Ter245Ter
C321.ΔA.2.104	1E+06	C	T	intergenic_region	icd-ymfD		n.1203578C>T	
C321.ΔA.2.104	1E+06	C	A	missense_variant	ymfE	protein_coding	c.443G>T	p.Trp148Leu
C321.ΔA.2.104	1E+06	G	A	synonymous_variant	intE	protein_coding	c.543C>T	p.Ala181Ala
C321.ΔA.2.104	1E+06	A	G	synonymous_variant	dhaR	protein_coding	c.789A>G	p.Lys263Lys
C321.ΔA.2.104	1E+06	C	T	missense_variant	hemA	protein_coding	c.587C>T	p.Pro196Leu
C321.ΔA.2.104	1E+06	C	T	missense_variant	chaA	protein_coding	c.233G>A	p.Ser78Asn
C321.ΔA.2.104	1E+06	A	G	missense_variant	puuA	protein_coding	c.457T>C	p.Phe153Leu

C321.ΔA.2.104	1E+06	C	T	synonymous_variant	insH1	protein_coding	c.453C>T	p.Gly151Gly
C321.ΔA.2.104	1E+06	A	G	missense_variant	uspE	protein_coding	c.152T>C	p.Leu51Pro
C321.ΔA.2.104	1E+06	T	C	missense_variant	trkG	protein_coding	c.565T>C	p.Trp189Arg
C321.ΔA.2.104	1E+06	C	T	synonymous_variant	ydbK	protein_coding	c.1170G>A	p.Ala390Ala
C321.ΔA.2.104	1E+06	T	C	synonymous_variant	tynA	protein_coding	c.702A>G	p.Glu234Glu
C321.ΔA.2.104	1E+06	A	G	synonymous_variant	ydbD	protein_coding	c.1239A>G	p.Gln413Gln
C321.ΔA.2.104	1E+06	G	T	missense_variant	hrpA	protein_coding	c.3136G>T	p.Asp1046Tyr
C321.ΔA.2.104	2E+06	A	G	synonymous_variant	ydcU	protein_coding	c.756A>G	p.Thr252Thr
C321.ΔA.2.104	2E+06	C	T	intergenic_region	curA-mcbR	n.1524889C>T		
C321.ΔA.2.104	2E+06	T	C	missense_variant	narZ	protein_coding	c.1838A>G	p.Asn613Ser
C321.ΔA.2.104	2E+06	G	A	synonymous_variant	fdnG	protein_coding	c.2418G>A	p.Thr806Thr
C321.ΔA.2.104	2E+06	A	C	synonymous_variant	ydeT	protein_coding	c.1191T>G	p.Gly397Gly
C321.ΔA.2.104	2E+06	T	C	missense_variant	ydfJ	protein_coding	c.5248A>G	p.Ile1750Val
C321.ΔA.2.104	2E+06	T	C	missense_variant	ydfX	protein_coding	c.46A>G	p.Ile16Val
C321.ΔA.2.104	2E+06	A	G	stop_lost	ydfJ	protein_coding	c.4084T>C	p.Ter1362Glnext*?
C321.ΔA.2.104	2E+06	A	G	missense_variant	ydfB	protein_coding	c.91A>G	p.Lys31Glu
C321.ΔA.2.104	2E+06	G	A	splice_region_variant&stop_retain	ydfB	protein_coding	c.128G>A	p.Ter43Ter

				ed_variant				
C321.ΔA.2.104	2E+06	G	A	synonymous_variant	ydfJ	protein_coding	c.4047C>T	p.Ile1349Ile
C321.ΔA.2.104	2E+06	G	A	synonymous_variant	ydfJ	protein_coding	c.1551C>T	p.Phe517Phe
C321.ΔA.2.104	2E+06	G	A	synonymous_variant	insD1	protein_coding	c.444G>A	p.Pro148Pro
C321.ΔA.2.104	2E+06	C	T	splice_region_variant&stop_retained_variant	ribC	protein_coding	c.641G>A	p.Ter214Ter
C321.ΔA.2.104	2E+06	T	C	synonymous_variant	ppsR	protein_coding	c.663T>C	p.Arg221Arg
C321.ΔA.2.104	2E+06	G	A	missense_variant	ydiU	protein_coding	c.982C>T	p.Arg328Trp
C321.ΔA.2.104	2E+06	C	T	splice_region_variant&stop_retained_variant	pheT	protein_coding	c.2387G>A	p.Ter796Ter
C321.ΔA.2.104	2E+06	T	C	missense_variant	pheS	protein_coding	c.124A>G	p.Met42Val
C321.ΔA.2.104	2E+06	T	C	missense_variant	astC	protein_coding	c.45A>G	p.Ile15Met
C321.ΔA.2.104	2E+06	A	G	missense_variant	ydjK	protein_coding	c.662T>C	p.Val221Ala
C321.ΔA.2.104	2E+06	T	C	synonymous_variant	fadD	protein_coding	c.1050A>G	p.Gly350Gly
C321.ΔA.2.104	2E+06	C	T	splice_region_variant&stop_retained_variant	tsaB	protein_coding	c.695G>A	p.Ter232Ter
C321.ΔA.2.104	2E+06	G	A	synonymous_variant	yobB	protein_coding	c.471G>A	p.Ser157Ser

C321.ΔA.2.104	2E+06	A	T	missense_variant	exoX	protein_coding	c.84A>T	p.Lys28Asn
C321.ΔA.2.104	2E+06	A	G	missense_variant	edd	protein_coding	c.1328T>C	p.Val443Ala
C321.ΔA.2.104	2E+06	C	T	splice_region_variant&stop_retained_variant	aspS	protein_coding	c.1772G>A	p.Ter591Ter
C321.ΔA.2.104	2E+06	T	C	missense_variant	araG	protein_coding	c.1130A>G	p.Asn377Ser
C321.ΔA.2.104	2E+06	C	T	splice_region_variant&stop_retained_variant	pgsA	protein_coding	c.548G>A	p.Ter183Ter
C321.ΔA.2.104	2E+06	G	A	synonymous_variant	fliK	protein_coding	c.15G>A	p.Ala5Ala
C321.ΔA.2.104	2E+06	A	G	missense_variant	yeeZ	protein_coding	c.340T>C	p.Ser114Pro
C321.ΔA.2.104	2E+06	C	T	synonymous_variant	wcaD	protein_coding	c.165G>A	p.Leu55Leu
C321.ΔA.2.104	2E+06	G	A	missense_variant	yegD	protein_coding	c.1186G>A	p.Ala396Thr
C321.ΔA.2.104	2E+06	A	G	synonymous_variant	yegL	protein_coding	c.42T>C	p.Ser14Ser
C321.ΔA.2.104	2E+06	G	A	stop_gained	bglX	protein_coding	c.1501C>T	p.Gln501*
C321.ΔA.2.104	2E+06	T	C	missense_variant	cdd	protein_coding	c.872T>C	p.Val291Ala
C321.ΔA.2.104	2E+06	C	T	splice_region_variant&stop_retained_variant	folE	protein_coding	c.668G>A	p.Ter223Ter
C321.ΔA.2.104	2E+06	C	T	synonymous_variant	fruA	protein_coding	c.1353G>A	p.Ala451Ala

C321.ΔA.2.104	2E+06	A	G	missense_variant	fruA	protein_coding	c.1349T>C	p.Val450Ala
C321.ΔA.2.104	2E+06	G	A	splice_region_variant&stop_retained_variant	yejF	protein_coding	c.1589G>A	p.Ter530Ter
C321.ΔA.2.104	2E+06	A	G	synonymous_variant	bcr	protein_coding	c.1005T>C	p.Phe335Phe
C321.ΔA.2.104	2E+06	G	A	splice_region_variant&stop_retained_variant	yejM	protein_coding	c.1760G>A	p.Ter587Ter
C321.ΔA.2.104	2E+06	C	T	synonymous_variant	yejO	protein_coding	c.3774G>A	p.Val1258Val
C321.ΔA.2.104	2E+06	C	G	missense_variant	yejO	protein_coding	c.3772G>C	p.Val1258Leu
C321.ΔA.2.104	2E+06	G	A	splice_region_variant&stop_retained_variant	nrdA	protein_coding	c.2285G>A	p.Ter762Ter
C321.ΔA.2.104	2E+06	G	A	splice_region_variant&stop_retained_variant	nrdB	protein_coding	c.1130G>A	p.Ter377Ter
C321.ΔA.2.104	2E+06	C	T	splice_region_variant&stop_retained_variant	accD	protein_coding	c.914G>A	p.Ter305Ter
C321.ΔA.2.104	2E+06	G	A	missense_variant	emrK	protein_coding	c.626C>T	p.Thr209Met
C321.ΔA.2.104	3E+06	C	T	splice_region_variant&stop_retained_variant	ligA	protein_coding	c.2015G>A	p.Ter672Ter

C321.ΔA.2.104	3E+06	T	C	missense_variant	ligA	protein_coding	c.1321A>G	p.Lys441Glu
C321.ΔA.2.104	3E+06	C	T	splice_region_variant&stop_retained_variant	zipA	protein_coding	c.986G>A	p.Ter329Ter
C321.ΔA.2.104	3E+06	T	C	missense_variant	rlmN	protein_coding	c.164A>G	p.Asp55Gly
C321.ΔA.2.104	3E+06	G	A	missense_variant	hcaE	protein_coding	c.491G>A	p.Arg164His
C321.ΔA.2.104	3E+06	C	T	splice_region_variant&stop_retained_variant	rpoE	protein_coding	c.575G>A	p.Ter192Ter
C321.ΔA.2.104	3E+06	G	A	synonymous_variant	intA	protein_coding	c.399G>A	p.Ser133Ser
C321.ΔA.2.104	3E+06	G	A	synonymous_variant	gabD	protein_coding	c.1347G>A	p.Val449Val
C321.ΔA.2.104	3E+06	T	C	synonymous_variant	proV	protein_coding	c.198T>C	p.Gly66Gly
C321.ΔA.2.104	3E+06	G	A	synonymous_variant	yqaA	protein_coding	c.81C>T	p.Val27Val
C321.ΔA.2.104	3E+06	G	C	synonymous_variant	alaS	protein_coding	c.588C>G	p.Gly196Gly
C321.ΔA.2.104	3E+06	T	C	missense_variant	norW	protein_coding	c.391T>C	p.Tyr131His
C321.ΔA.2.104	3E+06	T	C	missense_variant	ascG	protein_coding	c.460A>G	p.Asn154Asp
C321.ΔA.2.104	3E+06	A	G	missense_variant	hycC	protein_coding	c.902T>C	p.Leu301Pro
C321.ΔA.2.104	3E+06	G	A	missense_variant	hypE	protein_coding	c.241G>A	p.Ala81Thr
C321.ΔA.2.104	3E+06	C	T	synonymous_variant	cysI	protein_coding	c.1665G>A	p.Ala555Ala

C321.ΔA.2.104	3E+06	A	G	missense_variant	ygdG	protein_coding	c.532A>G	p.Lys178Glu
C321.ΔA.2.104	3E+06	T	C	missense_variant	amiC	protein_coding	c.14A>G	p.Asn5Ser
C321.ΔA.2.104	3E+06	T	C	missense_variant	recB	protein_coding	c.1217A>G	p.Glu406Gly
C321.ΔA.2.104	3E+06	C	T	splice_region_variant&stop_retained_variant	lgt	protein_coding	c.875G>A	p.Ter292Ter
C321.ΔA.2.104	3E+06	G	A	missense_variant	uacT	protein_coding	c.962G>A	p.Gly321Asp
C321.ΔA.2.104	3E+06	G	A	synonymous_variant	xerD	protein_coding	c.483C>T	p.Ser161Ser
C321.ΔA.2.104	3E+06	T	C	missense_variant	metK	protein_coding	c.34T>C	p.Ser12Pro
C321.ΔA.2.104	3E+06	T	G	missense_variant	mutY	protein_coding	c.780T>G	p.Cys260Trp
C321.ΔA.2.104	3E+06	T	C	missense_variant	glcB	protein_coding	c.1009A>G	p.Ile337Val
C321.ΔA.2.104	3E+06	A	G	synonymous_variant	nudF	protein_coding	c.465T>C	p.Gly155Gly
C321.ΔA.2.104	3E+06	C	T	splice_region_variant&stop_retained_variant	ribB	protein_coding	c.653G>A	p.Ter218Ter
C321.ΔA.2.104	3E+06	G	A	splice_region_variant&stop_retained_variant	cca	protein_coding	c.1238G>A	p.Ter413Ter
C321.ΔA.2.104	3E+06	G	A	missense_variant	fadH	protein_coding	c.203G>A	p.Gly68Asp
C321.ΔA.2.104	3E+06	C	T	synonymous_variant	fadH	protein_coding	c.558C>T	p.Gly186Gly

C321.ΔA.2.104	3E+06	T	C	missense_variant	tdcE	protein_coding	c.1261A>G	p.Ile421Val
C321.ΔA.2.104	3E+06	T	C	missense_variant	mtr	protein_coding	c.628A>G	p.Asn210Asp
C321.ΔA.2.104	3E+06	T	C	missense_variant	folP	protein_coding	c.652A>G	p.Met218Val
C321.ΔA.2.104	3E+06	G	A	intergenic_region	mlaF-yrbG		n.3340764G>A	
C321.ΔA.2.104	3E+06	A	G	missense_variant	degS	protein_coding	c.545A>G	p.Asn182Ser
C321.ΔA.2.104	3E+06	T	C	missense_variant	tldD	protein_coding	c.946A>G	p.Thr316Ala
C321.ΔA.2.104	3E+06	C	T	splice_region_variant&stop_retained_variant	rplV	protein_coding	c.332G>A	p.Ter111Ter
C321.ΔA.2.104	3E+06	C	T	splice_region_variant&stop_retained_variant	rpsG	protein_coding	c.539G>A	p.Ter180Ter
C321.ΔA.2.104	3E+06	G	A	synonymous_variant	cysG	protein_coding	c.1275G>A	p.Gln425Gln
C321.ΔA.2.104	4E+06	T	C	missense_variant	nikC	protein_coding	c.748T>C	p.Phe250Leu
C321.ΔA.2.104	4E+06	C	G	missense_variant	yhhI	protein_coding	c.675C>G	p.Ile225Met
C321.ΔA.2.104	4E+06	T	C	synonymous_variant	pitA	protein_coding	c.567T>C	p.Arg189Arg
C321.ΔA.2.104	4E+06	C	T	missense_variant	yhjD	protein_coding	c.365C>T	p.Thr122Met
C321.ΔA.2.104	4E+06	T	C	missense_variant	eptB	protein_coding	c.1115A>G	p.Tyr372Cys
C321.ΔA.2.104	4E+06	C	T	synonymous_variant	xylG	protein_coding	c.1275C>T	p.Arg425Arg

C321.ΔA.2.104	4E+06	C	T	missense_variant	yiaK	protein_coding	c.437C>T	p.Pro146Leu
C321.ΔA.2.104	4E+06	G	A	synonymous_variant	yiaS	protein_coding	c.93G>A	p.Ala31Ala
C321.ΔA.2.104	4E+06	C	A	missense_variant	hldD	protein_coding	c.612C>A	p.Ser204Arg
C321.ΔA.2.104	4E+06	C	T	splice_region_variant&stop_retained_variant	waaU	protein_coding	c.1073G>A	p.Ter358Ter
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	waaA	protein_coding	c.1277G>A	p.Ter426Ter
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	gmk	protein_coding	c.623G>A	p.Ter208Ter
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	rnpA	protein_coding	c.359G>A	p.Ter120Ter
C321.ΔA.2.104	4E+06	G	A	synonymous_variant	yidD	protein_coding	c.36G>A	p.Leu12Leu
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	yidC	protein_coding	c.1646G>A	p.Ter549Ter
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	tnaC	protein_coding	c.74G>A	p.Ter25Ter
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retain	yieE	protein_coding	c.749G>A	p.Ter250Ter

				ed_variant				
C321.ΔA.2.104	4E+06	C	T	splice_region_variant&stop_retained_variant	bglG	protein_coding	c.836G>A	p.Ter279Ter
C321.ΔA.2.104	4E+06	C	T	splice_region_variant&stop_retained_variant	pstB	protein_coding	c.773G>A	p.Ter258Ter
C321.ΔA.2.104	4E+06	C	T	splice_region_variant&stop_retained_variant	pstA	protein_coding	c.890G>A	p.Ter297Ter
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	rbsD	protein_coding	c.419G>A	p.Ter140Ter
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	rbsK	protein_coding	c.929G>A	p.Ter310Ter
C321.ΔA.2.104	4E+06	C	T	splice_region_variant&stop_retained_variant	yieP	protein_coding	c.692G>A	p.Ter231Ter
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	ilvM	protein_coding	c.263G>A	p.Ter88Ter
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	ubiB	protein_coding	c.1640G>A	p.Ter547Ter
C321.ΔA.2.104	4E+06	G	A	splice_region_var	ubiD	protein_coding	c.1493G>A	p.Ter498Ter

				iant&stop_retained_variant				
C321.ΔA.2.104	4E+06	T	C	synonymous_variant	pepQ	protein_coding	c.618T>C	p.Gly206Gly
C321.ΔA.2.104	4E+06	T	C	missense_variant	glnG	protein_coding	c.32A>G	p.Asp11Gly
C321.ΔA.2.104	4E+06	C	T	splice_region_variant&stop_retained_variant	ftsN	protein_coding	c.959G>A	p.Ter320Ter
C321.ΔA.2.104	4E+06	T	C	missense_variant	ptsA	protein_coding	c.811A>G	p.Met271Val
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	murI	protein_coding	c.857G>A	p.Ter286Ter
C321.ΔA.2.104	4E+06	C	A	intergenic_region	rrlB-rrfB		n.4173476C>A	
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	secE	protein_coding	c.383G>A	p.Ter128Ter
C321.ΔA.2.104	4E+06	G	A	missense_variant	hemE	protein_coding	c.255G>A	p.Met85Ile
C321.ΔA.2.104	4E+06	A	G	missense_variant	yjaB	protein_coding	c.430T>C	p.Tyr144His
C321.ΔA.2.104	4E+06	G	A	synonymous_variant	metA	protein_coding	c.540G>A	p.Thr180Thr
C321.ΔA.2.104	4E+06	A	G	missense_variant	metH	protein_coding	c.2620A>G	p.Thr874Ala
C321.ΔA.2.104	4E+06	C	T	synonymous_variant	metH	protein_coding	c.2950C>T	p.Leu984Leu
C321.ΔA.2.104	4E+06	C	T	missense_variant	lamB	protein_coding	c.725C>T	p.Ala242Val

C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	ubiA	protein_coding	c.872G>A	p.Ter291Ter
C321.ΔA.2.104	4E+06	A	G	synonymous_variant	dinF	protein_coding	c.453A>G	p.Glu151Glu
C321.ΔA.2.104	4E+06	G	A	splice_region_variant&stop_retained_variant	ssb	protein_coding	c.536G>A	p.Ter179Ter
C321.ΔA.2.104	4E+06	C	T	splice_region_variant&stop_retained_variant	alsK	protein_coding	c.929G>A	p.Ter310Ter
C321.ΔA.2.104	4E+06	C	G	intergenic_region	adiY-adiA	n.4340053C>G		
C321.ΔA.2.104	4E+06	C	T	synonymous_variant	yjdF	protein_coding	c.369G>A	p.Gly123Gly
C321.ΔA.2.104	4E+06	A	G	synonymous_variant	cadA	protein_coding	c.1866T>C	p.Ala622Ala
C321.ΔA.2.104	4E+06	A	G	synonymous_variant	cadC	protein_coding	c.1488T>C	p.Thr496Thr
C321.ΔA.2.104	4E+06	G	A	stop_gained	ampC	protein_coding	c.67C>T	p.Gln23*
C321.ΔA.2.104	4E+06	C	T	missense_variant	yjfM	protein_coding	c.430C>T	p.Arg144Cys
C321.ΔA.2.104	4E+06	T	C	intergenic_region	yjfC-aidB	n.4416174T>C		
C321.ΔA.2.104	4E+06	A	G	missense_variant	aidB	protein_coding	c.1523A>G	p.Gln508Arg
C321.ΔA.2.104	4E+06	A	G	missense_variant	ytfE	protein_coding	c.530T>C	p.Val177Ala
C321.ΔA.2.104	5E+06	A	G	missense_variant	insO	protein_coding	c.610A>G	p.Thr204Ala
C321.ΔA.2.104	5E+06	A	G	missense_variant	insI1	protein_coding	c.812T>C	p.Val271Ala

C321.ΔA.2.104	5E+06	C	T	missense_variant	insA	protein_coding	c.106C>T	p.Arg36Cys
C321.ΔA.2.104	5E+06	T	C	missense_variant	fimB	protein_coding	c.266T>C	p.Leu89Ser
C321.ΔA.2.104	5E+06	C	T	synonymous_variant	opgB	protein_coding	c.315G>A	p.Pro105Pro
C321.ΔA.2.104	5E+06	A	G	intergenic_region	fhuF-yjjZ	n.4607664A>G		
C321.ΔA.2.104	5E+06	C	T	synonymous_variant	yjjI	protein_coding	c.744G>A	p.Thr248Thr
C321.ΔA.2.104	20575	TG	T	intergenic_region	insA-rpsT	n.20576delG		
C321.ΔA.2.104	34112	GT	G	intergenic_region	carB-caiF	n.34113delT		
C321.ΔA.2.104	45463	T	TATA	start_lost&disruptive_inframe_insertion	fixX	protein_coding	c.2_3insAAT	
C321.ΔA.2.104	45463	T	TATA	splice_region_variant&disruptive_inframe_insertion	fixC	protein_coding	c.1285_1286insAAT	p.Ter429dup
C321.ΔA.2.104	52429	C	CATT	start_lost&disruptive_inframe_insertion	rsmA	protein_coding	c.-1_2dupAAT	
C321.ΔA.2.104	52429	C	CATT	splice_region_variant&conservative_inframe_insertion	pdxA	protein_coding	c.986_988dupAAT	p.Ter330dup
C321.ΔA.2.104	94480	C	CATA	start_lost&disruptive_inframe_insertion	murF	protein_coding	c.2_3insAAT	
C321.ΔA.2.104	94480	C	CATA	splice_region_variant&disruptive_inframe_insertion	murE	protein_coding	c.1486_1487insAAT	p.Ter496dup

C321.ΔA.2.104	99474	G	GATA	start_lost&disruptive_inframe_insertion	murG	protein_coding	c.2_3insAAT	
C321.ΔA.2.104	99474	G	GATA	splice_region_variant&disruptive_inframe_insertion	ftsW	protein_coding	c.1243_1244insAAT	p.Ter415dup
C321.ΔA.2.104	103811	C	CAAT	splice_region_variant&disruptive_inframe_insertion	ftsQ	protein_coding	c.827_829dupAAT	p.Ter277dup
C321.ΔA.2.104	136897	A	AG	intergenic_region	yacC-cueO	n.136897_136898insG		
C321.ΔA.2.104	204321	C	CAAT	splice_region_variant&disruptive_inframe_insertion	lpxB	protein_coding	c.1145_1147dupAAT	p.Ter383dup
C321.ΔA.2.104	214662	C	CATA	start_lost&disruptive_inframe_insertion	arfB	protein_coding	c.2_3insAAT	
C321.ΔA.2.104	214662	C	CATA	splice_region_variant&disruptive_inframe_insertion	yaeQ	protein_coding	c.544_545insAAT	p.Ter182dup
C321.ΔA.2.104	232880	CG	C	frameshift_variant	mltD	protein_coding	c.905delC	p.Pro302fs
C321.ΔA.2.104	298918	C	CG	frameshift_variant	paoC	protein_coding	c.1068dupC	p.Gly357fs
C321.ΔA.2.104	379063	A	AG	intergenic_region	frmR-yaiO	n.379063_379064insG		
C321.ΔA.2.104	414800	C	CATT	start_lost&disruptive_inframe_insertion	sbcC	protein_coding	c.2_3insAAT	
C321.ΔA.2.104	414800	C	CATT	splice_region_variant&co	sbcD	protein_coding	c.1201_1202insAAT	p.Ter401dup

				nservative_inframe_insertion				
C321.ΔA.2.104	440454	C	CA	intergenic_region	xseB-thiI	n.440454_440455insA		
C321.ΔA.2.104	457782	A	AG	intergenic_region	clpX-lon	n.457782_457783insG		
C321.ΔA.2.104	531493	C	CA	intergenic_region	allS-allA	n.531493_531494insA		
C321.ΔA.2.104	560739	T	TG	frameshift_variant	sfmD	protein_coding	c.2000dupG	p.Gly668fs
C321.ΔA.2.104	650998	GA	G	intergenic_region	citC-dpiB	n.650999delA		
C321.ΔA.2.104	719619	C	CT	intergenic_region	speF-ybfK	n.719619_719620insT		
C321.ΔA.2.104	722567	GT	G	frameshift_variant	kdpD	protein_coding	c.895delA	p.Thr299fs
C321.ΔA.2.104	775893	GA	G	frameshift_variant	tolA	protein_coding	c.510delA	p.Ala171fs
C321.ΔA.2.104	970912	CT	C	intergenic_region	rpsA-ihfB	n.970913delT		
C321.ΔA.2.104	975469	C	CAAT	splice_region_variant&disruptive_inframe_insertion	msbA	protein_coding	c.1745_1747dupAAT	p.Ter583dup
C321.ΔA.2.104	1E+06	C	CATA	start_lost&disruptive_inframe_insertion	holB	protein_coding	c.2_3insAAT	
C321.ΔA.2.104	1E+06	C	CATA	splice_region_variant&disruptive_inframe_insertion	tmk	protein_coding	c.640_641insAAT	p.Ter214dup
C321.ΔA.2.104	1E+06	TA	T	intergenic_region	ymfJ-cohE	n.1209324delA		
C321.ΔA.2.104	1E+06	A	AG	frameshift_variant	ycgB	protein_coding	c.969dupC	p.Tyr324fs
C321.ΔA.2.104	1E+06	C	CAAT	splice_region_variant&disruptive_i	prmC	protein_coding	c.830_832dupAAT	p.Ter278dup

				nframe_insertion				
C321.ΔA.2.104	1E+06	GCA	G	frameshift_variant	stfR	protein_coding	c.3068_3069delAC	p.His1023fs
C321.ΔA.2.104	1E+06	GT	G	frameshift_variant	cybB	protein_coding	c.449delT	p.Leu150fs
C321.ΔA.2.104	2E+06	A	AT	frameshift_variant	ynfF	protein_coding	c.118dupT	p.Ser40fs
C321.ΔA.2.104	2E+06	A	AT	frameshift_variant&stop_gained	exoX	protein_coding	c.81_82insT	p.Lys28fs
C321.ΔA.2.104	2E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	yefM	protein_coding	c.250_251insAAT	p.Ter84dup
C321.ΔA.2.104	2E+06	C	CATT	disruptive_inframe_insertion	yoeB	protein_coding	c.2_3insAAT	p.Val1_Lys2insMet
C321.ΔA.2.104	2E+06	GC	G	frameshift_variant	cpsG	protein_coding	c.340delG	p.Ala114fs
C321.ΔA.2.104	2E+06	G	GAAT	splice_region_variant&disruptive_inframe_insertion	yohJ	protein_coding	c.395_397dupAAT	p.Ter133dup
C321.ΔA.2.104	2E+06	AC	A	frameshift_variant	yohK	protein_coding	c.35delC	p.Thr12fs
C321.ΔA.2.104	2E+06	C	CCTGAGCTTTCATCCTATGAAATTAATTGCTGTTAAAAGCATTGGGTA	frameshift_variant&stop_gained	yejO	protein_coding	c.3781_3782insTACCCAATGCTTTTAACAGCAATTAATTTCATAGGATGAAAGCTCAG	p.Gly1261fs
C321.ΔA.2.104	2E+06	A	AAAATAC	conservative_inframe_insertion	yejO	protein_coding	c.3777_3778insGTATTT	p.Trp1259_Ser1260insValPhe

C321.ΔA.2.104	2E+06	CA	C	frameshift_variant	fadJ	protein_coding	c.870delT	p.Phe290fs
C321.ΔA.2.104	3E+06	G	GC	frameshift_variant	fryA	protein_coding	c.628dupG	p.Ala210fs
C321.ΔA.2.104	3E+06	GC	G	frameshift_variant	focB	protein_coding	c.685delC	p.Leu229fs
C321.ΔA.2.104	3E+06	CGCACTATG	C	intergenic_region	yfjL-yfjM	n.2769432_2769439delGCACTATG		
C321.ΔA.2.104	3E+06	G	GC	intergenic_region	yfjM-rnlA	n.2769807_2769808insC		
C321.ΔA.2.104	3E+06	C	CATT	start_lost&disruptive_inframe_insertion	truD	protein_coding	c.-1_2dupAAT	
C321.ΔA.2.104	3E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	ispF	protein_coding	c.476_478dupAAT	p.Ter160dup
C321.ΔA.2.104	3E+06	A	AG	frameshift_variant	ygeR	protein_coding	c.218dupC	p.Tyr74fs
C321.ΔA.2.104	3E+06	G	GCC	frameshift_variant	yggN	protein_coding	c.590_591dupGG	p.Leu198fs
C321.ΔA.2.104	3E+06	CCA	C	frameshift_variant	dnaG	protein_coding	c.1726_1727delCA	p.Gln576fs
C321.ΔA.2.104	3E+06	G	GCC	frameshift_variant	dnaG	protein_coding	c.1728_1729insCC	p.Glu577fs
C321.ΔA.2.104	3E+06	C	CATT	start_lost&disruptive_inframe_insertion	rplW	protein_coding	c.2_3insAAT	
C321.ΔA.2.104	3E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	rplD	protein_coding	c.604_605insAAT	p.Ter202dup

C321.ΔA.2.104	3E+06	C	CT	frameshift_variant	gspA	protein_coding	c.1336dupA	p.Ser446fs
C321.ΔA.2.104	4E+06	G	GC	frameshift_variant	yhhS	protein_coding	c.1018dupG	p.Ala340fs
C321.ΔA.2.104	4E+06	T	TCC	intergenic_region	bcsG-ldrD	n.3700603_3700604insCC		
C321.ΔA.2.104	4E+06	CG	C	frameshift_variant	waaA	protein_coding	c.1272delG	p.His425fs
C321.ΔA.2.104	4E+06	C	CT	frameshift_variant&start_lost	dgoD	protein_coding	c.2_3insA	p.Met1fs
C321.ΔA.2.104	4E+06	C	CT	frameshift_variant&splice_region_variant	dgoA	protein_coding	c.616_617insA	
C321.ΔA.2.104	4E+06	ACG	A	frameshift_variant	dgoA	protein_coding	c.596_597delCG	p.Ala199fs
C321.ΔA.2.104	4E+06	C	CATT	start_lost&disruptive_inframe_insertion	dgoK	protein_coding	c.2_3insAAT	
C321.ΔA.2.104	4E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	dgoR	protein_coding	c.688_689insAAT	p.Ter230dup
C321.ΔA.2.104	4E+06	GA	G	frameshift_variant	dgoR	protein_coding	c.661delT	p.Ser221fs
C321.ΔA.2.104	4E+06	AGG	A	intergenic_region	kup-rbsD	n.3935127_3935128delGG		
C321.ΔA.2.104	4E+06	C	CGG	frameshift_variant	wzzE	protein_coding	c.1002_1003dupGG	p.Ala335fs
C321.ΔA.2.104	4E+06	C	CATT	start_lost&disruptive_inframe_insertion	hemD	protein_coding	c.2_3insAAT	
C321.ΔA.2.104	4E+06	C	CATT	splice_region_var	hemC	protein_coding	c.940_941insAAT	p.Ter314dup

				iant&conservative_inframe_insertion				
C321.ΔA.2.104	4E+06	A	AAAT	splice_region_variant&disruptive_inframe_insertion	yigP	protein_coding	c.602_604dupAAT	p.Ter202dup
C321.ΔA.2.104	4E+06	A	AAAT	splice_region_variant&non_coding_transcript_exon_variant	esrE	lincRNA	n.248_250dupAAT	
C321.ΔA.2.104	4E+06	ATC	A	intergenic_region	ubiB-tatA	n.4023794_4023795delTC		
C321.ΔA.2.104	4E+06	C	CATA	start_lost&disruptive_inframe_insertion	birA	protein_coding	c.2_3insAAT	
C321.ΔA.2.104	4E+06	C	CATA	splice_region_variant&disruptive_inframe_insertion	murB	protein_coding	c.1027_1028insAAT	p.Ter343dup
C321.ΔA.2.104	4E+06	GC	G	frameshift_variant	phnP	protein_coding	c.59delG	p.Gly20fs
C321.ΔA.2.104	4E+06	GA	G	frameshift_variant	mutL	protein_coding	c.98delA	p.Asn33fs
C321.ΔA.2.104	4E+06	AC	A	frameshift_variant	msrA	protein_coding	c.275delG	p.Gly92fs
C321.ΔA.2.104	5E+06	AC	A	frameshift_variant	yjhG	protein_coding	c.1492delG	p.Val498fs
C321.ΔA.2.104	5E+06	C	CA	frameshift_variant	fimB	protein_coding	c.28dupA	p.Arg10fs
C321.ΔA.2.104	5E+06	C	CA	intergenic_region	yjjM-yjjN	n.4597823_4597824insA		

C321.ΔA.2.104	5E+06	A	AG	frameshift_variant	leuV	protein_coding	c.79dupC	p.Leu27fs
C321.ΔA.2.104	5E+06	A	AG	frameshift_variant	leuP	protein_coding	c.79dupC	p.Leu27fs
C321.ΔA.2.104	5E+06	A	AG	frameshift_variant	leuQ	protein_coding	c.79dupC	p.Leu27fs
C321.ΔA.2.64	5323	G	A	synonymous_variant	yaaX	protein_coding	c.90G>A	p.Pro30Pro
C321.ΔA.2.64	25207	G	A	splice_region_variant&stop_retained_variant	ileS	protein_coding	c.2816G>A	p.Ter939Ter
C321.ΔA.2.64	25295	A	G	missense_variant	lspA	protein_coding	c.88A>G	p.Ile30Val
C321.ΔA.2.64	30896	G	T	missense_variant	carB	protein_coding	c.79G>T	p.Asp27Tyr
C321.ΔA.2.64	44129	G	A	splice_region_variant&stop_retained_variant	fixB	protein_coding	c.941G>A	p.Ter314Ter
C321.ΔA.2.64	45750	G	A	splice_region_variant&stop_retained_variant	fixX	protein_coding	c.287G>A	p.Ter96Ter
C321.ΔA.2.64	54757	C	T	splice_region_variant&stop_retained_variant	lptD	protein_coding	c.2354G>A	p.Ter785Ter
C321.ΔA.2.64	58179	G	A	splice_region_variant&stop_retained_variant	djlA	protein_coding	c.815G>A	p.Ter272Ter
C321.ΔA.2.64	63431	C	T	splice_region_variant&stop_retain	polB	protein_coding	c.2351G>A	p.Ter784Ter

				ed_variant				
C321.ΔA.2.64	64263	C	T	missense_variant	polB	protein_coding	c.1519G>A	p.Ala507Thr
C321.ΔA.2.64	72743	C	T	missense_variant	thiQ	protein_coding	c.16G>A	p.Asp6Asn
C321.ΔA.2.64	72743	C	T	splice_region_variant&stop_retained_variant	thiP	protein_coding	c.1610G>A	p.Ter537Ter
C321.ΔA.2.64	85664	C	T	synonymous_variant	ilvI	protein_coding	c.204C>T	p.Val68Val
C321.ΔA.2.64	88188	T	C	synonymous_variant	cra	protein_coding	c.330T>C	p.Ile110Ile
C321.ΔA.2.64	98233	G	A	splice_region_variant&stop_retained_variant	murD	protein_coding	c.1316G>A	p.Ter439Ter
C321.ΔA.2.64	102070	G	A	missense_variant	ddlB	protein_coding	c.7G>A	p.Asp3Asn
C321.ΔA.2.64	102070	G	A	splice_region_variant&stop_retained_variant	murC	protein_coding	c.1475G>A	p.Ter492Ter
C321.ΔA.2.64	102995	G	A	missense_variant	ftsQ	protein_coding	c.10G>A	p.Ala4Thr
C321.ΔA.2.64	116886	T	C	synonymous_variant	hofB	protein_coding	c.45A>G	p.Gly15Gly
C321.ΔA.2.64	133742	T	A	missense_variant	acnB	protein_coding	c.2297T>A	p.Ile766Asn
C321.ΔA.2.64	156361	A	G	synonymous_variant	yadN	protein_coding	c.354T>C	p.Ser118Ser
C321.ΔA.2.64	160814	C	T	synonymous_variant	sfsA	protein_coding	c.504G>A	p.Ala168Ala

C321.ΔA.2.64	171836	A	G	missense_variant	fhuB	protein_coding	c.544A>G	p.Thr182Ala
C321.ΔA.2.64	173434	C	T	splice_region_variant&stop_retained_variant	hemL	protein_coding	c.1280G>A	p.Ter427Ter
C321.ΔA.2.64	185809	C	T	splice_region_variant&stop_retained_variant	glnD	protein_coding	c.2672G>A	p.Ter891Ter
C321.ΔA.2.64	193258	G	A	splice_region_variant&stop_retained_variant	frr	protein_coding	c.557G>A	p.Ter186Ter
C321.ΔA.2.64	197727	G	A	splice_region_variant&stop_retained_variant	rseP	protein_coding	c.1352G>A	p.Ter451Ter
C321.ΔA.2.64	202385	G	A	splice_region_variant&stop_retained_variant	fabZ	protein_coding	c.455G>A	p.Ter152Ter
C321.ΔA.2.64	213756	C	T	missense_variant	rof	protein_coding	c.7G>A	p.Asp3Asn
C321.ΔA.2.64	213756	C	T	splice_region_variant&stop_retained_variant	yaeP	protein_coding	c.200G>A	p.Ter67Ter
C321.ΔA.2.64	216888	C	T	splice_region_variant&stop_retained_variant	proS	protein_coding	c.1718G>A	p.Ter573Ter
C321.ΔA.2.64	221945	A	G	synonymous_variant	metN	protein_coding	c.531T>C	p.Arg177Arg

C321.ΔA.2.64	223237	G	A	splice_region_variant&stop_retained_variant	gmhB	protein_coding	c.575G>A	p.Ter192Ter
C321.ΔA.2.64	223835	C	T	non_coding_transcript_exon_variant	rrsH	rRNA	n.235C>T	
C321.ΔA.2.64	236026	A	G	missense_variant	dnaQ	protein_coding	c.130A>G	p.Thr44Ala
C321.ΔA.2.64	253368	T	C	missense_variant	ykfJ	protein_coding	c.61T>C	p.Trp21Arg
C321.ΔA.2.64	257670	A	G	missense_variant	crl	protein_coding	c.13A>G	p.Ser5Gly
C321.ΔA.2.64	264888	T	C	missense_variant	ykfB	protein_coding	c.253A>G	p.Thr85Ala
C321.ΔA.2.64	295866	C	T	missense_variant	intF	protein_coding	c.283G>A	p.Asp95Asn
C321.ΔA.2.64	296893	A	G	missense_variant	paoD	protein_coding	c.886T>C	p.Ser296Pro
C321.ΔA.2.64	299130	A	G	missense_variant	paoC	protein_coding	c.857T>C	p.Leu286Pro
C321.ΔA.2.64	311939	G	A	intergenic_region	ykgM-ykgR		n.311939G>A	
C321.ΔA.2.64	336301	T	C	missense_variant	yahF	protein_coding	c.473T>C	p.Leu158Pro
C321.ΔA.2.64	362284	A	G	intragenic_variant	lacZ		n.362284A>G	
C321.ΔA.2.64	369824	A	G	missense_variant	mhpB	protein_coding	c.497A>G	p.Asn166Ser
C321.ΔA.2.64	402649	G	A	splice_region_variant&stop_retained_variant	psiF	protein_coding	c.320G>A	p.Ter107Ter
C321.ΔA.2.64	406218	G	A	splice_region_variant&stop_retain	yaiA	protein_coding	c.191G>A	p.Ter64Ter

				ed_variant				
C321.ΔA.2.64	411506	C	T	missense_variant	araJ	protein_coding	c.25G>A	p.Ala9Thr
C321.ΔA.2.64	413600	C	T	synonymous_variant	sbcC	protein_coding	c.1203G>A	p.Ala401Ala
C321.ΔA.2.64	428542	G	A	splice_region_variant&stop_retained_variant	secD	protein_coding	c.1847G>A	p.Ter616Ter
C321.ΔA.2.64	433606	G	A	splice_region_variant&stop_retained_variant	ribD	protein_coding	c.1103G>A	p.Ter368Ter
C321.ΔA.2.64	434165	G	A	splice_region_variant&stop_retained_variant	ribE	protein_coding	c.470G>A	p.Ter157Ter
C321.ΔA.2.64	438536	C	T	missense_variant	dxs	protein_coding	c.691G>A	p.Glu231Lys
C321.ΔA.2.64	493071	G	A	splice_region_variant&stop_retained_variant	dnaX	protein_coding	c.1931G>A	p.Ter644Ter
C321.ΔA.2.64	495240	C	T	stop_gained	htpG	protein_coding	c.1072C>T	p.Gln358*
C321.ΔA.2.64	499472	A	G	missense_variant	gsk	protein_coding	c.299A>G	p.His100Arg
C321.ΔA.2.64	500836	C	T	stop_gained	ybaL	protein_coding	c.1452G>A	p.Trp484*
C321.ΔA.2.64	503766	G	A	intergenic_region	fsr-ushA	n.503766G>A		
C321.ΔA.2.64	623295	C	T	synonymous_variant	fepB	protein_coding	c.264G>A	p.Gln88Gln
C321.ΔA.2.64	668980	C	T	splice_region_variant&stop_retain	nadD	protein_coding	c.641G>A	p.Ter214Ter

				ed_variant				
C321.ΔA.2.64	669623	C	T	splice_region_variant&stop_retained_variant	holA	protein_coding	c.1031G>A	p.Ter344Ter
C321.ΔA.2.64	669665	A	G	missense_variant	holA	protein_coding	c.989T>C	p.Leu330Pro
C321.ΔA.2.64	670654	C	T	splice_region_variant&stop_retained_variant	lptE	protein_coding	c.581G>A	p.Ter194Ter
C321.ΔA.2.64	679212	A	G	missense_variant	ybeU	protein_coding	c.658A>G	p.Thr220Ala
C321.ΔA.2.64	696926	C	A	synonymous_variant	asnB	protein_coding	c.1299G>T	p.Leu433Leu
C321.ΔA.2.64	709983	C	T	splice_region_variant&stop_retained_variant	fldA	protein_coding	c.530G>A	p.Ter177Ter
C321.ΔA.2.64	722801	A	G	missense_variant	kdpD	protein_coding	c.662T>C	p.Leu221Pro
C321.ΔA.2.64	723403	C	T	synonymous_variant	kdpD	protein_coding	c.60G>A	p.Pro20Pro
C321.ΔA.2.64	740043	T	C	intergenic_region	phr-dtpD		n.740043T>C	
C321.ΔA.2.64	761440	G	A	missense_variant	sucB	protein_coding	c.871G>A	p.Gly291Ser
C321.ΔA.2.64	774870	G	A	missense_variant	tolQ	protein_coding	c.670G>A	p.Val224Ile
C321.ΔA.2.64	781486	C	T	synonymous_variant	nadA	protein_coding	c.354C>T	p.Gly118Gly
C321.ΔA.2.64	791831	G	A	missense_variant	modE	protein_coding	c.713C>T	p.Ala238Val
C321.ΔA.2.64	796547	T	C	synonymous_variant	pgl	protein_coding	c.63T>C	p.Asn21Asn

C321.ΔA.2.64	808519	C	T	missense_variant	N840_4502	protein_coding	c.181G>A	p.Asp61Asn
C321.ΔA.2.64	808519	C	T	synonymous_variant	N840_4503	protein_coding	c.9G>A	p.Glu3Glu
C321.ΔA.2.64	816521	C	A	missense_variant	bla	protein_coding	c.109C>A	p.Gln37Lys
C321.ΔA.2.64	816550	A	G	synonymous_variant	bla	protein_coding	c.138A>G	p.Glu46Glu
C321.ΔA.2.64	816640	C	T	synonymous_variant	bla	protein_coding	c.228C>T	p.Gly76Gly
C321.ΔA.2.64	816656	A	G	missense_variant	bla	protein_coding	c.244A>G	p.Ile82Val
C321.ΔA.2.64	816886	T	C	synonymous_variant	bla	protein_coding	c.474T>C	p.Thr158Thr
C321.ΔA.2.64	816957	T	C	missense_variant	bla	protein_coding	c.545T>C	p.Val182Ala
C321.ΔA.2.64	870757	T	C	missense_variant	fsaA	protein_coding	c.10T>C	p.Tyr4His
C321.ΔA.2.64	872152	C	T	synonymous_variant	moeB	protein_coding	c.84G>A	p.Ala28Ala
C321.ΔA.2.64	885338	A	G	intergenic_region	rimO-bssR	n.885338A>G		
C321.ΔA.2.64	889941	A	G	missense_variant	ybjG	protein_coding	c.554T>C	p.Val185Ala
C321.ΔA.2.64	905415	C	T	synonymous_variant	ybjO	protein_coding	c.321C>T	p.Tyr107Tyr
C321.ΔA.2.64	916196	G	A	missense_variant	ltaE	protein_coding	c.205C>T	p.His69Tyr
C321.ΔA.2.64	940119	C	T	synonymous_variant	lrp	protein_coding	c.420C>T	p.Gly140Gly
C321.ΔA.2.64	978135	C	T	missense_variant	kdsB	protein_coding	c.179C>T	p.Ala60Val
C321.ΔA.2.64	983284	C	T	synonymous_variant	mukE	protein_coding	c.558C>T	p.Arg186Arg

C321.ΔA.2.64	989893	T	C	missense_variant	ycbB	protein_coding	c.1742T>C	p.Val581Ala
C321.ΔA.2.64	1E+06	T	C	intergenic_region	pyrD-zapC	n.1012907T>C		
C321.ΔA.2.64	1E+06	T	C	synonymous_variant	helD	protein_coding	c.1527T>C	p.Gly509Gly
C321.ΔA.2.64	1E+06	T	C	missense_variant	clsC	protein_coding	c.1079T>C	p.Ile360Thr
C321.ΔA.2.64	1E+06	G	A	missense_variant	fabG	protein_coding	c.439G>A	p.Gly147Ser
C321.ΔA.2.64	1E+06	C	T	synonymous_variant	ycfJ	protein_coding	c.459C>T	p.Gly153Gly
C321.ΔA.2.64	1E+06	C	T	missense_variant	icd	protein_coding	c.334C>T	p.Arg112Cys
C321.ΔA.2.64	1E+06	G	A	synonymous_variant	intE	protein_coding	c.543C>T	p.Ala181Ala
C321.ΔA.2.64	1E+06	A	G	intergenic_region	stfE-pinE	n.1216743A>G		
C321.ΔA.2.64	1E+06	A	G	missense_variant	ychF	protein_coding	c.608T>C	p.Met203Thr
C321.ΔA.2.64	1E+06	C	T	missense_variant	hemA	protein_coding	c.587C>T	p.Pro196Leu
C321.ΔA.2.64	1E+06	C	T	missense_variant	chaA	protein_coding	c.271G>A	p.Ala91Thr
C321.ΔA.2.64	1E+06	C	T	missense_variant	chaA	protein_coding	c.233G>A	p.Ser78Asn
C321.ΔA.2.64	1E+06	C	T	missense_variant	chaA	protein_coding	c.200G>A	p.Arg67His
C321.ΔA.2.64	1E+06	T	C	missense_variant	narK	protein_coding	c.407T>C	p.Ile136Thr
C321.ΔA.2.64	1E+06	A	G	missense_variant	narG	protein_coding	c.2980A>G	p.Lys994Glu
C321.ΔA.2.64	1E+06	A	G	synonymous_variant	oppB	protein_coding	c.18A>G	p.Leu6Leu
C321.ΔA.2.64	1E+06	C	T	missense_variant	trpD	protein_coding	c.538G>A	p.Ala180Thr

C321.ΔA.2.64	1E+06	A	G	missense_variant	rluB	protein_coding	c.169A>G	p.Ile57Val
C321.ΔA.2.64	1E+06	C	T	synonymous_variant	mppA	protein_coding	c.1386C>T	p.Asn462Asn
C321.ΔA.2.64	1E+06	A	G	missense_variant	uspE	protein_coding	c.152T>C	p.Leu51Pro
C321.ΔA.2.64	1E+06	A	G	synonymous_variant	ydbD	protein_coding	c.1239A>G	p.Gln413Gln
C321.ΔA.2.64	2E+06	C	T	intergenic_region	curA-mcbR	n.1524889C>T		
C321.ΔA.2.64	2E+06	T	C	missense_variant	yncD	protein_coding	c.1924A>G	p.Thr642Ala
C321.ΔA.2.64	2E+06	G	A	synonymous_variant	ydfJ	protein_coding	c.1551C>T	p.Phe517Phe
C321.ΔA.2.64	2E+06	G	A	synonymous_variant	insD1	protein_coding	c.444G>A	p.Pro148Pro
C321.ΔA.2.64	2E+06	A	G	synonymous_variant	clcB	protein_coding	c.279A>G	p.Gly93Gly
C321.ΔA.2.64	2E+06	T	C	missense_variant	pntA	protein_coding	c.1028A>G	p.Asp343Gly
C321.ΔA.2.64	2E+06	G	A	missense_variant	ydgH	protein_coding	c.467G>A	p.Arg156His
C321.ΔA.2.64	2E+06	C	T	synonymous_variant	lhr	protein_coding	c.60C>T	p.Arg20Arg
C321.ΔA.2.64	2E+06	T	C	missense_variant	ydhS	protein_coding	c.1517T>C	p.Leu506Pro
C321.ΔA.2.64	2E+06	A	G	missense_variant	ydhV	protein_coding	c.1630T>C	p.Trp544Arg
C321.ΔA.2.64	2E+06	T	C	synonymous_variant	fadK	protein_coding	c.1341T>C	p.Arg447Arg
C321.ΔA.2.64	2E+06	T	C	synonymous_variant	ppsR	protein_coding	c.663T>C	p.Arg221Arg
C321.ΔA.2.64	2E+06	A	G	synonymous_variant	ydiY	protein_coding	c.375T>C	p.Tyr125Tyr

C321.ΔA.2.64	2E+06	G	A	missense_variant	katE	protein_coding	c.118G>A	p.Ala40Thr
C321.ΔA.2.64	2E+06	C	T	missense_variant	astB	protein_coding	c.898G>A	p.Val300Met
C321.ΔA.2.64	2E+06	C	T	missense_variant	yebZ	protein_coding	c.533G>A	p.Arg178His
C321.ΔA.2.64	2E+06	G	A	synonymous_variant	yobB	protein_coding	c.471G>A	p.Ser157Ser
C321.ΔA.2.64	2E+06	A	T	missense_variant	exoX	protein_coding	c.84A>T	p.Lys28Asn
C321.ΔA.2.64	2E+06	C	T	synonymous_variant	yecE	protein_coding	c.177C>T	p.Arg59Arg
C321.ΔA.2.64	2E+06	G	A	missense_variant	yedN	protein_coding	c.845C>T	p.Ala282Val
C321.ΔA.2.64	2E+06	G	A	synonymous_variant	fliK	protein_coding	c.15G>A	p.Ala5Ala
C321.ΔA.2.64	2E+06	T	C	missense_variant	mtfA	protein_coding	c.158T>C	p.Val53Ala
C321.ΔA.2.64	2E+06	A	G	synonymous_variant	yeeE	protein_coding	c.282T>C	p.Cys94Cys
C321.ΔA.2.64	2E+06	G	A	missense_variant	hisD	protein_coding	c.823G>A	p.Ala275Thr
C321.ΔA.2.64	2E+06	G	A	missense_variant	rfbD	protein_coding	c.763C>T	p.Pro255Ser
C321.ΔA.2.64	2E+06	C	T	missense_variant	wcaK	protein_coding	c.208G>A	p.Ala70Thr
C321.ΔA.2.64	2E+06	G	A	missense_variant	yegD	protein_coding	c.1186G>A	p.Ala396Thr
C321.ΔA.2.64	2E+06	T	C	missense_variant	gatD	protein_coding	c.631A>G	p.Asn211Asp
C321.ΔA.2.64	2E+06	T	C	missense_variant	yehT	protein_coding	c.191A>G	p.Glu64Gly
C321.ΔA.2.64	2E+06	G	A	stop_gained	bglX	protein_coding	c.1501C>T	p.Gln501*

C321.ΔA.2.64	2E+06	T	C	missense_variant	cdd	protein_coding	c.872T>C	p.Val291Ala
C321.ΔA.2.64	2E+06	C	T	stop_retained_variant	yejO	protein_coding	c.3266G>A	p.Ter1089Ter
C321.ΔA.2.64	2E+06	G	A	missense_variant	atoA	protein_coding	c.514G>A	p.Ala172Thr
C321.ΔA.2.64	2E+06	T	G	missense_variant	menE	protein_coding	c.1166A>C	p.Glu389Ala
C321.ΔA.2.64	2E+06	G	A	missense_variant	nuoG	protein_coding	c.215C>T	p.Ser72Phe
C321.ΔA.2.64	2E+06	T	C	synonymous_variant	alaA	protein_coding	c.801T>C	p.Gly267Gly
C321.ΔA.2.64	2E+06	A	G	missense_variant	yfcO	protein_coding	c.746T>C	p.Val249Ala
C321.ΔA.2.64	2E+06	G	A	missense_variant	fadJ	protein_coding	c.1594C>T	p.Pro532Ser
C321.ΔA.2.64	2E+06	G	A	missense_variant	emrK	protein_coding	c.626C>T	p.Thr209Met
C321.ΔA.2.64	3E+06	C	A	missense_variant	ypdB	protein_coding	c.353C>A	p.Ala118Asp
C321.ΔA.2.64	3E+06	A	G	missense_variant	yfeA	protein_coding	c.1873T>C	p.Tyr625His
C321.ΔA.2.64	3E+06	T	C	synonymous_variant	ucpA	protein_coding	c.63A>G	p.Gly21Gly
C321.ΔA.2.64	3E+06	A	G	missense_variant	tmcA	protein_coding	c.2009T>C	p.Phe670Ser
C321.ΔA.2.64	3E+06	G	A	missense_variant	ypfJ	protein_coding	c.185C>T	p.Pro62Leu
C321.ΔA.2.64	3E+06	T	C	synonymous_variant	ispG	protein_coding	c.228A>G	p.Lys76Lys
C321.ΔA.2.64	3E+06	T	C	intergenic_region	ndk-pbpC			n.2648908T>C
C321.ΔA.2.64	3E+06	G	A	missense_variant	hcaE	protein_coding	c.491G>A	p.Arg164His

C321.ΔA.2.64	3E+06	T	C	missense_variant	hmp	protein_coding	c.619T>C	p.Ser207Pro
C321.ΔA.2.64	3E+06	T	C	missense_variant	yfjP	protein_coding	c.637T>C	p.Cys213Arg
C321.ΔA.2.64	3E+06	T	C	missense_variant	ypjA	protein_coding	c.2437A>G	p.Thr813Ala
C321.ΔA.2.64	3E+06	G	C	missense_variant	ypjA	protein_coding	c.1664C>G	p.Thr555Arg
C321.ΔA.2.64	3E+06	G	A	synonymous_variant	pncC	protein_coding	c.135C>T	p.Ser45Ser
C321.ΔA.2.64	3E+06	G	A	missense_variant	srlR	protein_coding	c.376G>A	p.Val126Ile
C321.ΔA.2.64	3E+06	G	A	missense_variant	norW	protein_coding	c.1109G>A	p.Gly370Glu
C321.ΔA.2.64	3E+06	T	C	synonymous_variant	ygbI	protein_coding	c.465A>G	p.Glu155Glu
C321.ΔA.2.64	3E+06	T	C	intergenic_region	casA-ygcB		n.2886800T>C	
C321.ΔA.2.64	3E+06	T	C	missense_variant	ygcB	protein_coding	c.763A>G	p.Thr255Ala
C321.ΔA.2.64	3E+06	C	A	synonymous_variant	cysH	protein_coding	c.720G>T	p.Gly240Gly
C321.ΔA.2.64	3E+06	G	A	missense_variant	truC	protein_coding	c.166C>T	p.Pro56Ser
C321.ΔA.2.64	3E+06	A	G	missense_variant	hyuA	protein_coding	c.1184A>G	p.Asp395Gly
C321.ΔA.2.64	3E+06	G	A	synonymous_variant	lysS	protein_coding	c.585C>T	p.Arg195Arg
C321.ΔA.2.64	3E+06	A	G	missense_variant	fbaA	protein_coding	c.11T>C	p.Ile4Thr
C321.ΔA.2.64	3E+06	T	C	missense_variant	cmtB	protein_coding	c.359A>G	p.Glu120Gly
C321.ΔA.2.64	3E+06	T	C	missense_variant	glcB	protein_coding	c.1009A>G	p.Ile337Val

C321.ΔA.2.64	3E+06	T	C	synonymous_variant	pitB	protein_coding	c.147A>G	p.Gln49Gln
C321.ΔA.2.64	3E+06	A	C	synonymous_variant	gss	protein_coding	c.1779T>G	p.Gly593Gly
C321.ΔA.2.64	3E+06	G	A	missense_variant	gss	protein_coding	c.1328C>T	p.Ala443Val
C321.ΔA.2.64	3E+06	G	A	synonymous_variant	hybB	protein_coding	c.22C>T	p.Leu8Leu
C321.ΔA.2.64	3E+06	T	C	missense_variant	parC	protein_coding	c.1415A>G	p.Tyr472Cys
C321.ΔA.2.64	3E+06	T	C	missense_variant	cpdA	protein_coding	c.638A>G	p.Asn213Ser
C321.ΔA.2.64	3E+06	A	G	synonymous_variant	nudF	protein_coding	c.465T>C	p.Gly155Gly
C321.ΔA.2.64	3E+06	A	G	synonymous_variant	glnE	protein_coding	c.108T>C	p.Leu36Leu
C321.ΔA.2.64	3E+06	C	T	missense_variant	sstT	protein_coding	c.1232C>T	p.Ala411Val
C321.ΔA.2.64	3E+06	T	C	synonymous_variant	dcuD	protein_coding	c.72T>C	p.Pro24Pro
C321.ΔA.2.64	3E+06	A	G	missense_variant	degS	protein_coding	c.545A>G	p.Asn182Ser
C321.ΔA.2.64	3E+06	A	G	missense_variant	aaeB	protein_coding	c.1477T>C	p.Cys493Arg
C321.ΔA.2.64	3E+06	G	A	missense_variant	csrD	protein_coding	c.1739C>T	p.Thr580Met
C321.ΔA.2.64	3E+06	T	C	synonymous_variant	yhdW	protein_coding	c.564T>C	p.Asn188Asn
C321.ΔA.2.64	4E+06	C	T	synonymous_variant	frlA	protein_coding	c.1071C>T	p.Phe357Phe
C321.ΔA.2.64	4E+06	G	A	missense_variant	hofM	protein_coding	c.259C>T	p.Pro87Ser
C321.ΔA.2.64	4E+06	T	C	missense_variant	gntT	protein_coding	c.611T>C	p.Ile204Thr

C321.ΔA.2.64	4E+06	G	A	missense_variant	glgP	protein_coding	c.1721C>T	p.Ala574Val
C321.ΔA.2.64	4E+06	A	G	synonymous_variant	yhgN	protein_coding	c.480A>G	p.Leu160Leu
C321.ΔA.2.64	4E+06	T	C	missense_variant	yhhY	protein_coding	c.433T>C	p.Tyr145His
C321.ΔA.2.64	4E+06	A	G	intergenic_region	yrhD-yrhB		n.3585376A>G	
C321.ΔA.2.64	4E+06	C	T	synonymous_variant	yhhL	protein_coding	c.75C>T	p.Phe25Phe
C321.ΔA.2.64	4E+06	T	C	synonymous_variant	zntA	protein_coding	c.1758T>C	p.Asn586Asn
C321.ΔA.2.64	4E+06	T	C	missense_variant	yhiJ	protein_coding	c.547A>G	p.Ile183Val
C321.ΔA.2.64	4E+06	T	C	synonymous_variant	pitA	protein_coding	c.567T>C	p.Arg189Arg
C321.ΔA.2.64	4E+06	T	C	synonymous_variant	gadE	protein_coding	c.327T>C	p.His109His
C321.ΔA.2.64	4E+06	A	G	synonymous_variant	insK	protein_coding	c.588A>G	p.Arg196Arg
C321.ΔA.2.64	4E+06	C	T	missense_variant	yiaK	protein_coding	c.437C>T	p.Pro146Leu
C321.ΔA.2.64	4E+06	C	A	missense_variant	yiaM	protein_coding	c.257C>A	p.Ser86Tyr
C321.ΔA.2.64	4E+06	G	A	synonymous_variant	yiaS	protein_coding	c.93G>A	p.Ala31Ala
C321.ΔA.2.64	4E+06	G	A	synonymous_variant	aldB	protein_coding	c.553C>T	p.Leu185Leu
C321.ΔA.2.64	4E+06	G	A	missense_variant	yibI	protein_coding	c.260C>T	p.Ala87Val
C321.ΔA.2.64	4E+06	T	C	missense_variant	lldP	protein_coding	c.382T>C	p.Phe128Leu
C321.ΔA.2.64	4E+06	A	G	intergenic_region	lldD-trmL		n.3783077A>G	
C321.ΔA.2.64	4E+06	G	A	synonymous_variant	nepI	protein_coding	c.600C>T	p.Gly200Gly

C321.ΔA.2.64	4E+06	T	C	missense_variant	ade	protein_coding	c.1562T>C	p.Leu521Pro
C321.ΔA.2.64	4E+06	A	G	missense_variant	ilvB	protein_coding	c.1237T>C	p.Ser413Pro
C321.ΔA.2.64	4E+06	C	T	intergenic_region	ilvB-ivbL	n.3854792C>T		
C321.ΔA.2.64	4E+06	A	C	missense_variant	ivbL	protein_coding	c.59T>G	p.Val20Gly
C321.ΔA.2.64	4E+06	G	A	intergenic_region	dgoR-yidX	n.3877287G>A		
C321.ΔA.2.64	4E+06	G	A	synonymous_variant	gyrB	protein_coding	c.81C>T	p.Ile27Ile
C321.ΔA.2.64	4E+06	T	C	synonymous_variant	yieE	protein_coding	c.228T>C	p.Asn76Asn
C321.ΔA.2.64	4E+06	T	C	synonymous_variant	atpD	protein_coding	c.1365A>G	p.Glu455Glu
C321.ΔA.2.64	4E+06	T	C	missense_variant	atpH	protein_coding	c.329A>G	p.Asp110Gly
C321.ΔA.2.64	4E+06	G	A	splice_region_variant&stop_retained_variant	rbsD	protein_coding	c.419G>A	p.Ter140Ter
C321.ΔA.2.64	4E+06	G	A	splice_region_variant&stop_retained_variant	rbsK	protein_coding	c.929G>A	p.Ter310Ter
C321.ΔA.2.64	4E+06	C	T	splice_region_variant&stop_retained_variant	yieP	protein_coding	c.692G>A	p.Ter231Ter
C321.ΔA.2.64	4E+06	G	A	splice_region_variant&stop_retained_variant	ilvM	protein_coding	c.263G>A	p.Ter88Ter
C321.ΔA.2.64	4E+06	G	A	splice_region_variant&sto	ubiB	protein_coding	c.1640G>A	p.Ter547Ter

				p_retained_variant				
C321.ΔA.2.64	4E+06	G	A	splice_region_variant&stop_retained_variant	ubiD	protein_coding	c.1493G>A	p.Ter498Ter
C321.ΔA.2.64	4E+06	T	C	synonymous_variant	pepQ	protein_coding	c.618T>C	p.Gly206Gly
C321.ΔA.2.64	4E+06	T	C	missense_variant	glnG	protein_coding	c.32A>G	p.Asp11Gly
C321.ΔA.2.64	4E+06	C	T	splice_region_variant&stop_retained_variant	ftsN	protein_coding	c.959G>A	p.Ter320Ter
C321.ΔA.2.64	4E+06	T	C	missense_variant	ptsA	protein_coding	c.811A>G	p.Met271Val
C321.ΔA.2.64	4E+06	G	A	splice_region_variant&stop_retained_variant	murI	protein_coding	c.857G>A	p.Ter286Ter
C321.ΔA.2.64	4E+06	C	A	intergenic_region	rrlB-rrfB		n.4173476C>A	
C321.ΔA.2.64	4E+06	G	A	splice_region_variant&stop_retained_variant	secE	protein_coding	c.383G>A	p.Ter128Ter
C321.ΔA.2.64	4E+06	G	A	missense_variant	hemE	protein_coding	c.255G>A	p.Met85Ile
C321.ΔA.2.64	4E+06	A	G	missense_variant	yjaB	protein_coding	c.430T>C	p.Tyr144His
C321.ΔA.2.64	4E+06	G	A	synonymous_variant	metA	protein_coding	c.540G>A	p.Thr180Thr
C321.ΔA.2.64	4E+06	A	G	missense_variant	metH	protein_coding	c.2620A>G	p.Thr874Ala

C321.ΔA.2.64	4E+06	C	T	synonymous_variant	metH	protein_coding	c.2950C>T	p.Leu984Leu
C321.ΔA.2.64	4E+06	A	G	missense_variant	metH	protein_coding	c.3569A>G	p.Tyr1190Cys
C321.ΔA.2.64	4E+06	C	T	missense_variant	lamB	protein_coding	c.725C>T	p.Ala242Val
C321.ΔA.2.64	4E+06	G	A	splice_region_variant&stop_retained_variant	ubiA	protein_coding	c.872G>A	p.Ter291Ter
C321.ΔA.2.64	4E+06	A	G	synonymous_variant	dinF	protein_coding	c.453A>G	p.Glu151Glu
C321.ΔA.2.64	4E+06	G	A	splice_region_variant&stop_retained_variant	ssb	protein_coding	c.536G>A	p.Ter179Ter
C321.ΔA.2.64	4E+06	C	T	splice_region_variant&stop_retained_variant	alsK	protein_coding	c.929G>A	p.Ter310Ter
C321.ΔA.2.64	4E+06	C	G	intergenic_region	adiY-adiA		n.4340053C>G	
C321.ΔA.2.64	4E+06	C	T	synonymous_variant	yjdF	protein_coding	c.369G>A	p.Gly123Gly
C321.ΔA.2.64	4E+06	A	G	synonymous_variant	cadA	protein_coding	c.1866T>C	p.Ala622Ala
C321.ΔA.2.64	4E+06	A	G	synonymous_variant	cadC	protein_coding	c.1488T>C	p.Thr496Thr
C321.ΔA.2.64	4E+06	G	A	stop_gained	ampC	protein_coding	c.67C>T	p.Gln23*
C321.ΔA.2.64	4E+06	C	T	missense_variant	yjfM	protein_coding	c.430C>T	p.Arg144Cys
C321.ΔA.2.64	4E+06	T	C	intergenic_region	yjfC-aidB		n.4416174T>C	
C321.ΔA.2.64	4E+06	A	G	missense_variant	ytfE	protein_coding	c.530T>C	p.Val177Ala

C321.ΔA.2.64	5E+06	A	G	missense_variant	insO	protein_coding	c.610A>G	p.Thr204Ala
C321.ΔA.2.64	5E+06	A	G	missense_variant	insI1	protein_coding	c.812T>C	p.Val271Ala
C321.ΔA.2.64	5E+06	C	T	missense_variant	insA	protein_coding	c.106C>T	p.Arg36Cys
C321.ΔA.2.64	5E+06	T	C	missense_variant	fimB	protein_coding	c.266T>C	p.Leu89Ser
C321.ΔA.2.64	5E+06	C	T	synonymous_variant	opgB	protein_coding	c.315G>A	p.Pro105Pro
C321.ΔA.2.64	5E+06	A	G	intergenic_region	fhuF-yjjZ	n.4607664A>G		
C321.ΔA.2.64	5E+06	C	T	synonymous_variant	yjjI	protein_coding	c.744G>A	p.Thr248Thr
C321.ΔA.2.64	20575	TG	T	intergenic_region	insA-rpsT	n.20576delG		
C321.ΔA.2.64	34112	GT	G	intergenic_region	carB-caiF	n.34113delT		
C321.ΔA.2.64	45463	T	TATA	start_lost&disruptive_inframe_insertion	fixX	protein_coding	c.2_3insAAT	
C321.ΔA.2.64	45463	T	TATA	splice_region_variant&disruptive_inframe_insertion	fixC	protein_coding	c.1285_1286insAAT	p.Ter429dup
C321.ΔA.2.64	52429	C	CATT	start_lost&disruptive_inframe_insertion	rsmA	protein_coding	c.-1_2dupAAT	
C321.ΔA.2.64	52429	C	CATT	splice_region_variant&conservative_inframe_insertion	pdxA	protein_coding	c.986_988dupAAT	p.Ter330dup
C321.ΔA.2.64	94480	C	CATA	start_lost&disruptive_inframe_insertion	murF	protein_coding	c.2_3insAAT	

C321.ΔA.2.64	94480	C	CATA	splice_region_variant&disruptive_inframe_insertion	murE	protein_coding	c.1486_1487insAAT	p.Ter496dup
C321.ΔA.2.64	99474	G	GATA	start_lost&disruptive_inframe_insertion	murG	protein_coding	c.2_3insAAT	
C321.ΔA.2.64	99474	G	GATA	splice_region_variant&disruptive_inframe_insertion	ftsW	protein_coding	c.1243_1244insAAT	p.Ter415dup
C321.ΔA.2.64	103811	C	CAAT	splice_region_variant&disruptive_inframe_insertion	ftsQ	protein_coding	c.827_829dupAAT	p.Ter277dup
C321.ΔA.2.64	136897	A	AG	intergenic_region	yacC-cueO	n.136897_136898insG		
C321.ΔA.2.64	204321	C	CAAT	splice_region_variant&disruptive_inframe_insertion	lpxB	protein_coding	c.1145_1147dupAAT	p.Ter383dup
C321.ΔA.2.64	214662	C	CATA	start_lost&disruptive_inframe_insertion	arfB	protein_coding	c.2_3insAAT	
C321.ΔA.2.64	214662	C	CATA	splice_region_variant&disruptive_inframe_insertion	yaeQ	protein_coding	c.544_545insAAT	p.Ter182dup
C321.ΔA.2.64	232880	CG	C	frameshift_variant	mltD	protein_coding	c.905delC	p.Pro302fs
C321.ΔA.2.64	298918	C	CG	frameshift_variant	paoC	protein_coding	c.1068dupC	p.Gly357fs
C321.ΔA.2.64	379063	A	AG	intergenic_region	frmR-yaiO	n.379063_379064insG		
C321.ΔA.2.64	414800	C	CATT	start_lost&disrup	sbcC	protein_coding	c.2_3insAAT	

				tive_infr ame_ins ertion				
C321.ΔA.2.64	414800	C	CATT	splice_re gion_var iant&co nservativ e_infram e_inserti on	sbcD	protein_ coding	c.1201_1 202insA AT	p.Ter401 dup
C321.ΔA.2.64	440454	C	CA	intergenic_r egion	xseB-thiI		n.440454_4 40455insA	
C321.ΔA.2.64	560739	T	TG	frameshi ft_varian t	sfmD	protein_ coding	c.2000d upG	p.Gly668 fs
C321.ΔA.2.64	650998	GA	G	intergenic_r egion	citC-dpiB		n.650999del A	
C321.ΔA.2.64	719619	C	CT	intergenic_r egion	speF-ybfK		n.719619_7 19620insT	
C321.ΔA.2.64	722567	GT	G	frameshi ft_varian t	kdpD	protein_ coding	c.895del A	p.Thr299 fs
C321.ΔA.2.64	785427	G	GA	frameshi ft_varian t	aroG	protein_ coding	c.750du pA	p.Glu251 fs
C321.ΔA.2.64	1E+06	A	AG	frameshi ft_varian t	ycgB	protein_ coding	c.969du pC	p.Tyr324 fs
C321.ΔA.2.64	1E+06	G	GT	frameshi ft_varian t	rssB	protein_ coding	c.363du pT	p.Ala122 fs
C321.ΔA.2.64	1E+06	A	AG	frameshi ft_varian t	acnA	protein_ coding	c.2211d upG	p.Met73 8fs
C321.ΔA.2.64	1E+06	GA	G	frameshi ft_varian t	ydaY	protein_ coding	c.397del A	p.Arg133 fs
C321.ΔA.2.64	2E+06	TC	T	frameshi ft_varian t	yneO	protein_ coding	c.3739d elG	p.Glu124 7fs
C321.ΔA.2.64	2E+06	TA	T	frameshi ft_varian t	ydiQ	protein_ coding	c.591del A	p.Lys197 fs
C321.ΔA.2.64	2E+06	G	GA	frameshi ft_varian t	manX	protein_ coding	c.858du pA	p.Asp28 7fs
C321.ΔA.2.64	2E+06	A	AT	frameshi ft_varian t&stop_ gained	exoX	protein_ coding	c.81_82i nsT	p.Lys28f s

C321.ΔA.2.64	2E+06	GC	G	frameshift_variant	cheA	protein_coding	c.924delG	p.Gln309fs
C321.ΔA.2.64	2E+06	AC	A	frameshift_variant	yfbS	protein_coding	c.1530delG	p.Cys511fs
C321.ΔA.2.64	2E+06	G	GA	intergenic_region	yfcE-yfcF	n.2423830_2423831insA		
C321.ΔA.2.64	3E+06	G	GC	frameshift_variant	hyfB	protein_coding	c.200dupC	p.Glu69fs
C321.ΔA.2.64	3E+06	AT	A	intergenic_region	kgtP-rrfG	n.2730031delT		
C321.ΔA.2.64	3E+06	T	TC	frameshift_variant	yfiH	protein_coding	c.499dupG	p.Glu167fs
C321.ΔA.2.64	3E+06	A	AG	intergenic_region	yfjM-rnlA	n.2769889_2769890insG		
C321.ΔA.2.64	3E+06	C	CA	frameshift_variant	ygcB	protein_coding	c.2057dupT	p.Leu686fs
C321.ΔA.2.64	3E+06	A	AC	intergenic_region	bacA-folB	n.3204822_3204823insC		
C321.ΔA.2.64	3E+06	CCA	C	frameshift_variant	dnaG	protein_coding	c.1726_1727delCA	p.Gln576fs
C321.ΔA.2.64	3E+06	G	GCC	frameshift_variant	dnaG	protein_coding	c.1728_1729insCC	p.Glu577fs
C321.ΔA.2.64	3E+06	GA	G	frameshift_variant	ygjV	protein_coding	c.68delT	p.Phe23fs
C321.ΔA.2.64	3E+06	T	TC	frameshift_variant	yhcM	protein_coding	c.738dupG	p.Lys247fs
C321.ΔA.2.64	4E+06	G	GC	frameshift_variant	yhhS	protein_coding	c.1018dupG	p.Ala340fs
C321.ΔA.2.64	4E+06	G	GC	frameshift_variant	yhjA	protein_coding	c.650dupG	p.Val218fs
C321.ΔA.2.64	4E+06	CA	C	frameshift_variant	lyxK	protein_coding	c.1425delA	p.Lys475fs
C321.ΔA.2.64	4E+06	GC	G	frameshift_variant	bglF	protein_coding	c.41delG	p.Gly14fs
C321.ΔA.2.64	4E+06	AT	A	intergenic_region	ravA-kup	n.3933103delT		

C321.ΔA.2.64	4E+06	C	CGG	frameshift_variant	wzzE	protein_coding	c.1002_1003dupGG	p.Ala335fs
C321.ΔA.2.64	4E+06	C	CATT	start_lost&disruptive_inframe_insertion	hemD	protein_coding	c.2_3insAAT	
C321.ΔA.2.64	4E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	hemC	protein_coding	c.940_941insAAT	p.Ter314dup
C321.ΔA.2.64	4E+06	A	AAAT	splice_region_variant&disruptive_inframe_insertion	yigP	protein_coding	c.602_604dupAAT	p.Ter202dup
C321.ΔA.2.64	4E+06	A	AAAT	splice_region_variant&non_coding_transcript_exon_variant	esrE	lincRNA	n.248_250dupAAT	
C321.ΔA.2.64	4E+06	ATC	A	intergenic_region	ubiB-tatA		n.4023794_4023795delTC	
C321.ΔA.2.64	4E+06	C	CATA	start_lost&disruptive_inframe_insertion	birA	protein_coding	c.2_3insAAT	
C321.ΔA.2.64	4E+06	C	CATA	splice_region_variant&disruptive_inframe_insertion	murB	protein_coding	c.1027_1028insAAT	p.Ter343dup
C321.ΔA.2.64	4E+06	GC	G	frameshift_variant	phnP	protein_coding	c.59delG	p.Gly20fs
C321.ΔA.2.64	4E+06	GA	G	frameshift_variant	mutL	protein_coding	c.98delA	p.Asn33fs
C321.ΔA.2.64	4E+06	AC	A	frameshift_variant	msrA	protein_coding	c.275delG	p.Gly92fs

C321.ΔA.2.64	5E+06	C	CA	intergenic_region	yjjM-yjjN	n.4597823_4597824insA		
C321.ΔA.2.64	5E+06	A	AG	frameshift_variant	leuV	protein_coding	c.79dupC	p.Leu27fs
C321.ΔA.2.64	5E+06	A	AG	frameshift_variant	leuP	protein_coding	c.79dupC	p.Leu27fs
C321.ΔA.2.64	5E+06	A	AG	frameshift_variant	leuQ	protein_coding	c.79dupC	p.Leu27fs
C321.ΔA.2.39	39772	T	C	missense_variant	caiA	protein_coding	c.616A>G	p.Ser206Gly
C321.ΔA.2.39	67565	A	G	missense_variant	araA	protein_coding	c.604T>C	p.Phe202Leu
C321.ΔA.2.39	79002	C	T	missense_variant	leuD	protein_coding	c.283G>A	p.Asp95Asn
C321.ΔA.2.39	112739	A	G	synonymous_variant	coaE	protein_coding	c.312T>C	p.Thr104Thr
C321.ΔA.2.39	195468	A	G	missense_variant	ispU	protein_coding	c.736A>G	p.Thr246Ala
C321.ΔA.2.39	219056	A	G	synonymous_variant	tsaA	protein_coding	c.369T>C	p.Ser123Ser
C321.ΔA.2.39	264888	T	C	missense_variant	ykfB	protein_coding	c.253A>G	p.Thr85Ala
C321.ΔA.2.39	307631	A	G	missense_variant	ecpC	protein_coding	c.754T>C	p.Ser252Pro
C321.ΔA.2.39	339757	C	T	missense_variant	yahI	protein_coding	c.542C>T	p.Ala181Val
C321.ΔA.2.39	340807	A	G	missense_variant	yahJ	protein_coding	c.632A>G	p.Glu211Gly
C321.ΔA.2.39	351434	C	T	missense_variant	prpD	protein_coding	c.1169C>T	p.Ala390Val
C321.ΔA.2.39	401673	C	T	missense_variant	phoA	protein_coding	c.878C>T	p.Pro293Leu
C321.ΔA.2.39	464355	A	G	missense_variant	queC	protein_coding	c.242T>C	p.Val81Ala

C321.ΔA.2.39	527850	A	G	synonymous_variant	ybbD	protein_coding	c.162A>G	p.Gly54Gly
C321.ΔA.2.39	534559	C	T	missense_variant	gcl	protein_coding	c.1595C>T	p.Ala532Val
C321.ΔA.2.39	554082	C	T	synonymous_variant	cysS	protein_coding	c.423C>T	p.Asp141Asp
C321.ΔA.2.39	583801	T	C	missense_variant	ompT	protein_coding	c.881A>G	p.Asp294Gly
C321.ΔA.2.39	675236	A	G	missense_variant	ybeQ	protein_coding	c.359T>C	p.Val120Ala
C321.ΔA.2.39	679621	A	G	synonymous_variant	djlC	protein_coding	c.363A>G	p.Glu121Glu
C321.ΔA.2.39	693615	A	G	missense_variant	miaB	protein_coding	c.388T>C	p.Phe130Leu
C321.ΔA.2.39	697957	C	T	missense_variant	asnB	protein_coding	c.268G>A	p.Asp90Asn
C321.ΔA.2.39	715065	T	C	missense_variant	ybfG	protein_coding	c.580A>G	p.Thr194Ala
C321.ΔA.2.39	744222	G	A	missense_variant	ybgL	protein_coding	c.10G>A	p.Asp4Asn
C321.ΔA.2.39	744222	G	A	splice_region_variant&stop_retained_variant	ybgK	protein_coding	c.932G>A	p.Ter311Ter
C321.ΔA.2.39	754954	G	A	splice_region_variant&stop_retained_variant	sdhD	protein_coding	c.347G>A	p.Ter116Ter
C321.ΔA.2.39	757053	A	G	missense_variant	sdhB	protein_coding	c.317A>G	p.Asp106Gly
C321.ΔA.2.39	771998	G	A	missense_variant	cydA	protein_coding	c.1493G>A	p.Arg498His
C321.ΔA.2.39	774168	A	G	synonymous_variant	ybgC	protein_coding	c.369A>G	p.Ala123Ala

C321.ΔA.2.39	774870	G	A	missense_variant	tolQ	protein_coding	c.670G>A	p.Val224Ile
C321.ΔA.2.39	775990	G	A	missense_variant	tolA	protein_coding	c.601G>A	p.Glu201Lys
C321.ΔA.2.39	778079	G	A	splice_region_variant&stop_retained_variant	tolB	protein_coding	c.1292G>A	p.Ter431Ter
C321.ΔA.2.39	819327	C	T	stop_gained	bioC	protein_coding	c.700C>T	p.Gln234*
C321.ΔA.2.39	848553	C	T	missense_variant	fiu	protein_coding	c.85G>A	p.Ala29Thr
C321.ΔA.2.39	916196	G	A	missense_variant	ltaE	protein_coding	c.205C>T	p.His69Tyr
C321.ΔA.2.39	933332	C	T	splice_region_variant&stop_retained_variant	infA	protein_coding	c.218G>A	p.Ter73Ter
C321.ΔA.2.39	940119	C	T	synonymous_variant	lrp	protein_coding	c.420C>T	p.Gly140Gly
C321.ΔA.2.39	978135	C	T	missense_variant	kdsB	protein_coding	c.179C>T	p.Ala60Val
C321.ΔA.2.39	983284	C	T	synonymous_variant	mukE	protein_coding	c.558C>T	p.Arg186Arg
C321.ΔA.2.39	989019	A	G	missense_variant	ycbB	protein_coding	c.868A>G	p.Thr290Ala
C321.ΔA.2.39	1E+06	C	T	missense_variant	agp	protein_coding	c.170C>T	p.Thr57Met
C321.ΔA.2.39	1E+06	C	T	missense_variant	rutA	protein_coding	c.1138G>A	p.Glu380Lys
C321.ΔA.2.39	1E+06	G	A	splice_region_variant&stop_retained_variant	yceQ	protein_coding	c.320G>A	p.Ter107Ter
C321.ΔA.2.39	1E+06	G	A	splice_region_var	fabG	protein_coding	c.734G>A	p.Ter245Ter

				iant&stop_retained_variant				
C321.ΔA.2.39	1E+06	C	T	synonymous_variant	lolD	protein_coding	c.168C>T	p.Gly56Gly
C321.ΔA.2.39	1E+06	C	A	missense_variant	ymfE	protein_coding	c.443G>T	p.Trp148Leu
C321.ΔA.2.39	1E+06	G	A	synonymous_variant	intE	protein_coding	c.543C>T	p.Ala181Ala
C321.ΔA.2.39	1E+06	A	G	synonymous_variant	dhaR	protein_coding	c.789A>G	p.Lys263Lys
C321.ΔA.2.39	1E+06	C	T	missense_variant	hemA	protein_coding	c.587C>T	p.Pro196Leu
C321.ΔA.2.39	1E+06	A	G	missense_variant	trpE	protein_coding	c.1180T>C	p.Ser394Pro
C321.ΔA.2.39	1E+06	C	T	synonymous_variant	insH1	protein_coding	c.453C>T	p.Gly151Gly
C321.ΔA.2.39	1E+06	A	G	missense_variant	uspE	protein_coding	c.152T>C	p.Leu51Pro
C321.ΔA.2.39	1E+06	G	A	synonymous_variant	recE	protein_coding	c.2571C>T	p.Pro857Pro
C321.ΔA.2.39	1E+06	T	C	missense_variant	trkG	protein_coding	c.565T>C	p.Trp189Arg
C321.ΔA.2.39	1E+06	C	T	synonymous_variant	ydbK	protein_coding	c.1170G>A	p.Ala390Ala
C321.ΔA.2.39	1E+06	T	C	synonymous_variant	tynA	protein_coding	c.702A>G	p.Glu234Glu
C321.ΔA.2.39	1E+06	A	G	synonymous_variant	ydbD	protein_coding	c.1239A>G	p.Gln413Gln
C321.ΔA.2.39	1E+06	G	T	missense_variant	hrpA	protein_coding	c.3136G>T	p.Asp1046Tyr
C321.ΔA.2.39	2E+06	A	G	synonymous_variant	ydcU	protein_coding	c.756A>G	p.Thr252Thr
C321.ΔA.2.39	2E+06	C	T	intergenic_region	curA-mcbR	n.1524889C>T		

C321.ΔA.2.39	2E+06	T	C	missense_variant	narZ	protein_coding	c.1838A>G	p.Asn613Ser
C321.ΔA.2.39	2E+06	G	A	splice_region_variant&stop_retained_variant	ydfB	protein_coding	c.128G>A	p.Ter43Ter
C321.ΔA.2.39	2E+06	G	A	synonymous_variant	ydfJ	protein_coding	c.4047C>T	p.Ile1349Ile
C321.ΔA.2.39	2E+06	G	A	synonymous_variant	ydfJ	protein_coding	c.1551C>T	p.Phe517Phe
C321.ΔA.2.39	2E+06	G	A	synonymous_variant	insD1	protein_coding	c.444G>A	p.Pro148Pro
C321.ΔA.2.39	2E+06	T	C	missense_variant	pntB	protein_coding	c.1307A>G	p.Asn436Ser
C321.ΔA.2.39	2E+06	C	T	splice_region_variant&stop_retained_variant	ribC	protein_coding	c.641G>A	p.Ter214Ter
C321.ΔA.2.39	2E+06	T	C	synonymous_variant	ppsR	protein_coding	c.663T>C	p.Arg221Arg
C321.ΔA.2.39	2E+06	C	T	splice_region_variant&stop_retained_variant	pheT	protein_coding	c.2387G>A	p.Ter796Ter
C321.ΔA.2.39	2E+06	T	C	missense_variant	astC	protein_coding	c.45A>G	p.Ile15Met
C321.ΔA.2.39	2E+06	T	C	missense_variant	gdhA	protein_coding	c.271T>C	p.Tyr91His
C321.ΔA.2.39	2E+06	A	G	missense_variant	ydjK	protein_coding	c.662T>C	p.Val221Ala
C321.ΔA.2.39	2E+06	T	C	synonymous_variant	fadD	protein_coding	c.1050A>G	p.Gly350Gly
C321.ΔA.2.39	2E+06	C	T	splice_region_variant&stop_retain	tsaB	protein_coding	c.695G>A	p.Ter232Ter

				ed_variant				
C321.ΔA.2.39	2E+06	G	A	synonymous_variant	yobB	protein_coding	c.471G>A	p.Ser157Ser
C321.ΔA.2.39	2E+06	A	T	missense_variant	exoX	protein_coding	c.84A>T	p.Lys28Asn
C321.ΔA.2.39	2E+06	A	G	missense_variant	edd	protein_coding	c.1328T>C	p.Val443Ala
C321.ΔA.2.39	2E+06	C	T	splice_region_variant&stop_retained_variant	aspS	protein_coding	c.1772G>A	p.Ter591Ter
C321.ΔA.2.39	2E+06	C	T	splice_region_variant&stop_retained_variant	pgsA	protein_coding	c.548G>A	p.Ter183Ter
C321.ΔA.2.39	2E+06	G	A	synonymous_variant	fliK	protein_coding	c.15G>A	p.Ala5Ala
C321.ΔA.2.39	2E+06	G	A	missense_variant	yegD	protein_coding	c.1186G>A	p.Ala396Thr
C321.ΔA.2.39	2E+06	G	A	stop_gained	bglX	protein_coding	c.1501C>T	p.Gln501*
C321.ΔA.2.39	2E+06	T	C	missense_variant	cdd	protein_coding	c.872T>C	p.Val291Ala
C321.ΔA.2.39	2E+06	C	T	splice_region_variant&stop_retained_variant	folE	protein_coding	c.668G>A	p.Ter223Ter
C321.ΔA.2.39	2E+06	C	T	synonymous_variant	fruA	protein_coding	c.1353G>A	p.Ala451Ala
C321.ΔA.2.39	2E+06	A	G	missense_variant	fruA	protein_coding	c.1349T>C	p.Val450Ala
C321.ΔA.2.39	2E+06	G	A	splice_region_variant&stop_retained_variant	yejF	protein_coding	c.1589G>A	p.Ter530Ter

C321.ΔA.2.39	2E+06	G	A	splice_region_variant&stop_retained_variant	yejM	protein_coding	c.1760G>A	p.Ter587Ter
C321.ΔA.2.39	2E+06	C	T	synonymous_variant	yejO	protein_coding	c.3774G>A	p.Val1258Val
C321.ΔA.2.39	2E+06	A	G	missense_variant	dsbE	protein_coding	c.212T>C	p.Val71Ala
C321.ΔA.2.39	2E+06	G	A	splice_region_variant&stop_retained_variant	nrdA	protein_coding	c.2285G>A	p.Ter762Ter
C321.ΔA.2.39	2E+06	G	A	splice_region_variant&stop_retained_variant	nrdB	protein_coding	c.1130G>A	p.Ter377Ter
C321.ΔA.2.39	2E+06	A	G	synonymous_variant	menB	protein_coding	c.465T>C	p.Thr155Thr
C321.ΔA.2.39	2E+06	C	T	splice_region_variant&stop_retained_variant	accD	protein_coding	c.914G>A	p.Ter305Ter
C321.ΔA.2.39	2E+06	G	A	missense_variant	emrK	protein_coding	c.626C>T	p.Thr209Met
C321.ΔA.2.39	3E+06	C	T	splice_region_variant&stop_retained_variant	ligA	protein_coding	c.2015G>A	p.Ter672Ter
C321.ΔA.2.39	3E+06	C	T	splice_region_variant&stop_retained_variant	zipA	protein_coding	c.986G>A	p.Ter329Ter
C321.ΔA.2.39	3E+06	T	C	missense_variant	rlmN	protein_coding	c.164A>G	p.Asp55Gly

C321.ΔA.2.39	3E+06	G	A	missense_variant	hcaE	protein_coding	c.491G>A	p.Arg164His
C321.ΔA.2.39	3E+06	C	T	splice_region_variant&stop_retained_variant	rpoE	protein_coding	c.575G>A	p.Ter192Ter
C321.ΔA.2.39	3E+06	A	G	missense_variant	pheA	protein_coding	c.760A>G	p.Ser254Gly
C321.ΔA.2.39	3E+06	G	A	synonymous_variant	intA	protein_coding	c.399G>A	p.Ser133Ser
C321.ΔA.2.39	3E+06	T	C	missense_variant	yfjL	protein_coding	c.542A>G	p.Asn181Ser
C321.ΔA.2.39	3E+06	G	A	synonymous_variant	gabD	protein_coding	c.1347G>A	p.Val449Val
C321.ΔA.2.39	3E+06	T	C	synonymous_variant	proV	protein_coding	c.198T>C	p.Gly66Gly
C321.ΔA.2.39	3E+06	G	A	synonymous_variant	yqaA	protein_coding	c.81C>T	p.Val27Val
C321.ΔA.2.39	3E+06	G	C	synonymous_variant	alaS	protein_coding	c.588C>G	p.Gly196Gly
C321.ΔA.2.39	3E+06	A	G	missense_variant	hycC	protein_coding	c.902T>C	p.Leu301Pro
C321.ΔA.2.39	3E+06	G	A	missense_variant	hypE	protein_coding	c.241G>A	p.Ala81Thr
C321.ΔA.2.39	3E+06	C	T	synonymous_variant	cysI	protein_coding	c.1665G>A	p.Ala555Ala
C321.ΔA.2.39	3E+06	G	A	missense_variant	recC	protein_coding	c.2321C>T	p.Ala774Val
C321.ΔA.2.39	3E+06	C	T	splice_region_variant&stop_retained_variant	lgt	protein_coding	c.875G>A	p.Ter292Ter
C321.ΔA.2.39	3E+06	T	C	missense_variant	metK	protein_coding	c.34T>C	p.Ser12Pro

C321.ΔA.2.39	3E+06	T	G	missense_variant	mutY	protein_coding	c.780T>G	p.Cys260Trp
C321.ΔA.2.39	3E+06	T	C	missense_variant	glcB	protein_coding	c.1009A>G	p.Ile337Val
C321.ΔA.2.39	3E+06	G	A	missense_variant	qseC	protein_coding	c.893G>A	p.Arg298Gln
C321.ΔA.2.39	3E+06	A	G	synonymous_variant	nudF	protein_coding	c.465T>C	p.Gly155Gly
C321.ΔA.2.39	3E+06	C	T	splice_region_variant&stop_retained_variant	ribB	protein_coding	c.653G>A	p.Ter218Ter
C321.ΔA.2.39	3E+06	G	A	splice_region_variant&stop_retained_variant	cca	protein_coding	c.1238G>A	p.Ter413Ter
C321.ΔA.2.39	3E+06	C	T	synonymous_variant	fadH	protein_coding	c.558C>T	p.Gly186Gly
C321.ΔA.2.39	3E+06	T	C	missense_variant	mtr	protein_coding	c.628A>G	p.Asn210Asp
C321.ΔA.2.39	3E+06	T	C	missense_variant	folP	protein_coding	c.652A>G	p.Met218Val
C321.ΔA.2.39	3E+06	G	A	intergenic_region	mlaF-yrbG		n.3340764G>A	
C321.ΔA.2.39	3E+06	A	G	missense_variant	degS	protein_coding	c.545A>G	p.Asn182Ser
C321.ΔA.2.39	3E+06	T	C	missense_variant	tldD	protein_coding	c.946A>G	p.Thr316Ala
C321.ΔA.2.39	3E+06	T	C	missense_variant	rpsD	protein_coding	c.422A>G	p.Asp141Gly
C321.ΔA.2.39	3E+06	C	T	splice_region_variant&stop_retained_variant	rplV	protein_coding	c.332G>A	p.Ter111Ter
C321.ΔA.2.39	3E+06	C	T	splice_region_var	rpsG	protein_coding	c.539G>A	p.Ter180Ter

				iant&stop_retained_variant				
C321.ΔA.2.39	3E+06	G	A	synonymous_variant	cysG	protein_coding	c.1275G>A	p.Gln425Gln
C321.ΔA.2.39	4E+06	C	G	missense_variant	yhhI	protein_coding	c.675C>G	p.Ile225Met
C321.ΔA.2.39	4E+06	T	C	synonymous_variant	pitA	protein_coding	c.567T>C	p.Arg189Arg
C321.ΔA.2.39	4E+06	C	T	missense_variant	yhjD	protein_coding	c.365C>T	p.Thr122Met
C321.ΔA.2.39	4E+06	T	C	missense_variant	eptB	protein_coding	c.1115A>G	p.Tyr372Cys
C321.ΔA.2.39	4E+06	C	T	synonymous_variant	xylG	protein_coding	c.1275C>T	p.Arg425Arg
C321.ΔA.2.39	4E+06	C	T	missense_variant	yiaK	protein_coding	c.437C>T	p.Pro146Leu
C321.ΔA.2.39	4E+06	G	A	synonymous_variant	yiaS	protein_coding	c.93G>A	p.Ala31Ala
C321.ΔA.2.39	4E+06	C	A	missense_variant	hldD	protein_coding	c.612C>A	p.Ser204Arg
C321.ΔA.2.39	4E+06	C	T	splice_region_variant&stop_retained_variant	waaU	protein_coding	c.1073G>A	p.Ter358Ter
C321.ΔA.2.39	4E+06	G	A	splice_region_variant&stop_retained_variant	waaA	protein_coding	c.1277G>A	p.Ter426Ter
C321.ΔA.2.39	4E+06	G	A	splice_region_variant&stop_retained_variant	gmk	protein_coding	c.623G>A	p.Ter208Ter
C321.ΔA.2.39	4E+06	G	A	splice_region_variant&sto	rnpA	protein_coding	c.359G>A	p.Ter120Ter

				p_retained_variant				
C321.ΔA.2.39	4E+06	G	A	synonymous_variant	yidD	protein_coding	c.36G>A	p.Leu12Leu
C321.ΔA.2.39	4E+06	G	A	splice_region_variant&stop_retained_variant	yidC	protein_coding	c.1646G>A	p.Ter549Ter
C321.ΔA.2.39	4E+06	G	A	splice_region_variant&stop_retained_variant	tnaC	protein_coding	c.74G>A	p.Ter25Ter
C321.ΔA.2.39	4E+06	G	A	splice_region_variant&stop_retained_variant	yieE	protein_coding	c.749G>A	p.Ter250Ter
C321.ΔA.2.39	4E+06	T	C	synonymous_variant	bglF	protein_coding	c.1269A>G	p.Gly423Gly
C321.ΔA.2.39	4E+06	C	T	splice_region_variant&stop_retained_variant	bglG	protein_coding	c.836G>A	p.Ter279Ter
C321.ΔA.2.39	4E+06	C	T	splice_region_variant&stop_retained_variant	pstB	protein_coding	c.773G>A	p.Ter258Ter
C321.ΔA.2.39	4E+06	C	T	splice_region_variant&stop_retained_variant	pstA	protein_coding	c.890G>A	p.Ter297Ter
C321.ΔA.2.39	4E+06	C	T	splice_region_variant&stop_retained_variant	glmU	protein_coding	c.1370G>A	p.Ter457Ter

C321.ΔA.2.39	4E+06	T	C	missense_variant	wecB	protein_coding	c.1097T>C	p.Leu366Pro
C321.ΔA.2.39	4E+06	T	C	synonymous_variant	pepQ	protein_coding	c.618T>C	p.Gly206Gly
C321.ΔA.2.39	4E+06	A	G	non_coding_transcript_exon_variant	rrlA	rRNA	n.1427A>G	
C321.ΔA.2.39	4E+06	C	T	missense_variant	polA	protein_coding	c.1709C>T	p.Ala570Val
C321.ΔA.2.39	4E+06	G	A	missense_variant	cpxA	protein_coding	c.550C>T	p.Arg184Trp
C321.ΔA.2.39	4E+06	T	C	missense_variant	pflD	protein_coding	c.1982T>C	p.Val661Ala
C321.ΔA.2.39	4E+06	A	G	missense_variant	argE	protein_coding	c.22T>C	p.Phe8Leu
C321.ΔA.2.39	4E+06	C	A	intergenic_region	rrlB-rrfB	n.4173476C>A		
C321.ΔA.2.39	4E+06	A	G	intergenic_region	rrlB-rrfB	n.4173545A>G		
C321.ΔA.2.39	4E+06	A	C	non_coding_transcript_exon_variant	rrfB	rRNA	n.12A>C	
C321.ΔA.2.39	4E+06	A	G	missense_variant	yjaB	protein_coding	c.430T>C	p.Tyr144His
C321.ΔA.2.39	4E+06	C	T	synonymous_variant	yjcS	protein_coding	c.1314G>A	p.Pro438Pro
C321.ΔA.2.39	4E+06	A	G	synonymous_variant	rpiR	protein_coding	c.336T>C	p.Phe112Phe
C321.ΔA.2.39	4E+06	A	G	missense_variant	yjdA	protein_coding	c.1334A>G	p.Asn445Ser
C321.ΔA.2.39	4E+06	C	T	missense_variant	rnr	protein_coding	c.758C>T	p.Pro253Leu
C321.ΔA.2.39	4E+06	C	T	missense_variant	yjfM	protein_coding	c.430C>T	p.Arg144Cys

C321.ΔA.2.39	4E+06	C	T	synonymous_variant	ulaA	protein_coding	c.228C>T	p.Val76Val
C321.ΔA.2.39	4E+06	A	G	intergenic_region	fklB-cycA		n.4431570A>G	
C321.ΔA.2.39	4E+06	G	A	synonymous_variant	bdcR	protein_coding	c.156G>A	p.Glu52Glu
C321.ΔA.2.39	4E+06	C	T	missense_variant	pepA	protein_coding	c.1114G>A	p.Glu372Lys
C321.ΔA.2.39	5E+06	A	G	intergenic_region	intB-insC1		n.4500118A>G	
C321.ΔA.2.39	5E+06	G	C	missense_variant	insO	protein_coding	c.641G>C	p.Arg214Thr
C321.ΔA.2.39	5E+06	G	C	missense_variant	insI1	protein_coding	c.781C>G	p.Leu261Val
C321.ΔA.2.39	5E+06	T	C	synonymous_variant	insO	protein_coding	c.729T>C	p.Pro243Pro
C321.ΔA.2.39	5E+06	T	C	synonymous_variant	insI1	protein_coding	c.693A>G	p.Leu231Leu
C321.ΔA.2.39	5E+06	T	C	synonymous_variant	fimI	protein_coding	c.315T>C	p.Asp105Asp
C321.ΔA.2.39	5E+06	C	T	synonymous_variant	yjjI	protein_coding	c.744G>A	p.Thr248Thr
C321.ΔA.2.39	5E+06	C	T	synonymous_variant	yjjK	protein_coding	c.1461G>A	p.Leu487Leu
C321.ΔA.2.39	224265	A	AG	non_coding_transcript_exon_variant	rrsH	rRNA	n.671dupG	
C321.ΔA.2.39	379063	AG	A	intergenic_region	frmR-yaiO		n.379064delG	
C321.ΔA.2.39	582432	C	CT	intergenic_region	tfaX-appY		n.582432_582433insT	
C321.ΔA.2.39	583303	CA	C	frameshift_variant	appY	protein_coding	c.582delA	p.Lys194fs
C321.ΔA.2.39	722567	GT	G	frameshift_variant	kdpD	protein_coding	c.895delA	p.Thr299fs
C321.ΔA.2.39	738554	T	TATA	start_lost&disruptive_infr	phr	protein_coding	c.2_3insAAT	

				ame_insertion				
C321.ΔA.2.39	738554	T	TATA	splice_region_variant&disruptive_inframe_insertion	ybgA	protein_coding	c.508_509insAAT	p.Ter170dup
C321.ΔA.2.39	775893	GA	G	frameshift_variant	tolA	protein_coding	c.510delA	p.Ala171fs
C321.ΔA.2.39	970912	CT	C	intergenic_region	rpsA-ihfB		n.970913delT	
C321.ΔA.2.39	975469	C	CAAT	splice_region_variant&disruptive_inframe_insertion	msbA	protein_coding	c.1745_1747dupAAT	p.Ter583dup
C321.ΔA.2.39	1E+06	C	CATA	start_lost&disruptive_inframe_insertion	holB	protein_coding	c.2_3insAAT	
C321.ΔA.2.39	1E+06	C	CATA	splice_region_variant&disruptive_inframe_insertion	tmk	protein_coding	c.640_641insAAT	p.Ter214dup
C321.ΔA.2.39	1E+06	TA	T	intergenic_region	ymfJ-cohE		n.1209324delA	
C321.ΔA.2.39	1E+06	A	AG	frameshift_variant	ycgB	protein_coding	c.969dupC	p.Tyr324fs
C321.ΔA.2.39	1E+06	C	CAAT	splice_region_variant&disruptive_inframe_insertion	prmC	protein_coding	c.830_832dupAAT	p.Ter278dup
C321.ΔA.2.39	1E+06	GCA	G	frameshift_variant	stfR	protein_coding	c.3068_3069delAC	p.His1023fs
C321.ΔA.2.39	2E+06	G	GCC	frameshift_variant	yneO	protein_coding	c.3862_3863dupGG	p.Thr1289fs
C321.ΔA.2.39	2E+06	A	AT	frameshift_variant	ynfF	protein_coding	c.118dupT	p.Ser40fs
C321.ΔA.2.39	2E+06	A	AT	frameshift_varian	exoX	protein_coding	c.81_82insT	p.Lys28fs

				t&stop_gained				
C321.ΔA.2.39	2E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	yefM	protein_coding	c.250_251insAAT	p.Ter84dup
C321.ΔA.2.39	2E+06	C	CATT	disruptive_inframe_insertion	yoeB	protein_coding	c.2_3insAAT	p.Val1_Lys2insMet
C321.ΔA.2.39	2E+06	GC	G	frameshift_variant	cpsG	protein_coding	c.340delG	p.Ala114fs
C321.ΔA.2.39	2E+06	G	GC	frameshift_variant	yehA	protein_coding	c.377dupG	p.Gly127fs
C321.ΔA.2.39	2E+06	G	GAAT	splice_region_variant&disruptive_inframe_insertion	yohJ	protein_coding	c.395_397dupAAT	p.Ter133dup
C321.ΔA.2.39	2E+06	AC	A	frameshift_variant	yohK	protein_coding	c.35delC	p.Thr12fs
C321.ΔA.2.39	2E+06	C	CCTGAGCTTTCATCCTATGAAATTAATTGCTGTTAAAAGCATTGGGTA	frameshift_variant&stop_gained	yejO	protein_coding	c.3781_3782insTACCCAATGCTTTTAACAGCAATTAATTTCATAGGATGAAAGCTCAG	p.Gly1261fs
C321.ΔA.2.39	2E+06	A	AAAATACCCATAGCT	frameshift_variant	yejO	protein_coding	c.3777_3778insAGCTATGGGTATTT	p.Gly1261fs
C321.ΔA.2.39	3E+06	G	GC	frameshift_variant	hyfB	protein_coding	c.200dupC	p.Glu69fs
C321.ΔA.2.39	3E+06	GCC	G	frameshift_variant	focB	protein_coding	c.684_685delCC	p.Leu229fs
C321.ΔA.2.39	3E+06	CGCACTATG	C	intergenic_region	yfjL-yfjM		n.2769432_2769439delGCACTATG	

C321.ΔA.2.39	3E+06	G	GC	intergenic_region	yfjM-rnlA	n.2769807_2769808insC		
C321.ΔA.2.39	3E+06	C	CATT	start_lost&disruptive_inframe_insertion	truD	protein_coding	c.-1_2dupAAT	
C321.ΔA.2.39	3E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	ispF	protein_coding	c.476_478dupAAT	p.Ter160dup
C321.ΔA.2.39	3E+06	A	AG	frameshift_variant	ygeR	protein_coding	c.218dupC	p.Tyr74fs
C321.ΔA.2.39	3E+06	G	GCC	frameshift_variant	yggN	protein_coding	c.590_591dupGG	p.Leu198fs
C321.ΔA.2.39	3E+06	CCA	C	frameshift_variant	dnaG	protein_coding	c.1726_1727delCA	p.Gln576fs
C321.ΔA.2.39	3E+06	G	GCC	frameshift_variant	dnaG	protein_coding	c.1728_1729insCC	p.Glu577fs
C321.ΔA.2.39	3E+06	C	CATT	start_lost&disruptive_inframe_insertion	rplW	protein_coding	c.2_3insAAT	
C321.ΔA.2.39	3E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	rplD	protein_coding	c.604_605insAAT	p.Ter202dup
C321.ΔA.2.39	4E+06	G	GC	frameshift_variant	yhhS	protein_coding	c.1018dupG	p.Ala340fs
C321.ΔA.2.39	4E+06	T	TC	intergenic_region	bcsG-ldrD	n.3700603_3700604insC		
C321.ΔA.2.39	4E+06	CG	C	frameshift_variant	waaA	protein_coding	c.1272delG	p.His425fs
C321.ΔA.2.39	4E+06	C	CT	frameshift_variant&start_lost	dgoD	protein_coding	c.2_3insA	p.Met1fs

C321.ΔA.2.39	4E+06	C	CT	frameshift_variant&splice_region_variant	dgoA	protein_coding	c.616_617insA	
C321.ΔA.2.39	4E+06	ACG	A	frameshift_variant	dgoA	protein_coding	c.596_597delCG	p.Ala199fs
C321.ΔA.2.39	4E+06	C	CATT	start_lost&disruptive_inframe_insertion	dgoK	protein_coding	c.2_3insAAT	
C321.ΔA.2.39	4E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	dgoR	protein_coding	c.688_689insAAT	p.Ter230dup
C321.ΔA.2.39	4E+06	GA	G	frameshift_variant	dgoR	protein_coding	c.661delT	p.Ser221fs
C321.ΔA.2.39	4E+06	A	AG	non_coding_transcript_exon_variant	rrlC	rRNA	n.1426dupG	
C321.ΔA.2.39	4E+06	C	CG	frameshift_variant	proP	protein_coding	c.175dupG	p.Ala59fs
C321.ΔA.2.39	4E+06	C	CG	frameshift_variant	yjgL	protein_coding	c.126dupG	p.Gln43fs
C321.ΔA.2.39	5E+06	C	CA	intergenic_region	yjjM-yjjN	n.4597823_4597824insA		
C321.ΔA.2.39	5E+06	A	AG	frameshift_variant	leuP	protein_coding	c.79dupC	p.Leu27fs
C321.ΔA.2.39	5E+06	A	AG	frameshift_variant	leuQ	protein_coding	c.79dupC	p.Leu27fs
C321.ΔA.2.39	5E+06	T	TC	intergenic_region	yjjG-prfC	n.4611239_4611240insC		
C321.ΔA.2.109.F	495240	C	T	stop_gained	htpG	protein_coding	c.1072C>T	p.Gln358*
C321.ΔA.2.109.F	500836	C	T	stop_gained	ybaL	protein_coding	c.1452G>A	p.Trp484*
C321.ΔA.2.109.F	819327	C	T	stop_gained	bioC	protein_coding	c.700C>T	p.Gln234*

C321.ΔA.2.109.F	2E+06	A	G	stop_lost	ydfJ	protein_coding	c.4084T>C	p.Ter1362Glnext*?
C321.ΔA.2.109.F	2E+06	G	A	stop_gained	bglX	protein_coding	c.1501C>T	p.Gln501*
C321.ΔA.2.109.F	4E+06	G	A	stop_gained	ampC	protein_coding	c.67C>T	p.Gln23*
C321.ΔA.2.109.F	5323	G	A	synonymous_variant	yaaX	protein_coding	c.90G>A	p.Pro30Pro
C321.ΔA.2.109.F	25207	G	A	splice_region_variant&stop_retained_variant	ileS	protein_coding	c.2816G>A	p.Ter939Ter
C321.ΔA.2.109.F	44129	G	A	splice_region_variant&stop_retained_variant	fixB	protein_coding	c.941G>A	p.Ter314Ter
C321.ΔA.2.109.F	45750	G	A	splice_region_variant&stop_retained_variant	fixX	protein_coding	c.287G>A	p.Ter96Ter
C321.ΔA.2.109.F	54757	C	T	splice_region_variant&stop_retained_variant	lptD	protein_coding	c.2354G>A	p.Ter785Ter
C321.ΔA.2.109.F	58179	G	A	splice_region_variant&stop_retained_variant	djlA	protein_coding	c.815G>A	p.Ter272Ter
C321.ΔA.2.109.F	63431	C	T	splice_region_variant&stop_retained_variant	polB	protein_coding	c.2351G>A	p.Ter784Ter
C321.ΔA.2.109.F	72743	C	T	splice_region_variant&stop_retained_variant	thiP	protein_coding	c.1610G>A	p.Ter537Ter

C321.ΔA.2.109.F	85664	C	T	synonymous_variant	ilvI	protein_coding	c.204C>T	p.Val68Val
C321.ΔA.2.109.F	88188	T	C	synonymous_variant	cra	protein_coding	c.330T>C	p.Ile110Ile
C321.ΔA.2.109.F	98233	G	A	splice_region_variant&stop_retained_variant	murD	protein_coding	c.1316G>A	p.Ter439Ter
C321.ΔA.2.109.F	102070	G	A	splice_region_variant&stop_retained_variant	murC	protein_coding	c.1475G>A	p.Ter492Ter
C321.ΔA.2.109.F	116886	T	C	synonymous_variant	hofB	protein_coding	c.45A>G	p.Gly15Gly
C321.ΔA.2.109.F	156361	A	G	synonymous_variant	yadN	protein_coding	c.354T>C	p.Ser118Ser
C321.ΔA.2.109.F	160814	C	T	synonymous_variant	sfsA	protein_coding	c.504G>A	p.Ala168Ala
C321.ΔA.2.109.F	173434	C	T	splice_region_variant&stop_retained_variant	hemL	protein_coding	c.1280G>A	p.Ter427Ter
C321.ΔA.2.109.F	185809	C	T	splice_region_variant&stop_retained_variant	glnD	protein_coding	c.2672G>A	p.Ter891Ter
C321.ΔA.2.109.F	193258	G	A	splice_region_variant&stop_retained_variant	frr	protein_coding	c.557G>A	p.Ter186Ter
C321.ΔA.2.109.F	194546	G	A	splice_region_variant&stop_retained_variant	dxr	protein_coding	c.1196G>A	p.Ter399Ter

C321.ΔA.2.109.F	195493	G	A	splice_region_variant&stop_retained_variant	ispU	protein_coding	c.761G>A	p.Ter254Ter
C321.ΔA.2.109.F	197727	G	A	splice_region_variant&stop_retained_variant	rseP	protein_coding	c.1352G>A	p.Ter451Ter
C321.ΔA.2.109.F	202385	G	A	splice_region_variant&stop_retained_variant	fabZ	protein_coding	c.455G>A	p.Ter152Ter
C321.ΔA.2.109.F	213756	C	T	splice_region_variant&stop_retained_variant	yaeP	protein_coding	c.200G>A	p.Ter67Ter
C321.ΔA.2.109.F	216888	C	T	splice_region_variant&stop_retained_variant	proS	protein_coding	c.1718G>A	p.Ter573Ter
C321.ΔA.2.109.F	221945	A	G	synonymous_variant	metN	protein_coding	c.531T>C	p.Arg177Arg
C321.ΔA.2.109.F	223237	G	A	splice_region_variant&stop_retained_variant	gmhB	protein_coding	c.575G>A	p.Ter192Ter
C321.ΔA.2.109.F	322153	C	T	synonymous_variant	ykgF	protein_coding	c.765C>T	p.Ile255Ile
C321.ΔA.2.109.F	402649	G	A	splice_region_variant&stop_retained_variant	psiF	protein_coding	c.320G>A	p.Ter107Ter
C321.ΔA.2.109.F	406218	G	A	splice_region_variant&stop_retain	yaiA	protein_coding	c.191G>A	p.Ter64Ter

				ed_variant				
C321.ΔA.2.109.F	413600	C	T	synonymous_variant	sbcC	protein_coding	c.1203G>A	p.Ala401Ala
C321.ΔA.2.109.F	428542	G	A	splice_region_variant&stop_retained_variant	secD	protein_coding	c.1847G>A	p.Ter616Ter
C321.ΔA.2.109.F	433606	G	A	splice_region_variant&stop_retained_variant	ribD	protein_coding	c.1103G>A	p.Ter368Ter
C321.ΔA.2.109.F	434165	G	A	splice_region_variant&stop_retained_variant	ribE	protein_coding	c.470G>A	p.Ter157Ter
C321.ΔA.2.109.F	473260	C	T	synonymous_variant	amtB	protein_coding	c.1246C>T	p.Leu416Leu
C321.ΔA.2.109.F	493071	G	A	splice_region_variant&stop_retained_variant	dnaX	protein_coding	c.1931G>A	p.Ter644Ter
C321.ΔA.2.109.F	611931	A	G	synonymous_variant	fes	protein_coding	c.69A>G	p.Val23Val
C321.ΔA.2.109.F	622874	A	G	synonymous_variant	fepB	protein_coding	c.685T>C	p.Leu229Leu
C321.ΔA.2.109.F	623295	C	T	synonymous_variant	fepB	protein_coding	c.264G>A	p.Gln88Gln
C321.ΔA.2.109.F	668980	C	T	splice_region_variant&stop_retained_variant	nadD	protein_coding	c.641G>A	p.Ter214Ter
C321.ΔA.2.109.F	669623	C	T	splice_region_variant&stop_retain	holA	protein_coding	c.1031G>A	p.Ter344Ter

				ed_variant				
C321.ΔA.2.109.F	670654	C	T	splice_region_variant&stop_retained_variant	lptE	protein_coding	c.581G>A	p.Ter194Ter
C321.ΔA.2.109.F	696926	C	A	synonymous_variant	asnB	protein_coding	c.1299G>T	p.Leu433Leu
C321.ΔA.2.109.F	709983	C	T	splice_region_variant&stop_retained_variant	fldA	protein_coding	c.530G>A	p.Ter177Ter
C321.ΔA.2.109.F	723403	C	T	synonymous_variant	kdpD	protein_coding	c.60G>A	p.Pro20Pro
C321.ΔA.2.109.F	743296	G	A	splice_region_variant&stop_retained_variant	ybgJ	protein_coding	c.656G>A	p.Ter219Ter
C321.ΔA.2.109.F	743296	G	A	synonymous_variant	ybgK	protein_coding	c.6G>A	p.Leu2Leu
C321.ΔA.2.109.F	744222	G	A	splice_region_variant&stop_retained_variant	ybgK	protein_coding	c.932G>A	p.Ter311Ter
C321.ΔA.2.109.F	754954	G	A	splice_region_variant&stop_retained_variant	sdhD	protein_coding	c.347G>A	p.Ter116Ter
C321.ΔA.2.109.F	773356	G	A	splice_region_variant&stop_retained_variant	ybgT	protein_coding	c.113G>A	p.Ter38Ter
C321.ΔA.2.109.F	778079	G	A	splice_region_variant&stop_retain	tolB	protein_coding	c.1292G>A	p.Ter431Ter

				ed_variant				
C321.ΔA.2.109.F	816550	A	G	synonymous_variant	bla	protein_coding	c.138A>G	p.Glu46Glu
C321.ΔA.2.109.F	816886	T	C	synonymous_variant	bla	protein_coding	c.474T>C	p.Thr158Thr
C321.ΔA.2.109.F	817129	G	A	synonymous_variant	bla	protein_coding	c.717G>A	p.Gly239Gly
C321.ΔA.2.109.F	820646	C	T	synonymous_variant	uvrB	protein_coding	c.15C>T	p.Phe5Phe
C321.ΔA.2.109.F	823721	A	G	synonymous_variant	ybhK	protein_coding	c.33T>C	p.Arg11Arg
C321.ΔA.2.109.F	907034	C	T	synonymous_variant	artJ	protein_coding	c.648G>A	p.Leu216Leu
C321.ΔA.2.109.F	933332	C	T	splice_region_variant&stop_retained_variant	infA	protein_coding	c.218G>A	p.Ter73Ter
C321.ΔA.2.109.F	940119	C	T	synonymous_variant	lrp	protein_coding	c.420C>T	p.Gly140Gly
C321.ΔA.2.109.F	982745	G	A	splice_region_variant&stop_retained_variant	mukF	protein_coding	c.1322G>A	p.Ter441Ter
C321.ΔA.2.109.F	983284	C	T	synonymous_variant	mukE	protein_coding	c.558C>T	p.Arg186Arg
C321.ΔA.2.109.F	1E+06	C	T	splice_region_variant&stop_retained_variant	fabA	protein_coding	c.518G>A	p.Ter173Ter
C321.ΔA.2.109.F	1E+06	G	A	synonymous_variant	ycbZ	protein_coding	c.915C>T	p.Gly305Gly
C321.ΔA.2.109.F	1E+06	G	A	splice_region_variant&stop_retain	yceQ	protein_coding	c.320G>A	p.Ter107Ter

				ed_variant				
C321.ΔA.2.109.F	1E+06	G	A	splice_region_variant&stop_retained_variant	fabG	protein_coding	c.734G>A	p.Ter245Ter
C321.ΔA.2.109.F	1E+06	G	A	synonymous_variant	intE	protein_coding	c.543C>T	p.Ala181Ala
C321.ΔA.2.109.F	1E+06	A	G	synonymous_variant	dhaR	protein_coding	c.789A>G	p.Lys263Lys
C321.ΔA.2.109.F	1E+06	A	G	splice_region_variant&stop_retained_variant	hemA	protein_coding	c.1257A>G	p.Ter419Ter
C321.ΔA.2.109.F	1E+06	C	T	synonymous_variant	rssB	protein_coding	c.726C>T	p.Gly242Gly
C321.ΔA.2.109.F	1E+06	C	T	synonymous_variant	puuC	protein_coding	c.1395C>T	p.Gly465Gly
C321.ΔA.2.109.F	1E+06	C	T	synonymous_variant	ydbK	protein_coding	c.1170G>A	p.Ala390Ala
C321.ΔA.2.109.F	1E+06	T	C	synonymous_variant	tynA	protein_coding	c.702A>G	p.Glu234Glu
C321.ΔA.2.109.F	1E+06	A	G	synonymous_variant	ydbD	protein_coding	c.1239A>G	p.Gln413Gln
C321.ΔA.2.109.F	2E+06	A	G	synonymous_variant	ydcU	protein_coding	c.756A>G	p.Thr252Thr
C321.ΔA.2.109.F	2E+06	G	A	synonymous_variant	fdnG	protein_coding	c.2418G>A	p.Thr806Thr
C321.ΔA.2.109.F	2E+06	A	C	synonymous_variant	ydeT	protein_coding	c.1191T>G	p.Gly397Gly
C321.ΔA.2.109.F	2E+06	G	A	splice_region_variant&stop_retained_variant	ydfB	protein_coding	c.128G>A	p.Ter43Ter

C321.ΔA.2.109.F	2E+06	G	A	synonymous_variant	ydfJ	protein_coding	c.4047C>T	p.Ile1349Ile
C321.ΔA.2.109.F	2E+06	G	A	synonymous_variant	ydfJ	protein_coding	c.1551C>T	p.Phe517Phe
C321.ΔA.2.109.F	2E+06	G	A	synonymous_variant	insD1	protein_coding	c.444G>A	p.Pro148Pro
C321.ΔA.2.109.F	2E+06	C	T	splice_region_variant&stop_retained_variant	ribC	protein_coding	c.641G>A	p.Ter214Ter
C321.ΔA.2.109.F	2E+06	T	C	synonymous_variant	ppsR	protein_coding	c.663T>C	p.Arg221Arg
C321.ΔA.2.109.F	2E+06	C	T	splice_region_variant&stop_retained_variant	pheT	protein_coding	c.2387G>A	p.Ter796Ter
C321.ΔA.2.109.F	2E+06	G	A	synonymous_variant	yeaM	protein_coding	c.534C>T	p.Gly178Gly
C321.ΔA.2.109.F	2E+06	T	C	synonymous_variant	fadD	protein_coding	c.1050A>G	p.Gly350Gly
C321.ΔA.2.109.F	2E+06	C	T	splice_region_variant&stop_retained_variant	tsaB	protein_coding	c.695G>A	p.Ter232Ter
C321.ΔA.2.109.F	2E+06	G	A	synonymous_variant	yobB	protein_coding	c.471G>A	p.Ser157Ser
C321.ΔA.2.109.F	2E+06	C	T	splice_region_variant&stop_retained_variant	aspS	protein_coding	c.1772G>A	p.Ter591Ter
C321.ΔA.2.109.F	2E+06	C	T	splice_region_variant&stop_retained_variant	pgsA	protein_coding	c.548G>A	p.Ter183Ter

C321.ΔA.2.109.F	2E+06	G	A	synonymous_variant	fliK	protein_coding	c.15G>A	p.Ala5Ala
C321.ΔA.2.109.F	2E+06	C	T	synonymous_variant	wcaD	protein_coding	c.165G>A	p.Leu55Leu
C321.ΔA.2.109.F	2E+06	A	G	synonymous_variant	yegL	protein_coding	c.42T>C	p.Ser14Ser
C321.ΔA.2.109.F	2E+06	C	T	splice_region_variant&stop_retained_variant	folE	protein_coding	c.668G>A	p.Ter223Ter
C321.ΔA.2.109.F	2E+06	C	T	synonymous_variant	fruA	protein_coding	c.1353G>A	p.Ala451Ala
C321.ΔA.2.109.F	2E+06	G	A	splice_region_variant&stop_retained_variant	yejF	protein_coding	c.1589G>A	p.Ter530Ter
C321.ΔA.2.109.F	2E+06	A	G	synonymous_variant	bcr	protein_coding	c.1005T>C	p.Phe335Phe
C321.ΔA.2.109.F	2E+06	G	A	splice_region_variant&stop_retained_variant	yejM	protein_coding	c.1760G>A	p.Ter587Ter
C321.ΔA.2.109.F	2E+06	G	A	splice_region_variant&stop_retained_variant	nrdA	protein_coding	c.2285G>A	p.Ter762Ter
C321.ΔA.2.109.F	2E+06	G	A	splice_region_variant&stop_retained_variant	nrdB	protein_coding	c.1130G>A	p.Ter377Ter
C321.ΔA.2.109.F	2E+06	C	T	synonymous_variant	pta	protein_coding	c.1779C>T	p.Thr593Thr
C321.ΔA.2.109.F	2E+06	C	T	splice_region_variant&stop_retain	accD	protein_coding	c.914G>A	p.Ter305Ter

				ed_variant				
C321.ΔA.2.109.F	3E+06	C	T	splice_region_variant&stop_retained_variant	ligA	protein_coding	c.2015G>A	p.Ter672Ter
C321.ΔA.2.109.F	3E+06	C	T	splice_region_variant&stop_retained_variant	zipA	protein_coding	c.986G>A	p.Ter329Ter
C321.ΔA.2.109.F	3E+06	C	T	splice_region_variant&stop_retained_variant	rpoE	protein_coding	c.575G>A	p.Ter192Ter
C321.ΔA.2.109.F	3E+06	G	A	synonymous_variant	intA	protein_coding	c.399G>A	p.Ser133Ser
C321.ΔA.2.109.F	3E+06	G	A	synonymous_variant	gabD	protein_coding	c.1347G>A	p.Val449Val
C321.ΔA.2.109.F	3E+06	T	C	synonymous_variant	proV	protein_coding	c.198T>C	p.Gly66Gly
C321.ΔA.2.109.F	3E+06	G	A	synonymous_variant	yqaA	protein_coding	c.81C>T	p.Val27Val
C321.ΔA.2.109.F	3E+06	G	C	synonymous_variant	alaS	protein_coding	c.588C>G	p.Gly196Gly
C321.ΔA.2.109.F	3E+06	C	T	synonymous_variant	cysI	protein_coding	c.1665G>A	p.Ala555Ala
C321.ΔA.2.109.F	3E+06	C	T	splice_region_variant&stop_retained_variant	lgt	protein_coding	c.875G>A	p.Ter292Ter
C321.ΔA.2.109.F	3E+06	G	A	synonymous_variant	xerD	protein_coding	c.483C>T	p.Ser161Ser
C321.ΔA.2.109.F	3E+06	A	G	synonymous_variant	nudF	protein_coding	c.465T>C	p.Gly155Gly
C321.ΔA.2.109.F	3E+06	C	T	splice_region_var	ribB	protein_coding	c.653G>A	p.Ter218Ter

				iant&stop_retained_variant				
C321.ΔA.2.109.F	3E+06	G	A	splice_region_variant&stop_retained_variant	cca	protein_coding	c.1238G>A	p.Ter413Ter
C321.ΔA.2.109.F	3E+06	T	C	synonymous_variant	ebgA	protein_coding	c.2967T>C	p.Arg989Arg
C321.ΔA.2.109.F	3E+06	C	T	synonymous_variant	fadH	protein_coding	c.558C>T	p.Gly186Gly
C321.ΔA.2.109.F	3E+06	G	A	splice_region_variant&stop_retained_variant	fmt	protein_coding	c.947G>A	p.Ter316Ter
C321.ΔA.2.109.F	3E+06	C	T	splice_region_variant&stop_retained_variant	rplV	protein_coding	c.332G>A	p.Ter111Ter
C321.ΔA.2.109.F	3E+06	C	T	splice_region_variant&stop_retained_variant	rpsG	protein_coding	c.539G>A	p.Ter180Ter
C321.ΔA.2.109.F	3E+06	G	A	synonymous_variant	cysG	protein_coding	c.1275G>A	p.Gln425Gln
C321.ΔA.2.109.F	4E+06	C	T	synonymous_variant	yhiN	protein_coding	c.1059G>A	p.Thr353Thr
C321.ΔA.2.109.F	4E+06	T	C	synonymous_variant	pitA	protein_coding	c.567T>C	p.Arg189Arg
C321.ΔA.2.109.F	4E+06	C	T	synonymous_variant	xylG	protein_coding	c.1275C>T	p.Arg425Arg
C321.ΔA.2.109.F	4E+06	G	A	synonymous_variant	yiaS	protein_coding	c.93G>A	p.Ala31Ala
C321.ΔA.2.109.F	4E+06	C	T	splice_region_variant&sto	waaU	protein_coding	c.1073G>A	p.Ter358Ter

				p_retained_variant				
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	waaA	protein_coding	c.1277G>A	p.Ter426Ter
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	gmk	protein_coding	c.623G>A	p.Ter208Ter
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	rnpA	protein_coding	c.359G>A	p.Ter120Ter
C321.ΔA.2.109.F	4E+06	G	A	synonymous_variant	yidD	protein_coding	c.36G>A	p.Leu12Leu
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	yidC	protein_coding	c.1646G>A	p.Ter549Ter
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	tnaC	protein_coding	c.74G>A	p.Ter25Ter
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	yieE	protein_coding	c.749G>A	p.Ter250Ter
C321.ΔA.2.109.F	4E+06	C	T	splice_region_variant&stop_retained_variant	bglG	protein_coding	c.836G>A	p.Ter279Ter
C321.ΔA.2.109.F	4E+06	C	T	splice_region_variant&stop_retain	pstB	protein_coding	c.773G>A	p.Ter258Ter

				ed_variant				
C321.ΔA.2.109.F	4E+06	C	T	splice_region_variant&stop_retained_variant	pstA	protein_coding	c.890G>A	p.Ter297Ter
C321.ΔA.2.109.F	4E+06	C	T	splice_region_variant&stop_retained_variant	glmU	protein_coding	c.1370G>A	p.Ter457Ter
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	rbsD	protein_coding	c.419G>A	p.Ter140Ter
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	rbsK	protein_coding	c.929G>A	p.Ter310Ter
C321.ΔA.2.109.F	4E+06	C	T	splice_region_variant&stop_retained_variant	yieP	protein_coding	c.692G>A	p.Ter231Ter
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	ilvM	protein_coding	c.263G>A	p.Ter88Ter
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	ubiB	protein_coding	c.1640G>A	p.Ter547Ter
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	ubiD	protein_coding	c.1493G>A	p.Ter498Ter

C321.ΔA.2.109.F	4E+06	T	C	synonymous_variant	pepQ	protein_coding	c.618T>C	p.Gly206Gly
C321.ΔA.2.109.F	4E+06	C	T	splice_region_variant&stop_retained_variant	ftsN	protein_coding	c.959G>A	p.Ter320Ter
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	murI	protein_coding	c.857G>A	p.Ter286Ter
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	secE	protein_coding	c.383G>A	p.Ter128Ter
C321.ΔA.2.109.F	4E+06	G	A	synonymous_variant	metA	protein_coding	c.540G>A	p.Thr180Thr
C321.ΔA.2.109.F	4E+06	G	A	synonymous_variant	aceB	protein_coding	c.735G>A	p.Ala245Ala
C321.ΔA.2.109.F	4E+06	C	T	synonymous_variant	metH	protein_coding	c.2950C>T	p.Leu984Leu
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	ubiA	protein_coding	c.872G>A	p.Ter291Ter
C321.ΔA.2.109.F	4E+06	A	G	synonymous_variant	dinF	protein_coding	c.453A>G	p.Glu151Glu
C321.ΔA.2.109.F	4E+06	G	A	splice_region_variant&stop_retained_variant	ssb	protein_coding	c.536G>A	p.Ter179Ter
C321.ΔA.2.109.F	4E+06	C	T	splice_region_variant&stop_retained_variant	alsK	protein_coding	c.929G>A	p.Ter310Ter

C321.ΔA.2.109.F	4E+06	C	T	synonymous_variant	yjdF	protein_coding	c.369G>A	p.Gly123Gly
C321.ΔA.2.109.F	4E+06	A	G	synonymous_variant	cadA	protein_coding	c.1866T>C	p.Ala622Ala
C321.ΔA.2.109.F	4E+06	A	G	synonymous_variant	cadC	protein_coding	c.1488T>C	p.Thr496Thr
C321.ΔA.2.109.F	5E+06	C	T	synonymous_variant	opgB	protein_coding	c.315G>A	p.Pro105Pro
C321.ΔA.2.109.F	5E+06	C	T	synonymous_variant	yjjI	protein_coding	c.744G>A	p.Thr248Thr
C321.ΔA.2.109.F	64263	C	T	missense_variant	polB	protein_coding	c.1519G>A	p.Ala507Thr
C321.ΔA.2.109.F	72743	C	T	missense_variant	thiQ	protein_coding	c.16G>A	p.Asp6Asn
C321.ΔA.2.109.F	79908	T	C	missense_variant	leuC	protein_coding	c.788A>G	p.Lys263Arg
C321.ΔA.2.109.F	102070	G	A	missense_variant	ddlB	protein_coding	c.7G>A	p.Asp3Asn
C321.ΔA.2.109.F	133742	T	A	missense_variant	acnB	protein_coding	c.2297T>A	p.Ile766Asn
C321.ΔA.2.109.F	161592	C	T	missense_variant	ligT	protein_coding	c.271G>A	p.Gly91Arg
C321.ΔA.2.109.F	171836	A	G	missense_variant	fhuB	protein_coding	c.544A>G	p.Thr182Ala
C321.ΔA.2.109.F	213756	C	T	missense_variant	rof	protein_coding	c.7G>A	p.Asp3Asn
C321.ΔA.2.109.F	253368	T	C	missense_variant	ykfJ	protein_coding	c.61T>C	p.Trp21Arg
C321.ΔA.2.109.F	257670	A	G	missense_variant	crl	protein_coding	c.13A>G	p.Ser5Gly
C321.ΔA.2.109.F	264888	T	C	missense_variant	ykfB	protein_coding	c.253A>G	p.Thr85Ala
C321.ΔA.2.109.F	295866	C	T	missense_variant	intF	protein_coding	c.283G>A	p.Asp95Asn

C321.ΔA.2.109.F	296893	A	G	missense_variant	paoD	protein_coding	c.886T>C	p.Ser296Pro
C321.ΔA.2.109.F	299130	A	G	missense_variant	paoC	protein_coding	c.857T>C	p.Leu286Pro
C321.ΔA.2.109.F	336301	T	C	missense_variant	yahF	protein_coding	c.473T>C	p.Leu158Pro
C321.ΔA.2.109.F	369824	A	G	missense_variant	mhpB	protein_coding	c.497A>G	p.Asn166Ser
C321.ΔA.2.109.F	411506	C	T	missense_variant	araJ	protein_coding	c.25G>A	p.Ala9Thr
C321.ΔA.2.109.F	438536	C	T	missense_variant	dxs	protein_coding	c.691G>A	p.Glu231Lys
C321.ΔA.2.109.F	443297	T	C	missense_variant	panE	protein_coding	c.268A>G	p.Thr90Ala
C321.ΔA.2.109.F	474984	C	T	missense_variant	ybaY	protein_coding	c.557C>T	p.Ala186Val
C321.ΔA.2.109.F	486909	T	C	missense_variant	mscK	protein_coding	c.1325T>C	p.Ile442Thr
C321.ΔA.2.109.F	499472	A	G	missense_variant	gsk	protein_coding	c.299A>G	p.His100Arg
C321.ΔA.2.109.F	669665	A	G	missense_variant	holA	protein_coding	c.989T>C	p.Leu330Pro
C321.ΔA.2.109.F	679212	A	G	missense_variant	ybeU	protein_coding	c.658A>G	p.Thr220Ala
C321.ΔA.2.109.F	722801	A	G	missense_variant	kdpD	protein_coding	c.662T>C	p.Leu221Pro
C321.ΔA.2.109.F	744222	G	A	missense_variant	ybgL	protein_coding	c.10G>A	p.Asp4Asn
C321.ΔA.2.109.F	757053	A	G	missense_variant	sdhB	protein_coding	c.317A>G	p.Asp106Gly
C321.ΔA.2.109.F	771998	G	A	missense_variant	cydA	protein_coding	c.1493G>A	p.Arg498His
C321.ΔA.2.109.F	774870	G	A	missense_variant	tolQ	protein_coding	c.670G>A	p.Val224Ile

C321.ΔA.2.109.F	775990	G	A	missense_variant	tolA	protein_coding	c.601G>A	p.Glu201Lys
C321.ΔA.2.109.F	816521	C	A	missense_variant	bla	protein_coding	c.109C>A	p.Gln37Lys
C321.ΔA.2.109.F	816957	T	C	missense_variant	bla	protein_coding	c.545T>C	p.Val182Ala
C321.ΔA.2.109.F	848553	C	T	missense_variant	fiu	protein_coding	c.85G>A	p.Ala29Thr
C321.ΔA.2.109.F	916196	G	A	missense_variant	ltaE	protein_coding	c.205C>T	p.His69Tyr
C321.ΔA.2.109.F	978135	C	T	missense_variant	kdsB	protein_coding	c.179C>T	p.Ala60Val
C321.ΔA.2.109.F	982745	G	A	missense_variant	mukE	protein_coding	c.19G>A	p.Glu7Lys
C321.ΔA.2.109.F	1E+06	C	T	missense_variant	agp	protein_coding	c.170C>T	p.Thr57Met
C321.ΔA.2.109.F	1E+06	C	T	missense_variant	rutA	protein_coding	c.1138G>A	p.Glu380Lys
C321.ΔA.2.109.F	1E+06	C	A	missense_variant	ymfE	protein_coding	c.443G>T	p.Trp148Leu
C321.ΔA.2.109.F	1E+06	C	T	missense_variant	hemA	protein_coding	c.587C>T	p.Pro196Leu
C321.ΔA.2.109.F	1E+06	C	T	missense_variant	chaA	protein_coding	c.233G>A	p.Ser78Asn
C321.ΔA.2.109.F	1E+06	A	G	missense_variant	puuA	protein_coding	c.457T>C	p.Phe153Leu
C321.ΔA.2.109.F	1E+06	A	G	missense_variant	uspE	protein_coding	c.152T>C	p.Leu51Pro
C321.ΔA.2.109.F	1E+06	T	C	missense_variant	trkG	protein_coding	c.565T>C	p.Trp189Arg
C321.ΔA.2.109.F	1E+06	G	T	missense_variant	hrpA	protein_coding	c.3136G>T	p.Asp1046Tyr
C321.ΔA.2.109.F	2E+06	T	C	missense_variant	narZ	protein_coding	c.1838A>G	p.Asn613Ser

C321.ΔA.2.109.F	2E+06	T	C	missense_variant	ydfJ	protein_coding	c.5248A>G	p.Ile1750Val
C321.ΔA.2.109.F	2E+06	T	C	missense_variant	ydfX	protein_coding	c.46A>G	p.Ile16Val
C321.ΔA.2.109.F	2E+06	A	G	missense_variant	ydfB	protein_coding	c.91A>G	p.Lys31Glu
C321.ΔA.2.109.F	2E+06	G	A	missense_variant	ydiU	protein_coding	c.982C>T	p.Arg328Trp
C321.ΔA.2.109.F	2E+06	T	C	missense_variant	pheS	protein_coding	c.124A>G	p.Met42Val
C321.ΔA.2.109.F	2E+06	T	C	missense_variant	astC	protein_coding	c.45A>G	p.Ile15Met
C321.ΔA.2.109.F	2E+06	A	G	missense_variant	ydjK	protein_coding	c.662T>C	p.Val221Ala
C321.ΔA.2.109.F	2E+06	A	T	missense_variant	exoX	protein_coding	c.84A>T	p.Lys28Asn
C321.ΔA.2.109.F	2E+06	A	G	missense_variant	edd	protein_coding	c.1328T>C	p.Val443Ala
C321.ΔA.2.109.F	2E+06	T	C	missense_variant	araG	protein_coding	c.1130A>G	p.Asn377Ser
C321.ΔA.2.109.F	2E+06	A	G	missense_variant	yeeZ	protein_coding	c.340T>C	p.Ser114Pro
C321.ΔA.2.109.F	2E+06	G	A	missense_variant	yegD	protein_coding	c.1186G>A	p.Ala396Thr
C321.ΔA.2.109.F	2E+06	T	C	missense_variant	cdd	protein_coding	c.872T>C	p.Val291Ala
C321.ΔA.2.109.F	2E+06	A	G	missense_variant	fruA	protein_coding	c.1349T>C	p.Val450Ala
C321.ΔA.2.109.F	2E+06	T	C	missense_variant	gyrA	protein_coding	c.1667A>G	p.Glu556Gly
C321.ΔA.2.109.F	2E+06	G	A	missense_variant	emrK	protein_coding	c.626C>T	p.Thr209Met
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	ligA	protein_coding	c.1321A>G	p.Lys441Glu

C321.ΔA.2.109.F	3E+06	T	C	missense_variant	hyfC	protein_coding	c.214T>C	p.Phe72Leu
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	rlmN	protein_coding	c.164A>G	p.Asp55Gly
C321.ΔA.2.109.F	3E+06	G	A	missense_variant	hcaE	protein_coding	c.491G>A	p.Arg164His
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	norW	protein_coding	c.391T>C	p.Tyr131His
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	ascG	protein_coding	c.460A>G	p.Asn154Asp
C321.ΔA.2.109.F	3E+06	A	G	missense_variant	hycC	protein_coding	c.902T>C	p.Leu301Pro
C321.ΔA.2.109.F	3E+06	G	A	missense_variant	hypE	protein_coding	c.241G>A	p.Ala81Thr
C321.ΔA.2.109.F	3E+06	A	G	missense_variant	ygdG	protein_coding	c.532A>G	p.Lys178Glu
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	amiC	protein_coding	c.14A>G	p.Asn5Ser
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	recB	protein_coding	c.1217A>G	p.Glu406Gly
C321.ΔA.2.109.F	3E+06	G	A	missense_variant	uacT	protein_coding	c.962G>A	p.Gly321Asp
C321.ΔA.2.109.F	3E+06	T	G	missense_variant	mutY	protein_coding	c.780T>G	p.Cys260Trp
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	glcB	protein_coding	c.1009A>G	p.Ile337Val
C321.ΔA.2.109.F	3E+06	G	A	missense_variant	fadH	protein_coding	c.203G>A	p.Gly68Asp
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	tdcE	protein_coding	c.1261A>G	p.Ile421Val
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	mtr	protein_coding	c.628A>G	p.Asn210Asp
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	folP	protein_coding	c.652A>G	p.Met218Val

C321.ΔA.2.109.F	3E+06	A	G	missense_variant	degS	protein_coding	c.545A>G	p.Asn182Ser
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	tldD	protein_coding	c.946A>G	p.Thr316Ala
C321.ΔA.2.109.F	4E+06	T	C	missense_variant	gph	protein_coding	c.328A>G	p.Thr110Ala
C321.ΔA.2.109.F	4E+06	T	C	missense_variant	nikC	protein_coding	c.748T>C	p.Phe250Leu
C321.ΔA.2.109.F	4E+06	C	G	missense_variant	yhhI	protein_coding	c.675C>G	p.Ile225Met
C321.ΔA.2.109.F	4E+06	C	T	missense_variant	yhjD	protein_coding	c.365C>T	p.Thr122Met
C321.ΔA.2.109.F	4E+06	T	C	missense_variant	eptB	protein_coding	c.1115A>G	p.Tyr372Cys
C321.ΔA.2.109.F	4E+06	T	C	missense_variant	wecH	protein_coding	c.104T>C	p.Val35Ala
C321.ΔA.2.109.F	4E+06	C	T	missense_variant	yiaK	protein_coding	c.437C>T	p.Pro146Leu
C321.ΔA.2.109.F	4E+06	C	A	missense_variant	hldD	protein_coding	c.612C>A	p.Ser204Arg
C321.ΔA.2.109.F	4E+06	T	C	missense_variant	glnG	protein_coding	c.32A>G	p.Asp11Gly
C321.ΔA.2.109.F	4E+06	T	C	missense_variant	ptsA	protein_coding	c.811A>G	p.Met271Val
C321.ΔA.2.109.F	4E+06	G	A	missense_variant	hemE	protein_coding	c.255G>A	p.Met85Ile
C321.ΔA.2.109.F	4E+06	A	G	missense_variant	yjaB	protein_coding	c.430T>C	p.Tyr144His
C321.ΔA.2.109.F	4E+06	A	G	missense_variant	metH	protein_coding	c.2620A>G	p.Thr874Ala
C321.ΔA.2.109.F	4E+06	C	T	missense_variant	lamB	protein_coding	c.725C>T	p.Ala242Val
C321.ΔA.2.109.F	4E+06	C	T	missense_variant	mdtO	protein_coding	c.875G>A	p.Cys292Tyr

C321.ΔA.2.109.F	4E+06	C	T	missense_variant	yjfM	protein_coding	c.430C>T	p.Arg144Cys
C321.ΔA.2.109.F	4E+06	A	G	missense_variant	aidB	protein_coding	c.1523A>G	p.Gln508Arg
C321.ΔA.2.109.F	4E+06	A	G	missense_variant	ytfE	protein_coding	c.530T>C	p.Val177Ala
C321.ΔA.2.109.F	4E+06	G	A	missense_variant	ytfL	protein_coding	c.569C>T	p.Ala190Val
C321.ΔA.2.109.F	4E+06	T	C	missense_variant	pyrB	protein_coding	c.767A>G	p.His256Arg
C321.ΔA.2.109.F	5E+06	C	T	missense_variant	insA	protein_coding	c.106C>T	p.Arg36Cys
C321.ΔA.2.109.F	5E+06	T	C	missense_variant	fimB	protein_coding	c.266T>C	p.Leu89Ser
C321.ΔA.2.109.F	311939	G	A	intergenic_region	ykgM-ykgR		n.311939G>A	
C321.ΔA.2.109.F	362284	A	G	intragenic_variant	lacZ		n.362284A>G	
C321.ΔA.2.109.F	503766	G	A	intergenic_region	fsr-ushA		n.503766G>A	
C321.ΔA.2.109.F	740043	T	C	intergenic_region	phr-dtpD		n.740043T>C	
C321.ΔA.2.109.F	780109	C	T	intergenic_region	lysW-valZ		n.780109C>T	
C321.ΔA.2.109.F	1E+06	C	T	intergenic_region	icd-ymfD		n.1203578C>T	
C321.ΔA.2.109.F	2E+06	C	T	intergenic_region	curA-mcbR		n.1524889C>T	
C321.ΔA.2.109.F	2E+06	G	A	intergenic_region	mglB-galS		n.2244378G>A	
C321.ΔA.2.109.F	3E+06	G	A	intergenic_region	mlaF-yrbG		n.3340764G>A	
C321.ΔA.2.109.F	3E+06	G	A	intergenic_region	fmt-rsmB		n.3435869G>A	
C321.ΔA.2.109.F	4E+06	C	T	intergenic_region	glmU-atpC		n.3917229C>T	
C321.ΔA.2.109.F	4E+06	G	A	non_coding_transcript_exon_variant	rrlC	rRNA	n.1619G>A	
C321.ΔA.2.109.F	4E+06	C	T	non_coding_transcript_ex	rrsB	rRNA	n.235C>T	

				on_variant			
C321.ΔA.2.109.F	4E+06	C	A	intergenic_region	rrlB-rrfB	n.4173476C>A	
C321.ΔA.2.109.F	4E+06	C	G	intergenic_region	adiY-adiA	n.4340053C>G	
C321.ΔA.2.109.F	4E+06	T	C	intergenic_region	yjfC-aidB	n.4416174T>C	
C321.ΔA.2.109.F	4E+06	T	C	intergenic_region	tabA-yjgL	n.4477340T>C	
C321.ΔA.2.109.F	5E+06	A	G	intergenic_region	fhuF-yjjZ	n.4607664A>G	
C321.ΔA.2.109.F	5E+06	A	G	intergenic_region	yjjK-slt		n.4632611A>G
C321.ΔA.2.109.F	45463	T	TATA	start_lost&disruptive_inframe_insertion	fixX	protein_coding	c.2_3insAAT
C321.ΔA.2.109.F	52429	C	CATT	start_lost&disruptive_inframe_insertion	rsmA	protein_coding	c.-1_2dupAAT
C321.ΔA.2.109.F	59171	T	TC	frameshift_variant	yabP	protein_coding	c.704dupC	p.Leu238fs
C321.ΔA.2.109.F	94480	C	CATA	start_lost&disruptive_inframe_insertion	murF	protein_coding	c.2_3insAAT
C321.ΔA.2.109.F	99474	G	GATA	start_lost&disruptive_inframe_insertion	murG	protein_coding	c.2_3insAAT
C321.ΔA.2.109.F	214662	C	CATA	start_lost&disruptive_inframe_insertion	arfB	protein_coding	c.2_3insAAT
C321.ΔA.2.109.F	232880	CG	C	frameshift_variant	mltD	protein_coding	c.905delC	p.Pro302fs
C321.ΔA.2.109.F	298918	C	CG	frameshift_variant	paoC	protein_coding	c.1068dupC	p.Gly357fs
C321.ΔA.2.109.F	414800	C	CATT	start_lost&disruptive_infr	sbcC	protein_coding	c.2_3insAAT

				ame_insertion				
C321.ΔA.2.109.F	560739	T	TG	frameshift_variant	sfmD	protein_coding	c.2000dupG	p.Gly668fs
C321.ΔA.2.109.F	722567	GT	G	frameshift_variant	kdpD	protein_coding	c.895delA	p.Thr299fs
C321.ΔA.2.109.F	748028	C	CATT	start_lost&disruptive_inframe_insertion	ybgO	protein_coding	c.-1_2dupAAT	
C321.ΔA.2.109.F	775893	GA	G	frameshift_variant	tolA	protein_coding	c.510delA	p.Ala171fs
C321.ΔA.2.109.F	1E+06	C	CATA	start_lost&disruptive_inframe_insertion	holB	protein_coding	c.2_3insAAT	
C321.ΔA.2.109.F	1E+06	A	AG	frameshift_variant	ycgB	protein_coding	c.969dupC	p.Tyr324fs
C321.ΔA.2.109.F	1E+06	GCA	G	frameshift_variant	stfR	protein_coding	c.3068_3069delAC	p.His1023fs
C321.ΔA.2.109.F	1E+06	GT	G	frameshift_variant	cybB	protein_coding	c.449delT	p.Leu150fs
C321.ΔA.2.109.F	2E+06	A	AT	frameshift_variant	ynfF	protein_coding	c.118dupT	p.Ser40fs
C321.ΔA.2.109.F	2E+06	A	AT	frameshift_variant&stop_gained	exoX	protein_coding	c.81_82insT	p.Lys28fs
C321.ΔA.2.109.F	2E+06	GC	G	frameshift_variant	cpsG	protein_coding	c.340delG	p.Ala114fs
C321.ΔA.2.109.F	2E+06	AC	A	frameshift_variant	yohK	protein_coding	c.35delC	p.Thr12fs
C321.ΔA.2.109.F	2E+06	CA	C	frameshift_variant	fadJ	protein_coding	c.870delT	p.Phe290fs
C321.ΔA.2.109.F	3E+06	G	GC	frameshift_variant	fryA	protein_coding	c.628dupG	p.Ala210fs
C321.ΔA.2.109.F	3E+06	AC	A	frameshift_variant	eutE	protein_coding	c.28delG	p.Val10fs

C321.ΔA.2.109.F	3E+06	C	CATT	start_lost&disruptive_inframe_insertion	truD	protein_coding	c.-1_2dupAAT	
C321.ΔA.2.109.F	3E+06	A	AG	frameshift_variant	ygeR	protein_coding	c.218dupC	p.Tyr74fs
C321.ΔA.2.109.F	3E+06	G	GCC	frameshift_variant	yggN	protein_coding	c.590_591dupGG	p.Leu198fs
C321.ΔA.2.109.F	3E+06	CCA	C	frameshift_variant	dnaG	protein_coding	c.1726_1727delCA	p.Gln576fs
C321.ΔA.2.109.F	3E+06	G	GCC	frameshift_variant	dnaG	protein_coding	c.1728_1729insCC	p.Glu577fs
C321.ΔA.2.109.F	3E+06	C	CATT	start_lost&disruptive_inframe_insertion	rplW	protein_coding	c.2_3insAAT	
C321.ΔA.2.109.F	3E+06	C	CT	frameshift_variant	gspA	protein_coding	c.1336dupA	p.Ser446fs
C321.ΔA.2.109.F	3E+06	G	GC	frameshift_variant	argD	protein_coding	c.889dupG	p.Ala297fs
C321.ΔA.2.109.F	4E+06	G	GC	frameshift_variant	yhhS	protein_coding	c.1018dupG	p.Ala340fs
C321.ΔA.2.109.F	4E+06	C	CT	frameshift_variant&start_lost	dgoD	protein_coding	c.2_3insA	p.Met1fs
C321.ΔA.2.109.F	4E+06	C	CT	frameshift_variant&splice_region_variant	dgoA	protein_coding	c.616_617insA	
C321.ΔA.2.109.F	4E+06	ACG	A	frameshift_variant	dgoA	protein_coding	c.596_597delCG	p.Ala199fs
C321.ΔA.2.109.F	4E+06	C	CATT	start_lost&disruptive_inframe_insertion	dgoK	protein_coding	c.2_3insAAT	
C321.ΔA.2.109.F	4E+06	GA	G	frameshift_variant	dgoR	protein_coding	c.661delT	p.Ser221fs

C321.ΔA.2.109.F	4E+06	C	CATT	start_lost&disruptive_inframe_insertion	hemD	protein_coding	c.2_3insAAT
C321.ΔA.2.109.F	4E+06	C	CATA	start_lost&disruptive_inframe_insertion	birA	protein_coding	c.2_3insAAT
C321.ΔA.2.109.F	4E+06	GC	G	frameshift_variant	phnP	protein_coding	c.59delG	p.Gly20fs
C321.ΔA.2.109.F	4E+06	GA	G	frameshift_variant	mutL	protein_coding	c.98delA	p.Asn33fs
C321.ΔA.2.109.F	4E+06	AC	A	frameshift_variant	msrA	protein_coding	c.275delG	p.Gly92fs
C321.ΔA.2.109.F	5E+06	AC	A	frameshift_variant	yjhG	protein_coding	c.1492delG	p.Val498fs
C321.ΔA.2.109.F	5E+06	C	CA	frameshift_variant	fimB	protein_coding	c.28dupA	p.Arg10fs
C321.ΔA.2.109.F	5E+06	A	AG	frameshift_variant	leuV	protein_coding	c.79dupC	p.Leu27fs
C321.ΔA.2.109.F	5E+06	A	AG	frameshift_variant	leuP	protein_coding	c.79dupC	p.Leu27fs
C321.ΔA.2.109.F	5E+06	A	AG	frameshift_variant	leuQ	protein_coding	c.79dupC	p.Leu27fs
C321.ΔA.2.109.F	4E+06	A	AAAT	splice_region_variant&non_coding_transcript_exon_variant	esrE	lincRNA	n.248_250dupAAT
C321.ΔA.2.109.F	45463	T	TATA	splice_region_variant&disruptive_inframe_insertion	fixC	protein_coding	c.1285_1286insAAT	p.Ter429dup
C321.ΔA.2.109.F	52429	C	CATT	splice_region_variant&conservative_infram	pdxA	protein_coding	c.986_988dupAAT	p.Ter330dup

				e_insertion				
C321.ΔA.2.109.F	94480	C	CATA	splice_region_variant&disruptive_inframe_insertion	murE	protein_coding	c.1486_1487insAAT	p.Ter496dup
C321.ΔA.2.109.F	99474	G	GATA	splice_region_variant&disruptive_inframe_insertion	ftsW	protein_coding	c.1243_1244insAAT	p.Ter415dup
C321.ΔA.2.109.F	103811	C	CAAT	splice_region_variant&disruptive_inframe_insertion	ftsQ	protein_coding	c.827_829dupAAT	p.Ter277dup
C321.ΔA.2.109.F	204321	C	CAAT	splice_region_variant&disruptive_inframe_insertion	lpxB	protein_coding	c.1145_1147dupAAT	p.Ter383dup
C321.ΔA.2.109.F	214662	C	CATA	splice_region_variant&disruptive_inframe_insertion	yaeQ	protein_coding	c.544_545insAAT	p.Ter182dup
C321.ΔA.2.109.F	414800	C	CATT	splice_region_variant&conservative_inframe_insertion	sbcD	protein_coding	c.1201_1202insAAT	p.Ter401dup
C321.ΔA.2.109.F	748028	C	CATT	splice_region_variant&conservative_inframe_insertion	ybgP	protein_coding	c.725_727dupAAT	p.Ter243dup
C321.ΔA.2.109.F	975469	C	CAAT	splice_region_variant&disruptive_inframe_insertion	msbA	protein_coding	c.1745_1747dupAAT	p.Ter583dup

C321.ΔA.2.109.F	1E+06	C	CATA	splice_region_variant&disruptive_inframe_insertion	tmk	protein_coding	c.640_641insAAT	p.Ter214dup
C321.ΔA.2.109.F	1E+06	C	CAAT	splice_region_variant&disruptive_inframe_insertion	prmC	protein_coding	c.830_832dupAAT	p.Ter278dup
C321.ΔA.2.109.F	2E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	yefM	protein_coding	c.250_251insAAT	p.Ter84dup
C321.ΔA.2.109.F	2E+06	C	CATT	disruptive_inframe_insertion	yoeB	protein_coding	c.2_3insAAT	p.Val1_Lys2insMet
C321.ΔA.2.109.F	2E+06	G	GAAT	splice_region_variant&disruptive_inframe_insertion	yohJ	protein_coding	c.395_397dupAAT	p.Ter133dup
C321.ΔA.2.109.F	3E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	ispF	protein_coding	c.476_478dupAAT	p.Ter160dup
C321.ΔA.2.109.F	3E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	rplD	protein_coding	c.604_605insAAT	p.Ter202dup
C321.ΔA.2.109.F	4E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	dgoR	protein_coding	c.688_689insAAT	p.Ter230dup

C321.ΔA.2.109.F	4E+06	C	CGGG	disruptive_inframe_insertion	wzzE	protein_coding	c.1001_1003dupGGG	p.Gly334dup
C321.ΔA.2.109.F	4E+06	C	CATT	splice_region_variant&conservative_inframe_insertion	hemC	protein_coding	c.940_941insAAT	p.Ter314dup
C321.ΔA.2.109.F	4E+06	A	AAAT	splice_region_variant&disruptive_inframe_insertion	yigP	protein_coding	c.602_604dupAAT	p.Ter202dup
C321.ΔA.2.109.F	4E+06	C	CATA	splice_region_variant&disruptive_inframe_insertion	murB	protein_coding	c.1027_1028insAAT	p.Ter343dup
C321.ΔA.2.109.F	20575	TG	T	intergenic_region	insA-rpsT	n.20576delG		
C321.ΔA.2.109.F	34112	GT	G	intergenic_region	carB-caiF	n.34113delT		
C321.ΔA.2.109.F	136897	A	AG	intergenic_region	yacC-cueO	n.136897_136898insG		
C321.ΔA.2.109.F	379063	A	AGG	intergenic_region	frmR-yaiO	n.379063_379064insGG		
C321.ΔA.2.109.F	440454	C	CA	intergenic_region	xseB-thiI	n.440454_440455insA		
C321.ΔA.2.109.F	457782	A	AG	intergenic_region	clpX-lon	n.457782_457783insG		
C321.ΔA.2.109.F	531493	C	CA	intergenic_region	allS-allA	n.531493_531494insA		
C321.ΔA.2.109.F	650998	GA	G	intergenic_region	citC-dpiB	n.650999delA		
C321.ΔA.2.109.F	719619	C	CT	intergenic_region	speF-ybfK	n.719619_719620insT		
C321.ΔA.2.109.F	779889	C	CT	intergenic_region	valT-lysW	n.779889_779890insT		
C321.ΔA.2.109.F	805228	C	CT	intergenic_region	ybhC-int	n.805228_805229insT		
C321.ΔA.2.109.F	970912	CT	C	intergenic_region	rpsA-ihfB	n.970913delT		
C321.ΔA.2.109.F	1E+06	TA	T	intergenic_region	ymfJ-cohE	n.1209324delA		
C321.ΔA.2.109.F	2E+06	T	TG	intergenic_region	yoeI-yeeY	n.2091312_2091313insG		

C321.ΔA.2.109.F	3E+06	CGCACTATG	C	intergenic_region	yfjL-yfjM	n.2769432_2769439delGCACTATG	
C321.ΔA.2.109.F	3E+06	G	GC	intergenic_region	yfjM-rnlA	n.2769807_2769808insC	
C321.ΔA.2.109.F	4E+06	T	TC	intergenic_region	bcsG-ldrD	n.3700603_3700604insC	
C321.ΔA.2.109.F	4E+06	AGG	A	intergenic_region	kup-rbsD	n.3935127_3935128delGG	
C321.ΔA.2.109.F	4E+06	AGT	A	non_coding_transcript_exon_variant	rrlA	rRNA	n.291_292delGT
C321.ΔA.2.109.F	5E+06	C	CA	intergenic_region	yjjM-yjjN	n.4597823_4597824insA	
C321.ΔA.2.109.F	3E+06	G	A	synonymous_variant	agaS	protein_coding	c.312G>A	p.Pro104Pro
C321.ΔA.2.109.F	816640	C	T	synonymous_variant	bla	protein_coding	c.228C>T	p.Gly76Gly
C321.ΔA.2.109.F	816656	A	G	missense_variant	bla	protein_coding	c.244A>G	p.Ile82Val
C321.ΔA.2.109.F	950693	T	C	missense_variant	dmsB	protein_coding	c.175T>C	p.Phe59Leu
C321.ΔA.2.109.F	4E+06	G	A	synonymous_variant	emrD	protein_coding	c.438G>A	p.Leu146Leu
C321.ΔA.2.109.F	379063	A	AG	intergenic_region	frmR-yaiO	n.379063_379064insG	
C321.ΔA.2.109.F	1E+06	G	C	intergenic_region	hemA-prmC	n.1272099G>C	
C321.ΔA.2.109.F	1E+06	A	ATTTACGCC	intergenic_region	hemA-prmC	n.1272100_1272101insTTTACGCC	
C321.ΔA.2.109.F	1E+06	GGAGGAATAA	G	intergenic_region	hemA-prmC	n.1272102_1272110delGAGGAATAA	
C321.ΔA.2.109.F	3E+06	G	A	synonymous_variant	ispF	protein_coding	c.48C>T	p.Gly16Gly
C321.ΔA.2.109.F	2E+06	C	A	missense_variant	leuE	protein_coding	c.610G>T	p.Ala204Ser

C321.ΔA.2.109.F	3E+06	G	A	synonymous_variant	mtr	protein_coding	c.63C>T	p.Gly21Gly
C321.ΔA.2.109.F	3E+06	G	A	missense_variant	murP	protein_coding	c.1189G>A	p.Ala397Thr
C321.ΔA.2.109.F	808519	C	T	missense_variant	N840_4502	protein_coding	c.181G>A	p.Asp61Asn
C321.ΔA.2.109.F	808519	C	T	synonymous_variant	N840_4503	protein_coding	c.9G>A	p.Glu3Glu
C321.ΔA.2.109.F	2E+06	T	C	missense_variant	nuoJ	protein_coding	c.521A>G	p.Asp174Gly
C321.ΔA.2.109.F	2E+06	G	A	missense_variant	patD	protein_coding	c.1283G>A	p.Gly428Asp
C321.ΔA.2.109.F	1E+06	GGAGGAATAA	G	frameshift_variant&start_lost	prmC	protein_coding	c.-8_1delGAGGAATAA	p.Met1fs
C321.ΔA.2.109.F	223835	C	T	non_coding_transcript_exon_variant	rrsH	rRNA	n.235C>T	
C321.ΔA.2.109.F	3E+06	T	C	intergenic_region	sseA-ryfA		n.2657361T>C	
C321.ΔA.2.109.F	4E+06	C	CGGGG	frameshift_variant	wzzE	protein_coding	c.1000_1003dupGGGG	p.Ala335fs
C321.ΔA.2.109.F	736478	T	A	synonymous_variant	ybfL	protein_coding	c.606T>A	p.Leu202Leu
C321.ΔA.2.109.F	801006	G	A	missense_variant	ybhI	protein_coding	c.1221G>A	p.Met407Ile
C321.ΔA.2.109.F	2E+06	A	AGCTTTCAT	frameshift_variant	yejO	protein_coding	c.3777_3778insATGAAAGC	p.Ser1260fs
C321.ΔA.2.109.F	2E+06	C	CTATGAAATTAATTGCTGTTAAAAGCATTGGGT	frameshift_variant	yejO	protein_coding	c.3775_3776insACCCAATGCTTTTAACAGCAATTAATTTCATA	p.Trp1259fs
C321.ΔA.2.109.F	2E+06	A	T	synonymous_variant	yejO	protein_coding	c.3780T>A	p.Ser1260Ser

C321.ΔA.2.109.F	3E+06	G	A	synonymous_variant	ygjJ	protein_coding	c.732G>A	p.Gln244Gln
C321.ΔA.2.109.F	3E+06	C	T	intergenic_region	yhcN-yhcO		n.3386519C>T	
C321.ΔA.2.109.F	1E+06	A	AATGAAATG	intergenic_region	ymfE-lit		n.1205553_1205554insATGAAATG	
C321.ΔA.2.109.F	2E+06	C	T	synonymous_variant	yneF	protein_coding	c.240G>A	p.Leu80Leu
C321.ΔA.2.109.F	3E+06	T	C	missense_variant	yqiA	protein_coding	c.7A>G	p.Thr3Ala

Supplemental Table 7. Non-synonymous mutations in essential or conditionally (minimal media) essential genes. N signifies the mutation is absent from the strain, while Y means it is present.

Mutation	C321.ΔA2..39	C321.ΔA.2.64	C321.ΔA.2.104	C321.ΔA.2.109.F	Function
folP M218V	Mutant	WT	Mutant	Mutant	Tetrahydrafolate biosynthesis
waaA frameshift	Mutant	WT	Mutant	WT	LPS biosynthesis
trpE S394P	Mutant	WT	Mutant	WT	Tryptophan biosynthesis
metK S12P	Mutant	WT	Mutant	WT	SAM biosynthesis
kdsB A60V	Mutant	Mutant	Mutant	Mutant	LPS biosynthesis
lspA I30V	WT	Mutant	Mutant	WT	Lipoprotein posttranslational modification
carB D27Y	WT	Mutant	Mutant	WT	catalyzes the first committed step in the separate biosynthetic pathways for the production of arginine, and pyrimidine nucleotides
ftsQ A4T	WT	Mutant	Mutant	WT	Cell division
dnaQ T44A	WT	Mutant	Mutant	WT	Proofreading exonuclease of polIII
purK F228S	WT	Mutant	Mutant	WT	Ribonucleotide biosynthesis
polB A507T	WT	Mutant	Mutant	Mutant	Pol II is required for a nucleotide excision repair pathway
holA L330P	WT	Mutant	Mutant	Mutant	DNA polIII
hemE M85I	WT	Mutant	Mutant	Mutant	Heme biosynthesis
Dxs E231K	WT	Mutant	Mutant	Mutant	Isoprenoid biosynthesis
cydA R498H	WT	WT	Mutant	Mutant	Cyctochrome biosynthesis
leuC K263R	WT	WT	Mutant	Mutant	Leucine biosynthesis
pheS M42V	WT	WT	Mutant	Mutant	Phenylalanine-tRNA synthetase

Supplemental Table 8. Mass spec peptide data.

MG1655=MG1655 glyT; C321.ΔA.2.109.F=C321.ΔA.2.109.F glyT; C321.ΔB.2.109.F=C321.ΔB.2.109.F pBAD prfB, glyT)

Sequence	Modifications	Modified sequence	Missed cleavages	Experiment (	MS/MS m/z	Charge	m/z	Mass	Resolution	Mass Error [ppm]	PEP	gene	suppressoraa
KLLNLR	Deamidation (NQ)	_KLLN(de)LR_	1	MG1655	379.25	2	379.25	756.49	57260	0.0199	0.0428	bcsF	l
SSIWCLPR	Unmodified	_SSIWCLPR_	0	MG1655	339.68	3	340.18	1017.5	56457	2.7684	0.0591	bcsF	l
FVAHPGCVFAGCDDAS	Unmodified	_FVAHPGCVFAGCDDAS_	0	MG1655	855.35	2	855.36	1708.7	38299	-0.731	1E-06	cbpM	g
FVAHPGCVFAGCDDAS	Unmodified	_FVAHPGCVFAGCDDAS_	0	C321.ΔA.2.109.F	855.36	2	855.36	1708.7	38506	-0.51	1E-06	cbpM	g
EITKQLATSQLTGDR	Phospho (STY)	_EITKQLAT(ph)SQLTGDR_	1	MG1655	870.93	2	870.93	1739.8	37205	-2.062	0.0525	dgoR	k
VLVLAEGVK	Unmodified	_VLVLAEGVK_	0	MG1655	464.3	2	464.3	926.58	50642	0.7379	0.0046	elbB	g
VLVLAEGVK	Unmodified	_VLVLAEGVK_	0	PRE_C321.ΔA.2.109.F	464.3	2	464.3	926.58	53617	0.5348	0.0582	elbB	g

VLVLAEGVK	Unmodified	_VLVLAEGVK_	0	C321.ΔA.2.109.F	464.3	2	464.3	926.58	50474	0.2658	0.00005	elbB	g
VLVLAEGVK	Unmodified	_VLVLAEGVK_	0	PRE_C321.ΔB.2.109.F	464.3	2	464.3	926.58	53320	0.548	0.03036	elbB	g
ITIAGYALYEQHHPVGR	Unmodified	_ITIAGYALYEQHHPVGR_	0	C321.ΔA.2.109.F	482.01	4	482.01	1924	49244	0.539	0.00006	fbaB	g
KISINR	Acetyl (Protein N-term)	_(ac)KISINR_	1	MG1655	386.74	2	386.74	771.46	56593	0.3901	0.004	flgN	n
PSQGAGQCWTK	Unmodified	_PSQGAGQCWTK_	0	MG1655	610.28	2	610.28	1218.5	45390	-0.762	0.0025	galK	w
VLLIPGNR	Unmodified	_VLLIPGNR_	0	MG1655	441.28	2	441.28	880.55	53164	-0.01	0.0008	gatD	g
VLLIPGNR	Unmodified	_VLLIPGNR_	0	PRE_C321.ΔA.2.109.F	441.28	2	441.28	880.55	52265	1.3245	0.0055	gatD	g
VLLIPGNR	Unmodified	_VLLIPGNR_	0	C321.ΔA.2.109.F	441.28	2	441.28	880.55	53386	0.0452	0.0025	gatD	g
VLLIPGNR	Unmodified	_VLLIPGNR_	0	PRE_C321.ΔB.2.109.F	441.28	2	441.28	880.55	52461	-0.223	0.0055	gatD	g

ANFESLR	Unmodified	_ANFESLR_	0	C321.ΔA.2.109.F	418.72	2	418.72	835.42	53192	0.9068	0.001	glnE	?
ANFESLR	Unmodified	_ANFESLR_	0	C321.ΔB.2.109.F	418.72	2	418.72	835.42			0.0191	glnE	?
ILHLNLK	Deamidation (NQ)	_ILHLN(de)LK_	0	C321.ΔA.2.109.F	426.27	2	426.27	850.53	54356	0.1932	0.059	glnP	h
ILHLNLK	Deamidation (NQ)	_ILHLN(de)LK_	0	C321.ΔB.2.109.F	426.27	2	426.27	850.53	54021	-0.181	0.0423	glnP	h
TALSAPSAPPPAR	Unmodified	_TALSAPSAPPPAR_	0	MG1655	618.34	2	618.34	1234.7	44523	0.3635	0.00001	glpB	?
TALSAPSAPPPAR	Unmodified	_TALSAPSAPPPAR_	0	C321.ΔA.2.109.F	618.34	2	618.34	1234.7	44624	0.3647	4E-06	glpB	?
EILAEVYNGDAPAGVSLSPAR	Unmodified	_EILAEVYNGDAPAGVSLSPAR_	0	MG1655	1065	2	1065	2128.1	45734	0.3161	5E-11	hemC	g
QQESEAGIISNSPTVAAAR	Unmodified	_QQESEAGIISNSPTVAAAR_	0	MG1655	964.99	2	964.99	1928		-1.186	2E-07	hypD	g
QLGVIAGGCEMK	Unmodified	_QLGVIAGGCE	0	PRE_C321.ΔB.	631.82	2	631.82	1261.6	43487	2.7054	0.0581	kbl	g

		MK_		2.109.F									
KGQSIKVK	Acetyl (Protein N-term), Deamidation (NQ)	_(ac)KGQ(de)SIKVK_	2	C321.ΔA.2.109.F	465.79	2	465.78	929.55	50621	0.3633	0.0097	kdsC	k
KGQSIKVK	Acetyl (Protein N-term), Deamidation (NQ)	_(ac)KGQ(de)SIKVK_	2	C321.ΔB.2.109.F	465.79	2	465.78	929.55	50158	0.3317	0.0285	kdsC	k
LASIAIK	Unmodified	_LASIAIK_	0	C321.ΔB.2.109.F	358.24	2	358.24	714.46	57045	-0.007	0.001	mngB	s
LHPAPHAASAK	Phospho (STY)	_LHPAPHAAS(ph)AK_	0	MG1655	590.29	2	590.29	1178.6	44951	1.7434	0.0162	nrdH	k
YGEVSQVR	Unmodified	_YGEVSQVR_	0	MG1655	469.24	2	469.24	936.47	50057	0.8126	0.0006	nrfC	v
KEAQQVVGYHR	2 Deamidation (NQ)	_KEAQ(de)Q(de)VVGYHR_	1	C321.ΔA.2.109.F	658.84	2	658.83	1315.7	42505	0.0523	0.0052	nuoG	v
KEAQQ	2 Dea	_KEAQ(	1	C321.Δ	658.83	2	658.83	1315.7	42238	1.5896	0.0163	nuoG	v

VVGYHR	midation (NQ)	de)Q(de)VVGYHR_		B.2.109.F									
LIAAIR	Unmodified	_LIAAIR_	0	MG1655	328.73	2	328.73	655.44	58212	0.1955	0.0013	rfbC	a
LIAAIR	Unmodified	_LIAAIR_	0	PRE_C321.ΔA.2.109.F	328.73	2	328.73	655.44	58945	-1.267	0.0538	rfbC	a
LIAAIR	Unmodified	_LIAAIR_	0	C321.ΔA.2.109.F	328.73	2	328.73	655.44	58269	0.3263	0.0011	rfbC	a
LIAAIR	Unmodified	_LIAAIR_	0	PRE_C321.ΔB.2.109.F	328.73	2	328.73	655.44	57233	-0.896	0.0279	rfbC	a
LIAAIR	Unmodified	_LIAAIR_	0	C321.ΔB.2.109.F	328.73	2	328.73	655.44	57785	-0.101	0.0015	rfbC	a
LIAHIR	Unmodified	_LIAHIR_	0	C321.ΔA.2.109.F	361.74	2	361.74	721.46	59002	-0.163	0.0019	rfbC	h
LIAHIR	Unmodified	_LIAHIR_	0	C321.ΔB.2.109.F	361.25	2	361.74	721.46	57790	-0.413	0.047	rfbC	h
ASGKPLLNLLP	Unmodified	_ASGKPLLNLLP_	1	MG1655	561.85	2	561.85	1121.7	47438	0.1942	0.0037	ribB	g
HVGPDVCLHLVGA	Unmodified	_HVGPDVCLHLVGAGSP	0	MG1655	629.99	3	629.66	1885.9	47794	-0.051	2E-07	ribD	g

GSPGSK		GSK_											
QPALGYLNGTPK	Unmodified	_QPALGYLNGTPK_	0	MG1655	629.85	2	629.84	1257.7	42933	0.3485	0.0002	rpsG	g
QPALGYLNGTPK	Deamidation (NQ)	_QPALGYLN(de)GTPK_	0	MG1655	630.34	2	630.34	1258.7	42810	-1.001	0.0072	rpsG	g
QPALGYLNWTPK	Unmodified	_QPALGYLNWTPK_	0	MG1655	694.37	2	694.37	1386.7	41831	-0.055	0.0005	rpsG	w
LVRLNRTK	Phospho (STY)	_LVRLNRT(ph)K_	2	MG1655	540.31	2	540.31	1078.6	48740	1.1388	0.0294	speF	?
QWINNSPGAPSSATASVTSPITSPSQWGR	Unmodified	_QWINNSPGAPSSATASVTSPITSPSQWGR_	0	C321.ΔB.2.109.F	996.16	3	995.49	2983.4	34899	0.6045	2E-20	uidA	w
SVLVDGGQYTGQDWCR	Unmodified	_SVLVDGGQYTGQDWCR_	0	MG1655	920.91	2	920.92	1839.8	44685	-2.155	7E-10	wecE	?
SVLVDGGQYTGQDWCR	Unmodified	_SVLVDGGQYTGQDWCR_	0	C321.ΔA.2.109.F	920.92	2	920.92	1839.8	38245	-0.805	1E-10	wecE	?
RQFAGF	Deamidation	_RQ(de)FAGF_	1	C321.ΔA.2.	363.68	2	363.68	725.35			0.0464	ybbP	f

	(NQ)			109.F									
QGMPVTVQFGDEAGHEGCR	Oxidation (M)	_QGM(ox)PVTVQFGDEAGHEGCR_	0	C321.ΔA.2.109.F	697.98	3	697.64	2089.9	41836	0.1682	5E-05	ybhG	g
QGMPVTVQFGDEAGHEGCR	Unmodified	_QGMPVTVQFGDEAGHEGCR_	0	C321.ΔA.2.109.F	692.64	3	692.31	2073.9	41350	-0.218	1E-07	ybhG	g
QGMPVTVQFGDEAGHEGCR	Unmodified	_QGMPVTVQFGDEAGHEGCR_	0	C321.ΔA.2.109.F	1038.5	2	1038	2073.9	34350	-1.245	1E-22	ybhG	g
YHAERPGK	Unmodified	_YHAERPGK_	1	MG1655	319.84	3	319.83	956.48			0.0421	ybhG	g
YHAERPGK	Unmodified	_YHAERPGK_	1	C321.ΔA.2.109.F	319.84	3	319.83	956.48	58844	0.7983	0.0578	ybhG	g
YHAERPGK	Unmodified	_YHAERPGK_	1	C321.ΔA.2.109.F	479.25	2	479.25	956.48	50266	-0.912	0.0018	ybhG	g
GGTILGLK	Unmodified	_GGTILGLK_	0	MG1655	379.74	2	379.74	757.47	59123	0.7462	0.001	ydcY	g
GGTILGLK	Unmodified	_GGTILGLK_	0	C321.ΔA.2.109.F	379.74	2	379.74	757.47	57687	0.4015	0.0027	ydcY	g

KGGTILGLK	Acetyl (Protein N-term)	_(ac)KGGTILGLK_	1	MG1655	464.8	2	464.79	927.58	50174	-0.106	0.00007	ydcY	g
KGGTILGLK	Acetyl (Protein N-term)	_(ac)KGGTILGLK_	1	C321.ΔA.2.109.F	464.8	2	464.79	927.58	50352	-0.389	0.00005	ydcY	g
ALNPSLR	Deamidation (NQ)	_ALN(de)PSLR_	0	MG1655	386.22	2	386.22	770.43	56447	-0.165	0.057	ydiQ	p
ALNVSLR	Deamidation (NQ)	_ALN(de)VSLR_	0	C321.ΔB.2.109.F	387.23	2	387.23	772.44	58021	0.4148	0.051	ydiQ	v
LPSPQVVGAESEEDASHAAGSNLR	Unmodified	_LPSPQVVGAESEEEDASHAAGSNLR_	0	MG1655	851.08	3	850.74	2549.2	38401	-0.547	2E-18	yebG	g
IVASGENPLR	Deamidation (NQ)	_IVASGEN(de)PLR_	0	MG1655	528.79	2	528.79	1055.6			0.0296	ygfS	r
VAEQEVR	Unmodified	_VAEQEVR_	0	C321.ΔB.2.109.F	415.72	2	415.72	829.43	52559	-0.064	0.00026	yidG	e
SLAGWSCFR	Unmodified	_SLAGWSCFR_	0	MG1655	542.26	2	542.26	1082.5	47394	-2.982	0.00003	yiiS	g

KNNINR	Acetyl (Protein N-term), Deamidation (NQ)	_(ac)KNNIN(de)R_	1	C321.ΔB.2.109.F	401.23	2	401.21	800.41	58148	-1.092	0.0314	ymfE	r
KGIINIR	Deamidation (NQ)	_KGIIN(de)IR_	1	C321.ΔB.2.109.F	408.24	2	407.76	813.51		-0.513	0.0425	yodC	r

Plasmid sequences

glyT plasmid sequence

CCGCGGCCGGGAAGCCGATCTCGGCTTGAACGAATTGTTAGGTGGCGGTACTTGGGTCGATATCAAAGT
GCATCACTTCTTCCCGTATGCCCAACTTTGTATAGAGAGCCACTGCGGGATCGTCACCGTAATCTGCTTGC
ACGTAGATCACATAAGCACCAAGCGCGTTGGCCTCATGCTTGAGGAGATTGATGAGCGCGGTGGCAATG
CCCTGCCTCCGGTGCTCGCCGGAGACTGCGAGATCATAGATATAGATCTCACTACGCGGCTGCTCAAACC
TGGGCAGAACGTAAGCCGCGAGAGCGCCAACAACCGCTTCTTGGTCGAAGGCAGCAAGCGCGATGAAT
GTCTTACTACGGAGCAAGTTCCCGAGGTAATCGGAGTCCGGCTGATGTTGGGAGTAGGTGGCTACGTCT
CCGAACTCACGACCGAAAAGATCAAGAGCAGCCCGCATGGATTTGACTTGGTCAGGGCCGAGCCTACAT
GTGCGAATGATGCCCATACTTGAGCCACCTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTT
GCTGCTGCGTAACATCGTTGCTGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTT
GGATGCCCGAGGCATAGACTGTACAAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGTT
ACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTGAGCGCATACGCTACTTGCATTACAGTTTA
CGAACCGAACAGGCTTATGTCAACTGGGTTCGTGCCTTCATCCGTTTCCACGGTGTGCGTCCATGGCTGT
CAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAatttaaaaggatctaggtgaagatcc
tttttgataatcATAAATTGCACTGAAATCTAGAAATATTTTATCTGATTAATAAGAtgatcttcttgagatcgttttggt
ctgcgcgtaatctcttgctctgaaaacgaaaaaaccgccttgcagggcggtttttcgaaggttctctgagctaccaactctttgaaccgagg
taactggcttggaggagcgcagtcaccaaaacttgtcctttcagtttagccttaaccggcgcatgacttcaagactaactcctctaaatcaa
ttaccagtggctgctgccagtggtgcttttgcatgtctttccgggttggactcaagacgatagttaccggataaggcgcagcggtcggactg
aacgggggttcgtgcatacagtccagcttggagcgaactgcctacccggaactgagtgtcaggcgtggaatgagacaaacgcggccat
aacagcggaatgacaccggtaaaccgaaaggcaggaacaggagagcgcacgagggagccgccaggggaaacgcctggtatctttat
agtcctgtcgggtttcgccaccactgatttgagcgtcagatttcgtgatgcttgtcaggggggcggagcctatggaaaaacggctttgccgc
ggccctctcacttccctgttaagtatcttcctggcatcttccaggaaatctccgccccgttcgtaagccatttccgctcgccgcagtcgaacg
accgagcgtagcgagtcagtgagcgaggaagcggaatatatcctgtatcacatattctgCTGACGCACCGGTGCAGTTTTTAC
GGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTaagcttaaaaaaaatccttagctttcgctaaggatctgcaggtcga
ctctagAGGATCCTGGAGCGGGCAGCGGGAATCGAACCCGCATCATCAGCTTtGAAGGCTGAGGTAATAG
CCATTATACGATGCCCGCGAATTCagcgttacaagtattacacaaagtttttatgttgagaatatttttttgatggggcgccactt
atttttgatcgttcgctcaaagaagcggcgcc

Gent resistance cassette

P15a origin

rrnC terminator

glyT

lpp promoter

mCherry-TAG plasmid

TAGAAGCTTAATTAGCTGAGTCTAGAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCT
TTCGTTTTATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGCCCTAGACCTAGGG
TACGGGTTTTGCTGCCCGCAAACGGGCTGTTCTGGTGTTGCTAGTTTGTTATCAGAATCGCAGATCCGGC
TTCAGGTTTGCCGGCTGAAAGCGCTATTTCTTCCAGAATTGCCATGATTTTTTCCCCACGGGAGGCGTCAC
TGGCTCCCGTGTTGTCGGCAGCTTTGATTCGATAAGCAGCATCGCCTGTTTCAGGCTGTCTATGTGTGACT
GTTGAGCTGTAACAAGTTGTCTCAGGTGTTCAATTTCATGTTCTAGTTGCTTTGTTTTACTGGTTTCACCTG
TTCTATTAGGTGTTACATGCTGTTCATCTGTTACATTGTCGATCTGTTCATGGTGAACAGCTTTAAATGCAC
CAAAAACTCGTAAAAGCTCTGATGTATCTATCTTTTTTACACCGTTTTCATCTGTGCATATGGACAGTTTTC
CCTTTGATATCTAACGGTGAACAGTTGTTCTACTTTTGTTTGTTAGTCTTGATGCTTCACTGATAGATACAA
GAGCCATAAGAACCTCAGATCCTTCCGTATTTAGCCAGTATGTTCTCTAGTGTGGTTCGTTGTTTTTGCGT
GAGCCATGAGAACGAACCATTGAGATCATGCTTACTTTGCATGTCACTCAAAAATTTTGCCTCAAAACTG
GTGAGCTGAATTTTTGCAGTTAAAGCATCGTGTAGTGTTTTTCTTAGTCCGTTACGTAGGTAGGAATCTGA
TGTAATGGTTGTTGGTATTTTGTCACCATTCATTTTTATCTGGTTGTTCTCAAGTTCGGTTACGAGATCCAT
TTGTCTATCTAGTTCAACTTGGAAAATCAACGTATCAGTCGGGCGGCCTCGCTTATCAACCACCAATTTCA
TATTGCTGTAAGTGTTTAAATCTTTACTTATTGGTTTCAAAACCCATTGGTTAAGCCTTTTAAACTCATGGT
AGTTATTTTCAAGCATTAACATGAACTTAAATTCATCAAGGCTAATCTCTATATTTGCCTTGTGAGTTTTCT
TTTGTGTTAGTTCTTTTAATAACCACTCATAAATCCTCATAGAGTATTTGTTTTCAAAAGACTTAACATGTT
CCAGATTATATTTTATGAATTTTTTTAACTGGAAAAGATAAGGCAATATCTCTTCACTAAAAACTAATTCTA
ATTTTTCGCTTGAGAACTTGGCATAGTTTGTCCACTGGAAAATCTCAAAGCCTTTAACCAAAGGATTCCTG
ATTTCCACAGTTCTCGTCATCAGCTCTCTGGTTGCTTTAGCTAATACACCATAAGCATTTTCCCTACTGATG
TTCATCATCTGAGCGTATTGGTTATAAGTGAACGATACCGTCCGTTCTTTCCTTGTAGGGTTTTCAATCGT
GGGGTTGAGTAGTGCCACACAGCATAAAATTAGCTTGGTTTCATGCTCCGTTAAGTCATAGCGACTAATC
GCTAGTTCATTTGCTTTGAAAACAACTAATTCAGACATACATCTCAATTGGTCTAGGTGATTTTAATCACTA
TACCAATTGAGATGGGCTAGTCAATGATAATTACTAGTCCTTTTCCCGGGAGATCTGGGTATCTGTAAATT
CTGCTAGACCTTTGCTGGAAAACTTGTAAATTCTGCTAGACCCTCTGTAAATTCCGCTAGACCTTTGTGTG
TTTTTTTTGTTTATATTCAAGTGGTTATAATTTATAGAATAAAGAAAGAATAAAAAAAGATAAAAAGAATA
GATCCCAGCCCTGTGTATAACTCACTACTTTAGTCAGTTCCGCAGTATTACAAAAGGATGTCGCAAACGCT
GTTTGCTCCTCTACAAAACAGACCTTAAAACCCTAAAGGCTTAAGTAGCACCCTCGCAAGCTCGGGCAAA
TCGCTGAATATTCCTTTTGTCTCCGACCATCAGGCACCTGAGTCGCTGTCTTTTTCGTGACATTCAGTTCGC
TGCGCTCACGGCTCTGGCAGTGAATGGGGGTAAATGGCACTACAGGCGCCTTTTATGGATTCATGCAAG
GAAACTACCCATAATACAAGAAAAGCCCGTCACGGGCTTCTCAGGGCGTTTTATGGCGGGTCTGCTATGT
GGTGCTATCTGACTTTTTGCTGTTCAGCAGTTCCTGCCCTCTGATTTTCCAGTCTGACCACTTCGGATTATC
CCGTGACAGGTCATTCAGACTGGCTAATGCACCCAGTAAGGCAGCGGTATCATCAACAGGCTTACCCGTC
TTACTGTCCCTAGTGCTTGGATTCTCACCAATAAAAAACGCCCGGCGGCAACCGAGCGTTCTGAACAAAT
CCAGATGGAGTTCTGAGGTCATTACTGGATCTATCAACAGGAGTCCAAGCGAGCTCGTAAACTTGGTCTG
ACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCT
GACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACC
GCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG
AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTT
CGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGT
ATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAG
CGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATG
GCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAAC
CAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACC

GCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGA
TCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTT
TCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACA
CGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATG
AGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAG
TGCCACCTGACGTCCAACATATGGCACCGGAGGCTTTCGTCTTCACCTCGAGTCCCTATCAGTGATAGAG
ATTGACATCCCTATCAGTGATAGAGATACTGAGCACATCAGCAGGACGCACTGACCGAATTCCAATAAT
TTTGTTTAACTTTAAGAAGGAGATCTGCAGATGGTTTCCAAGGGCGAGGAGGATAACATGGCTATCATTA
AAGAGTTCATGCGCTTCAAAGTTCACATGGAGGGTTCTGTTAACGGTCACGAGTTCGAGATCGAAGGCG
AAGGCGAGGGCCGTCCGTATGAAGGCACCCAGACCGCCAAACTGAAAGTGACTAAAGGCGGCCCGCTG
CCTTTTGCGTGGGACATCCTGAGCCCGCAATTTATGTACGGTTCTAAAGCGTATGTTAAACACCCAGCGG
ATATCCCGGACTATCTGAAGCTGTCTTTTCCGGAAGGTTTCAAGTGGGAACGCGTAATGAATTTTGAAGA
TGGTGGTGTCGTGACCGTCACTCAGGACTCCTCCCTGCAGGATGGCGAGTTCATCTATAAAGTTAAACTG
CGTGGTACTAATTTTCCATCTGATGGCCCGGTGATGCAGAAAAAGACGATGGGTTGGGAGGCGTCTAGC
GAACGCATGTATCCGGAAGATGGTGCGCTGAAAGGCGAAATTAAACAGCGCCTGAAACTGAAAGATGG
CGGCCATTATGACGCTGAAGTGAAAACCACGTACAAAGCCAAGAAACCTGTGCAGCTGCCTGGCGCGTA
CAATGTGAATATTAAACTGGACATCACCTCTCATAATGAAGATTATACGATCGTAGAGCAATATGAGCGC
GCGGAGGGTCGTCATTCTACCGGTGGCATGGATGAGCTGTACAAATAG

mCherry

rrnB terminator

ampR gene

ampR promoter

lambda t0 promoter

pLtet promoter

pZE21G pBAD prfB

CTCGAGAAGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGTCTTTTACTGGCTCTTCTCG
CTAACCAAACCGGTAACCCCGCTTATTAAAAGCATTCTGTAACAAAGCGGGACCAAAGCCATGACAAAAA
CGCGTAACAAAAGTGTCTATAATCACGGCAGAAAAGTCCACATTGATTATTTGCACGGCGTCACACTTTG
CTATGCCATAGCATTTTTATCCATAAGATTAGCGGATCCTACCTGACGCTTTTTATCGCAACTCTCTACTGT
TTCTCCATACCCGTTTTTTGGGCTAACAGGAGGAATTAACCGGTACCatgTTTGAAATTAATCCGGTAAATA
ATCGCATTCAGGACCTCACGGAACGCTCCGACGTTCTTAGGGGGTATCTTTGACTACGACGCCAAGAAAG
AGCGTCTGGAAGAAGTAAACGCCGAGCTGGAACAGCCGGATGTCTGGAACGAACCCGAACGCGCACAG
GCGCTGGGTAAAGAGCGTTCCTCCCTCGAAGCCGTTGTCGACACCCTCGACCAAATGAAACAGGGGCTG
GAAGATGTTTCTGGTCTGCTGGAACTGGCTGTAGAAGCTGACGACGAAGAAACCTTTAACGAAGCCGTT
GCTGAACTCGACGCCCTGGAAGAAAAACTGGCGCAGCTTGAGTTCCGCCGTATGTTCTCTGGCGAATATG
ACAGCGCCGACTGCTACCTCGATATTCAGGCGGGGTCTGGCGGTACGGAAGCACAGGACTGGGCGAGC
ATGCTTGAGCGTATGTATCTGCGCTGGGCAGAATCGCGTGGTTTCAAAACTGAAATCATCGAAGAGTCG

GAAGGTGAAGTGGCGGGTATTAAATCCGTGACGATCAAAATCTCCGGCGATTACGCTTACGGCTGGCTG
CGTACAGAAACCGGCGTTCACCGCCTGGTGCGTAAAAGCCCGTTTGACTCCGGCGGTCGTCGCCACACGT
CGTTCAGCTCCGCGTTTGTTTATCCGGAAGTTGATGATGATATTGATATCGAAATCAACCCGGCGGATCT
GCGCATTGACGTTTATCGCACGTCCGGCGCGGGCGGTCAGCACGTTAACCGTACCGAATCTGCGGTGCG
TATTACCCACATCCCGACCGGGATCGTGACCCAGTGCCAGAACGACCGTTCCCAGCACAAGAACAAAGAT
CAGGCCATGAAGCAGATGAAAGCGAAGCTTTATGAACTGGAGATGCAGAAGAAAAATGCCGAGAAACA
GGCGATGGAAGATAACAAATCCGACATCGGCTGGGGCAGCCAGATTCGTTCTTATGTCCTTGATGACTCC
CGCATTAAAGATCTGCGCACCGGGGTAGAAACCCGCAACACGCAGGCCGTGCTGGACGGCAGCCTGGA
TCAATTTATCGAAGCAAGTTTGAAAGCAGGGTTAtgaAAGCTTGATATCGAATTCCTGCAGCCCGGGGGA
TCCCATGGTACGCGTGCTAGAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTT
TATCTGTTGTTTGTCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCGCCCTAGACCTAGGGCGTTCGG
CTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCA
GGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT
TTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCG
ACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCC
GCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGT
ATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCG
CTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCA
GCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCT
AACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAA
GAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA
GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGG
AACGAAAACTCACGTTAAGGGATTTTGGTCATGACTAGTGCTTGGATTCTCACCAATAAAAAACGCCCGG
CGGCAACCGAGCGTTCTGAACAAATCCAGATGGAGTTCTGAGGTCATTACTGGATCTATCAACAGGAGTC
CAAGCGAGCTCTCGAACCCCAGAGTCCCGCTCAGAAGAACTCGTCAAGAAGGCGATAGAAGGCGATGC
GCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTT
CAGCAATATCACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGAT
GAATCCAGAAAAGCGGCCATTTTCCACCATGATATTCGGCAAGCAGGCATCGCCATGGGTCACGACGAG
ATCCTCGCCGTCGGGCATGCGCGCCTTGAGCCTGGCGAACAGTTCGGCTGGCGCGAGCCCCTGATGCTC
TTCGTCCAGATCATCCTGATCGACAAGACCGGCTTCCATCCGAGTACGTGCTCGCTCGATGCGATGTTTCG
CTTGGTGGTCGAATGGGCAGGTAGCCGGATCAAGCGTATGCAGCCGCCGCATTGCATCAGCCATGATGG
ATACTTTCTCGGCAGGAGCAAGGTGAGATGACAGGAGATCCTGCCCCGGCACTTCGCCCAATAGCAGCC
AGTCCCTTCCCGCTTCAGTGACAACGTCGAGCACAGCTGCGCAAGGAACGCCCGTCGTGGCCAGCCACG
ATAGCCGCGCTGCCTCGTCCTGCAGTTCATTCAGGGCACCGGACAGGTCGGTCTTGACAAAAAGAACCG
GGCGCCCCTGCGCTGACAGCCGGAACACGGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCCAGTCAT
AGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAATCCATCTTGTTCAATCATGCGAAA
CGATCCTCATCCTGTCTCTTGATCAGATCTTGATCCCCTGCGCCATCAGATCCTTGGCGGCAAGAAAGCCA
TCCAGTTTACTTTGCAGGGCTTCCCAACCTTACCAGAGGGCGCCCCAGCTGGCAATTCCGACGTCTAAGA
AACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAC

colE1 origin

Kan resistance cassette

prfB

pBAD promoter

Appendix B

Chapter 4 supplemental data

Sequences

In vitro assay sfGFP template (TAA)

atataGCGAATTAATACGACTCACTATAGGGCTTAAGTATAAGGAGGAAAAAAT**ATGCGTAAAG**
GTGAAGAACTATTCACCGGTGTTGTTCCGATCCTGGTTGAACTGGACGGTGACGTTAACGGTC
ACAAATTCTCTGTTCGTGGTGAAGGTGAAGGTGACGCTACCAACGGTAAACTGACCCTGAAA
TTCATCTGCACCACCGGTAAACTGCCGGTTCCGTGGCCGACCCTGGTTACCACCCTGACCTACG
GTGTTCAGTGCTTCGCTCGTTACCCGGACCACATGAAACAGCACGACTTCTTCAAATCTGCTAT
GCCGGAAGGTTACGTTCAGGAACGTACCATCTCTTTCAAAGACGACGGTACCTACAAAACCCG
TGCTGAAGTTAAATTCGAAGGTGACACCCTGGTTAACCGTATCGAACTGAAAGGTATCGACTT
CAAAGAAGACGGTAACATCCTGGGTCACAAACTGGAATACAACTTCAACTCTCACAACGTTTA
CATCACCGCTGACAAACAGAAAAACGGTATCAAAGCTAACTTCAAAATCCGTCACAACGTTGA
AGACGGTTCTGTTCAGCTGGCTGACCACTACCAGCAGAACACCCCGATCGGTGACGGTCCGGT
TCTGCTGCCGGACAACCACTACCTGTCTACCCAGTCTGTTCTGTCTAAAGACCCGAACGAAAAA
CGTGACCACATGGTTCTACTAGAATTCGTTACCGCTGCTGGTATCACCCACGGTATGGACGAA
CTATAA*TAATAACTAGCATAACCCCTCTCTAAACGGAGGGGTTT*

bolded sequence is sfGFP ORF

yellow highlight is the T7 promoter

blue highlight is the RBS

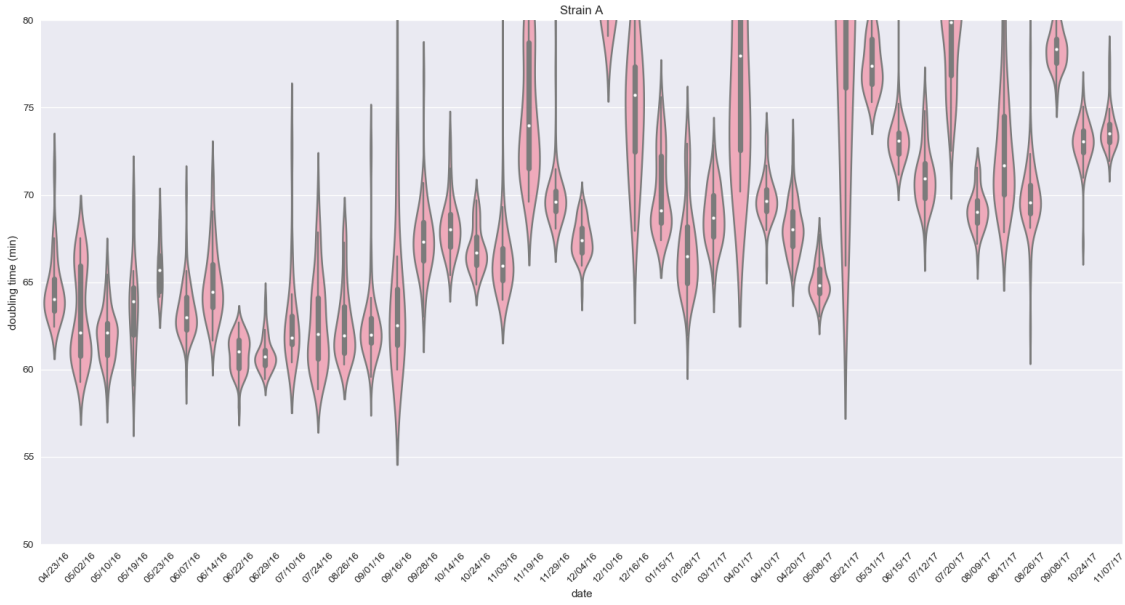
italicized is the terminator sequence

In vitro assay RF1 template (wt)

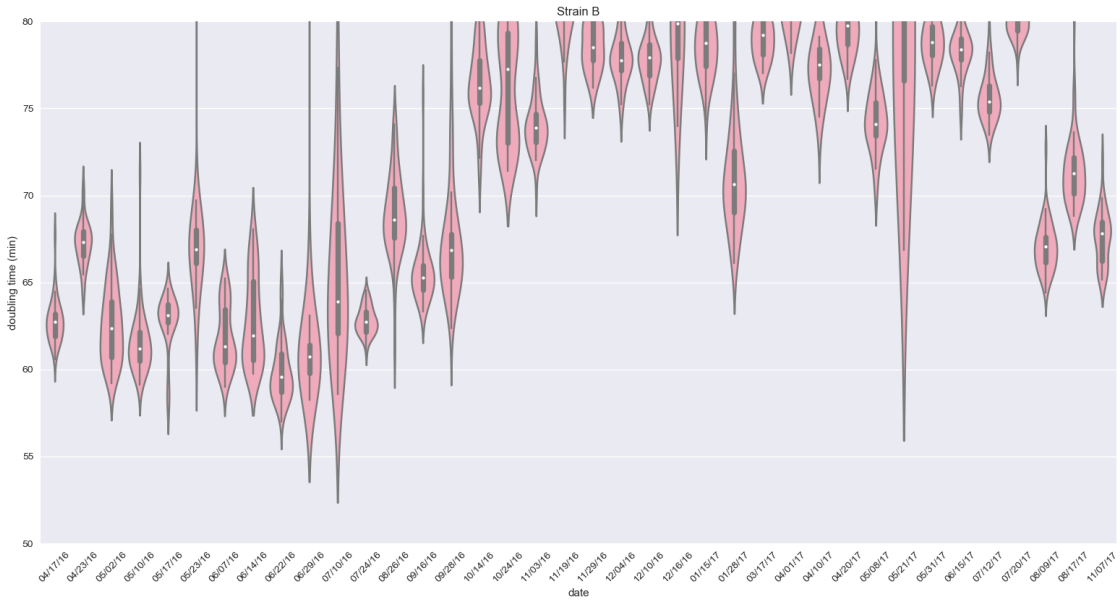
atataGCGAATTAATACGACTCACTATAGGGCTTAAGTATAAGGAGGAAAAAAT**atgaagccttctat**
cgttgccaaactggaagccctgcatgaacgccatgaagaagttcaggcgttgctgggtgacgcgcaaactatcgccgacca
ggaacgttttcgcgcattatcacgcgaatatgcgcagttaagtgatgtttcgcgctgttttaccgactggcaacaggttcagg
aagatatcgaaaccgcacagatgatgctcgatgatcctgaaatgcgtgagatggcgcaggatgaactgcgcgaagctaaa
gaaaaaagcgagcaactggaacagcaattacaggttctgttactgccaaaagatcctgatgacgaacgtaacgccttcctc
gaagtccgagccggaaccggcggcgacgaagcggcgctgttcgcgggcgatctgttccgtatgtacagccgttatgccgaa
gcccgccgctggcgggtagaaatcatgagcgccagcgagggtgaacatggtggttataaagagatcatcgccaaaattag
cggtgatggtgtgtatggtcgtctgaaatttgaatccggcggtcatcgcgtgcaacgtgttcctgctacggaatcgcagggtc
gtattcatacttctgcttgtaccgttgcggtaatgccagaactgcctgacgcagaactgccggacatcaacccagcagattta
cgcattgatactttccgctcgtcaggggcgggtggtcagcacgttaacaccaccgattcggcaattcgtattactcacttgccg
accgggattgttgttgaatgtcaggacgaacgttcacaacataaaaacaaagctaaagcactttctgttctcggtgctcgca
tccacgctgctgaaatggcaaaacgccaacaggccgaagcgtctacccgtcgtaacctgctggggagtggcgatcgcagcg
accgtaaccgtacttacaacttcccgcaggggcgcgttaccgatcaccgcatcaacctgacgctctaccgcctggatgaagt
gatggaaggtaagctggatatgctgattgaaccgattatccaggaacatcaggccgaccaactggcggcgttgtccgagca
ggaataa *TAACTAGCATAACCCCTCTCTAAACGGAGGGGTTT*

Appendix C

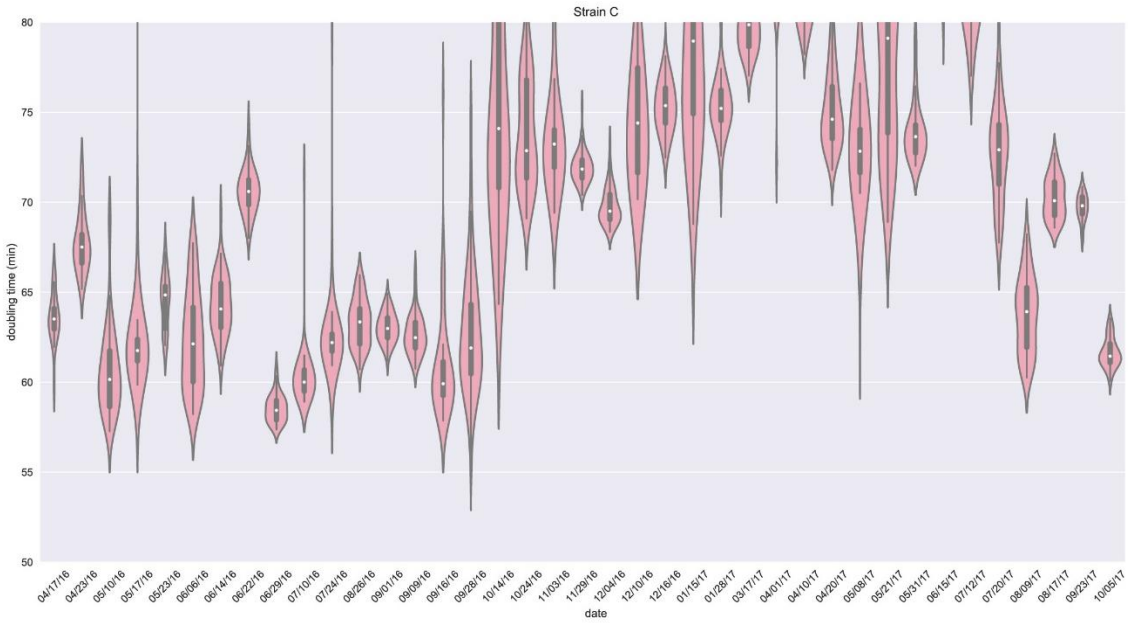
Chapter 5 supplemental data



Supplemental Figure 1. Violin plot of doubling time of population of picked colonies for strain A over time.

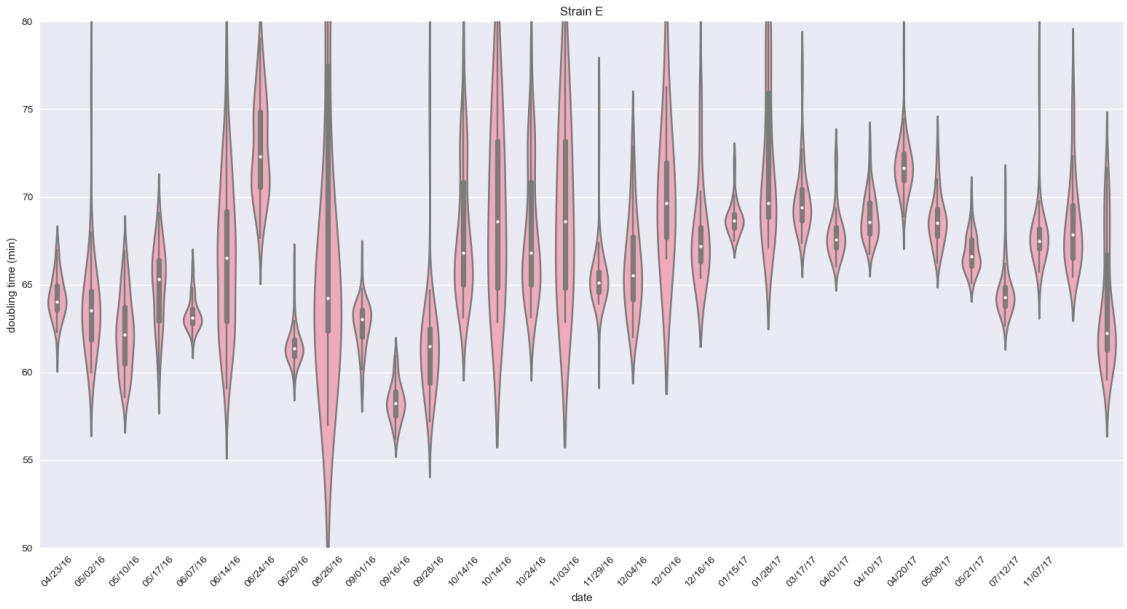


Supplemental Figure 2. Violin plot of doubling time of population of picked colonies for strain B over time.

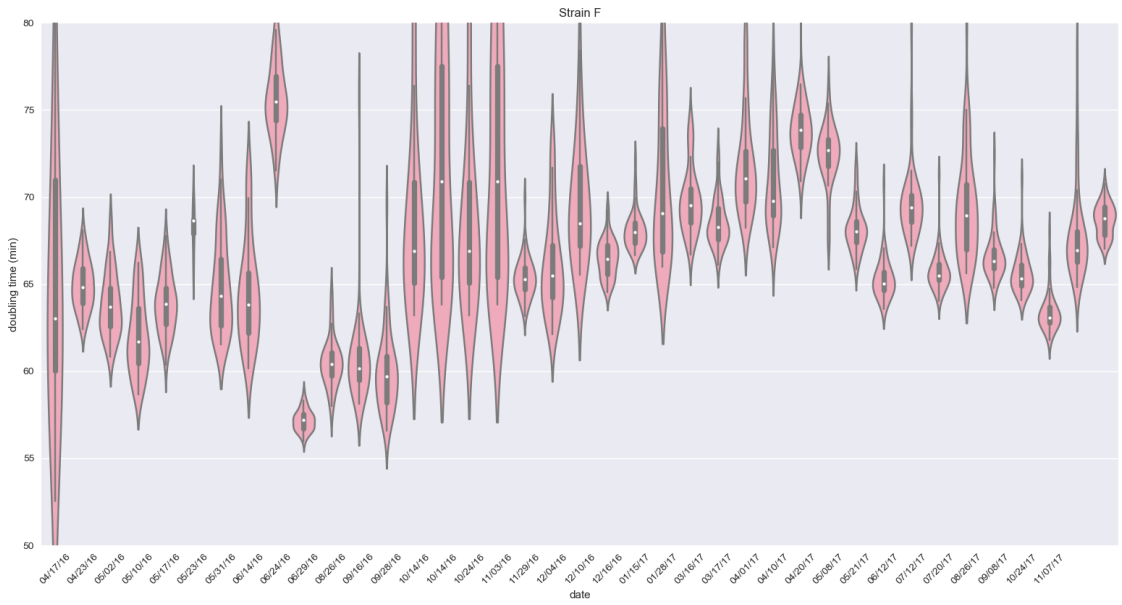


Supplemental Figure 3. Violin plot of doubling time of population of picked colonies for strain C over time.

Supplemental Figure 4. Violin plot of doubling time of population of picked colonies for strain D over time.



Supplemental Figure 5. Violin plot of doubling time of population of picked colonies for strain E over time.



Supplemental Figure 6. Violin plot of doubling time of population of picked colonies for strain F over time.

References

1 Crick, F. Central dogma of molecular biology. *Nature* **227**, 561-563 (1970).

2 Jukes, T. H. & Osawa, S. Evolutionary changes in the genetic code. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* **106**, 489-494, doi:10.1016/0305-0491(93)90122-l (1993).

3 Ibba, M. & Soll, D. Aminoacyl-tRNA synthesis. *Annu Rev Biochem* **69**, 617-650, doi:10.1146/annurev.biochem.69.1.617 (2000).

4 Chapeville, F. *et al.* On the role of soluble ribonucleic acid in coding for amino acids. *Proc Natl Acad Sci U S A* **48**, 1086-1092 (1962).

5 Brown, C. M. & Tate, W. P. Direct recognition of mRNA stop signals by Escherichia coli polypeptide chain release factor two. *J Biol Chem* **269**, 33164-33170 (1994).

6 Craigen, W. J., Lee, C. C. & Caskey, C. T. Recent advances in peptide chain termination. *Mol Microbiol* **4**, 861-865 (1990).

7 Youngman, E. M., McDonald, M. E. & Green, R. Peptide release on the ribosome: mechanism and implications for translational control. *Annu Rev Microbiol* **62**, 353-373, doi:10.1146/annurev.micro.61.080706.093323 (2008).

8 Capecchi, M. R. & Klein, H. A. Release factors mediating termination of complete proteins. *Nature* **226**, 1029-1033 (1970).

9 Dittmar, K. A., Sorensen, M. A., Elf, J., Ehrenberg, M. & Pan, T. Selective charging of tRNA isoacceptors induced by amino-acid starvation. *EMBO Rep* **6**, 151-157, doi:10.1038/sj.embor.7400341 (2005).

10 Elf, J., Nilsson, D., Tenson, T. & Ehrenberg, M. Selective charging of tRNA isoacceptors explains patterns of codon usage. *Science* **300**, 1718-1722, doi:10.1126/science.1083811 (2003).

11 Knight, R. D., Freeland, S. J. & Landweber, L. F. A simple model based on mutation and selection explains trends in codon and amino-acid usage and GC composition within and across genomes. *Genome Biology* **2**, doi:ARTN 0010.1

DOI 10.1186/gb-2001-2-4-research0010 (2001).

12 Chen, S. L., Lee, W., Hottes, A. K., Shapiro, L. & McAdams, H. H. Codon usage between genomes is constrained by genome-wide mutational processes. *Proc Natl Acad Sci U S A* **101**, 3480-3485, doi:10.1073/pnas.0307827100 (2004).

13 Tuller, T. *et al.* An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* **141**, 344-354, doi:10.1016/j.cell.2010.03.031 (2010).

14 Cannarozzi, G. *et al.* A role for codon order in translation dynamics. *Cell* **141**, 355-367, doi:10.1016/j.cell.2010.02.036 (2010).

15 Brule, C. E. & Grayhack, E. J. Synonymous Codons: Choose Wisely for Expression. *Trends Genet* **33**, 283-297, doi:10.1016/j.tig.2017.02.001 (2017).

16 Elf, J. & Ehrenberg, M. What makes ribosome-mediated transcriptional attenuation sensitive to amino acid limitation? *PLoS Comput Biol* **1**, e2, doi:10.1371/journal.pcbi.0010002 (2005).

17 Chaney, J. L. *et al.* Widespread position-specific conservation of synonymous rare codons within coding sequences. *PLoS Comput Biol* **13**, e1005531, doi:10.1371/journal.pcbi.1005531 (2017).

18 Pechmann, S. & Frydman, J. Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat Struct Mol Biol* **20**, 237-243, doi:10.1038/nsmb.2466 (2013).
19 Wei, Y., Wang, J. & Xia, X. Coevolution between Stop Codon Usage and Release Factors in Bacterial Species. *Mol Biol Evol* **33**, 2357-2367, doi:10.1093/molbev/msw107 (2016).
20 Povolotskaya, I. S., Kondrashov, F. A., Ledda, A. & Vlasov, P. K. Stop codons in bacteria are not selectively equivalent. *Biol Direct* **7**, 30, doi:10.1186/1745-6150-7-30 (2012).
21 Sambrook, J. F., Fan, D. P. & Brenner, S. A strong suppressor specific for UGA. *Nature* **214**, 452-453 (1967).
22 Strigini, P. & Brickman, E. Analysis of specific misreading in Escherichia coli. *J Mol Biol* **75**, 659-672 (1973).
23 Ryden, S. M. & Isaksson, L. A. A Temperature-Sensitive Mutant of Escherichia-Coli That Shows Enhanced Misreading of Uag/a and Increased Efficiency for Some Transfer-Rna Nonsense Suppressors. *Mol Gen Genet* **193**, 38-45, doi:Doi 10.1007/Bf00327411 (1984).
24 Fan, Y. *et al.* Heterogeneity of Stop Codon Readthrough in Single Bacterial Cells and Implications for Population Fitness. *Mol Cell* **67**, 826-836 e825, doi:10.1016/j.molcel.2017.07.010 (2017).
25 Baggett, N. E., Zhang, Y. & Gross, C. A. Global analysis of translation termination in E. coli. *PLoS Genet* **13**, e1006676, doi:10.1371/journal.pgen.1006676 (2017).
26 Lajoie, M. J. *et al.* Genomically recoded organisms expand biological functions. *Science* **342**, 357-360, doi:10.1126/science.1241459 (2013).
27 Isaacs, F. J. *et al.* Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science* **333**, 348-353, doi:10.1126/science.1205822 (2011).
28 Amiram, M. *et al.* Evolution of translation machinery in recoded bacteria enables multi-site incorporation of nonstandard amino acids. *Nat Biotechnol* **33**, 1272-1279, doi:10.1038/nbt.3372 (2015).
29 Ma, N. J. & Isaacs, F. J. Genomic Recoding Broadly Obstructs the Propagation of Horizontally Transferred Genetic Elements. *Cell Syst* **3**, 199-207, doi:10.1016/j.cels.2016.06.009 (2016).
30 Ma, N. J., Moonan, D. W. & Isaacs, F. J. Precise manipulation of bacterial chromosomes by conjugative assembly genome engineering. *Nat Protoc* **9**, 2285-2300, doi:10.1038/nprot.2014.081 (2014).
31 Mandell, D. J. *et al.* Biocontainment of genetically modified organisms by synthetic protein design. *Nature* **518**, 55-60, doi:10.1038/nature14121 (2015).
32 Rovner, A. J. *et al.* Recoded organisms engineered to depend on synthetic amino acids. *Nature* **518**, 89-93, doi:10.1038/nature14095 (2015).
33 Nikic, I. & Lemke, E. A. Genetic code expansion enabled site-specific dual-color protein labeling: superresolution microscopy and beyond. *Curr Opin Chem Biol* **28**, 164-173, doi:10.1016/j.cbpa.2015.07.021 (2015).
34 Wang, K. *et al.* Optimized orthogonal translation of unnatural amino acids enables spontaneous protein double-labelling and FRET. *Nat Chem* **6**, 393-403, doi:10.1038/nchem.1919 (2014).
35 Xiao, H. *et al.* Genetic incorporation of multiple unnatural amino acids into proteins in mammalian cells. *Angew Chem Int Ed Engl* **52**, 14080-14083, doi:10.1002/anie.201308137 (2013).
36 Zheng, Y., Gilgenast, M. J., Hauc, S. & Chatterjee, A. Capturing Post-Translational Modification-Triggered Protein-Protein Interactions Using Dual Noncanonical Amino

Acid Mutagenesis. *ACS Chem Biol* **13**, 1137-1141, doi:10.1021/acschembio.8b00021 (2018).
37 Alkatib, S. *et al.* The contributions of wobbling and superwobbling to the reading of the genetic code. *PLoS Genet* **8**, e1003076, doi:10.1371/journal.pgen.1003076 (2012).
38 Crick, F. H. Codon--anticodon pairing: the wobble hypothesis. *J Mol Biol* **19**, 548-555 (1966).
39 Tuorto, F. & Lyko, F. Genome recoding by tRNA modifications. *Open Biol* **6**, doi:10.1098/rsob.160287 (2016).
40 Mukai, T. *et al.* Reassignment of a rare sense codon to a non-canonical amino acid in Escherichia coli. *Nucleic Acids Res* **43**, 8111-8122, doi:10.1093/nar/gkv787 (2015).
41 Wang, K. *et al.* Defining synonymous codon compression schemes by genome recoding. *Nature* **539**, 59-64, doi:10.1038/nature20124 (2016).
42 Lajoie, M. J. *et al.* Probing the limits of genetic recoding in essential genes. *Science* **342**, 361-363, doi:10.1126/science.1241460 (2013).
43 Napolitano, M. G. *et al.* Emergent rules for codon choice elucidated by editing rare arginine codons in Escherichia coli. *Proc Natl Acad Sci U S A* **113**, E5588-5597, doi:10.1073/pnas.1605856113 (2016).
44 Ostrov, N. *et al.* Design, synthesis, and testing toward a 57-codon genome. *Science* **353**, 819-822, doi:10.1126/science.aaf3639 (2016).
45 Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**, 2006 0008, doi:10.1038/msb4100050 (2006).
46 Richmond, C. Genome-wide expression profiling in Escherichia coli K-12. *Nucleic Acids Research* **27**, 3821-3835, doi:10.1093/nar/27.19.3821 (1999).
47 Yang, L. *et al.* Genome-wide inactivation of porcine endogenous retroviruses (PERVs). *Science* **350**, 1101-1104, doi:10.1126/science.aad1191 (2015).
48 Mukai, T. *et al.* Highly reproductive Escherichia coli cells with no specific assignment to the UAG codon. *Sci Rep* **5**, 9699, doi:10.1038/srep09699 (2015).
49 Annaluru, N. *et al.* Total synthesis of a functional designer eukaryotic chromosome. *Science* **344**, 55-58, doi:10.1126/science.1249252 (2014).
50 Gibson, D. G. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329**, 52-56, doi:10.1126/science.1190719 (2010).
51 Karp, R. M. Reducibility among Combinatorial Problems. 85-103, doi:10.1007/978-1-4684-2001-2_9 (1972).
52 Vazirani, V. V. *Approximation Algorithms*. (Springer-Verlag Berlin Heidelberg, 2003).
53 Chvatal, V. A Greedy Heuristic for the Set-Covering Problem. *Mathematics of Operations Research* **4**, 233-235, doi:10.1287/moor.4.3.233 (1979).
54 Boeke, J. D. *et al.* GENOME ENGINEERING. The Genome Project-Write. *Science* **353**, 126-127, doi:10.1126/science.aaf6850 (2016).
55 *gBlocks® Gene Fragments*, <https://www.idtdna.com/pages/products/genes-and-gene-fragments/gblocks-gene-fragments> (
56 *Gene Synthesis Service*, <https://www.genscript.com/gene_synthesis.html> (
57 Lau, Y. H. *et al.* Large-scale recoding of a bacterial genome by iterative recombineering of synthetic DNA. *Nucleic Acids Res* **45**, 6971-6980, doi:10.1093/nar/gkx415 (2017).
58 Curry, K. A. & Tomich, C. S. Effect of ribosome binding site on gene expression in Escherichia coli. *DNA* **7**, 173-179, doi:10.1089/dna.1988.7.173 (1988).
59 Espah Borujeni, A., Channarasappa, A. S. & Salis, H. M. Translation rate is controlled by coupled trade-offs between site accessibility, selective RNA unfolding and sliding at

upstream standby sites. *Nucleic Acids Res* **42**, 2646-2659, doi:10.1093/nar/gkt1139 (2014).
60 Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* **27**, 946-950, doi:10.1038/nbt.1568 (2009).
61 Gallagher, R. R., Li, Z., Lewis, A. O. & Isaacs, F. J. Rapid editing and evolution of bacterial genomes using libraries of synthetic DNA. *Nat Protoc* **9**, 2301-2316, doi:10.1038/nprot.2014.082 (2014).
62 Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**, 894-898, doi:10.1038/nature08187 (2009).
63 Gregg, C. J. *et al.* Rational optimization of tolC as a powerful dual selectable marker for genome engineering. *Nucleic Acids Res* **42**, 4779-4790, doi:10.1093/nar/gkt1374 (2014).
64 Kolisnychenko, V. *et al.* Engineering a reduced Escherichia coli genome. *Genome Res* **12**, 640-647, doi:10.1101/gr.217202 (2002).
65 Tsuchihashi, Z. & Kornberg, A. Translational frameshifting generates the gamma subunit of DNA polymerase III holoenzyme. *Proc Natl Acad Sci U S A* **87**, 2516-2520 (1990).
66 Prather, N. E., Murgola, E. J. & Mims, B. H. Primary structure of an unusual glycine tRNA UGA suppressor. *Nucleic Acids Res* **9**, 6421-6428 (1981).
67 Heinemann, I. U. *et al.* Enhanced phosphoserine insertion during Escherichia coli protein synthesis via partial UAG codon reassignment and release factor 1 deletion. *FEBS Lett* **586**, 3716-3722, doi:10.1016/j.febslet.2012.08.031 (2012).
68 Zhou, J. & Rudd, K. E. EcoGene 3.0. *Nucleic Acids Res* **41**, D613-624, doi:10.1093/nar/gks1235 (2013).
69 Lorenz, R. *et al.* ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**, 26, doi:10.1186/1748-7188-6-26 (2011).
70 SantaLucia, J. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences* **95**, 1460-1465, doi:10.1073/pnas.95.4.1460 (1998).
71 <http://arep.med.harvard.edu/kzhang/cgi-bin/myOligoTm.cgi> (
72 Lajoie, M. J., Gregg, C. J., Mosberg, J. A., Washington, G. C. & Church, G. M. Manipulating replisome dynamics to enhance lambda Red-mediated multiplex genome engineering. *Nucleic Acids Res* **40**, e170, doi:10.1093/nar/gks751 (2012).
73 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).
74 H, L. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997v1 [q-bio.GN]* (2013).
75 *https://github.com/broadinstitute/picard*.
76 Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**, 11 10 11-33, doi:10.1002/0471250953.bi1110s43 (2013).
77 Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80-92, doi:10.4161/fly.19695 (2012).
78 Ito, K., Uno, M. & Nakamura, Y. A tripeptide 'anticodon' deciphers stop codons in messenger RNA. *Nature* **403**, 680-684, doi:10.1038/35001115 (2000).

79 Uno, M., Ito, K. & Nakamura, Y. Polypeptide release at sense and noncognate stop codons by localized charge-exchange alterations in translational release factors. *Proc Natl Acad Sci U S A* **99**, 1819-1824, doi:10.1073/pnas.032457599 (2002).
80 Kisselev, L. Polypeptide Release Factors in Prokaryotes and Eukaryotes. *Structure* **10**, 8-9, doi:10.1016/s0969-2126(01)00703-1 (2002).
81 Lekomtsev, S. *et al.* Different modes of stop codon restriction by the Stylonychia and Paramecium eRF1 translation termination factors. *Proc Natl Acad Sci U S A* **104**, 10824-10829, doi:10.1073/pnas.0703887104 (2007).
82 Seit-Nebi, A., Frolova, L. & Kisselev, L. Conversion of omnipotent translation termination factor eRF1 into ciliate-like UGA-only unipotent eRF1. *EMBO Rep* **3**, 881-886, doi:10.1093/embo-reports/kvf178 (2002).
83 Lekomtsev, S. A. Nonstandard genetic codes and translation termination. *Molecular Biology* **41**, 878-885, doi:10.1134/s0026893307060027 (2007).
84 McCutcheon, J. P., McDonald, B. R. & Moran, N. A. Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. *PLoS Genet* **5**, e1000565, doi:10.1371/journal.pgen.1000565 (2009).
85 Campbell, J. H. *et al.* UGA is an additional glycine codon in uncultured SR1 bacteria from the human microbiota. *Proc Natl Acad Sci U S A* **110**, 5540-5545, doi:10.1073/pnas.1303090110 (2013).
86 Inagaki, Y., Bessho, Y. & Osawa, S. Lack of peptide-release activity responding to codon UGA in Mycoplasma capricolum. *Nucleic Acids Res* **21**, 1335-1338 (1993).
87 Kannan, T. R. & Baseman, J. B. Expression of UGA-Containing Mycoplasma Genes in Bacillus subtilis. *Journal of Bacteriology* **182**, 2664-2667, doi:10.1128/jb.182.9.2664-2667.2000 (2000).
88 Grosjean, H. *et al.* Predicting the minimal translation apparatus: lessons from the reductive evolution of mollicutes. *PLoS Genet* **10**, e1004363, doi:10.1371/journal.pgen.1004363 (2014).
89 Duarte, I., Nabuurs, S. B., Magno, R. & Huynen, M. Evolution and diversification of the organellar release factor family. *Molecular biology and evolution* **29**, 3497-3512, doi:10.1093/molbev/mss157 (2012).
90 Young, D. J., Edgar, C. D., Poole, E. S. & Tate, W. P. The codon specificity of eubacterial release factors is determined by the sequence and size of the recognition loop. *Rna* **16**, 1623-1633, doi:10.1261/rna.2117010 (2010).
91 Young, D. J. *et al.* Bioinformatic, structural, and functional analyses support release factor-like MTRF1 as a protein able to decode nonstandard stop codons beginning with adenine in vertebrate mitochondria. *Rna* **16**, 1146-1155, doi:10.1261/rna.1970310 (2010).
92 Korostelev, A., Zhu, J., Asahara, H. & Noller, H. F. Recognition of the amber UAG stop codon by release factor RF1. *EMBO J* **29**, 2577-2585, doi:10.1038/emboj.2010.139 (2010).
93 Laurberg, M. *et al.* Structural basis for translation termination on the 70S ribosome. *Nature* **454**, 852-857, doi:10.1038/nature07115 (2008).
94 Nakamura, Y. & Ito, K. A tripeptide discriminator for stop codon recognition. *FEBS letters* **514**, 30-33, doi:10.1016/s0014-5793(02)02330-x (2002).
95 Ma N, Hemez C, Barber K, Rinehart J & FJ., I. *Organisms with alternative genetic codes resolve unassigned codons vis mistranslation and ribosomal rescue* (2018).
96 Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nat Biotechnol* **19**, 751-755, doi:10.1038/90802 (2001).

97 Uemura, S. *et al.* Single-molecule imaging of full protein synthesis by immobilized ribosomes. *Nucleic acids research* **36**, e70, doi:10.1093/nar/gkn338 (2008).
98 Johnson, D. B. *et al.* Release factor one is nonessential in Escherichia coli. *ACS Chem Biol* **7**, 1337-1344, doi:10.1021/cb300229q (2012).
99 Johnson, D. B. *et al.* RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. *Nat Chem Biol* **7**, 779-786, doi:10.1038/nchembio.657 (2011).
100 Dirla, S., Chien, J. Y. & Schleif, R. Constitutive mutations in the Escherichia coli AraC protein. *J Bacteriol* **191**, 2668-2674, doi:10.1128/JB.01529-08 (2009).
101 Uno, M., Ito, K. & Nakamura, Y. Functional specificity of amino acid at position 246 in the tRNA mimicry domain of bacterial release factor 2. *Biochimie* **78**, 935-943, doi:10.1016/s0300-9084(97)86715-6 (1996).
102 Chadani, Y., Ito, K., Kutsukake, K. & Abo, T. ArfA recruits release factor 2 to rescue stalled ribosomes by peptidyl-tRNA hydrolysis in Escherichia coli. *Mol Microbiol* **86**, 37-50, doi:10.1111/j.1365-2958.2012.08190.x (2012).
103 Petropoulos, A. D., McDonald, M. E., Green, R. & Zaher, H. S. Distinct roles for release factor 1 and release factor 2 in translational quality control. *J Biol Chem* **289**, 17589-17596, doi:10.1074/jbc.M114.564989 (2014).
104 Zaher, H. S. & Green, R. Quality control by the ribosome following peptide bond formation. *Nature* **457**, 161-166, doi:10.1038/nature07582 (2009).
105 Nakahigashi, K. *et al.* HemK, a class of protein methyl transferase with similarity to DNA methyl transferases, methylates polypeptide chain release factors, and hemK knockout induces defects in translational termination. *Proc Natl Acad Sci U S A* **99**, 1473-1478, doi:10.1073/pnas.032488499 (2002).
106 Dreyfus, M. & Heurgue-Hamard, V. Termination troubles in Escherichia coli K12. *Mol Microbiol* **79**, 288-291, doi:10.1111/j.1365-2958.2010.07468.x (2011).
107 Mora, L., Heurgue-Hamard, V., de Zamaroczy, M., Kervestin, S. & Buckingham, R. H. Methylation of bacterial release factors RF1 and RF2 is required for normal translation termination in vivo. *J Biol Chem* **282**, 35638-35645, doi:10.1074/jbc.M706076200 (2007).
108 Florin, T. *et al.* An antimicrobial peptide that inhibits translation by trapping release factors on the ribosome. *Nat Struct Mol Biol* **24**, 752-757, doi:10.1038/nsmb.3439 (2017).
109 Freistroffer, D. V., Kwiatkowski, M., Buckingham, R. H. & Ehrenberg, M. The accuracy of codon recognition by polypeptide release factors. *Proc Natl Acad Sci U S A* **97**, 2046-2051, doi:10.1073/pnas.030541097 (2000).
110 Hetrick, B., Lee, K. & Joseph, S. Kinetics of stop codon recognition by release factor 1. *Biochemistry* **48**, 11178-11184, doi:10.1021/bi901577d (2009).
111 Field, A., Hetrick, B., Mathew, M. & Joseph, S. Histidine 197 in release factor 1 is essential for a site binding and peptide release. *Biochemistry* **49**, 9385-9390, doi:10.1021/bi1012047 (2010).
112 He, S. L. & Green, R. Visualization of codon-dependent conformational rearrangements during translation termination. *Nat Struct Mol Biol* **17**, 465-470, doi:10.1038/nsmb.1766 (2010).
113 Svidritskiy, E. & Korostelev, A. A. Conformational Control of Translation Termination on the 70S Ribosome. *Structure* **26**, 821-828 e823, doi:10.1016/j.str.2018.04.001 (2018).
114 Trappl, K. & Joseph, S. Ribosome Induces a Closed to Open Conformational Change in Release Factor 1. *J Mol Biol* **428**, 1333-1344, doi:10.1016/j.jmb.2016.01.021 (2016).

115 Trappl, K., Mathew, M. A. & Joseph, S. Thermodynamic and kinetic insights into stop codon recognition by release factor 1. *PLoS One* **9**, e94058, doi:10.1371/journal.pone.0094058 (2014).
116 Hoernes, T. P. *et al.* Atomic mutagenesis of stop codon nucleotides reveals the chemical prerequisites for release factor-mediated peptide release. *Proc Natl Acad Sci U S A* **115**, E382-E389, doi:10.1073/pnas.1714554115 (2018).
117 Csorgo, B., Feher, T., Timar, E., Blattner, F. R. & Posfai, G. Low-mutation-rate, reduced-genome Escherichia coli: an improved host for faithful maintenance of engineered genetic constructs. *Microb Cell Fact* **11**, 11, doi:10.1186/1475-2859-11-11 (2012).
118 Umenhoffer, K. *et al.* Reduced evolvability of Escherichia coli MDS42, an IS-less cellular chassis for molecular and synthetic biology applications. *Microb Cell Fact* **9**, 38, doi:10.1186/1475-2859-9-38 (2010).
119 DeVito, J. A. Recombineering with tolC as a selectable/counter-selectable marker: remodeling the rRNA operons of Escherichia coli. *Nucleic Acids Res* **36**, e4, doi:10.1093/nar/gkm1084 (2008).
120 Sharan, S. K., Thomason, L. C., Kuznetsov, S. G. & Court, D. L. Recombineering: a homologous recombination-based method of genetic engineering. *Nat Protoc* **4**, 206-223, doi:10.1038/nprot.2008.227 (2009).
121 Ghosh, H. P., Ghosh, K. & Ganoza, M. C. Mechanism of polypeptide chain termination. Translation of tandem amber termination codons by an amber suppressor transfer ribonucleic acid. *J Biol Chem* **247**, 5322-5326 (1972).
122 Poole, E. S., Major, L. L., Mannering, S. A. & Tate, W. P. Translational termination in Escherichia coli: three bases following the stop codon crosslink to release factor 2 and affect the decoding efficiency of UGA-containing signals. *Nucleic Acids Res* **26**, 954-960 (1998).
123 Major, L. L., Edgar, T. D., Yee Yip, P., Isaksson, L. A. & Tate, W. P. Tandem termination signals: myth or reality? *FEBS Letters* **514**, 84-89, doi:10.1016/s0014-5793(02)02301-3 (2002).
124 Brown, C. M., Stockwell, P. A., Trotman, C. N. & Tate, W. P. The signal for the termination of protein synthesis in procaryotes. *Nucleic Acids Res* **18**, 2079-2086 (1990).
125 Poole, E. S., Brown, C. M. & Tate, W. P. The identity of the base following the stop codon determines the efficiency of in vivo translational termination in Escherichia coli. *EMBO J* **14**, 151-158 (1995).
126 Rudd, K. E. Novel intergenic repeats of Escherichia coli K-12. *Res Microbiol* **150**, 653-664 (1999).
127 Zhang, Y. & Gladyshev, V. N. An algorithm for identification of bacterial selenocysteine insertion sequence elements and selenoprotein genes. *Bioinformatics* **21**, 2580-2589, doi:10.1093/bioinformatics/bti400 (2005).
128 Zinoni, F., Birkmann, A., Stadtman, T. C. & Bock, A. Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate dehydrogenase (formate-hydrogen-lyase-linked) from Escherichia coli. *Proc Natl Acad Sci U S A* **83**, 4650-4654 (1986).
129 Heurgue-Hamard, V., Champ, S., Engstrom, A., Ehrenberg, M. & Buckingham, R. H. The hemK gene in Escherichia coli encodes the N(5)-glutamine methyltransferase that modifies peptide release factors. *EMBO J* **21**, 769-778, doi:10.1093/emboj/21.4.769 (2002).

130 Betney, R., de Silva, E., Krishnan, J. & Stansfield, I. Autoregulatory systems controlling translation factor expression: thermostat-like control of translational accuracy. *RNA* **16**, 655-663, doi:10.1261/rna.1796210 (2010).
131 Wannier, T. M. *et al.* Adaptive evolution of genomically recoded Escherichia coli. *Proc Natl Acad Sci U S A* **115**, 3090-3095, doi:10.1073/pnas.1715530115 (2018).
132 Pierson, W. E. *et al.* Uniformity of Peptide Release Is Maintained by Methylation of Release Factors. *Cell Rep* **17**, 11-18, doi:10.1016/j.celrep.2016.08.085 (2016).
133 Schmidt, A. *et al.* The quantitative and condition-dependent Escherichia coli proteome. *Nat Biotechnol* **34**, 104-110, doi:10.1038/nbt.3418 (2016).
134 Coco, W. M. *et al.* DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nat Biotechnol* **19**, 354-359, doi:10.1038/86744 (2001).
135 Chin, J. W., Martin, A. B., King, D. S., Wang, L. & Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proc Natl Acad Sci U S A* **99**, 11020-11024, doi:10.1073/pnas.172226299 (2002).
136 Italia, J. S. *et al.* An orthogonalized platform for genetic code expansion in both bacteria and eukaryotes. *Nat Chem Biol* **13**, 446-450, doi:10.1038/nchembio.2312 (2017).
137 Kuznetsov, G. *et al.* Optimizing complex phenotypes through model-guided multiplex genome engineering. *Genome Biol* **18**, 100, doi:10.1186/s13059-017-1217-z (2017).
138 Tibshirani, R. Regression Shrinkage and Selection Via the Lasso *Journal of the Royal Statistical Society* **58**, 267-288 (1994).