

Proteostasis modulates gene dosage evolution in antibiotic-resistant bacteria

Reviewed Preprint

v2 • March 4, 2025

Revised by authors

Reviewed Preprint

v1 • October 31, 2024

Chinmaya Jena, Saillesh Chinnaraj, Soham Deolankar, Nishad Matange 

Department of Biology, Indian Institute of Science Education and Research, Pune, India

 https://en.wikipedia.org/wiki/Open_access

 Copyright information

eLife Assessment

This **important** study explores the interplay between gene dosage and gene mutations in the evolution of antibiotic resistance. The authors provide **compelling** evidence connecting proteostasis with gene duplication during experimental evolution in a model system. This paper is likely to be of interest to researchers studying antibiotic resistance, proteostasis, and bacterial evolution.

<https://doi.org/10.7554/eLife.99785.2.sa3>

Abstract

Evolution of gene expression frequently drives antibiotic resistance in bacteria. We had previously (Patel and Matange, *eLife*, 2021) shown that, in *Escherichia coli*, mutations at the *mgrB* locus were beneficial under trimethoprim exposure and led to overexpression of dihydrofolate reductase (DHFR), encoded by the *folA* gene. Here, we show that DHFR levels are further enhanced by spontaneous duplication of a genomic segment encompassing *folA* and spanning hundreds of kilobases. This duplication was rare in wild type *E. coli*. However, its frequency was elevated in a *lon*-knockout strain, altering the mutational landscape early during trimethoprim adaptation. We then exploit this system to investigate the relationship between trimethoprim pressure and *folA* copy number. During long-term evolution, *folA* duplications were frequently reversed. Reversal was slower under antibiotic pressure, first requiring the acquisition of point mutations in DHFR or its promoter. Unexpectedly, despite resistance-conferring point mutations, some populations under high trimethoprim pressure maintained *folA* duplication to compensate for low abundance DHFR mutants. We find that evolution of gene dosage depends on expression demand, which is generated by antibiotic and exacerbated by proteolysis of drug-resistant mutants of DHFR. We propose a novel role for proteostasis as a determinant of copy number evolution in antibiotic-resistant bacteria.

Introduction

The dosage of a gene, i.e., relative copy number, directly impacts its level of expression and can alter organismal phenotype and fitness [1–3]. Mutations that increase copy numbers, termed as gene duplications and amplifications (GDAs), are implicated in diverse biological phenomena including cancer progression [4–6], speciation [7–9], reproductive isolation [10–12] and antimicrobial resistance (AMR) [2, 13]. GDAs are important for rapid adaptation since their frequencies of occurrence can be higher by several orders of magnitude than point mutations [2, 13–16]. Following the initial GDA mutation, however, amplified genes could be lost over time, pseudogenized, sub-functionalized or neo-functionalized [17–19]. Which of these possibilities is realized depends on the nature and strength of selection imposed by the environment, as well as the fitness cost associated with maintaining multiple gene copies [17–19].

In bacteria, GDA mutations are frequently encountered during adaptation to antibiotics [13] or heavy metals [20, 21]. They occur by RecA-dependent nonequal recombination as well as RecA-independent processes and often involve repeat sequences such as in IS-elements [22–25]. Amplified genes in drug-resistant bacteria code for effectors of resistance such as drug inactivating enzymes, drug target modifiers or “by-pass” mechanisms [26–28], but may also compensate for resistance-associated fitness costs [29]. More recently, the view that GDAs are less stable than point mutations has emerged. For instance, transient alterations in gene dosage are a common mechanism of heteroresistance, i.e. resistant sub-populations within a drug-sensitive bacterial isolate [20, 30, 31]. Further, it is estimated that as many as 20% of cells in a bacterial population may harbour a duplication of some genomic region [32]. These copy number polymorphisms create within-population heterogeneities in gene expression, serving as a bet-hedging strategy with adaptive value in fluctuating environments [33].

Dihydrofolate reductase (DHFR) is a ubiquitous enzyme crucial for growth and division. It catalyzes the NADPH-dependent reduction of tetrahydrofolate, an important cofactor in nucleotide and amino acid biosynthesis. DHFR is coded by an essential gene in most organisms and its deletion or pharmacological inhibition results in cell death or growth stasis [34]. Not surprisingly, anti-folates are effective therapeutic agents for diverse pathologies [35, 36]. For instance, methotrexate, used in the treatment of cancers and autoimmune conditions [37], pyrimethamine, used as an anti-malarial [38], and trimethoprim, used as an anti-bacterial [39], are all competitive inhibitors of DHFR. Interestingly, copy number amplification and consequent overexpression of DHFR confers resistance in all three systems [40–43]. Thus, the gene dosage of DHFR appears to be tightly regulated across evolutionarily distant organisms, making it a useful system to understand the principles underlying selection and evolutionary fate of GDA mutations.

In *Escherichia coli* DHFR is coded by *folA*. Several point mutations that confer trimethoprim resistance have been identified in this enzyme [44–47]. Additionally, mechanisms that transcriptionally up-regulate *folA* expression are also known in trimethoprim resistant *E. coli*. These include cis-regulatory mutations in the *folA* promoter and trans-acting loss of function mutations in *mgrB*, which upregulate *folA* transcription in a PhoQP-dependent manner [44–46, 48]. Earlier work from our group identified Lon, a conserved quality control protease in bacteria, as a post-translational regulator of trimethoprim resistance. Lon degraded mutant DHFRs and reduced their half-lives *in vivo*, directly affecting genotype-phenotype relationships of resistance-conferring mutations [49, 50].

In this study, we first compare early mutational events in wild type and Lon-deficient *E. coli* challenged with trimethoprim. We show that a large genomic duplication encompassing *folA* confers trimethoprim resistance in both genetic backgrounds, though its frequency of occurrence is elevated in *lon*-knockout *E. coli*. We then use this system to ask what factors determine the fate of duplicated *folA* gene copies during evolution at different trimethoprim pressures.

Results

A large IS-flanked duplication encompassing *folA* mediates rapid adaptation to trimethoprim in *E. coli*

We had previously evolved 5 replicate populations of wild type *E. coli* K-12 MG1655 in 300 ng/mL of trimethoprim ($\sim 0.3 \times \text{MIC}$) for 25 generations. We then isolated and sequenced the genome of a dominant trimethoprim-resistant clone from each population (designated as WTMPR1-5) (**Figure 1 A**). These sequencing results have been reported earlier [46]. Briefly, all isolates harboured loss-of-function mutations in the *mgrB* gene (**Figure 1 B**) that resulted in PhoQP-dependent overexpression of *folA* [46]. Additionally, isolates WTMPR1, 3 and 5 had mutations at the active site of DHFR, while WTMPR4 harboured a large genomic duplication of ~ 0.6 Mb encompassing *folA* (**Figure 1 B, C**). We performed similar evolution experiments starting with an isogenic *E. coli* Δlon strain to ask if early adaptation to trimethoprim was altered by Lon-deficiency (**Figure 1 A**). Wild type and *lon*-deficient *E. coli* have similar MIC for trimethoprim, though the *lon*-knockout has a higher IC₅₀ [49]. Among trimethoprim-resistant isolates derived from *E. coli* Δlon , designated as LTMPR1-5, three harboured mutations at the *mgrB* locus (LTMPR1-3, **Figure 1 A, B**, **Table 1**). LTMPR4 and LTMPR5 harboured mutations in *pitA*, coding for a metal-phosphate symporter, that have been reported earlier in trimethoprim adapted *E. coli* [51]. LTMPR5 had a deletion encompassing *pitA*, indicating that loss of PitA was beneficial in trimethoprim. A knockout of *pitA* indeed showed greater colony formation on trimethoprim-supplemented media demonstrating that mutations in this gene were adaptive (**Figure S1**), though the underlying molecular mechanisms are unknown at present. Interestingly, none of the LTMPR isolates had mutations in DHFR or its promoter. Instead, LTMPR2-4 had evolved GDA mutations similar to WTMPR4, while LTMPR1 had a duplication in an overlapping but shorter genomic stretch (**Figure 1 B, C**, **Table 1**).

Duplicated genomic regions in WTMPR4 and LTMPR1-4 were flanked by IS-elements (**Figure 1 C**). Seeking to explain the higher frequency of GDA mutations in the trimethoprim-adapted *lon*-knockout, we asked whether IS-transposases were sensitive to proteolysis by Lon. Indeed, heterologously expressed InsB and InsL transposases had greater stability in the *lon*-knockout (**Figure S1**), suggesting that their accumulation may contribute to higher GDA frequency in *E. coli* Δlon . It would follow that a higher proportion of GDA mutants would be isolated from *E. coli* Δlon during adaptation in other environments as well. In agreement, GDA mutations were detected after evolution of *E. coli* Δlon in 2 of the 5 additional growth conditions tested (**Figure 1 D**, **S2**). In all cases duplications were flanked by IS elements (**Figure S2**). Further, duplicated genomic regions were similar among isolates derived from the same growth condition, but different across environments (**Figure 1 B**, **S2**). On the other hand, wild type *E. coli* did not show GDA mutations in any of the other tested conditions (**Figure 1 D**).

The duplicated region in trimethoprim-resistant isolates included genes previously associated with resistance to trimethoprim, namely, *folA*, and *acrA* and *B*, which code for components of the AcrAB-TolC efflux pump (**Figure 1 C**). In one of the isolates, i.e. LTMPR1, the amplified stretch encompassed *folA* but not *acrAB* without affecting trimethoprim IC₅₀ values (**Figure 1 B, C**). We therefore reasoned that duplication of *folA* was the main effector of adaptation to trimethoprim. Immunoblotting showed that DHFR protein was indeed overproduced in isolates LTMPR1 and WTMPR4 over and above the increase mediated by inactivation of *mgrB* alone (**Figure 2 A**).

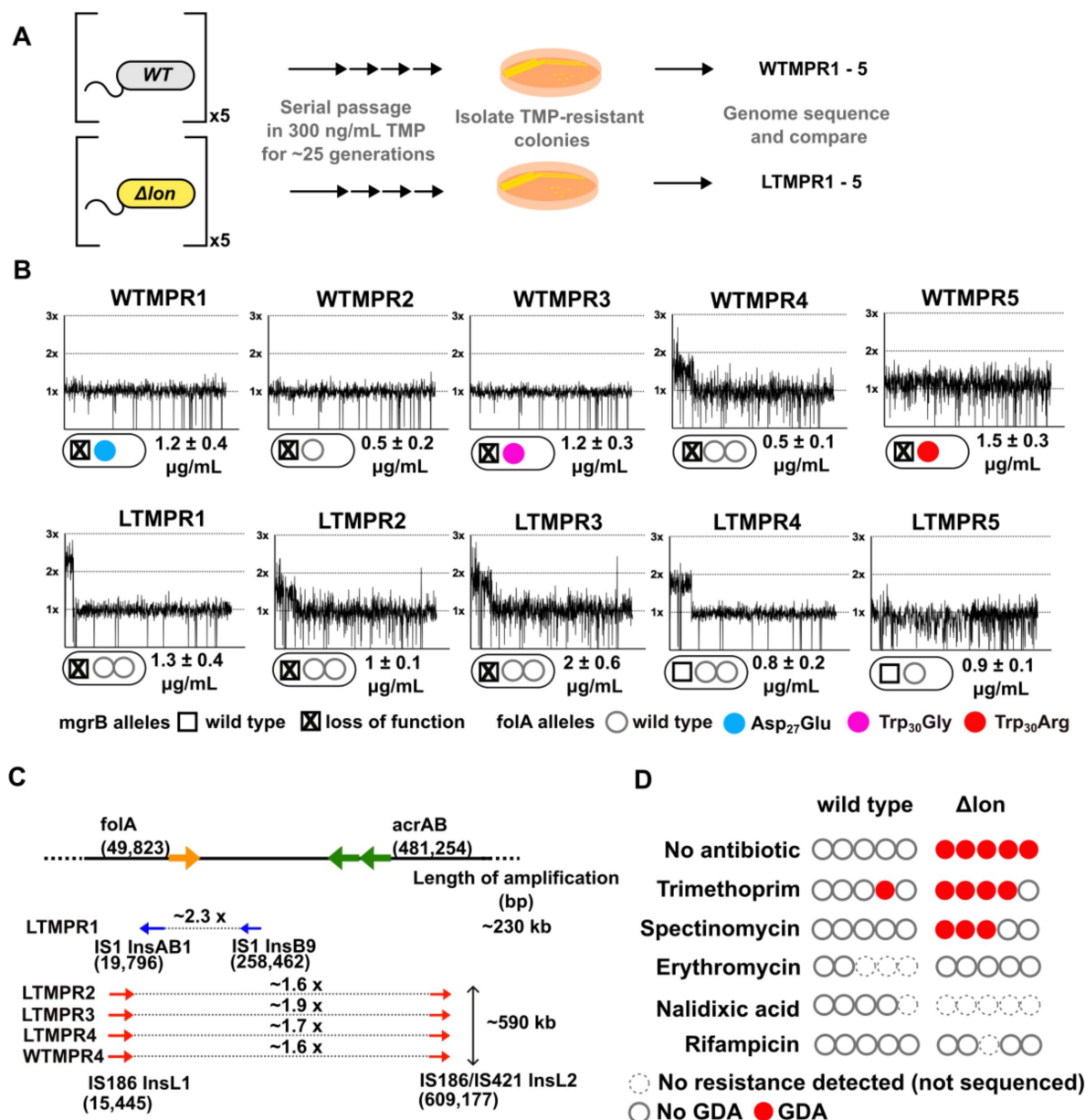


Figure 1.

Adaptation to antibiotics by large genomic duplications is favoured in *lon*- deficient *E. coli*.

A. Schematic of the experimental pipeline used to compare the mutation repertoire of adaptation to trimethoprim in wild type or *lon*-deficient *E. coli*. **B.** Coverage depth plots and mutations at the *folA* and *mgrB* loci of 5 independently evolved trimethoprim-resistant isolates derived from wild type (WTMPR1-5) or Δlon *E. coli* (LTMPR1-5). Coverage depth plots show the number of reads from Illumina short-read sequencing (y-axis) mapped to each genomic coordinate of the reference genome (x-axis). Coverage of 1x, 2x and 3x are marked by dotted lines for reference. Mutations in *mgrB* and *folA*, and trimethoprim IC50 values (mean $\pm$ S.D. from 3 independent measurements) are provided for each isolate below the appropriate coverage depth plot. **C.** Cartoon (not to scale) showing the duplicated genomic stretch and flanking IS-elements in trimethoprim-resistant isolates. Blue arrows represent IS1, while red arrows represent IS186. The positions of *folA* and *acrAB* genes, implicated in trimethoprim resistance in earlier studies, relative to the duplicated stretch are shown. Fold increase in the number of reads corresponding to the *folA* gene are indicated for each isolate. **D.** Summary of the prevalence of GDA mutations detected in wild type or *E. coli* Δlon after ~25 generations of evolution in control media (no antibiotic) or in media supplemented with sub-MIC antibiotics. Each circle represents one of the 5 replicates in the evolution experiment. Empty circles represent replicates in which no resistance was detected (dotted perimeter) or where the whole population or resistant isolates were sequenced but no GDA mutation was detected (solid perimeter). Red filled circles indicate that a GDA mutation was identified.

	LTMPR1	LTMPR2	LTMPR3	LTMPR4	LTMPR5	
Gene/Locus						Description
<i>mgrB</i>	IS5 insertion at -46	IS2 insertion at +82	IS5 insertion at -36	-	-	Negative feedback regulator of PhoQ sensor kinase
<i>ybcK</i>	-	IS2 insertion at +118	-	-	-	Prophage recombinase
<i>pitA</i>	-	-	-	Ser ₄₂₅ Pro	Δ 18 genes [yhiN, <i>pitA</i> ... <i>yhiS</i>]. IS5- mediated	Metal/Phosphate symporter
<i>insH21</i>	Δ <i>insH21</i>	Δ <i>insH21</i>	Δ <i>insH21</i>	Δ <i>insH21</i>	Δ <i>insH21</i>	IS5 transposase
GDA (encompassing <i>folA</i>)	19,796 to 258,462 bp (~2.3x)	15,445 to 609,177 bp (~1.6x)	15,445 to 609,177 bp (~1.9x)	15,445 to 609,177 bp (~1.7x)	-	

Table 1.

Complete list of mutations identified in isolates LTMPR1-5.

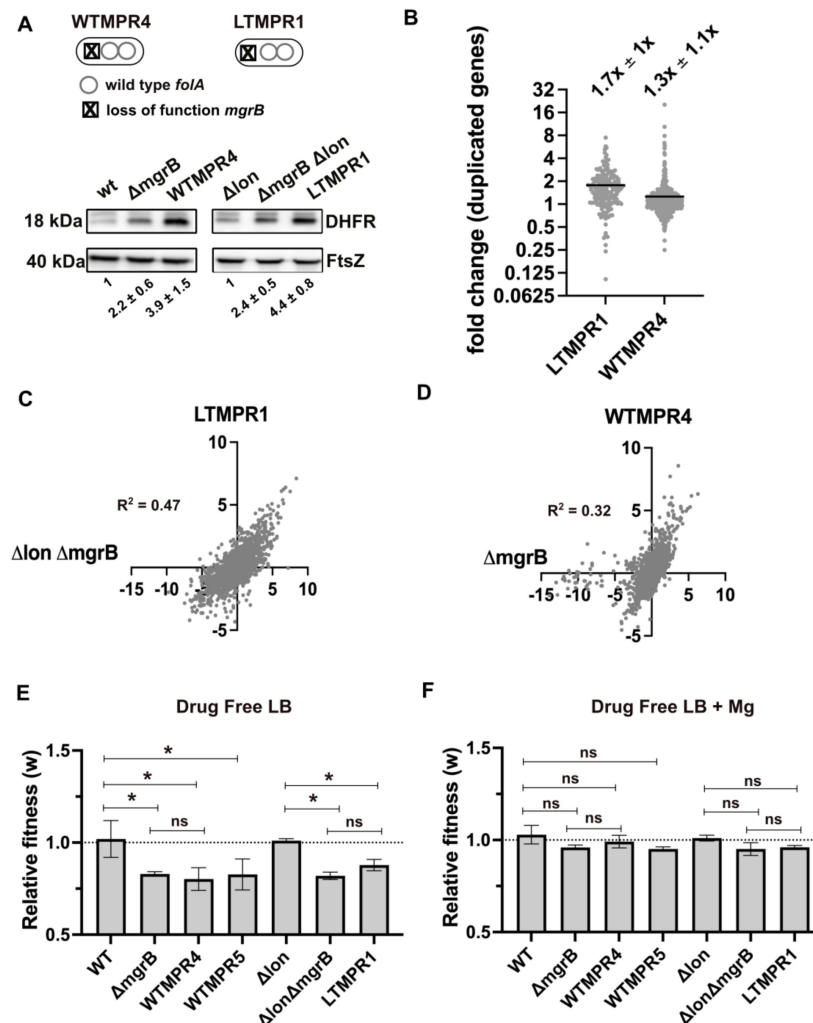


Figure 2.

Overexpression of DHFR in trimethoprim-resistant isolates with genomic duplication.

A. DHFR protein expression in trimethoprim-resistant isolates WTMPR4 and LTMPR1, compared with respective ancestors and *mgrB*-knockout strains. DHFR protein was detected by immunoblotting using an anti-DHFR polyclonal antibody. FtsZ was used as loading control. A representative immunoblot from 3 biological replicates is shown. Quantitation was performed by calculating band intensities using image analysis. Fold change values of DHFR levels over controls (wild type or *E. coli* Δlon , set to 1) is shown as mean $\pm$ S.D. from triplicate measurements. Alleles of *mgrB* and *folA* in WTMPR4 and LTMPR1 are shown diagrammatically. **B.** Expression level of duplicated genes in LTMPR1 and WTMPR4, determined by RNA-sequencing, expressed as fold over *E. coli* $\Delta lon \Delta mgrB$ and *E. coli* $\Delta mgrB$ respectively. Each point represents fold change value for a single gene. Mean of the scatter is shown as a black line, and its value is provided along with standard deviation above the plot. **C.** **D.** Correlation between expression levels of genes in LTMPR1 (C) and WTMPR4 (D) with *E. coli* $\Delta lon \Delta mgrB$ and *E. coli* $\Delta mgrB$ respectively. Gene expression levels were estimated using RNA-sequencing and fold changes were calculated using wild type *E. coli* as reference for WTMPR4 and *E. coli* $\Delta mgrB$, and *E. coli* Δlon as reference for LTMPR1 and *E. coli* $\Delta lon \Delta mgrB$. Each point represents $\log_2$ (fold change) value for a single gene. The resulting scatter was fit to a simple linear regression and the obtained R^2 values are provided. **E.** **F.** Relative fitness (w) of indicated *E. coli* strains calculated using competition experiments with *E. coli* $\Delta lacZ$ as reference in the absence (E) or presence (F) of 10 mM $MgSO_4$. Presence of high concentration of Mg^{2+} alleviates the fitness cost of *mgrB* mutations by inhibiting the PhoQP system [51]. For strains derived from *E. coli* Δlon , *E. coli* $\Delta lon \Delta lacZ$ was used as reference. Neutrality of $\Delta lacZ$ was established by competition with the unmarked ancestral strains, i.e. wild type (WT) and *E. coli* Δlon . No change in relative fitness compared to the reference strain (w = 1) is indicated by a dotted line. Mean $\pm$ S.D. from three independent measurements is plotted. Statistical significance was tested using an unpaired t-test and a p-value of < 0.5 was considered statistically significantly different (*). A p-value ≥ 0.5 was considered not significantly different (ns).

In addition to *folA*, ~173 and ~531 other genes were also duplicated in LTMPR1 and WTMPR4, respectively. To understand the global transcriptional burden imposed by this amplification, we analysed the gene expression profiles of WTMPR4 and LTMPR1 using RNA-sequencing (Supplementary File 1). Since both isolates had loss-of-function mutations in *mgrB*, which cause large changes in global gene expression [51], we also analysed the transcriptomes of *E. coli* $\Delta mgrB$ and *E. coli* $\Delta lon\Delta mgrB$ strains for comparison (Supplementary File 1). Transcript level of *folA* was higher in LTMPR1 (2.3-fold) and WTMPR4 (5.7-fold) over respective controls (Supplementary File 1). Other duplicated genes were mildly overexpressed by ~1.7-fold on average in LTMPR1 (over *E. coli* $\Delta lon\Delta mgrB$) and ~1.3-fold in WTMPR4 (over *E. coli* $\Delta mgrB$) (Figure 2B, Supplementary File 1). However, there was large variation in expression level among duplicated genes (Figure 2B), indicating that increase in gene dosage did not necessarily result in enhanced expression levels. Instead, global gene expression profiles of WTMPR4 and LTMPR1 were correlated with *E. coli* $\Delta mgrB$ and *E. coli* $\Delta lon\Delta mgrB$ respectively (Figure 2C, D) and most of the top 50 up/down-regulated genes in WTMPR4 and LTMPR1 could be attributed to loss of *mgrB* alone (Figure S3). In line with this result, strains with and without genomic amplifications had comparable fitness costs (~20 %) in drug-free media, similar in magnitude to the cost of *mgrB* inactivation (Figure 2E). Further, the fitness cost was alleviated in growth media supplemented with high Mg^{2+} , i.e. by inhibiting the PhoQP pathway, confirming the source of this cost to be loss of *mgrB* rather than the GDA (Figure 2F). Thus, we concluded that large GDA mutations encompassing *folA* led to DHFR overexpression and represented an alternate mechanism for trimethoprim-resistance in *E. coli* with undetectable additional cost.

Maintenance of *folA* duplication is linked to antibiotic-generated demand for DHFR expression

We next investigated the evolutionary fate of duplicate *folA* gene copies by evolving the LTMPR1 isolate, which harboured the shortest duplication, in the absence of antibiotic. We also evolved this isolate at sustained drug pressure, i.e. trimethoprim at the same concentration at which it had originally emerged (300 ng/mL). Using three replicate populations in each condition, we asked how DHFR expression and trimethoprim IC₅₀ values changed over ~252 generations of evolution and used genome sequencing to identify the underlying mutational basis. As controls, we performed similar evolution experiments with WTMPR4 (which harboured a GDA and a functional Lon protease) and WTMPR5 (which harboured a *folA* point mutation and a functional Lon protease) (Figure 3A).

Evolution in trimethoprim-free media

In drug-free conditions, trimethoprim IC₅₀ of LTMPR1 rapidly declined and a concomitant loss of colony formation on trimethoprim-supplemented plates was observed (Figure 3B, C). Evolved WTMPR4, but not WTMPR5, showed similar loss of colony formation on trimethoprim indicating that resistance mediated by GDA mutations was less stable than point mutations regardless of the presence of Lon (Figure 3C). For LTMPR1, these trends were explained by a decline in DHFR expression levels over evolution and reversion to a single *folA* copy (Figure 3D, E, S4). Mutation in *mgrB* was, however, consistently maintained till the end of the experiment (Supplementary File 2). In agreement with genomics data, trimethoprim IC₅₀s of evolving populations stabilized at a value similar to *E. coli* $\Delta lon\Delta mgrB$ (Figure 3B). Interestingly, a second GDA mutation (“GDA-2”, ~2.2 Mb to 3.3 Mb) emerged in these populations (Figure 3E, S4). The GDA-2 mutation was apparently lost and regained multiple times during the experiment (Figure 3E, S4), suggesting a dynamic maintenance in the population rather than fixation. Further, the region of the genome amplified in GDA-2 was similar to what we had observed when *E. coli* Δlon had evolved in drug-free media (Figure S2, S4), indicating that GDA-2 was likely a fitness enhancing mutation unrelated to trimethoprim adaptation.

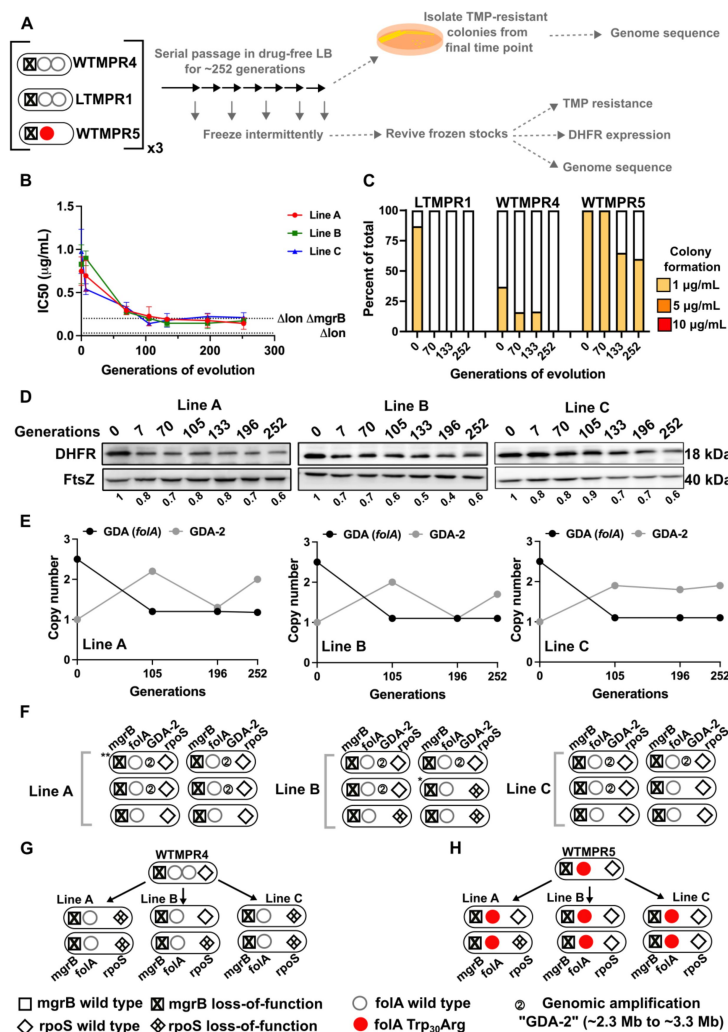


Figure 3.

Evolutionary fate of *folA*-duplication in the absence of drug pressure.

A. Schematic of the experimental pipeline used to investigate the impact of evolution in drug-free media on resistance level, *folA* copy number and expression level of different trimethoprim resistant *E. coli* populations. **B.** Trimethoprim IC₅₀ values of 3 evolving lineages derived from LTMPR1 (A, B, C) over 252 generations of evolution in antibiotic-free medium. Mean ± S.D. from three measurements is plotted at each time point. IC₅₀ values of *E. coli* Δlon and ΔlonΔmgrB are shown as dotted lines for reference. **C.** Colony formation of LTMPR1, WTMPR4 and WTMPR5 evolved in antibiotic-free medium at different generations. Fraction of the population capable of forming colonies at 1, 5 and 10 μg/mL trimethoprim was calculated across 3 replicate lines. Mean value from the three lines is plotted. The results of similar experiments performed on ancestors (0 generations) are also provided for reference. **D.** DHFR expression during evolution in the absence of trimethoprim in LTMPR1 lineages A, B and C, measured by immunoblotting using anti-DHFR polyclonal antibody. FtsZ was used as a loading control. Quantitation was performed by calculating band intensities using image analysis. DHFR expression at each time point was normalized to the ancestor (i.e. 0 generations, set to 1). Mean of three independent measurements is shown below each lane. **E.** Copy number of the ancestral GDA encompassing *folA* (GDA(*folA*)) and GDA-2 at different time points of evolution in the absence of trimethoprim are plotted. For GDA(*folA*), copy number was determined by dividing number of reads from an Illumina sequencing experiment corresponding to *folA* by the average number of reads mapping to the rest of the genome. **F.** Point mutations in *folA*, *rpoS* and *mgrB*, and the "GDA-2" genomic duplication in 6 randomly picked colonies from each of the LTMPR1 lines at 252 generations of evolution are shown schematically using the appropriate symbols. Asterisk (*)/** marks the genotypes that were carried forward for further analyses. **G, H.** Point mutations in *folA*, *mgrB* and *rpoS* in isolates derived from 252 generations of evolution of WTMPR4 (**G**) and WTMPR5 (**H**) in no antibiotic. From each of the evolving lines 2 random isolates were picked for genome sequencing. Various point mutants at the 3 gene loci of interest are represented by appropriate symbols as shown in the legend.

To further investigate the emergence of GDA-2 during evolution of LTMPR1 in drug-free media we sequenced 6 randomly picked colonies from each of the lineages at the final time point (**Figure 3 F**). All the sequenced isolates had a single copy of *folA* and a majority of isolates showed the presence of GDA-2 (**Figure 3 F**, Supplementary File 2). Interestingly, some of the isolates that did not harbour GDA-2 had evolved mutations in *rpoS* (**Figure 3 F**, Supplementary File 2). Loss-of-function mutations in RpoS are fixed at late time points during adaptation to trimethoprim and enhance the fitness of *E. coli* without conferring resistance [46, 51]. We also sequenced a few randomly picked isolates from WTMPR4 and WTMPR5 evolved lines (2 from each of the 3 replicate lineages in each case) to check for the presence of GDA-2 and *rpoS* mutations. As expected, clones derived from WTMPR4 had reverted to a single copy of *folA*, while WTMPR5 retained its ancestrally inherited *folA* mutation. None of them had evolved the GDA-2 mutation, while a few of the isolates from WTMPR4 and WTMPR5 lines had acquired *rpoS* mutations (**Figure 3 G, H**).

Consistent with earlier studies, and the role of mutations in *rpoS* in general fitness enhancement in *E. coli*, we found that an LTMPR1-derived isolate with mutant *rpoS* recovered fitness in drug-free media (**Figure 4 A**). Interestingly, an isolate with the GDA-2 mutation had also recovered fitness (**Figure 4 A**). Since GDA-2 occurred only in a *lon*-deficient background, and GDA-2 and *rpoS* mutations were mutually exclusive, we wondered whether the fitness enhancing effects of GDA-2 were also mediated by reduced RpoS activity. Lon is known to reduce the levels of RpoS during the exponential growth phase of *E. coli* [52]. Therefore, we reasoned that at least some targets of RpoS would be overexpressed in a *lon*-knockout. To test this prediction, we analysed the transcriptome of *E. coli* Δlon and specifically looked at the expression levels of 168 known targets of RpoS (Supplementary File 3). Of these, 41 genes were overexpressed in the *lon*-knockout by at least 1.5-fold over wild type, including genes induced in the stationary phase and by various stresses such as *poxB*, *aidB* and *cbpA* (**Figure 4 B, C**). Levels of *rpoS* mRNA were however unaffected in this strain. Next, we compared the expression levels of these 41 RpoS-targets in LTMPR1 and *E. coli* $\Delta lon \Delta mgrB$ with *E. coli* Δlon as the reference and found that most had similar expression levels in these three strains (**Figure 4 C**). Indeed a few of these genes showed higher expression levels in LTMPR1 and *E. coli* $\Delta lon \Delta mgrB$, which was not unexpected since loss of *mgrB* is known to stabilize RpoS [53]. We checked their expression in isolates derived from LTMPR1 which harboured either GDA-2 or *rpoS* mutations. In both isolates RpoS target genes were down-regulated indicating a re-setting of RpoS activity (**Figure 4 C**). Further, in the isolate harbouring GDA-2, *rpoS* mRNA levels were themselves lowered to 0.49x of *E. coli* Δlon (Supplementary File 3). Finally, deletion of *rpoS* from *E. coli* Δlon enhanced the fitness of the strain (**Figure 4 D**). Thus, evolution of LTMPR1 in antibiotic-free media was characterized by rapid reversal of GDA-mediated resistance and acquisition of mutations that enhanced fitness by reducing the levels of RpoS-transcribed genes (**Figure 4 E**).

Evolution at constant trimethoprim pressure

In the presence of trimethoprim (300 ng/mL), antibiotic resistance level and DHFR expression remained high over the course of evolution of LTMPR1 (**Figure 5 A, B, C, D**). WTMPR4 and WTMPR5 too remained resistant, and in the case of WTMPR4 an increased colony forming efficiency was observed at a late time point (**Figure 5 C**). These data suggested that under drug pressure high *folA* copy number was likely to be maintained. Indeed, population sequencing of LTMPR1 lineages showed that the ancestral GDA mutation was maintained until 196 generations of evolution in all three replicates (**Figure 5 E**, **S5**). By 252 generations, however, a majority of bacteria in Lines A and C had reverted to a single *folA* copy (**Figure 5 E**, **S5**) and had acquired a point mutation either in the *folA* gene (Line C) or its promoter (Lines A) (**Figure 5 E**, Supplementary File 4). Line B, on the other hand, retained high *folA* copy number, but with a reduction in the length of the genomic duplication (**Figure 5 D**, **S5**). Line B had also acquired a mutation in the *folA* promoter, however its frequency in the population was low (**Figure 5 D**, Supplementary File 4).

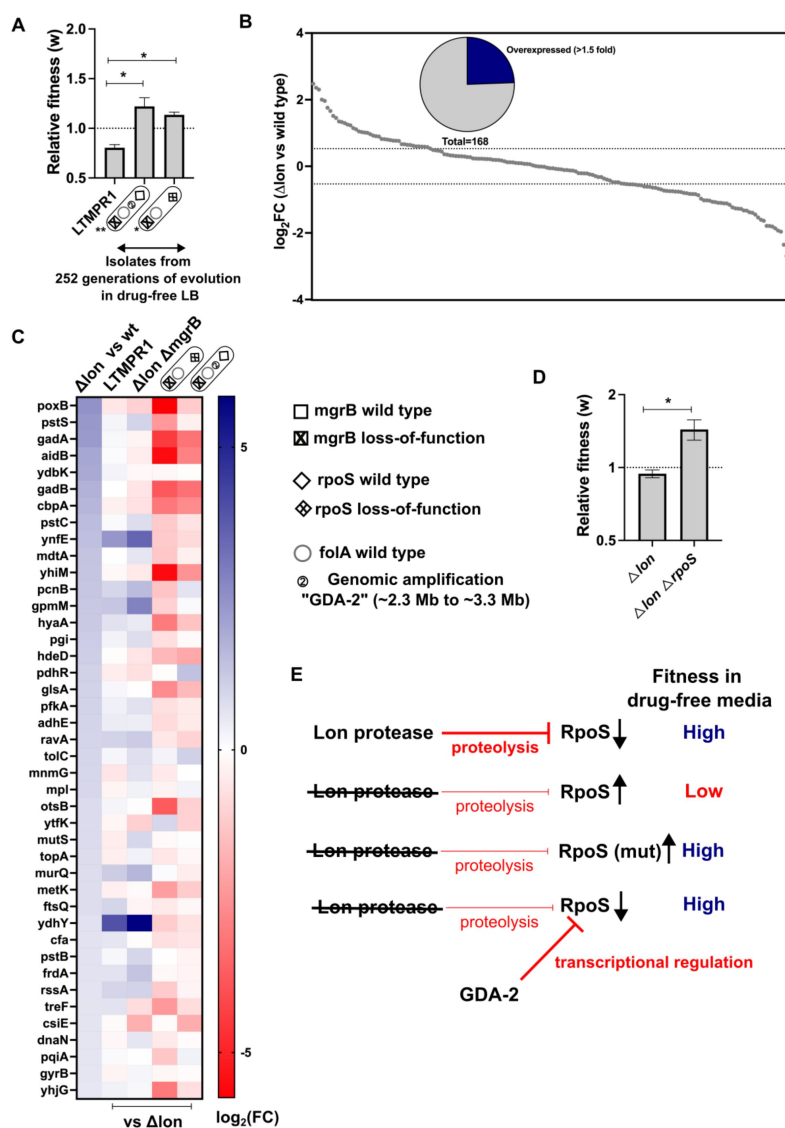


Figure 4.

Lower RpoS activity leads to fitness enhancement of LTMPR1 in drug-free media.

A. Relative fitness (w) of isolates with the indicated genotypes derived from LTMPR1 evolution in antibiotic-free media (see also **Figure 3**) calculated using a competitive growth assay using *E. coli* Δ lon Δ lacZ as the reference strain. Mean $\pm$ S.D. from three independent measurements is plotted. No change in relative fitness compared to the reference (w = 1) is shown as a dotted line for reference. Statistical significance was tested using an unpaired t- test. A p-value of < 0.05 was considered as a statistically significant difference (*). **B.** Expression level of 168 known targets of RpoS in *E. coli* Δ lon compared to wild type expressed as $\log_2$ (fold change) values determined by RNA sequencing. Dotted lines at 1.5- fold higher and lower than wild type represent cut-offs used to identify overexpressed and under-expressed genes. The pie chart shows the fraction of RpoS targets that were overexpressed in *E. coli* Δ lon by at least 1.5-fold. **C.** Expression levels of 41 de-regulated RpoS targets in *E. coli* Δ lon compared to wild type are shown as a heat map in the first vertical. The expression levels of these genes in LTMPR1, *E. coli* Δ lon Δ mgrB or LTMPR1- derived isolates compared to *E. coli* Δ lon are provided in the subsequent verticals. **D.** Relative fitness of *E. coli* Δ lon Δ rpoS in antibiotic-free media calculated using a competitive growth assay using *E. coli* Δ lon Δ lacZ as the reference strain. Neutrality of the Δ lacZ genetic marker was verified by competition between *E. coli* Δ lon and *E. coli* Δ lon Δ lacZ. Mean $\pm$ S.D. from three independent measurements is plotted. No change in relative fitness compared to the reference (w = 1) is shown as a dotted line for reference. Statistical significance was tested using an unpaired t-test. A p-value of < 0.5 was considered as a statistically significant difference (*). **E.** Model for the effects of Lon deficiency on RpoS levels and bacterial fitness. The roles of mutations within RpoS or the GDA-2 mutation in compensating for the defects of RpoS overproduction are shown.

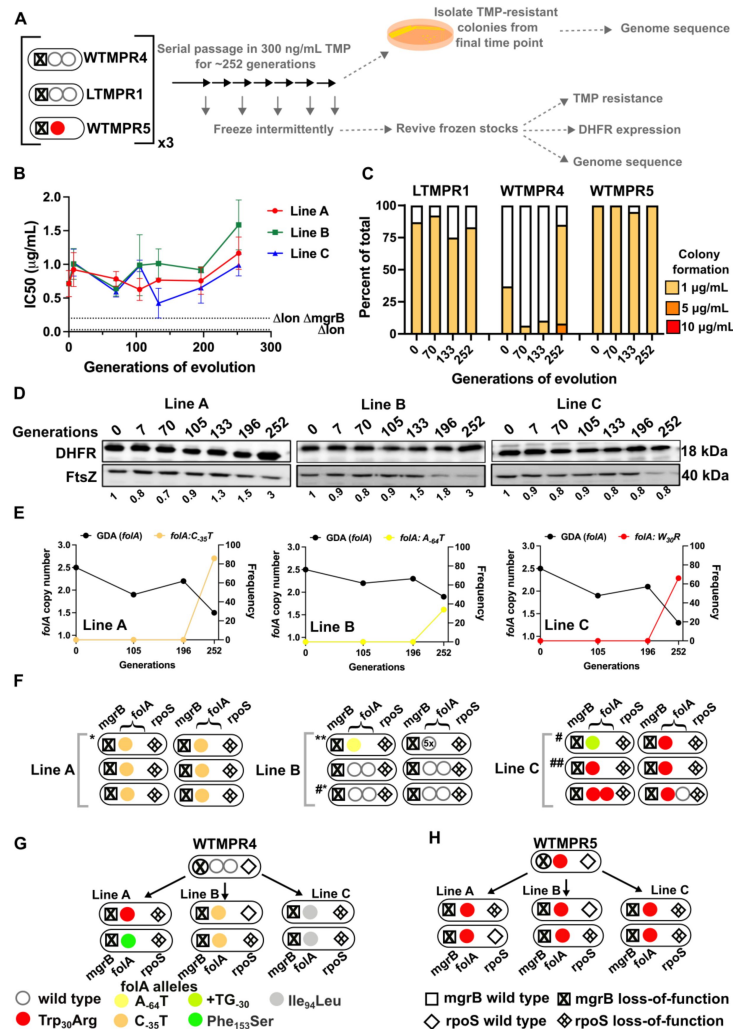


Figure 5.

Evolutionary fate of *folA*-encompassing GDA in LTMPR1 at sustained trimethoprim pressure.

A. Schematic of the experimental pipeline used to investigate the impact of evolution in constant drug-pressure on resistance level, *folA* copy number, point mutations in *folA* and expression level of different trimethoprim resistant *E. coli* populations. **B.** Trimethoprim IC50 values of 3 evolving lineages starting from LTMPR1 (A, B, C) over 252 generations of evolution at 300 ng/mL trimethoprim. Mean ± S.D. from three measurements is plotted at each time point. IC50 values of *E. coli* Δlon and $\Delta lon \Delta mgrB$ are shown as dotted lines for reference. **C.** Colony formation of LTMPR1, WTMPR4 and WTMPR5 evolved in 300 ng/mL trimethoprim at indicated time points. Fraction of the population capable of forming colonies at 1, 5 and 10 μ g/mL trimethoprim was calculated across 3 replicate lines. Mean value from the three lines is plotted. The results of similar experiments performed on ancestors (0 generations) are also provided for reference. **D.** DHFR expression in LTMPR1 lineages A, B and C evolved in 300 ng/mL trimethoprim, measured by immunoblotting using anti-DHFR polyclonal antibody. FtsZ was used as a loading control. Quantitation was performed by calculating band intensities using image analysis. DHFR expression at each time point was normalized to the ancestor (i.e. 0 generations, set to 1). Mean of three independent measurements is shown below each lane. **E.** Copy number of *folA* (GDA(*folA*)) at different time points of evolution of LTMPR1 in trimethoprim is plotted on the left y-axis. Copy number was determined by dividing number of reads from an Illumina sequencing experiment corresponding to *folA* by the average number of reads mapping to the rest of the genome. Frequency of various *folA* alleles in the evolving populations at each of the time points is plotted on the right y-axis. **F.** Point mutations in *folA*, *rpoS* and *mgrB* in 6 randomly picked trimethoprim-resistant colonies from each of the LTMPR1 lines at 252 generations of evolution are shown diagrammatically. Genotypes that were carried forward for further analysis are marked (* or #). **G, H.** Point mutations in *folA*, *mgrB* and *rpoS* in isolates derived from 252 generations of evolution of WTMPR4 (G) and WTMPR5 (H) in 300 ng/mL of trimethoprim. From each of the evolving lineages 2 random isolates were picked for genome sequencing. Various point mutants at the 3 gene loci are represented by appropriate symbols as shown in the legend.

In agreement with population sequencing, none of the resistant isolates derived from LTMPR1 Line A at the final time point showed the ancestral GDA. Instead, they all harboured a C₃₅T mutation in the *folA* promoter that is known to up-regulate DHFR expression [51] (Figure 5 F, Supplementary File 4). Among Line B isolates, many had retained the ancestral duplication, without acquiring any additional mutations in *folA*. One of the isolates from this Line that had reverted to a single *folA* copy had gained the A₆₄T *folA* promoter mutation (Figure 5 F, Supplementary File 4). In Line C too we found that isolates that had lost the ancestral GDA had evolved a point mutation in the *folA* promoter (+TG₃₀) or a Trp₃₀Arg mutation in DHFR (Figure 5 F, Supplementary File 4). We found similar results in sequenced isolates from WTMPR4 lineages, i.e. loss of GDA and gain of a point mutation either in *folA* or its promoter (Figure 5 G). WTMPR5-derived isolates did not accumulate additional mutations in *folA* but did acquire mutations in *rpoS* (Figure 5 H). Interestingly, two isolates from LTMPR1 Lines B and C had evolved new GDA mutations encompassing *folA*. In one of the Line B isolates (Isolate 2) the ancestral GDA was replaced by 5x amplification of a shorter genomic region (~12 kb) encompassing *folA* (Figure S6). Both ends of this newly amplified region were flanked by IS1 elements (Figure S6). On the other hand, one of the Line C isolates (Isolate 5) had an expanded duplication beyond what was originally found in LTMPR1 (Figure S6). Also, both copies of *folA* in this isolate coded for DHFR-Trp₃₀Arg (Figure 5 E, S6), suggesting that the original GDA was lost and an extended GDA mutation was gained after fixation of the mutant *folA* allele.

Based on these results we concluded that even under drug pressure, GDA mutations were reversed in most bacteria. Notably, reversion to a single *folA* copy was slower than in drug-free media as it first required the fixation of point mutations at the *folA* locus. However, in a smaller sub-population, more than one *folA* gene copy was maintained. This maintenance was dynamic and likely involved repeated loss and gain of GDA mutations.

To further investigate the replacement of GDA by point mutants, we checked to see if point mutants could invade a population harbouring GDA mutations under drug pressure. We allowed LTMPR1-derived isolates with point mutations in *folA* (marked with $\Delta lacZ::Cat$) to compete with a GDA-harboring isolate from the same time point in evolution and asked how rapidly the former could invade a population of the latter. We performed these competitions with a contemporary GDA-harboring isolate rather than the ancestor to nullify the effects of other media adaptations such as *rpoS* mutations that may have been acquired over evolution. Since for this assay we used point mutant strains genetically marked with $\Delta lacZ::Cat$ we verified the neutrality of the $\Delta lacZ::Cat$ marker by checking if the marked isolate could invade its unmarked parent. In the presence of trimethoprim (300 ng/mL) point mutants were enriched over the GDA mutant in co-culture by about 100-fold (Figure 6 A) recreating the results of our evolution experiment. A possible explanation for this finding may be higher resistance level conferred by *folA* point mutations than GDA. Indeed, some isolates with *folA* point mutations had higher IC₅₀ values compared to a contemporary GDA-harboring isolate (Figure 6 B). However, regardless of IC₅₀ values, all isolates also had similar doubling times in the presence of 300 ng/mL trimethoprim (Figure 6 B). Thus, though there were no large observable differences in growth rate between isolates, point mutants could indeed invade and outcompete GDA mutations, reflecting either subtle differences in fitness between them or their relative stabilities in culture.

Maintenance of high *folA* gene copy number under drug pressure is driven by unstable mutants of DHFR

Though sustained drug pressure did extend the lifetime of *folA* gene duplication, acquisition of trimethoprim-resistant point mutations facilitated reversion to a single gene copy. We therefore asked whether *folA* duplication could be maintained under conditions of increasing antibiotic pressure, i.e., trimethoprim concentrations starting at 1 µg/mL and ending at 10 µg/mL over 252 generations of evolution (Figure 7 A). Under these conditions, trimethoprim IC₅₀ values

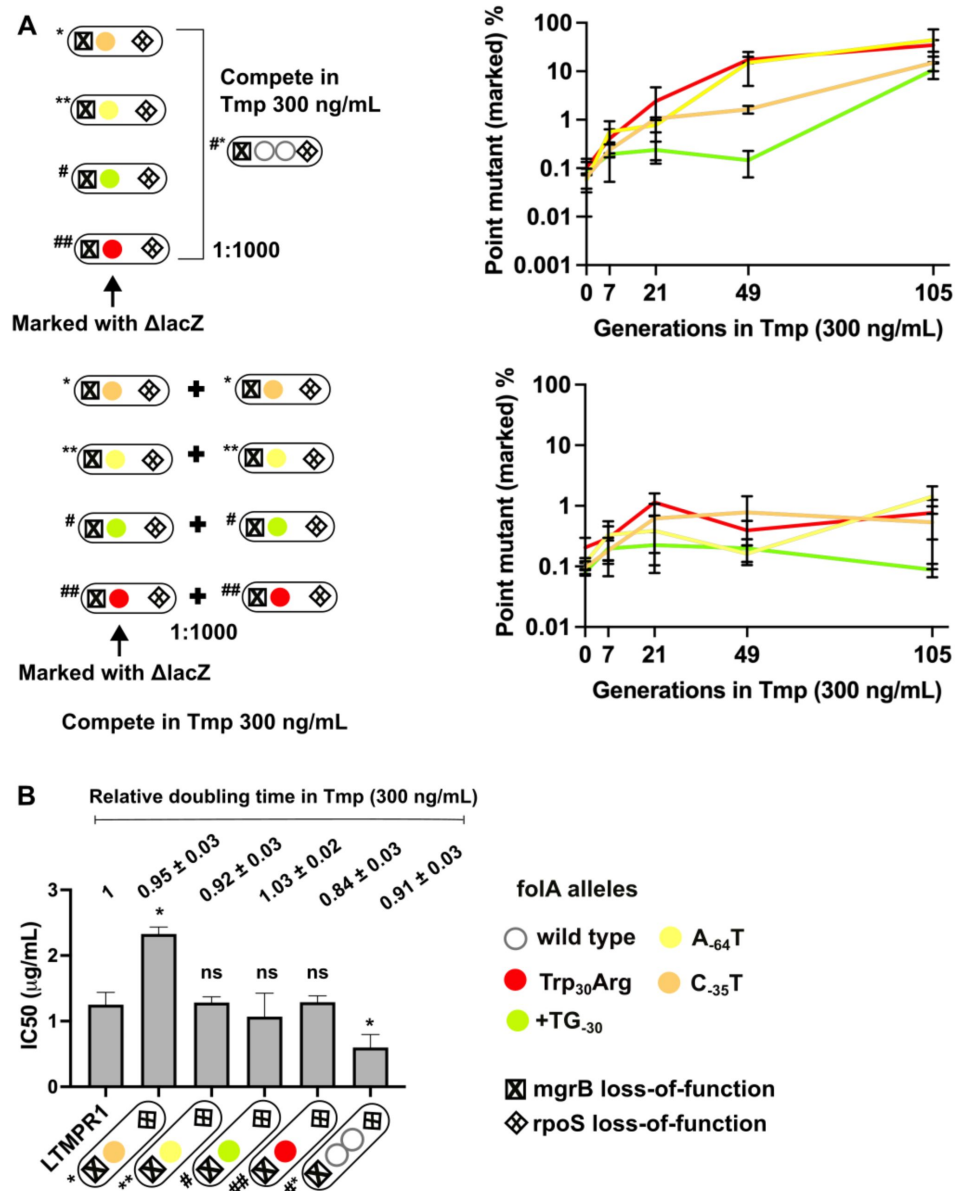


Figure 6.

Invasion of GDA-dominant populations by point mutations under trimethoprim pressure.

A. Upper panel. Competition between isolates with point mutations in *folA* and an isolate that harboured a *folA* duplication derived from the same time point of LTMPR1 evolution (upper panel). The genotypes of the isolates (i.e. alleles at *mgrB*, *folA* and *rpoS* loci) used are shown diagrammatically (see also **Figure 5**, Supplementary File 4). The initial mixing ratio was 1000:1 in favour of the GDA mutant. Point mutants were marked genetically with $\Delta lacZ::Cat$. Neutrality of the $\Delta lacZ::Cat$ marker was verified by competing marked and unmarked point mutants (lower panel). Percentage of the population constituted by the marked point mutants over ~105 generations of serial transfer in media supplemented with trimethoprim (300 ng/mL) is plotted. Mean $\pm$ S.D. from 3 replicates are plotted and traces for individual mutants are appropriately colored. **B.** Trimethoprim IC₅₀ values of the LTMPR1 ancestor or evolved isolates from 252 generations in trimethoprim with indicated genotypes (see also **Figure 5**, Supplementary File 4) are plotted as bars. Mean $\pm$ S.D. from three independent measurements are plotted. Statistical significance was tested using an unpaired t-test and LTMPR1 was used as the reference. A p-value < 0.05 was considered significantly different (*) while ≥ 0.5 was considered not statistically significant (ns). Doubling times relative to LTMPR1 (set to 1) in the presence of 300 ng/mL of trimethoprim are provided above the graph as mean $\pm$ S.D. from three independent measurements.

increased sharply over the course of evolution eventually reaching ~40 to ~160-fold higher than ancestral LTMPR1 (**Figure 7 B**). Colony formation at high trimethoprim concentrations also increased at later generations of evolution (**Figure 7 C**). Similarly evolved WTMPR4 and WTMPR5 also showed enhanced colony formation at high trimethoprim concentrations (**Figure 7 C**).

Curiously, we observed that the 3 replicate lineages of LTMPR1 showed different trends in DHFR expression. Line A maintained a high level of DHFR protein for the duration of the experiment. On the other hand, Lines B and C showed high DHFR expression until 196 generations, after which it sharply declined (**Figure 7 D**). Genome sequencing revealed that all 3 lines convergently evolved the Trp₃₀Arg mutation in DHFR early during evolution and had fixed this mutation by 196 generations. Until this time point, all 3 lines retained *folA* duplication (**Figure 7 E**, **S7**). Subsequently, Line A evolved an additional C₃₅T *folA* promoter mutation and had partially lost the ancestral GDA (**Figure 7 E**, **S7**, Supplementary File 5). In Lines B and C, a second missense mutation in DHFR, i.e. Tyr₁₅₁Asp, appeared between 196 and 252 generations (**Figure 7 D**, Supplementary File 5). Line B also showed an Ile₉₄Leu mutation in DHFR at low frequency at generation 252 (**Figure 7 D**, Supplementary File 5).

Unexpectedly, despite acquiring additional mutations and lower DHFR expression level, LTMPR1 Lines B and C retained two copies of *folA* in contrast to Line A (**Figure 7 E**, **S7**, Supplementary File 5).

Individual resistant clones from each of the lineages provided greater insight into the genetic trajectories of evolution. In LTMPR1 Line A, 2 of the 6 sequenced isolates had lost gene duplication and acquired the Trp₃₀Arg coding and C₃₅T promoter mutations in *folA* (**Figure 7 F**, Supplementary File 5). Interestingly, the other 4 sequenced isolates from Line A retained the ancestral GDA and harboured the same combination of *folA* point mutations, but only in 1 of the 2 gene copies (**Figure 7 F**, Supplementary File 5). Ten of the twelve isolates from LTMPR1 Lines B and C, on the other hand, retained *folA* duplication. In all ten isolates, both copies of *folA* had acquired Trp₃₀Arg and Tyr₁₅₁Asp mutations (**Figure 7 F**, Supplementary File 5). Unlike LTMPR1, all sequenced isolates of WTMPR4 at 252 generations of evolution had lost their GDA mutation and gained 2 *folA* mutations (**Figure 7 G**). WTMPR5 too had gained an additional *folA* mutation (**Figure 7 H**). Notably, none of the isolates from WTMPR4 and WTMPR5 lineages had the Trp₃₀Arg and Tyr₁₅₁Asp combination of mutations that occurred repeatedly in LTMPR1 lineages.

We next sought to understand why some LTMPR1 bacteria had retained high *folA* copy number despite acquiring multiple point mutations in the drug target. To address this question, we first asked what the advantage of *folA* duplication was in the two different mutational trajectories. Isolates from Line A had similar IC₅₀ values for trimethoprim regardless of whether they had one or two copies of *folA* (**Figure 8 A**). On the contrary, isolates from Line B that harboured only one copy of mutant *folA* (coding for DHFR-Trp₃₀Arg-Tyr₁₅₁Asp) displayed ~3.5-fold lower IC₅₀ than those that harboured two copies (**Figure 8 A**). Immunoblotting showed that isolates from Line A had similar expression level of DHFR (**Figure 8 B**). However, DHFR protein level was significantly compromised by the Trp₃₀Arg-Tyr₁₅₁Asp mutation combination and double the number of *folA* gene copies restored this defect (**Figure 8 B**). Thus, higher gene dosage of DHFR was dispensable once a promoter up-regulatory mutation was fixed. On the other hand, bacteria with alleles of DHFR harbouring multiple missense mutations continued to benefit from overexpression due to *folA* duplication.

Mutations at position Trp₃₀ in DHFR render the protein resistant to inhibition by trimethoprim but are accompanied by lower *in vivo* proteolytic stability due to perturbation of the hydrophobic core of the protein [54]. This affects the expression level of DHFR, resulting in lower cellular abundance that limits the resistance level conferred by mutations at Trp₃₀ [49, 54]. Interestingly, residue Tyr₁₅₁ lies in the same hydrophobic pocket as Trp₃₀ (**Figure 8 C**). To test

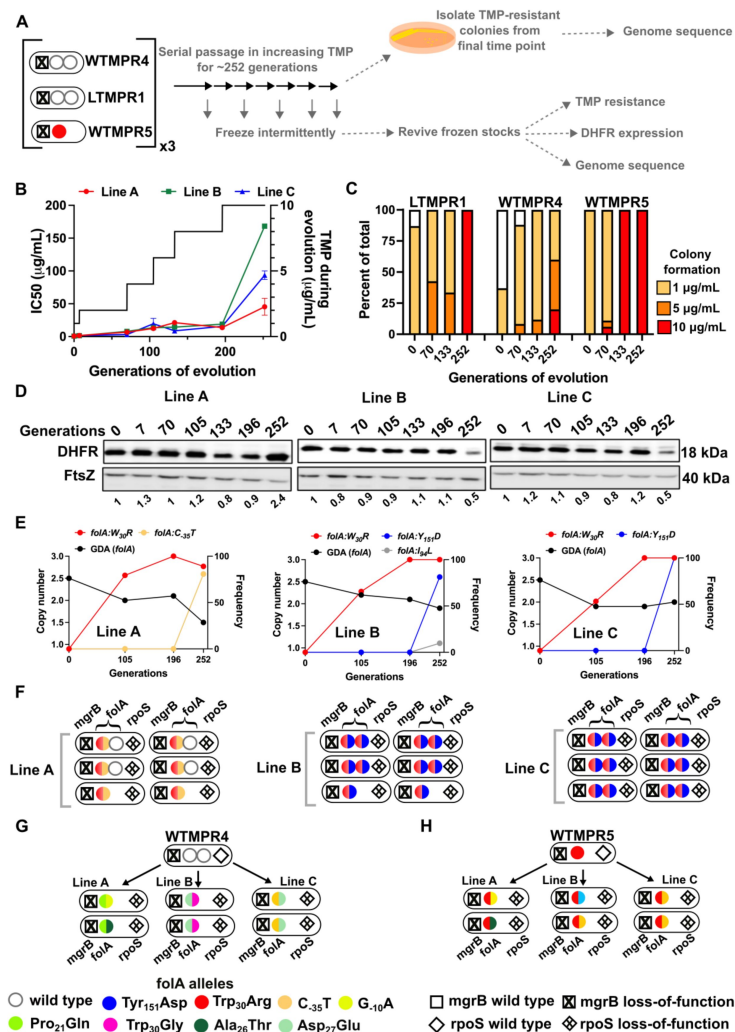


Figure 7.

Evolutionary fate of *folA*-encompassing GDA in LTMPR1 at increasing trimethoprim pressure.

A. Schematic of the experimental pipeline used to investigate the impact of evolution in increasing drug-pressure on resistance level, *folA* copy number, point mutations in *folA* and expression level of different trimethoprim resistant *E. coli* populations. **B.** Trimethoprim IC₅₀ values of 3 evolving lineages starting from LTMPR1 (A, B, C) over 252 generations of evolution are plotted on the left Y-axis. Mean $\pm$ S.D. from three measurements is plotted at each time point. Trimethoprim concentrations used during evolution are plotted in the right Y-axis. **C.** Colony formation of LTMPR1, WTMPR4 and WTMPR5 evolved in increasing trimethoprim at indicated time points. Fraction of the population capable of forming colonies at 1, 5 and 10 μ g/mL trimethoprim was calculated across 3 replicate lines. Mean value from the three lines is plotted. The results of similar experiments performed on ancestors (0 generations) are also provided for reference. **D.** DHFR expression in LTMPR1 lineages A, B and C evolved in increasing trimethoprim, measured by immunoblotting using anti-DHFR polyclonal antibody. FtsZ was used as a loading control. Quantitation was performed by calculating band intensities using image analysis. DHFR expression at each time point was normalized to the ancestor (i.e. 0 generations, set to 1). Mean of three independent measurements is shown below each lane. **E.** Copy number of *folA* (GDA(*folA*)) at different time points of LTMPR1 evolution in increasing trimethoprim is plotted on the left y-axis. Copy number was determined by dividing number of reads from an Illumina sequencing experiment corresponding to *folA* by the average number of reads mapping to the rest of the genome. Frequency of various *folA* alleles in the evolving LTMPR1 populations at each of the time points is shown on the right y-axis. **F.** Point mutations in *folA*, *rpoS* and *mgrB* in 6 randomly-picked trimethoprim-resistant colonies from each of the LTMPR1 lines at 252 generations of evolution are shown schematically. **G, H.** Point mutations in *folA*, *mgrB* and *rpoS* in isolates derived from 252 generations of evolution of WTMPR4 (G) and WTMPR5 (H) in increasing trimethoprim. From each of the evolving lineages 2 random isolates were picked for genome sequencing. Various point mutants at the 3 gene loci are represented by appropriate symbols as shown in the legend.

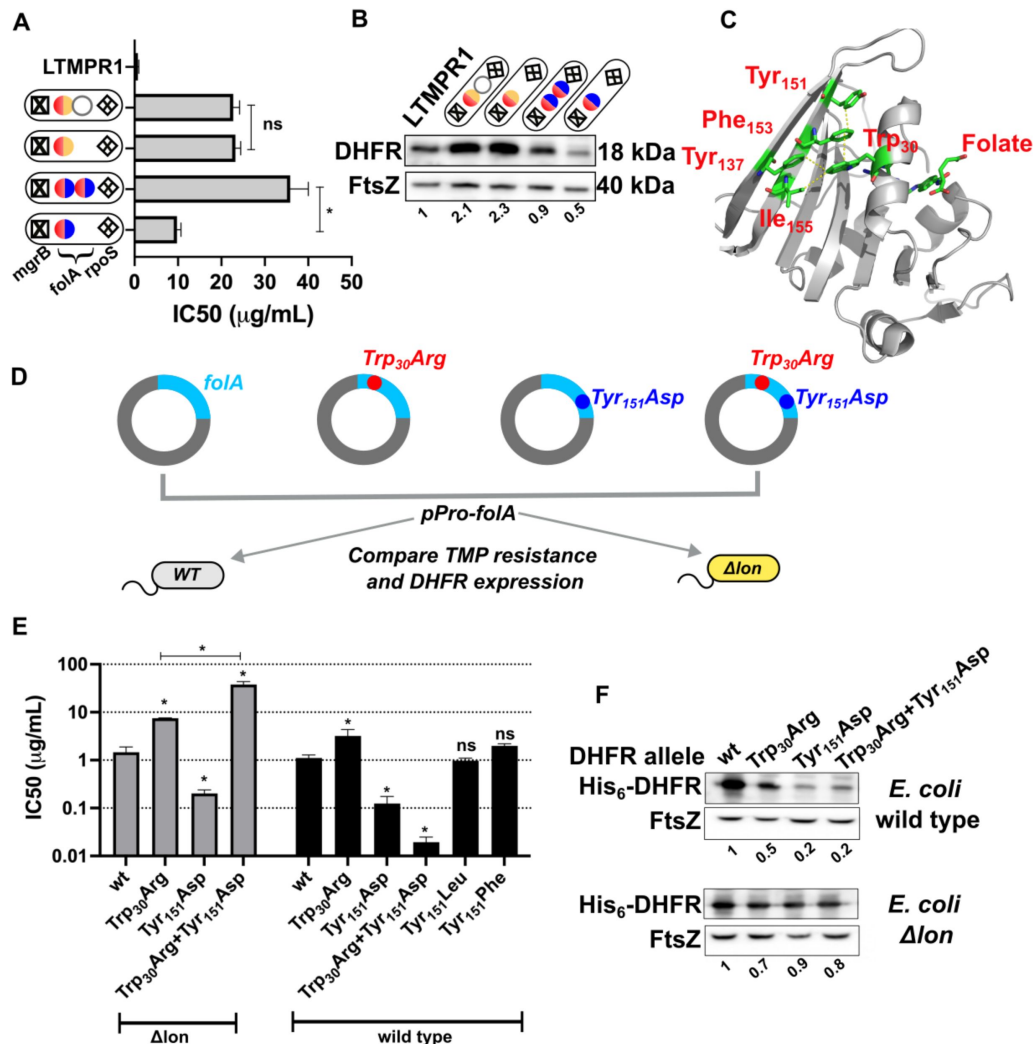


Figure 8.

Proteostatic pressure facilitates maintenance of *folA* duplication

A. Trimethoprim-IC₅₀ values of resistant isolates with the indicated genotypes (see also [Figure 7](#)) derived from LTMPR1 evolution in increasing antibiotic pressure. Mean ± S.D. from three independent measurements is plotted. Statistical significance was tested using an unpaired t-test. A p-value of <0.05 was considered statistically significant (*). A p-value > 0.05 was considered not statistically significant (ns). **B.** DHFR expression level in LTMPR1-derived trimethoprim resistant isolates with indicated genotypes assessed by immunoblotting using anti-DHFR polyclonal antibody. FtsZ was used as a loading control. Quantitation was performed by calculating band intensities using image analysis. DHFR expression was normalized to the ancestor (LTMPR1). Mean of three independent measurements is provided. **C.** Cartoon representation of the structure of *E. coli* DHFR (PDB: 7DFR) bound to folate. Residues Trp₃₀, Tyr₁₅₁, Phe₁₅₃ and Ile₁₅₅ which form hydrophobic interactions and are required for proteolytic stability are shown as sticks and coloured by element (C: green, O:red, N:blue). Distances of less than or equal to 4 Å, which indicates possible interactions, are shown as dotted yellow traces. Folate bound in the active site of DHFR is shown as sticks. **D.** Schematic of the experiment used to test trimethoprim resistance and expression level of various DHFR mutants in *E. coli* wild type (WT) or *Δlon*. **E.** Trimethoprim IC₅₀ values of *E. coli* wild type or *E. coli* *Δlon* heterologously expressing DHFR (wt) or its mutants. Mean ± S.D. from three independent measurements is plotted. Statistical significance was tested using an unpaired t-test. Comparisons were between mutant and wild type DHFR, unless otherwise indicated. A p-value of <0.05 was considered statistically significant (*). A p-value > 0.05 was considered not statistically significant (ns). **F.** Expression level of plasmid-borne DHFR (wt) or its mutants in wild type or *Δlon* *E. coli* assessed by immunoblotting using anti-DHFR polyclonal antibody. FtsZ was used as a loading control. Quantitation was performed by calculating band intensities using image analysis. Expression of mutants was normalized to wild type DHFR (set to 1). Mean of three independent measurements is shown.

the effect of combination of mutations at these sites on trimethoprim resistance and DHFR expression, we first heterologously expressed wild type DHFR, Trp₃₀Arg or Tyr₁₅₁Asp single mutants and the Trp₃₀Arg-Tyr₁₅₁Asp double mutant in wild type *E. coli* (**Figure 8 D** [↗](#)). As expected, DHFR Trp₃₀Arg expression increased the IC₅₀ of trimethoprim. Surprisingly, neither DHFR Tyr₁₅₁Asp, nor the Trp₃₀Arg-Tyr₁₅₁Asp double mutant conferred resistance to the antibiotic (**Figure 8 E** [↗](#)). Indeed, both mutants conferred lower IC₅₀ values than wild type DHFR. This effect was due to loss of hydrophobicity at the Tyr₁₅₁ position, as DHFR-Tyr₁₅₁Leu and DHFR-Tyr₁₅₁Phe conferred similar IC₅₀ values as wild type DHFR (**Figure 8 E** [↗](#)). Since the Trp₃₀Arg-Tyr₁₅₁Asp mutation combination had evolved in a *lon*-deficient background, we next expressed these mutants in *E. coli Δlon* (**Figure 8 D** [↗](#)). Here too, DHFR Trp₃₀Arg conferred trimethoprim resistance, while the Tyr₁₅₁Asp mutant did not (**Figure 8 E** [↗](#)). Strikingly, the combination of Trp₃₀Arg and Tyr₁₅₁Asp conferred very high-level trimethoprim resistance (**Figure 8 E** [↗](#)) demonstrating synergistic epistasis that was contingent on the absence of Lon protease. The levels of expression of DHFR mutants in wild type and *lon*-deficient bacteria mirrored trimethoprim resistance levels indicating that clearance by Lon altered their genotype-phenotype relationship (**Figure 8 F** [↗](#)). Based on these data, we concluded that the action of proteostatic machinery generated additional demand for expression by clearing mutant-DHFR proteins. This additional demand drove the maintenance of *folA* duplications in bacteria evolving at very high drug pressures.

Discussion

Ohno's hypothesis, proposed in 1970, provided a conceptual framework to think about the evolution and eventual fate of duplicated genes [55 [↗](#)]. Several theoretical and experimental studies have built upon the original idea to delineate the evolutionary outcomes of gene duplication [18 [↗](#), 56 [↗](#), 57 [↗](#)]. More recently, experimental evolution of microbes has been used to test some of the suggestions made by Ohno and others. This approach has the inherent advantage of greater control over the selection pressure applied and availability of genetic tools to interrogate molecular mechanisms. In most studies, an artificial system was constructed and then evolved under defined selection pressures [33 [↗](#), 58 [↗](#)–61 [↗](#)]. Our study began with serendipitous isolation of spontaneous trimethoprim-resistant mutants of *E. coli* that overproduced DHFR because of a *folA* gene duplication. We exploited this system to understand the relationship between selection pressure and gene duplication in the context of drug-resistance evolution. The selection pressure for higher gene expression, i.e. “expression demand”, can be understood as the amount of a gene product that is required to maintain fitness in a particular environment. Antibiotics directly increase expression demand by inhibiting the activity of their targets. Our study shows that increasing gene dosage is a rapid, but transient evolutionary response to ensure that expression demand is met. On the other hand, point mutations are relatively stable routes to mitigating expression demand. Over short time scales, changes in gene dosage can counter the effects of antibiotic. However, their inherent instability results in almost inevitable replacement with phenotypically equivalent point mutations. Thus, continued antibiotic exposure delayed but could not prevent reversion to low gene copy number in most of our experiments. Further, we discover a novel association between proteostasis and gene dosage evolution. In *E. coli* evolving at high trimethoprim pressure, bacterial protein quality control exacerbated expression demand by clearing mutant DHFR-s. As a result, drug-resistant mutants of DHFR with low proteolytic stability drove the maintenance of *folA* gene duplication (**Figure 9** [↗](#)). Based on these findings we propose that a combination of factors, which include environmental and physiological, alter the relationship between gene copy number and fitness, and drive gene dosage evolution.

A key finding from our study was the two-fold influence that proteostatic machinery had on gene copy number. Firstly, Lon protease-deficiency promoted GDA mutations, possibly by allowing accumulation of IS-transposase proteins in *E. coli*. This observation resonates with earlier work by Nicoloff and co-workers who found that large IS-element flanked genomic duplications in

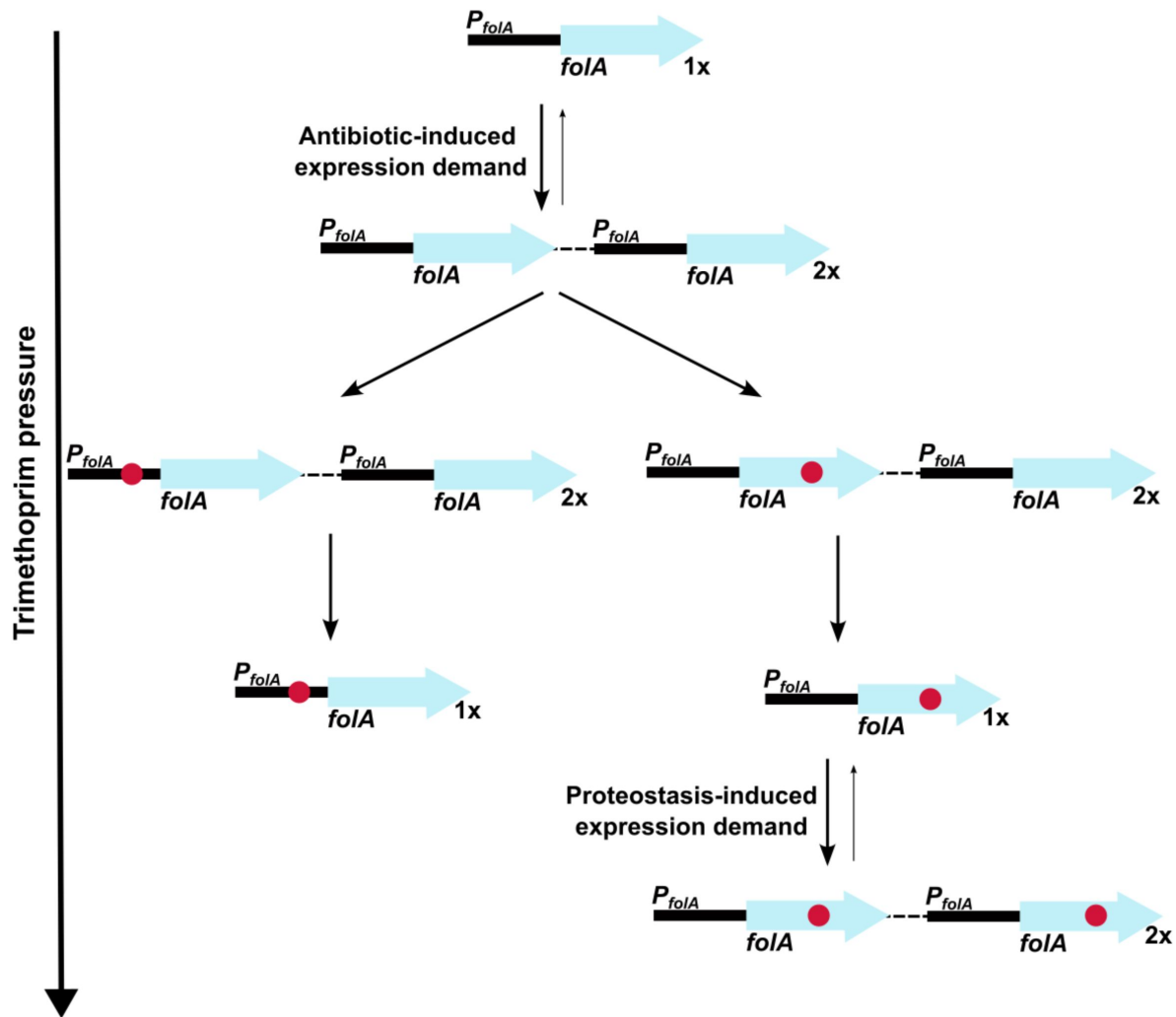


Figure 9.

Model of expression demand induced selection of *folA* gene duplication, followed by replacement by phenotypically equivalent point mutants. Expression demand for DHFR is generated by trimethoprim pressure, which results in the selection of *folA* gene duplications. Gene duplication is inherently unstable and reverses to a single copy of wild type *folA* when drug-pressure is withdrawn. Under drug pressure, *folA* gene duplication is maintained until a point mutation that confers resistance arises in one of the copies of *folA* (promoter or coding region). Upon acquisition of a point mutation, reversal of gene duplication ensues unless additional expression demand is generated by the action of proteostatic machinery on unstable drug-resistant DHFR mutants.

tetracycline-resistant *E. coli* were commonly preceded by spontaneous mutations that inactivated *lon* [62, 63]. Further, the role of Lon protease in regulating IS-element mobility by a similar mechanism has been previously demonstrated for IS1 and IS903 [64, 65]. We note however that other mechanisms by which Lon may modulate the frequency of IS-mediated GDAs in *E. coli* cannot be ruled out. For instance, Lon protease homologs from several organisms directly bind DNA [66–70], though the role of this binding in DNA transactions is unclear. Similarly, Lon is a regulator of the recombination protein, RecA, by its action on LexA [71] and in *Salmonella* RecA stimulated the rate of gene duplications mediated by IS3 [23]. The second impact of protein quality control on GDA was the action of proteases on trimethoprim-resistant mutants of DHFR, which promoted gene duplication by generating additional demand for expression (**Figure 9**). Importantly, which of these two effects dominated was determined by a combination of environment, i.e. drug concentrations, and other mutations acquired at the *folA* locus during evolution. In particular, mutations in the *folA* promoter could alleviate expression demand and hence reduce selection for high gene copy number.

The interplay between GDA and point mutations during evolution that we report here agrees in many respects with a recent study by Tomanek and Guet [61]. They use an elegantly designed artificial system [33, 61] in which expression of the *galK* gene in *E. coli* was driven by a very low activity promoter. This strain was then allowed to evolve in different concentrations of galactose. Like the present study, they found that the activity of IS-elements biased evolutionary trajectories in favor of duplications. They also showed that co-occurrence of duplications and point mutations was selected only at high expression demand, in agreement with our findings. Finally, they argued that gene duplication events may constrain evolution by point mutations. Our study corroborates these results at the short term in that the *lon*-knockout which showed a higher propensity for GDA mutations did not evolve point mutations in *folA* at early time points of evolution. However, given the relative low stability of large GDAs, these were inevitably lost over longer duration of evolution and replaced with point mutations. Thus, it is likely that the different mutational profiles of GDA-biased populations may be relevant only at short durations of evolution. It is important to note that though we have focused on the effects of mutations at the *folA* locus, which are directly under trimethoprim pressure, many of the isolates from our evolution experiments also harboured additional mutations at other loci. For instance, mutations in *rpoS* were seen throughout our evolution lines and at least in drug-free media were significant contributors to bacterial fitness. It is not yet clear what the role of other mutations are likely to be under antibiotic selection, though it may be envisaged that these additional mutations play the role of media adaptations. We did not take their effects into account since we expected them to be minor contributors to fitness especially during evolution under trimethoprim selection. However, their role in modulating the stability and selection of GDA mutations in drug-free/low drug environments may be more significant and remains to be investigated in the future.

In summary, the present study uncovers the complex relationship between gene copy number, point mutations and environment, and identifies proteostasis as a regulator of gene dosage evolution in bacteria. These findings open up the possibility of alternative explanations for how gene duplications may be maintained over long periods of evolution and provides insight into how gene expression demand is mitigated by multiple mechanisms in bacteria. Moreover, our findings relating proteostasis in general and Lon in particular, to adaptive mutations add to a growing body of literature that define a causal link between protein quality control and mutational paths of evolution. For DHFR itself, decreasing the stringency of protein quality control changes the fitness landscape of single mutations and their combinations, altering the mutational repertoires accessible during adaptive evolution [49, 50, 72, 73]. Similar effects of protein homeostatic machinery on mutant fitness are reported from a diverse range of proteins that influence phenomena like viral evolution [74] and the fate of horizontally acquired genes [75–77]. Indeed, a central role for protein quality control machinery in molecular and organismal evolution is reflected in conserved signatures of the proteostatic network across the different kingdoms of life [78]. These links may be important to consider for a new wave of

potential antibacterials that target proteostasis [79–81]. It appears critical to ensure that the impact of changes in evolvability are considered when designing such strategies for therapeutic intervention.

Materials and methods

Strains, plasmids and culture conditions

Escherichia coli K-12 MG1655 was used as wild type for all experiments. *E. coli* DH10B was used for plasmid construction and manipulation. *E. coli* strains were cultured in Luria-Bertani (LB) broth with shaking at 180–200 rpm or on LB agar plates at 37 °C. Ampicillin was used at a concentration of 100 µg/mL for selection of plasmid-transformed strains. Trimethoprim was added to growth media at required concentrations as needed. The strains and plasmids used in this study are listed in Table 2.

Laboratory evolution of antibiotic resistance

Short-term evolution of antibiotic resistance

For short term laboratory evolution of antibiotic resistance, wild type and *E. coli* *Δlon* populations were cultured in 96-well plates in LB medium incorporating sub-inhibitory concentrations of different antibiotics. Following concentrations of antibiotics were used: trimethoprim (300 ng/mL; MIC/3); spectinomycin (5 µg/mL; MIC/10); nalidixic acid (4.25 µg/mL; MIC/3); erythromycin (10 µg/mL; MIC/10); rifampicin (4.25 µg/mL; MIC/3). Control populations were also set up in antibiotic-free LB. Five replicates were included for each strain in each antibiotic. Total volume of culture in each well was 150 µL. 10 % of the population was passaged to fresh growth media every 24 hours, for 7 days, resulting in ~25 generations of evolution. At the end of 7 days, the evolved populations were frozen as glycerol stocks.

To check for the evolution of resistance and isolation of resistant clones for further analyses, frozen populations were revived in drug-free LB broth and grown overnight (14–16 hours). Saturated cultures were serially diluted and spotted on LB agar plates supplemented with the appropriate antibiotic at MIC. Colonies were randomly picked, cultured overnight in drug-free LB and frozen in 50% glycerol at -80 °C until further use.

Long-term evolution at different trimethoprim pressures

For long-term evolution, trimethoprim-resistant isolates LTMPR1, WTMPR4 and WTMPR5 were first revived from frozen stocks in drug-free LB overnight. Triplicate evolving lineages were then established by sub-culturing 1% of the saturated culture of each ancestor into 2 mL of appropriate medium. For relaxed selection drug-free LB was used. For constant antibiotic pressure, trimethoprim was added at a concentration of 300 ng/mL. For increasing antibiotic pressure, trimethoprim was increased from 1 µg/mL to 10 µg/mL over course of the experiment as shown in Figure 7B. Cultures were grown for 10–14 hours (~7 generations) and then passaged (1%) into fresh media. This was continued for 36 passages, i.e. ~252 generations. Aliquots of cultures were frozen periodically for further analyses.

For isolating individual clones, populations were revived in drug-free LB from frozen stocks at 252 generations of evolution in drug-free LB, serially diluted and then spotted onto LB agar plates supplemented with trimethoprim (1, 5 or 10 µg/mL). Plates were incubated for 18–20 hours at 37 °C and resistant colonies were picked randomly for further analyses.

Sr. No.	Strain	Description	Source
1.	<i>Escherichia coli</i> K-12 MG1655	Wild type <i>E. coli</i> ; used as ancestor for evolution experiments and as reference strains for phenotypic and genotypic comparisons	Laboratory stocks
<i>Gene knockouts generated in E. coli K-12 MG1655</i>			
1.	<i>E. coli</i> $\Delta mgrB$	Isogenic knockout of <i>mgrB</i> gene ($\Delta mgrB::Kan$)	[46]
2.	<i>E. coli</i> $\Delta lacZ$	Isogenic knockout of <i>lacZ</i> gene ($\Delta lacZ::Cat$)	[46]
3.	<i>E. coli</i> Δlon	Isogenic knockout of <i>lon</i> gene ($\Delta lon::Kan$)	[49]
4.	<i>E. coli</i> $\Delta lon \Delta lacZ$	Double knockout of <i>lon</i> ($\Delta lon::Kan$) and <i>lacZ</i> ($\Delta lacZ::Cat$) genes	[46]
5.	<i>E. coli</i> $\Delta mgrB \Delta lon$	Double knockout of <i>lon</i> ($\Delta lon::Cat$) and <i>mgrB</i> ($\Delta mgrB::Kan$) genes	[46]
6.	<i>E. coli</i> $\Delta lon \Delta rpoS$	Double knockout of <i>lon</i> ($\Delta lon::Cat$) and <i>rpoS</i> ($\Delta rpoS::Kan$) genes	[46]
7.	<i>E. coli</i> $\Delta pitA$	Isogenic knockout of <i>pitA</i> gene ($\Delta pitA::Kan$)	This study.
<i>Strains/Isolates derived from short-term adaptive laboratory evolution</i>			

Table 2.

List of strains and plasmids used in the study

1.	WTMPR1, WTMPR2, WTMPR3, WTMPR4 and WTMPR5	Trimethoprim-resistant derivatives of <i>E. coli</i> wild-type obtained by laboratory evolution at 300 ng/ml of trimethoprim for ~25 generation, followed by selection on agar plates supplemented with 1 µg/mL of trimethoprim	[46] Note: These strains are labelled as TMPR1-5 in the original publication. They have been renamed WTMPR1-5 to indicate that they were derived after evolution of wild type.
2.	LTMPR1, LTMPR2, LTMPR3, LTMPR4 and LTMPR5	Trimethoprim-resistant derivatives of <i>E. coli</i> Δlon obtained by laboratory evolution at 300 ng/ml of trimethoprim for ~25 generation, followed by selection on agar plates supplemented with 1 µg/mL of trimethoprim	This study
<i>Plasmids</i>			
1.	pPRO- <i>insB</i>	<i>InsB</i> gene cloned into the pPROEx-HtC expression plasmid in frame-with an N-terminal hexa-histidine tag.	This study
2.	pPRO- <i>insL</i>	<i>InsL</i> gene cloned into the pPROEx-HtA expression plasmid in frame-with an N-terminal hexa-histidine tag.	This study
3.	pPRO- <i>folA</i>	<i>FolA</i> gene cloned into the pPROEx-HtB expression plasmid in frame-with an N-terminal hexa-histidine tag.	[54]
4.	pPRO- <i>folA</i> Trp ₃₀ Arg	<i>FolA</i> gene cloned into the pPROEx-HtB expression plasmid, harbouring the Trp ₃₀ Arg mutation in frame-with an N-terminal hexa-histidine tag.	[54]

Table 2. (continued)

5.	pPRO- <i>folA</i> Tyr ₁₅₁ Asp	<i>FolA</i> gene cloned into the pPROEx-HtB expression plasmid, harbouring the Tyr ₁₅₁ Asp mutation in frame-with an N-terminal hexa-histidine tag.	This study
6.	pPRO- <i>folA</i> Trp ₃₀ Arg Tyr ₁₅₁ Asp	<i>FolA</i> gene cloned into the pPROEx-HtB expression plasmid, harbouring the Trp ₃₀ Arg and Tyr ₁₅₁ Asp mutations in frame-with an N-terminal hexa-histidine tag.	This study
7.	pPRO- <i>folA</i> Tyr ₁₅₁ Phe	<i>FolA</i> gene cloned into the pPROEx-HtB expression plasmid, harbouring the Tyr ₁₅₁ Phe mutation in frame-with an N-terminal hexa-histidine tag.	This study
8.	pPRO- <i>folA</i> Tyr ₁₅₁ Leu	<i>FolA</i> gene cloned into the pPROEx-HtB expression plasmid, harbouring the Tyr ₁₅₁ Leu mutation in frame-with an N-terminal hexa-histidine tag.	This study

Table 2. (continued)

Measuring trimethoprim resistance

Trimethoprim IC50

Trimethoprim Inhibitory Concentration - 50 (IC50), i.e. the antibiotic concentration needed for 50 % growth inhibition, was calculated using a broth dilution assay. Frozen glycerol stocks of appropriate bacterial strains were revived in 1-3 mL LB for 6-8 hours. In a 96-well plate, 150 µL of LB was added to individual wells. Trimethoprim stocks were added, and the antibiotic was serially diluted. No antibiotic was added to a few wells, which served as controls. Next, 1.5 µL of the appropriate bacterial culture was added to each well. The peripheral wells of the plate were filled with water to prevent dehydration. The 96-well plate was then incubated at 37°C for 15-18 hrs. Optical density at 600 nm was measured for each well using a plate reader (Ensign, Perkin Elmer). Obtained data were first normalised to growth in drug-free wells (set to 1) and then fitted to a sigmoidal log(inhibitor)-response curve using Graphpad Prism software. The equation for this model is as follows:

$$Y = \text{Min} + (\text{Max} - \text{Min}) / (1 + 10^{(X - \text{Log}(\text{IC}_{50}))})$$

where, Min: Minimum value of normalised OD and Max: maximum value of normalised OD Values of IC50 were estimated from the above curve fit and averaged across 3 independent measurements for every strain.

Colony forming efficiency

To calculate the colony forming efficiency of evolving populations at different trimethoprim concentrations, frozen glycerol stocks were first revived in 1 mL LB for 14-15 hours. Saturated cultures were serially diluted (10-fold). Diluted cultures were then spotted (10 µL) on LB agar plates containing 0, 1, 5 and 10 µg/mL of trimethoprim. Spots were allowed to dry completely and the plates were incubated at 37 °C for ~18 hours. Colonies were counted from the least dense spot and multiplied by the appropriate dilution factor to estimate colony forming units per mL (CFUs/mL) of culture.

Growth curves

Appropriate frozen stocks were revived overnight in drug-free LB (1-3 mL) and 1 µL of the saturated culture was added to 200 µL of media supplemented with trimethoprim (300 ng/mL) in the wells of a sterile uncoated 96-well plate. Optical density at 600 nm was measured during growth at 37 °C at 15-minute intervals over a 24-hour period on a Perkin-Elmer Ensign Multimode Plate Reader. Doubling time (DT) (in mins) was estimated from log phase, using the following formula:

$$DT = \ln(2) / \left(\frac{\ln(\frac{OD_2}{OD_1})}{T_2 - T_1} \right)$$

Where OD1 and OD2 represent optical densities at the onset and end of log phase respectively, and T1 and T2 represent time elapsed (in mins), corresponding to OD1 and OD2.

Genome sequencing, identification of mutations and detection of GDA mutations

Genomic DNA was extracted from late log-phase cultures of appropriate strain/resistant isolates/evolving populations using phenol:chloroform:isoamyl alcohol (25:24:1) extraction as previously described [82]. Extracted genomic DNA was then cleaned-up using HiGemoMB DNA

isolation and purification kit (Himedia, India) as per manufacturer's instructions and quantitated spectrophotometrically. The integrity was checked by running purified DNA on an agarose gel before further analyses.

For whole-genome resequencing, paired-end whole-genome next-generation sequencing (NGS) was performed on a MiSeq system (Illumina, USA) with read lengths of 150-250 bps. Library preparation and sequencing services were provided by Eurofins (Bangalore, India) and Genotypic (Bangalore, India). Processed reads were aligned to the reference genome *E. coli* K-12 MG1655 (U00096) using bowtie2. All sequenced genomes had an average coverage depth of at least 100x. Variant calling and prediction of new junctions in the genome to identify large structural mutations was done using Breseq [83] using default settings in population or clonal modes as needed. All variants and new junctions that were already present in the ancestor were excluded. Variants unique to the test sample were classified as "mutations".

For detection of GDA mutations, 'coverage' vs 'coordinate in reference genome' plot generated by the bam.cov command in Breseq was used [83]. Contiguous stretches greater than 10 kb with coverage depth of 1.5x or more, compared to the rest of the genome, were classified as GDA mutations. The presence of IS-elements or repeat sequences was manually checked using the coordinates of the GDA.

Calculating relative fitness of *E. coli* strains

Relative fitness of gene knockouts or evolved clones was estimated using competition assays. Saturated cultures of test strains and reference strains (*E. coli* $\Delta lacZ$ or *E. coli* $\Delta lacZ \Delta lon$) were mixed 1:1 to a total of 1% in 3 mL LB. Strains were allowed to compete for ~7 generations (24 hours incubation). Mixed cultures were diluted and plated on LB agar supplemented with IPTG (1 mM) and X-gal (50 μ g/mL) at the start and end of the competition experiment. Plates were incubated for 18-24 hrs, and blue and white colonies were counted to determine the CFUs/mL at initial and final time points.

Relative fitness (w) was calculated using the following formula:

$$w = \frac{\ln \left(\frac{T_f}{T_i} \right)}{\ln \left(\frac{R_f}{R_i} \right)}$$

where T_f and T_i are the final and initial CFUs/mL of the test strain and R_f and R_i are the final and initial CFUs/mL of the reference strain, respectively.

Estimation of DHFR expression by immunoblotting

Immunoblotting was used to determine the expression level of DHFR in various strains/mutants of *E. coli* as described in Patel and Matange (2021) [46]. Briefly, equal lysate protein (5 μ g) was subjected to SDS-PAGE on a 15% gel and electroblotted onto PVDF membranes. The membrane was blocked with 5% BSA solution and cut into 2 halves at the 25 kDa molecular weight band. The lower part of the membrane was probed with anti-DHFR polyclonal IgG (100 ng/mL) and the upper part was probed with anti-FtsZ serum (1:50,000). HRP-linked anti-rabbit IgG (1:10,000) was used as the secondary antibody. Hybridised antibodies were detected using chemiluminescence. Band intensities were quantitated in ImageJ and normalised to appropriate control (set to 1). Fold change in DHFR was calculated as follows:

$$\text{Fold change}_{DHFR} = \frac{\text{Normalised intensity}_{DHFR}}{\text{Normalised intensity}_{FtsZ}}$$

Gene expression analysis using RNA-sequencing

Transcriptome analyses were performed using RNA sequencing. For RNA-seq, saturated overnight culture of appropriate strains/isolates was inoculated (1 %) into 3 mL LB in triplicate and incubated at 37 °C for 3 hours with shaking at 180 rpm. The cultures were centrifuged and bacterial pellets were resuspended in 1 mL of RNeasy lysis buffer (Qiagen) for storage. RNA extraction, sequencing and preliminary data analyses were performed by Eurofins (Bangalore, India). Total RNA was extracted using RNeasy spin columns and quantitated using Qubit and Bioanalyzer. Bacterial rRNA was depleted using RiboZero kit. Bacterial mRNA was then reverse transcribed, library prepared and quality control performed using TapeStation platform. Paired end sequencing was performed on Illumina platform with read lengths of 150 bp read length. Processed reads were aligned to the reference genome (U00096.3) using HISAT2 (version 2.1.0). Abundance estimation was done using featureCounts (version 1.34.0). Differential gene expression (DGE) analysis was done by comparing the expression of individual genes in test and reference samples. The output of the DGE analyses were $\log_2$ (fold change) values for each gene and P-values for statistical significance.

Cloning and site-directed mutagenesis

Cloning of *InsB* and *InsL* transposases

InsB was amplified from *E. coli* K-12 MG1655 genomic DNA by PCR using primers *insB1_fwd_NcoI* (5'-GGCGCCATGGATATGCCGGGCAACAGCCG-3') *insB1_rev_EcoRI* (5'-CCTTTGAATTCTTATTGATAGTGTGTTTATG-3'). For amplification of *insL*, primers *insL_EcoRI_fwd* (5'-ATGCCGAATTCATGAATTACTCTCACGATAAC-3') and *insL_HindIII_rev* (5'-ATGCCAAGCTTTAGTTCTTCTTTTCGGATCC-3') were used. Amplified *insB* gene was digested with *NcoI* and *EcoRI* restriction endonuclease and ligated with similarly digested pPROEx-HtC plasmid. Amplified *insL* was digested with *EcoRI* and *HindIII* restriction endonucleases and ligated with similarly digested pPROEX- HTa plasmid. Positive clones were screening using restriction digestion. Clones were transformed into *E. coli* wild type or Δlon for further analyses.

Site-directed mutagenesis for DHFR

Site-directed mutagenesis was performed to introduce mutations into the coding sequence of *folA* in pPRO-folA as described in [54]. Briefly, PCR was performed using pPRO-folA or pPRO-folA *Trp₃₀Arg* as template and the appropriate forward and reverse mutagenic primers given below:

folA_Tyr₁₅₁Asp_fwd_BglII:

5'-CTCTCACAGCGATTGCTTTGAGATCTTGGAGCGGCGGGGCC-3'

folA_Tyr₁₅₁Asp_rev_BglII:

5'-CGCCGCTCCAAGATCTCAAAGCAATCGCTGTGAGAGTTCTGCG-3'

folA_Tyr₁₅₁Leu_fwd_BglII:

5'-TCTCACAGCTTATGCTTTGAGATCTTGGAGCGGCGGGGCC-3'

folA_Tyr₁₅₁Leu_rev_BglII:

5'-CGCCGCTCCAAGATCTCAAAGCATAAGCTGTGAGAGTTCTGCG-3'

folA_Tyr₁₅₁Phe_fwd_BgIII:

5'- CTCACAGCTTTTGCTTTGAGATCTTGGAGCGGCGGGGCC-3'

folA_Tyr₁₅₁Phe_rev_BgIII:

5'- GCCGCTCCAAGATCTCAAAGCAAAAGCTGTGAGAGTTCTGCG-3'

PCR products were digested with *DpnI* overnight, transformed into chemically competent *E. coli* DH10B and selected on LB agar plates containing ampicillin. Plasmids harbouring the desired mutations were screened by restriction digestion and confirmed by sequencing. Sequenced clones were transformed into *E. coli* wild type or *Δlon* for further analyses.

In-cell stability of InsB and InsL

A chloramphenicol-chase assay was used to assess the in-cell stability of heterologously expressed InsB and InsL transposases as described in Matange (2020) [49] with a few modifications. Overnight cultures of *E. coli* wild type or *Δlon* harbouring plasmids pPRO- insB or pPRO-insL were diluted 1% in 5 mL LB supplemented with ampicillin and IPTG (0.1mM). After 4 hours of growth at 37 °C, 1 mL of culture was pelleted. The cell pellet was lysed in 100 μL of 1x Laemmli sample loading buffer by heating at 95°C for 10-20 mins. Chloramphenicol was added to the remaining culture at a final concentration of 25 μg/mL to stop protein synthesis and the culture was placed back in the incubator. After 15, 45, and 90 minutes 1 mL aliquots of the culture were pelleted and lysed. Lysates were subjected to SDS- PAGE followed by immunoblotting to detect the levels of InsB and InsL. Anti-Hexahis monoclonal antibody (Invitrogen) was used as the primary antibody for InsL. Anti-Hexahis polyclonal antibody (Invitrogen) was used as the primary antibody for InsB. Band intensities were calculated in ImageJ and normalised to control (set to 1) to calculate fraction of protein remaining after chloramphenicol treatment.

Supplementary Figures

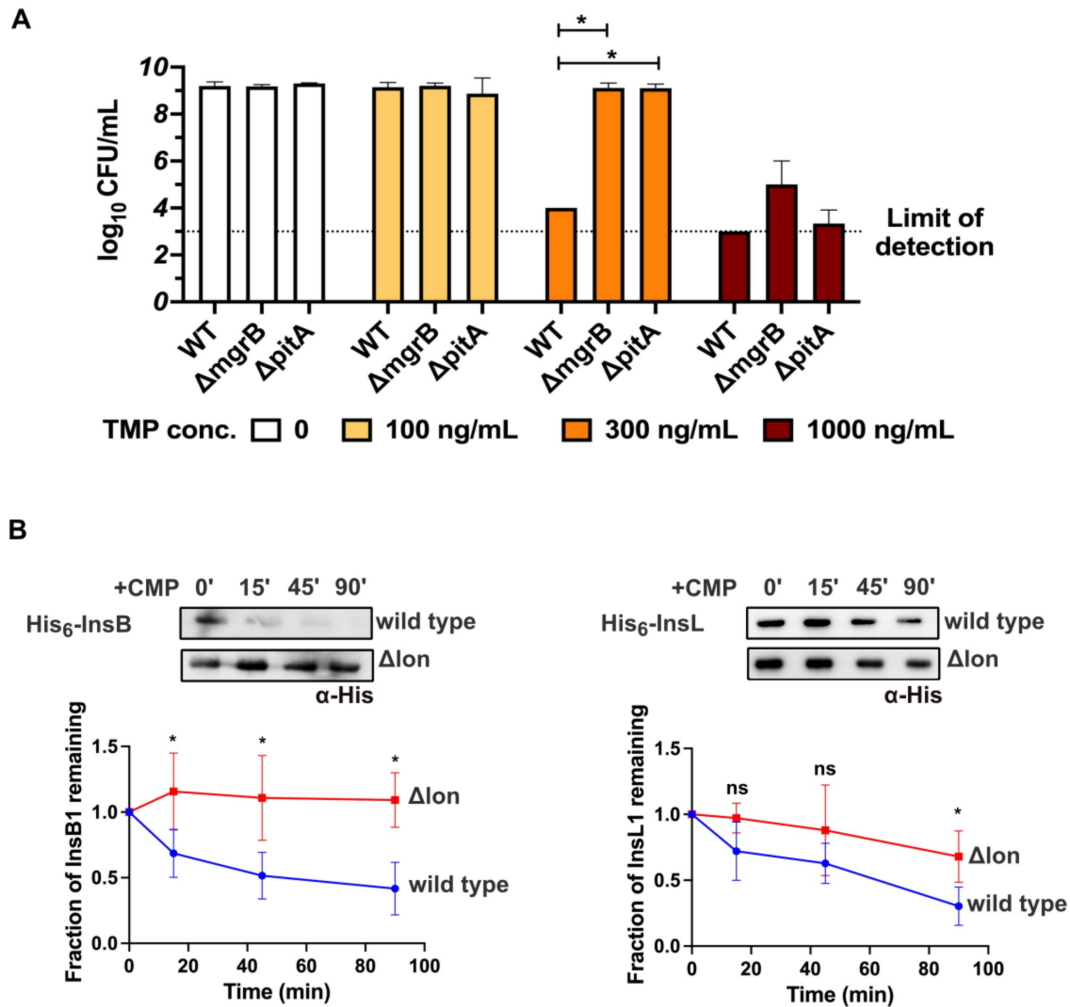


Figure S1.

A. Loss of *mgrB* or *pitA* genes is beneficial in trimethoprim. Colony formation of *E. coli* wild type, *ΔmgrB* and *ΔpitA* on media supplemented with trimethoprim at the indicated concentrations. Mean $\pm$ S.D. from three independent replicates are plotted. Statistical significance was tested using a t-test and p-value < 0.05 is indicated with an asterisk (*). **B.** Stability of heterologously expressed InsB and InsL transposases in wild type or *Δlon* *E. coli* measured using a chloramphenicol (CMP) chase assay. Levels of the plasmid-expressed transposase were measured using immunoblotting with an anti-His monoclonal/polyclonal antibody in lysates prepared from cells at indicated time points following treatment with 50 μ g/mL of CMP. A representative image from three independent experiments is shown. Band intensities were quantitated by image analyses and normalized to 0' (set to 1). The fraction of protein remaining is plotted in blue for wild type and red for *E. coli* *Δlon*. Values from 3 independent experiments are plotted as mean $\pm$ S.D. Statistical significance was tested using an unpaired t-test. A p-value of < 0.05 was considered statistically significant (*). A p-value > 0.05 was considered not statistically significant (ns).

Representative coverage plots of *E. coli* Δlon evolved in no antibiotic or spectinomycin

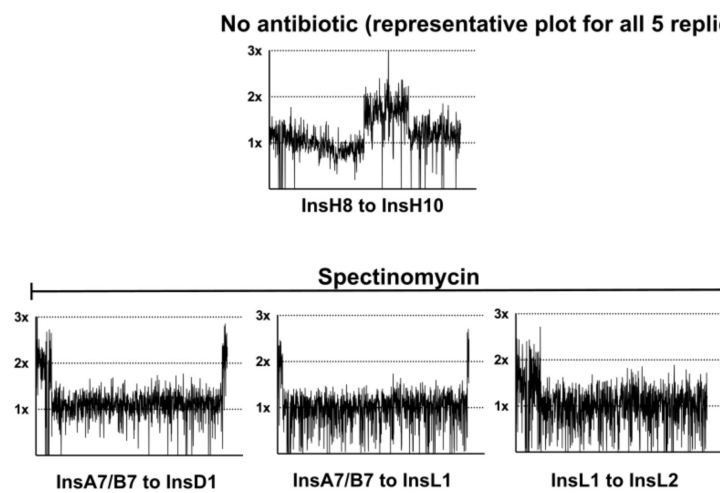


Figure S2.

Coverage plots and flanking IS elements for *E. coli* Δlon populations evolving in no antibiotic or spectinomycin that showed GDA mutations. For no-antibiotic, only one of the 5 replicates is shown as similar coverage plots were obtained for all 5 replicates.

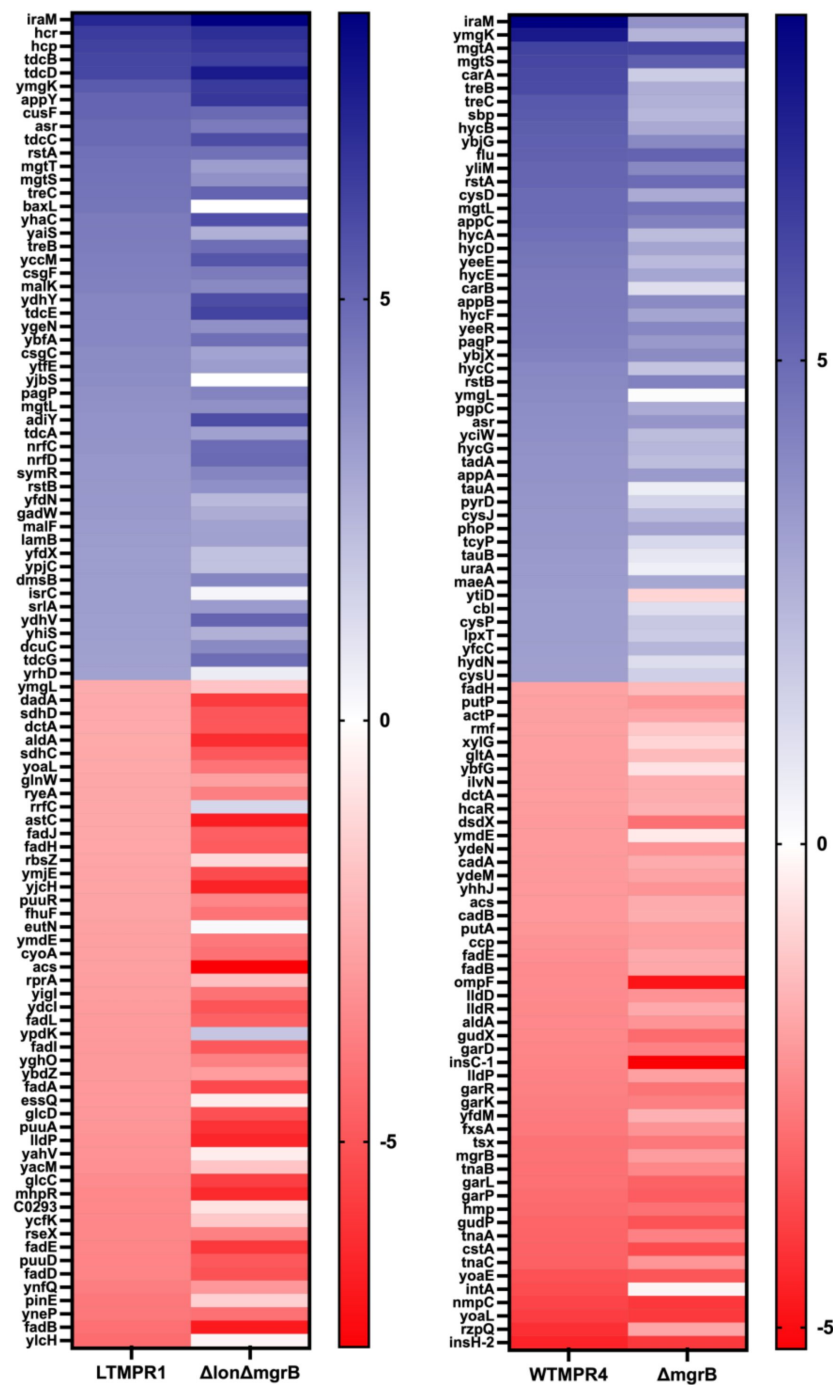


Figure S3.

Gene expression changes associated with *foIA*-encompassing duplication.

Heat map comparing the top 50 up/downregulated genes in LTMPR1 and WTMPR4 with *E. coli* $\Delta lon\Delta mgrB$ and *E. coli* $\Delta mgrB$ respectively. $\log_2$ (Fold change) values were calculated based on RNA-sequencing with respect to wild type or Δlon *E. coli* and represented on a green/red scale as indicated. Gene names are shown adjacent to the appropriate cell. Genes are ranked according to their fold change in LTMPR1 and WTMPR4.

Figure S4.

Coverage depth plots for population sequencing at 105, 196 and 252 generations of the 3 lineages of LTMPR1 (A, B, C) evolving in trimethoprim-free media. Coverage depth plots show the number of reads from Illumina sequencing (y-axis) mapped to each genomic coordinate of the reference genome (x-axis). Coverage of 1x, 2x and 3x are marked by dotted lines for reference.

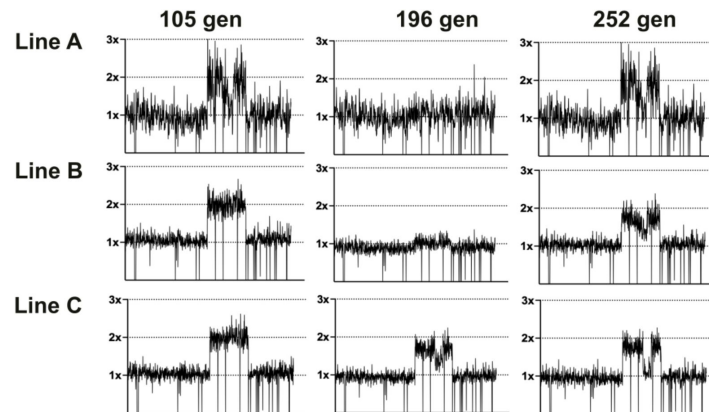


Figure S5.

Coverage depth plots for population sequencing at 105, 196 and 252 generations of the 3 lineages of LTMPR1 (A, B, C) evolving in trimethoprim-supplemented (300 ng/mL) media. Coverage depth plots show the number of reads from Illumina sequencing (y-axis mapped) to each genomic coordinate of the reference genome (x-axis). Coverage of 1x, 2x and 3x are marked by dotted lines for reference.

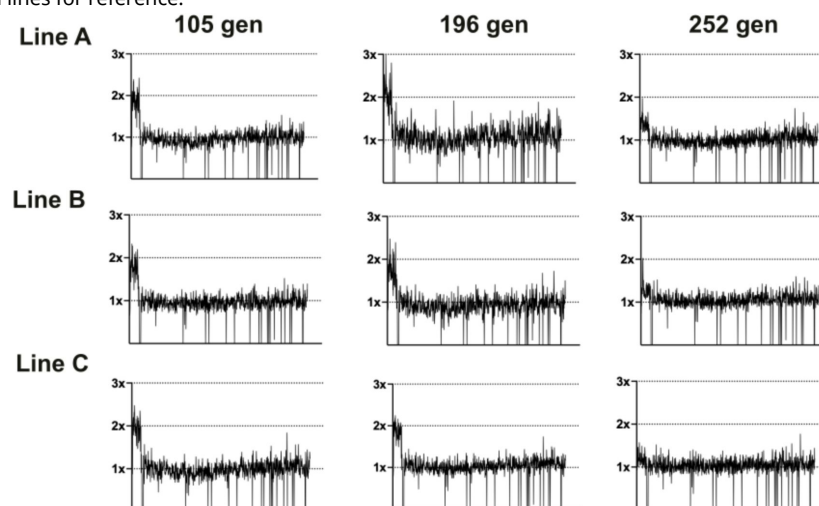


Figure S6.

Coverage plots for Line C-Isolate 5 and Line B-Isolate 2. Line B Isolate 3 showed 5x amplification of a shorter genomic stretch encompassing *folA* which is shown diagrammatically next to the coverage plot. Line C Isolate 5 showed an expanded GDA encompassing *folA* compared to the LTMPR1 ancestor, the coordinates for which are below the coverage plot.

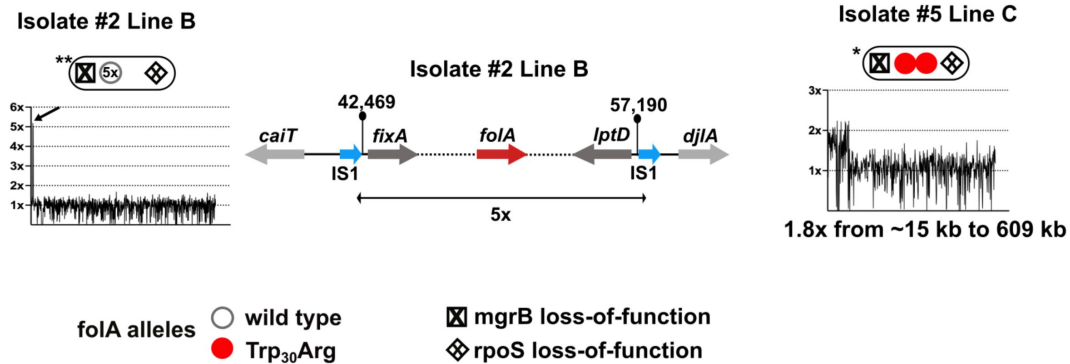
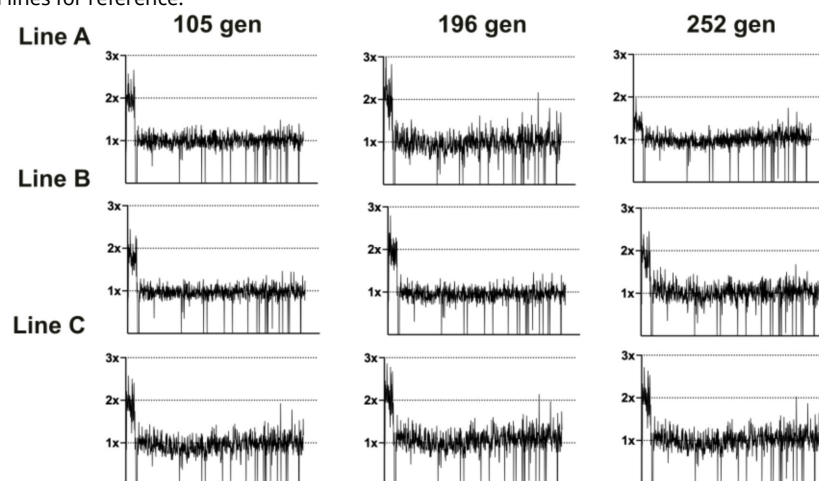


Figure S7.

Coverage depth plots for population sequencing at 105, 196 and 252 generations of the 3 lineages of LTMPR1 (A, B, C) evolving in increasing trimethoprim concentrations. Coverage depth plots show the number of reads from Illumina sequencing (y-axis) mapped to each genomic coordinate of the reference genome (x-axis). Coverage of 1x, 2x and 3x are marked by dotted lines for reference.



Acknowledgements

We acknowledge Dr. Manjula Reddy for providing us with the anti-FtsZ polyclonal antibody. We acknowledge Bruhad Dave for his valuable inputs during discussions at early stages of this project and Prof. Satyajit Rath, Prof. Sandhya Visweswariah and Dr. Girish Deshpande for their comments on the manuscript. Funding for this work was provided by DBT/Wellcome Trust India Alliance. NM is a recipient of the India Alliance Intermediate Fellowship. CJ is a recipient of a Senior Research Fellowship from the Council of Scientific and Industrial Research (CSIR), Govt. of India.

Additional files

Supplementary File 1. *Gene expression changes in LTMPR1 and WTMPR4 isolates.* [↗](#)

Supplementary File 2. *Genomic changes associated with evolution of LTMPR1 in drug-free media.* [↗](#)

Supplementary File 3. *Gene expression changes in isolates derived from LTMPR1 evolution in drug-free media containing mutation in *rpoS* or *GDA-2*.* [↗](#)

Supplementary File 4. *Genomic changes associated with evolution of LTMPR1 in trimethoprim (300 ng/mL).* [↗](#)

Supplementary File 5. *Genomic changes associated with evolution of LTMPR1 in increasing trimethoprim pressure.* [↗](#)

Supplementary File 6. *Genome sequencing of isolates derived from WTMPR4 evolution at different trimethoprim pressures.* [↗](#)

Supplementary File 7. *Genome sequencing of isolates derived from WTMPR5 evolution at different trimethoprim pressures.* [↗](#)

References

1. Andersson DI, Hughes D (2009) **Gene Amplification and Adaptive Evolution in Bacteria** *Annu Rev Genet* **43**:167–95 <https://doi.org/10.1146/annurev-genet-102108-134805>
2. Elliott KT, Cuff LE, Neidle EL (2013) **Copy number change: evolving views on gene amplification** *Future Microbiol* **8**:887–99 <https://doi.org/10.2217/fmb.13.53>
3. Kondrashov FA (2012) **Gene duplication as a mechanism of genomic adaptation to a changing environment** *P Roy Soc B-Biol Sci* **279**:5048–57 <https://doi.org/10.1098/rspb.2012.1108>
4. Dürrebaum M, Storchová Z (2016) **Effects of aneuploidy on gene expression: implications for cancer** *Febs J* **283**:791–802 <https://doi.org/10.1111/febs.13591>
5. Jiang Z, Li CX, Li FY, Wang XS (2013) **EGFR Gene Copy Number as a Prognostic Marker in Colorectal Cancer Patients Treated with Cetuximab or Panitumumab: A Systematic Review and Meta Analysis** *Plos One* **8** <https://doi.org/10.1371/journal.pone.0056205>
6. Spielmann M, Lupiáñez DG, Mundlos S (2018) **Structural variation in the 3D genome** *Nat Rev Genet* **19**:453–67 <https://doi.org/10.1038/s41576-018-0007-0>
7. Pascual-Anaya J, D’Aniello S, Kuratani S, Garcia-Fernández J (2013) **Evolution of gene clusters in deuterostomes** *Bmc Dev Biol* **13** <https://doi.org/10.1186/1471-213x-13-26>
8. Taylor JS, Van de Peer Y, Meyer A (2001) **Genome duplication, divergent resolution and speciation** *Trends Genet* **17**:299–301 [https://doi.org/10.1016/S0168-9525\(01\)02318-6](https://doi.org/10.1016/S0168-9525(01)02318-6)
9. Ting CT, Tsaur SC, Sun S, Browne WE, Chen YC, Patel NH, et al. (2004) **Gene duplication and speciation in Drosophila: Evidence from the Odysseus locus** *P Natl Acad Sci USA* **101**:12232–5 <https://doi.org/10.1073/pnas.0401975101>
10. Jiao WB, Patel V, Klasen J, Liu F, Pecinkova P, Ferrand M, et al. (2021) **The Evolutionary Dynamics of Genetic Incompatibilities Introduced by Duplicated Genes in Arabidopsis thaliana** *Mol Biol Evol* **38**:1225–40 <https://doi.org/10.1093/molbev/msaa306>
11. Maclean CJ, Greig D (2011) **Reciprocal Gene Loss Following Experimental Whole-Genome Duplication Causes Reproductive Isolation in Yeast** *Evolution* **65**:932–45 <https://doi.org/10.1111/j.1558-5646.2010.01171.x>
12. Zuellig MP, Sweigart AL (2018) **Gene duplicates cause hybrid lethality between sympatric species of Mimulus** *Plos Genet* **14** <https://doi.org/10.1371/journal.pgen.1007130>
13. Sandegren L, Andersson DI (2009) **Bacterial gene amplification: implications for the evolution of antibiotic resistance** *Nat Rev Microbiol* **7**:578–88 <https://doi.org/10.1038/nrmicro2174>
14. Anderson P, Roth J (1981) **Spontaneous tandem genetic duplications in Salmonella typhimurium arise by unequal recombination between rRNA (rrn) cistrons** *Proc Natl Acad Sci U S A* **78**:3113–7 <https://doi.org/10.1073/pnas.78.5.3113>

15. Seaton SC, Elliott KT, Cuff LE, Laniohan NS, Patel PR, Neidle EL (2012) **Genome-wide selection for increased copy number in *Acinetobacter baylyi* ADP1: locus and context-dependent variation in gene amplification** *Mol Microbiol* **83**:520–35 <https://doi.org/10.1111/j.1365-2958.2011.07945.x>
16. Sonti RV, Roth JR (1989) **Role of gene duplications in the adaptation of *Salmonella typhimurium* to growth on limiting carbon sources** *Genetics* **123**:19–28 <https://doi.org/10.1093/genetics/123.1.19>
17. Hahn MW (2009) **Distinguishing Among Evolutionary Models for the Maintenance of Gene Duplicates** *J Hered* **100**:605–17 <https://doi.org/10.1093/jhered/esp047>
18. Innan H, Kondrashov F (2010) **The evolution of gene duplications: classifying and distinguishing between models** *Nat Rev Genet* **11**:97–108 <https://doi.org/10.1038/nrg2689>
19. Kuzmin E, Taylor JS, Boone C (2022) **Retention of duplicated genes in evolution** *Trends Genet* **38**:59–72 <https://doi.org/10.1016/j.tig.2021.06.016>
20. Hufnagel DA, Choby JE, Hao S, Johnson AF, Burd EM, Langelier C, et al. (2021) **Antibiotic-Selected Gene Amplification Heightens Metal Resistance** *mBio* **12** <https://doi.org/10.1128/mBio.02994-20>
21. von Rozycki T, Nies DH (2009) ***Cupriavidus metallidurans*: evolution of a metal-resistant bacterium** *Antonie Van Leeuwenhoek* **96**:115–39 <https://doi.org/10.1007/s10482-008-9284-5>
22. Bi X, Liu LF (1994) **RecA-independent and recA-dependent intramolecular plasmid recombination: Differential homology requirement and distance effect** *J Mol Biol* **235**:414–23 <https://doi.org/10.1006/jmbi.1994.1002>
23. Reams AB, Kofoed E, Kugelberg E, Roth JR (2012) **Multiple Pathways of Duplication Formation with and Without Recombination (RecA) in *Salmonella enterica*** *Genetics* **192**:397–415 <https://doi.org/10.1534/genetics.112.142570>
24. Reams AB, Kofoed E, Savageau M, Roth JR (2010) **Duplication frequency in a population of *Salmonella enterica* rapidly approaches steady state with or without recombination** *Genetics* **184**:1077–94 <https://doi.org/10.1534/genetics.109.111963>
25. Reams AB, Neidle EL (2003) **Genome plasticity in *Acinetobacter*: new degradative capabilities acquired by the spontaneous amplification of large chromosomal segments** *Molecular Microbiology* **47**:1291–304 <https://doi.org/10.1046/j.1365-2958.2003.03342.x>
26. Min YH, Kwon AR, Yoon JM, Yoon EJ, Shim MJ, Choi EC (2008) **Molecular analysis of constitutive mutations in *ermB* and *ermA* selected in vitro from inducibly MLSB-resistant enterococci** *Arch Pharm Res* **31**:377–80 <https://doi.org/10.1007/s12272-001-1167-8>
27. Sun S, Berg OG, Roth JR, Andersson DI (2009) **Contribution of gene amplification to evolution of increased antibiotic resistance in *Salmonella typhimurium*** *Genetics* **182**:1183–95 <https://doi.org/10.1534/genetics.109.103028>
28. Vinella D, Cashel M, D'Ari R (2000) **Selected amplification of the cell division genes *ftsQ*-*ftsA*-*ftsZ* in *Escherichia coli*** *Genetics* **156**:1483–92 <https://doi.org/10.1093/genetics/156.4.1483>
29. Nilsson AI, Zorzet A, Kanth A, Dahlstrom S, Berg OG, Andersson DI (2006) **Reducing the fitness cost of antibiotic resistance by amplification of initiator tRNA genes** *Proc Natl Acad Sci U S A*

A 103:6976–81 <https://doi.org/10.1073/pnas.0602171103>

30. Andersson DI, Nicoloff H, Hjort K (2019) **Mechanisms and clinical relevance of bacterial heteroresistance** *Nature Reviews Microbiology* **17**:479–96 <https://doi.org/10.1038/s41579-019-0218-1>
31. Nicoloff H, Hjort K, Levin BR, Andersson DI (2019) **The high prevalence of antibiotic heteroresistance in pathogenic bacteria is mainly caused by gene amplification** *Nat Microbiol* **4**:504–14 <https://doi.org/10.1038/s41564-018-0342-0>
32. Sun S, Ke R, Hughes D, Nilsson M, Andersson DI (2012) **Genome-wide detection of spontaneous chromosomal rearrangements in bacteria** *Plos One* **7**:e42639 <https://doi.org/10.1371/journal.pone.0042639>
33. Tomanek I, Grah R, Lagator M, Andersson AMC, Bollback JP, Tkacik G, et al. (2020) **Gene amplification as a form of population-level gene expression regulation** *Nat Ecol Evol* **4**:612–25 <https://doi.org/10.1038/s41559-020-1132-7>
34. Ahmad SI, Kirk SH, Eisenstark A (1998) **Thymine metabolism and thymineless death in prokaryotes and eukaryotes** *Annu Rev Microbiol* **52**:591–625 <https://doi.org/10.1146/annurev.micro.52.1.591>
35. Gangjee A, Jain HD, Kurup S (2007) **Recent advances in classical and non-classical Antifolates as antitumor and antiopportunistic infection agents: part I** *Anti-Cancer Agent Me* **7**:524–42 <https://doi.org/10.2174/187152007781668724>
36. Gangjee A, Jain HD, Kurup S (2008) **Recent advances in classical and non-classical antifolates as antitumor and antiopportunistic infection agents: Part II** *Anti-Cancer Agent Me* **8**:205–31 <https://doi.org/10.2174/187152008783497064>
37. Huennekens FM (1994) **The methotrexate story: a paradigm for development of cancer chemotherapeutic agents** *Adv Enzyme Regul* **34**:397–419 [https://doi.org/10.1016/0065-2571\(94\)90025-6](https://doi.org/10.1016/0065-2571(94)90025-6)
38. Plowe CV (2022) **Malaria chemoprevention and drug resistance: a review of the literature and policy implications** *Malar J* **21**:104 <https://doi.org/10.1186/s12936-022-04115-8>
39. Gleckman R, Blagg N, Joubert DW (1981) **Trimethoprim: mechanisms of action, antimicrobial activity, bacterial resistance, pharmacokinetics, adverse reactions, and therapeutic indications** *Pharmacotherapy* **1**:14–20 <https://doi.org/10.1002/j.1875-9114.1981.tb03548.x>
40. Brochet M, Couve E, Zouine M, Poyart C, Glaser P (2008) **A naturally occurring gene amplification leading to sulfonamide and trimethoprim resistance in *Streptococcus agalactiae*** *J Bacteriol* **190**:672–80 <https://doi.org/10.1128/JB.01357-07>
41. Goker E, Waltham M, Kheradpour A, Trippett T, Mazumdar M, Elisseyeff Y, et al. (1995) **Amplification of the Dihydrofolate-Reductase Gene Is a Mechanism of Acquired-Resistance to Methotrexate in Patients with Acute Lymphoblastic-Leukemia and Is Correlated with P53 Gene-Mutations** *Blood* **86**:677–84 <https://doi.org/10.1182/blood.V86.2.677.bloodjournal862677>
42. Inselburg J, Bzik DJ, Horii T (1987) **Pyrimethamine resistant *Plasmodium falciparum*: overproduction of dihydrofolate reductase by a gene duplication** *Mol Biochem Parasitol* **26**:121–34 [https://doi.org/10.1016/0166-6851\(87\)90136-8](https://doi.org/10.1016/0166-6851(87)90136-8)

43. Morales C, Garcia MJ, Ribas M, Miro R, Munoz M, Caldas C, et al. 2009) **Dihydrofolate reductase amplification and sensitization to methotrexate of methotrexate-resistant colon cancer cells** *Mol Cancer Ther* **8**:424–32 <https://doi.org/10.1158/1535-7163.MCT-08-0759>
44. Baym M, Lieberman TD, Kelsic ED, Chait R, Gross R, Yelin I, et al. 2016) **Spatiotemporal microbial evolution on antibiotic landscapes** *Science* **353**:1147–51 <https://doi.org/10.1126/science.aag0822>
45. Palmer AC, Toprak E, Baym M, Kim S, Veres A, Bershtein S, et al. 2015) **Delayed commitment to evolutionary fate in antibiotic resistance fitness landscapes** *Nat Commun* **6**:7385 <https://doi.org/10.1038/ncomms8385>
46. Patel V, Matange N 2021) **Adaptation and compensation in a bacterial gene regulatory network evolving under antibiotic selection** *eLife* **10** <https://doi.org/10.7554/eLife.70931>
47. Watson M, Liu JW, Ollis D 2007) **Directed evolution of trimethoprim resistance in Escherichia coli** *Febs J* **274**:2661–71 <https://doi.org/10.1111/j.1742-4658.2007.05801.x>
48. Toprak E, Veres A, Michel JB, Chait R, Hartl DL, Kishony R 2011) **Evolutionary paths to antibiotic resistance under dynamically sustained drug selection** *Nat Genet* **44**:101–5 <https://doi.org/10.1038/ng.1034>
49. Matange N 2020) **Highly Contingent Phenotypes of Lon Protease Deficiency in Escherichia coli upon Antibiotic Challenge** *J Bacteriol* **202** <https://doi.org/10.1128/JB.00561-19>
50. Rodrigues JV, Bershtein S, Li A, Lozovsky ER, Hartl DL, Shakhnovich EI 2016) **Biophysical principles predict fitness landscapes of drug resistance** *Proc Natl Acad Sci U S A* **113**:E1470–8 <https://doi.org/10.1073/pnas.1601441113>
51. Vinchhi R, Yelpure C, Balachandran M, Matange N 2023) **Pervasive gene deregulation underlies adaptation and maladaptation in trimethoprim-resistant E. coli** *mBio* **14**:e0211923 <https://doi.org/10.1128/mbio.02119-23>
52. Ranquet C, Gottesman S 2007) **Translational regulation of the Escherichia coli stress factor RpoS: a role for SsrA and Lon** *J Bacteriol* **189**:4872–9 <https://doi.org/10.1128/JB.01838-06>
53. Xu J, Li T, Gao Y, Deng J, Gu J 2019) **MgrB affects the acid stress response of Escherichia coli by modulating the expression of iraM** *FEMS Microbiol Lett* **366** <https://doi.org/10.1093/femsle/fnz123>
54. Matange N, Bodkhe S, Patel M, Shah P 2018) **Trade-offs with stability modulate innate and mutationally acquired drug resistance in bacterial dihydrofolate reductase enzymes** *Biochem J* **475**:2107–25 <https://doi.org/10.1042/BCJ20180249>
55. Ohno S. 1970) **Evolution by gene duplication** London, New York: Allen & Unwin; Springer-Verlag
56. Bergthorsson U, Andersson DI, Roth JR. 2007) **Ohno's dilemma: Evolution of new genes under continuous selection** *P Natl Acad Sci USA* **104**:17004–9 <https://doi.org/10.1073/pnas.0707158104>
57. Long MY, VanKuren NW, Chen SD, Vibranovski MD. 2013) **New Gene Evolution: Little Did We Know** *Annual Review of Genetics* **47**:307–33 <https://doi.org/10.1146/annurev-genet-111212-133301>

58. Holloway AK, Palzkill T, Bull JJ 2007) **Experimental evolution of gene duplicates in a bacterial plasmid model** *J Mol Evol* **64**:215–22 <https://doi.org/10.1007/s00239-006-0087-x>
59. Kugelberg E, Kofoed E, Reams AB, Andersson DI, Roth JR 2006) **Multiple pathways of selected gene amplification during adaptive mutation** *P Natl Acad Sci USA* **103**:17319–24 <https://doi.org/10.1073/pnas.0608309103>
60. Naseeb S, Ames RM, Delneri D, Lovell SC 2017) **Rapid functional and evolutionary changes follow gene duplication in yeast** *Proc Biol Sci* **284**:20171393 <https://doi.org/10.1098/rspb.2017.1393>
61. Tomanek I, Guet CC 2022) **Adaptation dynamics between copy-number and point mutations** *eLife* **11** <https://doi.org/10.7554/eLife.82240>
62. Nicoloff H, Perreten V, Levy SB 2007) **Increased genome instability in Escherichia coli lon mutants: relation to emergence of multiple-antibiotic-resistant (Mar) mutants caused by insertion sequence elements and large tandem genomic amplifications** *Antimicrob Agents Chemother* **51**:1293–303 <https://doi.org/10.1128/AAC.01128-06>
63. Nicoloff H, Perreten V, McMurry LM, Levy SB 2006) **Role for tandem duplication and lon protease in AcrAB-TolC-dependent multiple antibiotic resistance (Mar) in an Escherichia coli mutant without mutations in marRAB or acrRAB** *Journal of Bacteriology* **188**:4413–23 <https://doi.org/10.1128/Jb.01502-05>
64. Derbyshire KM, Kramer M, Grindley ND 1990) **Role of instability in the cis action of the insertion sequence IS903 transposase** *Proc Natl Acad Sci U S A* **87**:4048–52 <https://doi.org/10.1073/pnas.87.11.4048>
65. Rouquette C, Serre MC, Lane D 2004) **Protective role for H-NS protein in IS1 transposition** *J Bacteriol* **186**:2091–8 <https://doi.org/10.1128/JB.186.7.2091-2098.2004>
66. Hua CF, Wang TT, Shao XL, Xie YP, Huang H, Liu JG, et al. 2020) **Pseudomonas syringae dual-function protein Lon switches between virulence and metabolism by acting as both DNA-binding transcriptional regulator and protease in different environments** *Environ Microbiol* **22**:2968–88 <https://doi.org/10.1111/1462-2920.15067>
67. Karłowicz A, Węgrzyn K, Gross M, Kaczynska D, Ropelewska M, Siemiakowska M, et al. 2017) **Defining the crucial domain and amino acid residues in bacterial Lon protease for DNA binding and processing of DNA-interacting substrates** *J Biol Chem* **292**:7507–18 <https://doi.org/10.1074/jbc.M116.766709>
68. Liu T, Lu B, Lee I, Ondrovicová G, Kutejová E, Suzuki CK 2004) **DNA and RNA binding by the mitochondrial Lon protease is regulated by nucleotide and protein substrate** *J Biol Chem* **279**:13902–10 <https://doi.org/10.1074/jbc.M309642200>
69. Lu B, Yadav S, Shah PG, Liu T, Tian B, Puksza S, et al. 2007) **Roles for the human ATP-dependent Lon protease in mitochondrial DNA maintenance** *J Biol Chem* **282**:17363–74 <https://doi.org/10.1074/jbc.M611540200>
70. Sonezaki S, Okita K, Oba T, Ishii Y, Kondo A, Kato Y 1995) **Protein substrates and heat shock reduce the DNA binding ability of Escherichia coli Lon protease** *Appl Microbiol Biot* **44**:484–8
71. Breidenstein EB, Bains M, Hancock RE 2012) **Involvement of the lon protease in the SOS response triggered by ciprofloxacin in Pseudomonas aeruginosa PAO1** *Antimicrob Agents*

Chemother **56**:2879–87 <https://doi.org/10.1128/AAC.06014-11>

72. Guerrero RF, Scarpino SV, Rodrigues JV, Hartl DL, Ogbunugafor CB (2019) **Proteostasis Environment Shapes Higher-Order Epistasis Operating on Antibiotic Resistance** *Genetics* **212**:565–75 <https://doi.org/10.1534/genetics.119.302138>
73. Thompson S, Zhang Y, Ingle C, Reynolds KA, Kortemme T (2020) **Altered expression of a quality control protease in E. coli reshapes the in vivo mutational landscape of a model enzyme** *eLife* **9** <https://doi.org/10.7554/eLife.53476>
74. Yoon J, Patrick JE, Ogbunugafor CB, Shoulders MD (2023) **Viral Evolution Shaped by Host Proteostasis Networks** *Annu Rev Virol* **10**:77–98 <https://doi.org/10.1146/annurev-virology-100220-112120>
75. Bershtein S, Serohijos AW, Bhattacharyya S, Manhart M, Choi JM, Mu W, et al. (2015) **Protein Homeostasis Imposes a Barrier on Functional Integration of Horizontally Transferred Genes in Bacteria** *Plos Genet* **11**:e1005612 <https://doi.org/10.1371/journal.pgen.1005612>
76. Brown NG, Pennington JM, Huang W, Ayvaz T, Palzkill T (2010) **Multiple global suppressors of protein stability defects facilitate the evolution of extended-spectrum TEM beta-lactamases** *J Mol Biol* **404**:832–46 <https://doi.org/10.1016/j.jmb.2010.10.008>
77. Standley M, Blay V, Beleva Guthrie V, Kim J, Lyman A, Moya A, et al. (2022) **Experimental and In Silico Analysis of TEM beta-Lactamase Adaptive Evolution** *ACS Infect Dis* **8**:2451–63 <https://doi.org/10.1021/acsinfectdis.2c00216>
78. Koutsandreas T, Felden B, Chevet E, Chatziioannou A (2024) **Protein homeostasis imprinting across evolution** *NAR Genom Bioinform* **6**:lqae014 <https://doi.org/10.1093/nargab/lqae014>
79. Khodaparast L, Khodaparast L, Wu GQ, Michiels E, Gallardo R, Houben B, et al. (2023) **Exploiting the aggregation propensity of beta-lactamases to design inhibitors that induce enzyme misfolding** *Nature Communications* **14** <https://doi.org/10.1038/s41467-023-41191-z>
80. Lupoli TJ, Vaubourgeix J, Burns-Huang K, Gold B (2018) **Targeting the Proteostasis Network for Mycobacterial Drug Discovery** *ACS Infect Dis* **4**:478–98 <https://doi.org/10.1021/acsinfectdis.7b00231>
81. Won HI, Zinga S, Kandror O, Akopian T, Wolf ID, Schweber JTP, et al. (2024) **Targeted protein degradation in mycobacteria uncovers antibacterial effects and potentiates antibiotic efficacy** *Nat Commun* **15**:4065 <https://doi.org/10.1038/s41467-024-48506-8>
82. Matange N, Hegde S, Bodkhe S (2019) **Adaptation Through Lifestyle Switching Sculpt the Fitness Landscape of Evolving Populations: Implications for the Selection of Drug-Resistant Bacteria at Low Drug Pressures** *Genetics* **211**:1029–44 <https://doi.org/10.1534/genetics.119.301834>
83. Deatherage DE, Barrick JE (2014) **Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq** *Methods Mol Biol* **1165**:85–98 https://doi.org/10.1007/978-1-4939-0554-6_12

Author information

Chinmaya Jena

Department of Biology, Indian Institute of Science Education and Research, Pune, India

ORCID iD: [0000-0002-4401-3570](https://orcid.org/0000-0002-4401-3570)

Saillesh Chinnaraj

Department of Biology, Indian Institute of Science Education and Research, Pune, India

Soham Deolankar

Department of Biology, Indian Institute of Science Education and Research, Pune, India

Nishad Matange

Department of Biology, Indian Institute of Science Education and Research, Pune, India

ORCID iD: [0000-0002-2947-3931](https://orcid.org/0000-0002-2947-3931)

For correspondence: nishad@iiserpune.ac.in

Editors

Reviewing Editor

Daria Van Tyne

University of Pittsburgh, Pittsburgh, United States of America

Senior Editor

Alan Moses

University of Toronto, Toronto, Canada

Reviewer #1 (Public review):

Summary:

The study by Jena et al. addresses important questions on the fundamental mechanisms of genetic adaptation, specifically, does adaptation proceed via changes of copy number (gene duplication and amplification "GDA") or by point mutation. While this question has been worked on (for example by Tomanek and Guet) the authors add several important aspects relating to resistance against antibiotics and they clarify the ability of Lon protease to reduce duplication formation (previous work was more indirect).

A key finding Jena et al. present is that point mutations after significant competition displace GDA. A second one is that alternative GDA constantly arise and displace each other (see work on GDA-2 in Figure 3). Finally, the authors found epistasis between resistance allele that was contingent on lon. Together this shows an intricate interplay of lon proteolysis for the evolution and maintenance of antibiotic resistance by gene duplication.

Strengths:

The study has several important strengths: (i) the work on GDA stability and competition of GDA with point mutations is a very promising area of research and the authors contribute new aspects to it, (ii) rigorous experimentation, (iii) very clearly written introduction and discussion sections. To me, the best part of the data is that deletion of lon stimulates GDA, which has not been shown with such clarity until now.

Weaknesses:

Previously raised minor weaknesses and technical questions have been adequately resolved in the revised manuscript. As the experiments and their results are described in great detail the interested reader needs stamina. The details will, however, be informative to the specialist.

<https://doi.org/10.7554/eLife.99785.2.sa2>

Reviewer #3 (Public review):

Summary:

This is an important paper that investigates the relationship between proteolytic stability of an antibiotic target enzyme and the evolution of antibiotic resistance via increased gene copy number. The target of the antibiotic trimethoprim is dihydrofolate reductase (DHFR). In *Escherichia coli*, DHFR is encoded by *folA* and the major proteolysis housekeeping protease is Lon (lon). In this manuscript, the authors report the result of the experimental evolution of a lon mutant strain of *E. coli* in response to sub-inhibitory concentrations of the antibiotic trimethoprim then investigate the relationship between proteolytic stability of DHFR mutants and the evolution of *folA* gene duplication. After 25 generations of serial passaging in a fixed concentration of trimethoprim, the authors found that *folA* duplication events were more common during evolution of the lon strain, than the wt strain. However, with continued passaging, some *folA* duplications were replaced by a single copy of *folA* containing a trimethoprim resistance-conferring point mutation. Interestingly, evolution of the lon strain in the setting of increasing concentrations of trimethoprim resulted in evolved strains with different levels of DHFR expression. In particular, some strains maintained two copies of a mutant *folA* that encoded an unstable DHFR. In a lon⁺ background, this mutant *folA* did not express well and did not confer trimethoprim resistance. However, in the lon⁻ background, it displayed higher expression and conferred high-level trimethoprim resistance. The authors concluded that maintenance of the gene duplication event (and the absence of Lon) compensated for the proteolytic instability of this mutant DHFR. In summary, they provide evidence that the proteolytic stability of an antibiotic target protein is an important determinant of the evolution of target gene copy number in the setting of antibiotic selection.

Strengths:

The major strength of this paper is identifying an example of antibiotic resistance evolution that illustrates the interplay between the proteolytic stability and copy number of an antibiotic target in the setting of antibiotic selection. The results are rigorous and convincingly support the conclusions. This paper will be of interest to any biologist that studies the evolution of resistance mechanisms or gene duplication.

Weaknesses:

The impact of this finding is somewhat limited given that it is a single example that occurred in a lon strain of *E. coli*. Although the specific mechanism is unlikely to occur naturally, this study represents an important and convincing proof of the principle that gene duplication can provide increased expression demand for an unstable resistance determinant in the setting of antibiotic selection.

<https://doi.org/10.7554/eLife.99785.2.sa1>

Author response:

The following is the authors' response to the original reviews.

Public Reviews:

Reviewer #1 (Public review):

Summary:

The study by Jena et al. addresses important questions on the fundamental mechanisms of genetic adaptation, specifically, does adaptation proceed via changes of copy number (gene duplication and amplification "GDA") or by point mutation. While this question has been worked on (for example by Tomanek and Guet) the authors add several important aspects relating to resistance against antibiotics and they clarify the ability of Lon protease to reduce duplication formation (previous work was more indirect).

A key finding Jena et al. present is that point mutations after significant competition displace GDA. A second one is that alternative GDA constantly arise and displace each other (see work on GDA-2 in Figure 3). Finally, the authors found epistasis between resistance alleles that was contingent on Lon. Together this shows an intricate interplay of Lon proteolysis for the evolution and maintenance of antibiotic resistance by gene duplication.

Strengths:

The study has several important strengths: (i) the work on GDA stability and competition of GDA with point mutations is a very promising area of research and the authors contribute new aspects to it, (ii) rigorous experimentation, (iii) very clearly written introduction and discussion sections. To me, the best part of the data is that deletion of Lon stimulates GDA, which has not been shown with such clarity until now.

Weaknesses:

The minor weaknesses of the manuscript are a lack of clarity in parts of the results section (Point 1) and the methods (Point 2).

We thank the reviewer for their comments and suggestions on our manuscript. We also appreciate the succinct summary of primary findings that the Reviewer has taken cognisance of in their assessment, in particular the association of the Lon protease with the propensity for GDAs as well as its impact on their eventual fate. We have now revised the manuscript for greater clarity as suggested by Reviewer #1.

Reviewer #2 (Public review):

Summary:

In this strong study, the authors provide robust evidence for the role of proteostasis genes in the evolution of antimicrobial resistance, and moreover, for stabilizing the proteome in light of gene duplication events.

Strengths:

This strong study offers an important interaction between findings involving GDA, proteostasis, experimental evolution, protein evolution, and antimicrobial resistance. Overall, I found the study to be relatively well-grounded in each of these literatures, with experiments that spoke to potential concerns from each arena. For example, the

literature on proteostasis and evolution is a growing one that includes organisms (even micro-organisms) of various sorts. One of my initial concerns involved whether the authors properly tested the mechanistic bases for the role of Lon in promoting duplication events. The authors assuaged my concern with a set of assays (Figure 8).

More broadly, the study does a nice job of demonstrating the agility of molecular evolution, with responsible explanations for the findings: gene duplications are a quick-fix, but can be out-competed relative to their mutational counterparts. Without Lon protease to keep the proteome stable, the cell allows for less stable solutions to the problem of antibiotic resistance.

The study does what any bold and ambitious study should: it contains large claims and uses multiple sorts of evidence to test those claims.

Weaknesses:

While the general argument and conclusion are clear, this paper is written for a bacterial genetics audience that is familiar with the manner of bacterial experimental evolution. From the language to the visuals, the paper is written in a boutique fashion. The figures are even difficult for me - someone very familiar with proteostasis - to understand. I don't know if this is the fault of the authors or the modern culture of publishing (where figures are increasingly packed with information and hard to decipher), but I found the figures hard to follow with the captions. But let me also consider that the problem might be mine, and so I do not want to unfairly criticize the authors.

For a generalist journal, more could be done to make this study clear, and in particular, to connect to the greater community of proteostasis researchers. I think this study needs a schematic diagram that outlines exactly what was accomplished here, at the beginning. Diagrams like this are especially important for studies like this one that offer a clear and direct set of findings, but conduct many different sorts of tests to get there. I recommend developing a visual abstract that would orient the readers to the work that has been done.

The reviewer's comments regarding data presentation are well-taken. Since we already had a diagrammatic model that sums up the chief findings of our study (Figure 9), we have now provided schematics in Figures 1, 3, 5 and 8 to clarify the workflow of smaller sections of the study. We hope that these diagrams provide greater clarity with regards to the experiments we have conducted.

Next, I will make some more specific suggestions. In general, this study is well done and rigorous, but doesn't adequately address a growing literature that examines how proteostasis machinery influences molecular evolution in bacteria.

While this paper might properly test the authors' claims about protein quality control and evolution, the paper does not engage a growing literature in this arena and is generally not very strong on the use of evolutionary theory. I recognize that this is not the aim of the paper, however, and I do not question the authors' authority on the topic. My thoughts here are less about the invocation of theory in evolution (which can be verbose and not relevant), and more about engagement with a growing literature in this very area.

The authors mention Rodrigues 2016, but there are many other studies that should be engaged when discussing the interaction between protein quality control and evolution.

A 2015 study demonstrated how proteostasis machinery can act as a barrier to the usage of novel genes: Bershtein, S., Serohijos, A. W., Bhattacharyya, S., Manhart, M., Choi, J. M., Mu, W., ... & Shakhnovich, E. I. (2015). Protein homeostasis imposes a barrier to

functional integration of horizontally transferred genes in bacteria. PLoS genetics, 11(10), e1005612

A 2019 study examined how Lon deletion influenced resistance mutations in DHFR specifically: Guerrero RF, Scarpino SV, Rodrigues JV, Hartl DL, Ogbunugafor CB. The proteostasis environment shapes higher-order epistasis operating on antibiotic resistance. Genetics. 2019 Jun 1;212(2):565-75.

A 2020 study did something similar: Thompson, Samuel, et al. "Altered expression of a quality control protease in E. coli reshapes the in vivo mutational landscape of a model enzyme." Elife 9 (2020): e53476.

And there's a new review (preprint) on this very topic that speaks directly to the various ways proteostasis shapes molecular evolution:

Arenas, Carolina Diaz, Maristella Alvarez, Robert H. Wilson, Eugene I. Shakhnovich, C. Brandon Ogbunugafor, and C. Brandon Ogbunugafor. "Proteostasis is a master modulator of molecular evolution in bacteria."

I am not simply attempting to list studies that should be cited, but rather, this study needs to be better situated in the contemporary discussion on how protein quality control is shaping evolution. This study adds to this list and is a unique and important contribution. However, the findings can be better summarized within the context of the current state of the field. This should be relatively easy to implement.

We thank the reviewer for their encouraging assessment of our manuscript as well as this important critique regarding the context of other published work that relates proteostasis and molecular evolution. Indeed, this was a particularly difficult aspect for us given the different kinds of literature that were needed to make sense of our study. We have now added the references suggested by the reviewer as well as others to the manuscript. We have also added a paragraph in the discussion section (Lines 463-476) that address this aspect and hopefully fill the lacuna that the reviewer points out in this comment.

Reviewer #3 (Public review):

Summary:

This paper investigates the relationship between the proteolytic stability of an antibiotic target enzyme and the evolution of antibiotic resistance via increased gene copy number. The target of the antibiotic trimethoprim is dihydrofolate reductase (DHFR). In Escherichia coli, DHFR is encoded by folA and the major proteolysis housekeeping protease is Lon (lon). In this manuscript, the authors report the results of the experimental evolution of a lon mutant strain of E. coli in response to sub-inhibitory concentrations of the antibiotic trimethoprim and then investigate the relationship between proteolytic stability of DHFR mutants and the evolution of folA gene duplication. After 25 generations of serial passaging in a fixed concentration of trimethoprim, the authors found that folA duplication events were more common during the evolution of the lon strain, than the wt strain. However, with continued passaging, some folA duplications were replaced by a single copy of folA containing a trimethoprim resistance-conferring point mutation. Interestingly, the evolution of the lon strain in the setting of increasing concentrations of trimethoprim resulted in evolved strains with different levels of DHFR expression. In particular, some strains maintained two copies of a mutant folA that encoded an unstable DHFR. In a lon+ background, this mutant folA did not express well and did not confer trimethoprim resistance. However, in the lon- background, it displayed higher expression and conferred high-level trimethoprim resistance. The authors concluded that maintenance of the gene duplication event (and the absence of

Lon) compensated for the proteolytic instability of this mutant DHFR. In summary, they provide evidence that the proteolytic stability of an antibiotic target protein is an important determinant of the evolution of target gene copy number in the setting of antibiotic selection.

Strengths:

The major strength of this paper is identifying an example of antibiotic resistance evolution that illustrates the interplay between the proteolytic stability and copy number of an antibiotic target in the setting of antibiotic selection. If the weaknesses are addressed, then this paper will be of interest to microbiologists who study the evolution of antibiotic resistance.

Weaknesses:

*Although the proposed mechanism is highly plausible and consistent with the data presented, the analysis of the experiments supporting the claim is incomplete and requires more rigor and reproducibility. The impact of this finding is somewhat limited given that it is a single example that occurred in a *lon* strain and compensatory mutations for evolved antibiotic resistance mechanisms are described. In this case, it is not clear that there is a functional difference between the evolution of copy number versus any other mechanism that meets a requirement for increased "expression demand" (e.g. promoter mutations that increase expression and protein stabilizing mutations).*

We thank the reviewer for their in-depth assessment of our work and appreciate their concerns regarding reproducibility and rigor in analysis of our data. We have now incorporated this feedback and provided necessary clarifications/corrections in the revised version of our manuscript.

Recommendations for the authors:

Reviewer #1 (Recommendations for the authors):

Major Points:

*(1) The authors show that a deletion of *lon* increases the ability for GDA and they argue that this is adaptive during TMP treatment because it increases the dosage of *folA* (L. 129). However, the highest frequency of GDA occurred in drug-free conditions (see Figure 1C). This indicates either that GDA is selected in drug-free media and potentially selected against by certain antibiotics. It would help for the authors to discuss this possibility more clearly.*

We thank the reviewer for this astute observation. It is indeed striking that the GDA mutation (i.e. the GDA-2 mutation) selected in a *lon*-deficient background does not come up in presence of antibiotics. To probe this further, we have now measured the relative fitness of a representative population of *lon*-knockout from short-term evolution in drug-free LB (population #3) that harbours GDA-2 against its ancestor (marked with *DlacZ*). These competition experiments were performed in LB (in which GDA-2 emerged spontaneously), as well as in LB supplemented with antibiotics at the concentrations used during the short term evolution.

Values of relative fitness, *w* (mean $\pm$ SD from 3 measurements), are provided below:

LB: 1.4 ± 0.2

LB + Trimethoprim: 1.6 ± 0.2

LB + Spectinomycin: 0.9 ± 0.2

LB + Erythromycin: 1.3 ± 0.3

LB + Nalidixic acid: 1.5 ± 0.2

LB + Rifampicin: 1.4 ± 0.2

These data show an increase in relative fitness in drug-free LB as would be expected. Interestingly, we also observe an increase in relative fitness in LB supplemented with antibiotics, except spectinomycin. This result supports the idea that GDA-2 is a “media adaptation” and provides a general fitness advantage to the lon knockout. However, as the reviewer pointed out, we should expect to see GDA-2 emerge spontaneously in antibiotic-supplemented media as well. We think that this does not happen as the fitness advantage of drug-specific mutations (GDAs or point mutations) far exceed the advantage of a media adaptation GDA. As a result, we only see the specific mutations that provide high benefit against the antibiotic at least over the relatively short duration of 20-25 generations. It is noteworthy the GDA-2 mutation does come up in LTMPR1 when it is passaged over >200 generations in drug-free media, but shows fluctuating frequency over time. We expect, therefore, that given enough time we may detect the GDA-2 mutations even in antibiotic-supplemented media.

We note, however, that a major caveat in the above fitness calculations is that we cannot be sure that the competing ancestor has no GDA-2 mutations during the course of the experiment. Thus, the above fitness values are only indicative and not definitive. We have therefore not included these data in the revised manuscript.

(2) It is unclear if the isolates WTMPR1 - 5 and LTMPR1 - 5 were pure clones. The authors write in L.488 "Colonies were randomly picked, cultured overnight in drug-free LB and frozen in 50% glycerol at -80C until further use." And in L. 492 "For long-term evolution, trimethoprim-resistant isolates LTMPR1, WTMPR4 and WTMPR5 were first revived from frozen stocks in drug-free LB overnight." From these descriptions, it is possible that the isolates contained a fraction of cells of other genotypes since colonies are often formed by more than one cell and thus, unless pure-streaked, a subpopulation is present and would in drug-free media be maintained. The possibility of pre-existing subpopulations is important for all statements relating to "reversal".

This is indeed a valid concern. As far as we can tell all our initial isolates (i.e. WTMPR1-5 and LTMPR1-5) are pure clones at least as far as SNPs are concerned. This is based on whole genome sequencing data that we have reported earlier in Patel and Matange, eLife (2021), where we described the evolution and isolation of WTMPR1-5 and the present study for LTMPR1-5. All SNPs detected were present at a frequency of 100%. For clones with GDAs, however, there is no way to eliminate a sub-population that has a lower or higher gene copy number than average from an isolate. This is because of the inherent instability of GDAs that will inevitably result in heterogeneous gene copy number during standard growth. In this sense, there is most certainly a possibility of a pre-existing subpopulation within each of the clones that may have reversed the GDA. Indeed, we believe that it is this inherent instability that contributes to their rapid loss during growth in drug-free media.

Minor Points:

(1) L. 406. "allowing accumulation of IS transposases in E. coli" Please specify that it is the accumulation of transposase proteins (and not genes).

We have made this change.

(2) L. 221 typo. Known "to" stabilize.

We have made this change.

Reviewer #2 (Recommendations for the authors):

Most of my suggestions are found in the public review. I believe this to be a strong study, and some slight fixes can solidify its presence in the literature.

We have attempted to address the two main critiques by Reviewer 2. To simplify the understanding of our data, we have provided small schematics at various points in the paper to clarify the experimental pipelines used by us. We have also provided additional discussion situating our study in the emerging area of proteostasis and molecular evolution. We hope that our revisions have addressed these lacunae in our manuscript.

Reviewer #3 (Recommendations for the authors):

Major Points:

(1) The manuscript is generally a bit difficult to follow. The writing is overly complicated and lacks clarity at times. It should be simplified and improved.

We have made several revisions to the text, as well as provided schematics in some of our figures which hopefully make our paper easier to understand.

(2) I cannot find the raw variant summary data for the lon strain evolution experiment in trimethoprim (after 25 generations). Were there any other mutations identified? If not, this should be explicitly stated in the text and the variant output summary from sequencing included as supplemental data.

We apologise for this oversight. We have now provided these data as Table 1.

(3) What is the trimethoprim IC50 of the starting (pre-evolution) strains (i.e. wt and lon)? I can't find this information, but it is critical to interpretation.

We had reported these values earlier in Matange N., J Bact (2020). Wild type and lon-knockout have similar MIC values for trimethoprim, though the lon mutant shows a higher IC50 value. We have now mentioned this in the results section (Line 100-101) and also provided the reference for these data.

(4) What was the average depth of coverage for WGS? This information is necessary to assess the quality of the variant calling, especially for the population WGS.

All genome sequencing data has a coverage at least 100x. We have added this detail to the methods section (Line 580-581).

(5) Five replicate evolution experiments (25 generations, or 7x 10% daily batch transfers) were performed in trimethoprim for the wt and lon strains. Duplication of the folA locus occurred in 1/5 and 4/5 experiments, respectively. It is not entirely clear what type of sampling was actually done to arrive at these numbers (this needs to be stated more clearly), but presumably 1 random colony was chosen at the end of the passaging protocol for each replicate. Based on this result, the authors conclude that folA duplication occurred more frequently in the lon strain, however, this is not rigorously supported by a statistical evaluation. With N=5, one cannot rigorously conclude that a

20% frequency and 80% frequency are significantly different. Furthermore, it's not entirely clear what the mechanism of resistance is for these strains. For example, in one colony sequenced (LTMPR5), it appears no known resistance mechanism (or mutations?) were identified, and yet the IC50 = 900 nM, which is also similar to other strains.

Indeed, we agree with the reviewer that we don't have the statistical power to rigorously make this claim. However, since the lon-knockout showed us a greater frequency of GDA across 3 different environments we are fairly confident that loss of lon enhances the overall frequency for GDA mutations. This idea is also supported by a number of previous papers that related GDAs and IS-element transpositions with Lon, viz. Nicoloff et al, Antimicrob Agent Chemother (2007), Derbyshire et al. PNAS (1990), Derbyshire and Grindley, Mol Microbiol (1996). We have therefore not provided further justification in the revised manuscript.

We had indeed sampled a random isolate from each of the 5 populations and have added a schematic to figure 1 that provides greater clarity.

Having relooked at the sequencing data for LTMPR1-5 isolates (Table 1), we realised that both LTMPR4 and LTMPR5 harbour mutations in the pitA gene. We had missed this locus during the previous iteration of this manuscript and misidentified an mgrB mutations in LTMPR4. PitA codes for a metal-phosphate symporter. We have observed mutations in pitA in earlier evolution experiments with trimethoprim as well (Vinchhi and Yelpure et al. mBio 2023). Interestingly, in LTMPR5 there was a deletion of pitA, along with 17 other contiguous genes mediated by IS5. To test if loss of pitA is beneficial in trimethoprim, we tested the ability of a pitA knockout to grow on trimethoprim supplemented plates. Indeed, loss of pitA conferred a growth advantage to E. coli on trimethoprim, comparable to loss of mgrB, indicating that the mechanism of resistance of LTMPR5 may be due to loss of pitA. We have added these data to the Supplementary Figure 1 of the revised manuscript and provided a brief description in Lines 103-108. How pitA deficiency confers trimethoprim resistance is yet to be investigated. The mechanism is likely to be by activating some intrinsic resistance mechanism as loss of pitA also conferred a fitness benefit against other antibiotics. This work is currently underway in our lab and hence we do not provide any further mechanism in the present manuscript.

(6) Although measurement error/variance is reported, statistical tests were not performed for any of the experiments. This is critical to support the rigor and reproducibility of the conclusions.

We have added statistical testing wherever appropriate to the revised manuscript.

(7) Lines 150-155 and Figure 2E: Putting a wt copy of mgrB back into the WTMPR4 and LTMPR1 strains would be a better experiment to dissect out the role of mgrB versus the other gene duplications in these strains on fitness. Without this experiment, you cannot confidently attribute the fitness costs of these strains to the inactivation of mgrB alone.

We agree with the reviewer that our claim was based on a correlation alone. We have now added some new data to confirm our model (Figure 2 E, F). The costs of mgrB mutations come from hyperactivation of PhoQP. In earlier work we have shown that the costs (and benefit) of mgrB mutations can be abrogated in media supplemented with Mg^{2+} , which turns off the PhoQ receptor (Vinchhi and Yelpure et al. mBio, 2023). We use this strategy to show that like the mgrB-knockout, the costs of WTMPR4, WTMPR5 and LTMPR1 can be almost completely alleviated by adding Mg^{2+} to growth media. These results confirm that the source of fitness cost of TMP-resistant bacteria was not linked to GDA mutations, but to hyperactivation of PhoQP.

(8) Figure 3F and G: Does the top symbol refer to the starting strain for the 'long-term' evolution? If so, why does WTMPR4 not have the *mgrB* mutation (it does in Figure 1)? Based on your prior findings, it seems odd that this strain would evolve an *mgrB* loss of function mutation in the absence of trimethoprim exposure.

We thank the reviewer for pointing this error out. We have made the correction in the revised manuscript.

(9) Figure 6A: If the marker is neutral, it should be maintained at 0.1% throughout the 'neutrality' experiment. In both plots, the proportion of some marked strains goes up and then down. This suggests either ongoing evolution (these competitions take place over 105 generations), or noisy data. I suspect these data are just inherently noisy. I don't see error bars in the plots. Were these experiments ever replicated? It seems that replicating the experiments might be able to separate out noise from signal and perhaps clarify this point and better confirm the hypothesis that the point mutants are more fit.

These experiments were indeed noisy and the apparent enrichment is most likely a measurement error rather than a real change in frequency of competing genotypes. We have now provided individual traces for each of the competing pairs with mean and SD from triplicate observations at each time point.

(10) Figure 6A: Please indicate which plotted line refers to which 'point mutant' using different colors. These mutants have different trimethoprim IC₅₀s and doubling times, so it would be nice to be able to connect each mutant to its specific data plot.

We thank the reviewer for this suggestion. We have now colour coded the different strain combinations as suggested.

(11) Lines 284-285: I disagree that the IC₅₀s are similar. The C-35T mutant has IC₅₀ that is 2x that of LTMPR1. Perhaps more telling is that, compared to the *folA* duplication strain from the same time-point (which also carries the *rpoS* mutation), all of the point mutants have greater IC₅₀s (~2x greater). 2-fold changes in IC₅₀ are significant. It would seem that the point-mutants were likely not competing against LTMPR1 at the time they arose, so LTMPR1 might not be the best comparator if it was extinguished from the population early. I'm assuming this is why you chose a contemporary isolate (and, also, *rpoS* mutant) for the competition experiments. This should be explained more clearly.

We thank the reviewer for this comment. Indeed, the reviewer is correct about the rationale behind the use of a contemporary isolate and we have provided this clarification in the revised manuscript (Line 287-289). Also, the reviewer is correct in pointing out that a two-fold difference in IC₅₀ cannot be ignored. However, the key point here would be in assessing the differences in growth rates at the antibiotic concentration used during competition (i.e. 300 ng/mL). We are unable to see a direct correlation between the growth rates and enrichment in culture indicating that the observed trends are unlikely to be driven by 'level of resistance' alone. We have added these clarifications to the modified manuscript (Lines 299-301)

Minor Points:

(1) Line 13: Add a comma before 'Escherichia'

We have made this change.

(2) Line 14: Consider changing "mutations...were beneficial in trimethoprim" to "mutations...were beneficial under trimethoprim exposure"

We have made this change.

(3) Line 32: Is gene dosage really only "relative to the genome"? Is it not simply its relative copy number generally? Consider changing to "The dosage of a gene, or its relative copy number, can impact its level of expression..."

We have made this change.

(4) Line 38: The idea that GDAs are 1000x more frequent than point mutations seems an overgeneralization.

We agree with the reviewer and have softened our claim.

(5) Line 50: The term "hard-wired" is confusing. Please be more specific.

We have modified this statement to "...GDAs are less stable than point mutations....".

(6) Line 52-53: What do you mean by "there is also evidence to suggest that...more common in bacteria than appreciated"? Are you implying the field is naïve to this fact? If there is "evidence" of this, then a reference should be included. However, it's not clear why this is important to state in the article. I would consider simply removing this sentence. Less is more in this case.

We have removed this statement.

(7) Lines 59-60: Enzymes catalyze reactions. Please also state the substrates for DHFR. Consider, "It catalyzes the NADPH-dependent reduction of dihydrofolate to tetrahydrofolate, and important co-factor for..."

We have made this change.

(8) Line 72: Please change to, "In *E. coli*, DHFR is encoded by *folA*." You do not need to state this is a gene, as it is implicit with lowercase italics.

We have made this change.

(9) Lines 72-86: This paragraph is a bit confusing to read, as it has several different ideas in it. Consider breaking it into two paragraphs at Line 80, "In this study,...". The first paragraph could just review the trimethoprim resistance mechanisms in *E. coli* and so would change the first sentence (Line 72) to reflect this topic: "In *E. coli*, DHFR is encoded by *folA* and several different resistance mechanisms have been characterized." Then, just describe each mechanism in turn. Also, by "hot spots" it would seem you are referring to "point mutations" in the gene that alter the protein sequence and cluster onto the 3D protein structure when mapped? Please be more specific with this sentence for clarity.

We have made these changes.

(10) Lines 92-93: Please also state the MIC value of the strain to specifically define "sub-MIC". Alternatively, you could also state the fraction MIC (e.g. 0.1 x MIC).

We have modified this statement to “...in 300 ng/mL of trimethoprim (corresponding to ~0.3 x MIC) for 25 generations.”

(11) Lines 95-96. Remove, “These sequencing have been reported earlier, ...(2021)”. You just need to cite the reference.

We have made this change.

(12) Line 96: Remove the word “gene”.

We have made this change.

(13) Figure 1 and Figure 4C: The color scheme is tough for those with the most common type of color blindness. Red/green color deficiency causes a lot of difficulty with Red/gray, red/green, green/gray. Consider changing.

We thank the reviewer for bringing this to our notice. We have modified the colour scheme throughout the manuscript.

(14) Figure 1: Was there a trimethoprim resistance mechanism identified for LTMPR5?

As stated by us in response to major comment #7, LTMPR5’s resistance seems to come from a novel mechanism involving loss of the pitA gene.

(15) Line 349-351: Please briefly define “lower proteolytic stability” as a relative susceptibility to proteolytic degradation and make sure it is clear to the reader that this causes less DHFR. This needs to be clarified because it is confusing how a mutation that causes DHFR proteolytic instability would lead to an increase in trimethoprim IC50. So, you also need to mention that some mutations can cause both increased trimethoprim inhibition and lower proteolytic stability simultaneously. It seems the Trp30Arg mutation is an example of this, as this mutation is associated with a net increase in trimethoprim resistance despite the competing effects of the mutation on enzyme inhibition and DHFR levels.

We thank the reviewer for this comment and agree that the text in the original manuscript did not fully convey the message. We have made modifications to this section (Lines 359-363) in the revised manuscript in agreement with the reviewer’s suggestions.

<https://doi.org/10.7554/eLife.99785.2.sa0>