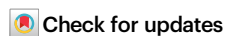


RNA polymerase stalling-derived genome instability underlies ribosomal antibiotic efficacy and resistance evolution

Received: 1 September 2023

Accepted: 24 July 2024

Published online: 03 August 2024

Yayun Zheng^{1,5}, Ruochen Chai^{1,5}, Tianmin Wang^{1,2,4,5}✉, Zeqi Xu¹,
Yihui He¹, Ping Shen¹ & Jintao Liu^{1,2,3}✉

Bacteria often evolve antibiotic resistance through mutagenesis. However, the processes causing the mutagenesis have not been fully resolved. Here, we find that a broad range of ribosome-targeting antibiotics cause mutations through an underexplored pathway. Focusing on the clinically important aminoglycoside gentamicin, we find that the translation inhibitor causes genome-wide premature stalling of RNA polymerase (RNAP) in a loci-dependent manner. Further analysis shows that the stalling is caused by the disruption of transcription-translation coupling. Anti-intuitively, the stalled RNAPs subsequently induce lesions to the DNA via transcription-coupled repair. While most of the bacteria are killed by genotoxicity, a small subpopulation acquires mutations via SOS-induced mutagenesis. Given that these processes are triggered shortly after antibiotic addition, resistance rapidly emerges in the population. Our work reveals a mechanism of action of ribosomal antibiotics, illustrates the importance of dissecting the complex interplay between multiple molecular processes in understanding antibiotic efficacy, and suggests new strategies for countering the development of resistance.

Antibiotics play a decisive role in combating bacterial infections. However, many antibiotics became ineffective due to the rapid emergence of drug resistance. Given the stagnation in developing new drugs, a deeper understanding of the effects of existing antibiotics on bacteria and their responses is urgently needed. The primary targets and modes of action of many antibiotics have been revealed. However, only knowing the interaction with their primary targets is often insufficient to fully explain the mechanism of killing or resistance evolution. Beyond the direct effect of inhibiting the primary targets, the resulting array of downstream consequences often matter as much¹.

Among the downstream consequences caused by antibiotics, DNA damage is of extensive interest, as repairing the damage often introduces mutations that lead to resistance. A well-studied example is

quinolones, which are known for their DNA-damaging capacity by inhibiting DNA gyrase or topoisomerase IV^{2,3}. Curiously, DNA damage has also been observed with ribosome inhibitors^{4,5}. Since their primary target is not directly involved with DNA, the underlying mechanism is less clear. DNA damage can be induced by various endogenous processes, such as perturbation to the homeostasis of nucleotide pool^{6,7} and blockage of the chromosome by R-loops^{8–11} or transcription factors¹²; these processes usually give rise to DNA lesions by interfering with transcription or DNA replication, especially when DNA repair system is not as competent^{12–15}. Among these mechanisms, oxidized nucleotides derived from reactive oxygen species (ROS) have been proposed to contribute to the DNA-damaging effect of various antibiotics^{16,17}. The ROS theory exemplified the significance of secondary effects and provoked broad discussions in the field of DNA

¹Center for Infection Biology, School of Basic Medical Sciences, Tsinghua University, Beijing, China. ²Tsinghua-Peking Center for Life Sciences, Beijing, China.

³SXU-Tsinghua Collaborative Innovation Center for Frontier Medicine, Shanxi Medical University, Taiyuan, Shanxi Province, China. ⁴Present address: School of Life Science and Technology, ShanghaiTech University, Shanghai, China. ⁵These authors contributed equally: Yayun Zheng, Ruochen Chai, Tianmin Wang.

✉ e-mail: wangtm@shanghaitech.edu.cn; jintaoliu@tsinghua.edu.cn

damage and mutagenesis¹⁸. However, some studies suggest that there might be alternative mechanisms that also contribute to DNA damage^{1,13}. Uncovering those mechanisms is important, as they will inform new strategies for the control of bacterial infections and resistance evolution.

Here, we find that RNAP stalling plays an important role in causing DNA damage and promoting resistance evolution. In many prokaryotes, the leading ribosome on nascent mRNA is in physical contact with the transcribing RNAP, which contributes to the normal processivity of the latter^{19,20}. This transcription-translation coupling is usually well-orchestrated under normal growth^{21,22}. We find that by inhibiting translation during active growth, the coupling is disrupted: ribosomal antibiotics give rise to the global occurrence of premature RNAP stalling and subsequently results in genome instability. While most of the bacteria are killed by genotoxicity, a small subpopulation acquires mutations via SOS-induced mutagenesis. Given that these processes are triggered shortly after antibiotic addition, mutations rapidly accumulate in the population. Under constant selection, mutants showing increased survival are quickly enriched, leading to the rapid evolution of antibiotic resistance.

Results

E. coli evolved resistance after a short period of gentamicin

We focused on gentamicin, a commonly used aminoglycoside antibiotic that targets bacterial ribosomes and exerts strong bactericidal effects. Specifically, we measured the effect of gentamicin on *E. coli* during exponential growth. We used a concentration of 6 $\mu\text{g/ml}$, which

is around two-folds of the minimum inhibitory concentration (MIC) in our experimental setting and is clinically relevant²³. Based on the bacterial growth curve, we found that the treatment arrested bacterial growth within one hour, and then caused a slow decline of biomass, indicating cell death and lysis (Fig. 1a). We measured the viable cell counts at various time points after gentamicin treatment ("Methods" section), and found that bacterial death started at ~40 min after drug treatment; then the number of surviving bacteria exhibited a gradual decrease until ~90 min, by which time ~99.9% cells had been killed (Fig. 1b).

Interestingly, cell death ceased after ~90 min of gentamicin treatment, and a small percentage of the population persisted (Fig. 1b). This phenomenon is typically implicated as phenotypic persistence²⁴. However, when we profiled isolates from the persisted population by measuring bacterial growth under various concentrations of gentamicin (Fig. 1c), we suspected that most of the isolates had mutated, as indicated by either defected growth without gentamicin (Fig. 1d, $t_5 - t_{11}$, labeled in gray; Supplementary Fig. 1a) or increased resistance to gentamicin (Fig. 1d, $t_5 - t_{11}$, labeled in black; Supplementary Fig. 1b). We also profiled isolates from earlier time points, and found that suspected mutants appeared at least as early as 1 h after gentamicin treatment, as indicated by defected growth without gentamicin (Fig. 1d, No. 11 in t_3). Note that those with defected growth may have increased survival under gentamicin, as antibiotic efficacy often depends on bacterial growth²⁵. These results showed that these suspected mutants accumulated quickly after gentamicin treatment, and contributed significantly to the survival of

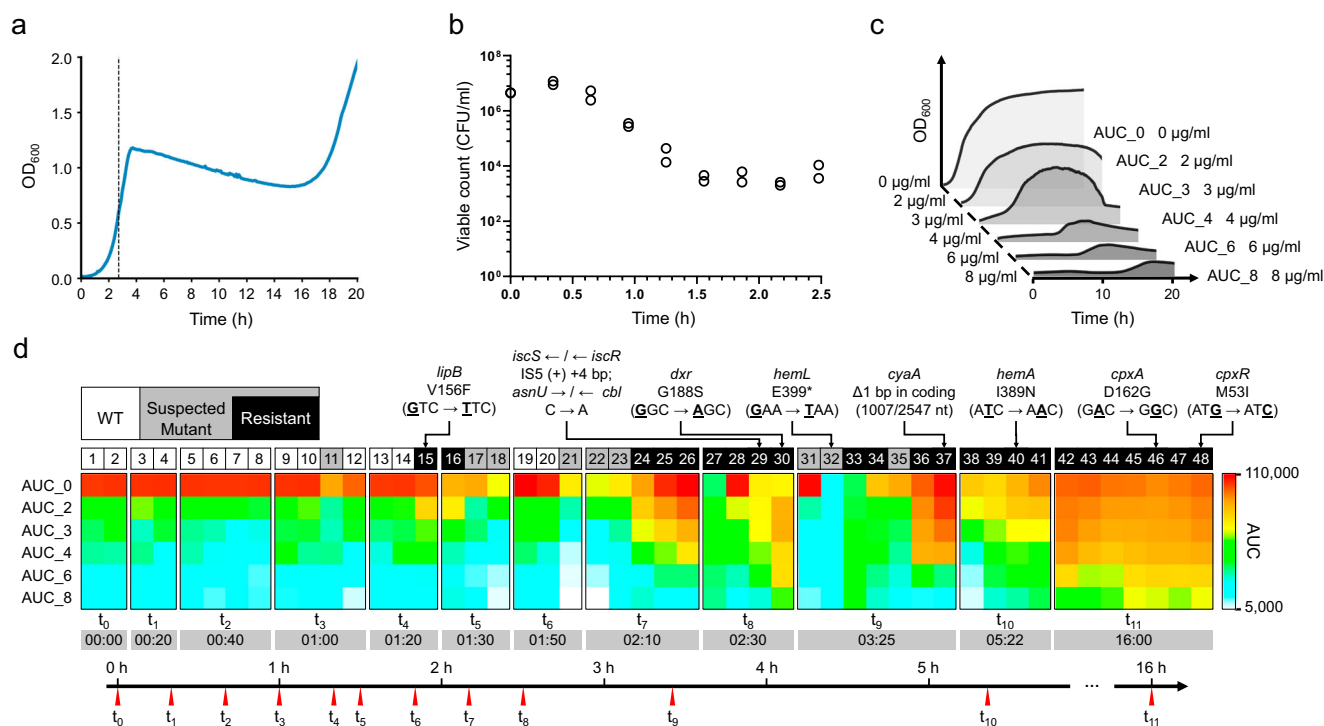


Fig. 1 | Bacterial growth and resistance evolution under gentamicin treatment. **a** Growth curve of *E. coli* treated with 6 $\mu\text{g/ml}$ gentamicin. The dashed line indicates the time of antibiotic addition. Three more biological replicates are shown in Supplementary Fig. 1c. **b** Lethality of gentamicin quantified by the time-kill curve. The viable bacteria counts are indicated as colony-forming units (CFUs). **c** Schematic of drug susceptibility assay. Bacterial isolates were cultured in a medium containing gentamicin of different concentrations. For each concentration, area under the growth curve (AUC, integration of the growth curve from 0 to 20 h, see "Methods" section) was quantified and indexed as "AUC_x" ($x = 0, 2, 3, 4, 6$ and 8, corresponding to gentamicin concentration in $\mu\text{g/ml}$). **d** Drug susceptibility

of bacterial isolates recovered during gentamicin treatment. The heatmaps show AUCs of the isolates under 6 different gentamicin concentrations (as indicated in c). The time points labeled below the heatmaps and on the axis indicate the duration of gentamicin treatment before the antibiotic was washed out and the clones were isolated (by plating the culture on antibiotic-free agar plates). The isolates are categorized into wild type, suspected mutant (altered AUC with or without gentamicin) including resistant (increased AUC under any concentration of gentamicin); these categories are highlighted by colors white, gray, and black respectively above the heatmaps. Mutations identified by resequencing from eight selected isolates are listed above. Source data are provided as a Source Data file.

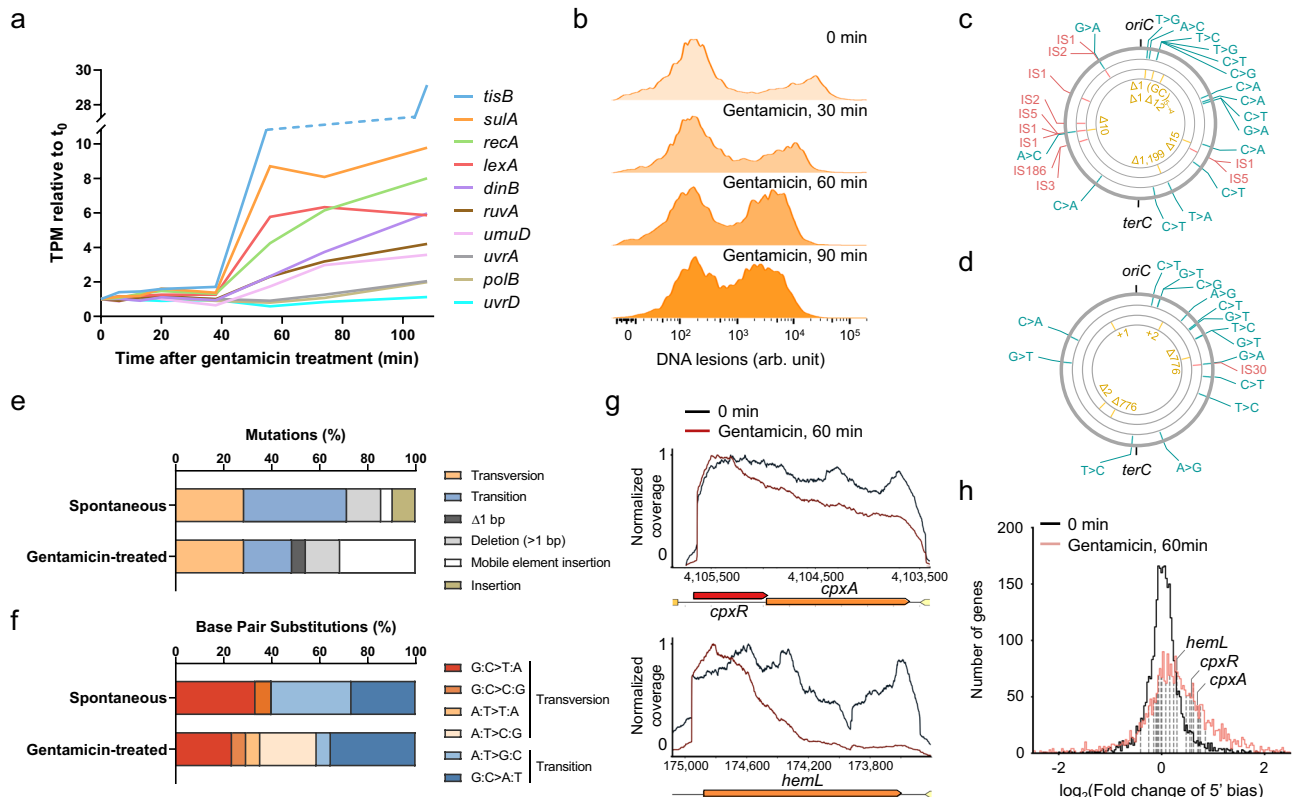


Fig. 2 | Genomic instability and induced mutagenesis caused by gentamicin.

a Activation of SOS response by gentamicin treatment. Shown are the fold changes of transcripts per million (TPM) of representative genes in SOS regulon measured by RNA-seq. **b** DNA damage induced by gentamicin. DNA damage levels at 0 min, 30 min, 60 min, and 90 min post gentamicin addition were measured by the TUNEL assay and flow cytometry. Damage-positive cell ratio: 0 min, 31.3%; 30 min, 36.6%; 60 min, 48.2%; 90 min, 50.9%. Representative of 2 biological replicates.

c, d Mutations isolated after gentamicin treatment (**c**) and spontaneous mutations (**d**). The positions of the mutations on the genome were indicated, along with the positions of *oriC* (3,925,744–3,925,975) and *terC* (1,609,157–1,609,179) for reference. Different rings represent different types of mutations: base pair substitutions (cyan), mobile element insertions (red), deletions and insertions (yellow).

e Spectrums of spontaneous mutations and of mutations isolated after gentamicin treatment. Shown are the percentages of different mutation categories.

f Spectrums of the base pair substitutions category shown in (**e**). **g** Transcript coverage of exemplified operons *cpxRA* and *hemL* measured by RNA-seq at 0 min (black outline) and 60 min (red outline) after gentamicin treatment. **h** Distributions of fold changes of 5' coverage bias at 0 min and 60 min after gentamicin treatment. Shown are histograms of the fold changes of 5' coverage bias from 3126 protein-coding genes (see “Methods” section), and the fold changes at 0 min were calculated from two replicates without gentamicin ($P = 9.07 \times 10^{-88}$, one-sided MWU test). The values of 20 protein-coding genes with detected mutations in (**c**) were marked by dashed lines. Source data are provided as a Source Data file.

the population under gentamicin. Consistent with the emergence of resistance, we repeatedly (>90% of the cases) observed recovery of growth after 12 h of gentamicin treatment (Supplementary Fig. 1c). To reveal the type of genetic alternations that happened in this process, we performed genome resequencing to identify the mutations in the isolates. Specifically, we selected eight representative isolates from five representative time points (Fig. 1d, marked by arrows). We identified mutations in all the isolates. The mutations resided on genes ranging from stress response to respiration, and many of the genes are associated with altered susceptibility or resistance to gentamicin^{23,26,27}. Taken together, under gentamicin treatment, the *E. coli* population rapidly accumulated mutants with high genetic diversity and is finally dominated by resistant strains within a few hours.

Gentamicin-induced genome instability and mutagenesis

We wondered why resistant mutants emerged within such short treatment by gentamicin. Many of the resistant isolates exhibited severe growth defects in the absence of gentamicin (Fig. 1d and Supplementary Fig. 1b). If those mutants emerged spontaneously, they should be overwhelmed by their siblings who grow normally and should only consist of a very tiny fraction of the population. Therefore, we speculated that induced mutagenesis might be underlying the

emergence of the resistant mutants. Consistent with our speculation, transcriptomic analysis showed that gentamicin treatment activated SOS response in 40 min (Fig. 2a). SOS response is normally triggered by DNA damage, which activates error-prone DNA damage repair. Using a fluorometric terminal deoxynucleotidyl transferase-dUTP nick end labeling (TUNEL) assay (Supplementary Fig. 2), we measured the amount of DNA lesions in the bacteria⁶. We found that the ratio of lesion-positive cells increased substantially after gentamicin treatment and plateaued in 60 min (Fig. 2b). These results showed that gentamicin treatment caused dramatic DNA damage and triggered SOS response.

Since the SOS response performs DNA repair in an error-prone manner, we should be able to find the trace of the mutagenic repair in the mutants. Therefore, we treated the bacteria with gentamicin for 3 h, cultured the surviving ones on an antibiotic-free LB agar plate, and profiled them by genome resequencing. Among the 100 clones, 98 carry mutations. After removing duplicates, there are 33 mutants and 35 distinct mutations in total (Fig. 2c; Supplementary Data 1). As a reference, our wild-type strain contains 21 genetic variation loci compared with the reference genome of *E. coli* K12 (NC000913.3) (Fig. 2d; Supplementary Data 2), and likely accumulated those mutations spontaneously during strain propagation. Interestingly, we observed distinct mutation

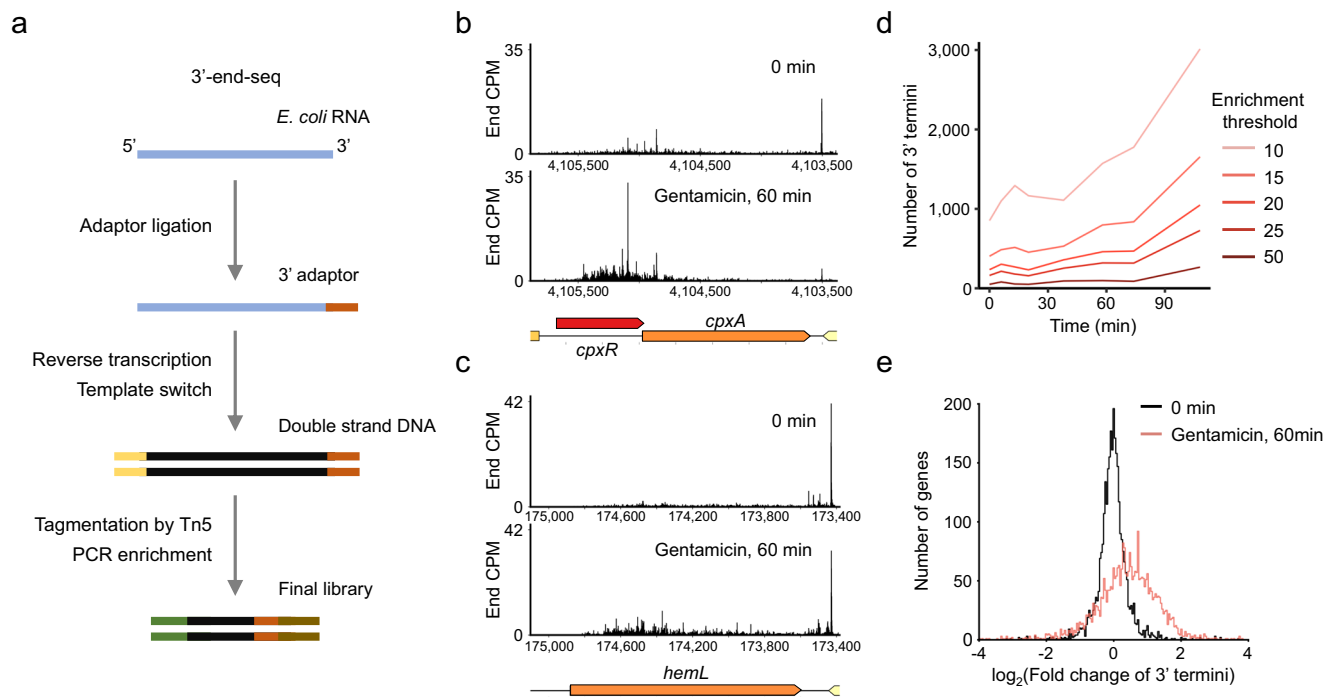


Fig. 3 | 3'-end-seq analysis revealed pervasive occurrence of 3' termini induced by gentamicin. a Schematic of the 3'-end-seq workflow. **b, c.** 3'-end-seq profiles across the *cpxRA* (**b**) and *hemL* (**c**) genes before (upper) and 60 min after gentamicin treatment (lower). The plotted spikes indicate counts per million (CPM) of 3' end at each nucleotide analyzed from 3'-end-seq with strand specificity. The arrow-shaped boxes below the plots show the corresponding genes and their coordinates on the genome. **d** Dynamics of the

dominant 3' termini upon gentamicin treatment under five different thresholds. **e** Distributions of the fold changes of 3' termini at 0 min and 60 min after gentamicin treatment (shown in black and red respectively, $P = 2.68 \times 10^{-159}$, one-sided Mann-Whitney U (MWU) test). Shown are histograms of the fold changes of 3' termini from 3126 protein-coding genes, and the fold changes at 0 min were calculated from two replicates without gentamicin. Source data are provided as a Source Data file.

patterns between our wild-type strain and the isolates obtained after gentamicin treatment. Specifically, the mutants isolated after gentamicin treatment exhibited an increase of structural variants, including insertion sequences (ISs) and deletions (Fig. 2e). Those variants are associated with recombination during DNA damage repair^{28,29}. Consistently, we observed activation of transposase genes corresponding to the relevant ISs after 60 min of gentamicin treatment (Supplementary Fig. 3a), which was right before the rise of mutants (Fig. 1d). In addition, the emerging $\Delta 1$ bp deletion (Supplementary Fig. 3b) is known to be frequently generated during translesion synthesis by DNA polymerase IV (DinB)³⁰, which was activated during the SOS response after gentamicin treatment (Fig. 2a).

Besides structural variants, the patterns of base pair substitutions were also distinct between our wild-type strain and the isolates after gentamicin treatment (Fig. 2f). In spontaneous mutations, transitions (60%) were more prevalent than transversions (40%). By contrast, we identified more transversions (59%) than transitions (41%), which has been reported as a hallmark of SOS-induced mutagenesis³¹. Among the transversions in the gentamicin-treated isolates, akin to the $\Delta 1$ bp deletions mentioned above, A:T $\rightarrow$ C:G is known to be derived from DinB-dependent translesion synthesis³². G:C $\rightarrow$ T:A is known to be caused by genotoxic 8-oxo-guanine (8-oxo-G) during oxidative stress¹⁵, and it has been reported that gentamicin could cause oxidative stress and thus DNA damage¹⁶. However, in our case, 8-oxo-G only showed a minor contribution to the mutations recovered after gentamicin treatment (Fig. 2f, 4 out of the 10 transversions). Together, these results suggest that gentamicin triggered genome instability and induced mutagenesis in *E. coli*. Consistent with this claim, deleting the error-prone DNA polymerase coding genes *polB*, *dinB*, and *umuC* in the SOS regulon decreased bacterial survival under gentamicin treatment (Supplementary Fig. 3c, d).

The genome instability and mutagenesis were caused by the global occurrence of premature RNAP stalling

Gentamicin targets ribosomes and disrupts translation. We wondered how a ribosome-targeting antibiotic could trigger genome instability and induce mutagenesis. Interestingly, when we analyzed the transcriptomic changes upon gentamicin treatment, we made an intriguing observation – gentamicin caused pronounced decrease of transcript coverage toward the 3' ends of many genes (Fig. 2g). Based on a measurement of coverage bias (“Methods” section), we found that the coverage bias of 29% of protein-coding genes increased by more than 1.5-fold (Fig. 2h). To further investigate this phenomenon, we utilized a method to dissect the 3' ends of bacterial transcripts with single-nucleotide resolution on the genome level, here referred to as 3'-end-seq (Fig. 3a, “Methods” section). First, we validated the accuracy of the method: 3'-end-seq reliably identified known intrinsic termination sites during normal RNAP elongation in the absence of gentamicin (upper panels in Fig. 3b, c and Supplementary Fig. 4). Then, using this method, we identified substantial amount of transcripts terminated within the gene-coding regions in the presence of gentamicin (lower panels in Fig. 3b, c); the termini may appear dominantly at particular sites, or may spread across the entire coding region, depending on the specific gene.

To quantitatively dissect the 3' termini within the protein-coding regions (i.e., excluding intrinsic termination, Supplementary Fig. 5) caused by gentamicin treatment at the genome scale, we performed two statistical analyses. The first analysis counted the number of termini surpassing a predefined “enrichment” threshold, named dominant 3' termini (“Methods” section); the second analysis quantified the normalized fold change of 3' termini in each gene upon gentamicin treatment (“Methods” section). Using these two analyses, we showed that both the number of dominant 3' termini and the fold change of 3'

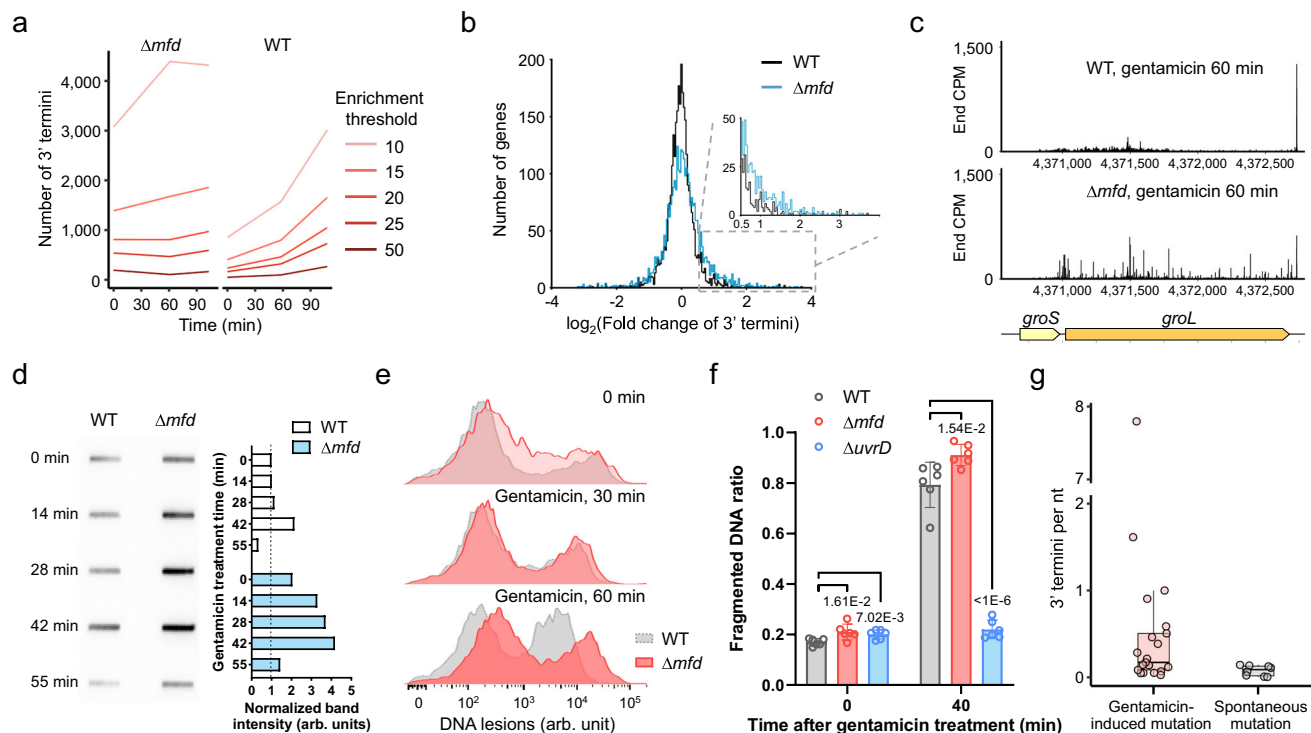


Fig. 4 | Genome-wide PRSEs, genome instability, and mutagenesis triggered by gentamicin. **a** Dynamics of the dominant 3' termini in the Δmfd strain (left), together with the wild-type strain (right, the same data as in Fig. 3d). **b** Distributions of fold changes of 3' termini of the wild-type and the Δmfd strain without gentamicin treatment (shown in black and blue respectively, $P = 3.60 \times 10^{-13}$, one-sided MWU test). The inset zooms in on the right tail of the distribution. **c** 3'-end-seq profiles across the *groSL* operon in the wild-type strain (upper) and the Δmfd strain (lower) after 60 min of gentamicin treatment. **d** R-loop accumulation after gentamicin treatment in the wild-type and the Δmfd strain. Left panel shows the Western blot of the R-loop in the genomic DNA of both strains across gentamicin treatment. Right panel shows the quantification of bands using ImageJ, with all the signals normalized to the wild-type strain's signal at 0 min (white for wild type, blue for Δmfd). A biological replicate is shown in Supplementary Fig. 7. **e** DNA damage level

of the Δmfd strain (red) overlaid with the wild-type strain (gray, the same data as in Fig. 2b) after gentamicin treatment. Representative of 2 biological replicates. **f** DNA damage in the wild-type, Δmfd , and $\Delta uvrD$ strains after gentamicin treatment. DNA damage levels after 0 min and 40 min gentamicin treatment were measured by PFGE analysis ($n = 6$ biological replicates). P -values (two-tailed t -test) were labeled on the columns respectively. Error bars, $\pm$ std. **g** Distributions of PRSE intensities of coding genes carrying gentamicin-induced mutations (21 genes from Fig. 2c) and spontaneous mutations (8 genes from Fig. 2d) respectively ($P = 6.87 \times 10^{-3}$, one-sided MWU test). Shown are PRSE intensities after 60 min of gentamicin treatment, calculated from 3' termini numbers per nt of each gene (see "Methods" section). Box plots: horizontal lines within boxes, medians; upper and lower box edges, first and third quartiles, respectively; whiskers, 1.5 times the interquartile range; circles, data for mutation genes. Source data are provided as a Source Data file.

termini showed a significant increase after gentamicin treatment (Fig. 3d, e). Quantitatively, the total number of dominant 3' termini increased by 3-fold after 110 min of treatment (Fig. 3d); moreover, the fold change of 3' termini increased by more than 1.5-fold in 42% of the genes after 60 min of gentamicin treatment (Fig. 3e). These results showed that gentamicin caused pervasive occurrence of 3' termini on the genome level.

Partial transcripts can be caused by arrested transcription elongation^{33,34}. Therefore, we speculated that the wide occurrence of 3' termini under gentamicin was caused by RNAP stalling. To this end, we deleted *mfd*, which encodes a DNA translocase that recognizes and rescues stalled RNAP^{35,36}. First of all, consistent with the function of Mfd, the knockout promoted the frequency of 3' termini occurring in protein-coding regions before gentamicin treatment (Fig. 4a, b). In addition, Δmfd significantly boosted the occurrence of dominant 3' termini following exposure to gentamicin (Fig. 4a, c and Supplementary Fig. 6). Therefore, we propose that the 3' termini located within the protein-coding regions were mainly caused by premature RNAP stalling events (PRSEs).

It was suggested that RNAP stalling could increase the occurrence of R-loops – transcription-derived DNA:RNA hybrid structures that persist after premature dissociation of the RNAP³⁷. Indeed, we observed an increase of R-loop formation 30–40 min after gentamicin treatment (Fig. 4d and Supplementary Fig. 7), which is consistent with the increase of PRSEs as characterized by 3'-end-seq (Fig. 3d).

Furthermore, *mfd* knockout also led to higher R-loop formation than wild type both before and after gentamicin treatment (Fig. 4d and Supplementary Fig. 7). These results suggested that gentamicin treatment triggered widespread occurrence of PRSEs.

Given the widespread occurrence of PRSEs, we wondered whether it was the cause of gentamicin-induced genome instability. According to this speculation, boosting PRSEs should result in an increase in DNA damage. Indeed, using both the TUNEL assay and pulse-field gel electrophoresis (PFGE), we found that knocking out *mfd* amplified damage in genomic DNA under gentamicin treatment (Fig. 4e, f and Supplementary Fig. 8a, d). Furthermore, we performed ChIP-seq to profile RNAP stalling and DNA damage (double-strand breaks, DSBs) simultaneously in vivo. The results suggested that many of the observed DSBs were likely caused by RNAP stalling (Supplementary Note 1; Supplementary Fig. 9; Supplementary Methods).

Remarkably, the genome instability induced by gentamicin can be abolished by deleting the RNAP-bound helicase UvrD: after 40 min of gentamicin treatment, >70% of genomic DNA in the wild-type strain was fragmented; in contrast, no significant increase in DNA damage was detected in the $\Delta uvrD$ strain (Fig. 4f and Supplementary Fig. 8b). UvrD has recently been implicated to be involved in transcription-coupled DNA repair^{38,39}. However, our result showed that it counter-intuitively promoted DNA damage under gentamicin. UvrD normally functions by backtracking the RNAPs when they are stalled at DNA lesions or other blockage sites, after which nucleotide excision repair

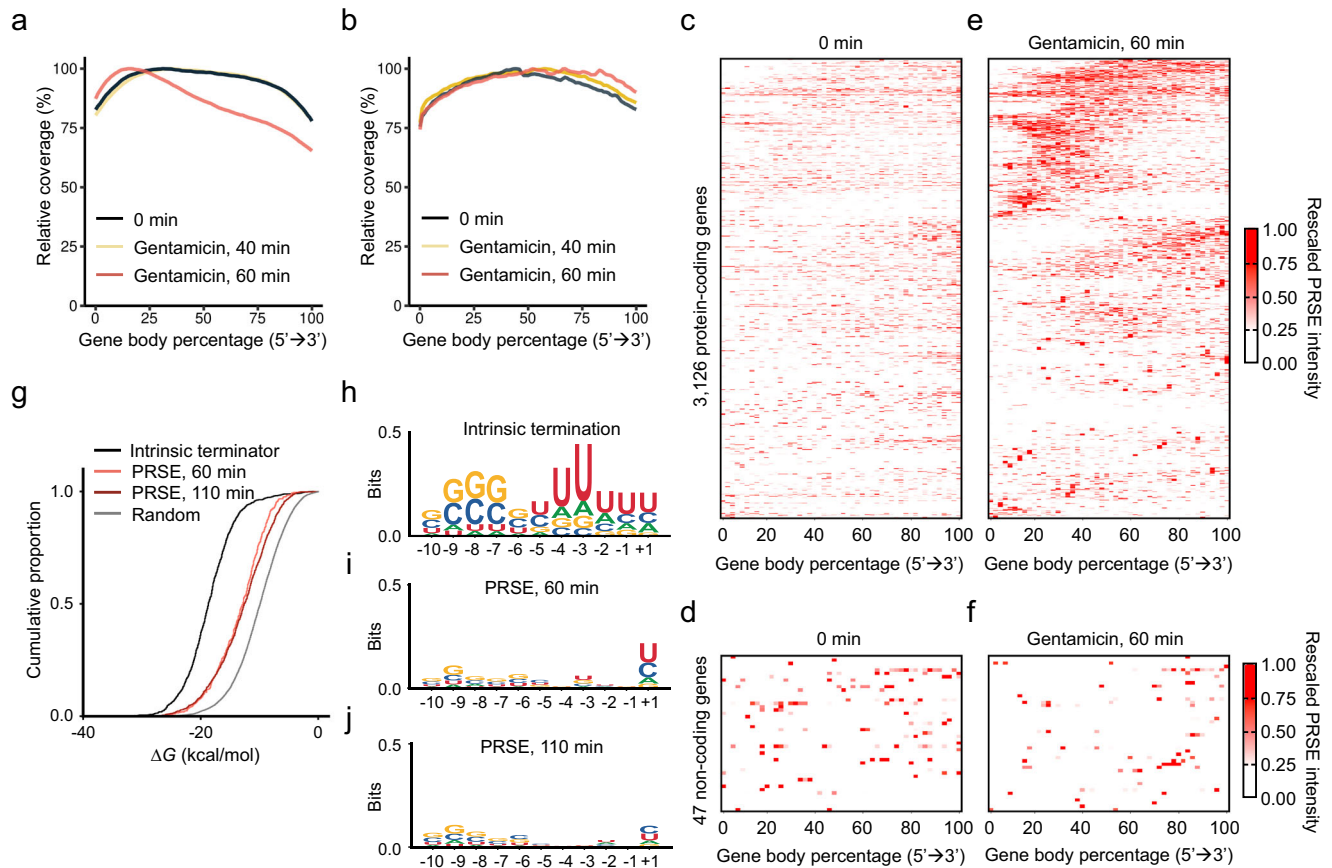


Fig. 5 | Genome-wide analysis of PRSEs caused by disruption of transcription-translation coupling. **a, b** Averaged transcript coverage of 3126 protein-coding genes (**a**) and 47 noncoding genes (**b**). Read coverages measured by RNA-seq after 0 min (black), 40 min (yellow), and 60 min (red) of gentamicin treatment are shown. **c–f** Distributions of PRSEs along the gene body. Shown are heatmaps of rescaled PRSE intensities (see “Methods” section), in protein-coding genes and noncoding genes respectively, after 0 min and 60 min of gentamicin treatment, with each row representing a single gene. **g** Cumulative distribution of minimum folding free energy (MFE) of dominant PRSEs. As a negative control, the MFE of random 50 mers on the *E. coli* genome was calculated ($n = 10,000$). As a positive

control, the MFE of 50 mers upstream of intrinsic terminators was calculated ($n = 1330$). MFE of PRSEs after 60 min ($n = 460$) and 110 min ($n = 1048$) of gentamicin treatment were obtained similarly, with enrichment cutoff all set as 20. Two-sided Kolmogorov–Smirnov test for group-wise comparison: 60 min ($P = 4.42 \times 10^{-84}$ vs. negative control; $P = 2.33 \times 10^{-15}$ vs. positive control) and 110 min ($P = 1.41 \times 10^{-105}$ vs. negative control; $P = 2.78 \times 10^{-15}$ vs. positive control). **h–j** Motif analysis of sequences upstream of intrinsic terminators (**h**) and of dominant PRSEs after 60 min (**i**) and 110 min (**j**) of gentamicin treatment. Enrichment cutoffs and sample sizes are as in (**g**). Source data are provided as a Source Data file.

factors UvrABC are recruited to perform the repair^{38,40}. In our scenario, the gentamicin-induced PRSEs may also induce backtracking of the RNAPs and subsequent recruiting of UvrABC; given the excision activity of UvrABC, they may have caused unintended damage to the DNA. Consistent with this speculation, we found that knocking out *uvrB* or *uvrC* led to a significant decrease in gentamicin-induced DNA damage (Supplementary Fig. 8b, c). Additionally, our data suggested the DNA damage was largely replication-independent, since replication inhibition did not cause a significant decrease in DNA damage under gentamicin treatment (Supplementary Fig. 10a, b). Our experiment also suggested that ROS were unlikely to be the primary cause of genomic instability: The accumulation of ROS lagged behind DNA damage. Specifically, the subpopulation with increased intracellular ROS did not appear until 60 min after gentamicin treatment (Supplementary Fig. 10c), while DNA damage had already plateaued by that time (Fig. 2b).

Given that genome instability activates SOS-induced mutagenic repair, we speculated that regions harboring PRSEs may serve as mutational hotspots within the genome. We quantified the normalized amount of 3' termini in individual genes (“Methods” section). Consistent with our speculation, we found that the genes harboring gentamicin-induced mutations experienced significantly higher rates of PRSEs than the genes harboring spontaneous mutations (Fig. 4g).

Collectively, these results suggest that gentamicin-induced genome instability and mutagenesis by triggering the global occurrence of premature RNAP stalling.

Premature RNAP stalling was caused by disruption of transcription-translation coupling

Next, we wondered how gentamicin, a ribosome-targeting antibiotic, triggered the widespread occurrence of premature RNAP stalling. In many bacteria, including *E. coli*, RNAP is closely followed by a trailing ribosome, which ensures that the movement of RNAP and ribosome is precisely coordinated^{20,41}. This is known as transcription-translation coupling, which contributes to RNAP elongation⁴¹. It has been demonstrated on a few highly expressed operons that inhibiting translation can compromise the RNAP processivity¹⁹. Therefore, we speculated that gentamicin triggered genome-wide PRSEs by disrupting transcription-translation coupling.

Our speculation predicts that, following gentamicin treatment, large-scale PRSEs would only be observed in protein-coding genes, while noncoding genes, which are not translated, would not exhibit significant PRSEs. To test this prediction, we analyzed the read coverage of our RNA-seq data (“Methods” section). Prior to gentamicin treatment, both protein-coding and noncoding genes displayed a uniform read coverage (Fig. 5a, b), indicating normal RNAP elongation.

After 60 min of gentamicin treatment, the read coverage on protein-coding genes showed a pronounced skewness towards the 5' ends (Fig. 5a), suggesting interruption of RNAP elongation; in contrast, no such coverage polarity was observed in the noncoding genes (Fig. 5b). These results suggest that gentamicin led to genome-wide PRSEs via disrupting transcription-translation coupling. Consistent with this claim, we also observed substantial PRSEs in several leader-peptide-regulated genes (Supplementary Fig. 11), where transcription-translation coupling has been shown to play critical roles in transcription regulation⁴².

We further looked into the precise locations of the PRSEs along individual genes. First of all, PRSEs were infrequent and randomly distributed in the absence of gentamicin (Fig. 5c, d), confirming the normal RNAP processivity under steady growth. Under gentamicin treatment, we found a significant increase of PRSEs in certain regions of the protein-coding genes (Fig. 5c, e; see Supplementary Note 2 and Supplementary Fig. 12 for functional and protein-level analyses of PRSE-affected genes); in contrast, there was no significant increase or enrichment of PRSEs in the noncoding genes (Fig. 5d, f). The non-uniform distribution of PRSEs in the protein-coding genes suggests that their occurrence may be sequence-specific. It is known that RNA hairpins can cause RNAP stalling⁴³, and this effect can be inhibited by transcription-translation coupling^{20,22,44}. We calculated the minimum folding free energy (MFE) of the sequences 50 nt upstream of the dominant PRSEs ("Methods" section). The results showed that they lead to significantly lower MFE than random sequences (Fig. 5g), suggesting that the RNAs they encode tend to form hairpins. We further analyzed whether there were any patterns in the sequences preceding the PRSEs. We found that in contrast to the canonical GC-rich hairpin-forming plus polyU sequence for intrinsic transcription termination (Fig. 5h, "Methods" section), there was no obvious consensus motif preceding the dominant PRSEs caused by gentamicin (Fig. 5i, j), which is consistent with the fact that their MFE was higher than that of the intrinsic terminators (Fig. 5g). These results suggest that, following gentamicin treatment, RNAP processivity became susceptible to weak RNA structures that were normally inhibited by transcription-translation coupling.

An overlooked pathway underlying the efficacy of ribosomal antibiotics and resistance evolution

Based on the results and analysis discussed above, we propose the following model (Fig. 6a, see also Supplementary Note 3 and Supplementary Fig. 13): gentamicin disrupts transcription-translation coupling, which results in genome-wide premature RNAP stalling at locations where RNA hairpin structures form; the stalled RNAPs promote DNA damage and cause genome instability via transcription-coupled repair factors; error-prone repair of the damage causes introduction of mutations, which promotes rapid emergence of antibiotic resistance in a small percentage of the cells.

Our model also suggests a mechanism that affects the efficacy of ribosome inhibitors^{1,45}: in addition to the known modes of actions⁴⁵, RNAP stalling, which causes genome instability, also contributes to the killing of the bacteria. Indeed, deleting *mfd* and any of the other PRSE-rescuing factors (*greA*, *greB*, *rnhA*, and *rnhB*) gave rise to fitness defect under gentamicin treatment (Fig. 6b). Since genome instability can be repaired through SOS response, the activation of SOS should be crucial for the survival of the bacteria. We constructed a fluorescent RNA reporter⁴⁶ to monitor the promoter activity of an SOS response master regulator gene *recA* ("Methods" section). Using this reporter, we found that the level of SOS response was highly heterogeneous in different cells. Using fluorescence-activated cell sorting (FACS), we extracted the high and low SOS responders respectively after 40 min of gentamicin treatment (Fig. 6c). Consistent with the

known function of SOS response, the high SOS responders exhibited three-fold higher survival than the low ones (Fig. 6d), which might act as a reservoir for the subsequent induced evolution toward antibiotic resistance.

Finally, we monitored the response of the bacteria under a comprehensive list of ribosome-targeting antibiotics. We found that all the tested drugs caused rapid activation of SOS response (Fig. 6e), which indicated DNA damage. To see whether RNAP stalling contributed to the efficacy of those antibiotics, we profiled the fitness impact of *mfd* knockout. We found that increasing RNAP stalling gave rise to substantial fitness defects under all the ribosome inhibitors, independent of their binding sites (Fig. 6f). This effect was not observed with other classes of antibiotics (Supplementary Fig. 14). These results suggest that the mechanism uncovered here applies to translation inhibition in general.

Discussion

Our work revealed a mechanism on how ribosome-targeting antibiotics could trigger genome instability and induced mutagenesis. The mutagenesis led to the rapid evolution of antibiotic resistance. Our results suggest that the regions harboring PRSEs served as mutational hotspots within the genome. We discussed the intrinsic factor determining the location of the PRSEs, namely, the DNA sequences that promote secondary RNA structures impeding RNAP elongation. This intrinsic factor makes some genes more prone to PRSEs than others, and therefore more prone to mutations, giving rise to biased mutagenesis. In addition to intrinsic factors, extrinsic factors may also play a role. Specifically, ribosomes serve as primary sensors for nutrient availability and various stresses. Our work showed that they could modulate mutagenesis by influencing the dynamics of RNAP and the occurrence of PRSE in the genome. Therefore, biased mutagenesis is also shaped by the environment. It is an open question whether the mutagenesis could be biased toward genome locations that benefit the survival of the bacteria⁴⁷.

It was known that aminoglycosides could cause genome instability, thus activating the SOS response^{4,5}. This has not been reported in *E. coli* mainly due to the low concentrations tested or short treatment durations in related studies^{48,49} (Supplementary Fig. 15). The underlying mechanism by which aminoglycosides cause DNA damage was often attributed to ROS¹⁶. However, in our experiments, ROS had a limited contribution to the early induction of genome instability (Supplementary Fig. 10c). Our results showed that genome instability was caused by RNAP stalling via induction of transcription-coupled nucleotide excision repair (NER). In support of our claim, a similar process was also found in mammalian cells⁵⁰. Normally, NER is induced when the RNAP is stalled by DNA lesion³⁸. Our results suggested that NER can also be induced when RNAP is stalled by the decoupling of transcription and translation. In addition, NER counter-intuitively promoted DNA damage under gentamicin. We proposed that the excision activity of UvrABC contributed to the cause of the damage. This coincides with the previous study reporting a tendency of NER machinery to excise undamaged DNA⁵¹. More biochemical experiments will be needed to further validate these claims, and other DNA repair factors with nuclease activity should also be taken into consideration when discussing DNA damage generated endogenously.

It was reported that knocking out *mfd* inhibited the evolution of resistance under various antibiotics⁵². Our work further confirmed inhibiting genes that rescue stalled RNAPs as a promising strategy for countering the development of resistance. We showed that DNA lesions were generated by RNAP stalling. Since unmasking the lesion from RNAP is required for the onset of SOS response⁴, inhibiting factors that help to dislodge RNAP can potentially delay the evolution of antibiotic resistance. In addition, our results also suggest a new strategy for improving antibiotic efficacy. Genome integrity, which assures the individuality, life-sustaining function, and normal replication of the

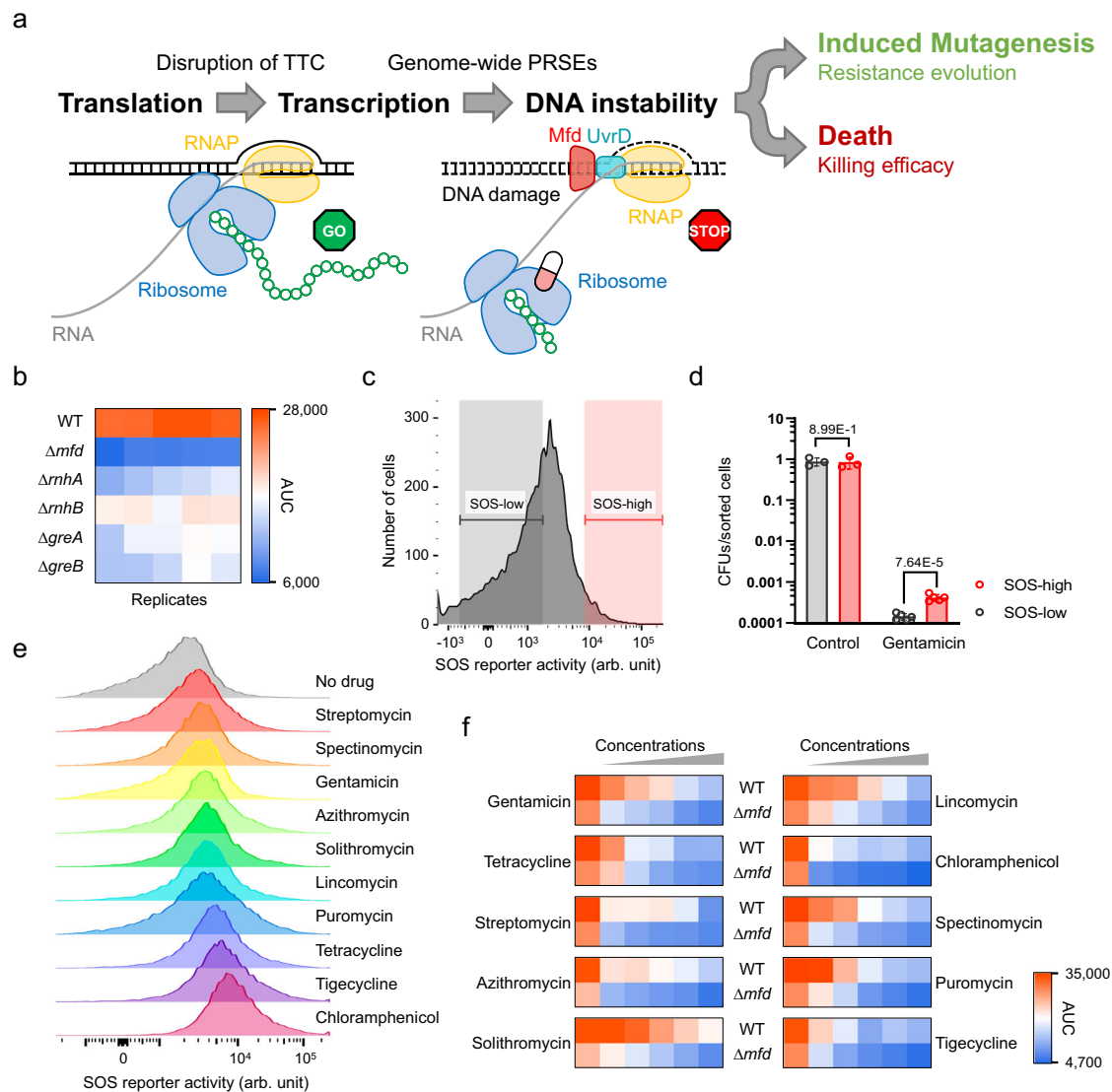


Fig. 6 | An overlooked pathway underlying the efficacy of ribosomal antibiotics and resistance evolution. **a** Schematic of the model. Potential mechanisms contributing to RNAP stalling upon transcription-translation decoupling were discussed in Supplementary Note 3 and Supplementary Fig. 13. **b** Gentamicin susceptibilities of PRSE rescue factor knockouts. The heatmaps show AUCs of bacterial growth curves during the first 10 h under 6 $\mu\text{g/ml}$ gentamicin. The calculation method used is similar to that employed in Fig. 1c. Each row represents a single strain with 5 biological replicates. **c** An example distribution of SOS response activity as measured by the *recA* transcription reporter. Red and gray areas indicate the gating for SOS-high (top 5%) and -low (bottom 40%) responders, respectively (see “Methods” section). **d** Survival rates of high and low SOS responders sorted from (c) after 40 min of gentamicin treatment. As for the control group, cells without gentamicin treatment were also sorted with the same criteria and tested for

survival. *P*-values (two-tailed *t*-test) were labeled on the columns respectively. Error bars, $\pm$ std ($n = 3$ in the control group, $n = 5$ in the gentamicin-treated group; biological replicates). **e** Activation of the SOS response under a comprehensive list of ribosomal antibiotics. SOS response was measured 40 min after drug addition as in (c). Representative of 2 biological replicates. **f** The susceptibilities of wild-type and Δmfd strains under a comprehensive list of ribosomal antibiotics. The concentrations are as below: Gentamicin, 0, 2.5, 5, 10, 20, 40 $\mu\text{g/ml}$; Tetracycline, 0, 1.25, 2.5, 5, 10, 20 $\mu\text{g/ml}$; Streptomycin, 0, 14.6, 29.3, 58.6, 117.2, 234.4 $\mu\text{g/ml}$; Azithromycin, 0, 7.8, 15.6, 31.1, 62.5, 125 $\mu\text{g/ml}$; Solithromycin, 0, 7.8, 15.6, 31.3, 62.5, 125 $\mu\text{g/ml}$; Lincomycin, 0, 390.6, 781.3, 1562.5, 3125, 6250 $\mu\text{g/ml}$; Chloramphenicol, 0, 11.7, 23.4, 46.9, 93.8, 187.5 $\mu\text{g/ml}$; Spectinomycin, 0, 39.1, 78.1, 156.3, 312.5, 625 $\mu\text{g/ml}$; Puromycin, 0, 97.7, 195.3, 390.6, 781.3, 1562.5 $\mu\text{g/ml}$; Tigecycline, 0, 1, 2, 4, 8, 16 $\mu\text{g/ml}$. Source data are provided as a Source Data file.

cells, plays a central role in bacterial survival. According to our model, inhibiting genes that rescue stalled RNAPs, such as R-loop removal and anti-backtracking factors of RNAP, may give rise to elevated levels of DNA damage. If the amount of lesions overwhelms DNA repair capacity, bacteria can be quickly killed. Drug efficacy and resistance evolution are intertwined – promoting drug efficacy could reduce the emergence of resistance. Here, we identified promoting RNAP stalling as a one-stone-for-two-birds strategy, which simultaneously boosts antibiotic efficacy and reduces the development of resistance.

Apart from Mfd, a recent study by Carvajal-Garcia et al. further discussed the contribution of the NER pathway to the evolution of antibiotic resistance⁵³. It has been demonstrated that Mfd, UvrABC,

and translesion DNA polymerases filling the gaps produced by NER are crucial for the development of rifampicin resistance in *B. subtilis* and several other bacterial species. Consistent with their finding, our work also suggests the significance of transcription-coupled NER pathway and mutagenic repair system in resistance evolution, though the primary cause of induction may vary: Carvajal-Garcia et al. discussed long-term antibiotic treatments, where mutations originating from NER were mainly driven by oxidation stress. In our study, genome instability and mutagenesis were induced in a much shorter time frame (within 1 h, Fig. 2b), when intracellular ROS has not yet accumulated to a significant level (Supplementary Fig. 10c). We propose that the genome instability and mutagenesis is due to pervasive RNAP stalling

following sudden translation inhibition during the active growth phase of the bacteria.

Finally, bacterial response to antibiotics is a complex issue. A systems perspective could help us to tackle this challenge. By dissecting the complex interplay between multiple molecular processes, we found that antibiotic efficacy and resistance evolution could be understood through an underexplored pathway. Furthermore, our work revealed PRSE as an overlook factor in understanding bacterial response. The PRSEs induced by antibiotics disrupt the normal transcription of genes, leading to genome-wide failure in the expression of various functional proteins and dysregulation of crucial pathways. Understanding the consequences and how bacteria respond will inform better use of existing antibiotics and promote the development of new ones¹.

Methods

Bacterial strains and cultivation conditions

In all experiments, *E. coli* K12 BW25113 was used (originally from CICC, 23872) as the wild-type strain. Single colony was picked up from a stock LB agar plate and was grown aerobically with shaking at 220 rpm in Luria-Bertani (LB) broth (10 g/l tryptone (Oxoid, LP0042), 10 g/l NaCl (Macklin, S805275), 5 g/l yeast extract (Oxoid, LP0021)) for 12 h as the seed culture. All cultures were carried out at 37 °C with shaking at 220 rpm. When necessary, 100 mg/l ampicillin (TargetMol, T6386) was supplemented to maintain the plasmid. All knockout strains were derived from the Keio collection⁵⁴. The excision of FRT-flanked *kan* cassette in Keio knockout strains was carried out by FLP recombinase-expressing pCP20 plasmid⁵⁵; the removal of the *kan* cassette was further confirmed by PCR amplification before using the strain (Supplementary Fig. 16). Besides, the *tolC* F16* (TTC→TGA) strain was an *E. coli* K12 BW25113 derivative constructed by CRISPR genome editing⁵⁶. *E. coli* strain DH10B (BioMed) was used as a host during molecular cloning.

Plasmid construction

The construction of the *P_{recA}pepper* plasmid used to surveil SOS responses was described below. We first adopted PCR to amplify *recA* promoter from *E. coli* K12 BW25113 genomic DNA, ampicillin resistance cassette, p15A plasmid origin of replication from p15A_PpyrLBI_mCherry⁵⁷, giving rise to three DNA fragments. We also purchased a synthetic pepper construct consisted of four repeated HBC620 binding motif sequences from GENEWIZ Int. Co., which was designed according to the sequence reported by Chen et al.⁴⁶. All these four DNA fragments were constructed into a *P_{recA}pepper* plasmid by Gibson assembly. The nucleotide sequence of the constructed plasmid was confirmed by Sanger sequencing and deposited at benchling (<https://benchling.com/s/seq-oujd4XqzHbQnUPkf85hy?m=slm-nDZVa3rVTOgNMWLMQW>).

Drug treatment

In most cases, overnight seed cultures were diluted by 1:350 into 700 µl of fresh medium and grown in a 48-well plate in a microplate reader (Tecan, Spark). When the cultures grew to OD₆₀₀~0.4, we carried out another dilution of 1:50 into fresh medium and cultivated with the same condition. When subcultures reached an OD₆₀₀ of ~0.5, 10 µl gentamicin (Selleck, S4030) was supplemented to a final concentration of 6 µg/ml.

When gentamicin treatments were needed for many mutants (as in Fig. 6b), we adopted a 96-well plate to perform these experiments, while keeping all other settings as mentioned above. This protocol was also adopted to treat bacterial cultures by other antibiotics (as in Fig. 6f and Supplementary Fig. 14): Chloramphenicol (Solarbio, C8050); Lincomycin (TargetMol, T1277); Purromycin (Macklin, P816466); Azithromycin (Selleck, S1835); Solithromycin (Selleck, S5268); Streptomycin (inalco, 1758-9319);

Spectinomycin (Macklin, S6106); Tigecycline (Macklin, T830543); Tetracycline (Solarbio, ST8750); Ampicillin (Sigma, A9518); Ofloxacin (Solarbio, O8211); Trimethoprim (Solarbio, T9070).

Killing assay to determine drug lethality

At specific time points after drug (gentamicin) treatment, aliquots of 100 µl culture were taken from the well in microplate readers; We only performed one sampling from each well. Bacteria in the culture were immediately pelleted by centrifugation at 5000 × *g* for 2 min. Pellets were resuspended with fresh medium and properly diluted, then plated onto LB agar plates containing no antibiotics. The plates were incubated at 37 °C for 48 h before determining CFUs. The colonies recovered from varied lengths of gentamicin treatment were subjected to other experiments, including drug susceptibility assay and genome resequencing (described below).

Drug susceptibility assay by AUC of bacterial growth

We adopted a drug susceptibility assay to identify survival modes of strains upon drug treatment (see definitions below) with high sensitivity and throughput. Strains to be tested were randomly picked among survived clones from agar plates in a killing assay. Their overnight seed cultures were diluted 1:350 into 700 µl of fresh medium and grown in a 48-well plate in the microplate reader. When cultures grew to OD₆₀₀ ~0.4, they were diluted 1:15 into wells of a 96-well plate containing gentamicin of 0 µg/ml, 2 µg/ml, 3 µg/ml, 4 µg/ml, 6 µg/ml, 8 µg/ml, respectively. Cultivation was carried out in the microplate reader for 20 h, and OD₆₀₀ was monitored every 5 min. Next, each strain's AUCs under a drug of specific concentration were calculated from the growth curve (OD₆₀₀ × second).

TUNEL assay to quantify DNA damage

The chemistry of the TUNEL assay was described in Supplementary Fig. 2a. At specific time points after gentamicin treatment, aliquots of 100 µl culture were taken from the culture. Bacteria were collected by centrifugation at 700 × *g* for 2 min and washed once with PBS. The bacteria were then fixed with 3% PFA (Sigma-Aldrich, F8775) and gently rotated for 15 min at room temperature. After fixation, samples were pelleted by centrifugation at 1000 × *g* for 2 min and washed twice with PBS, then permeabilized with 70% ethanol (Sigma-Aldrich, 459844) at -20 °C overnight. The next day, permeabilized bacteria samples were pelleted by centrifugation at 1000 × *g* for 2 min and washed twice with water.

Labeling for DNA lesion was performed using the Click-iT TUNEL Alexa Fluor 488 kit (Thermo Fisher, C10245). For the TdT reaction, the pellets were resuspended with 50 µl systems, containing 1 µl EdUTP and 2 µl TdT in 1 × TdT reaction buffer. Then the reaction systems were incubated at 37 °C for 30 min. When the reactions finished, cells were pelleted and washed once with water. Next, the Click-iT reactions were performed by incubating in 100 µl systems, consisting of 2.5 µl additive in 1 × Click-iT reaction buffer, at room temperature for 30 min, protected from light. After the reaction cells were pelleted at 1000 × *g* for 2 min and washed once with PBS. As a negative control, we skipped the TdT reaction step.

Finally, labeled samples were resuspended in 0.5 ml PBS and analyzed by flow cytometry (BD, Aria II SORP). Propidium iodide staining (Thermo Fisher, P1304MP, 1 µg/ml supplemented prior to flow cytometry) was performed to quantify cell permeabilization. The optical setting of the machine is described as below. FITC (for DNA lesion labeling), 488 nm blue laser, plus a 505 nm long pass filter and a 530/30 nm band-pass filter; PE-Texas Red channel (for PI staining), a 561 nm green-to-yellow laser, plus a 600 nm long pass filter and a 610/20 nm band-pass filter. We gated populations showing PI signal (PE-Texas Red channel) above 10³ A.U. for the analysis of DNA lesion profile (FITC channel).

Slot-blot hybridization analysis of RNA:DNA hybrids (R-loops)

After different durations of gentamicin treatment, genomic DNA of bacteria from 2.5 ml culture was extracted with MasterPure Complete DNA/RNA Purification Kit (LGC (Lucigen), MC85200), and their concentrations were measured by Qubit 1xdsDNA HS Assay Kit (Thermo Fisher, Q33230). For each sample, 4 µg DNA was treated with 2 U ShortCut RNase III (NEB, M0245S) in 100 µl reaction system at 37 °C overnight, then purified with the abovementioned MasterPure Kit and titrated by Qubit again. After digestion, 200 ng DNA was slotted onto the Hybond N+ membrane (RPN303B, GE Healthcare). The membrane was cross-linked twice with 1200 µJ UV light. Following UV cross-linking, the membrane was blocked in TBS-Tween (0.05% Tween20 in 20 mM Tris pH 7.5, 150 mM NaCl) containing 5% milk (Sangon, A600669) at room temperature for one hour, and incubated at 4 °C overnight with anti-DNA:RNA hybrid S9.6 antibody (1:1000 dilution; abcam, ab234957). After 10 min washing in TBS-Tween for three times, the membrane was incubated for one hour at room temperature with HRP Goat Anti-Mouse IgG (1:5000 dilution; abclonal, AS003). Next, the membrane was washed by TBS-Tween for another three times and then developed with an enhanced chemiluminescence substrate (Easybio, BE6706) for Western Blotting. S9.6 signals were detected by ImageQuant LAS 4000 (GE Healthcare).

FACS to separate bacterial subpopulations with different SOS responses

To monitor the SOS responses, we constructed a *P_{recA}pepper* plasmid (see “Plasmid construction” section) to surveil the promoter activity of *recA*, the master regulator of SOS responses via triggering self-cleavage of LexA transcriptional repressor. Under ribosomal antibiotics, the typical reporters based on fluorescent proteins would be inhibited in translation. Here, we circumvented this problem by using a fluorescent RNA reporter instead named *pepper*, which allowed robust imaging of RNA species in living cells⁴⁶. Mechanistically, *pepper* RNAs became fluorescent when forming a complex with a small molecule HBC620 (TargetMol, T19352), which needs to be supplemented to medium. We experimentally confirmed that both externally added HBC620 and endogenously expressed *pepper* RNA were essential to the fluorescence signal, consistent with the report by Chen et al.⁴⁶. To maximize the sensitivity of this reporter, we used a *tolC* F16* (TTC→TGA) strain of *E. coli* K12 BW25113 as the host in this assay (see “Bacterial strains and cultivation conditions” section): this strain allowed the enhanced accumulation of intracellular HBC620. All these designs together enabled sensitive quantification of the SOS response with a high signal-to-noise ratio under translation inhibition. The SOS reporter strain used in this work was the *tolC* F16* (TTC→TGA) strain transformed with the *P_{recA}pepper* reporter plasmid. During cultivation of this reporter strain, ampicillin was always supplemented to a medium or agar plate to maintain the plasmid.

The drug treatment of this SOS reporter strain was similar to the above description (see “Drug treatment” section), except that 5 µM HBC620 needed to be supplemented to the medium during the second subculture. After gentamicin treatment for 40 min, 50 µl culture was diluted in 950 µl PBS for FACS. Non-cell particles were excluded via FSC-SSC gating. Cells were analyzed for their SOS responses in the PE-Texas Red channel (for optical setting, see “TUNEL assay” section). Among whole population, the 5% of cells with the strongest signals were gated as SOS-high group, while the ones ranked bottom 40% were assigned as SOS-low. For each group, 50,000 cells were sorted in a four-way setting and collected into 1.5 ml EP-tube containing 0.5 ml LB medium. After recovery in a shaker for 20 min at 37 °C, all cultures were plated onto LB agar in 6-well plates. After incubating at 37 °C for 24 h, CFU counting can be performed. The survival rate of each group was next calculated by dividing the CFU counts by 50,000, the original number of cells collected during FACS. We also showed that without gentamicin treatment, the SOS-high and SOS-low groups exhibited

comparable 100% survival rates after cell sorting (Fig. 6d), suggesting that the sorting process did not cause cell damages or group-wise biases in this assay.

Genome resequencing

TIANamp Bacteria DNA kit (TIANGEN, DP302) was used to extract genomic DNA (gDNA) from strains. Paired-end libraries were prepared by a Tn5-based library construction method⁵⁸. Tn5 transposome was assembled by transposase and adapters according to the manufacturer's instruction (Robust Tn5 Transposase, Robustnique Corporation Ltd., Tianjin, China). For tagmentation, 12 ng gDNA was tagmented in a 20 µl reaction containing 1 × Tn5 transposome reaction buffer (LM buffer) and 1 µl Tn5 transposome at 55 °C for 15 min. The lengths of tagmented gDNA libraries were between 250 and 600 bp (a pre-test to titrate gDNA input amount is highly recommended to generate library of optimal lengths). To generate Illumina-compatible sequencing libraries, the tagmented gDNA libraries were amplified with primers containing P5 or P7 adapters as well as specific indexes. In 25 µl PCR reaction system (Vazyme, TD601), 8 µl of tagmentation product was input as the template. The PCR program was as follows: (i) 72 °C for 3 min; (ii) 98 °C for 30 s; (iii) 98 °C for 15 s; 60 °C for 30 s; 72 °C for 3 min; 8 cycles; (iv) 72 °C for 5 min. Finally, the indexed gDNA libraries were pooled together and purified with 1.0 × AMPure XP beads (Beckman Coulter). All oligonucleotides used in genome resequencing library preparation could be found at Supplementary Data 3.

After library construction, we subjected the samples to sequencing by Novogene Bioinformatics Institute via Illumina NovaSeq 6000 system with 2 × 150 pair-end configuration. Typically, we got three million pair-end reads for each sample, corresponding to ~50-fold genome coverage. After demultiplexing, we trimmed adapter sequences by cutadapt (v0.11.9) and adopted breseq (0.35.3)⁵⁹ to identify mutations from the first read of the NGS data in a default clonal mode against the reference genome (NC000913.3). Those identified mutations carried by our wild-type strain were regarded as spontaneous mutations. Note that we excluded mutations derived from genetic manipulations during the construction history of *E. coli* K12 MG1655, such as deletions of several sugar utilization operons⁵⁴. Except for spontaneous mutations, any other genetic alternation identified in the survived isolates was regarded as mutations that emerged after gentamicin treatment and were then subjected to mutation spectrum analysis.

Pulse-field gel electrophoresis

We adopted a protocol reported previously⁶⁰. Around 10⁸ *E. coli* cells were collected at a series of time points after gentamicin treatment from 48-well plate. After pelleted by a soft spin of 700 × *g* for 2 min (to minimize DNA damage caused by centrifugation), bacteria were resuspended by 100 µl LB and 1:1 thoroughly mixed with 2% low melting point agarose (Promega, V2111) preheated to 42 °C, and then molded into plugs (Bio-Rad, 1703713). When solidified, the agarose plugs were treated with lysozyme for two hours at 37 °C. The lysozyme solution contained 1 mg/ml lysozyme (Sigma, 62970), 10 mM Tris pH 7.0 (Solarbio, T1130), 50 mM NaCl (Thermo Fisher, AM9760G), 0.2% sodium deoxycholate (Sigma, 30970) and 0.5% sodium lauryl sarcosine (Sigma, 61745). Next, the plugs were rinsed with sterile water and incubated in Proteinase K solution for 12 h at 50 °C with mild agitation. The Proteinase K solution contained 1 mg/ml Proteinase K (Sigma, 539480), 100 mM EDTA (Solarbio, E1170), 0.2% sodium deoxycholate, 1% sodium lauryl sarcosine. After digestion, the plugs were washed for six times (10 min each with mild agitation) in wash buffer, containing 20 mM Tris pH 8.0 (Solarbio, T1150), 50 mM EDTA. The plugs were inserted into wells of 1% agarose gel in 0.5 × TBE (huaxingbio, HX1872) and sealed by agarose solution. Genomic DNA fragments were then separated by Bio-Rad CHEF Mapper® pulsed-field electrophoresis system. Running conditions were generated by the Auto-Algorithm

mode using a size range of 1–100 kb: 0.5 × TBE at 14 °C for 07:53 h at 6.0 V/cm using a 120° included angle with a 0.06–8.53 s linear switch time ramp. Gels were stained with Genecolor I (Gene-bio, GBY-1) overnight and visualized with UV-trans-illumination. The intensities of DNA bands were quantified using ImageJ software.

Total RNA extraction

At specific time points before or after drug treatment, total RNA was extracted as below. 150 µl culture was mixed with 150 µl RNeasy lysis buffer (Thermo Fisher, AM7020) and incubated at room temperature for at least 15 min. Then, the sample was diluted with 600 µl ice-cold PBS (SolarBio, P1020) and centrifuged at 5000 × *g* for 15 min. The obtained cell pellet was resuspended by 50 µl TE buffer (Thermo Fisher, AM9849) containing 12.5 mg/ml of lysozyme and 16.7% (v:v) proteinase K (QIAGEN, 19131) to lyse cells. The mixture was incubated at 37 °C for 15 min and subjected to purification. 100 µl of RNeasy spin columns (Beckman, A63987) was added to the lysate and incubated at room temperature for 15 min. Next, the tube was placed on a magnetic rack for 10 min, then the supernatant was discarded. The beads were washed twice by 200 µl fresh 80% (v:v) ethanol (Sigma-Aldrich, 459844) and air-dried for 15 min on the magnetic rack. Finally, the beads were thoroughly resuspended by 20 µl nuclease-free water (Thermo Fisher, AM9932) and incubated at room temperature for 15 min. After incubation, place the tube on a magnetic rack for 5 min. Finally, purified total RNA was obtained from the supernatant. The concentration of total RNA was measured by Qubit RNA HS Assay Kit (Thermo Fisher, Q32852) and stored at –80 °C before use.

3′-end-seq and RNA-seq library construction

For each purified sample, total RNA was subjected to 3′-end-seq and RNA-seq library preparation respectively.

The 3′-end-seq library preparation method was adopted from a multiplexed RNA-seq library preparation method previously reported by Avraham et al.⁶¹. Systematic modifications were made to repurpose it to keep 3′ end information of bacterial transcripts. First, without any fragmentation, 100 ng total RNA was dephosphorylated by FastAP (Thermo Fisher, EF0651) and depleted of genomic DNA by Turbo DNase (Thermo Fisher, AM2238). Then, RNA adapters were ligated to 3′ ends of total bacterial RNA by T4 RNA ligase I (NEB, M0437). Each RNA adapter carried a 5′ phosphate, a 3′ C3 blocking group, a 3′ constant sequence containing an i5 adapter used for priming in reverse transcription, and a unique 5′ barcode for multiplexing. Next, four to five samples were pooled into together and purified by the RNeasy column (QIAGEN, 74104). Purified RNA was subjected to an RNase H-based (NEB, M0523) reaction to deplete rRNA and the noncoding element *ssrA*⁶². RNA was purified by RNeasy spin columns, and then converted to cDNA by reverse transcription (NEB, M0466) with a primer targeting the constant sequence in the ligated adapter. Then, second strand cDNA was synthesized by template switch using a specific oligo, and pre-amplification by KAPA DNA polymerase (Roche, KK2602) for 8 cycles. The DNA was purified by 1.0 × AMPure XP beads. Finally, 2 ng purified DNA was tagged with N7 adapter-incubating transposase (Vazyme, TD513) and followed by 11 cycles of PCR amplification yielding around 100 ng DNA product. It should be noted that this final PCR step led to a biased sampling towards the 3′ ends of bacterial transcripts and completely abandoned the 5′-end information. All oligonucleotides used in 3′-end-seq library preparation could be found in Supplementary Data 4 (all ordered from GenScript, RNase-free HPLC purification).

When preparing the RNA-seq library, in order to generate comparable data to 3′-end-seq libraries, we applied a customized RNA-seq library construction method derived from the 3′-end-seq. To be specific, the protocol was generally the same with the 3′-end-seq only with two critical modifications. First, 100 ng total RNA of the same sample was firstly fragmented at 97 °C in FastAP buffer (Thermo Fisher,

EF0651) for 3 min. Second, the pre-PCR amplified double-strand DNA after the template switch was directly subjected to final PCR without fragmentation by transposases. These two modifications can help to eliminate the 3′-end-biased sampling during 3′-end-seq library construction, generating more uniform and unbiased RNA-seq data for transcript counting and coverage analysis. All oligonucleotides stated above were adopted from the 3′-end-seq protocol and could also be found in Supplementary Data 4.

All libraries were sequenced by GENEWIZ Int. Co. via Illumina NovaSeq 6000 system with 2 × 150 pair-end configuration. About five million pair-end reads were yielded for each demultiplexed library.

NGS data processing for 3′-end-seq

Firstly, sequencing raw data was subjected to FastQC (v0.11.9) for quality control and cutadapt (2.10) for adapter trimming. The data was then demultiplexed by cutadapt based on the unique RNA adapters. Paired-end reads were aligned to *E. coli* reference genome (NC_000913.3) by bowtie2 (2.4.1) with parameters: -X 1000 -I 18 --no-mixed --no-discordant. The mapped reads were converted to bam files by samtools (1.10). To minimize the effect of sequencing depth on further analysis, we subsampled every bam file to the equal number of 4.5 million paired-end reads. Reads of bam files were counted using R (4.0.2) method summarizeOverlaps from GenomicAlignments package (v1.26.0) in a strand-specific manner with the following options: “mode = “Union”, ignore.strand = FALSE, single-End = FALSE, preprocess.reads = invertStrand”.

As for further statistical analysis and visualization of 3′ end, we used a customized R script to process bam files. Due to the strand-ness, the 5′ ends from the first read in NGS raw data were actual 3′ end positions of RNA transcripts. Based on this principle, the termini signals were counted with single-nucleotide resolutions and normalized by Count Per Million (CPM) summed from all 3′ ends, and this was regarded as 3′ end intensity and used in further analysis and plotting.

NGS data processing for RNA-seq

RNA-seq data processing was similar to our previous report⁶². Briefly, adapter trimming, alignment to genome and gene read count table were all the same with 3′-end-seq as shown above. RNA-seq raw count was normalized by transcripts per million (TPM) as a quantification of RNA expression, then a series of TPM matrixes over drug treatment times were derived. As for data quality control, genes with average TPM < 5 or maximum TPM < 10 were excluded from further analysis, resulting in 3173 genes remained (3126 protein-coding genes plus 47 noncoding genes, Supplementary Data 5). Genome-wide strand-specific coverage files were also generated from RNA-seq data for each bam file by bedtools (v2.30.0) genomecov, with options ‘-strand -/+ -d -pc’, which was used in the gene expression polarity analysis.

Calculation of 5′ coverage bias

Here we explain the calculation of 5′ coverage bias to quantify the 5′ and 3′ transcript coverage bias upon gentamicin treatment. Specifically, for each gene, we first trimmed 10% from the 5′ and 3′ ends to avoid coverage changes caused by interference from neighboring genes. We then divided the remaining region equally into 5′ and 3′ halves and calculated the average coverage of the two halves. Finally, we calculated the ratio of the average 5′ coverage to the average 3′ coverage, and defined this as the 5′ coverage bias.

Technical definition of intrinsic terminator

Here we present the identification of intrinsic terminators from 3′-end-seq data. Briefly, the 3′ ends occurred 100 nt downstream the stop codon were taken into intrinsic terminator candidates. Those candidates were filtered by thresholds in CPM values (>1) and enrichment factor (>20) of the 3′ ends (see “Dominant 3′ termini calling” section for the definitions of enrichment factor). Two sets of candidates were

generated from two replicates, and their intersection was the final set of intrinsic terminators. During the intersection, we took candidates from replicate one if they fall into the ± 10 nt range of candidates in replicate two. Finally, the intersected candidates were defined as intrinsic terminators. The quality of this computational pipeline was verified by the high accuracy to capture known intrinsic terminators (Supplementary Fig. 4). All intrinsic terminators (at 0 min of drug treatment) defined here were excluded from 3' termini analyses as below.

Technical definition of 3' termini

Here we present the quantification of 3' termini within protein-coding regions (i.e., excluding intrinsic terminators and 3' ends in untranslated regions) (Supplementary Fig. 5). After NGS data processing, we obtained CPM values of 3' end at each nucleotide on both strands of bacterial genome. All analysis should take strand specificity into consideration. All 3' ends that occurred in gene-coding regions were taken as 3' termini candidates. As bacterial genes are proximal and even overlapped with each other, some candidates might actually derive from the intrinsic termination events of upstream genes. Thus, we first excluded intrinsic termination events identified by the computational pipeline described above, and next eliminated other candidates flanking (5-nt range) each intrinsic terminator to further reduce noise. The remaining candidates were finally defined as 3' termini and subjected to following analysis or calculation. A comprehensive list of 3' termini for all samples produced in our work was deposited to GEO (accession number: GSE240474, see also in "Data availability" section). In our work, 3' termini were further regarded as potential premature RNAP stalling events.

Dominant 3' termini calling

Here we clarify the definition of dominant 3' termini (Supplementary Fig. 17). Briefly, dominant 3' termini are defined as 3' termini surpassing a customized intensity cutoff, which is called enrichment factor. The enrichment factor was defined to distinguish 3' termini from the background noise in its flanking region, its calculation method was as below: we first found the nucleotide and the strand where this 3' terminus located at. Then we calculated the average of CPM of the nucleotide and two flanking nucleotides with strand specificity, and the CPM average was referred as half-peak height of this 3' terminus. The corresponding background noise was calculated by averaging the CPMs of 3' ends flanking (50-nt regions each side) those above-mentioned three nucleotides. Given these two parameters, we finally obtained the enrichment factor for this 3' terminus by dividing the half-peak height by the background noise. Note that the statistic "enrichment" was also used in intrinsic termination identification as described above.

After calculating enrichment factors for all 3' termini, dominant 3' termini were filtered according to different thresholds, such as 10, 15, 20, 25 and 50 in our work. Note that if more than one dominant 3' termini occurred within 10 nt, we only kept the one with biggest CPM. If there were more than one values kept, we finally select the one with a higher enrichment factor as the dominant 3' terminus.

Calculation of fold change of 3' termini in each gene

Here we explain the calculation of normalized fold change of total 3' termini in each protein-coding gene upon gentamicin treatment, simplified as fold change of 3' termini. This static excludes the effect exerted by transcription level to 3' termini by normalization: before calculate fold changes, the 3' termini CPM values were first normalized by the transcription level of the corresponding gene (TPM). The calculation process was as below:

First, for one particular gene ($gene_j$), we defined a transcription level normalization factor (equation 1), to control the gene expression

change before and after drug treatment.

$$\text{Normalization factor}_j = \frac{\min(\text{TPM}_{j,\text{treated}}, \text{TPM}_{j,\text{untreated}})}{\max(\text{TPM}_{j,\text{treated}}, \text{TPM}_{j,\text{untreated}})} \quad (1)$$

Next, the raw count of each 3' terminus ($3' \text{ terminus}_k$) associated with this gene was normalized by this factor (equation 2); and therefore 3' end count was reset to a new level assuming the same transcription level before and after drug treatment.

$$\begin{aligned} \text{End count}'_{j,k,\text{treated}} &= (\text{End count}_{j,k,\text{treated}} \times \text{Normalization factor}_j, \\ &\quad \text{if } \text{TPM}_{j,\text{treated}} = \max(\text{TPM}_{j,\text{treated}}, \text{TPM}_{i,\text{untreated}})) \\ &\quad \text{or } (\text{End count}_{j,k,\text{treated}}, \text{if } \text{TPM}_{j,\text{treated}} \\ &\quad = \min(\text{TPM}_{j,\text{treated}}, \text{TPM}_{i,\text{untreated}})) \end{aligned} \quad (2)$$

This normalization might lose some low abundance read count, due to the bottleneck effect during NGS library preparation and sequencing. To address this, we adopted a Poisson sampling simulation to mimic such bottleneck effect, so as to recover some potential loss. The end count' from equation 2 was further processed by equation 3, where end count' ≤ 2 was subjected to the Poisson sampling process.

$$\begin{aligned} \text{End count}''_{j,k,\text{treated}} &= (\text{Poisson sampling}(\lambda = \text{End count}'_{j,k,\text{treated}}), \\ &\quad \text{if } \text{End count}'_{j,k,\text{treated}} \leq 2) \text{ or } (\text{End count}'_{j,k,\text{treated}}, \\ &\quad \text{if } \text{End count}'_{j,k,\text{treated}} > 2) \end{aligned} \quad (3)$$

After these steps, the normalized end count value was converted to the CPM value to normalize the effect from sequencing depth by equation 4.

$$\text{End CPM}''_{j,k,\text{treated}} = (\text{End Count}''_{j,k,\text{treated}} \times 10^6) / \text{total depth}_{\text{treated}} \text{ in } 3' \text{-end-seq library} \quad (4)$$

Finally, normalized total 3' termini of one gene from one sample (such as drug treatment) was calculated by summing up all normalized CPM values of associated 3' termini by equation 5.

$$\text{Normalized total } 3' \text{ termini}_{j,\text{treated}} = \text{sum}(\text{End CPM}''_{j,k,\text{treated}}) \quad (5)$$

To calculate the fold change of 3' termini of one gene upon drug treatment, we used equation 6.

$$\text{Fold change of } 3' \text{ termini}_{j,\text{treated}} = (\text{Normalized total } 3' \text{ termini}_{j,\text{treated}}) / (\text{Normalized total } 3' \text{ termini}_{j,\text{treated},\text{untreated}}) \quad (6)$$

To quantify whether such change was significant or not, we assessed the technical noise of this process (from 3'-end-seq library construction, sequencing to computation) via calculating the fold change of 3' termini of one gene between two replicates without drug treatment. In particular, we replaced all "treated" terms mentioned above with the corresponding term from another "untreated" replicate, resulting in "Fold change of 3' termini_{j, untreated}". As expected, the fold change of 3' termini between two drug-free replicates from all genes followed a normal distribution centered around zero (Fig. 3e, "0 min"). In our work, "fold change of 3' termini" was regarded as a potential indicator of RNAP processivity in each gene.

RNA-seq coverage normalization to show genome-wide gene expression polarity

To quantify gene expression polarity at genome scale from RNA-seq data, we adopted a computational process to normalize RNA-seq coverage across all genes⁶². RNA-seq coverage per nucleotide was first calculated by bedtools in a strand-specific manner (see “NGS data processing” section). For noncoding genes, 5% length from 3′ was trimmed (according to annotations in NC_000913.3) in this analysis to eliminate noise caused by intrinsic transcription termination; as for protein-coding genes, whole coding sequences were kept as intrinsic transcription termination occurred outside of them. The length of each gene was rescaled to 0 to 1, then 101 equally spaced points were generated along transcription direction. For each gene, the value of each point was computed from the read coverages of the two nearest nucleotides via linear interpolation. Then, values of each point were averaged across all corresponding gene, and finally the averaged 101 values were normalized by dividing their maximum value (Fig. 5a, b).

Analysis of PSREs on mutation genes

We assumed that gentamicin-induced mutations are prone to reside on PRSE-prone genes, but spontaneous mutations are irrelevant to PRSE. To test this hypothesis, we focused on coding genes whose coding regions harboring mutations (while mobile elements are excluded for being inherently unstable): there are 21 genes harboring 29 gentamicin-induced mutations (out of 35), and 8 genes harboring spontaneous mutations (out of 24). Next, we calculated PRSEs frequency in the coding regions of these genes. To this end, for each gene we summed up the CPM of 3′ termini and normalized by the length of coding sequence (equation 7). Therefore, the obtained metric reflects the averaged 3′ termini signal per nucleotide, so we termed it as 3′ termini per nt (Equation 7).

$$3' \text{ termini per nt} = \text{sum}(\text{End CPM of } 3' \text{ termini}) / \text{gene length} \quad (7)$$

PRSE profile rescaling in a gene-wise manner

To visualize PRSEs of each gene before and after drug treatment in an unbiased manner (Fig. 5c–f), we adopted a computational approach to rescale the PRSE profile for each gene across different conditions. For each gene, the PRSE profile was defined as the vector of the CPM values of the 3′ end at each nucleotide along coding region of this gene. As mentioned in the “RNA-seq coverage normalization to show genome-wide gene expression polarity” section, for noncoding genes, nucleotides from the last 5% length proximal to their 3′ ends were trimmed.

We plotted two PRSE profiles at 0 min and 60 min after gentamicin treatment. For each gene, we first processed its PRSE profile at 60 min by dividing gene expression change between these two conditions ($\text{TPM}_{60 \text{ min}} / \text{TPM}_{0 \text{ min}}$ in RNA-seq data); the corresponding PRSE profile at 0 min was kept constant. Therefore, we removed the undesired impact of gene expression change on PRSE profile. Next, for each gene, its nucleotide sequence was equally divided into 50 segments along the transcription direction; within each segment, the processed CPM values of all nucleotides were summed. Thus, we normalized lengths of different genes and converted to new profiles with uniformly 50 elements. Finally, for each gene, the processed PRSE profiles of the two samples (e.g., 0 min and 60 min) were rescaled to [0, 1] by dividing the maximum value among all 100 values. The rescaled PRSE profiles were used for visualization in Fig. 5c–f.

RNA structure and sequence analysis

For RNA secondary structure analysis, RNA sequence of 50 nt upstream of the location of each dominant 3′ terminus (PRSE) was extracted from *E. coli* genome NC000913.3 and subjected to RNAfold

command line tool (2.4.18)⁶³ with parameters “--noPS --noconv -C” to predict minimal free energy (MFE) of RNA folding. For RNA sequence analysis, a home-made Python script was adopted to calculate and plot the motif logo of 10 nt sequences upstream the dominant 3′ terminus (PRSE). Here we defined the nucleotide positions of dominant PRSEs or intrinsic termination sites as +1 (Fig. 5h–j).

Statistics & reproducibility

In this study, statistical significance (*p*-value) was determined by *t*-test, Mann–Whitney U test, or Kolmogorov–Smirnov test. We reported the statistical test used, *p*-value, and sample size (*n*) for each test in the associated figure legends. Low-quality reads were excluded from the analysis of RNA-seq, 3′-end-seq and ChIP-seq data (see “NGS processing” sections in “Methods” and “Supplementary Methods” sections). Since we found that even slight differences in sequencing depth could affect the 3′-end-seq analysis, we sampled the reads from 3′-end-seq data to normalize all BAM files to the same read count, trimming only 2–4% of reads. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The raw data and processed data of RNA-seq, 3′-end-seq and ChIP-seq data generated in this study have been deposited in GEO (RNA-seq and 3′-end-seq data under accession code [GSE240474](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE240474); ChIP-seq data under the accession code [GSE267245](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE267245)). The raw data of genome resequencing have been deposited in SRA under accession code [SRP455999](https://www.ncbi.nlm.nih.gov/sra/acc.cgi?acc=SRP455999). Lists of all 3′ termini position and their abundance (CPM) upon gentamicin exposure in wild-type and *Δmfd* strains are available. These lists have been compiled as supplementary files and can be accessed under the aforementioned accession code GSE240474 in GEO. A comprehensive list of dominant 3′ termini for all samples can be found in Supplementary Data 6. A comprehensive list of fold changes in 3′ termini for each gene, comparing all drug-treated samples to untreated samples (0 min) in both wild-type and *Δmfd* strains, can be found in Supplementary Data 7. A list of Gam peak locations generated from ChIP-seq data is available in Supplementary Data 8. Source data are provided with this paper.

Code availability

The custom scripts for NGS data processing, including RNA-seq, 3′-end-seq and ChIP-seq, are freely available (https://github.com/RussellChai/Gentamicin_NGS).

References

- Baquero, F. & Levin, B. R. Proximate and ultimate causes of the bactericidal action of antibiotics. *Nat. Rev. Microbiol.* **19**, 123–132 (2021).
- Pribis, J. P. et al. Gamblers: an antibiotic-induced evolvable cell subpopulation differentiated by reactive-oxygen-induced general stress response. *Mol. Cell* **74**, 785–800.e7 (2019).
- Drlica, K., Malik, M., Kerns, R. J. & Zhao, X. Quinolone-mediated bacterial death. *Antimicrob. Agents Chemother.* **52**, 385–392 (2008).
- Baharoglu, Z., Babosan, A. & Mazel, D. Identification of genes involved in low aminoglycoside-induced SOS response in *Vibrio cholerae*: a role for transcription stalling and Mfd helicase. *Nucleic Acids Res.* **42**, 2366–2379 (2014).
- Baharoglu, Z., Krin, E. & Mazel, D. RpoS plays a central role in the SOS induction by sub-lethal aminoglycoside concentrations in *Vibrio cholerae*. *PLoS Genet.* **9**, e1003421 (2013).

6. Khodursky, A., Guzmán, E. C. & Hanawalt, P. C. Thymineless death lives on: new insights into a classic phenomenon. *Annu. Rev. Microbiol.* **69**, 247–263 (2015).
7. Cooke, M. S., Evans, M. D., Dizdaroğlu, M. & Lunec, J. Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J. Publ. Fed. Am. Soc. Exp. Biol.* **17**, 1195–1214 (2003).
8. Dutta, D., Shatalin, K., Epshtein, V., Gottesman, M. E. & Nudler, E. Linking RNA polymerase backtracking to genome instability in *E. coli*. *Cell* **146**, 533–543 (2011).
9. Rinaldi, C. et al. Damage to safeguard genome stability. *Front. cell Dev. Biol.* **8**, 618157 (2020).
10. Negro, V. et al. RadD contributes to R-Loop avoidance in sub-MIC tobramycin. *MBio* **10**, e01173–19 (2019).
11. Wimberly, H. et al. R-loops and nicks initiate DNA breakage and genome instability in non-growing *Escherichia coli*. *Nat. Commun.* **4**, 2115 (2013).
12. Paulsen, R. D. et al. A genome-wide siRNA screen reveals diverse cellular processes and pathways that mediate genome stability. *Mol. Cell* **35**, 228–239 (2009).
13. Xia, J. et al. Bacteria-to-human protein networks reveal origins of endogenous DNA damage. *Cell* **176**, 127–143.e24 (2019).
14. Tubbs, A. & Nussenzweig, A. Endogenous DNA damage as a source of genomic instability in cancer. *Cell* **168**, 644–656 (2017).
15. Gruber, C. C. & Walker, G. C. Incomplete base excision repair contributes to cell death from antibiotics and other stresses. *DNA Repair* **71**, 108–117 (2018).
16. Foti, J. J., Devadoss, B., Winkler, J., Collins, J. & Walker, G. Oxidation of the guanine nucleotide pool underlies cell death by bactericidal antibiotics. *Science* **336**, 315–319 (2012).
17. Belenky, P. et al. Bactericidal antibiotics induce toxic metabolic perturbations that lead to cellular damage. *Cell Rep.* **13**, 968–980 (2015).
18. Kohanski, M. A., Dwyer, D. J., Hayete, B., Lawrence, C. A. & Collins, J. J. A common mechanism of cellular death induced by bactericidal antibiotics. *Cell* **130**, 797–810 (2007).
19. Proshkin, S., Rahmouni, A. R., Mironov, A. & Nudler, E. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* **328**, 504–508 (2010).
20. Johnson, G. E., Lallane, J. B., Peters, M. L. & Li, G. W. Functionally uncoupled transcription–translation in *Bacillus subtilis*. *Nature* **585**, 124–128 (2020).
21. Iyer, S., Le, D., Park, B. R. & Kim, M. Distinct mechanisms coordinate transcription and translation under carbon and nitrogen starvation in *Escherichia coli*. *Nat. Microbiol.* **3**, 741–748 (2018).
22. Blaha, G. M. & Wade, J. T. Transcription-translation coupling in bacteria. *Annu. Rev. Genet.* **56**, 187–205 (2022).
23. Kohanski, M. A., Dwyer, D. J., Wierzbowski, J., Cottarel, G. & Collins, J. J. Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. *Cell* **135**, 679–690 (2008).
24. Brauner, A., Fridman, O., Gefen, O. & Balaban, N. Q. Distinguishing between resistance, tolerance and persistence to antibiotic treatment. *Nat. Rev. Microbiol.* **14**, 320–330 (2016).
25. Levin-Reisman, I. et al. Antibiotic tolerance facilitates the evolution of resistance. *Science* **355**, 826–830 (2017).
26. Ibacache-Quiroga, C., Oliveros, J. C., Couce, A. & Blázquez, J. Parallel evolution of high-level aminoglycoside resistance in *Escherichia coli* under low and high mutation supply rates. *Front. Microbiol.* **9**, 1–14 (2018).
27. Girgis, H. S., Hottes, A. K. & Tavazoie, S. Genetic architecture of intrinsic antibiotic susceptibility. *PLoS ONE* **4**, e5629 (2009).
28. Guerin, É. et al. The SOS response controls integron recombination. *Science* **324**, 1034 (2009).
29. Siguier, P., Gourbeyre, E. & Chandler, M. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol. Rev.* **38**, 865–891 (2014).
30. Wagner, J. & Nohmi, T. *Escherichia coli* DNA polymerase IV mutator activity: genetic requirements and mutational specificity. *J. Bacteriol.* **182**, 4587–4595 (2000).
31. Niccum, B. A. et al. New complexities of SOS-induced “untargeted” mutagenesis in *Escherichia coli* as revealed by mutation accumulation and whole-genome sequencing. *DNA Repair* **90**, 102852 (2020).
32. Lewis, E. B., Mudipalli, R., Eghbal, M. M. & Culyba, M. J. Effect of mismatch repair on the mutational footprint of the bacterial SOS mutator activity. *DNA Repair* **103**, 103130 (2021).
33. Larson, M. H. et al. A pause sequence enriched at translation start sites drives transcription dynamics in vivo. *Science* **344**, 1042–1047 (2014).
34. Vvedenskaya, I. O. et al. Interactions between RNA polymerase and the ‘core recognition element’ counteract pausing. *Science* **344**, 1285–1289 (2014).
35. Lindsey-Boltz, L. A. & Sancar, A. The transcription-repair coupling factor Mfd prevents and promotes mutagenesis in a context-dependent manner. *Front. Mol. Biosci.* **8**, 1–6 (2021).
36. Ragheb, M. N., Merrikh, C., Browning, K. & Merrikh, H. Mfd regulates RNA polymerase association with hard-to-transcribe regions in vivo, especially those with structured RNAs. *Proc. Natl Acad. Sci. USA* **118**, e2008498118 (2020).
37. Leela, J. K., Syeda, A. H., Anupama, K. & Gowrishankar, J. Rho-dependent transcription termination is essential to prevent excessive genome-wide R-loops in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **110**, 258–263 (2013).
38. Bharati, B. K. et al. Crucial role and mechanism of transcription-coupled DNA repair in bacteria. *Nature* **604**, 152–159 (2022).
39. Epshtein, V. et al. UvrD facilitates DNA repair by pulling RNA polymerase backwards. *Nature* **505**, 372–377 (2014).
40. Agapov, A., Olina, A. & Kulbachinskiy, A. RNA polymerase pausing, stalling and bypass during transcription of damaged DNA: from molecular basis to functional consequences. *Nucleic Acids Res.* **50**, 3018–3041 (2022).
41. Wee, L. M. et al. A trailing ribosome speeds up RNA polymerase at the expense of transcript fidelity via force and allostery. *Cell* **186**, 1244–1262.e34 (2023).
42. Wang, T. et al. Dynamics of transcription–translation coordination tune bacterial indole signaling. *Nat. Chem. Biol.* **16**, 440–449 (2020).
43. Kang, J. Y. et al. RNA polymerase accommodates a pause RNA hairpin by global conformational rearrangements that prolong pausing. *Mol. Cell* **69**, 802–815.e1 (2018).
44. Artsimovitch, I. & Landick, R. Interaction of a nascent RNA structure with RNA polymerase is required for hairpin-dependent transcriptional pausing but not for transcript release. *Genes Dev.* **12**, 3110–3122 (1998).
45. Lin, J., Zhou, D., Steitz, T. A., Polikanov, Y. S. & Gagnon, M. G. Ribosome-targeting antibiotics: modes of action, mechanisms of resistance, and implications for drug design. *Annu. Rev. Biochem.* **87**, 451–478 (2018).
46. Chen, X. et al. Visualizing RNA dynamics in live cells with bright and stable fluorescent RNAs. *Nat. Biotechnol.* **37**, 1287–1293 (2019).
47. Rosenberg, S. M. Evolving responsively: adaptive mutation. *Nat. Rev. Genet.* **2**, 504–515 (2001).
48. Babosan, A. et al. A qnr-plasmid allows aminoglycosides to induce SOS in *Escherichia coli*. *Elife* **11**, e69511 (2022).
49. Baharoglu, Z. & Mazel, D. *Vibrio cholerae* triggers SOS and mutagenesis in response to a wide range of antibiotics: a route towards multiresistance. *Antimicrob. Agents Chemother.* **55**, 2438–2441 (2011).

50. Sollier, J. et al. Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol. Cell* **56**, 777–785 (2014).
51. Branum, M. E., Reardon, J. T. & Sancar, A. DNA repair excision nuclease attacks undamaged DNA: a potential source of spontaneous mutations. *J. Biol. Chem.* **276**, 25421–25426 (2001).
52. Ragheb, M. N. et al. Inhibiting the evolution of antibiotic resistance. *Mol. Cell* **73**, 157–165.e5 (2019).
53. Carvajal-Garcia, J., Samadpour, A. N., Hernandez Viera, A. J. & Merrih, H. Oxidative stress drives mutagenesis through transcription-coupled repair in bacteria. *Proc. Natl Acad. Sci. USA* **120**, e2300761120 (2023).
54. Baba, T. et al. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006–0008 (2006).
55. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97**, 6640–6645 (2000).
56. Jiang, Y. et al. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl. Environ. Microbiol.* **81**, 2506–2514 (2015).
57. Wang, T., Shen, P., He, Y., Zhang, Y. & Liu, J. Spatial transcriptome uncovers rich coordination of metabolism in *E. coli* K12 biofilm. *Nat. Chem. Biol.* **19**, 940–950 (2023).
58. Wu, J., Dai, W., Wu, L. & Wang, J. SALP, a new single-stranded DNA library preparation method especially useful for the high-throughput characterization of chromatin openness states. *BMC Genom.* **19**, 143 (2018).
59. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* **1151**, 165–188 (2014).
60. Gunderson, K. & Chu, G. Pulsed-field electrophoresis of megabase-sized DNA. *Mol. Cell. Biol.* **11**, 3348–3354 (1991).
61. Avraham, R. et al. A highly multiplexed and sensitive RNA-seq protocol for simultaneous analysis of host and pathogen transcriptomes. *Nat. Protoc.* **11**, 1477–1491 (2016).
62. Wang, T., Shen, P., Chai, R., He, Y. & Liu, J. Profiling of bacterial transcriptome from ultra-low input with MiniBac-seq. *Environ. Microbiol.* **24**, 5774–5787 (2022).
63. Lorenz, R. et al. ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 26 (2011).

Acknowledgements

We thank Babak Javid, Mo Chen, and Hongwei Su for their comments during the writing of the manuscript. Keyong Sun for providing protocol of genome resequencing; Chengxia Zheng and Weifeng Zhang for their technical suggestion on hybridization assay of R-loop and PFGE; Yiwen Zhao for assistance with strain construction and proteomics experiments. We thank Haiteng Deng and Xianbin Meng of the Proteinomics Facility at Tsinghua University Technology Center for Protein Sciences for assistance in protein mass spectrometry experiments and analysis. J.L. was supported by the National Key R&D Program of China (2023YFC2306300), the National Natural Science Foundation of China (32170099), the Tsinghua University Initiative Scientific Research Program (20221080020, 20231080040), and the Tsinghua-Peking Center for Life Sciences. T.W. was supported by the National Natural Science

Foundation of China (32100071), the Chinese Postdoctoral Science Foundation (BX20190176 and 2020M670288), and the Tsinghua-Peking Center for Life Sciences.

Author contributions

T.W., Y.Z., R.C., and J.L. designed the research. Y.Z. performed genetics, biochemistry, cell sorting, genome resequencing, ChIP-seq, and proteomics experiments. Z.X. and T.W. helped with some of the experiments. R.C. established the 3'-end-seq method and profiled transcription dynamics. Y.H., Z.X., and Y.Z. were involved in RNA-seq library construction. P.S. helped with the establishment of the 3'-end-seq method. R.C. and T.W. performed computational analysis of the 3'-end-seq, ChIP-seq, and other transcriptomic data. Y.Z. and T.W. analyzed data from other experiments. P.S. helped with some of the analysis. Y.Z., R.C., T.W., and J.L. wrote the manuscript. Y.Z., R.C., and J.L. revised the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-50917-6>.

Correspondence and requests for materials should be addressed to Tianmin Wang or Jintao Liu.

Peer review information *Nature Communications* thanks Shan Yan, who co-reviewed with Jia Li; Houra Merrih for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024