

# A Cre Transcription Fidelity Reporter Identifies GreA as a Major RNA Proofreading Factor in *Escherichia coli*

Mikhail G. Bubunenko, Carolyn B. Court, Alison J. Rattray, Deanna R. Gotte, Maria L. Kireeva, Jorge A. Irizarry-Caro, Xintian Li, Ding J. Jin,<sup>1</sup> Donald L. Court,<sup>1</sup> Jeffrey N. Strathern,<sup>1</sup> and Mikhail Kashlev<sup>1</sup>

The RNA Biology Laboratory, National Cancer Institute-Frederick, Center for Cancer Research, Maryland 21702

**ABSTRACT** We made a coupled genetic reporter that detects rare transcription misincorporation errors to measure RNA polymerase transcription fidelity in *Escherichia coli*. Using this reporter, we demonstrated *in vivo* that the transcript cleavage factor GreA, but not GreB, is essential for proofreading of a transcription error where a riboA has been misincorporated instead of a riboG. A *greA* mutant strain had more than a 100-fold increase in transcription errors relative to wild-type or a *greB* mutant. However, overexpression of GreB in  $\Delta greA$  cells reduced the misincorporation errors to wild-type levels, demonstrating that GreB at high concentration could substitute for GreA in RNA proofreading activity *in vivo*.

**KEYWORDS** RNA polymerase; transcription errors; Gre proteins; Cre/lox fidelity reporter

**T**HE potential impact of rare transcription errors on normal cell physiology and disease has been the subject of debate for several decades [for review see Strathern *et al.* (2012)]. The transient nature of messenger RNA (mRNA) and a relatively high ( $10^{-3}$ – $10^{-4}$ ) rate of translation misreading make detection of transcription errors extremely difficult (Edelmann and Gallant 1977; Rosenberger and Foskett 1981; Bouadloun *et al.* 1983; Kramer and Farabaugh 2007). To circumvent this problem, several genetic reporter assays were developed in bacteria and yeast based on a nonsense codon introduced into the open reading frame (ORF) of “fidelity reporters” such as the *lacZ* or luciferase (*luc*) genes (Rosenberger and Hilton 1983; Blank *et al.* 1986; Taddei *et al.* 1997; Shaw *et al.* 2002; Koyama *et al.* 2003; Roghanian *et al.* 2015). Transcription errors, transiently eliminating the nonsense mutation, would restore the full-length ORF of such a reporter, resulting in the production of a functional protein, the activity of which could be quantitatively determined. Although being broadly used, this approach suffers from an intrinsic problem of distinguishing transcription

errors from a substantially more frequent ribosome misreading of a nonsense codon (Shaw *et al.* 2002; Kramer and Farabaugh 2007). The mRNA carrying an internal stop codon is also subjected to a nonsense-mediated decay in eukaryotes (Lykke-Andersen and Jensen 2015) and rho-dependent termination in prokaryotes (Gussin *et al.* 1987), further complicating the use of nonsense codons to detect transcription errors.

The high-resolution RNA sequencing (RNA-seq) method was initially developed for mammalian cells (Li *et al.* 2011) and then applied to study transcription misincorporation errors in *Escherichia coli* (Imashimizu *et al.* 2013) and *Caenorhabditis elegans* (Gout *et al.* 2013). The mammalian studies showed unexpectedly high levels of errors in mRNA (Li *et al.* 2011), many of which were later reported to be artifacts of RNA-seq (Hayden 2012; Pickrell *et al.* 2012), highlighting a technical problem in the detection of rare transcription errors due to the presence of higher levels ( $10^{-3}$ / $10^{-4}$ /bp) of errors introduced by RNA-seq and PCR methods (Pickrell *et al.* 2012). More recently, a modified RNA-seq protocol was implemented that allowed separation of transcription errors from those made during reverse transcription and PCR during the processing of RNA samples. This more reliable method found  $\sim 4 \times 10^{-6}$  transcription misincorporations per base, with a highly-biased spectrum of error types (Gout *et al.* 2013).

Nascent elongating transcript sequencing in *E. coli* [NET-seq (Churchman and Weissman 2011)] has also been applied for detection (Imashimizu *et al.* 2015) and, more recently,

Copyright © 2017 by the Genetics Society of America

doi: <https://doi.org/10.1534/genetics.116.198960>

Manuscript received December 7, 2016; accepted for publication March 4, 2017; published Early Online March 23, 2017.

Available freely online through the author-supported open access option.

Supplemental material is available online at [www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.198960/-/DC1](http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.198960/-/DC1).

<sup>1</sup>Corresponding authors: NCI-Frederick, Post Office Box B, Building 560, Frederick, MD 21702. E-mail: [jjind@mail.nih.gov](mailto:jjind@mail.nih.gov); [courtld@mail.nih.gov](mailto:courtld@mail.nih.gov); [strathej@mail.nih.gov](mailto:strathej@mail.nih.gov); and [kashlevm@mail.nih.gov](mailto:kashlevm@mail.nih.gov)

for bioinformatics analysis (James *et al.* 2016) of transcription errors. This approach is based on the assumption that transcription errors at the 3'-end of nascent RNAs generally inhibit transcription elongation, leading to enrichment of nascent RNAs with transcription errors near the 3'-RNA end that had been isolated as part of *in vivo* paused RNA polymerase (RNAP) complexes (Kireeva *et al.* 2008; Sydow *et al.* 2009; Walmacq *et al.* 2009; Irvin *et al.* 2014). This method allows the identification of several different types of errors across the *E. coli* transcriptome, with the most frequent being G-to-A or C-to-U errors (Imashimizu *et al.* 2015). Despite these technical improvements, detection of transcription errors *in vivo* represents a very challenging and laborious task.

This work describes a coupled reporter assay to detect transcription errors in *E. coli* based on a Cre recombinase that is catalytically inactive due a missense mutation (Irvin *et al.* 2014). Restoration of active Cre depends upon a transcription misincorporation error. The transient mRNA is then translated into a functional Cre tetramer, whose activity in turn can be monitored by a Cre-dependent DNA recombination event in a *galk-loxP-INV* reporter that converts a *gal* mutant gene to *gal*<sup>+</sup>, giving rise to Gal<sup>+</sup> colonies on selective or indicator agar. Due to the high sensitivity of the assay, we successfully monitored the intrinsic level of transcription errors in wild-type cells and, for the first time, identified GreA protein as a major transcription proofreading factor in *E. coli*.

## Materials and Methods

### Strains and plasmids

All strains are listed in Supplemental Material, Table S1 in File S1. The strains are derivatives of MG1655 except for recombinogenic strains used for recombineering that are derivatives of W3110. Oligonucleotides used for strain construction and PCRs are listed in Table S2 in File S1. Standard procedures for cell growth, P1 transduction, and media preparation were employed.

### Strain construction by recombineering

The *greA*<>*spc* and *greB*<>*amp* knockouts were made by double-stranded DNA recombineering as described (Bubunenko *et al.* 2007). Briefly, the *spc* and *amp* resistance drug cassettes were constructed to precisely replace *gre* ORFs with the respective *spc* or *amp* ORFs. The drug cassettes were PCR amplified using the *greA*-*spcF* and -*spcR* and *greB*-*ampF* and -*ampR* sets of primers (see Table S2 in File S1), purified with a PCR Purification Kit (QIAGEN, Valencia, CA), and ~3 ng electroporated into recombinogenic DY330 cells that were induced for Red functions for 15 min at 42° at 0.5 A<sub>600</sub> optical units (o.u.). After electroporation, they were grown in LB at 32° for 3 hr, plated on L agar with ampicillin (Amp) 30 μg/ml or spectinomycin (Spc) 50 μg/ml, and incubated at 34° for 2 days to reveal recombinant cell colonies. The gene knockouts were identified by PCR amplification with “checking primers”

that flank the respective chromosome regions and were verified by sequencing.

Insertion of *cre* into the *ara* operon was done by a counterselection procedure (Sawitzke *et al.* 2013b). The strain XT191, carrying the *cat-sacB amp* cassette located in the *araD* gene of the arabinose operon, expresses Red recombination functions and is used for counterselection against the *sacB* gene conferring sucrose sensitivity (Li *et al.* 2013). The *cat-sacB amp* region was replaced by the PCR product of *cre*<sup>+</sup> flanked by DNA homology arms to the region upstream and downstream of the P<sub>BAD</sub> *ara* region (primers *creF* and *creR*, Table S2 in File S1) by selecting for sucrose resistance (Suc<sup>R</sup>) on LB agar plates containing 6% sucrose and no NaCl. Finally, a *cat* cassette was made with PCR primers *catF* and *catR* and inserted downstream from *cre* replacing the nonessential *yabQ* (see Figure S4 in File S1). The entire P<sub>BAD</sub> -*cre yabQ*<>*cat* segment could then be moved between different strains by P1 transduction. The primers used to amplify the *cat* cassette with homology arms flanking the *yabQ* region are shown in Table S2 in File S1. The knockout was identified by PCR amplification with “checking primers” that flank *yabQ*.

The luciferase gene *luc* was inserted under P<sub>BAD</sub> control in MG1655 by also replacing *cat-sacB* using the same homology arms for recombineering as used to insert the *cre*<sup>+</sup> gene under P<sub>BAD</sub> control (primers *lucF* and *lucR*). This, like for *cre*, results in the AUG of *luc* being positioned at the *araB* gene start codon (Figure S4E in File S1). See Figure S4, Figure S5, and Table S2 in File S1 for details of primers used and construction. The *greA*<>*spec* and *greB*<>*amp* knockouts were each crossed into the MG1655 P<sub>BAD</sub>-*luc* construct CC1234 to generate, by P1 transduction, CC1235 and CC1236, respectively. These three strains were compared for their luciferase activity before and after induction with arabinose.

The TGT mutation in codon 324 of *cre* was created by oligo-recombineering (Sawitzke *et al.* 2013a). The 70-mer oligo containing the single-base change in codon TAT<sub>324</sub> to generate TGT or TGC was electroporated into W3110 cells containing *araD*<>*cre* and pSIM18 carrying the RED recombinogenic functions at 0.35 A<sub>600</sub> o.u. The cells were grown at 32° for 2 hr in LB, nonselectively plated on L agar, and incubated at 32° for 2 days. The TGT/TGC mutations were identified by Mismatch Amplification Mutation Assay PCR with the respective antisense primers (see Table S2 in File S1) and verified by sequencing (Thomason *et al.* 2014).

### Construction of the *galk-loxP-INV* reporter system

The construction of the *galk-loxP-INV* reporter system is shown in Figure S1 in File S1. The *galk* gene is highly conserved from *E. coli* to humans. Alignment of all GalK protein homologs revealed regions of the protein that are not conserved. We focused on one of these regions. At a site located between amino acid residues 239 and 240 of the native *E. coli* protein, we inserted a 36-bp DNA segment that contained a 34-bp *loxP* sequence (see below and Figure S2 in File S1). Despite the 12-codon insertion, cells from this recombination formed Gal<sup>+</sup> colonies on minimal galactose agar (Figure S1B in File S1).

Next, we deleted the region from just downstream of the *loxP* site to beyond the *galM* gene, and replaced it with a *cat-sacB* cassette rendering the cells *galK*<sup>-</sup>, chloramphenicol-resistant, and sucrose-sensitive (Figure S1C in File S1). We then used PCR to amplify DNA from the construct shown in Figure S1B in File S1, so as to include *loxP* and the distal *galK* segment through to the end of *galM*. The PCR primers used for amplification carried homologies on each 5'-end designed to insert the *loxP galK galM* DNA segment in an inverted orientation by homologous recombination. This segment replaced the *cat-sacB* cassette by a recombination that generated Suc<sup>R</sup> and chloramphenicol-sensitive (Cm<sup>S</sup>) cells. The *loxP\** on the left contains an engineered single-base pair change that does not affect Cre-mediated recombination efficiency, but reduces homologous recombination between these two *loxP* sites. The two *loxP* sites, indicated by carets, are in inverted orientation relative to each other (Figure S1D in File S1). Lastly, the *tetA* gene from the Tn10 transposon was inserted just beyond the upstream *loxP\** site in *galK* so that translation of *galK* was prematurely terminated beyond the *loxP\** site at a translational stop codon, UGA, engineered into the construct (Figure S1E in File S1). Upon addition of active Cre protein, the DNA between the two *loxP* sites is inverted, rendering cells Gal<sup>+</sup>. The cells remain Tet resistant because the *tetA* gene in this inverted orientation relative to the original construct is juxtaposed next to a fortuitous promoter generated by the fusion of *loxP* and the *gpmA* segment (Figure S1F in File S1). The regions shown in Figure S1, E and F in File S1 have been sequenced and those sequences are shown in Figure S2 and Figure S3 in File S1.

#### **The *cre/galK-loxP-INV* transcription fidelity reporter assay**

The *cre*<sup>TGT</sup> or *cre*<sup>TGC</sup> constructs under P<sub>BAD</sub> control were introduced into strain CC942, which is MG1655 with the *galK-loxP-INV*, by P1 transduction selecting for Cm<sup>R</sup> and screening on MacConkey Galactose (MacGal) agar for the Gal<sup>-</sup> phenotype. The respective recombinant strains CC945 and CC947 were purified two times on LB plates with Na citrate and appropriate antibiotic and stored in glycerol at -80°. For qualitative analysis, a single colony from glycerol stock was patched, streaked, or spotted on MacGal agar or on minimal M63Gal agar/agarose, and incubated at 37° for 2–3 days on MacGal or for 2–4 and up to 6 days on M63Gal agar and agarose plates, respectively. The typical results are shown in Figure 2A. For quantitative analysis, to determine the apparent frequency of the G → A transcription error, 2 ml LB was inoculated with a single colony and incubated at 37° for 20–25 hr. Nine cell cultures, each from an independent cell colony, were taken for each tested strain. Cell cultures were concentrated by centrifugation for 2 min in a microfuge, washed once with M63 medium, and resuspended to their original volume in M63. Cells were plated in duplicate at different dilutions on M63Gal agarose or MacGal agar to reveal Gal<sup>+</sup> colonies, and on L agar

to count viable cells. MacGal plates were incubated for 2 days, whereas M63Gal agarose plates were incubated for 5–6 days at 37°.

#### **Overexpression of *GreA* and *GreB* transcription factors**

pGreA and pGreB (Koulich *et al.* 1997) were kind gifts from Sergei Borukhov. Both plasmids are pTrc88A-based with *gre* genes expressed from the *P<sub>trc</sub>* hybrid promoter. The plasmids were electroporated into the NB1001  $\Delta greA$  cells and, after 2 hr of incubation at 37°, the cell cultures were plated on L agar with Amp 100  $\mu$ g/ml. The *cre*<sup>TGT</sup>-*cat* reporter was introduced into the pGre-containing cells by P1 transduction and the resulting cells were patched on M63Gal agar with Amp 50  $\mu$ g/ml and then incubated at 37° for 3 days. Cells transformed with the related pBR322 vector were used as a control.

#### **Luciferase assays**

Luciferase activity was measured by using the reagents and methods described in the Promega Luciferase Assay System according to the company's directions (Promega, Madison, WI). Cells were grown in LB to an OD<sub>650</sub> of 0.4. Culture aliquots (0.1 ml) were centrifuged for 2 min at 21,000 × g, and pellets were resuspended in 0.4 ml of Cell Culture Lysis Reagent with 2.5 mg/ml BSA and 1.25 mg/ml lysozyme. Twenty microliters of cell lysate were mixed with 0.1 ml of the luciferin substrate, incubated for 120 sec, and read in and EG&G Berthold Lumat LB 9507 single sample luminometer for 10 sec. The relative light unit was normalized to A<sub>650</sub>.

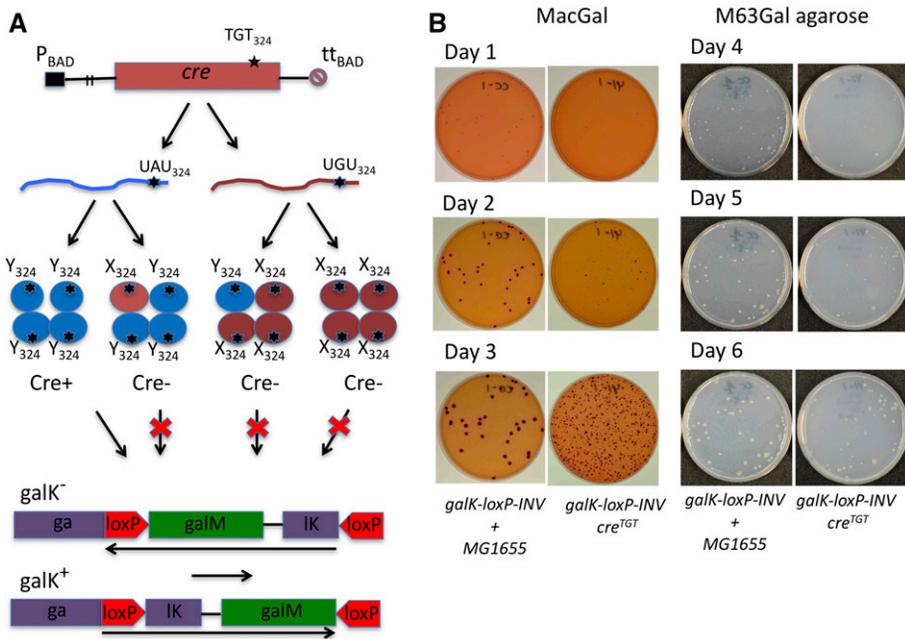
#### **Data availability**

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article. All strains and reagents are available upon request.

## **Results**

#### ***Cre/lox* reporter assay for detection of transcription errors in bacterial cells**

The Cre/Lox reporter assay described in this work was designed on the basis of our previous assay to study transcription fidelity in yeast cells (Irvin *et al.* 2014). The assay involved two sequential events that allowed the conversion of rare, transient transcription errors at a unique codon position of the *cre* gene to be detected as Cre-mediated recombination (Figure 1A). It was previously shown that the tyrosine residue in the active site of Cre was required in all the subunits of the active tetramer and that no other residue can substitute (Wierzbicki *et al.* 1987). In the bacterial system described here, Cre mediates recombination elsewhere in the genome causing Gal<sup>-</sup> cells to become Gal<sup>+</sup>. The two hallmarks of the assay are: (i) a catalytically inactive Cre recombinase caused by a missense mutation replacing the codon for the active site tyrosine (Y<sub>324</sub>), and (ii) a Cre recombination substrate, *i.e.*, two inverted LoxP sites flanking an inverted *galK* segment of



**Figure 1** The principles of the *cre*<sup>TGT</sup>/*galk-loxP-INV* assay designed for detection of G → A transcription errors in *E. coli*. (A) Flowchart of the assay. A *cre* gene with a TAT to TGT mutation in codon 324 was placed into the *ara* operon of the *E. coli* chromosome to ensure low-level expression of the totally inactive Cre Y324C protein from the P<sub>BAD</sub> promoter. Infrequent G → A transcription misincorporation events restore the wild-type Tyr324 residue in Cre yielding transient Cre recombination activity. (B) Plating efficiency assay used to quantitate preexisting Gal<sup>+</sup> as a measure of transcription misincorporation errors. The apparent frequency of Gal<sup>-</sup> → Gal<sup>+</sup> events is quantitated as a number of red colonies on indicator MacConkey Galactose (MacGal) plates or Gal<sup>+</sup> prototrophic colonies on selective M63Gal agarose plates. The time window for detection of preexisting Gal<sup>+</sup> is determined using a reference culture (shown as *galk-loxP-INV* + MG1655), which is composed of the defined small number of Gal<sup>+</sup> wild-type MG1655 cells (~60–80 cells) and the Gal<sup>-</sup> *galk-loxP-INV* cells (10<sup>4</sup>–10<sup>6</sup>) containing no Cre reporter

and, therefore, yielding no Gal<sup>+</sup> colonies. In this setup, MG1655 cells produce red Gal<sup>+</sup> colonies on an outgrowth of the Gal<sup>-</sup> *galk-loxP-INV* cells (the first column). Time when these colonies appear provides an important information on a time window for counting the Gal<sup>+</sup> colonies derived from Cre-mediated recombination in *galk-loxP-INV/cre*<sup>TGT</sup> cells (shown in the second column). A number of the Gal<sup>+</sup> colonies significantly increased after 3-day plating of *galk-loxP-INV/cre*<sup>TGT</sup> cells, but not the reference *galk-loxP-INV* + MG1655 cells, which was expected because the Cre-mediated recombination continued in *galk-loxP-INV/cre*<sup>TGT</sup> cells after they have already been plated on MacGal. In the example representing herein the quantitation of the G → A transcription error in the wild-type background (*galk-loxP-INV/cre*<sup>TGT</sup> cells), it takes 2 days to detect preexisting Gal<sup>+</sup> colonies on MacGal agar plates (the middle row) and up to 6 days to detect the colonies on the M63Gal agarose plates (the bottom row).

the galactose operon such that Cre-mediated recombination between the inverted LoxP sites restores an intact *galk* gene. The Gal<sup>+</sup> phenotype can be detected as cells able to form colonies on minimal M63 galactose or as red colonies on indicator MacGal agar (Figure 1B). The same recombination event also activates transcription of an inverted *tetA* gene located in the inversion site causing the cells to also become resistant to 12.5 μg/ml tetracycline.

The *cre*<sup>Y324C</sup> mutation changes the UAU codon to UGU and thus requires a G → A transcription error to create an mRNA encoding functional Cre (Irvin *et al.* 2014). The G → A error has been shown to be the major class of transcription misreading in *E. coli* (Imashimizu *et al.* 2013; James *et al.* 2016) and yeasts (Kireeva *et al.* 2008). When the mRNA is translated, a now functional Cre catalyzes recombination via a Lox-mediated DNA flipping event (Nagy 2000) at a *galk-loxP-INV* reporter gene. In this manner, a transient transcription error in the Cre mRNA is converted into a stable genetic change as a restored *galk* gene, the activity of which is assessed using Gal<sup>+</sup> selection or indicator agar. Importantly, the Cre protein acts as a tetramer strictly requiring Tyr<sub>324</sub> to be present in all four subunits (Figure 1) (Hartung and Kisters-Woike 1998; Ennifar *et al.* 2003; Lee and Sadowski 2003; Ghosh *et al.* 2007; Gibb *et al.* 2010). This property of Cre ensures that an active tetramer can arise from suppression of the codon 324 mutation at the transcription, rather than the translation, level. Indeed, translational suppression

would require ribosomes to misincorporate tyrosine for cysteine at the UGU codon four times in the same cell to have a chance of creating an active Cre tetramer, whereas only a single transcription error would be sufficient to produce an active Cre tetramer by multi-round translation of a single UAU-containing mRNA transcript. The estimated ribosomal error rate in *E. coli* is ~10<sup>-3</sup> per codon (Kramer and Farabaugh 2007), which makes four independent translation errors (10<sup>-12</sup>) happening at the UGU codon in the same cell and in a short period of time extremely unlikely. On the other hand, transcription errors, though rarer than translation errors, occur with rates of 10<sup>-5</sup>–10<sup>-7</sup>/nt in bacteria and eukaryotes (Rosenberger and Hilton 1983; Blank *et al.* 1986; Ninio 1991; Gout *et al.* 2013). We used this unique property of the Cre recombinase to reduce the impact of translation errors on the assay.

### Construction of *galk-loxP-INV* reporter

The construction of the *galk-loxP-INV* reporter system is shown in Figure S1 in File S1. The *galk* gene is highly conserved from *E. coli* to humans. Alignment of several homologs indicated four regions of the protein with reduced conservation, variable spacing, and located in loops on the protein surface. Using recombinering technology (Court *et al.* 2002), we first inserted DNA at these four positions to put the *galk* gene out of frame. We then restored the reading frame by substituting a 36-bp in-frame sequence containing

the 34-bp *loxP* site, and selected for Gal<sup>+</sup> cells. In all four cases, the 36-bp insertions resulted in a Gal<sup>+</sup> phenotype. For pragmatic reasons, we focused on a site located between amino acid residues 239 and 240 of the native *E. coli* sequence (Figure S1B in File S1). Next, we deleted the region from just downstream of the *loxP* site to beyond the *galM* gene, and replaced it with a *cat sacB* cassette (Sawitzke *et al.* 2013b), rendering the cells Gal<sup>-</sup>, chloramphenicol-resistant, and sucrose-sensitive (Figure S1C in File S1). Inversion of the downstream portion of *galK* including *galM* and a second *loxP* site was inserted by recombineering and selecting for sucrose resistance. (Figure S1D in File S1). This *loxP* site contained a single-base pair change that does not affect *cre*-mediated recombination efficiency, but reduces unwanted homologous recombination between *loxP* sites. Lastly, the *tetA* gene from Tn10 was inserted just beyond the upstream *loxP* site in *galK* to cause translation of the *galK* mRNA to terminate beyond the *loxP* site at a translational stop codon UGA engineered into the construct (Figure 1C). In this configuration, the cells have a Gal<sup>-</sup> phenotype. Upon providing active Cre protein, recombination between the two *loxP* sites inverts the DNA to generate an intact *galK* gene and restoring a Gal<sup>+</sup> phenotype (Figure S1F in File S1).

#### **Construction of the *cre* gene under *pBAD* promoter control**

The P1 Cre recombinase gene *cre* was PCR amplified with an optimized ribosome-binding site (Shine-Dalgarno, SD) upstream and inserted into the *E. coli* arabinose operon, replacing the native *araBAD* genes and their native SD by recombineering (Sawitzke *et al.* 2013b). This single copy of *cre* is under the control of the *ara* promoter region and AraC repressor, which provides very low-level transcription of Cre in the noninduced state used here. Any rare Cre transcripts are translated efficiently by the optimized SD (Guzman *et al.* 1995) ensuring that multiple peptides are made from one transcript. By keeping the level of transcription low, we hoped to minimize subunit mixing of Cre translated from functional and mutant transcripts. A *cat* gene conferring chloramphenicol resistance was placed downstream of the *cre* gene in the construct as a selectable marker to move the *cre* construct to other strains, and to also enable further *cre* gene modifications. These constructions are outlined in Figure S4 in File S1.

#### **Characterization of the coupled Cre and *galK-loxP-INV* reporter system**

The Cre and *galK-loxP-INV* reporter system was constructed by P1 transduction of *cre<sup>TGT</sup>* or *cre<sup>TGC</sup>* into the *galK-loxP-INV* reporter strain selecting for the linked Cm<sup>R</sup> marker (*yabQ<>cat*). These Cm<sup>R</sup> transductants were screened to be defective for utilization of either arabinose or galactose on MacConkey arabinose or MacGal agar plates, respectively. We note that a fraction of the transduced Cm<sup>R</sup> cells had a Gal<sup>+</sup> phenotype, as indicated by red colonies on

MacGal plates. This was caused by the presence of the P1 phage during the transduction. P1 carries the wild-type *cre* gene in its native genome, allowing expression of Cre and flipping of the *LoxP* sites in some of the cells being transduced. The cells remaining Gal<sup>-</sup> retained the inverted *galK* region as defined by PCR analysis.

#### **Optimization of the Cre/*galK-loxP-INV* reporter assay**

We tested various growth conditions and levels of the *cre<sup>TGT</sup>* expression to increase the sensitivity of the assay. Two different *cre<sup>TGT</sup>/galK-loxP-INV* clones kept at -80° were plated on LB medium at 37°, and four independent colonies of each clone were taken for these experiments.

For qualitative analysis, we employed cell patching and streaking techniques from single colonies on MacGal and M63Gal plates incubated at 30–42°. When the Gal<sup>-</sup> *cre<sup>TGT</sup>/galK-loxP-INV* cells were patched/streaked on MacGal indicator plates, infrequent Gal<sup>+</sup> red papillae appeared on top of a dense lawn of white Gal<sup>-</sup> cells. On minimal M63Gal plates, the rare Gal<sup>+</sup> colonies were formed because only Gal<sup>+</sup> cells were able to grow on a minimal medium with galactose. Thus, both types of the selection plates effectively detected *LoxP* recombinants. Three and five days of incubation at 37° on either MacGal or M63Gal plates, respectively, were optimal to reliably detect the Gal<sup>+</sup> colonies. These conditions were chosen as a standard for the patching and streaking assays throughout the work. Importantly, when the *cre* gene was not present with the *galK-loxP-INV* reporter (strain CC942), Gal<sup>+</sup> colonies were not detected even when > 10<sup>9</sup> cells were spread on the agar plates. This demonstrated that the *galK-loxP-INV* reporter was stable and did not undergo detectable recombination in the absence of Cre protein.

#### **Quantifying the apparent frequency of Gal<sup>-</sup> → Gal<sup>+</sup> events**

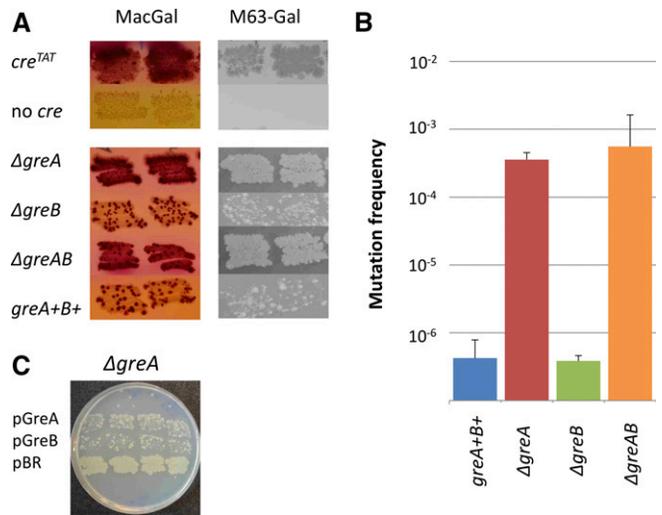
We used two approaches to detect Gal<sup>+</sup> cells: (1) MacGal plates on which Gal<sup>+</sup> colonies are red while Gal<sup>-</sup> colonies are white or (2) minimal plates in which galactose is a carbon source. Both approaches have features that require special attention to detail. On MacGal plates, all the Gal<sup>+</sup> and Gal<sup>-</sup> cells can grow into colonies and can continue to produce Gal<sup>+</sup> recombinants during that growth. Preexisting Gal<sup>+</sup> cells in LB were expected to form colonies more rapidly and be larger than their counterparts that became Gal<sup>+</sup> only after a growth period on the agar plate. We observed the appearance of two different Gal<sup>+</sup> populations when plating the *cre<sup>TGT</sup>/galK-loxP-INV* reporter strain on MacGal plates: (1) red colonies appearing between 24 and 48 hr of growth and (2) red colonies of variable sizes appearing after 48-hr incubation. The colonies coming up after 48-hr likely occur due to Cre-mediated recombination on the plate after these cells were put down. We have adopted a time window (24–48 hr) to identify red colonies derived by Cre-mediated recombination to *gal<sup>+</sup>* during the assay, from those occurring after the experiment. This was validated by control experiments with mixed plating of MG1655 Gal<sup>+</sup> cells in the presence of a lawn of Gal<sup>-</sup> cells from a *gal* deletion mutant. Similarly, we observed substantial growth of Gal<sup>-</sup> cells on

M63Gal agar plates due to an unknown carbon source in the Difco agar. On these plates, we could even detect residual growth of a *gal* deletion mutant. Because the *cre<sup>TGT</sup>/galK-loxP-INV* reporter strain could grow slowly on M63Gal agar, it generated an ever-increasing number of Cre-generated Gal<sup>+</sup> colonies with increased time of incubation. However, plates made with agarose instead of the Difco agar did not show this growth, and Gal<sup>+</sup> colonies derived during the experimental assay and before plating were readily detected at ~5–6 days and their number did not dramatically increase with time. Based on this finding, we have used M63Gal agarose and scored Gal<sup>+</sup> colonies after 5-days incubation at 37° throughout this work to determine the frequency of Gal<sup>+</sup> colonies. On the other hand, the MacGal plates were very useful for rapid and preliminary screening in *E. coli* strains with mutants affecting transcription fidelity and for wild-type cells growing under various conditions and stresses.

The MacGal method reproducibly yielded a ~10<sup>-6</sup> frequency of preexisting Gal<sup>+</sup> cells for the *cre<sup>TGT</sup>/galK-loxP* reporter strain when grown overnight in LB at 37°. A fluctuation test performed on M63 agarose yielded a similar Cre-mediated *galK*<sup>+</sup> recombination frequency (~10<sup>-6</sup>/cell) (see *Materials and Methods* for details). The scheme and the timeline for screening the Gal<sup>+</sup> colonies and the typical appearance of the colonies on MacGal and M63Gal agarose plates are shown in Figure 1B.

### Role of GreA and GreB in correcting transcription errors in *E. coli*

A role of GreA and GreB proteins in proofreading of transcription errors is widely postulated [Fish and Kane 2002; Borukhov *et al.* 2005; Zenkin and Yuzenkova 2015; and demonstrated *in vitro* for GreA (Erie *et al.* 1993)]. Recently, we gained evidence for an *in vivo* role of Gre factors by NET-seq, which detects transcription errors localized at the 3'-end of the nascent RNA (Imashimizu *et al.* 2015). We found that G→A transcription errors were substantially enriched at the 3' RNA ends in the *E. coli* cells lacking both GreA and GreB. The *cre<sup>TGT</sup>/galK-loxP-INV* reporter is designed to detect these G→A errors. Using our assay, we attempted to assess the role of Gre factors in modulating the G→A-type errors in *E. coli*. The  $\Delta greAB$  mutant used in the RNA-seq study was reconstructed for this experiment by introducing the *greA*<>*spc* and *greB*<>*amp* knockouts in the *galK-loxP-INV* background followed by the introduction of the *cre<sup>TGT</sup>* reporter. The  $\Delta greAB$  and *greA<sup>+</sup>B<sup>+</sup>* strains were patched on MacGal and M63Gal plates to observe the formation of Gal<sup>+</sup>. As shown by cell patching (Figure 2A),  $\Delta greAB$  significantly increased the accumulation of Gal<sup>+</sup> revertants, as compared to the *greA<sup>+</sup>B<sup>+</sup>/cre<sup>TGT</sup>* cells. Quantification of error frequencies in the  $\Delta greAB$  cells (Figure 2B) using M63Gal agarose plates solidly supported the patch analysis (Figure 2B). The number of Gal<sup>+</sup> revertants in  $\Delta greAB$  cells was ~2 orders of magnitude higher than in the *gre<sup>+</sup>* cells. Thus, the *cre<sup>TGT</sup>/galK-loxP-INV* reporter assay can easily detect transcription errors revealed in the absence of Gre factors.



**Figure 2** GreA is a major transcription proofreading factor identified by the *cre/galK-loxP-INV* genetic reporter assay. (A) Cells patched on MacConkey Galactose (MacGal) plates on M63Gal agar/agarose plates all show similar results that the G→A transcription error, identified by the density of Gal<sup>+</sup> colonies, dramatically increases if GreA ( $\Delta greA$  cells) or both GreA and GreB ( $\Delta greAB$  cells) are deleted from the wild-type cells (*greA<sup>+</sup>B<sup>+</sup>* cells). Deletion of only GreB ( $\Delta greB$  cells) has no effect. (B) Quantitation of the effect of Gre factors on the apparent frequency of G→A transcription error measured by the plating efficiency assay on M63Gal agarose. Each bar represents a value of Gal<sup>+</sup> counts from nine independent cell cultures plated in duplicate with the error bars shown. (C) Patches of  $\Delta greA$  cells on M63Gal agar show that overexpression of GreB (pGreB) or GreA (pGreA) almost equally suppresses the frequency of G→A transcription error in GreA-lacking cells as compared to the same cells containing pBR322 control plasmid.

### GreA, but not GreB, acts as an RNA proofreading factor

The relative impact of GreA vs. GreB factor on proofreading of transcription errors *in vitro* and *in vivo* has never been approached experimentally. To examine their impact on fidelity, we constructed single *greA*<>*spc* and *greB*<>*amp* knockouts in the *galK-loxP-INV* background. As with the  $\Delta greAB$  strain, these strains did not form Gal<sup>+</sup> colonies either on MacGal or M63Gal plates in the absence of a *cre* gene. When the *cre<sup>TGT</sup>* allele was present, the  $\Delta greA$  but not the  $\Delta greB$  strain showed an increased accumulation of Gal<sup>+</sup> revertants as compared to the *greA<sup>+</sup>B<sup>+</sup>/cre<sup>TGT</sup>* strain (Figure 2A). These data demonstrated a strong effect of GreA and no apparent effect of GreB on transcription fidelity of RNAP in correcting a G-to-A mistake in transcription.

Quantification of error frequencies in the  $\Delta greA$  cells (Figure 2B) using M63Gal agarose plates confirmed the results from the cell patching analysis (Figure 2B). Similarly to the  $\Delta greAB$  cells, the number of Gal<sup>+</sup> revertants in the  $\Delta greA$  cells was ~2 orders of magnitude higher than in the *gre<sup>+</sup>* cells, but was virtually unaffected in the  $\Delta greB$  cells. Clearly, a lack of GreB did not change the G→A error frequency either in the *greA<sup>+</sup>* or in the  $\Delta greA$  cells. Thus, the quantitative analysis revealed a critical role of GreA but not GreB in controlling the transcription errors in *E. coli* cells under growth conditions in rich media.

The lack of a GreB effect on transcription fidelity was surprising, especially in contrast to the substantial effect of its counterpart GreA (Figure 2, A and B). The two Gre factors have always been considered as having similar roles that would enhance fidelity of RNAP transcription (Fish and Kane 2002; Borukhov *et al.* 2005; Zenkin and Yuzenkova 2015). Both stimulate the intrinsic RNA cleavage activity of RNAP when it is stalled in the backtracked state (Borukhov *et al.* 1993; Toulme *et al.* 2000). Both also interact with the same site on the RNA polymerase (Laptenko *et al.* 2003; Opalka *et al.* 2003; Sosunova *et al.* 2003), and both can be efficient in proofreading of G→A errors, as demonstrated by our *in vitro* misincorporation experiments (Figure S6 in File S1). Therefore, we reasoned that the apparent lack of an impact of GreB on transcription fidelity *in vivo* might be due either to its lower affinity for the RNAP elongation complex than that of GreA, or to a relatively lower concentration of GreB in the cell (Rutherford *et al.* 2007; Lamour *et al.* 2008; Vinella *et al.* 2012), or both.

We tested this by overexpressing the GreB protein in the  $\Delta greA$  background. The experiment was done with the  $\Delta greA$   $cre^{TGT}/galK-loxP-INV$  strain transformed with plasmids that expressed either GreA or GreB protein from a leaky P<sub>trc</sub> promoter (Hsu *et al.* 1995) without IPTG induction. The pBR322 vector was taken as a control. Gal<sup>+</sup> revertants were scored by patching on M63Gal. The results showed (Figure 2C) that, in comparison to the  $\Delta greA$  cells containing pBR322, GreB overexpressed from the plasmid in  $\Delta greA$  cells reduced the intensity of Gal<sup>+</sup> patches almost to the level of cells overexpressing GreA. Thus, transcription fidelity in the  $\Delta greA$  mutant can be complemented by either GreA or GreB overexpression with efficient correction of G→A transcription errors, at least in the sequence context of our *cre* reporter.

The Gre factors also play a role in the global regulation of gene expression (Rutherford *et al.* 2007; Stepanova *et al.* 2007; Vinella *et al.* 2012). To rule out that the effect of Gre factors on Gal<sup>+</sup> accumulation in the assay is caused by their regulation of *cre* expression from the arabinose promoter, *luc* gene fusions were constructed by precisely replacing the *cre* ORF in the *ara* operon with the *luc* ORF in the strains with  $\Delta greA greB^+$ ,  $greA^+ \Delta greB$ , and  $greA^+ greB^+$  backgrounds. All three strains were inducible by arabinose from  $\sim 2 \times 10^5$  light units to  $2 \times 10^8$  light units. Less than a twofold variation in luciferase activity was seen among the three strains, indicating that the effect of Gre in the assay is attributable to modulating transcription errors, but not the level of *cre* expression from the arabinose promoter. This is consistent with the *ara* operon genes not having been reported as targets for Gre factors (Stepanova *et al.* 2007; Vinella *et al.* 2012).

## Discussion

In the present work, we describe a genetic assay to detect transcription misincorporation of G by A in *E. coli*. The genetic assay (Figure 1) consists of two main modules residing in the bacterial chromosome: (1) a gene for Cre recombinase with a

single-nucleotide change (A→G) in the second position of the TAT<sub>324</sub> codon that inserts cysteine for the essential tyrosine residue at position 324, and (2) a substrate for Cre recombinase, the inverted portion of *galK* flanked by a pair of *loxP* sites oriented so as to confer a Gal<sup>-</sup> phenotype until inverted by Cre-mediated recombination. A rare transcription misincorporation event converting the mutated codon back to UAU in the *Cre* mRNA results in the production of active Cre tetramers via multiple rounds of translation of the unique Cre<sup>+</sup> mRNA molecule. The recombination activity of the Cre tetramer at the *loxP* sites in *galK-loxP-INV* restores the *galK* gene, yielding Gal<sup>+</sup> colonies on minimal Gal-selective or MacGal indicator agar. In this manner, misincorporation in the *Cre* mRNA is detected by a secondary event, *i.e.*, a genetic change restoring the *galK* gene and the Gal<sup>+</sup> phenotype to the cell. The virtual absence of Cre-independent recombination events allows detection of transcription errors with apparent frequencies as low as  $10^{-8}$  per cell.

The TGT<sub>324</sub> *cre* reporter mutant was designed for detection of the G→A transition errors that had been previously found to be the most frequent type of transcription misincorporation error both in yeast and bacterial cells (Kireeva *et al.* 2008; Imashimizu *et al.* 2013). Our reporter system has been used to optimize some of the conditions for the assay and to analyze the natural level of the G→A error at this position in wild-type *E. coli*, and to determine the role of Gre proofreading factors in control of error frequency. GreA and GreB were long suspected to have a role in proofreading of transcription misincorporation errors *in vivo* (Erie *et al.* 1993). Both factors (Erie *et al.* 1993), as well as their eukaryotic analog TFIIS (Thomas *et al.* 1998; Sydow *et al.* 2009; Irvin *et al.* 2014; Schweikhard *et al.* 2014), reduce errors *in vitro* by stimulating intrinsic RNA cleavage by RNAP in elongation complexes that are arrested by backtracking. Errors at the 3'-end of nascent RNA isolated from stalled elongation complexes *in vitro* and *in vivo* (Imashimizu *et al.* 2013) have been shown to promote RNAP backtracking. Gre-mediated proofreading activity has been directly demonstrated *in vitro* for GreA (Erie *et al.* 1993) and more recently *in vivo* over the entire transcriptome by RNA-seq and NET-seq analyses of transcription errors when comparing wild-type vs. a  $\Delta greAB$  double mutant (Imashimizu *et al.* 2013, 2015), and by bioinformatics analysis (James *et al.* 2016).

In the present work, G→A errors were found to occur with a relatively low  $10^{-6}$  frequency in wild-type *E. coli* using our reporter assay. The error frequency was dramatically increased in the  $\Delta greAB$  mutant, consistent with our previously published data on genome-wide NET-seq analysis comparing transcription errors in wild-type and  $\Delta greAB$  cells (Imashimizu *et al.* 2015). Here, we show that GreA has the major impact on transcription fidelity because a similar increase of magnitude in G→A misincorporation occurred in either a  $\Delta greA$  single or  $\Delta greAB$  double knockouts, whereas wild-type and  $\Delta greB$  mutant cells were virtually indistinguishable in error frequency.

Our *in vivo* results for GreB were surprising since GreB has always been assigned a more versatile and potent function from *in vitro* studies (Borukhov *et al.* 2005). For example, GreB is more efficient at catalyzing RNA cleavage for *in vitro* transcription assays, causing cleavage from the RNA 3'-end of up to 18-nt long RNA fragments, whereas GreA cleaves off only 2–4-nt long fragments (Borukhov *et al.* 1993). However, an overall examination of the characteristics of the two Gre proteins revealed that, while both Gre proteins bind to a coincident site within the secondary pore of RNAP to exert their effect (Laptenko *et al.* 2003; Opalka *et al.* 2003; Sosunova *et al.* 2003), GreB is five times less abundant in cells (Rutherford *et al.* 2007; Lamour *et al.* 2008). These properties might disfavor GreB binding to RNAP. Indeed, we observed that overexpression of GreB in the  $\Delta greA$  cells reduced G→A errors to wild-type levels, demonstrating that GreB at high concentration could substitute for GreA in RNA proofreading activity *in vivo*. Thus, one cannot exclude the possibility that GreB may act as a proofreading factor in wild-type cells under certain growth conditions or on other sequence contexts, or perhaps with rare RNAP elongation conformations that are not accessible by GreA. It has been shown *in vitro* that when RNAP backtracks at distances > 2–3-bp, RNA catalysis by GreA is inhibited while that of GreB is stimulated (Borukhov *et al.* 1993). GreB may primarily use RNA proofreading activity for transcriptional emergencies that are distinct from transcription fidelity, *e.g.*, the rescue of RNAP from a prolonged pausing after backtracking (James *et al.* 2016). For these reasons, the limited proofreading capabilities of GreB under normal *E. coli* growth and its role in cell physiology will require further examination.

In conclusion, the assay that we have developed on the basis of the Cre/Lox reporter system has shown great promise for detecting transcription errors and for uncovering new details of Gre-mediated transcription proofreading in living cells. Importantly, the results of its application are consistent with our whole transcriptome analysis of misincorporation errors in *E. coli* (Imashimizu *et al.* 2013, 2015). The reporter assay will be useful in furthering our understanding of the role and mechanisms of transcription fidelity in bacterial cells, as well as for identifying RNAP mutants with altered error rates. Other proteins, *e.g.*, DksA and RNK, in addition to GreA and GreB also bind to the secondary channel of RNAP and are known to affect RNAP elongation (Fish and Kane 2002; Borukhov *et al.* 2005; Zenkin and Yuzenkova 2015) and transcription misincorporation (Roghianian *et al.* 2015; Satory *et al.* 2015). An evaluation of the impact of these proteins on transcription fidelity in *E. coli* with the Cre/Lox reporter is currently underway.

## Acknowledgments

We thank Thomas Schneider for bioinformatics support in identification of the fortuitous promoter for the *tetA* gene. We are also grateful to the other members of the RNA Biology Laboratory for discussion. This work was funded by

the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research (to M.K., J.N.S., D.L.C., and D.J.J.). Funding for open access charge: The Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research.

## Literature Cited

- Blank, A., J. A. Gallant, R. R. Burgess, and L. A. Loeb, 1986 An RNA polymerase mutant with reduced accuracy of chain elongation. *Biochemistry* 25: 5920–5928.
- Borukhov, S., V. Sagitov, and A. Goldfarb, 1993 Transcript cleavage factors from *E. coli*. *Cell* 72: 459–466.
- Borukhov, S., J. Lee, and O. Laptenko, 2005 Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Mol. Microbiol.* 55: 1315–1324.
- Bouadloun, F., D. Donner, and C. G. Kurland, 1983 Codon-specific missense errors *in vivo*. *EMBO J.* 2: 1351–1356.
- Bubunenko, M., T. Baker, and D. L. Court, 2007 Essentiality of ribosomal and transcription antitermination proteins analyzed by systematic gene replacement in *Escherichia coli*. *J. Bacteriol.* 189: 2844–2853.
- Churchman, L. S., and J. S. Weissman, 2011 Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* 469: 368–373.
- Court, D. L., J. A. Sawitzke, and L. C. Thomason, 2002 Genetic engineering using homologous recombination. *Annu. Rev. Genet.* 36: 361–388.
- Edelmann, P., and J. Gallant, 1977 Mistranslation in *E. coli*. *Cell* 10: 131–137.
- Ennifar, E., J. E. Meyer, F. Buchholz, A. F. Stewart, and D. Suck, 2003 Crystal structure of a wild-type Cre recombinase-loxP synapse reveals a novel spacer conformation suggesting an alternative mechanism for DNA cleavage activation. *Nucleic Acids Res.* 31: 5449–5460.
- Erie, D. A., O. Hajiseyedjavadi, M. C. Young, and P. H. von Hippel, 1993 Multiple RNA polymerase conformations and GreA: control of the fidelity of transcription. *Science* 262: 867–873.
- Fish, R. N., and C. M. Kane, 2002 Promoting elongation with transcript cleavage stimulatory factors. *Biochim. Biophys. Acta* 1577: 287–307.
- Ghosh, K., F. Guo, and G. D. Van Duyne, 2007 Synapsis of loxP sites by Cre recombinase. *J. Biol. Chem.* 282: 24004–24016.
- Gibb, B., K. Gupta, K. Ghosh, R. Sharp, J. Chen *et al.*, 2010 Requirements for catalysis in the Cre recombinase active site. *Nucleic Acids Res.* 38: 5817–5832.
- Gout, J. F., W. K. Thomas, Z. Smith, K. Okamoto, and M. Lynch, 2013 Large-scale detection of *in vivo* transcription errors. *Proc. Natl. Acad. Sci. USA* 110: 18584–18589.
- Gussin, G. N., S. Brown, and K. Matz, 1987 Translational polarity of a mutation in the initiator AUG codon of the  $\lambda$  *cl* gene. *Genetics* 117: 173–179.
- Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith, 1995 Tight regulation, modulation, and high-level expression by vectors containing the arabinose *PBAD* promoter. *J. Bacteriol.* 177: 4121–4130.
- Hartung, M., and B. Kisters-Woike, 1998 Cre mutants with altered DNA binding properties. *J. Biol. Chem.* 273: 22884–22891.
- Hayden, E. C., 2012 RNA studies under fire. *Nature* 484: 428.
- Hsu, L. M., N. V. Vo, and M. J. Chamberlin, 1995 *Escherichia coli* transcript cleavage factors GreA and GreB stimulate promoter escape and gene expression *in vivo* and *in vitro*. *Proc. Natl. Acad. Sci. USA* 92: 11588–11592.

- Imashimizu, M., T. Oshima, L. Lubkowska, and M. Kashlev, 2013 Direct assessment of transcription fidelity by high-resolution RNA sequencing. *Nucleic Acids Res.* 41: 9090–9104.
- Imashimizu, M., H. Takahashi, T. Oshima, C. McIntosh, M. Bubunenکو *et al.*, 2015 Visualizing translocation dynamics and nascent transcript errors in paused RNA polymerases *in vivo*. *Genome Biol.* 16: 98.
- Irvin, J. D., M. L. Kireeva, D. R. Gotte, B. K. Shafer, I. Huang *et al.*, 2014 A genetic assay for transcription errors reveals multilayer control of RNA polymerase II fidelity. *PLoS Genet.* 10: e1004532.
- James, K., P. Gamba, S. J. Cockell, and N. Zenkin, 2016 Misincorporation by RNA polymerase is a major source of transcription pausing *in vivo*. *Nucleic Acids Res.* 45: 1105–1113.
- Kireeva, M. L., Y. A. Nedialkov, G. H. Cremona, Y. A. Purtov, L. Lubkowska *et al.*, 2008 Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. *Mol. Cell* 30: 557–566.
- Koulich, D., M. Orlova, A. Malhotra, A. Sali, S. A. Darst *et al.*, 1997 Domain organization of *Escherichia coli* transcript cleavage factors GreA and GreB. *J. Biol. Chem.* 272: 7201–7210.
- Koyama, H., T. Ito, T. Nakanishi, N. Kawamura, and K. Sekimizu, 2003 Transcription elongation factor S-II maintains transcriptional fidelity and confers oxidative stress resistance. *Genes Cells* 8: 779–788.
- Kramer, E. B., and P. J. Farabaugh, 2007 The frequency of translational misreading errors in *E. coli* is largely determined by tRNA competition. *RNA* 13: 87–96.
- Lamour, V., S. T. Rutherford, K. Kuznedelov, U. A. Ramagopal, R. L. Gourse *et al.*, 2008 Crystal structure of *Escherichia coli* Rnk, a new RNA polymerase-interacting protein. *J. Mol. Biol.* 383: 367–379.
- Laptenko, O., J. Lee, I. Lomakin, and S. Borukhov, 2003 Transcript cleavage factors GreA and GreB act as transient catalytic components of RNA polymerase. *EMBO J.* 22: 6322–6334.
- Lee, L., and P. D. Sadowski, 2003 Identification of Cre residues involved in synapsis, isomerization, and catalysis. *J. Biol. Chem.* 278: 36905–36915.
- Li, M., I. X. Wang, Y. Li, A. Bruzel, A. L. Richards *et al.*, 2011 Widespread RNA and DNA sequence differences in the human transcriptome. *Science* 333: 53–58.
- Li, X. T., L. C. Thomason, J. A. Sawitzke, N. Costantino, and D. L. Court, 2013 Positive and negative selection using the *tetA-sacB* cassette: recombineering and P1 transduction in *Escherichia coli*. *Nucleic Acids Res.* 41: e204.
- Lykke-Andersen, S., and T. H. Jensen, 2015 Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat. Rev. Mol. Cell Biol.* 16: 665–677.
- Nagy, A., 2000 Cre recombinase: the universal reagent for genome tailoring. *Genesis* 26: 99–109.
- Ninio, J., 1991 Connections between translation, transcription and replication error-rates. *Biochimie* 73: 1517–1523.
- Opalka, N., M. Chlenov, P. Chacon, W. J. Rice, W. Wriggers *et al.*, 2003 Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell* 114: 335–345.
- Pickrell, J. K., Y. Gilad, and J. K. Pritchard, 2012 Comment on “Widespread RNA and DNA sequence differences in the human transcriptome.” *Science* 335: 1302, author reply 1302.
- Roghianian, M., N. Zenkin, and Y. Yuzenkova, 2015 Bacterial global regulators DksA/ppGpp increase fidelity of transcription. *Nucleic Acids Res.* 43: 1529–1536.
- Rosenberger, R. F., and G. Foskett, 1981 An estimate of the frequency of *in vivo* transcriptional errors at a nonsense codon in *Escherichia coli*. *Mol. Gen. Genet.* 183: 561–563.
- Rosenberger, R. F., and J. Hilton, 1983 The frequency of transcriptional and translational errors at nonsense codons in the *lacZ* gene of *Escherichia coli*. *Mol. Gen. Genet.* 191: 207–212.
- Rutherford, S. T., J. J. Lemke, C. E. Vrentas, T. Gaal, W. Ross *et al.*, 2007 Effects of DksA, GreA, and GreB on transcription initiation: insights into the mechanisms of factors that bind in the secondary channel of RNA polymerase. *J. Mol. Biol.* 366: 1243–1257.
- Satory, D., A. J. E. Gordon, M. Y. Wang, J. A. Halliday, I. Golding *et al.*, 2015 DksA involvement in transcription fidelity buffers stochastic epigenetic change. *Nucleic Acids Res.* 43: 10190–10199.
- Sawitzke, J. A., L. C. Thomason, M. Bubunenکو, X. Li, N. Costantino *et al.*, 2013a Recombineering: highly efficient *in vivo* genetic engineering using single-strand oligos. *Methods Enzymol.* 533: 157–177.
- Sawitzke, J. A., L. C. Thomason, M. Bubunenکو, X. Li, N. Costantino *et al.*, 2013b Recombineering: using drug cassettes to knock out genes *in vivo*. *Methods Enzymol.* 533: 79–102.
- Schweikhard, V., C. Meng, K. Murakami, C. D. Kaplan, R. D. Kornberg *et al.*, 2014 Transcription factors TFIIF and TFIIS promote transcript elongation by RNA polymerase II by synergistic and independent mechanisms. *Proc. Natl. Acad. Sci. USA* 111: 6642–6647.
- Shaw, R. J., N. D. Bonawitz, and D. Reines, 2002 Use of an *in vivo* reporter assay to test for transcriptional and translational fidelity in yeast. *J. Biol. Chem.* 277: 24420–24426.
- Sosunova, E., V. Sosunov, M. Kozlov, V. Nikiforov, A. Goldfarb *et al.*, 2003 Donation of catalytic residues to RNA polymerase active center by transcription factor Gre. *Proc. Natl. Acad. Sci. USA* 100: 15469–15474.
- Stepanova, E., J. Lee, M. Ozerova, E. Semenova, K. Datsenko *et al.*, 2007 Analysis of promoter targets for *Escherichia coli* transcription elongation factor GreA *in vivo* and *in vitro*. *J. Bacteriol.* 189: 8772–8785.
- Strathern, J. N., D. J. Jin, D. L. Court, and M. Kashlev, 2012 Isolation and characterization of transcription fidelity mutants. *Biochim. Biophys. Acta* 1819: 694–699.
- Sydow, J. F., F. Brueckner, A. C. Cheung, G. E. Damsma, S. Dengl *et al.*, 2009 Structural basis of transcription: mismatch-specific fidelity mechanisms and paused RNA polymerase II with frayed RNA. *Mol. Cell* 34: 710–721.
- Taddei, F., H. Hayakawa, M. Bouton, A. Cirinesi, I. Matic *et al.*, 1997 Counteraction by MutT protein of transcriptional errors caused by oxidative damage. *Science* 278: 128–130.
- Thomas, M. J., A. A. Platas, and D. K. Hawley, 1998 Transcriptional fidelity and proofreading by RNA polymerase II. *Cell* 93: 627–637.
- Thomason, L. C., J. A. Sawitzke, X. Li, N. Costantino, and D. L. Court, 2014 Recombineering: genetic engineering in bacteria using homologous recombination. *Curr. Protoc. Mol. Biol.* 106: 1.16.1–1.16.39.
- Toulme, F., C. Mosrin-Huaman, J. Sparkowski, A. Das, M. Leng *et al.*, 2000 GreA and GreB proteins revive backtracked RNA polymerase *in vivo* by promoting transcript trimming. *EMBO J.* 19: 6853–6859.
- Vinella, D., K. Potrykus, H. Murphy, and M. Cashel, 2012 Effects on growth by changes of the balance between GreA, GreB, and DksA suggest mutual competition and functional redundancy in *Escherichia coli*. *J. Bacteriol.* 194: 261–273.
- Walmacq, C., M. L. Kireeva, J. Irvin, Y. Nedialkov, L. Lubkowska *et al.*, 2009 Rpb9 subunit controls transcription fidelity by delaying NTP sequestration in RNA polymerase II. *J. Biol. Chem.* 284: 19601–19612.
- Wierzbicki, A., M. Kendall, K. Abremski, and R. Hoess, 1987 A mutational analysis of the bacteriophage P1 recombinase Cre. *J. Mol. Biol.* 195: 785–794.
- Zenkin, N., and Y. Yuzenkova, 2015 New insights into the functions of transcription factors that bind the RNA polymerase secondary channel. *Biomolecules* 5: 1195–1209.

Communicating editor: M. Hampsey