Genome-Wide Study of mRNA Degradation and Transcript Elongation in Escherichia coli

Genome-wide study of mRNA degradation and transcript elongation in *Escherichia coli*

Running title: Genome-wide mRNA dynamics study in *E.coli*

Huiyi Chen[1,2,*,a],Katsuyuki Shiroguchi[2,§,a], Hao Ge[3,4,b], X. Sunney Xie[2,3,b]

1. Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA
2. Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA
3. Biodynamic Optical Imaging Center (BIOPIC), School of Life Science, Peking University, Beijing 100871, China
4. Beijing International Center for Mathematical Research (BICMR), Peking University, Beijing 100871, China

*** Current Address: School of Biological Sciences, Nanyang Technological University, Singapore 637551**
§ Current Address: Laboratory for Quantitative Omics, RIKEN Center for Integrative Medical Sciences (IMS), Yokohama, Kanagawa 230-0045, Japan; where K.S. completed the work

a. equal contribution by authors
b. correspondence to be addressed to H.G. (haoge@pku.edu.cn) and X.S.X. (xie@chemistry.harvard.edu)

Abstract

An essential part of gene expression is the coordination of RNA synthesis and degradation, which occurs in the same cellular compartment in bacteria. Here, we report a genome-wide RNA degradation study in *Escherichia coli* using RNA-seq, and present evidence that the stereotypical exponential RNA decay curve obtained using initiation inhibitor, rifampicin, consists of two phases: residual RNA synthesis, a delay in the interruption of steady-state that is dependent on distance relative to the mRNA's 5'-end, and the exponential decay. This gives a more accurate RNA lifetime and RNA polymerase elongation rate simultaneously genome-wide. Transcripts typically have a single RNA decay constant along all positions, which is distinct between different operons, indicating that RNA stability is unlikely determined by local sequences. These measurements allowed us to establish a model for RNA processing involving co-transcriptional degradation, providing quantitative description of the macromolecular coordination in gene expression in bacteria on a system-wide level.

Keywords: modeling/mRNA decay/ RNAP elongation/ RNA-seq/ transcription

Subject Category: Quantitative Biology & Dynamical Systems/ Methods & Resources/ RNA Biology

Introduction

Descriptions of RNA synthesis and degradation are essential for understanding the dynamics of gene expression. The spatial localization of RNA in bacterial cells has been studied in detail by microscopy (Llopis et al., 2010; Nevo-Dinur et al., 2011). The kinetics of RNA synthesis have been measured directly in live *Escherichia coli (E. coli)* cells using fluorescently tagged RNA molecules (Golding & Cox, 2004). More traditionally, studies of RNA lifetimes of lacZ have contributed to describing the interaction of transcription, translation and RNA degradation (Schwartz et al, 1970; Yarchuk et al, 1992). Studies were also performed within RNA transcripts, leading to the identification of unusually stable RNA segments (Von Gabain et al, 1983). However, in these cases, only one or few RNAs were studied, so it is difficult to identify global rules that can be applied genome wide.

High throughput techniques, like microarray, have increased the amount of information through genome-wide measurements, allowing researchers to observe the global behaviors for identifying the typical and the unusual in gene expression systems. The first microarray study of RNA lifetime in *E.coli* increased the number of measured RNA lifetimes from by two orders of magnitude (Bernstein et al, 2002). The availability of such data allowed researchers to perform systemwide analysis and modeling (Bon et al, 2006; Deneke et al, 2013). Selinger *et al.* designed a microarray using multiple probes along operons to study degradation at sub-genic resolution, yielding the insight that 5' ends are generally less stable than 3' ends (Selinger et al, 2003). High throughput sequencing technology that has emerged in recent years allows

researchers to study nucleic acids with even higher base resolution. One of its applications is RNA sequencing (RNA-seq), which has been successfully used in genome-wide studies to investigate details in the mechanisms of processes like RNA polymerase transcription, and mRNA splicing (Ameur et al, 2011; Churchman & Weissman, 2011; Taggart et al, 2012). RNA-seq has been used to study RNA degradation in *Bacillus cereus* and *Bacillus subtilis*, although with limited or no time resolution (Kristoffersen et al, 2012; Liu et al, 2014). We now apply RNA-seq to the study of the kinetics of RNA dynamics to obtain high resolution data genome-wide so that we can further understand the interplay between degradation and synthesis.

The RNAP initiation inhibitor, rifampicin, has been used in genome-wide studies of RNA degradation (Bernstein et al, 2002; Selinger et al, 2003). The degradation of RNA is observed and measured after synthesis is inhibited. While rifampicin binds all free RNAPs, it does not affect RNAPs that are already bound and transcribing on the DNA, resulting in residual RNA synthesis (Pato & Von Meyenburg, 1970). The effect of such residual RNA synthesis upon RNA decay has already been indirectly inferred from the observed delays in RNA decay (Chow & Dennis, 1994; Von Gabain et al, 1983; Yarchuk et al, 1992) and modeled mathematically (Deneke et al, 2013), although others have attributed the delay to rifampicin permeability (Llopis et al, 2010). If the resolution of the measurement is high enough to capture such a distance-dependent delay and it is proved to be caused by residual synthesis, then RNA elongation rates can be measured genome-wide.

In this study, we obtained details of global RNA dynamics in exponentially-growing and stationary phase *E. coli* cells using RNA-seq, with sub-minute time resolution to allow observation of RNA synthesis and degradation. We observed, similar to previous studies (Chow & Dennis, 1994; Deneke et al, 2013; Llopis et al, 2010), that there is a delay before RNA abundance decays after addition of rifampicin in both growth conditions. We show through direct experiment that elongating RNAPs are primarily responsible for this delay. As a result, we extracted RNAP elongation rate and RNA degradation rate simultaneously from our high resolution data. The fitted exponential lifetimes for RNA degradation at all positions of a transcript are similar for most mono-cistronic, and poly-cistronic RNAs. RNA dynamics in stationary phase *E.coli* exhibit the same trends as in exponentially growing cells, suggesting that any coordination between transcription and degradation remains essentially unchanged despite changes in transcription initiation rates. Finally, we revisited the meaning of the observed exponential decay of RNA and propose that co-transcriptional degradation is prevalent among RNAs, underscoring the interlinked nature of transcription, translation and RNA degradation.

Result

Simultaneous measurement of RNA chain elongation and degradation genome wide using rifampicin

RNA-seq was used to measure the abundance of mRNA in *E.coli* cells growing exponentially (OD_{600nm} = 0.3) at various time intervals after treatment with rifampicin, an initiation inhibitor of RNAP. The data was segmented into 300-nucleotide bins according to transcription unit annotation (Keseler et al, 2012), and includes both mRNA and non-coding RNA (Supplementary

Table S1, S2). *In vitro* synthesized RNAs, used as a non-degrading control, were added at defined concentrations to the bacterial lysate, and used to normalize relative abundance within time points (Supplementary Fig. S1). The abundance of RNA at different positions is further normalized with respect to the data at the zero minute time point. After the normalization, our genome-wide dataset revealed that there is less RNA on the 5' ends compared to the 3' ends of RNA transcription units at subsequent time points (Fig. 1a). This difference between the 5' and 3' ends was previously noted by Selinger *et al.*, who concluded that there was a difference in 5' and 3' stabilities on RNAs (Selinger et al, 2003). This observation was thought to support the idea that degradation generally happens in a 5' to 3' direction (Selinger et al, 2003). We confirmed the RNA-seq measurements by quantitative PCR (qPCR) (Supplementary Fig. S2, S3).

A closer examination of the data revealed that the differences in 5' and 3' abundances is due to a delay in the appearance of degradation (Fig. 1b), while the exponential decay was similar across the operon. This argues against the existence of differential lifetime along the transcript. Instead, it suggests that transcription elongation is visible in the data. Similar decay curves have been reported in literature, and are thought, but not confirmed, to be the result of elongating RNAPs (Chow & Dennis, 1994). To check for the contribution of elongating RNAPs, we compared the RNA degradation measured by adding an RNAP elongation inhibitor, streptolydigin, to RNA degradation measured by adding rifampicin. By qPCR, we observed that the 5' end of our 12kb transcript (Fig. 1c, d) showed degradation immediately after the cells were treated with rifampicin, similar to the RNA-seq data. Furthermore, the downstream segments had stable RNA abundance for an amount of time proportional to their distance from the 5' end (Fig. 1d). These positional differences in the onset of declining RNA abundance were not apparent when the cells were treated with streptolydigin (Fig. 1e), confirming that the products of elongating RNAPs are visible in the rifampicin data. The 5' to 3' differences in RNA degradation are caused through RNA synthesis by RNAPs that had initiated transcription prior to rifampicin addition.

Therefore, we fitted our data differently from previous papers by including a delay before the exponential decay (Fig. 1b). Each delay represents the time when the last average RNAP went through the transcript segment. Because most promoters in bacteria are weak, there is usually one or no RNAP transcribing each transcription unit per cell (Bon et al, 2006; Proshkin et al, 2010). The last average RNAP is thus representative of RNAPs in their physiological condition. Then, linear fitting of the delay time across the length of the RNA was performed to extract the elongation rate (Fig. 1f). Unlike previous methods of measuring RNAP elongation rates that relied on specific small molecule induction of a single RNA (Rose et al, 1970; Vogel & Jensen, 1994), our approach, using RNA-seq and rifampicin, is basically capable of measuring the native elongation rate of all actively synthesized RNAs in the cell.

The 5' to 3' differences in RNA abundance is general across the genome, evident from the averaged genome view (Fig. 1g). It also confirms that the averaged delay before start of RNA degradation scales proportionally to the distance from the 5' end, and that the relative abundance on average decreases at a similar rate from the 5' end to the 3' end. The expected delay of rifampicin action is less than30 seconds (Pato & Von Meyenburg, 1970).

Global behavior of RNA degradation

The RNA lifetimes extracted from our fitting, which corrects for the delay in RNA degradation, showed that all positions along an RNA generally have a lifetime similar to other segments on the same transcript (Fig. 2). For all mono-cistronic, and most poly-cistronic RNAs that we measured, consecutive segments along a transcript have lifetimes that are within two-fold of each other (Supplementary Table S1). We found examples of poly-cistronic transcripts that had different lifetimes across the operon. In agreement with a previous report, we found that the malEFG RNA has segments with different lifetimes (Newbury et al, 1987) (Supplementary Table S3). We also found five other poly-cistronic RNAS with different segment lifetimes (>2-fold difference) that have not been reported in literature (Supplementary Table S3). We also note that in a few cases, all segments of some very long transcripts, smtA-mukFEB and hybOABCDEFG, had a single exponential lifetime that was much shorter than the time estimated for the synthesis of the full length transcript from the measured elongation rate, which suggests co-transcriptional degradation for at least these transcripts (discussed later).

We average the lifetimes of the 300nt segments along each transcript, and report these lifetimes (Fig. 3a, Supplementary Table S4). Our measurements gave an average lifetime of 2.5min, shorter than the previous genome-wide microarray measurements which have an average lifetime of 6min for genes, probably because of our correction for residual RNAP activity (Bernstein et al, 2002; Selinger et al, 2003). The distribution of lifetimes of operons and genes are similar (Supplementary Fig. S4). The previously reported correlation between gene function and lifetime (Bernstein et al, 2002; Selinger et al, 2003) was not observed in our data (Supplementary Table S5).

RNA chain elongation rates genome-wide

We were able to measure RNA chain elongation rates for 263 transcripts of >1200nt length that have enough data for analysis, producing a survey of native mRNA chain elongation rates (Fig. 3b, Supplementary Table S6). Our average elongation rate of 25nt/s agrees with previous measurements of native mRNA transcription elongation rates. The trp operon was previously measured at 27nt/s, and bulk mRNA elongation rates were estimated at 29nt/s (Rose et al, 1970; Manor et al, 1969). A few of the mRNA transcription rates are in the range of speeds typically associated with ribosomal RNA transcription (~70nt/s), and may be the consequence of regulation by transcriptional elongation factors (Roberts et al, 2008). Many other elongation rates have been measured by cloning the gene of interest onto a plasmid behind an inducible promoter (Proshkin et al, 2010; Vogel & Jensen, 1994). However, it has shown that plasmid-based measurements of elongation rates can be changed by varying the concentration of IPTG (Epshtein & Nudler, 2003). Our measurements rely on native RNAP initiation rates, which are directly regulated by the cell.

Comparison of exponential and stationary phase RNA dynamics

We also performed measurements for cells growing in stationary phase (OD600nm=1.7). We found that lifetimes were still generally constant across transcripts, and averaged ~4.5min, with a smaller deviation than for exponential phase lifetimes (Fig. 4a, Supplementary Table S2, S7).

Elongation rates were slightly slower than those measured for RNAs expressed in exponential phase (Fig. 4b, Supplementary Table S8). The spike-in RNAs used for normalization also allowed us to estimate that average copy number of measured RNA was ~50 times lower in the stationary phase cells. Furthermore, knowing both steady-state abundance (time zero in our measurement) and RNA lifetimes allowed us to calculate synthesis rates, which revealed that the change in RNA abundance was better explained by changes in transcription initiation rates (R^2=0.67-0.88) than in lifetimes (R^2=0.05-0.13) (Fig. 4c, d).

A dynamic model incorporating transcription, translation and RNA degradation in *E.coli*

We observed that some transcripts, like the smtA-mukFEB and hybOABCDEFG operons mentioned earlier, are so long that the amount of time needed to synthesize the full length RNA is longer than the lifetime of the 5' end of the transcript. In other words, the 5' end is degraded before the 3' end is synthesized, which is called co-transcriptional degradation. The degradation of the 5' end of an RNA before the synthesis of its 3' end has previously been shown for the trp and lac operons (Cannistraro & Kennell, 1985; Morikawa & Imamoto, 1969; Morse et al, 1969). However, there has not been a systematic study of co-transcriptional degradation. For the 263 transcripts with a measured elongation rate in our dataset, we found that 88 (33%) in exponential phase have synthesis time that is longer than the lifetime of their 5' end (synthesis time/lifetime >1) (Fig. 5a, Supplementary Table S9), suggesting that co-transcriptional degradation is potentially common.

The coordination between transcription, translation and degradation in the cell is generally described. Transcription is initiated by RNAP binding, and the RNAP proceeds at a constant speed. As soon as the ribosome binding site is synthesized, ribosomes can bind the mRNA to start translation. Enzymes with RNase activity also target the 5' end of the mRNA, and compete with ribosomes to bind and degrade. The exact temporal coordination is, however, not well understood.

To better understand the prevalence of co-transcriptional degradation, we start with the simplest single compartment model. Although some RNases are limited by their membrane localization, the quantitative information needed to build a more complicated model is not yet available. Hence we use a minimal model which is consistent with our quantitative data, hence the following scenario based on our experimental observations is modeled: the initiation of transcription occurs with a single rate-limiting step with rate k and the elongation speed, v, is assumed to be uniform across the same transcript, which is consistent with our data. Since the lifetimes of RNA segments do not vary along the same transcript, we assume the binding of RNase at the 5' end of the transcript, competing with ribosomes, serves as the single rate limiting step with rate d for the whole RNA degradation process. This assumption does not distinguish between the two possible 5'-dependent degradation enzymes (supplementary). We further assume that RNase that is bound to the transcript closely trails behind ribosomes, rapidly cleaving the ribonucleic acid after the last ribosome molecule, instead of being able to cut between ribosomes (Fig. 5b) (Schneider et al, 1978). Although mRNA can be cleaved in between ribosomes to rescue stalled ribosomes, the occurrence of such an event is low, 0.4%, and does not affect our model (Moore & Sauer, 2007). Lastly, we assume the elongation

velocity of the ribosome on the transcript is similar to that of the RNAP, which has been observed experimentally (Proshkin et al, 2010). This ensures that the lifetime of the 5' and 3' ends are similar.

Next, we consider two different mechanisms: one is co-transcriptional degradation, which allows RNase to bind RNA at/near the 5' end at any time after RNA synthesis has started (Fig. 5b); while the other one is post-transcriptional degradation, which only allow RNase to bind at/near the 5' end only after the entire transcript is completely synthesized.

The co-transcriptional degradation model is solved by a delayed single exponential decay curve (supplementary). This means that RNA abundance exhibits a position-dependent delay followed by single exponential decay at all segments of an RNA, similar to our experimental observations. This is in contrast to the post-transcriptional model, which has a more complex solution. It has a delay (equilibrium), followed by linear decay and then finally an exponential decay. Like in the co-transcriptional model, the delay is dependent on the distance of the segment from the 5' end, while the linear decay is due to waiting for RNA synthesis to complete before degradation can commence (supplementary).

In the fastest case of post-transcriptional degradation, i.e. the binding rate *d* of RNase is extremely high, RNA instantaneously degrades the moment it is completely synthesized, the degradation of the RNA abundance at the 5' end is exactly linear (supplementary). When the binding rate *d* is finite, the decay curve of the RNA abundance at the 5'end should always be slower than that of instantaneous degradation. Therefore, in such a post-transcriptional case, the RNA lifetime at the 5'end should be greater than the synthesis time for the whole transcript.

On the contrary, when the ratio of synthesis time/lifetime>1, there are moments where the 5' end RNA abundance is lower than that expected from the fastest post-transcriptional decay mechanism , e.g. the smtA-mukFEB operon (Supplementary Fig. S5). It indicates that the degradation of these transcripts is co-transcriptional. Even though we do not have a synthesis time/lifetime ratio larger than 1 for all observed RNAs to identify co-transcriptionally degraded RNAs, by Occam's Razor, we think that co-transcriptional degradation is a simple mechanism that can explain observed degradation patterns.

We caution that a co-transcriptional degradation mechanism does not imply that every single produced RNA in the cell is being synthesized and degraded simultaneously. Both ribosome binding, which is proposed to be protective (Schneider at el, 1978), and RNase binding, which initiates degradation, are stochastic events on the molecular level. While there are instances where the RNase binds before the RNAP is done synthesizing the RNA, there are other instances when ribosomes happen to consecutively bind so that the full length RNA is synthesized before RNase has a chance to bind and initiate degradation. From our model, we can calculate the probability of having a transcript that is degraded before the full transcript has been synthesized to be $1-e^{dL/v}$ (supplementary), which is dependent on the ratio of RNA synthesis time to RNA lifetime. For transcripts with a synthesis time/lifetime (dL/V) > 1, the probability of each RNA molecule being co-transcriptionally degraded is >60%. If the co-transcriptional degradation model holds genome-wide, even for transcripts that have a synthesis time/lifetime <1, we can still calculate that a sizeable fraction of RNA that is simultaneously

being synthesized and degraded. For instance, with a synthesis time/lifetime of 0.5, 40% of RNAs are degraded before completely synthesized.

The identification of more transcripts, besides the lac and trp operons, that are likely to be co-transcriptionally degraded, has increased the amount of evidence supporting genome-wide co-transcriptional degradation. Moreover, co-transcriptional degradation likely does not happen exclusively to very long transcripts as in literature. The generality of constant life time across the same transcript implies that RNAs are at least degraded in a 5' to 3' direction. It would thus follow that determinants of RNA lifetime can be expected to preferentially exist on the 5' proximal end of the transcript. This conclusion ties in well with the observation that the 5' end contains important determinants that regulate RNA lifetime (Arnold et al, 1998).

Discussion

The fundamental observation that RNA decays exponentially with a single lifetime, established in early experiments on a few RNAs, indicates the existence of a single rate-determining step (Mosteller et al, 1970; Schwartz et al, 1970). This rate-determining step is thought to initiate degradation (Belasco, 2010; Carpousis, 2007). In our study, we find that segments within the same transcript genome-wide share a similar rate constant (Fig. 2). This suggests that the effect of the rate-determining step is shared by all parts of the RNA, and that transcription, translation and degradation work together in a defined manner to achieve such an outcome genome-wide. While our work does not comment on the specific mechanism of the decay pathway, it provides limits to the models of RNA degradation on a genome-wide scale, which has not yet been achieved in literature. Following the example of earlier studies of RNA degradation that have been more holistic, we integrate the kinetics of transcription, translation and degradation into a mathematical model that predicts our RNA-seq results. This detailed modeling of RNA dynamics in *E.coli* supports the generality of our assumptions, which are based on data obtained from gene-specific studies.

A consequence of exponential decay that has been only partially addressed in literature is that RNA can be broken down at any time after the initiation of synthesis, even before the entire transcript is synthesized (Cannistraro & Kennell, 1985; Morikawa & Imamoto, 1969; Morse et al, 1969). This means that RNAs can be co-transcriptionally degraded. This is counter-intuitive since full length RNA is the major observed mRNA species in the cell (Llopis et al, 2010), which has been shown to diffuse and localize to different parts of the cell in some cases (Nevo-Dinur et al, 2011). Moreover, the existence of sub-cellular compartmentalization, with separate zones of DNA, ribosomes and RNases that RNA has to diffuse through, make co-transcriptional degradation less plausible (Mackie, 2013). However co-transcriptional degradation is a stochastic process, not a deterministic one, which allows full length and partially degraded mRNA to co-exist via the same mechanism. In our work, we proposed a dynamic model to distinguish the mechanisms of co-transcriptional and post-transcriptional RNA degradation that allows us to estimate the distribution of full length and already degraded RNAs before the completion of transcription for a single transcript species.

In summary, RNA degradation was measured with high time and base resolution in exponentially-growing and stationary phase cells, leading to the clear observation of a lag in

RNA degradation that is dependent on the distance from the 5’ end after addition of transcription inhibitor, rifampicin. This allowed us to confirm simultaneous observation of transcription elongation and degradation, resulting in a first survey of native RNAP elongation rates, and more accurate extraction of RNA lifetimes along RNAs genome wide. The observed constant exponential lifetime along transcripts suggests 5 ' to 3‘ directed RNA degradation, and supports the mechanism of co-transcriptional degradation for each transcription unit. Taken together, the parameters provided by our measurement highlights the tight coordination between transcription and RNA degradation in *E.coli* cells.

Acknowledgements:

This work was supported by NIH grant GM096450 to X.S.X. and Xiaowei Zhuang. K.S. was supported by a Postdoctoral Fellowship for Research Abroad from the Japanese Society for the Promotion of Science and a fellowship from The Uehara Memorial Foundation. H.G. is supported by NSFC 10901040, 21373021 and the Foundation for the Author of National Excellent Doctoral Dissertation of China (No. 201119)

Contributions:
H.C., K.S. and X.S.X. conceived the project. H.C. and K.S. designed and executed experiments, and analyzed results. K.S., H.G., and X.S.X. conceived the model. H.C., K.S., H.G. and X.S.X. wrote the manuscript.

Conflict of Interest:
The authors declare that they have no competing financial interests.

References

Ameur A, Zaghlool A, Halvardson J,Wetterbom A, Gyllensten U, Cavelier L & Feuk L (2011) Total RNA Sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat. Struct. Mol. Biol.* **18**, 1435-1440

Arnold, TE, Yu J & Belasco JG (1998) mRNA stabilization by the ompA 5' untranslated region: two protective elements hinder distinct pathways for mRNA degradation. *RNA* **4**, 319-330.

Belasco JG (2010) All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nat. Rev. Mol. Cell. Biol.* **11**, 467-478

Bernstein JA, Khodursky AB, Lin P-H, Lin-Chao S & Cohen SN (2002) Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc. Natl. Acad. Sci.* **99**, 9697-9702

Bon M, McGowan S & Cook PR (2006) Many expressed genes in bacteria and yeast are transcribed only once per cell cycle. *FASEB J.* **20** E1071-E1074

Cannistraro V & Kennell D (1985) Evidence that the 5' end of lac mRNA starts to decay as soon as it is synthesized. *J. Bacteriol.* **161**, 820-822

Carpousis AJ (2007) The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E. *Annu. Rev. Microbiol.* **61**, 71-87

Chow J & Dennis PP (1994) Coupling between mRNA synthesis and mRNA stability in *Escherichia coli. Mol. Microbiol.* **11**, 919-931

Churchman LS & Weissman J (2011) Nascent Transcript Sequencing Visualizes Transcription at Nucleotide Resolution. *Nature* **469**, 368-373

Deneke C, Lipowsky R & Valleriani A (2013) Effect of ribosome shielding on mRNA stability. *Phys. Biol.* **10***,* 046008

Epshtein V & Nudler E (2003) Cooperation Between RNA Polymerase Molecules in Transcription Elongation. *Science* **300**, 801-805

Golding I & Cox EC (2004) RNA dynamics in live *Escherichia coli* cells. *Proc. Nat. Acad. Sci.* **101**, 11310-11315

Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martínez C, Fulcher C, Huerta AM, Kothari A, Krummenacker M, Latendresse M, Muñiz-Rascado L, Ong Q, Paley S, Schröder I, Shearer AG, Subhraveti P, Travers M, Weerasinghe D, Weiss V, *et al.* (2013) EcoCyc: fusing model organism databases with systems biology. *Nucl. Ac. Res.* **41** D605-D615

Kristoffersen SM, Hasse C, Weil MR, Passalacqua KD, Niazi F, Hutchison SK, Desany B, Kolstø A-B, Tourasse NJ, Read TD & Økstad OA (2012) Global mRNA decay analysis at single nucleotide resolution reveals segmental and positional degradation patterns in a Gram-positive bacterium. *Genome Biol.*, Vol. 13: R30

Liu B, Deikus G, Bree A, Durand S, Kearns DB & Bechhofer DH (2014) Global analysis of mRNA decay intermediates in Bacillus subtilis wild-type and polynucleotide phosphorylase-deletion strains. *Mol. Microbiol.*

Llopis, PM, Jackson AF, Sliusarenko O, Surovtsev I, Heinritz J, Emonet T & Jacobs-Wagner C (2010) Spatial organization of the flow of genetic information in bacteria. *Nature* **466**, 77-82

Mackie GA (2013) RNase E: at the interface of bacterial RNA processing and decay. *Nat. Rev. Microbiol.* **11**, 45 – 57

Manor H, Goodman D & Stent GS (1967) RNA chain growth rates in *Escherichia coli. J. Mol. Biol.* **39**, 1-29

Moore SD & Sauer RT (2007) The tmRNA System for Translation Surveillance and Ribosome Rescue. *Annu. Rev. Biochem.* **76**, 101-124

Morikawa N & Imamoto F (1969) Degradation of Tryptophan Messenger: On The Degradation of Messenger RNA for Tryptophan Operon in *Escherichia coli. Nature* **223**, 37-40

Morse, DE, Mosteller R, Baker RF & Yanofsky C (1969) Degradation of Tryptophan Messenger: Direction of *in vivo* Degradation of Tryptophan Messenger RNA–A Correction *Nature* **223**, 40-43

Mosteller RD, Rose JK & Yanofsky C (1970) Transcription Initiation and Degradation of trp mRNA. *Cold Spring Harb. Symp. Quant. Biol.* **35**, 461-466

Nevo-Dinur K, Nussbaum-Shochat A, Ben-Yehuda S & Amster-Choder O (2011) Translation-Independent Localization of mRNA in E. coli. *Science* **331**, 1081-1084

Newbury SF, Smith NH & Higgins CF (1987) Differential mRNA Stability Controls Relative Gene Expression within a Polycistronic Operon. *Cell* **51**, 1131-1143

Pato ML & Von Meyenburg K (1970) Residual RNA Synthesis in *Escherichia coli* after Inhibition of Initiation of Transcription by Rifampicin. *Cold Spring Harb. Symp. Quant. Biol.* **35**, 479-504

Proshkin S, Rahmouni A R, Mironov A & Nudler E (2010) Cooperation Between Translating Ribosomes and RNA Polymerase in Transcription Elongation. *Science* **328**, 504-508

Roberts JW, Shankar S & Filter JJ (2008) RNA polymerase elongation factors. *Annu. Rev. Microbiol.* **62**, 211-233

Rose JK, Mosteller RD & Yanofsky C (1970) Tryptophan messenger ribonucleic acid elongation rates and steady-state levels of tryptophan operon enzymes under various growth conditions. *J. Mol. Biol.* **51**, 541-550

Ryals J, Little R & Bremer H (1982) Temperature Dependence of RNA Synthesis Parameters in *Escherichia coli. J. Bacteriol.* **151**, 879-887

Schneider E, Blundell M & Kennell D (1978) Translation and mRNA decay. *Molec. gen. Genet.*, **160**, 121-129

Schwartz T, Craig E & Kennell D (1970) Inactivation and Degradation of Messenger Ribonucleic Acid from the Lactose Operon of *Escherichia coli*. *J. Mol. Biol.* **54**, 299-311

Selinger DW, Saxena RM, Cheung KJ, Church GM & Rosenow C (2003) Global RNA Half-Life Analysis in *Escherichia coli* Reveals Positional Patterns of Transcript Degradation. *Genome Res.* **13**, 216-223

Taggart AJ, DeSimone AM, Shih JS, Filloux ME & Fairbrother WG (2012) Large-scale mapping of branchpoints in human pre-mRNA transcripts *in vivo*. *Nat. Struct. Mol. Biol.* **19**, 719-721

Vogel U & Jensen KF (1994) The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *J. Bacteriol.* **176**, 2807-2813.

Von Gabain A, Belasco JG, Schottel JL, Chang AC & Cohen SN (1983) Decay of mRNA in *Escherichia coli*: Investigation of the fate of specific segments of transcripts. *Proc. Natl. Acad. Sci.* **80**, 653-657.

Yarchuk O, Jacques N, Guillerez J & Dreyfus M (1992) Interdependence of Translation, Transcription and mRNA Degradation in the lacZ Gene. *J. Mol. Biol.* **226**, 581-596

Figure 1

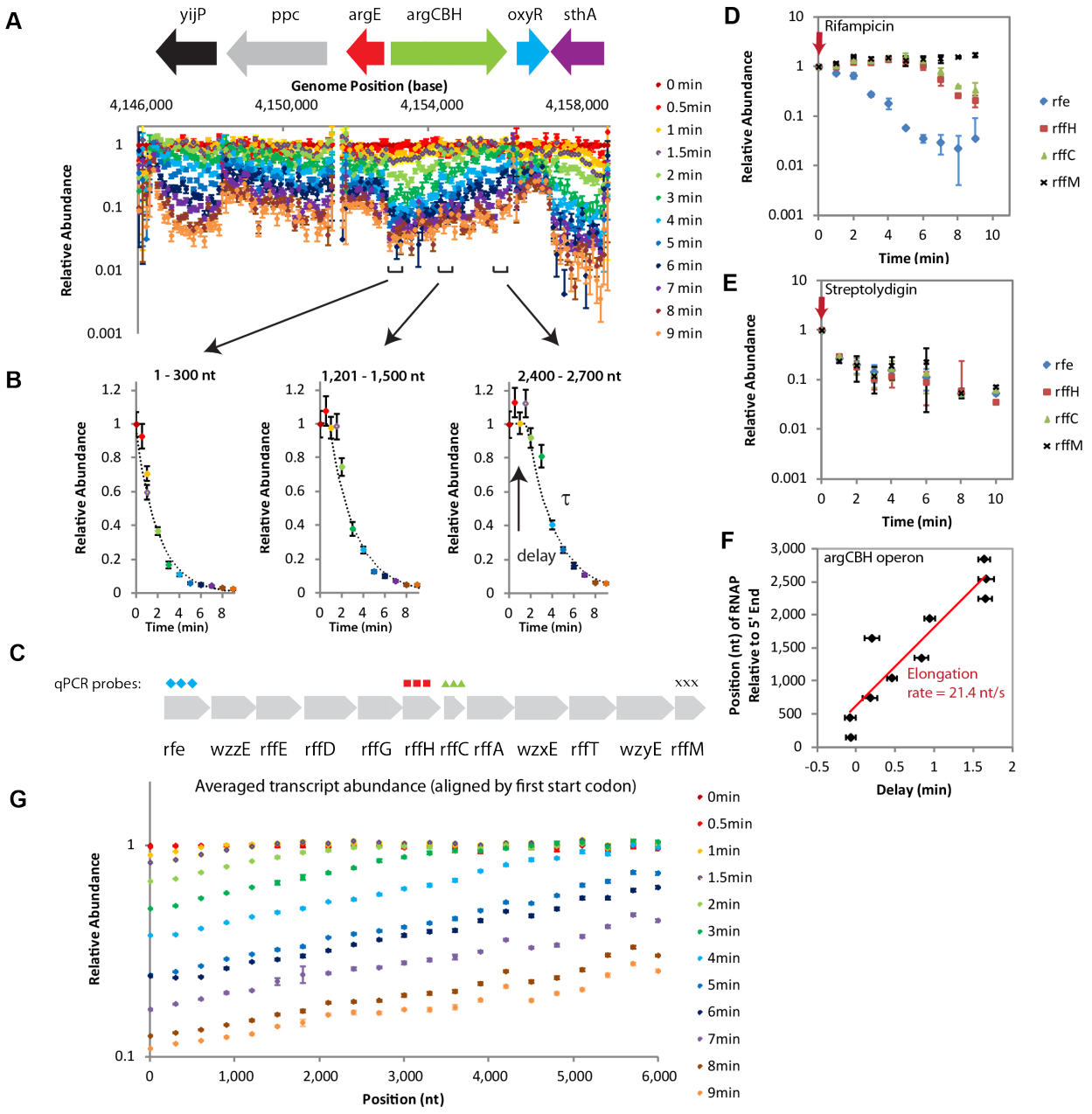


Figure 1: RNA-seq data allows detailed study of RNA synthesis and degradation genome-wide.

(a) RNA abundance data over time of exponential phase MG1655 cells shows trends correlating with direction of RNA synthesis. Representative normalized RNA abundance measured by RNA-seq (one experiment) for each time point is shown with corresponding genome annotation and root N error. Segments where the data exhibits a similar degradation pattern correspond

with the annotated operons. The small RNA, oxyS, has been omitted from the genomic annotation due to its small size.

(b) RNA abundance data from segments of the argCBH operon plotted against time reveal a delay in degradation. The data was fitted to extract the delay before RNA abundance decay, and the exponential lifetime, τ.

(c) 4 qPCR probes were designed against the 12kb rfe-wzzE-rffEDGHCA-wzxE-rffT-wzyE-rrfM transcript for the experiment shown in (D) and (E), which confirm that elongating RNAPs are responsible for the delay before RNA abundance decays.

(d) qPCR experiment in exponential phase AS19 cells treated with rifampicin, a RNAP initiation inhibitor, replicates trend observed in RNA-seq data for exponential phase MG1655 cells treated with rifampicin. RNA in the 5'-most position, rfe, declines in abundance immediately. Mid-transcript positions (rffH and rffC) maintained steady-state levels of RNA abundance for ~6min before showing degradation. The 3'-most position, rffM, maintained steady-state levels of RNA for the entire duration of the experiment. Data is from 2 biological samples with duplicates and error bars show s.d.

(e) qPCR experiment in exponential phase AS19 cells with streptolydigin, a RNAP elongation inhibitor, does not replicate trend observed in RNA-seq data for exponential phase MG1655 cells treated with rifampicin. RNA abundance at all probe positions declines immediately after streptolydigin addition. Data is from 2 biological samples with duplicates and error bars show s.d.

(f) The RNAP elongation rate, assumed constant and measured at 21.4nt/s (s.e. = 0.5nt/s), on the argCBH operon is extracted from the delay in degradation along the operon.

(g) Averaged genome-wide view of data. All RNAs were aligned by their first start codon, and averaged to obtain a genome-wide view of RNA abundance. The trends observed in figure 1a are reproduced on the average genome-wide scale, suggesting that they are typical.

Figure 2

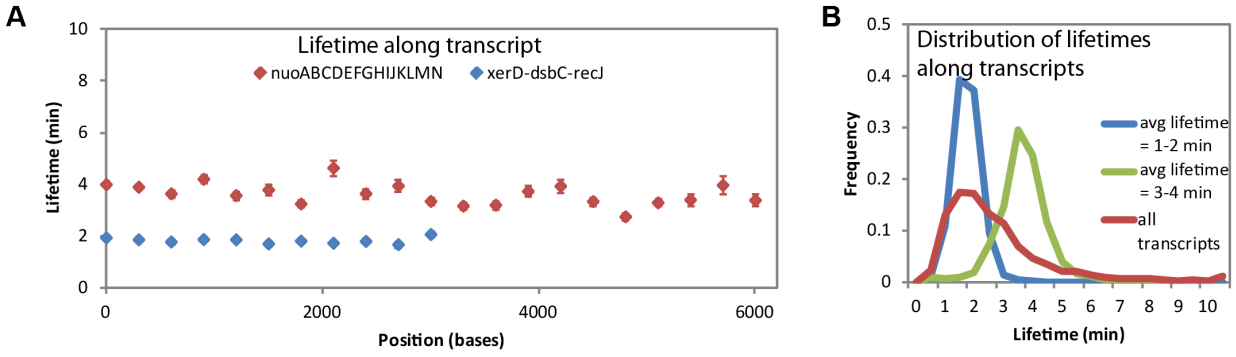


Figure 2: Lifetime is generally constant across RNAs.

(a) The lifetime along all segments of 3371-nt transcript, xerD-dsbC-recJ, is ~2min (s.e. shown). The lifetime along most segments of the 14kb nuo operon is ~4 min (s.e. shown).

(b) The distribution of lifetimes of all segments from transcripts longer than 600nt with an average lifetime of 1-2 min (blue), 3-4 min (green) are different from each other and they are more narrow compared to the distribution of lifetimes of all transcripts (red).

Figure 3

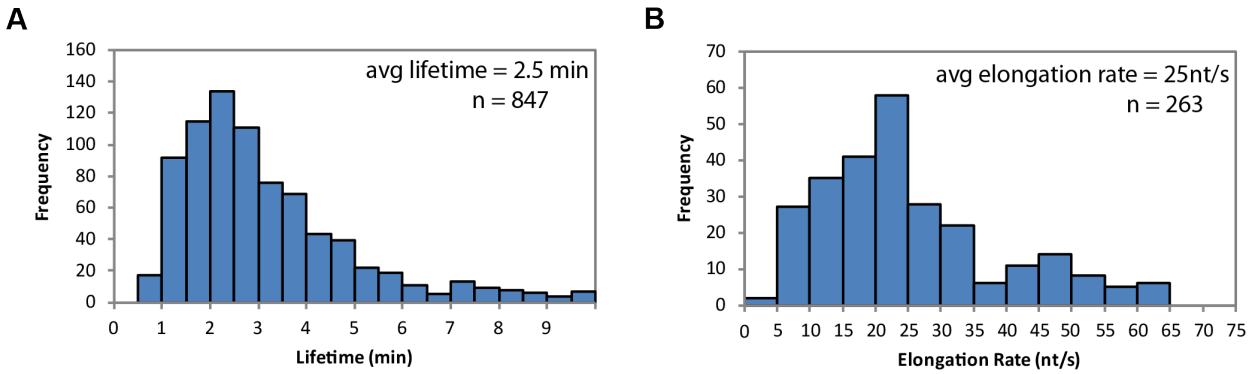


Figure 3: Distributions of lifetimes and elongation rates measured for exponentially growing cells.

(a) The lifetimes of 847 transcripts were measured, with an average lifetime of 2.5 min (standard deviation 2.5min).

(b) The elongation rates for 263 transcripts longer than 1200nt were measured. The average is 25nt/s (standard deviation 14nt/s).

Figure 4

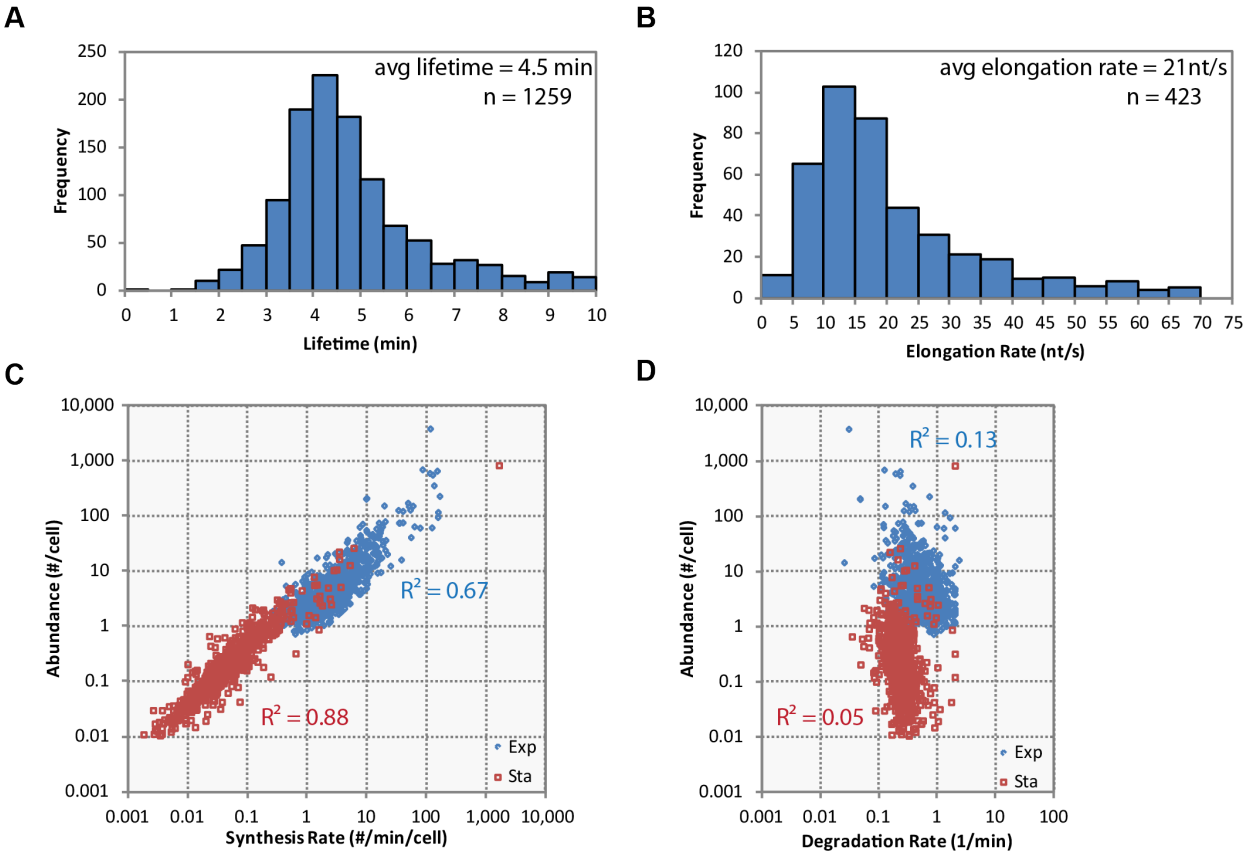


Figure 4. Comparison of stationary phase measurements with exponential phase measurements.

(a) Stationary phase lifetimes averaged 4.5 min (standard deviation = 2.4min). The lifetimes of 1259 transcripts in stationary phase *E.coli* were measured. Relative to exponential phase lifetimes which have a standard deviation as large as the mean (2.5 min/2.5min = 1), the distribution of lifetimes in stationary phase is narrow (2.4min/4.5min = 0.53)

(b) Stationary phase elongation rates average 21 nt/s (standard deviation = 14nt/s).The elongation rates for 423 transcripts longer than 1200nt for stationary phase *E.coli* were measured.

(c) RNA abundance can be explained by RNA synthesis rate for both exponential and stationary phase cells. RNA abundance and transcription initiation rates correlate well for the exponential phase measurements (R^2 = 0.67), and have an even stronger correlation in stationary phase (R^2 = 0.88).

(d) RNA abundance is poorly correlated to degradation rate (1/lifetime) in exponential and stationary phase *E.*coli. Degradation rates do not correlate well with RNA abundance in exponential (R^2=0.13) and correlates more poorly in stationary phase (R^2=0.05), in contrast to the correlation of RNA abundance with synthesis rates.

Figure 5

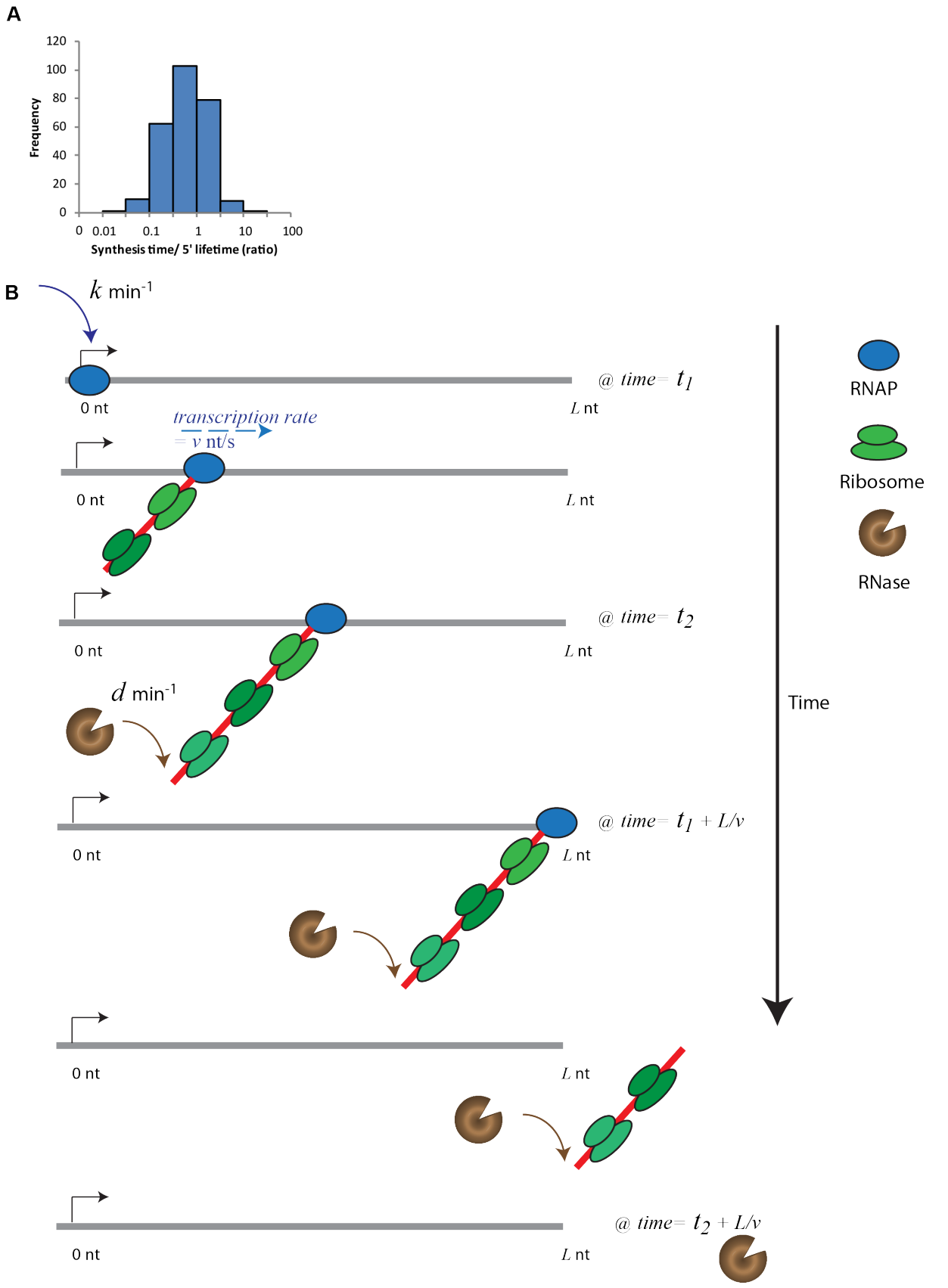

A
120
100
80
60
40
20
0
Frequency
0
0.01
0.1
1
10
100
Synthesis time/ 5' lifetime (ratio)
B
k min-1
@ time= t1
0 nt
L nt
transcription rate
= v nt/s
0 nt
L nt
@ time= t2
0 nt
L nt
d min-1
@ time= t1 + L/v
0 nt
L nt
0 nt
L nt
@ time= t2 + L/v
0 nt
L nt
Time
RNAP
Ribosome
RNase

Figure 5: Co-transcriptional degradation of RNA.

(a) 88 out of the 263 RNAs with a measured elongation rate in exponentially growing cells have a synthesis time (estimated from the length of the RNA and the elongation rate) that is longer than the lifetime of the 5' end.

(b) Model of temporal coordination of co-transcriptional degradation on RNA transcript of length *L* explains degradation curves and how synthesis time can be longer than lifetime. Initially at $T = t_1$, RNAP binds onto the genome with a single rate of transcriptional initiation, k min^{-1}, and proceeds to generate the RNA transcript with a constant elongation rate, v nt/s. At $T = t_2$, before full length RNA is generated, RNase binds to the 5' end in a rate-determining step with rate *d* min^{-1}, which is affected by the competition between ribosome and RNase binding. Any subsequent binding is not rate-limiting. The RNase cannot degrade the RNA more quickly than the movement of the last ribosome (dark green) on the transcript. At $T = t_1 + L/v$, the RNAP generates the end of the RNA transcript. Because the RNase has already bound and started to degrade the transcript, the transcript length is less than *L* nts at $T = t_1 + L/v$.

1. Methods

E.coli strains K12 MG1655 or AS19 (gift of Peter Nielson, University of Copenhagen, Denmark) were grown in LB-Miller (EMD) at 30°C with shaking at 250 rpm. Overnight cultures were diluted 1:250 into fresh media and grown to OD_{600nm} 0.3 for the exponential phase sample, and OD_{600nm} 1.7 for the stationary phase sample. The doubling time during steady-state growth in exponential phase was ~40 min.

Rifampicin (Sigma R7382, previously R5777), dissolved in DMSO, was added to MG1655 and AS19 cultures at a final concentration of 500 ng/µl, and aliquots of the cultures were quenched in 10% volume of cold 9:1 ethanol: phenol solution at specified time points. The cells were pelleted and washed with cold PBS before storage in -80°C.

Streptolydigin (ChemCon GmbH) was dissolved in water, and added to AS19 cultures at a final concentration of 100ng/µl.

1.2 RNA Purification

Bacterial pellets were thawed on ice. Cells were first resuspended and incubated in 1mg/ml of lysozyme (Sigma) prepared in T.E. buffer, and then completely lysed by the addition of 1 volume cell lysis buffer (Puregene). Spike-in RNAs were introduced at this stage.

The samples were extracted with acidic phenol/chloroform (Sigma) 2 – 4 times, and then once with chloroform:isoamyl alcohol (24:1, Sigma). The RNA was then precipitated with addition of $1/10^{th}$ volume 5M NaCl and 1 volume isopropanol on ice or in -20°C. RNA was pelleted and washed twice with 70% ethanol, resuspended to 100-200ng/µl in nuclease-free water (Ambion or Cellgro), then digested with DnaseI (NEB) to remove contaminating DNA. The RNA was then purified as described. After resuspending in water, the RNA was used immediately.

1.3 RNAseq protocols and rRNA removal

rRNA was removed by Epicentre's Ribo-Zero Kit (for gram negative bacteria) following manufacturer's protocol. RNA was finally collected by RNA Clean-up and Concentration column (Zymo Research).

The RNA was fragmented, and prepared for Illumina sequencing as previously reported (Taniguchi, 2010). The concentration of the libraries were determined by qPCR (Applied Biosystems Fast 7500 Real-Time PCR System)according to Illumina's protocol, substituting DyNamo HS qPCR mix (Thermo Fisher, previously New England Biolabs) for the Kappa qPCR mix. Samples were run on an updated Genome Analyzer II, or a HiSeq2000. The raw data is available from NCBI Short Reads Archive (BioProject PRJNA258398).

1.4 Data analysis

RNAseq reads sorted by barcodes were mapped onto MG1655 (NC_000913.2) using Bowtie 0.12.7. The mapped reads were then collated and normalized across time points by the average of the total number of spike in reads, ssrA, ssrS, and rnpB, before normalizing by abundance at

time 0min. Within transcription units, the data was broken into 300 nucleotide (nt) bins. Transcription units that have overlapping annotations were not included in the dataset (Keseler et al, 2012). The half life for each fragment was obtained by fitting in Igor using the parameters

$$\text{if } 0 < x < P_i\text{: } y = 1$$

$$\text{else, } y = \exp(-(x - P_i)/ T_i)$$

where x is time (min), P_i is the polymerase passage time (delay) for fragment i, and T_i is the lifetime of fragment i. Elongation rates were fitted linearly using P_i for each fragment by a custom Igor program. To ensure proper fitting, only transcripts longer than 1200nt (giving 4 data points) were analyzed. Transcription units with large uncertainty in fitting ($error_{rate}$/rate> 1) or nonsensical results (0<$error_{rate}$/rate) were discarded. Other manipulations were performed in Excel. The relationship between rRNA chain elongation rate and culture temperature is known (Ryals et al, 1982), and may be used to estimate mRNA chain elongation rate at other temperatures.

2. Supplementary

2.1 Linearity of sequencing data by spike-in

Spike in RNA was prepared from phiX174 (NC001422). The genes D, F, G and H, and fragments 190 (pos 4501 - 4680) and 290 (pos 4031 - 4320) were cloned from ssDNA (New England Biolabs) with the addition of a T7 promoter. RNA was produced by *in vitro* transcription using Ambion's MegaScript kit, according to manufacturer's protocol. RNA integrity was verified by running on a TBE gel (BioRad). The in vitro transcribed RNAs were added to samples at the following ratios:

Gene D – 0.1 RNA/cell

Gene F – 1 RNA/cell

Gene G – 1 RNA/cell

Gene H - 0.01 RNA/cell

Fragment 190 – 100 RNA/cell

Fragment 290 – 10 RNA/cell

The number of cells in the sample was determined by counting the number of colony forming units on an LB agar plate. The RNA-seq reads were linear for the range of spike-ins (Supplementary Fig. S1).

2.2 Reproducibility of data by qPCR (exponential phase and stationary phase)

RNA from E.coli growing in exponential or stationary phase were collected as described in Methods, and processed. RNA abundance of two biological duplicates was measured by qPCR and compared with the corresponding RNA-seq data. The control RNAs were selected from

RNA-seq data as a representative of their class – for instance RNA length, RNAP transcription elongation rate, and RNA lifetime. In some instances, traces can look similar but give different fitting results (Supplementary Fig. S2A: arnBCADTEF-1). The delay and lifetimes measured by both methods were compared by a paired Student t test, and were not significantly different, confirming the RNAseq measurement method. (Supplementary Fig. S2, S3).

2.3 Differentially regulated genes

The lifetimes of segments within a transcript were analyzed to identify differentially regulated transcripts. Adjacent segments on transcription units that had a lifetime different by two-fold were identified. The final list (Supplementary Table S3) was compiled after manual verification to remove noisy data or poor fits. Of the 13 identified transcripts, MalEFG have been previously reported to have differential regulation within the transcription unit (Newbury et al, 1987).

2.4 Analysis of correlation between gene function and lifetime

Lifetimes of all 300nt bins within a transcription unit were averaged, with the exception of the 13 transcripts identified in section 2.3: Differentially regulated transcripts that were excluded. The transcript lifetime was then assigned to each gene on the transcript. Genes were grouped by gene categories obtained from J. Craig Venter Institute's Comprehensive Microbial Resource (Davidsen et al, 2010). The distributions were compared using a Kolmogonov-Smirnov test.

Of the 17 gene categories, Mobile and Extrachromosomal Elements was ignored as it contained only 1 gene. 5 categories had a different distribution than the starting distribution ($P<0.05$), of which 4 biased towards longer lifetimes, while one was biased towards shorter lifetimes (Supplementary Table S5).

Comparing the distribution of lifetime of genes instead of that of transcripts, the gene lifetime distribution has more entries with longer lifetimes (Supplementary Fig. S4). The distribution of gene lifetime, compared to transcript lifetime, is not significantly different by a Kolmogonov-Smirnov test ($D=0.0278$, p-value = 0.85).

Models:

Co-transcriptional degradation

We will first briefly summarize the current literature related to mRNA processes, and use the concepts to build a model that will predict the mRNA degradation curves that we have observed.

mRNA synthesis is initiated by the binding of RNAP onto DNA, and chain synthesis is widely assumed to proceed at a constant speed (Vogel & Jensen, 1994). After the ribosome binding site has been transcribed, ribosomes can begin translating the mRNA even if the mRNA is not complete. Multiple ribosomes can translate a single mRNA simultaneously, and indeed polysomes were observed for highly expressed proteins (Warner et al, 1963). It has been

observed that ribosomes and RNAP activity proceed at similar speeds (Proshkin et al, 2010). The presence of ribosomes is proposed to be protective of mRNA – chloramphenicol, which blocks peptide chain elongation, while leaving the ribosome on the mRNA stabilizes mRNA, while kasugamycin, which prevents initiation, destabilizes mRNA (Schneider at al, 1978). Decreasing the affinity of the RBS can also reduce the lifetime of an mRNA (Yarchuk et al, 1992).

The completion of the 5' end of the mRNA also allows degradation to commence parallel to translation. The enzyme, rppH, cleaves the 5' triphosphate cap to leave a single phosphate, which RNaseE can now bind to the 5' monophophate and endonucleolytically degrade the mRNA (Deana et al 2010). Both RNaseE and rppH can be the regulator of the rate determining step in mRNA degradation (Luciano et al, 2012). Removal of the ribosome binding site by RNaseE prevents any further ribosome initiation, but does not interfere with ribosomes that have initiated.

To date, only 3'-to-5' exonucleases have been found in *E.coli*. Because of the directionality of their action, these enzymes can only act after RNaseE cleavage, or after the full length RNA has been synthesized and released. The action of these downstream degradation enzymes are thought to be very rapid as degradation intermediates are not observable unless ribonuclease II and polynucleotide phosphorylase (PNPase) are knocked out (Donovan & Kushner, 1986). In *E.coli*, PNPase is thought to form a complex with RNaseE, termed the degradosome (Carpousis, 2007).

We incorporate these details into a mathematical model and introduce several parameters. Our model is concerned with the outcome of the assumption, and not the mechanism.

mRNA is synthesized at the rate of k, which is also the rate of RNAP initiation. Consistent with many modeling papers, we define k to be a first-order rate process (Bar-Even at el., 2006; Friedman et al, 2006)

For modeling purposes, we assume that only one RNase is rate-determining, and do not distinguish between the two 5' dependent RNases. The RNase degrades from the 5' end, and thus does not cut between ribosomes, or between a ribosome and the RNAP.

The binding of RNase occurs on the mRNA's 5' end, competing with ribosome binding, and is a first-order rate process, i.e. the waiting time for RNase to bind and initiate degradation is exponentially distributed with a mean value of 1/d. d decreases with increasing [ribosome]. Similar to ribosome binding, RNase binding can occur as soon as the 5' end is generated. Because of the sequence of events, ribosomes need to follow behind RNAP, and translocate at a similar rate to mRNA transcription.

The mRNA abundance at any position *n* consists of all the transcripts that started transcription (*n*/*v*) min prior, and may have degraded during the (*n*/*v*) min time period. Here, *v* is the velocity of transcription in nts/min. Hence the steady-state mRNA abundance at position n is

$$c(n) = \int_{s'=n/v}^{+\infty} k \int_{s'-n/v}^{+\infty} de^{-ds} ds ds' = \frac{k}{d}$$

When the drug inhibiting transcription initiation is added, after time *t*

$$c(n,t) = \int_{s'=\max(t,n/v)}^{+\infty} k \int_{s'-n/v}^{+\infty} de^{-ds} ds ds' = \begin{cases} \dfrac{k}{d} & if\ t < n/v \\ \dfrac{ke^{-d(t-n/v)}}{d} & if\ t > n/v \end{cases}$$

The relative abundance is

$$f(n,t) = \frac{c(n,t)}{c(n)} = \begin{cases} 1 & if\ t < n/v \\ e^{-d(t-n/v)} & if\ t > n/v \end{cases}$$

which is a delayed single exponential.

The mean waiting time for the degradation initiation of an mRNA is 1/*d*, but the mean waiting time for the complete degradation of an mRNA is 1/*d*+*L*/*v*, where *L* is the full length of the mRNA. The mean waiting time for the complete degradation of the mRNA is larger than the transcription elongation time, which explains how it is possible to have a lifetime that is shorter than the transcription elongation time – co-transcriptional degradation. Note that *d* is the rate of RNA degradation typically used in the central dogma kinetic model, because after degradation initiation, protein synthesis cannot be started.

From this model, the probability of having a transcript that is degraded before the full transcript has been synthesized is $1\text{-}e^{dL/v}$.

Post-transcriptional degradation: where the mRNA degrades only after it has been fully synthesized.

This model is similar to the co-transcriptional model except for the timing of degradation initiation.

The mRNA abundance at any position *n* consists of all the transcripts that started synthesis *n*/*v* min ago from now before, and will not degrade until at least (*L*/*v*) min. Hence the steady-state mRNA abundance at position n is

$$c(n) = \int_{s'=(L+n)/v}^{+\infty} k \int_{s'-n/v-L/v}^{+\infty} de^{-ds} ds ds' + \int_{s'=n/v}^{(L+n)/v} k ds' = \frac{k}{d} + kL/v$$

When the drug inhibiting transcription initiation is added, after time t

$$c(n,t) = \begin{cases} \displaystyle\int_{s'=(L+n)/v}^{+\infty} k \int_{s'-n/v-L/v}^{+\infty} de^{-ds} ds ds' + \int_{s'=n/v}^{(L+n)/v} k ds' = \frac{k}{d} + kL/v & if\ t < n/v \\ \displaystyle\int_{s'=(L+n)/v}^{+\infty} k \int_{s'-n/v-L/v}^{+\infty} de^{-ds} ds ds' + \int_{s'=t}^{(L+n)/v} k ds' = \frac{k}{d} + k\left(\frac{L+n}{v} - t\right) & if\ n/v < t < (L+n)/v \end{cases}$$

Hence the relative abundance over time in this case would be a plateau (equilibrium) followed by a linear decay, and further followed by a single exponential decay.

The length for the linear region is L/v, which is the total time for ~seconds to minutes. It is hard to distinguish from the delayed single exponential decay in most transcripts. For instance, if *L*=1kb, *v*=25bp/s, *d*=0.5/min, then the linear decay is from 100% to 75%. For some extremely long transcript (e.x. *L*=10kb), the decay in the linear region would be much more distinct (100% to 23%). We do not see this decay pattern in longer transcripts.

References

Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, Pilpel Y & Barkai N (2006) Noise in Protein Expression Scales With Natural Protein Abundance. *Nat. Genet.* **38**, 636-43

Davidsen T, Beck E, Ganapathy A, Montgomery R, Zafar N, Yang Q, Madupu R, Goetz P, Galinsky K, White O & Sutton G (2010) The Comprehensive Microbial Resource. *Nuc. Acids Res.* **38**, D340-345

Deana A, Celesnik H & Belasco JG (2008) The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* **451**(7176):355-8

Donovan WP & Kushner SR (1986) Polynucleotide phosphorylase and ribonuclease II are required for Cell Viability and mRNA Turnover in *Escherichia coli*. *Proc. Nat. Ac. Sci.* **83**, 120-124

Friedman N, Cai L & Xie XS (2006) Linking Stochastic Dynamics to Population Distribution: An Analytical Framework of Gene Expression". *Phys. Rev. Lett.* **97**, 168302

Luciano DJ, Hui MP, Deana A, Foley PL, Belasco KJ & Belasco JG (2012) Differential Control of the Rate of 5'-End-Dependent mRNA Degradation in *Escherichia coli. J. Bacteriol.* **194**, 6233-6239

Taniguchi Y, Choi PJ, Li G-W, Chen H, Babu M, Hearn J, Emili A & Xie XS (2010) Quantifying *E.coli* proteome and transcriptome in Single Cells with Single-Molecule Sensitivity. *Science* **329***,* 533-538

Warner JR, Knopf PM & Rich A (1963) A Multiple Ribosomal Structure in Protein Synthesis. *Proc. Natl. Acad. Sci.* **49**: 122–129

Supplementary Methods

qPCR protocols and primer sequences

1-2ug of total RNA was reverse transcribed (MMuLV, New England Biolabs) with random primers (Ambion) according to manufacturer's protocol. Reactions were diluted 2-4 times with nuclease-free water for measurement. The following qPCR primers were used:

ssrA1.1: AAAGAGATCGCGTGGAAGCC
ssrA1.2: ACCCGCGTCCGAAATTCCTA

290_1.1: GAG ATT CTG TCT TTT CGT ATG CAG GG
290_1.2: GGG AGC ACA TTG TAG CAT TGT GCC

crp1.1: TGGCAAACCGCAAACAGACC
crp1.2: CGATGTAGTACAGCGTTTCCGC

crp2.1: CAACCAGACGCTATGACTCACC
crp2.2: GAATGCGTCCCACGGTTTCA

ais1.1: TACTGGCGCTCGCTGCAATT
ais1.2: GTGCTGCTGTGCCAGTCTGG

ais2.1: TATTGCTAAAGATAAGCGTGACG
ais2.2: CCCATCCAGATAAACTTTGCC

lpxC1.1: AGAAAGTCACCCTGACGTTACG
lpxC1.2: ACGCACAGATTTGGCATCGG

lpxC2.1: CAAACTGCTGCAGGCTGTCC
lpxC2.2: AGCTGAAGGCGCTTTGAAGG

coaE1.1: TTAACGGGAGGCATTGGCAGT
coaE1.2: CCTGACGCGCAATAATATCGGC

coaE2.1: AACTTAAGCGCACCATGCAGC
coaE2.2: TCAATGACGTCATCTGCCACG

arnBCADTEF1.1: GGAGGAACTCGCTGCAGTTAAAG
arnBCADTEF 1.2: GCGATGGCATGCTGATTTCC

arnBCADTEF2.1: CTGACCTTCGGACCACAATGG
arnBCADTEF2.2: CCATGCCGATAAACTGAGCG

arnBCADTEF3.1: TTACCTTATCGGCCTCTTCG

arnBCADTEF3.2: CGACTGATTTCCGCATAACG

arnBCADTEF4.1: GTTAGCTTTTGGCCTGGATGC
arnBCADTEF4.2: AGGCATAGGCTTTGCTTAGC

rsxABCDGE-nth1.1: AACGTTTGTGATGACGCTGG
rsxABCDGE-nth1.2: CAGAATAAATGCCAGGGTGC

rsxABCDGE-nth2.1: CAC ATG GCG GGC GTT CAT C
rsxABCDGE-nth2.2: AGCGATTGCTTCGCCAGTCA

rsxABCDGE-nth3.1: CTACAGCGGTATTCTGGCGG
rsxABCDGE-nth3.2: GCGAATCGCTTTCTGCCATA

rsxABCDGE-nth4.1: ATCGTACTCAATTTGCGCCGG
rsxABCDGE-nth4.2: CCCGTGCAGGATCAACCAAT

epd-pgk-fbaA1.1: GTGCTTTGTATGAATCCGGACG
epd-pgk-fbaA1.2: TTCAACAAATGCGCCATGCC

epd-pgk-fbaA2.1: GATGACCGATCTGGATCTTGCTG
epd-pgk-fbaA2.2: CGCGTCGCTGGTTACTTTCC

epd-pgk-fbaA3.1: GGTTGGCGTGTTCGAATTCC
epd-pgk-fbaA3.2: ATTGCTGCCAGAGTGTCGCC

epd-pgk-fbaA4.1: CATCGATACCGATACCCAATGG
epd-pgk-fbaA4.2: TTCGCCTTTCGGGTTACCCA

damX-dam-rpe-gph-trpS1.1: AACCAGAAGACGAGCTGAAACC
damX-dam-rpe-gph-trpS1.2: AATTGATCTGCGGTTCGCCA

damX-dam-rpe-gph-trpS2.1: TAAACGGCATTTGCCCAAGG
damX-dam-rpe-gph-trpS2.2: GATCAGGTCGCTATTGATATCGG

damX-dam-rpe-gph-trpS3.1: CTTGCTGCTGCGGTAGATATGG
damX-dam-rpe-gph-trpS3.2: TCTGCGCCGTTACCAATCCA

damX-dam-rpe-gph-trpS4.1: GAAAGGTGAAGTGGCTGATGC
damX-dam-rpe-gph-trpS4.2: CAGGAAGGCTTCATCGTTGC

rfe1.1: AACCAAACTTCCGCAAACGTCACC
rfe1.2: AAAACAAGCA CACCGGCACA AGC

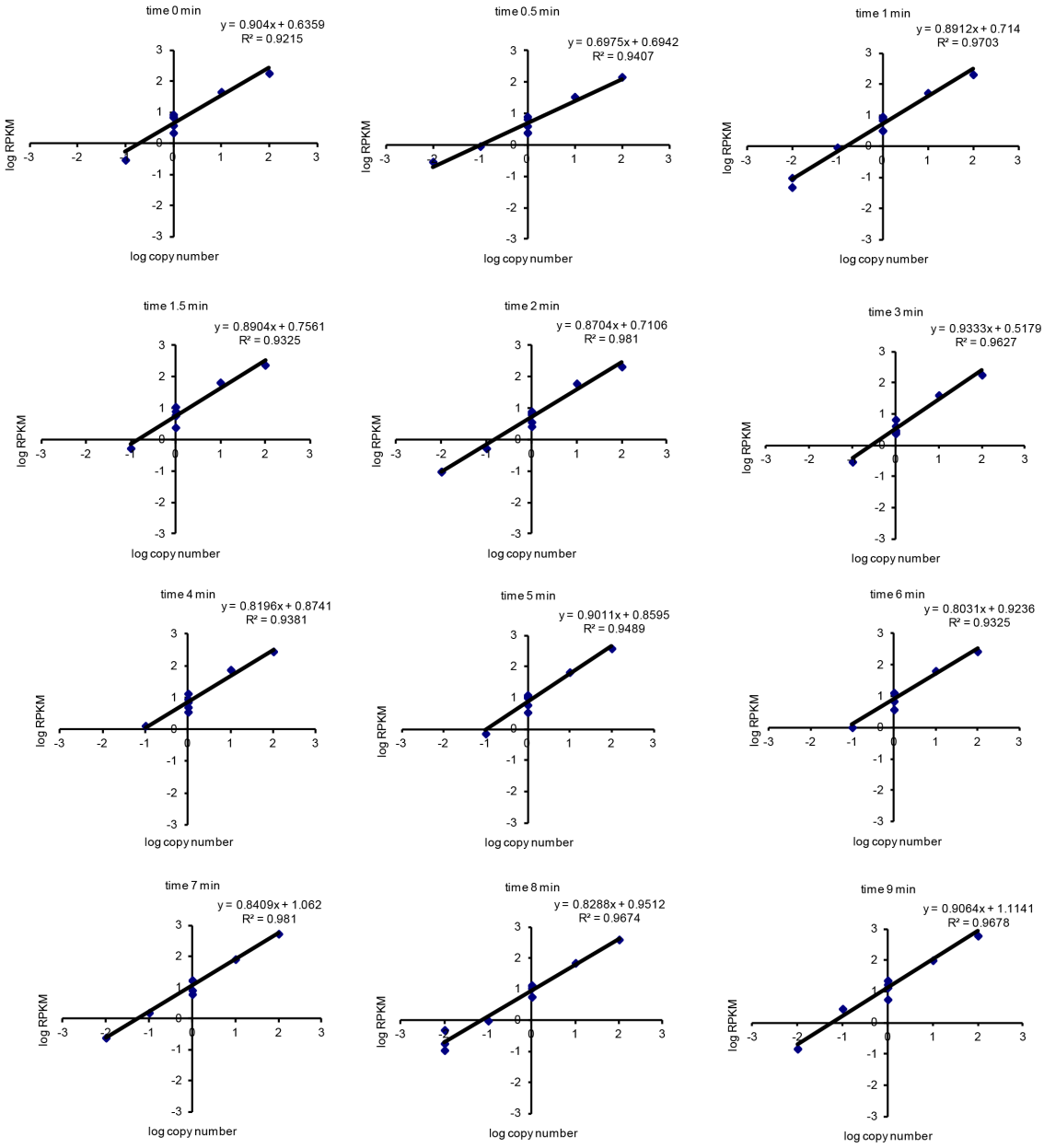
rffH1.1: ACTCACGACAGCCTGATTGAAGC
rffH1.2: TTCGCTAATG AACTGGCAGC ACG

rffC1.1: AGGTGAAGTTGATTTGGCGCTACC
rffC1.2: CGAAAACGGC TTTGCGCAAA TGC

rffM1.1: CCAAAGCAGGAGATCATCATGCG

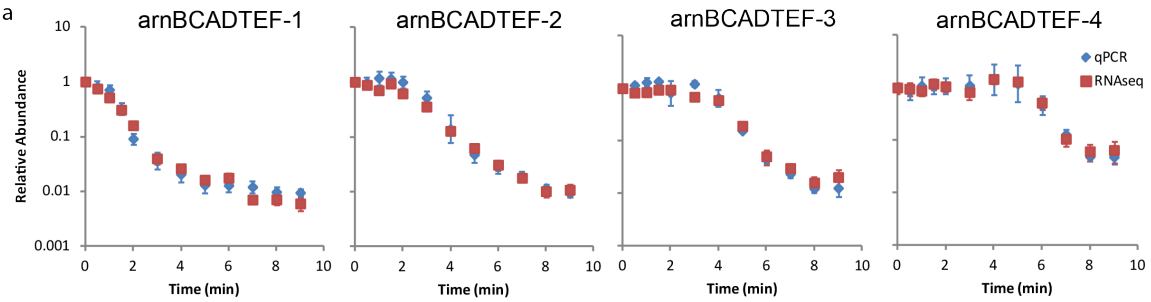
rffM1.2: AGCGTTTGCC AGATTTTCGG TGC

Supplementary Figures



Supplementary Figure 1. Spike-In RNAs confirm linear correspondence between number of sequencing reads and amount of RNA in samples. The RPKM of each 300nt bin within the spike-in RNAs correspond to the number of RNAs added to the sample at each time point.

a



b Delay (min) ± SE

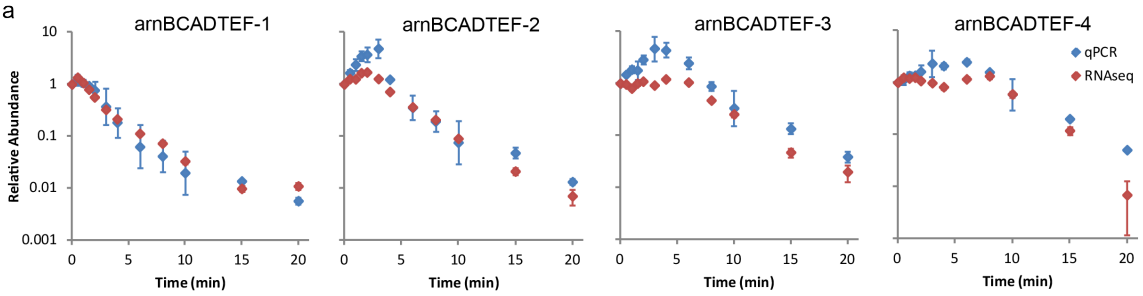
	5'							3'
probe	1		2		3		4	
	qPCR	RNAseq	qPCR	RNAseq	qPCR	RNAseq	qPCR	RNAseq
arn	-3.62 ± 1.38	0.78 ± 0.06	2.49 ± 0.52	3.30 ± 0.09	5.00 ± 0.21	5.39 ± 0.12	6.87 ± 0.59	7.00 ± 0.22
damX	2.57 ± 0.08	4.10 ± 0.16	3.81 ± 0.17	5.25 ± 0.16	7.13 ± 0.14	8.77 ± 0.41	8.09 ± 0.41	7.52 ± 0.25
epd	0.71 ± 0.26	0.25 ± 0.14	2.44 ± 0.23	1.85 ± 0.11	2.78 ± 0.21	2.65 ± 0.19	5.82 ± 0.38	7.71 ± 0.28
rsx	-0.86 ± 0.88	0.57 ± 0.44	3.96 ± 0.20	0.93± 0.27	5.11 ± 0.25	0.15 ± 0.47	7.57 ± 0.13	7.08 ± 0.51
ais	-1.35 ± 0.60	0.36 ± 0.04	0.12 ± 0.27	0.99 ± 0.04				
coaE	2.00 ± 0.20	0.83 ± 0.38	2.20 ± 0.25	0.25 ± 1.17				
crp	2.36 ± 0.27	1.92 ± 0.14	2.44 ± 0.45	2.67 ± 0.17				

c Lifetime (min) ± SE

	5'							3'
probe	1		2		3		4	
	qPCR	RNAseq	qPCR	RNAseq	qPCR	RNAseq	qPCR	RNAseq
arn	2.83 ± 0.35	1.57 ± 0.03	1.72 ± 0.13	1.38 ± 0.04	1.06 ± 0.08	1.01 ± 0.06	1.04 ± 0.20	1.01 ± 0.13
damX	1.38 ± 0.06	1.35 ± 0.07	1.45 ± 0.14	1.17 ± 0.08	1.64 ± 0.20	2.06 ± 0.66	1.24 ± 0.30	2.17 ± 0.27
epd	1.94 ± 0.28	1.71 ± 0.08	2.30 ± 0.18	2.42 ± 0.07	2.46 ± 0.20	2.87 ± 0.16	3.87 ± 1.74	4.20 ± 1.32
rsx	3.09 ± 0.35	2.04 ± 0.23	1.20 ± 0.10	2.02 ± 0.14	1.25 ± 0.12	2.85 ± 0.24	1.11 ± 0.13	1.03 ± 0.26
ais	1.79 ± 0.17	1.16 ± 0.03	1.35 ± 0.09	1.24 ± 0.03				
coaE	2.36 ± 0.20	2.88 ± 0.23	2.60 ± 0.29	4.14 ± 0.83				
crp	2.33 ± 0.09	3.16 ± 0.10	2.24 ± 0.17	2.69 ± 0.11				

Supplementary Figure 2: Exponential *E.coli* RNA-seq data is confirmed by qPCR measurements. (a) Degradation traces of regions in a representative operon, arnBCADTEF, measured by qPCR and RNA-seq are similar. (b and c) Summary of fitting data from qPCR and RNA-seq degradation traces of selected transcription units. (b) The delay time fitted from qPCR and RNA-seq data are not significantly different (p = 0.77, paired Student t test) (c) The lifetimes fitted from qPCR and RNA-seq data are not significantly different (p = 0.26, paired Student t test).

a



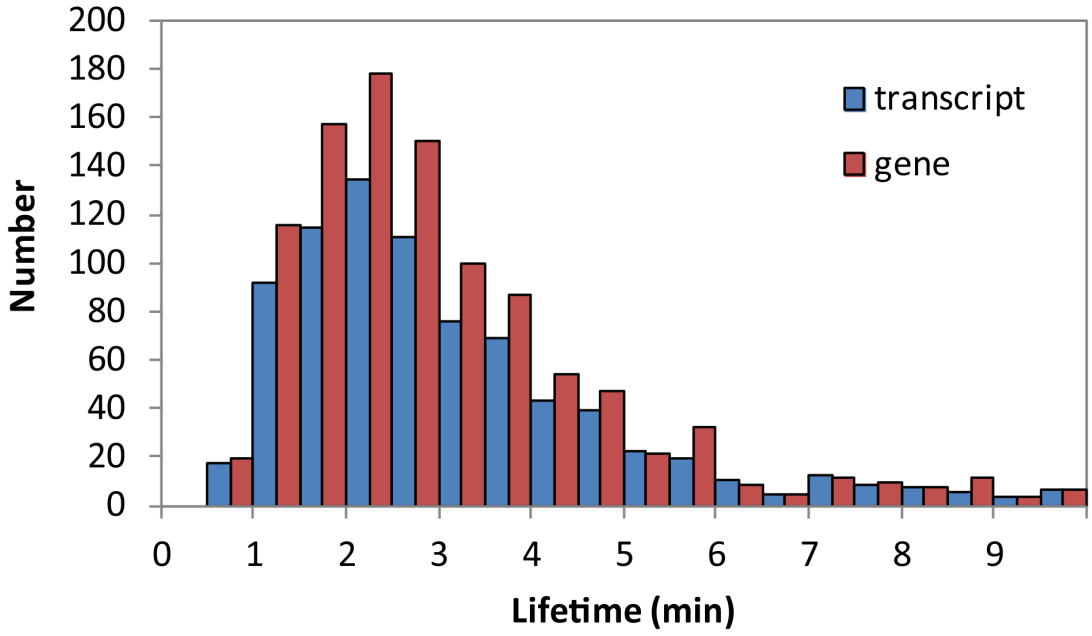
b Delay (min) ± SE

	5'				3'			
probe	1		2		3		4	
	qPCR	RNAseq	qPCR	RNAseq	qPCR	RNAseq	qPCR	RNAseq
arn	3.72 ± 0.52	3.11 ± 0.04	6.10 ± 0.62	7.90 ± 0.00	8.28 ± 0.61	8.10 ± 0.17	8.82 ± 0.13	8.75 ± 0.08
damX	6.36 ± 0.19	5.62 ± 0.10	7.54 ± 0.05	7.39 ± 0.14	8.86 ± 0.10	8.69 ± 0.08	9.05 ± 0.63	8.68 ± 0.10
epd	4.90 ± 0.94	5.26 ± 0.18	7.18 ± 0.70	6.54 ± 0.06	7.41 ± 0.43	5.84 ± 0.07	9.05 ± 0.59	6.74 ± 0.05
ais	2.22 ± 0.41	4.47 ± 0.06	4.38 ± 0.42	7.15 ± 0.18				
coaE	5.32 ± 0.34	3.46 ± 0.14	5.07 ± 0.12	4.33 ± 0.21				
crp	5.46 ± 0.16	4.40 ± 0.04	6.15 ± 0.08	4.07 ± 0.07				

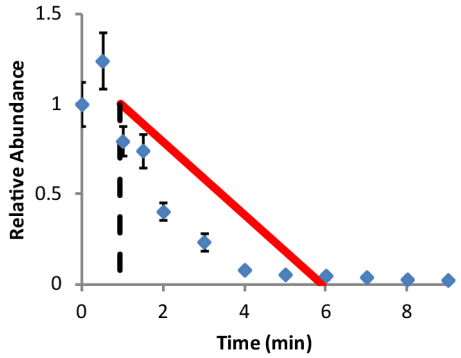
c Lifetime (min) ± SE

	5'				3'			
probe	1		2		3		4	
	qPCR	RNAseq	qPCR	RNAseq	qPCR	RNAseq	qPCR	RNAseq
arn	1.45 ± 0.12	1.69 ± 0.03	1.14 ± 0.15	0.064 ± 0.00	0.84 ± 0.22	0.65 ± 0.08	0.73 ± 0.04	0.52 ± 0.05
damX	1.32 ± 0.06	1.41 ± 0.07	1.13 ± 0.06	0.96 ± 0.08	0.89 ± 0.09	0.67 ± 0.05	0.77 ± 0.26	0.78 ± 0.07
epd	1.56 ± 0.24	1.60 ± 0.11	1.18 ± 0.27	1.44 ± 0.04	1.30 ± 0.21	2.13 ± 0.06	0.98 ± 0.31	2.29 ± 0.05
ais	2.10 ± 0.25	1.69 ± 0.04	1.78 ± 0.14	0.98 ± 0.09				
coaE	1.96 ± 0.25	1.69 ± 0.09	2.42 ± 0.13	1.66 ± 0.14				
crp	1.48 ± 0.05	1.64 ± 0.03	1.28 ± 0.02	1.86 ± 0.04				

Supplementary Figure 3: Stationary phase *E.coli* RNA-seq data is confirmed by qPCR measurements. (a) Degradation traces of regions in a representative operon, arnBCADTEF, measured by qPCR and RNA-seq are similar. (b and c) Summary of fitting data from qPCR and RNA-seq degradation traces of selected transcription units. (b) The delay time fitted from qPCR and RNA-seq data are not significantly different ($p = 0.38$, paired Student t test). (c) The lifetimes fitted from qPCR and RNA-seq data are not significantly different ($p = 0.81$, paired Student t test).



Supplementary Figure 4: Comparison of lifetime distribution of transcripts (n=847) and genes (n=1092). The lifetime distributions are not significantly different by a Kolmogonov-Smirnov test (D=0.0278, p-value=0.85).



Supplementary Figure 5: Abundance of smtA-mukFEB 5' most 300nt decays faster than predicted by post-transcriptional instantaneous degradation model (red line). We correct for the delay (dotted line) obtained from fitting.