

Dynamic insights on transcription initiation and RNA processing during bacterial adaptation

Caroline Lacoux^{1,#}, Aymeric Fouquier d'Hérouël², Françoise Wessner- Le Bohec¹, Nicolas Innocenti^{1,3,#}, Chantal Bohn⁴, Sean P. Kennedy^{5,#}, Tatiana Rochat⁶, Rémy A. Bonnin^{4,#}, Pascale Serron¹, Erik Aurell^{3,7}, Philippe Bouloc⁴ and Francis Repoila^{1,*}

¹Université Paris-Saclay, INRAE, AgroParisTech, MICALIS Institute, 78350, Jouy-en-Josas, France.

²Luxembourg Center for Systems Biomedicine, University of Luxembourg, 6 Avenue du Swing, 4367, Belvaux,

Luxembourg. ³Department of Computational Biology, Royal Institute of Technology, AlbaNova University

Center, SE-10691 Stockholm, Sweden. ⁴Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology

of the Cell (I2BC), 91198, Gif-sur-Yvette, France. ⁵Department of Computational Biology, USR3756 CNRS,

Pasteur Institute, Paris, France. ⁶VIM, INRA, Université Paris-Saclay, 78350 Jouy-en-Josas, France. ⁷Aalto

University, Department of Information and Computer Science, Espoo, Finland. [#]Present address: CL: Institut de

Pharmacologie Moléculaire et Cellulaire, UMR 7275, Sophia Antipolis-Nice, France; NI: Symbio Me AB,

Näktergalsvägen 2 E, 178 39 Ekerö, Stockholm county, Sweden; RB: EA7361, Faculty of Medicine, Université

Paris-Saclay, Le Kremlin-Bicêtre, France

***For correspondence:** francis.repoila@inra.fr

Running head: Dynamics of transcription and RNA processing

Keywords: Transcription initiation, RNA processing, RNA degradation, tagRNA-seq, bacterial adaptation

ABSTRACT

Transcription initiation and RNA processing govern gene expression and enable bacterial adaptation by reshaping the RNA landscape. The aim of this study was to simultaneously observe these two fundamental processes in a transcriptome responding to an environmental signal. A controlled σ^E system in *E. coli* was coupled to our previously-described tagRNA-seq method to yield process kinetics information. Changes in transcription initiation frequencies (TIF) and RNA processing frequencies (PF) were followed using 5' RNA tags. Changes in TIF showed a binary increased/decreased pattern that alternated between transcriptionally activated and repressed promoters, providing the bacterial population with transcriptional oscillation. PF variation fell into three categories of cleavage activity; i) constant and independent of RNA levels, ii) increased once RNA has accumulated, and iii) positively correlated to changes in TIF. This work provides a comprehensive and dynamic view of major events leading to transcriptomic reshaping during bacterial adaptation. It unveils an interplay between transcription initiation and the activity of specific RNA cleavage sites. This study utilized a well-known genetic system to analyze fundamental processes, and can serve as blueprint for comprehensive studies that exploit the RNA metabolism to decipher and understand bacterial gene expression control.

INTRODUCTION

In response to an environmental cue, gene expression reprogramming triggers changes in RNA synthesis, RNA processing and/or degradation, resulting in a physiological response. Transcription initiation and RNA processing frequencies for genes not involved in these switches remain unchanged during the adaptation process and their primary and processed RNAs remain at constant levels. In contrast, for genes mediating the adaptation process, the transcription initiation frequency and/or the RNA processing frequency vary resulting in novel or transient changes in RNA levels (e.g. (Reznikoff et al. 1985; Phadtare and Severinov 2010; Rochat et al. 2013; Bouloc and Repoila 2016)). These events occur in minutes, or less, and in a fraction of the bacterial population doubling time (Anderson and

Dunman 2009; Esquerre et al. 2014).

Transcriptomic approaches are suited to provide information on adaptation processes in a genome-wide manner as they enable comparisons of RNA landscape snapshots between different environmental conditions or genetic backgrounds (Croucher and Thomson 2010; Filiatrault 2011; Guell et al. 2011; Mader et al. 2011). The majority of these studies compare steady state RNA levels in bacterial offspring against the initial bacterial population. As such, they lack information on the dynamics of RNA remodeling, the chronology of events, and the mechanisms responsible for changes in gene expression. Pioneering studies addressing RNA dynamics during adaptation processes have been reported. For instance, in *Caulobacter crescentus*, activations and repressions of gene expression during the cell cycle development was visualized using DNA microarrays (Laub et al. 2000). In *Escherichia coli* K12, the expression kinetics of regulons under control of the extracytoplasmic sigma factor σ^E and the non-coding RNA RhyB were observed (Masse et al. 2005; Rhodius et al. 2006; Bury-Mone et al. 2009; Gogol et al. 2011). However, these studies remain unable to distinguish transcriptional or post-transcriptional effects on RNA levels. Similarly, genomic-scale measurements of RNA stability uncouple transcription from RNA processing and degradation by using rifampicin, an antibiotic blocking transcription initiation (Mosteller and Yanofsky 1970; Redon et al. 2005; Kristoffersen et al. 2012; Esquerre et al. 2014; Dar and Sorek 2018). This technique cannot map the RNA processing sites that are key actors in RNA decay (Mohanty and Kushner 2016), and may mask interplay between transcription, RNA processing and degradation as demonstrated in eukaryotic organisms (Dahan and Choder 2013; Singh et al. 2015; Peck et al. 2019). RNA-seq methods have been gradually adapted in an effort to probe regulatory mechanisms of bacterial transcriptomes. Differential RNA-seq (dRNA-seq) compares total transcriptomes to those enriched for 5' triphosphate RNA ends to aid in transcription start sites (TSSs) prediction (Sharma and Vogel 2014). Genome-wide single-nucleotide-resolution mapping of 5' RNA termini can be performed by methods such as tagging RNA-seq (tagRNA-seq) or EMOTE (Fouquier d'Herouel et al. 2011; Linder et al. 2014; Innocenti et al. 2015). We demonstrated, in *E. coli* K12 and *Enterococcus faecalis*, that tagRNA-seq can map and distinguish TSSs and RNA processing sites (PSSs), and that tagging efficacy is proportional to the

abundance of 5' termini (Fouquier d'Herouel et al. 2011; Innocenti et al. 2015). These new techniques, while powerful, come with several caveats. dRNA-seq relies on the 5'-phosphate-dependent exonuclease (TEX) to degrade RNAs from 5' monophosphate groups, yet TEX is sensitive to RNA folding and can generate artefactual internal TSSs (Szittyá et al. 2010; Conway et al. 2014; Prados et al. 2016). Along the same lines, tagRNA-seq generates a significant number of RNA ends, coined 'undetermined' (UND) that cannot be classified as TSS or PSS without additional and individual analysis (Innocenti et al. 2015). Nevertheless, tagRNA-seq provides the most promising means to investigate the dynamics of key events reshaping the transcriptome during the bacterial adaptation process.

The maintenance of envelop homeostasis in *E. coli* is ensured by regulatory networks including the σ^E regulon (Ades 2008; Silhavy et al. 2010; Grabowicz and Silhavy 2016). The extracytoplasmic σ factor (σ^E) encoded by *rpoE* is normally sequestered by RseA at the inner surface of the cytoplasmic membrane. Perturbations of surface proteins or envelop integrity cause σ^E to be released into the cytoplasm where it binds to RNA polymerase (RNAP). In turn, σ^E -RNAP binds to specific promoters directing transcription of ~100 encoding sequences (CDSs) including surface proteins, enzymes, envelop compounds, transcription and translation components, and *rpoE* itself. Regulatory factors, including three small RNAs (sRNAs), MicA, RybB and MicL, are also expressed. These sRNAs provide negative feedback to the σ^E regulon, modulating the translation and/or the stability of surface proteins mRNAs (Rhodius et al. 2006; Mutalik et al. 2009; Gogol et al. 2011; Rhodius et al. 2012; Guo et al. 2014; Shimada et al. 2017).

The aim of this study was to evidence at the genomic scale changes in transcription initiation frequency (TIF) and RNA processing frequency (PF) during an adaptation process, the activities of two key molecular events generating 5' RNA ends. The σ^E regulon exhibits the crucial aspects of RNA metabolism governing gene expression control: RNA synthesis, processing and degradation. It consists of manageable number of genes, making it an attractive model to simultaneously observe RNA transcription initiation and RNA processing for the first time at the genomic scale in bacteria. We used tagRNA-seq to study the kinetics of the RNA pool within a single bacterial generation in response to

σ^E induction. The results presented, including the synchrony between changes in the cleavage efficacy measured at specific PSSs, and the activity of transcription initiation at the cognate TSSs, provide evidence to support the hypothesis that interplay exists between transcription initiation and RNA cleavage.

RESULTS AND DISCUSSION

Dynamics of the RNA levels in response to σ^E induction

The σ^E regulon was used as a model system in this study to observe an evolving bacterial RNA landscape. The σ^E encoding sequence, *rpoE*, was expressed under the control of an inducible promoter responding to anhydrotetracycline (aTc) in *E. coli* K12 (Bury-Mone et al. 2009). After the addition of aTc, the dynamics of the RNA pool was monitored by collecting samples every 5 minutes (min) over a period of 20 min. Samples were then analyzed by tagRNA-seq (Fig. 1A) (Innocenti et al. 2015). RNA-seq raw data were processed as presented in the ‘Materials and Methods’ section. Supplemental Tables S1 and S2 provide a comprehensive list of normalized data, statistical treatments, and temporal changes in RNA levels. As detailed and discussed in Supplemental section S1, changes observed in RNA levels over the experimental time course demonstrated that the σ^E regulon was induced and functional, a prerequisite for this study. The bulk of gene reprogramming was observed within the 10 first min and was nearly complete at 20 min (Fig. 1B). The majority of changes in RNA levels showed a monotonic variation (increased or decreased) and a minority was ‘transient’. mRNA targets of sRNAs synthesized by σ^E -RNAP (MicA, RybB and MicL) showed differential sensitivity to sRNA-mediated degradation. These observations confirmed that our σ^E experimental system produced the expected gene expression reprogramming patterns that had been previously reported (Supplemental section S1 and Tables S2, S3). The results further demonstrated that the inducible σ^E system was functional and that resulting RNA landscape kinetics could be used to analyze changes in TIF and PF.

Frequencies of processes generating 5' RNA ends

TagRNA-seq discriminates 5' RNA ends generated by transcription initiation and RNA cleavage events by use of two short RNA sequences termed 'TSS-' and 'PSS-tag'. Criteria utilized to assign 5' RNA ends are found in Supplemental section S2 and the modulation of tags in the RNA pool is summarized in Supplemental Table S4. We identified 1147 5' RNA ends as TSSs, and 594 as PSSs, while 1706 remained UNDs (Table 1). We previously demonstrated that tag-counts provide good estimates for the relative abundance of 5' termini (Fouquier d'Herouel et al. 2011; Innocenti et al. 2015). Although current RNA-seq methodologies exhibit strong technical variability (McIntyre et al. 2011; Evans et al. 2018), the consistency analysis of tag-counts between kinetic series indicated that mean tag-counts reflect dynamics despite the variability between individual series (see Materials and Methods and Supplemental Table S4). Based on this result, we report mean tag-count values that relate to frequencies of events generating 5' RNA ends. We defined the 'tagging rate' (R) as the average number of tags appearing or disappearing over time, a magnitude coupled to the frequency of the process generating any given RNA terminus. An initial steady state of expression is assumed prior to the induction of σ^E where $R_0 = 0$ tag/min. The tagging-rate over an interval of time, R_{interval} is $R_i = [(n_i - n_{i-1}) / (t_i - t_{i-1})]$; where n_i and n_{i-1} are mean tag-counts at corresponding time points t_i and t_{i-1} . From R_i , we focused on changes in tagging frequency as a metric of gene reprogramming. Changes in tagging frequency are evaluated by comparing R_i at different time points to yield a 'tagging acceleration' (A). For consecutive experimental time intervals, A_{interval} is $A_i = [(R_i - R_{i-1}) / (t_i - t_{i-1})]$. As with R_0 , the initial tagging acceleration is assumed to be $A_0 = 0$ tag/min². Changes in tagging frequency (A_i) can be positive or negative. When $A_i > 0$, it indicates an increased frequency, when $A_i < 0$, it indicates a decreased frequency, when $A_i = 0$, it indicates a frequency equal to the one in the previous time interval ($R_i = R_{i-1}$). Thus, if a particular 5' end is a TSS, the A_i values correspond to changes in transcription initiation frequency (changes in TIF). A_i values for a PSS correspond to changes in the RNA processing frequency (changes in PF). Supplemental Table S5 summarizes mean values for R_i and A_i for each 5' RNA end mapped. Since only three biological replicates were used and that RNA treatments feature large technical variability, measured tag-counts and calculated R_i and A_i values were typified by high dispersion, as expected and shown by standard deviations (Supplemental Tables S4

and S5). Consequently, the robustness of mean values was evaluated by statistical methods apply to small numbers of samples (Supplemental Tables S4 and S5). We calculated confidence values pA_i , expressed as meta-analysis p-value using Fisher's method. pA_i values $< 1.45 \times 10^{-5}$ are considered significant, with A_i values at least 95% reliable. Out of 3447 total 5' termini mapped, this level of confidence applies for $\sim 40\%$ of RNA ends.

Assessment of a threshold for biological significance

Deducing biological relevance from changes in frequency (tagging-acceleration) calculations is a challenge without a similar guiding precedent. Here, we leveraged the well-studied σ^E system and the fact that an increased TIF was anticipated for σ^E -dependent TSSs (σ^E -TSSs) in order to establish an empirical 'biological significance' threshold for changes in A_i . Out of 88 reported (confirmed experimentally or predicted) σ^E -TSSs (Gama-Castro et al. 2016; Keseler et al. 2017), 40 were mapped, including 22 assigned as TSSs (Table 2). A_5 values (5 min following σ^E induction) for this set of promoters varied greatly from 0.1 tag/min^2 (P_{rmlB}) to 21.1 tag/min^2 (P_{bamE}) and prevented direct determination of a threshold. This result was unsurprising given the large dynamic range of bacterial promoters, including σ^E -dependent ones (Mutalik et al. 2009; Rhodius and Mutalik 2010). However, since it may that certain σ^E -TSS did not responded to the σ^E induction, we imposed an arbitrary threshold based on the lowest A_5 value: $|A_i - A_{i-1}| > 0.7 \text{ tag/min}^2$, or 7 times the value for P_{rmlB} . While arbitrary, such a threshold was effectively stringent as it excluded σ^E -dependent P_{rmlB} from promoters called as transcriptionally activated during the first five min (t_{5-0}). Empirical significance of this threshold is strengthened as over 80% of RNA ends with a significant biological change at t_5 were retained as statistically significant (Supplemental Table S5). This cutoff was subsequently applied to 5'RNA ends over the course of the experiment, including RNA cleavage sites (PSSs). Out of all mapped 5' RNA ends, approximately one third (38%) showed a significant change over 20 min of measurements. The bulk of significant TSSs changes were found at the earliest interval, t_{5-0} . In comparison, the bulk of responding PSSs peaked later at t_{10} (Table1). The observation that most gene

reprogramming occurred within the first 10 min following induction fits with the conclusion inferred from RNA levels (Supplemental section S1, Fig. 1B). Significant changes in UND 5' ends mirrored the pattern of PSSs, most likely due to our assignment method that favors the presence of true PSSs in this group (Supplemental section S2).

By applying the threshold value of $> |0.7| \text{ tag/min}^2$, we retrieved tendencies observed for changes in RNA levels, indicating that such an empirical threshold is suited for a global analysis. To pinpoint general features and highlight specific outputs from changes in TIF and PF, we then focused on patterns described by changes in A_i values. These are presented as selected examples. Unless otherwise mentioned, all satisfy the biological significance and the statistical confidence.

Pattern of changes in TIF at σ^E -dependent promoters

Significant changes in A_i values were recorded for 21/22 assigned σ^E -TSSs, the sole exception being P_{ydhIJK} (Table 2). Nine TSSs (P_{bamAp2} , P_{eptB} , P_{micA} , P_{plsB} , $P_{rpoEp2a}$, P_{rseAp3} , P_{rybB} , P_{sbmA} and P_{yhjJ}) exhibited a biphasic response, with increased TIF observed at intervals t_{5-0} and t_{15-10} , and decreased values at t_{10-5} and t_{20-15} . The second increase at t_{15-10} was of similar or higher magnitude. We also observed that decreases in TIF were similar or higher compared with increases, indicating that transcription initiation tended to return to the t_0 state (Table 2, Fig. 1C, 1D). The biphasic pattern reflects a 'transcriptional oscillation' within the bacterial population, initially synchronized by addition of the inducer at t_0 . This pattern could reflect transcriptional bursting occurring at the single cell level for highly expressed genes where initiation occurs in surges and promoters exhibit intermittent periods of inactivity possibly due to topological constraints (Golding et al. 2005; So et al. 2011; Chong et al. 2014).

Four additional promoters (P_{bamDp2} , P_{rnlB} , P_{ogpGH} and P_{ybfG}) showed a delayed response with increased TIF at t_{10} instead of t_5 , followed by decreased- and increased TIFs (Fig. 1C, 1E). Three others (P_{bamE} , P_{bepA} , and P_{yggN}) displayed an increased/decreased TIF pattern during t_{15-0} , but then continued in increase in t_{20-15} (Fig. 1C, 1F). The remaining five, from the original 21 (P_{clpX} , P_{psd} , P_{dsbCp2} , $P_{ygiMcca}$, and P_{yrap2}), showed an increased TIF that peaked at t_{10} or t_{15} , followed by a drastic decrease (Fig. 1C, 1G). Variations from the biphasic response most likely reflect additional controls modulating the use of σ^E -

TSSs by σ^E -RNAP. Changes in TIF for P_{dsbCp2} and P_{ydhIJK} , for instance, were rather weak compared to other σ^E -TSSs, indicating low transcriptional activation by σ^E induction (Supplemental Table S5). In contrast, RNA levels for *dsbC* and *ydhIJK* were highly increased at t_5 and remained 20-fold above baseline from t_{10} through t_{20} (Supplemental Table S2). These observations strongly suggest that in addition to transcription, stabilization of these transcripts occurs and can be responsible for the variations in the biphasic pattern.

Among other known σ^E -TSSs, 13 were assigned as UND in our study (Table 2). Six showed no significant variations in TIF, suggesting little to no activation. The remaining seven generally followed the biphasic pattern, indicating that these promoters were activated by σ^E induction (P_{bacA} , $P_{yieEchrR}$, P_{lpxP} , P_{micL} , P_{sixA} , P_{yfeKS} and P_{yiiS}), (Table 2, Supplemental Fig. S3).

Five previously mapped TSSs (Keseler et al. 2017) were assigned as ‘PSSs’ (Table 2), and only three of them (P_{degP} , P_{lpxD} and P_{fkpA}) showed significant variations in A_i , with similar patterns to those observed for mapped σ^E -TSSs (Fig. 1E and 1G, Supplemental Fig. S3). The assignment of these 5’ RNA ends as PSSs, may be due to the transcriptional organization revealed by tagging density at these loci (Supplemental Table S5). *lpxD* is embedded in a complex operon where at least three promoters are located upstream the mapped promoter (coordinate 200960; Table 2). Longer transcripts containing *lpxD* may favor the PSS assignment of the corresponding nucleotide due to RNA processing or degradation generating 5’ monophosphate ends. In line with this possibility, a TSS was mapped two nucleotides downstream, which may be the TSS of *lpxD* (coordinate 200962; Supplemental Table S5). Similarly, reported TSSs for *degP* and *fkpA* (Table 2) were embedded within strongly transcribed regions that featured an abundance of tags ligated to upstream and downstream nucleotides: 28 and 12 consecutive nucleotides were tagged for *degP* and *fkpA*, respectively. Based on our data, we predict TSSs at 180840/42 for *degP*, and 3477528 for *fkpA*.

In general, the tag-counts dynamics revealed specific patterns of changes in TIF at each σ^E -TSS under σ^E induction. Although each σ^E -TSS did not respond with the similar intensity and can exhibit variations, transcription activation features a general biphasic pattern originating with an increased TIF.

Biphasic patterns at σ^E -independent promoters

Half of mapped TSSs responded to σ^E induction, with $\sim 75\%$ of these responding within the first 10 min (Table 1). We observed two major TIF variations profiles for promoters responding to σ^E induction but reportedly not transcribed by σ^E -RNAP (Keseler et al. 2017). A biphasic pattern with increased TIFs at intervals t_{5-0} and t_{15-10} and decreases at t_{10-5} and t_{20-15} was observed (*dps*, *grxA*, *rpmHp2*, *rpsMp2*, *rpsT*, *secG*, *tufB*, *yabI* and *ybhQ2*) (Fig. 2A, Supplemental Fig. S4A). As with σ^E -TSSs, this indicates transcription activation, a conclusion corroborated by increased RNA levels at t_5 for most of the corresponding CDSs (Table S2). The second pattern was typified by decreases at intervals t_{5-0} and t_{15-10} (*cspD*, *gatY*, *glpT*, *hupA*, *lpp*, *polA*, *sodA*, *ssrSp1* and *treB*) (Fig. 2B, Supplemental Fig. S4B). This latter profile, with decreased TIF at the interval t_{5-0} is likely to indicate transcriptional repression. The half-life of most *E. coli* mRNAs is less than five minutes and so RNA degradation has a prominent impact on RNA levels (Chen et al. 2015; Dar and Sorek 2018). Reduced synthesis during t_{5-0} would, therefore, be sufficient to decrease the amount of RNA. Decreased levels of transcript corroborate this conclusion, except for relatively stable transcripts (*lpp* and *ssrS*) (Table S1, Supplementary section S1).

From the data, we hypothesize that the opposing patterns of changes in TIF could result from competition between sigma factors binding to core-RNAP, with contribution from promoter strengths, transcription regulators, and transcriptional burst that temporarily prevent promoters from being re-used by the RNAP (Jishage et al. 1996; Golding et al. 2005; Chong et al. 2014; Mauri and Klumpp 2014). In the simplest model, promoters with an increased activity at t_5 would outcompete those displaying a decreased TIF. During the first minutes after σ^E induction, the activated promoters would be occupied and/or refractory to re-initiation and the completion of ongoing transcription would release core-RNAP. The situation would then reverse, with the binding of free sigma factors to core-RNAP and the utilization of the ‘unoccupied’ promoters during the interval t_{5-0} . This would generate opposite changes in TIF between the two groups of promoters and transcription oscillation to the bacterial population, a hypothesis consistent with current modeling of transcriptionally active promoters

indicating ON/OFF states at the single cell level (Jones and Elf 2018).

About 15% of total mapped TSSs showed significant changes in TIF after t_{10} (Table 1). These promoters may themselves be a response to cellular variations provoked by the increase of σ^E . However, many of these TSSs are flanking nucleotides to TSSs detected in previous time intervals. For instance, the TSS for *yccA*, reported at 1031443 (Keseler et al. 2017) and encoding a modulator of FtsH protease, was also tagged at 1031444/5/6/7/8: Two positions showed no significant changes in TIF, two other decreased at t_{5-0} and one decreased at t_{15-10} (Supplemental Table S5). Other examples were also observed (*gatY*-2177230/1/2/3, *yobF*-1907615/6/7/8, *yceDpI*-1146648/49/50) (Supplemental Table S5), and may suggest that TIF changes can modify stringency of the holo-RNAP to utilize specific nucleotides within the DNA promoter.

Further analysis combining changes in TIF and RNA levels revealed additional transcriptional and post-transcriptional effects. For instance, the comparison of TSSs and RNA levels of *dps* and *gatY* (Fig. 2C, 2D) showed that RNA levels track changes in TIF for both genes in the interval t_{5-0} , but then diverge. *dps* RNA synthesis is accompanied by an increased degradation as marked by decreasing RNA levels. The decreased TIF for *gatY* during the first five minutes would be sufficient to decrease RNA levels without a change in RNA stability. After t_5 , changes in TIF marked a global increase that may be responsible for the increasing *gatY* levels observed onward from t_{10} (Fig. 2C, 2D). Both *dps* and *gatY* reach similar RNA levels at t_{20} (2- to 3-fold lower) but their respective paths likely involved differing controls. This highlights the interest of combining kinetics- and 5' tagging-approaches with RNA-seq to shed light on mechanistic aspects of gene expression control.

Selectivity of RNA processing sites

We examined the activity of RNA cleavage at PSSs and identified three distinct classes corresponding to PF changes. The first class, representing ~60% of total PSSs (Table 1), included PF changes below the threshold of biological significance ($|A_i - A_{i-1}| \leq 0.7 \text{ tag/min}^2$), indicating cleavage independent of any variation in RNA amounts. For activated genes, most of the newly synthesized RNA will be not

cleaved at these sites, which suggests stabilization. The second class, ~ 30% of total PSSs, included PSSs with late changes in PF relative to variations of measurable RNA amounts (Table 1, three right columns). At these PSSs, cleavage activity responded to RNA accumulation and PF peaked when RNA reached maximal amounts, primarily at t_{10} , t_{15} and t_{20} . The remaining third PSSs, ~10% of total PSSs, included early changes in PF during the first five minutes (column 't₅' in Table 1). These changes paralleled patterns of increasing RNA levels or variations in TIF when transcription activation was counterbalanced by RNA degradation. For transcriptionally activated genes, RNA cleavage most likely occurs concomitantly to RNA synthesis. Each of the three classes, with examples is discussed in the following sections.

RNA dynamics at the *rpsU-dnaG-rpoD* operon. In bacteria, polycistronic transcripts encode proteins which may be required in different amounts. A common solution is separation via cleavage coupled to differential RNA decay for each CDS component (Rochat *et al.* 2013). The *E. coli* σ^E -independent *rpsU-dnaG-rpoD* operon is one such example that encodes the ribosomal protein S21, the primase DnaG, and the vegetative sigma factor σ^D (Fig. 3A). The *dnaG* mRNA is rendered unstable in comparison to *rpoD* by RNase E cleavage, maintaining the amount of DnaG at about 100 molecules, versus a few thousands for σ^D (Burton *et al.* 1983; Yajnik and Godson 1993). Interestingly, amounts of σ^D increase under σ^E induction (Rhodius *et al.* 2006). We hypothesized that increased *dnaG-rpoD* RNA levels should be accompanied by a commensurate increased PF at *dnaG* PSSs such as to preserve the DnaG and σ^D protein ratios.

Three TSSs were mapped at the *rpsU-dnaG-rpoD* locus: Two σ^D -TSSs upstream *rpsU* at coordinates 3210646/7 and 3210716/7/8, and one σ^H -TSS within *dnaG*, assigned as UND, at position 3212688 (Fig. 3A). Only TSS-3210716 displayed a significantly increased TIF during the interval t_{10-5} , which was insufficient to explain the increased abundance of RNA levels observed at t_5 (Fig. 3B, 3C).

Eleven PSSs were mapped: Nine nested in transcripts harboring *dnaG* and *rpoD* sequences and two within the 5' untranslated region (5'UTR) of *rpsU* (Fig. 3A). Four (-3212541, -3212942, -3213927 and -3214086) had no significant changes in PF. PSSs-3213710 and -3213842 within *rpoD* and PSS-

3212941 in between *dnaG* and *rpoD* CDSs showed modest changes in PFs at t_{15} and t_{20} suggesting that RNA cleavage varies due to RNA accumulation (Fig. 3D). Changes in PF for PSS-3210759, within the 5'UTR of *rpsU*, also had a delayed activity relative to RNA synthesis, peaking at t_{15} once the global *rpsU* RNA amounts were declining (Fig. 3B, 3D). The final two (PSSs-3212852/53) within the *dnaG* stop codon, showed progressive and significant increased PF that paralleled the increasing amounts of *dnaG-rpoD* RNA. It should be noted that the RNase E site acting at the end of *dnaG* was previously reported at position 3212852 (Burton et al. 1983). This indicates that cleavages at these PSSs affects the newly synthesized *dnaG-rpoD* transcripts (Fig. 3B, 3D). Upon closer investigation, PF for PSS-3212853 also manifested a response that peaked once *dnaG-rpoD* RNAs had reached their maximal amounts at t_{15} , (Fig. 3B, 3D), strongly suggesting a cleavage activity linked to RNA accumulation. However, PF for PSS-3212852 mirrored the increasing amounts of *dnaG-rpoD* RNAs and peaked with RNA levels at t_{10} , which parallels the increased TIF observed for TSS-3210716 at t_{10-5} (Fig. 3C, 3D). The concomitant increase of RNA amounts, PF and TIF at t_{10-5} strongly suggests that PSS-3212852 responded to RNA synthesis and, possibly, to transcription activation.

Our results confirm the importance of the RNase E site (3212852) in the control of *rpsU-dnaG-rpoD* operon (Burton et al. 1983; Yajnik and Godson 1993). Beyond this, we are able to detect that PSS-3212852 responds to the increasing RNA synthesis, in contrast to other PSSs which depend on accumulated RNA or act independently of RNA levels. To our knowledge, this is the first observation of differential activity of PSSs within a bacterial operon using genome scale RNA-seq. Most importantly, detection was performed without altering transcription (by rifampicin) or by shifting environmental conditions to inactivate an essential RNase (RNase E).

Changes in PFs for PSSs in *dnaG-rpoD* transcripts revealed that, within a transcript, certain PSSs respond specifically and differently to changes in RNA levels and RNA synthesis. The evidence for such differentiation is supported by the independence of RNA level and responding PFs across the transcriptome. For instance, *mreC*, *phoBR*, *ptrA*, *sbmA* or *yabI* RNAs reached their highest amounts at t_5 and harbor PSSs whose changes in PF were not significant and/or occurred at later time intervals (Supplemental Table S1, S5). A further example of different classes of PSSs in the σ^E -dependent

bicistronic operon *bepA-yfgD* is presented in Supplemental section S4.

RNA dynamics at the *ahpCF* operon. We wished to test reports that transcription activation of certain genes is overbalanced by increased RNA degradation resulting in net lower levels of RNA (Redon et al. 2005; Esquerre et al. 2014; Nouaille et al. 2017). If such a control occurred during σ^E induction, we expected: 1) increasing RNA levels at early time points, followed by a degradation-mediated decrease, and 2) changes in TIF featuring activation coupled with increased PFs of newly synthesized RNA molecules. This implies synchronous changes in TIF and PF for the 5' ends of the two distinct RNA molecules, which originate from a common primary transcript.

Several transcripts satisfy these expectations, and the *ahpCF* operon is presented here as we mapped TSSs at locations previously reported (Fig. 4A) (Keseler et al. 2017). *ahpCF* encodes a hydroperoxidase involved in the detoxification of hydrogen peroxide. Under σ^E induction, the abundance of each CDS (*ahpC* and *ahpF*) increased > 4-fold in the interval t_{5-0} and then fell to ~2-fold below that of t_0 (Fig. 4B). In response to σ^E induction, a biphasic pattern was observed for only one of the σ^D -TSSs (638921) indicating that it was responsible for the majority of RNA increase measured at t_5 (Fig. 4C). Six PSSs were assigned within the transcription unit *ahpCF*. Four PSSs showed no significant (638762, 638907) or mild and late (638763, 639731) changes in PF indicating that RNA cleavage occurred independently of RNA amounts at these sites. In contrast, PSSs upstream *ahpC* (638922) and between *ahpC* and *ahpF* (639732) displayed changes in PFs that paralleled changes in TIFs measured for TSS-638921 (Fig. 4C, 4D). However, we are cautious in our conclusions here as TSS-638921 and PSS-638922 are separated by one nucleotide, and so we cannot fully rule out that the later might be a TSS explaining why we observed a biphasic pattern. In contrast, for PSS-639732, the similarity between changes in PF and in TIF patterns strongly suggests that RNA cleavages responded to increased TIF, and most likely occurred on nascent RNA molecules. Given the requirement for RNA degradation, we remarked on an absence of PSSs mapped inside CDSs *ahpC* and *ahpF*. This indicates that our method 'captured' only certain RNA cleavage sites and it is probable that short-lived or smaller RNA fragments eluded the tagging method or the RNA-seq protocol.

Both *rpsU-dnaG-rpoD* and *ahpCF* operons attest to the specific response of certain PSSs to RNA synthesis during the adaptation process. The difference between patterns presented Fig. 3D and Fig. 4D can be explained by the accumulation of cleaved RNA molecules between experimental time points for *dnaG-rpoD*, and the elimination of cleaved RNA molecules between time points for *ahpCF*. Most importantly, changes in PF at specific PSSs that parallel changes in TIF indicates that cleavage activity is somehow coupled to transcription. A few previous studies touched upon the relationship between RNA synthesis and stability in *E. coli* (Chow and Dennis 1994; Chen et al. 2015; Nouaille et al. 2017). Here we were able to confirm this observation while providing additional insight and examples.

A major role for RNA processing and degradation has been long recognized in single gene expression studies (Burton et al. 1983; Gorski et al. 1985; Newbury et al. 1987; Haugel-Nielsen et al. 1996; Ludwig et al. 2001; Repoila and Gottesman 2001; Winkler et al. 2004; Urban and Vogel 2008; Prevost et al. 2011). The contribution of RNA degradation in adjusting gene expression during bacterial adaptation has been highlighted and, more recently, differential mRNA decay was visualized in a genome wide manner as a key actor reshaping the expression of operonic organizations (Dar and Sorek 2018). In *E. coli* and *Lactococcus lactis*, Coccagn-Bousquet and colleagues established the major impact of RNA stability to counteract RNA synthesis and adjust growth to nutrient availability (Redon et al. 2005; Esquerre et al. 2014; Nouaille et al. 2017; Dressaire et al. 2018). In *S. aureus* responding to diverse cues, Dunman and colleagues showed that more than 80% of stress-modulated transcripts have modified stability during adaptation and that factors of RNA processing and degradation can be targets of antimicrobial molecules (Anderson et al. 2006; Olson et al. 2011). In *Salmonella*, Vogel and colleagues assigned ~ 22,000 RNase-E-mediated PSSs resulting from a 28°C to 44°C temperature shift, reinforcing the role of RNA processing and turnover in gene expression control (Chao et al. 2017). Our study also highlights the strong impact of RNA processing and stability on the bacterial adaptation response. Moreover, we improve on existing techniques through the dual mapping of TSSs and PSSs in a single RNA sequencing run. Combined measurement of changes in frequencies (TIF

and PF) allow us to discriminate transcriptional and posttranscriptional events and correlate these fundamental events to variations of RNA levels. In the case of PSSs, changes in PF unveils cleavage sites that figure prominently in the evolving RNA landscape, and hence, in the bacterial response to environmental cues.

CONCLUSION

Here we describe, for the first time in the literature, the direct and simultaneous observation of changes in the efficacies of transcription initiation and RNA cleavage at the genomic scale. This work provides the first comprehensive and dynamic view of fundamental processes modulating gene expression.

The σ^E regulon in *E. coli* served as an ideal model by providing a list of 5' RNA ends where frequencies were expected to vary upon σ^E induction: The σ^E -TSSs. From σ^E -TSSs, general patterns were inferred for activation and repression of transcription initiation at σ^E -independent TSSs. We revealed that a significant number of assigned TSSs showing changes in TIF displayed a biphasic pattern, conferring transcriptional oscillation to the bacterial population sensing and responding to an environmental signal. Within this same evolving transcriptome, we discovered three classes of PSSs including one whose activity respond to transcription. The factors providing selectivity to PSSs or able to couple transcription initiation to RNA cleavage remain to be established.

In the light of the data and the results, it is evident that additional experiments will be required to strengthen and refine conclusions on transcriptional and post-transcriptional events shaping the evolving RNA landscape. Yet, absent the insights provided by tagRNA-seq in this study, these avenues of investigation would have remained obscured. Directed mutagenesis on certain PSSs, or swapping promoters and cleaved RNA sequences, may clarify the coupling observed between transcription initiation and certain cleavage sites. Such a work will also further clarify the activity of PSSs as a function of RNA amounts or transcriptional activity.

Together, our data show that tagRNA-seq is suited for the direct and simultaneous analysis of transcriptional and posttranscriptional events taking place at sub-generation timeframes in response to

an environmental cue. We have laid the foundation to probe, at the genomic scale, major processes controlling gene expression during physiological adaptations in bacteria. This generalized approach will provide dynamic and mechanistic insights to studies that exploit changes in RNA levels to decipher and understand gene expression reprogramming or for engineering purposes.

MATERIALS AND METHODS

Bacterial growth and RNA preparation

Three individual clones of the *E. coli* strain MG1655/pZE21-rpoE (Bury-Mone et al. 2009) were grown 18 h in LB medium at 37°C with agitation (200 rpm). Each culture was diluted 1/1000 in pre-warmed LB medium and grown to an OD₆₀₀ ~ 0.3. Then, the t₀ sample was collected and aTc was immediately added to a final concentration of 10 nM. Samples of each culture were collected at 5, 10, 15 and 20 min after the addition of aTc. The sampling period was deliberately shorter than the doubling time of the bacterial strain and total RNA was prepared as previously described (Bury-Mone et al. 2009).

RNA tagging and sequencing

RNA tagging was performed as previously described (Fouquier d'Herouel et al. 2011; Innocenti et al. 2015). Briefly for each sample, 20 µg of total RNA was ligated with the RNA adaptor PSS-tag (5'-GCAUAGGGGUAAA-3') using the T4 RNA ligase I (New England Biolabs). Samples were then treated with the tobacco alkaline phosphatase (TAP; EPICENTRE Biotechnologies) and ligated to the TSS-tag RNA adaptor (5'-GCGAGACUGAGAA-3'). Since 5' RNA ends can be tagged by both RNA adaptors and may provide ambiguities in assignments (5' termini coined 'UNDs') (Innocenti et al. 2015), we performed at t₀ for each of the three experimental series, a transcriptome where the TAP treatment was omitted. The comparison between samples treated with TAP and untreated enabled us to assign TSSs otherwise classified as 'UND' (Supplemental section S2). The 18 RNA samples (5 time points for each of the 3 kinetic series, and the 3 untreated TAP at t₀) were sequenced on a SOLiD 5500W technology instrument (Thermo Fisher Scientific).

Alignment, coverage, gene expression level, 5' tag detection and normalization

We applied an updated version of our previously described procedure (Innocenti et al. 2015). 5' and 3' sequencing adapters (Supplemental Table S6) were first stripped in colorspace using cutadapt v1.16 (Martin 2011), retaining sequences of ≥ 30 nucleotides (options -c -m 30 -g file:Solid_5prime.fasta -a file:Solid_3prime.fasta). Remaining sequences were detagged in colorspace of TSS- and PSS-tags using 'cutadapt,' retaining sequences ≥ 20 nucleotides (option -c -m 20), resulting in reads sorted as TSS, PSS, or unknown. These reads were then aligned to the genome of the *E. coli* K12 strain MG1655 [GenBank: U00096.3] using Bowtie v1.2.2 (Langmead et al. 2009). Alignments were used for calculating expression levels with Cuffdiff v2.2.1 (Trapnell et al. 2010). Ribosomal RNA regions were masked and appropriate sequence corrections were applied (options -M U00096_rRNA.gtf --multi-read-correct --library-type fr-secondstrand --frag-bias-correct). The remaining analysis was done as in (Innocenti et al. 2015). Normalized data are available in numerical format in Supplemental Table S1 (RNA levels) and S5 (tag counts).

Assignments of transcription starts and RNA processing sites

A 5' RNA end was considered mapped when at least 42 tags in total were found in the 18 transcriptomes. As RNAP can initiate transcription at a few consecutive nucleotides within a unique promoter (Robb et al. 2013; Vvedenskaya et al. 2015), each 5' RNA end mapped was treated individually. Determination of the nature of 5' RNA ends, assigned as 'TSS', 'PSS' or 'UND', was based on tag-counts and TAP treatments. Further details are provided in Supplemental section S2.

Changes in transcription initiation and RNA processing frequencies

Changes in the frequency of processes generating 5' RNA ends was postulated as null before the induction of σ^E at t_0 . Tag-counts reflect the relative abundance of a given 5' RNA end (Innocenti et al. 2015). The tagging-rate, R_i , is the number of tags appearing or disappearing per time unit and reflects the frequency of the process generating a given 5' RNA end. Changes in the frequency was inferred by the average variation of R_i between two consecutive time points, a magnitude that we termed 'tagging-acceleration' (A_i).

Statistical analysis

Reported TSS and PSS numbers gave rise to corresponding empirical cumulative distribution functions, yielding empirical p-values for the number of sites detected at each genomic position. Rate p-values were obtained from Fisher's combined probability test by establishing the test statistic $X_4^2 \sim -2 (\ln p_n + \ln(1 - p_{n-1}))$, which then was evaluated against the chi-squared distribution with 4 degrees of freedom. $n \in \{1,2,3,4\}$ designates each time point following the initial one, i.e. 5, 10, 15, and 20 min, and $p_0 = 0$. Acceleration p-values were similarly calculated by aggregating rate p-values. Results with p-values $< 1.45 \cdot 10^{-5}$ were considered significant (Bonferroni correction).

Consistency between kinetic series

Log-ratios of total tag-counts between consecutive time points were calculated for each reported position within each kinetic series. Observations at a given position and time point were called 'point-wise consistent' if ratios from all three kinetic series showed the same trend (i.e. were all positive or zero and negative or zero, respectively, with at least one non-zero observation). We allowed for a pseudo-count of one in the calculation of total tag log-ratios to avoid numerical divergence without introducing bias. Observations at positions with all four time-points reproducible were called 'globally consistent'. Thus, for all 3447 mapped ends, we analyzed the consistency of tag-counts between kinetic series as provided in Supplemental Table S4.

DATA DEPOSITION

The data are available in NCBI SRA repository (PRJNA561076).

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article. Supplemental_Material.doc: Supplementary text containing sections S1 to S4, as mentioned in the main text. Supplemental_Table_S1.xls: RNA levels for individual series, mean and ratios at each experimental time point. Supplemental_Table_S2.xls: Changes in RNA levels at each experimental time point. Supplemental_Table_S3.xls: Changes in RNA levels for targets of MicA, RybB and MicL. Supplemental_Table_S4.xls: Tag-counts for individual

series, coordinates, 5' RNA ends assignment, statistics and mean values. Supplemental_Table_S5.xls: 'Ai' and 'Ri' values and statistics. Supplemental_Table_S6.xls: Sequencing adaptors.

ACKNOWLEDGMENTS

We are grateful to the anonymous reviewers for constructive comments. We acknowledge Benoit Quinquis for technical assistance in RNA-seq, the MetaGenoPolis platform for sequencing services, and members of 'CPE team' for comments and insightful discussions. We thank Nicolas Crapart and Claudia Bevilacqua at the @BRIDGE platform (INRAE, Jouy en Josas) for advices and guidance on quality control of RNA samples. We acknowledge administrative and technical staffs at Micalis (Jouy en Josas, France), KTH (Stockholm, Sweden) and I2BC-Paris-Saclay (Orsay, France). This work was supported by grant ANR-12-BSV6-0008 (ReadRNA) from the Agence Nationale de la Recherche for FR and PB, the Academy of Finland as part of Finland Distinguished Professor program (project 129024/Aurell) for EA, and institutions INRAE (France), CNRS (France) and KTH (Sweden).

Authors' contributions: CL, FW, NI, TR, PS, EA, PB, FR: Design the project and provide insights on experimental plans; CL, FW, CB, RB, FR: RNA preparation and RNA-tagging; CL, AFD, FW, FR: data analyses; AFD, NI, EA: RNA-seq raw data analysis; SPK: RNA sequencing and English editing; CL, FW, AFD, PS, TR, EA, PB, FR: article writing and illustrations; FR: project coordination. All authors read and approved the final manuscript.

REFERENCES

- Ades SE. 2008. Regulation by destruction: design of the sigmaE envelope stress response. *Curr Opin Microbiol* **11**: 535-540.
- Anderson KL, Dunman PM. 2009. Messenger RNA Turnover Processes in *Escherichia coli*, *Bacillus subtilis*, and Emerging Studies in *Staphylococcus aureus*. *Int J Microbiol* **2009**: 525491.
- Anderson KL, Roberts C, Disz T, Vonstein V, Hwang K, Overbeek R, Olson PD, Projan SJ, Dunman PM. 2006. Characterization of the *Staphylococcus aureus* heat shock, cold shock, stringent, and SOS responses and their effects on log-phase mRNA turnover. *J Bacteriol* **188**: 6739-6756.
- Bouloc P, Repoila F. 2016. Fresh layers of RNA-mediated regulation in Gram-positive bacteria. *Curr Opin Microbiol* **30**: 30-35.
- Burton ZF, Gross CA, Watanabe KK, Burgess RR. 1983. The operon that encodes the sigma subunit of RNA polymerase also encodes ribosomal protein S21 and DNA primase in *E. coli* K12. *Cell* **32**: 335-349.
- Bury-Mone S, Nomane Y, Reymond N, Barbet R, Jacquet E, Imbeaud S, Jacq A, Bouloc P. 2009. Global analysis of extracytoplasmic stress signaling in *Escherichia coli*. *PLoS Genet* **5**: e1000651.
- Chao Y, Li L, Girodat D, Forstner KU, Said N, Corcoran C, Smiga M, Papenfort K, Reinhardt R, Wieden HJ et al. 2017. *In Vivo* Cleavage Map Illuminates the Central Role of RNase E in Coding and Non-coding RNA Pathways. *Mol Cell* **65**: 39-51.
- Chen H, Shiroguchi K, Ge H, Xie XS. 2015. Genome-wide study of mRNA degradation and transcript elongation in *Escherichia coli*. *Mol Syst Biol* **11**: 781.
- Chong S, Chen C, Ge H, Xie XS. 2014. Mechanism of transcriptional bursting in bacteria. *Cell* **158**: 314-326.
- Chow J, Dennis PP. 1994. Coupling between mRNA synthesis and mRNA stability in *Escherichia coli*. *Mol Microbiol* **11**: 919-931.
- Conway T, Creecy JP, Maddox SM, Grissom JE, Conkle TL, Shadid TM, Teramoto J, San Miguel P, Shimada T, Ishihama A et al. 2014. Unprecedented high-resolution view of bacterial operon architecture revealed by RNA sequencing. *MBio* **5**: e01442-01414.
- Croucher NJ, Thomson NR. 2010. Studying bacterial transcriptomes using RNA-seq. *Curr Opin Microbiol* **13**: 619-624.
- Dahan N, Choder M. 2013. The eukaryotic transcriptional machinery regulates mRNA translation and decay in the cytoplasm. *Biochim Biophys Acta* **1829**: 169-173.
- Dar D, Sorek R. 2018. Extensive reshaping of bacterial operons by programmed mRNA decay. *PLoS Genet* **14**: e1007354.
- Dressaire C, Pobre V, Laguerre S, Girbal L, Arraiano CM, Coccagn-Bousquet M. 2018. PNPase is involved in the coordination of mRNA degradation and expression in stationary phase cells of *Escherichia coli*. *BMC Genomics* **19**: 848.
- Esquerre T, Laguerre S, Turlan C, Carpousis AJ, Girbal L, Coccagn-Bousquet M. 2014. Dual role of transcription and transcript stability in the regulation of gene expression in *Escherichia coli* cells cultured on glucose at different growth rates. *Nucleic Acids Res* **42**: 2460-2472.
- Evans C, Hardin J, Stoebel DM. 2018. Selecting between-sample RNA-Seq normalization methods from the perspective of their assumptions. *Brief Bioinform* **19**: 776-792.
- Filiatrault MJ. 2011. Progress in prokaryotic transcriptomics. *Curr Opin Microbiol* **14**: 579-586.
- Fouquier d'Herouel A, Wessner F, Halpern D, Ly-Vu J, Kennedy SP, Serror P, Aurell E, Repoila F. 2011. A simple and efficient method to search for selected primary transcripts: non-coding and antisense RNAs in the human pathogen *Enterococcus faecalis*. *Nucleic Acids Res* **39**: e46.
- Gama-Castro S, Salgado H, Santos-Zavaleta A, Ledezma-Tejeida D, Muniz-Rascado L, Garcia-Sotelo JS, Alquicira-

- Hernandez K, Martinez-Flores I, Pannier L, Castro-Mondragon JA et al. 2016. RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic Acids Res* **44**: D133-143.
- Gogol EB, Rhodius VA, Papenfort K, Vogel J, Gross CA. 2011. Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon. *Proc Natl Acad Sci U S A* **108**: 12875-12880.
- Golding I, Paulsson J, Zawilski SM, Cox EC. 2005. Real-time kinetics of gene activity in individual bacteria. *Cell* **123**: 1025-1036.
- Gorski K, Roch JM, Prentki P, Krisch HM. 1985. The stability of bacteriophage T4 gene 32 mRNA: a 5' leader sequence that can stabilize mRNA transcripts. *Cell* **43**: 461-469.
- Grabowicz M, Silhavy TJ. 2016. Envelope Stress Responses: An Interconnected Safety Net. *Trends Biochem Sci*.
- Guell M, Yus E, Lluch-Senar M, Serrano L. 2011. Bacterial transcriptomics: what is beyond the RNA hori-zome? *Nat Rev Microbiol* **9**: 658-669.
- Guisbert E, Rhodius VA, Ahuja N, Witkin E, Gross CA. 2007. Hfq modulates the sigmaE-mediated envelope stress response and the sigma32-mediated cytoplasmic stress response in *Escherichia coli*. *J Bacteriol* **189**: 1963-1973.
- Guo MS, Updegrove TB, Gogol EB, Shabalina SA, Gross CA, Storz G. 2014. MicL, a new sigmaE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. *Genes Dev* **28**: 1620-1634.
- Haugel-Nielsen J, Hajnsdorf E, Regnier P. 1996. The *rpsO* mRNA of *Escherichia coli* is polyadenylated at multiple sites resulting from endonucleolytic processing and exonucleolytic degradation. *Embo J* **15**: 3144-3152.
- Innocenti N, Golumbeanu M, Fouquier d'Herouel A, Lacoux C, Bonnin RA, Kennedy SP, Wessner F, Serron P, Bouloc P, Repoila F et al. 2015. Whole-genome mapping of 5' RNA ends in bacteria by tagged sequencing: a comprehensive view in *Enterococcus faecalis*. *Rna* **21**: 1018-1030.
- Jishage M, Iwata A, Ueda S, Ishihama A. 1996. Regulation of RNA polymerase sigma subunit synthesis in *Escherichia coli*: intracellular levels of four species of sigma subunit under various growth conditions. *J Bacteriol* **178**: 5447-5451.
- Jones D, Elf J. 2018. Bursting onto the scene? Exploring stochastic mRNA production in bacteria. *Curr Opin Microbiol* **45**: 124-130.
- Kaern M, Elston TC, Blake WJ, Collins JJ. 2005. Stochasticity in gene expression: from theories to phenotypes. *Nat Rev Genet* **6**: 451-464.
- Keseler IM, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martinez C, Caspi R, Fulcher C, Gama-Castro S, Kothari A, Krummenacker M et al. 2017. The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res* **45**: D543-D550.
- Kristoffersen SM, Haase C, Weil MR, Passalacqua KD, Niazi F, Hutchison SK, Desany B, Kolsto AB, Tourasse NJ, Read TD et al. 2012. Global mRNA decay analysis at single nucleotide resolution reveals segmental and positional degradation patterns in a Gram-positive bacterium. *Genome Biol* **13**: R30.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**: R25.
- Laub MT, McAdams HH, Feldblyum T, Fraser CM, Shapiro L. 2000. Global analysis of the genetic network controlling a bacterial cell cycle. *Science* **290**: 2144-2148.
- Levine E, Zhang Z, Kuhlman T, Hwa T. 2007. Quantitative characteristics of gene regulation by small RNA. *PLoS Biol* **5**: e229.
- Linder P, Lemeille S, Redder P. 2014. Transcriptome-Wide Analyses of 5'-Ends in RNase J Mutants of a Gram-Positive Pathogen Reveal a Role in RNA Maturation, Regulation and Degradation. *PLoS Genet* **10**: e1004207.

- Ludwig H, Homuth G, Schmalisch M, Dyka FM, Hecker M, Stulke J. 2001. Transcription of glycolytic genes and operons in *Bacillus subtilis*: evidence for the presence of multiple levels of control of the *gapA* operon. *Mol Microbiol* **41**: 409-422.
- Mader U, Nicolas P, Richard H, Bessieres P, Aymerich S. 2011. Comprehensive identification and quantification of microbial transcriptomes by genome-wide unbiased methods. *Curr Opin Biotechnol* **22**: 32-41.
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetJournal* **17**: 10-12.
- Masse E, Vanderpool CK, Gottesman S. 2005. Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol* **187**: 6962-6971.
- Mauri M, Klumpp S. 2014. A model for sigma factor competition in bacterial cells. *PLoS Comput Biol* **10**: e1003845.
- McAdams HH, Arkin A. 1997. Stochastic mechanisms in gene expression. *Proc Natl Acad Sci U S A* **94**: 814-819.
- McIntyre LM, Lopiano KK, Morse AM, Amin V, Oberg AL, Young LJ, Nuzhdin SV. 2011. RNA-seq: technical variability and sampling. *BMC Genomics* **12**: 293.
- Mohanty BK, Kushner SR. 2016. Regulation of mRNA Decay in Bacteria. *Annu Rev Microbiol* **70**: 25-44.
- Mosteller RD, Yanofsky C. 1970. Transcription of the tryptophan operon in *Escherichia coli*: rifampicin as an inhibitor of initiation. *J Mol Biol* **48**: 525-531.
- Mutalik VK, Nonaka G, Ades SE, Rhodius VA, Gross CA. 2009. Promoter strength properties of the complete sigma E regulon of *Escherichia coli* and *Salmonella enterica*. *J Bacteriol* **191**: 7279-7287.
- Newbury SF, Smith NH, Higgins CF. 1987. Differential mRNA stability controls relative gene expression within a polycistronic operon. *Cell* **51**: 1131-1143.
- Nouaille S, Mondeil S, Finoux AL, Moulis C, Girbal L, Coccain-Bousquet M. 2017. The stability of an mRNA is influenced by its concentration: a potential physical mechanism to regulate gene expression. *Nucleic Acids Res* **45**: 11711-11724.
- Olson PD, Kuechenmeister LJ, Anderson KL, Daily S, Beenken KE, Roux CM, Reniere ML, Lewis TL, Weiss WJ, Pulse M et al. 2011. Small molecule inhibitors of *Staphylococcus aureus* RnpA alter cellular mRNA turnover, exhibit antimicrobial activity, and attenuate pathogenesis. *PLoS Pathog* **7**: e1001287.
- Peck SA, Hughes KD, Victorino JF, Mosley AL. 2019. Writing a wrong: Coupled RNA polymerase II transcription and RNA quality control. *Wiley Interdiscip Rev RNA* **10**: e1529.
- Phadtare S, Severinov K. 2010. RNA remodeling and gene regulation by cold shock proteins. *RNA Biol* **7**: 788-795.
- Prados J, Linder P, Redder P. 2016. TSS-EMOTE, a refined protocol for a more complete and less biased global mapping of transcription start sites in bacterial pathogens. *BMC Genomics* **17**: 849.
- Prevost K, Desnoyers G, Jacques JF, Lavoie F, Masse E. 2011. Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes Dev* **25**: 385-396.
- Raj A, van Oudenaarden A. 2008. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**: 216-226.
- Redon E, Loubiere P, Coccain-Bousquet M. 2005. Role of mRNA stability during genome-wide adaptation of *Lactococcus lactis* to carbon starvation. *J Biol Chem* **280**: 36380-36385.
- Repoila F, Gottesman S. 2001. Signal transduction cascade for regulation of RpoS: temperature regulation of DsrA. *J Bacteriol* **183**: 4012-4023.
- Reznikoff WS, Siegele DA, Cowing DW, Gross CA. 1985. The regulation of transcription initiation in bacteria. *Annu Rev Genet* **19**: 355-387.
- Rhodius VA, Mutalik VK. 2010. Predicting strength and function for promoters of the *Escherichia coli* alternative sigma factor, sigmaE. *Proc Natl Acad Sci U S A* **107**: 2854-2859.

- Rhodium VA, Mutalik VK, Gross CA. 2012. Predicting the strength of UP-elements and full-length *E. coli* sigmaE promoters. *Nucleic Acids Res* **40**: 2907-2924.
- Rhodium VA, Suh WC, Nonaka G, West J, Gross CA. 2006. Conserved and variable functions of the sigmaE stress response in related genomes. *PLoS Biol* **4**: e2.
- Robb NC, Cordes T, Hwang LC, Gryte K, Duchi D, Craggs TD, Santoso Y, Weiss S, Ebright RH, Kapanidis AN. 2013. The transcription bubble of the RNA polymerase-promoter open complex exhibits conformational heterogeneity and millisecond-scale dynamics: implications for transcription start-site selection. *J Mol Biol* **425**: 875-885.
- Rochat T, Bouloc P, Repoila F. 2013. Gene expression control by selective RNA processing and stabilization in bacteria. *FEMS Microbiol Lett* **344**: 104-113.
- Sharma CM, Vogel J. 2014. Differential RNA-seq: the approach behind and the biological insight gained. *Curr Opin Microbiol* **19**: 97-105.
- Shimada T, Tanaka K, Ishihama A. 2017. The whole set of the constitutive promoters recognized by four minor sigma subunits of *Escherichia coli* RNA polymerase. *PLoS One* **12**: e0179181.
- Silhavy TJ, Kahne D, Walker S. 2010. The bacterial cell envelope. *Cold Spring Harb Perspect Biol* **2**: a000414.
- Singh G, Pratt G, Yeo GW, Moore MJ. 2015. The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion. *Annu Rev Biochem* **84**: 325-354.
- So LH, Ghosh A, Zong C, Sepulveda LA, Segev R, Golding I. 2011. General properties of transcriptional time series in *Escherichia coli*. *Nat Genet* **43**: 554-560.
- Szittyá G, Moxon S, Pantaleo V, Toth G, Rusholme Pilcher RL, Moulton V, Burgyan J, Dalmay T. 2010. Structural and functional analysis of viral siRNAs. *PLoS Pathog* **6**: e1000838.
- Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**: 511-515.
- Urban JH, Vogel J. 2008. Two seemingly homologous noncoding RNAs act hierarchically to activate *glmS* mRNA translation. *PLoS Biol* **6**: e64.
- Vvedenskaya IO, Zhang Y, Goldman SR, Valenti A, Visone V, Taylor DM, Ebright RH, Nickels BE. 2015. Massively Systematic Transcript End Readout, "MASTER": Transcription Start Site Selection, Transcriptional Slippage, and Transcript Yields. *Mol Cell* **60**: 953-965.
- Winkler WC, Nahvi A, Roth A, Collins JA, Breaker RR. 2004. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**: 281-286.
- Yajnik V, Godson GN. 1993. Selective decay of *Escherichia coli* *dnaG* messenger RNA is initiated by RNase E. *J Biol Chem* **268**: 13253-13260.

TABLE 1 _ Lacoux et al.,

	Ends	No Changes (%)	Changes (%)	Changes distribution (%)			
				t ₅	t ₁₀	t ₁₅	t ₂₀
TSS	1147	45.1	54.9	23.1	17.0	10.7	4.1
PSS	594	58.8	41.2	9.8	13.6	11.3	6.6
UND	1706	74.5	25.5	4.7	10.1	5.3	5.3
Total	3447	62.0	38.0	37.6	40.7	27.3	16

TABLE 1. 5' RNA ends showing changes in frequency. Column 'Ends': Number of 5' termini mapped and assigned as TSS, PSS or UND; 'No Changes (%)': Percentage of 5' ends with no changes in frequency during the course of the experiment ($|A_i - A_{i-1}| \leq 0.7 \text{ tag/min}^2$); 'Changes (%)': Percentage of ends with changes in frequency during the course of the experiment; 'Changes distribution (%)': Percentage of ends with changes in frequency selected at 5 (t₅), 10 (t₁₀), 15 (t₁₅) and 20 (t₂₀) min. All percentage values are relative to the number of 5' ends mapped. Line 'Total': Total in each column.

TABLE 2 _ Lacoux et al.

σ^E -TSS reported at EcoCyc and assigned as TSS in this study								
Prom.	Coord.	Strand	Assigned	A ₀	A ₅	A ₁₀	A ₁₅	A ₂₀
<i>P_{bamAp2}</i>	197026	+	TSS	0	13.2	-15	22.2	-35
<i>P_{sbnA}</i>	396553	+	TSS (396554)	0	3.1	-4.6	6.8	-1.3
<i>P_{clpX}</i>	457202	+	TSS	0	4.2	5.6	2.9	-2.6
<i>P_{ybfG}</i>	716641	-	TSS (716639)	0	0.5	3.3	-3	-0.8
<i>P_{rybB}</i>	888057	-	TSS (888055)	0	7.3	-5.1	21.9	-12
<i>P_{opgGH}</i>	1109257	+	TSS (1109255)	0	6.7	7.2	-5.2	9
<i>P_{ydhLJK}</i>	1720998	+	TSS	0	0.3	-0.1	0.3	0.3
<i>P_{bepA}</i>	2616068	+	TSS	0	4.7	-1.4	1.3	2.7
<i>P_{rseAp3}</i>	2709632	-	TSS (2709630)	0	20.1	-23	91.4	-94
<i>P_{rpoEp2a}</i>	2710087	-	TSS (2710086)	0	1.4	-1.5	0.6	-0.4
<i>P_{bamDp2}</i>	2735961	+	TSS (2735960)	0	1.8	3	-0.8	3.1
<i>P_{bamE}</i>	2753486	+	TSS (2753488)	0	21.1	-18	0.2	16
<i>P_{rmlB}</i>	2766707	+	TSS (2766706)	0	0.1	1.63	-2.1	2.4
<i>P_{micA}</i>	2814802	+	TSS	0	20	-12	60.7	-72
<i>P_{dsbCp2}</i>	3039633	-	TSS (3039631)	0	0.6	0.8	0.8	-1.2
<i>P_{yggM}</i>	3101798	-	TSS	0	18.4	-0.2	2.2	21
<i>P_{ygiMcca}</i>	3201042	+	TSS	0	0.4	0.1	0.6	-1.5
<i>P_{yrap2}</i>	3296072	+	TSS	0	3.7	2.8	6.4	-3.3
<i>P_{yhjJ}</i>	3682016	-	TSS (3682013)	0	0.2	-0.1	0.7	-1.2
<i>P_{eptB}</i>	3710581	-	TSS (3710579)	0	0.8	0	0.9	-1.3
<i>P_{plsB}</i>	4256598	-	TSS (4256596)	0	3.2	-1.9	6.6	-5.7
<i>P_{psd}</i>	4390638	-	TSS (4390634)	0	0.5	0.7	1.3	-2.9
σ^E -TSS reported at EcoCyc and assigned as UND in this study								
<i>P_{ahpF}</i>	639002	+	UND	0	0	0.1	-0.1	0.1
<i>P_{yeaY}</i>	1890561	-	UND (1890560)	0	0.1	0.2	-0.3	0.1
<i>P_{micL}</i>	1958747	+	UND (1958745)	0	0.8	0.3	-0.5	-1.7
<i>P_{sixA}</i>	2456999	-	UND (2456997)	0	2.5	-2.5	3.4	2.6
<i>P_{lpxP}</i>	2495578	+	UND (2495581)	0	2.2	-1.4	0.8	2.3
<i>P_{yfeKS}</i>	2537315	+	UND	0	4.6	-0.9	0.7	-5
<i>P_{yfeY}</i>	2551243	-	UND	0	0	0.2	0.1	-0.1
<i>P_{bamB der}</i>	2638975	-	UND	0	0	0.1	-0.2	0.2
<i>P_{bacA}</i>	3204175	-	UND (3204172)	0	0.6	-0.1	1.1	-2.2
<i>P_{greA}</i>	3328853	-	UND (3328850)	0	0	0.2	-0.3	0.1
<i>P_{lptAB rpoN}</i>	3343326	+	UND	0	0.2	0	0.5	0.2
<i>P_{yieE chrR}</i>	3893829	+	UND	0	3.6	0.7	-3	-1.3
<i>P_{yis}</i>	4112869	+	UND (4112871)	0	0.5	0.2	1.6	-0.7
σ^E -TSS reported at EcoCyc and assigned as PSS in this study								
<i>P_{degP}</i>	180845	+	PSS ^(#)	0	0.5	1.1	-0.2	-0.7
<i>P_{bamAp}</i>	197821	+	PSS (197820)	0	0.1	0	-0.1	0.2
<i>P_{lpxD}</i>	200960	+	PSS	0	0	0.6	-0.1	0.3
<i>P_{fkpA}</i>	3477525	-	PSS ^(#)	0	0.4	1.7	0	-1.3
<i>P_{rpoHp3}</i>	3600870	-	PSS ^(#)	0	0.5	-0.3	0	-0.3

TABLE. 2. Tagging-accelerations measured at 5' RNA ends mapped and previously reported as σ^E -TSSs. Columns '**Prom.**' and '**Coord.**' provide the promoter name and the nucleotide position of the TSS on the chromosome assigned at EcoCyc website, respectively. '**Strand**' indicates the transcription orientation provided by EcoCyc and found in this study. '**Assigned**' reports the assignment of 5' RNA ends determined in this study, and numbers in brackets provide the coordinates of the nucleotide tagged if different from the EcoCyc coordinate. '**A_i**': tagging-accelerations measured during kinetic intervals $t_{[i-(i-1)]}$. ^(#): Abundant tagging flanking the TSS reported at EcoCyc database.

FIGURE LEGENDS

FIGURE 1. Changes in transcription initiation frequency at σ^E -TSSs. **A.** Principle of the tagRNA-seq method used in a time course experiment. Aliquots of growing bacteria were harvested every 5 minutes over 20 minutes following the addition of aTc. Total RNA was extracted and sequenced by tagRNA-seq. RNA samples at t_0 were duplicated: one treated with tobacco alkaline phosphatase (+TAP); the other was not treated (TAP no). **B.** Number of CDS whose RNA level varies at least 4-fold compared to t_0 . The selection process is non-repetitive: A CDS selected at one time point was excluded from the selection process at subsequent time points. Total CDSs selected are shown by the blue histogram. Among those, the red histogram indicates the number of selected CDSs known or predicted to be transcribed by σ^E -RNAP. **C.** Heatmap correlating patterns of tagging-acceleration values (A_i) measured at σ^E -TSSs assigned over the course of the experiment (Table 2). *: A_i values are below the statistical confidence (pA_i) or differences in A_i values are below the biological threshold of variation imposed ($> |0.7| \text{ tag/min}^2$). Cap letters D, E, F and G refer to graphs on the left of the figure. Numbers attached to TSSs refer to chromosomal coordinates. **D.** Biphasic pattern of changes in TIF (A_i values expressed in tag/min^2) observed for about half of σ^E -TSSs mapped. **E, F and G,** variation from the biphasic patterns for certain σ^E -TSSs mapped. X-axis (min): Experimental time points of the kinetics. Y-axis: A_i values measured. Supplemental Fig. S3 shows the patterns of changes in TIF for reported σ^E -TSSs mapped as UNDs and PSSs in this study.

FIGURE 2. Changes in transcription initiation frequency at σ^E -independent TSSs. **A and B.** Heatmap correlating patterns of changes in TIF (A_i values) at selected TSSs mapped and also reported in (Gama-Castro et al. 2016; Keseler et al. 2017). **A.** Transcription activation; **B.** Transcription repression. **C.** Graphical representation of changes in TIF indicating transcription activation and repression for TSSs *dps*-848948 and *gatY*-2177231, respectively. **D.** RNA levels for *dps* and *gatY*. Y-axes (RNA): \log_2 of the ratio between RNA amounts at an experimental time point (t_i) and t_0 . Legends are otherwise identical to Fig. 1.

FIGURE 3. RNA dynamics at the *rpsU-dnaG-rpoD* locus. **A.** Locus organization. Black thick arrows indicate CDSs and transcription orientation. Thin horizontal arrows show TSSs mapped and coordinates on the *E. coli* K12 chromosome MG1655. The blue color for TSSs indicates significant changes in TIF during the σ^E -mediated adaptation; the grey color marks the absence of significant changes in TIF. Sigma factors known to be involved in the activity of promoters are in brackets. Vertical arrows indicate PSSs mapped, in bold and in color those with significant changes in PF, in grey, those with no significant changes. **B.** RNA levels over the course of the experiment. **C.** Pattern of changes in TIF at TSSs mapped; colors correspond to those used panel A. **D.** Pattern of changes in PF for PSSs mapped. Colors correspond to those used panel A. Legends are otherwise identical to Fig. 1 and Fig. 2.

FIGURE 4. RNA dynamics at the *ahpCF* locus. Legends are identical to Fig. 3. Note that the proximity of PSS-638922 and TSS-638921 does not allow us to rule out that the PSS is not a TSS. This PSS is presented in bold and grey in the panel A, and with dashed lines in the panel D.

Fig. 1: Lacoux et al

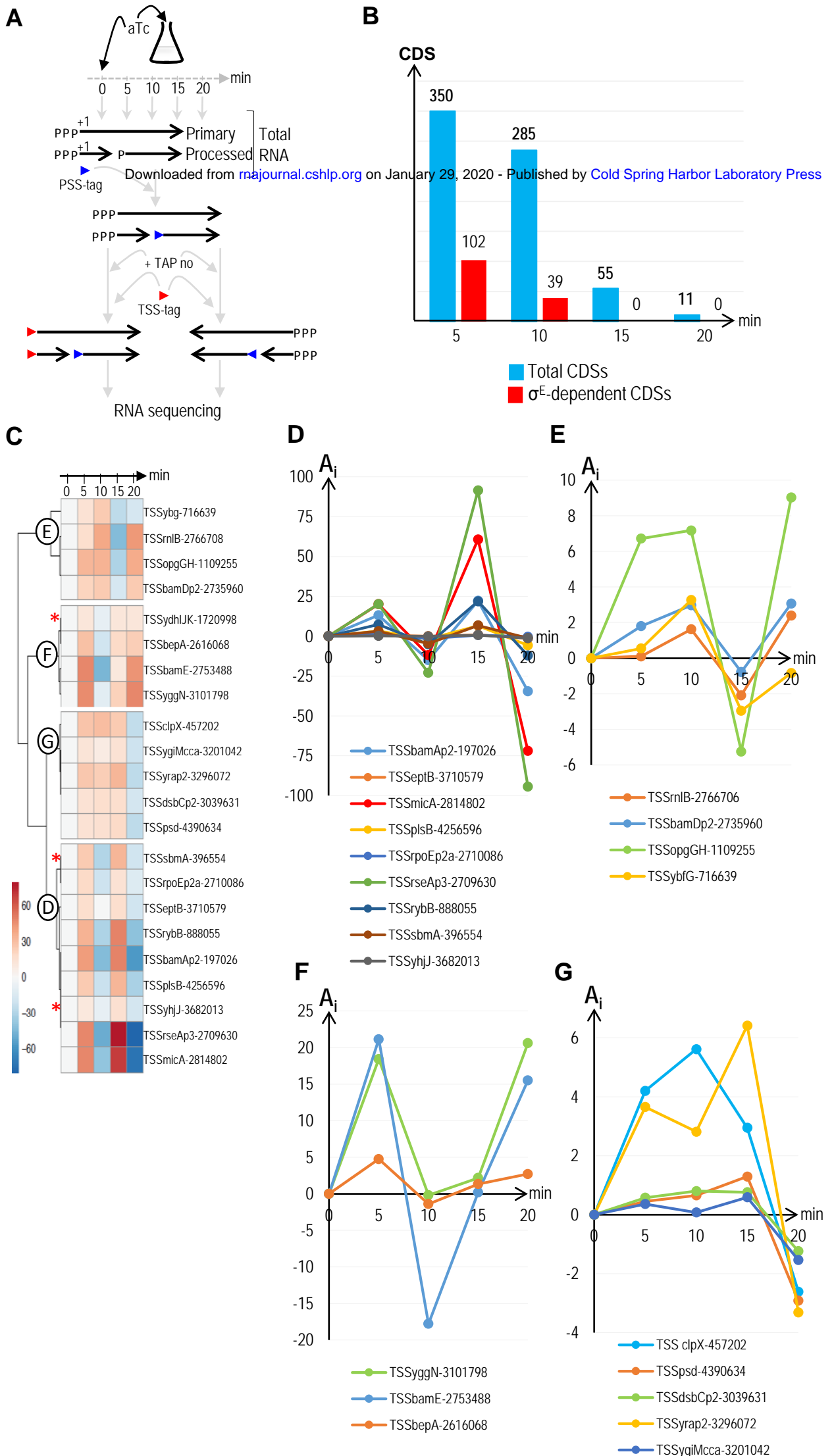


Fig. 2: Lacoux et al

Downloaded from rnajournal.cshlp.org on January 29, 2020 - Published by Cold Spring Harbor Laboratory Press

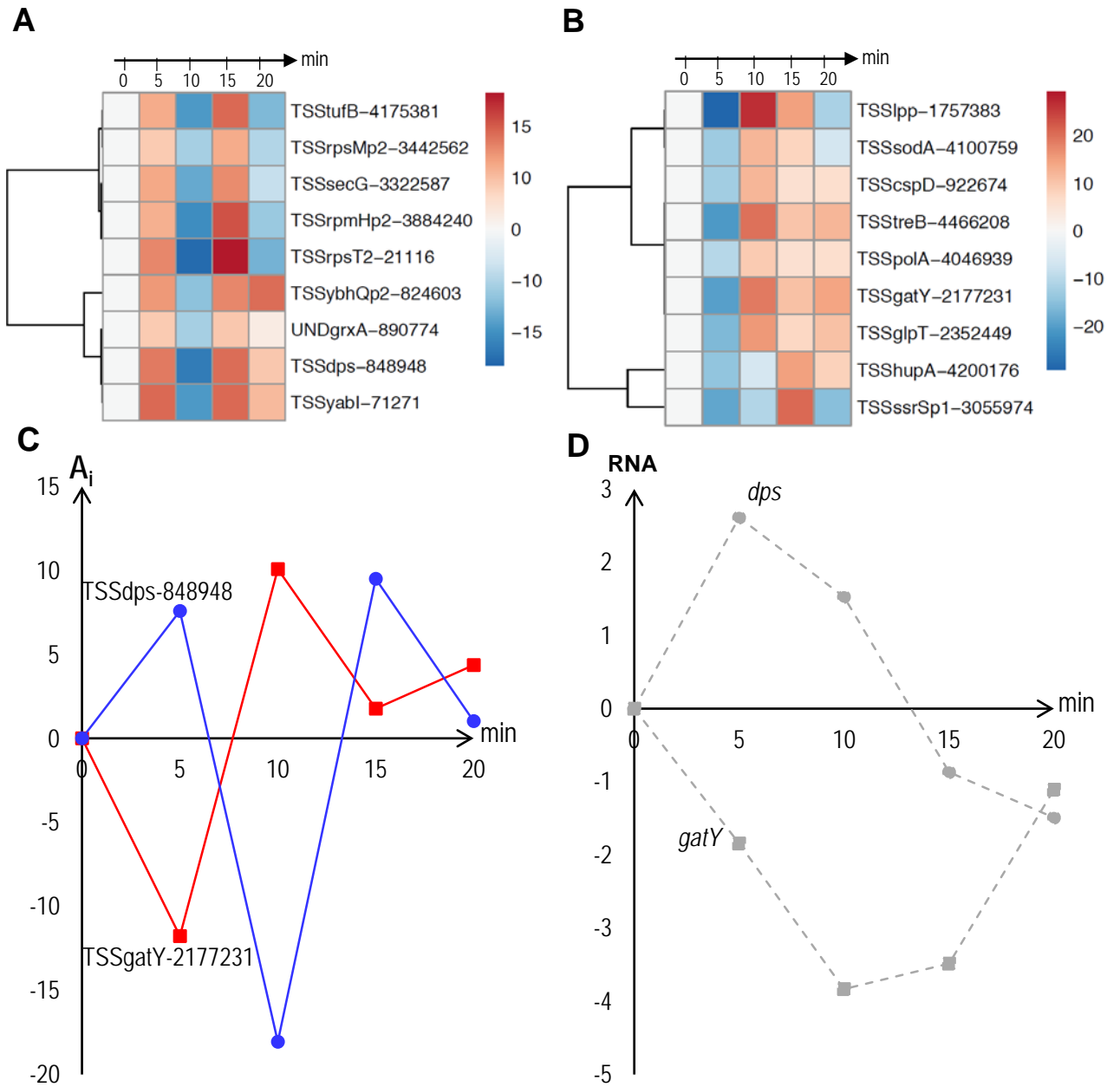


Fig. 3: Lacoux et al

Downloaded from rnajournal.cshlp.org on January 29, 2020 - Published by Cold Spring Harbor Laboratory Press

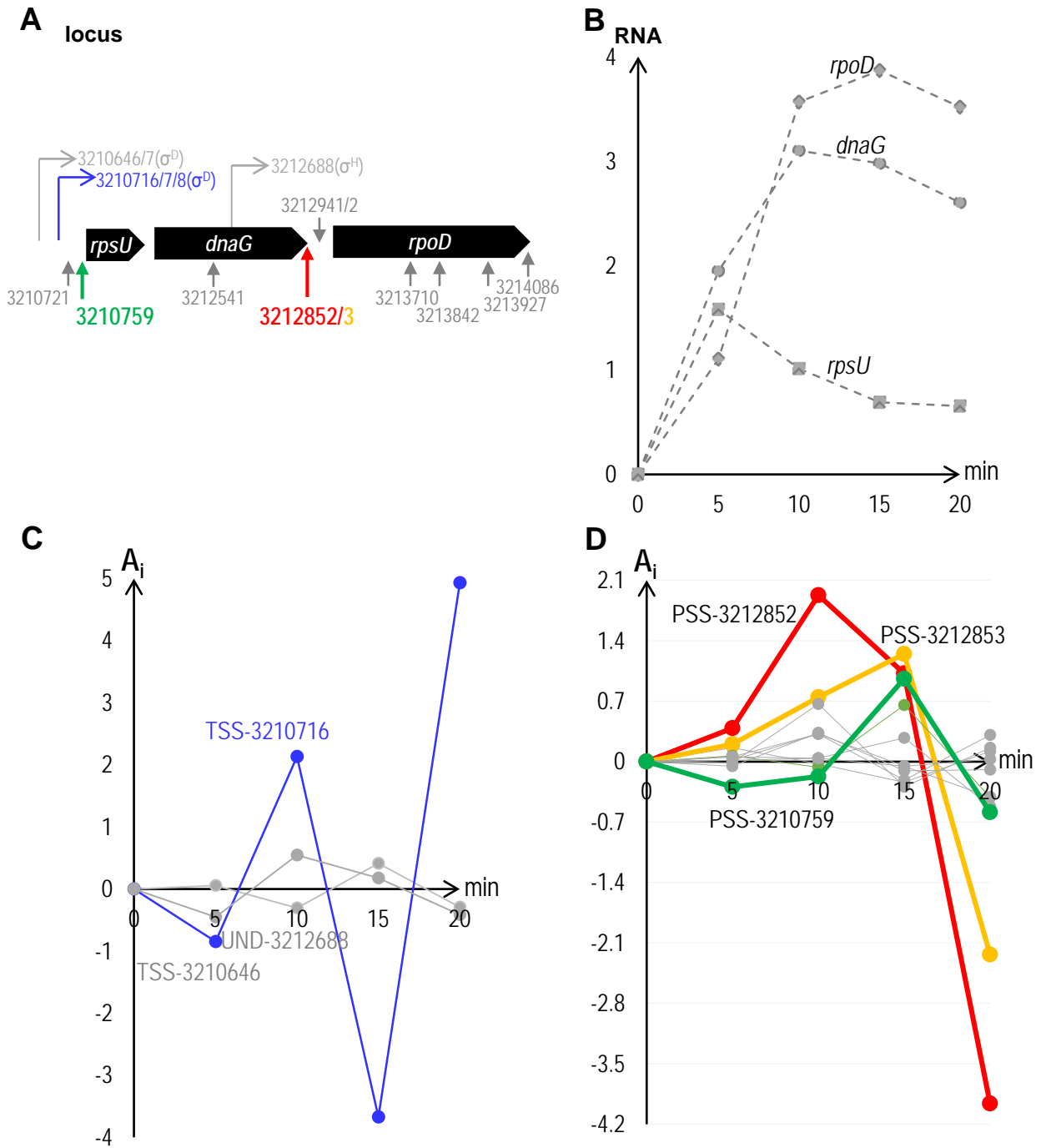
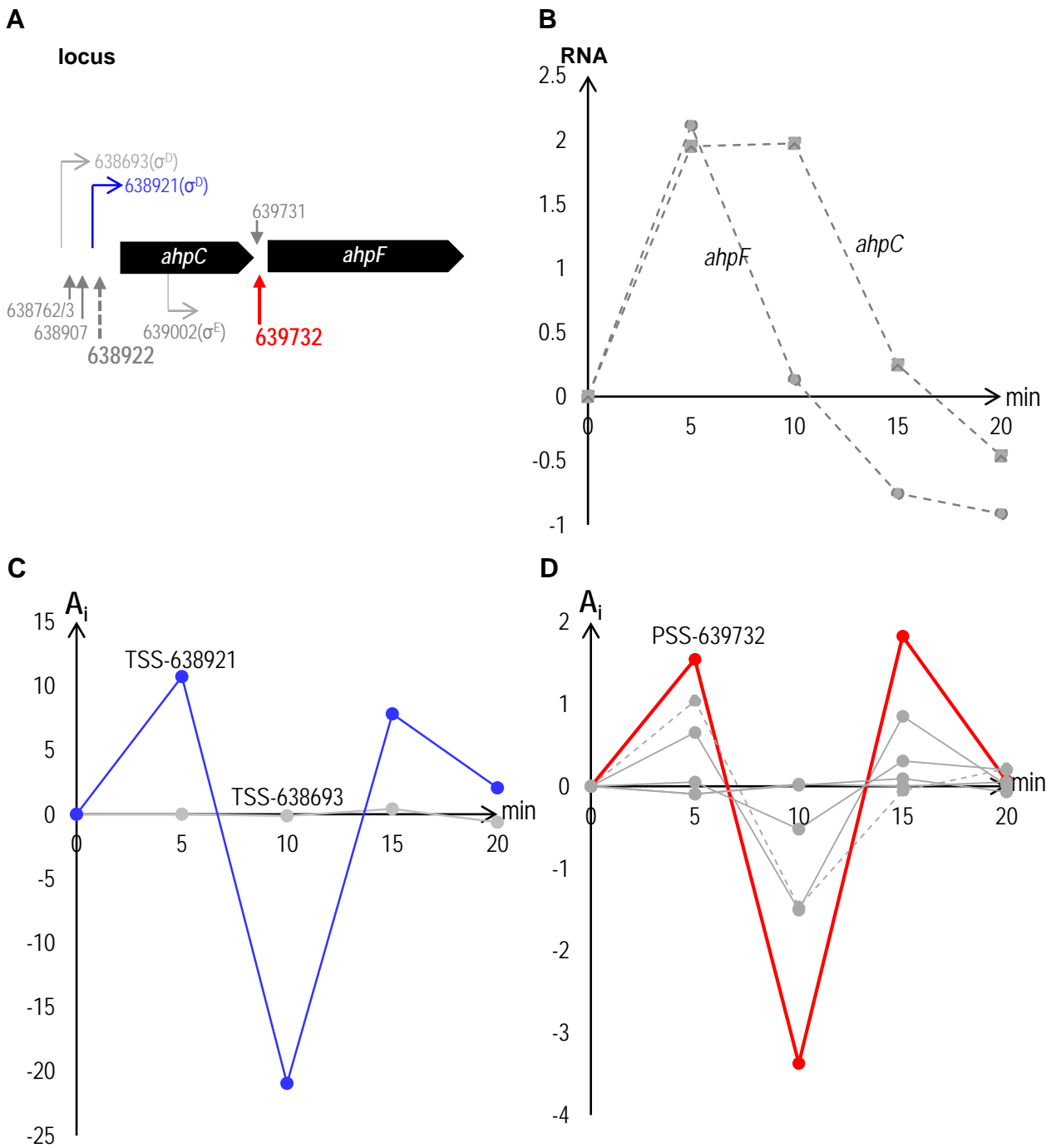


Fig. 4: Lacoux et al





RNA

A PUBLICATION OF THE RNA SOCIETY

Dynamic insights on transcription initiation and RNA processing during bacterial adaptation

Caroline Lacoux, Aymeric Fouquier d'Hérouël, Françoise Wessner- Le Bohec, et al.

RNA published online January 28, 2020

Supplemental Material <http://rnajournal.cshlp.org/content/suppl/2020/01/28/rna.073288.119.DC1>

P<P Published online January 28, 2020 in advance of the print journal.

Accepted Manuscript Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.

Creative Commons License This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see <http://rnajournal.cshlp.org/site/misc/terms.xhtml>). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Webinar Automation-friendly full-length scRNA-seq



To subscribe to *RNA* go to:
<http://rnajournal.cshlp.org/subscriptions>
