Molecular origins of transcriptional heterogeneity in diazotrophic *K. oxytoca*

Tufail Bashir^{1†}, Rowan D Brackston^{1†}, Christopher J Waite¹, Ioly Kotta-Loizou¹, Christoph Engl², Martin Buck^{1*}, Jörg Schumacher^{1*}

*For correspondence:

m.buck@imperial.ac.uk (MB);

j.schumacher@imperial.ac.uk (JS)

[†]These authors contributed equally to this work

¹Faculty of Natural Sciences, Department of Life Sciences, Imperial College London, London SW7 2AZ, UK; ²School of Biological & Chemical Sciences, Queen Mary University of London, London E1 4NS, UK

Abstract Phenotypic heterogeneity in clonal bacterial batch cultures is an important adaptive strategy to changing environments, including in diazotrophs with the unique capacity to convert di-nitrogen into bio-available ammonium. In diazotrophic *Klebsiella oxytoca* we simultaneously measured mRNA levels of key regulatory (*glnK-amtB*, *nifLA*) and structural (*nifHDK*) operons required for establishing nitrogen fixation, using dual molecule, single cell RNA-FISH. Through stochastic transcription models and mutual information analysis we revealed likely molecular origins for heterogeneity in nitrogenase expression. In wildtype and regulatory variant strains we inferred contributions from intrinsic and extrinsic noise, finding that *nifHDK* transcription is inherently bursty, but that noise propagation through signalling is also significant. The regulatory gene *glnK* had the highest discernible effect on *nifHDK* variance, while noise from factors outside of the regulatory pathway were negligible. Results provide evidence that heterogeneity is a fundamental property of this regulatory system, indicating potential constraints for engineering homogeneous nitrogenase expression.

Introduction

Cell to cell variability in transcription has been recognised across many different cell types, and attributed to a range of causes (Elowitz, 2002; Engl, 2019). In bacteria such variability has been suggested to underpin phenotypic differences, or heterogeneity, between otherwise genetically identical cells cultured under a particular condition. Understanding this phenotypic heterogeneity in bacterial clonal populations has important implications for medical and biotechnological applications (201, 2016).

Phenotypic heterogeneity may be particularly relevant in costly stress response systems, of which the response to nitrogen starvation is a key example (Schreiber et al., 2016). In organisms such as *Klebsiella oxytoca*, nitrogen starvation triggers a transition to diazotrophic behaviour in which bacterial cells use atmospheric di-nitrogen as their nitrogen source for growth (Dixon and Kahn, 2004). While this transition, and the associated transcriptional programme, is essential for continued growth under conditions deplete of fixed nitrogen, it is also very costly since the resultant ATP consuming nitrogenase enzyme may ultimately constitute up to 20 % of the proteome (Dixon and Kahn, 2004). As such, if fixed nitrogen sources soon become available again, it is likely advantageous to have not fully undergone the diazotrophic transition. Activation of the nitrogen stress response is potentially somewhat of a gamble in which the payoff depends strongly on future

41 conditions. Such a scenario may therefore indicate a requirement for bet-hedging strategies in
42 which heterogeneity proves advantageous at the population level (*van Boxtel et al., 2017*).

43 Regardless of whether there are advantages to heterogeneity, it is also important to establish
44 the underlying mechanistic causes, especially if there is an aim to ultimately modify cell behaviour.
45 Studies at single-cell resolution are crucial to this, enabling the full distribution of expression levels
46 to be obtained over a population of cells. Such data not only provides qualitative insight into the
47 variability between cells, but can also enable the inference of the underlying dynamical processes
48 (*Munsky et al., 2018*), through the combination of biological insight and quantitative models. This is
49 the approach we take here to investigate the sources of noise that lead to heterogeneity of *nifHDK*
50 gene expression at the transcriptional level.

51 In *K. oxytoca*, expression of a functional nitrogenase involves coordinated transcription of 18
52 *nif* genes that are organised in 5 operons. The *nifHDK* operon encodes the structural genes of the
53 nitrogenase and is the most highly expressed operon within the *nif* cluster. The core hierarchical
54 regulatory system of *nif* gene expression (Fig. 1(A)) consists of the nitrogen regulator NtrC activating
55 expression of *glnK-amtB* and *nifLA* operons, with NifA activating *nifHDK* gene expression when not
56 directly inhibited by NifL (*Dixon and Kahn, 2004*). Both NtrC and NifA are bacterial enhancer binding
57 proteins (EBP) that activate the major variant σ^{54} RNA polymerase, with a distinct ATPase dependent
58 activating mechanism compared with the canonical σ^{70} type RNA polymerases (*Schumacher et al.,*
59 *2006*). The inhibitory NifL-NifA complex is destabilised by (i) GlnK binding, (ii) a reduced state
60 of the NifL, (iii) ADP and (iv) α -ketoglutarate binding (reviewed in (*Dixon and Kahn, 2004*)). This
61 arrangement integrates signals conducive to nitrogen fixation, namely a reducing environment,
62 high energy levels and presumably low nitrogen levels, as high α -ketoglutarate levels indicate a
63 low nitrogen status in the closely related *E. coli* (*Schumacher et al., 2013*). A low nitrogen status,
64 defined as the ratio of glutamine/ α -ketoglutarate also increases the activity of NtrC and enhances
65 its expression by triggering uridylation of PII signalling proteins. Further, low glutamine levels affect
66 the post-translational uridylation state of GlnK, however it is unclear if the uridylation state affects
67 GlnK function in destabilising the NifL-NifA complex in *K. oxytoca* (*He et al., 1998*).

68 Phenotypic heterogeneity in nitrogen fixing *K. oxytoca* grown in chemostats has been demon-
69 strated under limiting ammonium availability and suggested to be caused by events downstream
70 of GlnK (*Schreiber et al., 2016*). A degree of phenotypic heterogeneity is known to result from the
71 inherent stochasticity in gene expression and has been widely observed in other systems (*Elowitz,*
72 *2002; Cai et al., 2006; Kiviet et al., 2014*). Such stochasticity is common to all chemical reaction
73 systems involving small numbers of molecules of which transcription is a key example. However
74 transcription is further observed to occur in bursts (*Jones and Elf, 2018; Golding et al., 2005; Suter*
75 *et al., 2011; Larson et al., 2013*), short periods of intense transcriptional activity, resulting in in-
76 creased levels of heterogeneity. Together, inherent stochasticity and burstiness lead to what is
77 often referred to as intrinsic noise which may be a fundamental property of transcription of a
78 given gene. However, it is understood that in addition to this intrinsic noise, other sources of noise
79 external to a particular gene may also be relevant. These additional contributions to heterogeneity,
80 referred to as extrinsic noise, have been observed experimentally by simultaneously measuring the
81 expression of two or more copies of the same gene at the single cell level (*Swain et al., 2002; Raser*
82 *and O'Shea, 2004; Gasch et al., 2017*). Correlations in the expression of these two gene copies
83 reflect perturbations that simultaneously affect both.

84 While noise contributions can sometimes be directly measured, it is also possible to make
85 inferences via a modelling approach. Intrinsic noise and bursty transcription have long been
86 modelled by the so-called Telegraph process that describes stochastic events within transcription
87 and predicts the probability distribution of mRNA copy numbers (*Ko, 1991; Peccoud and Ycart, 1995;*
88 *Raj et al., 2006; Iyer-Biswas et al., 2009*). This model and its variants have been widely used to
89 infer transcriptional properties (*Schreiber et al., 2016; Munsky et al., 2015; Sepúlveda et al., 2016;*
90 *Larsson et al., 2019*) and therefore provide improved analysis and understanding from experimental
91 data. Such models can also be readily extended to include the effect of extrinsic noise (*Lenive et al.,*

Figure 1. (A) Regulatory pathway governing *nifHDK* expression, adapted from *Dixon and Kahn (2004)*. Transcription of *nifHDK* is subject to a hierarchical regulatory system. (B) Population size during transition to diazotrophy in wild-type *K. oxytoca*. Following run-out of ammonia, cultures display arrested growth during the diazotrophic transition, marked in gray. Growth from ten hours onwards is achieved through nitrogen fixation. (C) Distribution of *nifHDK* transcript abundance at 8 hours. Almost all cells have non-zero expression levels, but there is significant variability across the population. The average expression level is not necessarily representative of the typical cell.

Figure 1-Figure supplement 1. Cumulative acetylene reduction comparison.

2016; Sherman et al., 2015; Ham et al., 2019), and subsequently quantify this contribution.

Here we are interested in the relative contributions of intrinsic and extrinsic noise to expression of the *nifHDK* operon, under conditions in which free living *K. oxytoca* cells transition to use atmospheric di-nitrogen as their nitrogen source for growth. Using smRNA-FISH and examining the relationship between genes at different levels in the regulatory cascade controlling nitrogenase gene expression, we are able to reveal the combined roles of intrinsic noise at the level of the *nifHDK* promoter and extrinsic noise arising from upstream regulation. By further fitting stochastic models for the transcription process, these contributions are quantified, along with additional details of the regulatory mechanisms.

Results

Transition to diazotrophy

To measure the full spectrum of heterogeneity pertaining to the transition into and establishment of full diazotrophy, precultures grown in nitrogen replete aerobic conditions were transferred to nitrogen free, anaerobic media at time zero, when no *nif* expression or acetylene reduction was detectable. We found that *K. oxytoca* (WT) populations subjected to these conditions first consumed residual ammonia then experienced a growth arrest, which lasted around 5 hours, (see Fig. 1(B)). This was followed by a resumption of growth using N_2 gas derived ammonia as their nitrogen source, evidenced using the established acetylene reduction activity profiles as a measure of bulk population nitrogenase enzyme activity in batch culture (Dilworth, 1966).

In order to investigate and verify the regulatory roles of *glnB*, *glnK* and *nifA*, we also grew cells lacking the positive regulator GlnK ($\Delta glnK$), cells lacking the enhancer binding protein NifA and its co-transcribed inhibitor NifL ($\Delta nifLA$), and cells lacking the upstream regulator GlnB ($\Delta glnB$), which controls the NtrB dependent phosphorylation of NtrC in the cells in diazotrophic behaviour. We conducted the same experiments on each mutant, tracking the population growth and Nitrogenase activity (see Fig. 1 S1). We found that the absence of GlnB had negligible effect. By contrast both the $\Delta glnK$ and $\Delta nifLA$ mutants showed markedly different behaviour to the wild-type, also confirming the known role of each. Some nitrogen fixation was evidenced in the absence of GlnK (30 % of that in WT at 9.5 hours), but the $\Delta nifLA$ mutant was unable to fix at any level.

Further to these population averaged observations, we established use of FISH in *K. oxytoca* to estimate single molecule mRNA levels for the *nifHDK* operon which encodes the iron protein complex of the nitrogenase enzyme. We followed the protocol outlined in (Skinner et al., 2013), measuring expression in two additional control samples, as detailed in materials and methods. Results indicated significant variation in *nifHDK* transcript abundance (see Fig. 1(C) for an example). As is common in gene expression data, the distribution is decidedly non-Gaussian, with a bulk of cells at low expression levels and a longer tail of much higher values. This means that the “average” cell has relatively high expression levels compared with the majority, and therefore that bulk measurements of *nifHDK* gene expression do not represent a typical cell.

Mutual information analysis confirms direct propagation of transcriptional noise to *nifHDK* via GlnK

Given the observation of significant heterogeneity in *nifHDK* gene expression, we sought to establish whether extrinsic noise is present, and if this noise arises from the regulatory pathway or from elsewhere. We therefore grew WT cells, $\Delta glnK$ and $\Delta glnB$ mutants, to levels where different bacterial cultures displayed equivalent cumulative levels of nitrogenase activity (as described in methods). Two sets of probes were used to measure *nifHDK* and simultaneously either *glnK* or *nifLA* expression in individual cells to monitor variability across these paired genes (see Fig. 2 S1). Generally speaking, expression co-variance between two genes can either occur because one gene has a direct (and relatively quick) influence on the other gene, or because both genes are simultaneously affected by another source of variability. This could be a shared upstream control protein, or “global” factors

Table 1. Mutual information with *nifHDK* transcript abundance. The entropy of *nifHDK* abundance gives a measure of the total variability, while the mutual information is the amount of variability that can be directly explained by the abundance of the other gene. Smaller significance levels indicate a stronger statistical indication of non-zero mutual information.

Genetic strain	Gene	Mutual information [bits]	Entropy <i>nifHDK</i> [bits]	Significance level
WT	<i>glnK</i>	0.188	2.25	0.00016
WT	<i>glnK</i>	0.340	2.44	0.00035
WT	<i>nifL</i>	0.275	2.87	0.62
WT	<i>nifL</i>	0.318	2.75	0.47
$\Delta glnB$	<i>glnK</i>	0.118	1.90	0.0
$\Delta glnB$	<i>glnK</i>	0.072	1.92	0.0074
$\Delta glnK$	<i>nifL</i>	0.067	0.69	0.28
$\Delta glnK$	<i>nifL</i>	0.096	1.30	0.047

such as RNA polymerase or sigma factor abundance that may influence many genes simultaneously and should therefore be evident as significant co-variance between any expressed pair of genes.

We used mutual information (MI) as a suitable metric for quantifying co-variance in gene expression. MI is commonly used to analyse single cell data in situations where relationships are complex, nonlinear and unknown, as no assumptions or prior knowledge are required (*Mc Mahon et al., 2014, 2015*). Here, we compute the MI between each pair of genes and a null distribution of MI values from which a corresponding significance level can be computed (see materials and methods). The MI can further be compared to the entropy of the distribution of *nifHDK* expression levels, in order to compare the mutual information with the total information content. The results of this analysis are displayed in Fig. 2 and Tab. 1.

The results demonstrate a modest level of MI between *glnK* and *nifHDK* but no statistically significant measure of MI between *nifLA* and *nifHDK*. Undetectable MI between *nifLA* and *nifHDK* suggests that global sources of extrinsic noise are not significant for these two genes, and by extension are unlikely to be significant for any other σ^{54} dependent genes. The measurable low-level MI between *glnK* and *nifHDK* is therefore suggestive of a direct propagation of noise from *glnK* to *nifHDK*. In other words, stochastic variability of *glnK* expression acts to increase the level of variability in *nifHDK* expression; those cells that have high levels of *glnK* mRNA at a given time, are also likely to have higher levels of *nifHDK* mRNA. While this is clearly intuitively consistent given the known direct role of *glnK* to de-repress the inactive NifL-NifA complex (*He et al., 1998; Little et al., 2002*), the implications for a lower bound on *nifHDK* variability are significant. Given that the cells are exposed to identical conditions within the culture, it is noteworthy that variability at the level of *glnK* acts to increase variability in *nifHDK*.

Stochastic models incorporating extrinsic noise

Next, we sought to understand the minimum level of variability that might be observed if the direct regulatory role of GlnK were bypassed. This was done by measuring expression levels in a $\Delta nifLA$ strain in which *nifA* is overexpressed ectopically on a plasmid (+*nifA*, see materials and methods), thus making *nifHDK* expression independent of GlnK control.

Fig. 3(A) displays a distribution for the +*nifA* mutant, showing significant remaining sources of heterogeneity. Given that fluctuations in GlnK are no longer anticipated to have an effect on *nifHDK* expression, and that global sources of extrinsic noise were found to be minimal, *nifHDK* transcription heterogeneity must be due to intrinsic variability at the level of the *nifHDK* promoter. Such variability is widely attributed to bursty transcription, in which the promoter is intermittently active and inactive (*Jones and Elf, 2018; Golding et al., 2005*). The sources of this intermittent activity may in turn be related to transcription factor binding/unbinding (*Jones et al., 2014*), or to mechanical supercoiling effects (*Chong et al., 2014; Sevier et al., 2016*). Regardless of the mechanism, it is likely

Figure 2. Mutual information analysis based on dual probe measurements for one set of biological replicates. See supplementary figure for exemplar images. For each case the marginal distributions of the two mRNA abundances are shown, in addition to the calculated mutual information and the total entropy in *nifHDK* abundance. Mutual information is compared to a null distribution obtained by randomly shuffling the *nifHDK* data 100,000 times, thereby providing a p -value for each pair. These values for the displayed data and a further biological replicate are displayed in Tab. 1.

Figure 2-Figure supplement 1. Raw images demonstrating the use of dual FISH probes.

Figure 3. Stochastic modelling of bursty transcription incorporating extrinsic noise. (A) mRNA copy number distributions for each of the +nifA, WT and ΔglnK strains. (B) Schematic of bursty transcription and its relation to the stochastic model. Extrinsic noise here is catered for by considering the burst frequency to itself be variable between cells. (C) Variation of the model parameters between mutants, plotted as a function of the mean expression level. Error bars denote 95% Bayesian credible intervals, obtained from the MCMC chains (see figure supplement.)

Figure 3-Figure supplement 1. Exemplar MCMC chain and posterior distributions from parameter inference.

that similar levels of intrinsic variability may also be relevant when other genetic variants are tested.

Based on the observations that both intrinsic and extrinsic sources of noise may be generally relevant, we sought to develop stochastic models for transcription that could incorporate both these effects, following the approach outlined in (Ham et al., 2019). When considering *nifHDK* transcription to be intrinsically bursty, as shown schematically in Fig. 3(B), this can be modelled by the Telegraph model (Peccoud and Ycart, 1995; Iyer-Biswas et al., 2009), where the promoter undergoes rapid activation and deactivation at rates λ and ν respectively. When active, transcription occurs at a rate K while the mRNA is degraded according to a first order degradation process with rate δ . It can be shown that under particular parameter ranges this process results in a negative binomial distribution for the transcript copy number (Ham et al., 2019; Paulsson and Ehrenberg, 2000), parametrized by the normalised burst frequency λ/δ and mean burst size K/ν . In this context we take the view that extrinsic noise arising from *glnK* variability acts as a variation between cells in the burst frequency. This leads to a model that is additionally parametrized by the normalised frequency variation σ/δ . Further details are given in the materials and methods.

For each of the three exemplar distributions shown in Fig. 3(A), a comparison is given with the model following the parameter fitting process. It is evident that the model can provide a good

Figure 4. (A) Calculated contributions of extrinsic noise to the variance in *nifHDK* transcript abundance. Contributions are calculated for thousands of parameters sampled from the MCMC chains, thereby providing the most probable value and 68% credible intervals. (B) Schematic showing propagation of noise through the signalling cascade. Only a contribution from *glnK* is supported directly by the experimental data, although contributions from *nifLA* and from *ntrC* cannot be excluded. Global noise sources that are found to be very small.

fit to the data in each case, thereby enabling us to draw meaning from the model parameters, inferred for a number of genetic variants and displayed in Fig. 3(C). We observed a clear relationship between the mean expression level and the burst frequency, which is contrasted by the very limited correlation between mean expression and burst size. Cells with very low expression levels such as the $\Delta glnK$ strain have a very low burst frequency, yet their burst size is within a factor of two of that for the most highly expressing $+nifA$ strain. For the lower expressing strains, the level of inferred variability in the burst frequency seems proportional to the average. The same is not true for the $+nifA$ strain in which extrinsic noise is very low.

Based on these model fits, we further calculated the contributions of intrinsic and extrinsic noise to the total variance in *nifHDK* expression levels (Sherman et al., 2015; Hilfinger and Paulsson, 2011) (see methods), as shown in Fig. 4(A). For the $\Delta glnK$ strains, the results demonstrate that there remains a significant contribution from extrinsic noise of around 30%, indicating that some variability in burst frequency may arise from noise sources other than cell to cell variation on GlnK activity. The extrinsic noise contribution increases to around 70% in the WT, $\Delta amtB$ and $\Delta glnB$ mutants in which noise propagation from *glnK* is known to be present, but is almost zero in the $+nifA$ mutant in which the regulatory pathway via GlnK is uncoupled.

Discussion

Bacteria can be subject to any of a multitude of stresses and respond in a multitude of ways. While the average behaviour of a clonal population is often of interest, the degree of variability between cells is also important. Such knowledge may provide both mechanistic biochemical insight, as well as an indication of typical survival strategies that leverage this variability. Here we have examined the nitrogen starvation response in a model diazotroph, determining behaviour at the population level and the expression of several genes in the regulatory cascade at single cell resolution.

From the bulk measurements, our data extends the existing understanding of the *nif* regulatory network in *K. oxytoca*. We find that a $\Delta glnB$ strain behaves in a similar manner to the WT in a number of respects, consistent with the highly homologous GlnK being able to compensate for the loss of GlnB. GlnB and GlnK were shown to have redundant functionalities in regulating nitrogen assimilation genes in *E. coli*, but this redundancy is not fully extended to *nifHDK* gene regulation in the closely related *Klebsiella pneumoniae* (Arcondéguy et al., 1999). Further, GlnB and GlnK post

translational uridylation plays an important role in regulating nitrogen assimilation genes but GlnK uridylation is not required for *nifHDK* gene regulation (He *et al.*, 1998). Our Δ *glnK* strain showing much slower *nifHDK* expression is consistent with these previous findings. Given that GlnK is largely required to dissociate NifA from the co-transcribed NifL, non-zero expression of *nifHDK* in Δ *glnK* must be indicative either of low-level spontaneous dissociation of NifA or compensatory action by GlnB. In *K. pneumoniae*, GlnB when overexpressed can functionally compensate for a deletion of *glnK* (Arcondéguy *et al.*, 1999), and because *glnB*, unlike *glnK*, is constitutively expressed, such a compensatory mechanism of *glnK* by *glnB* could explain the reduced contribution of extrinsic noise. Lastly, low-level *nifHDK* expression did not occur in the Δ *nifLA* strain. This is consistent with a scheme of control in which NifA provides the essential but leaky switch for *nifHDK* expression while GlnK is the regulatory modulator.

By examining Mutual Information between pairs of genes we can infer how noise is propagated through the regulatory cascade. MI is generally small compared with *nifHDK* entropy and is below the level of statistical significance between *nifLA* and *nifHDK*, suggesting that global sources of extrinsic noise are negligible. In other gene transcription systems it has been suggested that variability in sigma factor or RNA-polymerase abundance can provide a source of variability in transcription (Elowitz, 2002; Swain *et al.*, 2002; Taniguchi *et al.*, 2010), but our results suggest that this is not the case here. We cannot exclude the possibility that fluctuations do propagate from NifLA, since the differences in mRNA and protein lifetimes can act to reduce correlations between transcript and protein abundance (Taniguchi *et al.*, 2010). Fluctuations could also propagate from the master regulator NtrC, but any sources of noise above this level must be small and therefore all variability in *nifHDK* expression must arise from within, and not outside, the regulatory network (see Fig. 4(B)).

Unlike for *nifLA*, small but statistically significant MI is observed between *glnK* and *nifHDK*. This observation, together with the known positive regulatory role of GlnK, indicates that variability in *glnK* expression must propagate down to *nifHDK*. We therefore have a direct detection of extrinsic noise acting on *nifHDK* transcription.

While the MI analysis evidences the non-zero contribution of extrinsic noise to heterogeneity in *nifHDK* transcription, we further quantify this contribution through stochastic modelling. All single cell measurements for *nifHDK* expression are consistent with a model in which transcription is inherently bursty, even in the +*nifA* mutant in which the promotor should be constitutively active. This burstiness may result for example from unavoidable dissociation of the NifA EBP and acts to generate significant levels of intrinsic noise. The model also incorporates propagation of noise down the signalling cascade as variability in the frequency of the transcriptional bursts. Further to this, by examining the variation of model parameters across mutants we can deduce that the average expression level is principally determined by the burst frequency rather than the burst size. This is consistent with recent observations for the phage shock protein (Psp) membrane stress response which is also σ^{54} dependent (Engl *et al.*, 2019), and in contrast to the existing understanding for σ^{70} promoters (So *et al.*, 2011).

The model and inferred parameter variation displayed in Fig. 3(C) are consistent with a picture in which bursts are initiated by binding of the NifA EBP to its target promoter closed complexes and terminated by its unbinding. Extrinsic noise then relates to variability in the availability of the free NifA protein. In the Δ *glnK* strain most NifA proteins are bound to NifL meaning that bursts are rare, yet when they do occur, the duration of the burst is relatively unaffected. The absolute magnitude of the extrinsic noise in this case is small, since all cells are subject to similarly low NifA availability. For the opposite case of the +*nifA* mutant, high NifA availability enables frequent bursts of transcription, yet the burst size is similar to the WT and other mutants, perhaps because the average time before dissociation of NifA from the promotor is independent of its abundance within the cell. This may reflect the rather slow conversion of a closed promoter complex as bound by NifA to an open promoter complex from which NifA has dissociated (Friedman *et al.*, 2013). The level of extrinsic noise in this case is very low, since NifA availability is large enough for the system

to essentially be “saturated”: any variability in *NifA* has little impact if the burst frequency is already at a maximal value.

The modelling and quantitative analysis gives us mechanistic insight into the sources of heterogeneity, implying that the large variation in *nifHDK* expression is an inherent property of this system. What is not clear is the implication of this for evolutionary fitness. A first possibility is that the heterogeneity is either unavoidable or sufficiently benign that the extra regulatory effort that would be required to suppress it is not worthwhile (*Lestas et al., 2010; Yan et al., 2019*). In this scenario heterogeneity is no more than an interesting artefact of the regulatory system. The alternative is that heterogeneity is beneficial, perhaps in a bet-hedging sense, as has been indicated before in bacterial stress response systems (*Carey et al., 2018; Patange et al., 2018*). Actually testing for evidence of bet-hedging is generally challenging (*Simons, 2011; Grimbergen et al., 2015*), and the strength as a strategy depends both on the specifics of the stress response system and on the typical frequency of environmental changes (*Kussell and Liebler, 2005*). Since the response to nitrogen starvation incurs such a heavy metabolic cost and resource availability for enteric bacteria is highly unpredictable, bet-hedging remains plausible.

It is quite possible that the observations made here are applicable to many other stress response systems. However particular interest in the nitrogen starvation response arises from the desire to engineer higher bulk levels of nitrogen fixation (*Gasparotti et al., 2020*). In this context the significant heterogeneity observed here may impose particular limitations on industrial scale use of diazotrophs (*Delvigne et al., 2014*), as well as confound the efficient use of clonal populations of diazotrophs in the rhizosphere unless engineered to avoid variance. Our results suggest that heterogeneity can be somewhat reduced by bypassing *glnK* and expressing *nifA* heterologously, yet only to a degree. Further questions then arise as to the extent this would affect population level fitness. If the sustainable population size or geometric growth rate are affected too much, a homogeneously fixing population may perform no better in bulk. We hope therefore that the results presented here will motivate a more careful consideration of the challenges surrounding the use of diazotrophic bacteria as a source of fixed nitrogen.

Materials and methods

Bacterial Strains and growth conditions

All experiments were performed with *Klebsiella oxytoca* M5a1, obtained from (*Yu et al., 2018*). Whole gene knockout mutants, marked with a kanamycin resistance (*nptII*) gene, were derived from M5a1 using Lambda red recombineering (Datsenko Wanner). To generate a strain overexpressing *NifA*, the M5a1 *nifA* gene sequence was cloned into the pSEVA424 vector (seva.cnb.csic.es) under the control of the *P_{trc}* promoter and a synthetic ribosome binding site (BBa_B0032, Registry of Standard Biological Parts), prior to transformation into the $\Delta nifLA$ mutant background by electroporation. NH_4 run-out was used to de-repress *gln/nif* gene expression and stimulate a reproducible transition into diazotrophic growth. Briefly, *K. oxytoca* strains were cultured in Nitrogen-Free David and Mingioli (NFDM) medium (*Cannon et al., 1974*) [69 mM K_2HPO_4 , 25 mM KH_2PO_4 , 0.1 mM Na_2MoO_4 , 90 μM FeSO_4 , 0.8 mM MgSO_4 , 2 % w/v glucose] supplemented with NH_4Cl as a nitrogen source. To ensure replete cellular N status, seed cultures were supplemented with 20 mM NH_4Cl and grown to an OD_{600} of approximately 2-3. Cells were washed and resuspended in NFDM supplemented with 0.5 mM NH_4Cl to an OD_{600} of 0.1. Cultures were crimp-sealed in 70 ml glass serum bottles (Wheaton) and chilled on ice whilst sparged with N_2 gas for 45 minutes to establish a micro-aerobic atmosphere. Colorimetric O₂xyDot sensors (OxySense) fixed inside the bottles were used to verify O_2 concentration. Following injection of 1 mL pure, O_2 -free acetylene into the headspace, cultures were warmed to 25°C and shaken at 200 rpm for up to 24 hours.

Nitrogenase assay

Nitrogen fixation was assessed via the acetylene reduction assay (*Shah and Brill, 1973*): 500 µl of culture headspace was sampled via gas-tight syringe and subject to gas chromatography through a HayeSep N column (Agilent) at 90°C in N₂ carrier gas. Acetylene and ethylene were detected by flame ionisation (FID) at 300°C and ChemStation software (Agilent) was used to integrate signal peak areas. Periodically, 15 ml of oxygen-free N₂ gas was injected into sample bottles via gas-tight syringe prior to extraction of an equivalent volume of cell culture for analysis of OD₆₀₀ and smRNA-FISH. Accumulative nitrogenase activity are expressed as % acetylene consumption and ethylene production, normalised by OD₆₀₀.

RNA fluorescence in-situ hybridisation

mRNA-FISH was performed according to the protocol described by *Skinner et al. (2013)*. Briefly, bacterial cells were sampled anaerobically and collected by centrifugation. Pelleted cells were fixed in a buffer containing 3.7 % (v/v) formaldehyde prior to permeabilization in 70 % (v/v) ethanol. Hybridization and wash steps were performed in saline and sodium citrate buffer (SSC; 150 mM sodium chloride, 15 mM sodium citrate) supplemented with 40 % (v/v) formamide. All solutions were prepared using DEPC-treated water and RNase-free plasticware. DNA probes against the *nifHDK*, *nifLA* and *glnKamtB* structural operons were designed using the Stellaris® Probe Designer version 4.2; the oligo length was set at 20 nt, the minimal spacing length at 2 nt and the masking level at 1-2. The probes were purchased pre-labelled with 6-carboxytetramethylrhodamine, succinimidyl ester [6-TAMRA] (*nifLA*, *glnKamtB*) or Cy5 equivalent Quasar® 670 (*nifHDK*) from LGC Biosearch Technology. Hybridization was performed overnight at 30°C at a final concentration of 1 µM, in buffer containing 2 mM ribonucleoside-vanadyl complex (VRC), 1 mgml⁻¹ *E. coli* tRNA and 10 % (w/v) dextran sulphate. Following multiple wash steps, chromosomal DNA was stained with 10 µgml⁻¹ DAPI (4',6-diamidino-2-phenylindole) for 30 minutes before cells were immobilised using 1 % (w/v) agarose pads on 35 mm high µ-Dishes (ibidi) for imaging.

Microscopy and image analysis

Cells were harvested for carrying out RNA FISH and microscopy analysis as described in *Skinner et al. (2013)*, with slight modifications e.g. 35 mm IBIDI discs were used for imaging purpose with the help of a WF1 Zeiss Axio observer inverted microscope. Multiple fields of view were acquired for each sample, and within each field of view five z-slices were captured for further processing. Data stacks were converted to TIFF format using ImageJ, and cell segmentation masks from brightfield or DAPI images were generated using Schnitzcells. The protocol outlined in *Skinner et al. (2013)* was then followed to detect and quantify mRNA in each cell using the Spatzcells package in MATLAB. Output from this software consisted of an estimated number of mRNA in each imaged cell.

We harvested cells from different bacterial strains when they exhibited the same cumulative levels of nitrogenase activity corresponding to 5 % acetylene reduction. Time points selected to achieve this level of nitrogenase activity across different M5a1 bacterial strains were 14.5 h for *ΔglnB* and M5a1 (WT) strain, 19.5 h for *ΔglnK* strain and 14.5 h for *ΔamtB* strain.

Mutual information analysis

In order to assess the relationship between transcript abundance of two different genes, we evaluated the mutual information (MI). MI provides a method for evaluating statistical dependencies between two random variables from the joint and marginal probability densities, even when the underlying relationship is complex and nonlinear.

As with any statistic intended for classification, it is important to ascertain a confidence threshold in order to rule out false positives. This is achieved by evaluating a null distribution for the value of the test statistic in the case that there is no relationship. Here we achieve this by shuffling the data for one gene and evaluating the MI obtained in this case. By performing this shuffling and

evaluation 100,000 times, a null distribution for the MI is obtained. From this null distribution a significance level can be obtained for the measured MI.

Stochastic modelling and parameter inference

The stochastic model for transcription is based upon the Telegraph model in which a given gene transitions from inactive to active at rate λ and from active to inactive at rate ν . When the promotor is active, transcription occurs at rate K , while degradation of the mRNA occurs at rate δ independent of the promotor activity. If $\nu \gg \lambda$ and $K \gg \delta$, the distribution of transcript abundances is the negative binomial (Ham et al., 2019),

$$p\left(n; \frac{\lambda}{\delta}, \frac{K}{\nu}\right) = \text{NegBinom}\left(n; \frac{\lambda}{\delta}, \frac{K}{\nu}\right) \quad (1)$$

$$= \text{NegBinom}(n; r, p) = \binom{n+r-1}{n} (1-p)^r p^n. \quad (2)$$

We additionally take extrinsic noise arising from variation in g/nK (and potentially other factors) to act as a variable activation rate. This is incorporated by taking the parameter λ to itself vary between cells according to a log-normal distribution with mean λ and standard deviation σ . This leads to the compound distribution,

$$q\left(n; \frac{\lambda}{\delta}, \frac{K}{\nu}, \frac{\sigma}{\delta}\right) = \int_0^\infty \text{NegBinom}\left(n; r, \frac{\nu}{\nu+K}\right) \text{LogNormal}\left(r; \frac{\lambda}{\delta}, \frac{\sigma}{\delta}\right) dr. \quad (3)$$

This integral can be evaluated numerically by computing the distribution for a range of values for r then performing a numerical integration across them. The model therefore gives us an expected transcript abundance distribution in terms of three parameter ratios, $\frac{\lambda}{\delta}$ is the average normalised burst frequency, $\frac{K}{\nu}$ the average burst size and $\frac{\sigma}{\delta}$ the normalised standard deviation on burst frequency. This distribution can be calculated by numerically performing the integration above.

Given this model, we fitted the parameters via a Bayesian inference approach using an MCMC sampling scheme implemented in the programming language Julia. This enabled us to obtain posterior distributions for each of the three parameter ratios, from which we obtained maximum *a posteriori* (MAP) estimates for each parameter and 95% credible intervals, as plotted in Fig. 3(C). All code relating to the modelling and parameter inference is available at <https://github.com/rdbrockston/TranscriptionModels>.

Calculating extrinsic contributions to variance

Given the model fits to each dataset, we were able to calculate the extrinsic contributions to variance following the approach in Sherman et al. (2015); Hilfinger and Paulsson (2011). If n is the copy number of mRNA drawn from the compound distribution $q(n|\lambda)$, where λ is the variable burst frequency, then the total variance of n may be decomposed as:

$$\text{Var}[n] = \text{E}[\text{Var}[n|\lambda]] + \text{Var}[\text{E}[n|\lambda]]. \quad (4)$$

The first term is the average variance of the mRNA copy number distribution, where the average is over the distribution of values for λ . This term gives the contribution of intrinsic noise as it is essentially a weighted sum of the variation that arises for a fixed λ . The second term is the variance of the mean copy number, where the variance is again evaluated over the distribution of values for λ . This quantifies the contribution of the extrinsic noise since it is a measure of the variation in the copy number directly resulting from the variation in λ . We can calculate each of these terms numerically given a set of model parameters, thereby calculating the fractional extrinsic contribution to the total variance as,

$$\mathcal{E} = \frac{\text{Var}[\text{E}[n|\lambda]]}{\text{Var}[n]} \quad (5)$$

In practise the MCMC scheme yields a joint distribution over the three parameters $\frac{\lambda}{\delta}, \frac{K}{v}, \frac{\sigma}{\delta}$. In order to accurately assess a best estimate of the extrinsic contribution as well as confidence intervals we evaluate the contribution $\mathcal{E}$ for 4000 parameter triplets sampled from the chain. From this distribution we calculate a maximum *a posteriori* estimate and 68% credible intervals.

Acknowledgments

This work was supported through BBSRC grants BB/N003608/1 and BB/L027135/1. TB was recipient of a Commonwealth Rutherford Fellowship. RDB is grateful to Michael Stumpf, Lucy Ham and John Pinney for many helpful discussions.

References

- Single-Cell Microbiology. Nat Biotechnol. 2016; 34(11):1077–1077. doi: 10.1038/nbt.3728.
- Arcondéguy T, Heeswijk WC, Merrick M. Studies on the Roles of GlnK and GlnB in Regulating *Klebsiella Pneumoniae* NifL-Dependent Nitrogen Control. FEMS Microbiol Lett. 1999; 180(2):263–270. doi: 10.1111/j.1574-6968.1999.tb08805.x.
- van Boxtel C, van Heerden JH, Nordholt N, Schmidt P, Bruggeman FJ. Taking Chances and Making Mistakes: Non-Genetic Phenotypic Heterogeneity and Its Consequences for Surviving in Dynamic Environments. J R Soc Interface. 2017; 14(132):20170141. doi: 10.1098/rsif.2017.0141.
- Cai L, Friedman N, Xie XS. Stochastic Protein Expression in Individual Cells at the Single Molecule Level. Nature. 2006; 440(7082):358–362. doi: 10.1038/nature04599.
- Cannon FC, Dixon RA, Postgate JR, Primrose SB. Chromosomal Integration of *Klebsiella* Nitrogen Fixation Genes in *Escherichia Coli*. Journal of General Microbiology. 1974; 80(1):227–239. doi: 10.1099/00221287-80-1-227.
- Carey JN, Mettert EL, Roggiani M, Myers KS, Kiley PJ, Goulian M. Regulated Stochasticity in a Bacterial Signaling Network Permits Tolerance to a Rapid Environmental Change. Cell. 2018; 173(1):196–207.e14. doi: 10.1016/j.cell.2018.02.005.
- Chong S, Chen C, Ge H, Xie XS. Mechanism of Transcriptional Bursting in Bacteria. Cell. 2014; 158(2):314–326. doi: 10.1016/j.cell.2014.05.038.
- Delvigne F, Zune Q, Lara AR, Al-Soud W, Sørensen SJ. Metabolic Variability in Bioprocessing: Implications of Microbial Phenotypic Heterogeneity. Trends in Biotechnology. 2014; 32(12):608–616. doi: 10.1016/j.tibtech.2014.10.002.
- Dilworth MJ. Acetylene Reduction by Nitrogen-Fixing Preparations from *Clostridium Pasteurianum*. Biochimica et Biophysica Acta (BBA) - General Subjects. 1966; 127(2):285–294. doi: 10.1016/0304-4165(66)90383-7.
- Dixon R, Kahn D. Genetic Regulation of Biological Nitrogen Fixation. Nat Rev Microbiol. 2004; 2(8):621–631. doi: 10.1038/nrmicro954.
- Elowitz MB. Stochastic Gene Expression in a Single Cell. Science. 2002; 297(5584):1183–1186. doi: 10.1126/science.1070919.
- Engl C. Noise in Bacterial Gene Expression. Biochim Soc Trans. 2019; 47(1):209–217. doi: 10.1042/BST20180500.
- Engl C, Jovanovic G, Brackston RD, Kotta-Loizou I, Buck M. The Route to Transcription Initiation Determines the Mode of Transcriptional Bursting in Bacteria; 2019.
- Friedman LJ, Mumm JP, Gelles J. RNA Polymerase Approaches Its Promoter without Long-Range Sliding along DNA. Proceedings of the National Academy of Sciences. 2013; 110(24):9740–9745. doi: 10.1073/pnas.1300221110.
- Gasch AP, Yu FB, Hose J, Escalante LE, Place M, Bacher R, Kanbar J, Ciobanu D, Sandor L, Grigoriev IV, Kendzierski C, Quake SR, McClean MN. Single-Cell RNA Sequencing Reveals Intrinsic and Extrinsic Regulatory Heterogeneity in Yeast Responding to Stress. PLOS Biol. 2017; 15(12):e2004050. doi: 10.1371/journal.pbio.2004050.
- Gasperotti A, Brameyer S, Fabiani F, Jung K. Phenotypic Heterogeneity of Microbial Populations under Nutrient Limitation. Current Opinion in Biotechnology. 2020; 62:160–167. doi: 10.1016/j.copbio.2019.09.016.

- 437 **Golding I**, Paulsson J, Zawilski SM, Cox EC. Real-Time Kinetics of Gene Activity in Individual Bacteria. *Cell*. 2005;
438 123(6):1025–1036. doi: [10.1016/j.cell.2005.09.031](https://doi.org/10.1016/j.cell.2005.09.031).
- 439 **Grimbergen AJ**, Siebring J, Solopova A, Kuipers OP. Microbial Bet-Hedging: The Power of Being Different. *Curr*
440 *Opin Microbiol*. 2015; 25:67–72. doi: [10.1016/j.mib.2015.04.008](https://doi.org/10.1016/j.mib.2015.04.008).
- 441 **Ham L**, Brackston RD, Stumpf MPH. Extrinsic Noise and Heavy-Tailed Laws in Gene Expression. *bioRxiv*. 2019;
442 <https://www.biorxiv.org/content/10.1101/623371v1>, doi: 10.1101/623371.
- 443 **He L**, Soupene E, Ninfa A, Kustu S. Physiological Role for the GlnK Protein of Enteric Bacteria: Relief of NifL
444 Inhibition under Nitrogen-Limiting Conditions. *J Bacteriol*. 1998; 180(24):6661–6667.
- 445 **Hilfinger A**, Paulsson J. Separating Intrinsic from Extrinsic Fluctuations in Dynamic Biological Systems. *Proceed-*
446 *ings of the National Academy of Sciences*. 2011; 108(29):12167–12172. doi: [10.1073/pnas.1018832108](https://doi.org/10.1073/pnas.1018832108).
- 447 **Iyer-Biswas S**, Hayot F, Jayaprakash C. Stochasticity of Gene Products from Transcriptional Pulsing. *Phys Rev E*.
448 2009; 79(3):031911. doi: [10.1103/PhysRevE.79.031911](https://doi.org/10.1103/PhysRevE.79.031911).
- 449 **Jones D**, Elf J. Bursting onto the Scene? Exploring Stochastic mRNA Production in Bacteria. *Curr Opin Microbiol*.
450 2018; 45:124–130. doi: [10.1016/j.mib.2018.04.001](https://doi.org/10.1016/j.mib.2018.04.001).
- 451 **Jones DL**, Brewster RC, Phillips R. Promoter Architecture Dictates Cell-to-Cell Variability in Gene Expression.
452 *Science*. 2014; 346(6216):1533–1536. doi: [10.1126/science.1255301](https://doi.org/10.1126/science.1255301).
- 453 **Kiviet DJ**, Nghe P, Walker N, Boulineau S, Sunderlikova V, Tans SJ. Stochasticity of Metabolism and Growth at the
454 Single-Cell Level. *Nature*. 2014; 514(7522):376–379. doi: 10.1038/nature13582.
- 455 **Ko MSH**. A Stochastic Model for Gene Induction. *J Theor Biol*. 1991; 153(2):181–194. doi: 10.1016/S0022-
456 5193(05)80421-7.
- 457 **Kussell E**, Liebler S. Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments.
458 *Science*. 2005; 309(5743):2075–2078. doi: [10.1126/science.1114383](https://doi.org/10.1126/science.1114383).
- 459 **Larson DR**, Fritzsche C, Sun L, Meng X, Lawrence DS, Singer RH. Direct Observation of Frequency Modulated
460 Transcription in Single Cells Using Light Activation. *eLife*. 2013; 2. doi: [10.7554/eLife.00750](https://doi.org/10.7554/eLife.00750).
- 461 **Larsson AJM**, Johnsson P, Hagemann-Jensen M, Hartmanis L, Faridani OR, Reinius B, Segerstolpe Å, Rivera CM,
462 Ren B, Sandberg R. Genomic Encoding of Transcriptional Burst Kinetics. *Nature*. 2019; doi: 10.1038/s41586-
463 018-0836-1.
- 464 **Lenive O**, Kirk PDW, Stumpf MPH. Inferring Extrinsic Noise from Single-Cell Gene Expression Data Using
465 Approximate Bayesian Computation. *BMC Syst Biol*. 2016; 10(1). doi: 10.1186/s12918-016-0324-x.
- 466 **Lestas I**, Vinnicombe G, Paulsson J. Fundamental Limits on the Suppression of Molecular Fluctuations. *Nature*.
467 2010; 467(7312):174–178. doi: 10.1038/nature09333.
- 468 **Little R**, Colombo V, Leech A, Dixon R. Direct Interaction of the NifL Regulatory Protein with the GlnK Signal
469 Transducer Enables the *Azotobacter Vinelandii* NifL-NifA Regulatory System to Respond to Conditions Replete
470 for Nitrogen. *J Biol Chem*. 2002; 277(18):15472–15481. doi: [10.1074/jbc.M112262200](https://doi.org/10.1074/jbc.M112262200).
- 471 **Mc Mahon SS**, Lenive O, Filippi S, Stumpf MPH. Information Processing by Simple Molecular Motifs and
472 Susceptibility to Noise. *J R Soc Interface*. 2015; 12(110):20150597. doi: [10.1098/rsif.2015.0597](https://doi.org/10.1098/rsif.2015.0597).
- 473 **Mc Mahon SS**, Sim A, Filippi S, Johnson R, Liepe J, Smith D, Stumpf MPH. Information Theory and Signal
474 Transduction Systems: From Molecular Information Processing to Network Inference. *Seminars in Cell &*
475 *Developmental Biology*. 2014; 35:98–108. doi: [10.1016/j.semcdb.2014.06.011](https://doi.org/10.1016/j.semcdb.2014.06.011).
- 476 **Munsky B**, Fox Z, Neuert G. Integrating Single-Molecule Experiments and Discrete Stochastic Mod-
477 els to Understand Heterogeneous Gene Transcription Dynamics. *Methods*. 2015; 85:12–21. doi:
478 [10.1016/j.ymeth.2015.06.009](https://doi.org/10.1016/j.ymeth.2015.06.009).
- 479 **Munsky B**, Li G, Fox ZR, Shepherd DP, Neuert G. Distribution Shapes Govern the Discovery of Predictive Models
480 for Gene Regulation. *Proc Natl Acad Sci*. 2018; 115(29):7533–7538. doi: [10.1073/pnas.1804060115](https://doi.org/10.1073/pnas.1804060115).
- 481 **Patange O**, Schwall C, Jones M, Villava C, Griffith DA, Phillips A, Locke JCW. *Escherichia Coli* Can Survive Stress by
482 Noisy Growth Modulation. *Nat Commun*. 2018; 9(1). doi: 10.1038/s41467-018-07702-z.

- 483 **Paulsson J**, Ehrenberg M. Random Signal Fluctuations Can Reduce Random Fluctuations in Regulated Com-
484 ponents of Chemical Regulatory Networks. *Phys Rev Lett*. 2000; 84(23):5447–5450. doi: [10.1103/Phys-](https://doi.org/10.1103/Phys-)
485 [RevLett.84.5447](https://doi.org/10.1103/PhysRevLett.84.5447).
- 486 **Peccoud J**, Ycart B. Markovian Modelling of Gene Product Synthesis. *Theor Popul Biol*. 1995; 48:222–234.
- 487 **Raj A**, Peskin CS, Tranchina D, Vargas DY, Tyagi S. Stochastic mRNA Synthesis in Mammalian Cells. *PLoS Biol*.
488 2006; 4(10):e309. doi: [10.1371/journal.pbio.0040309](https://doi.org/10.1371/journal.pbio.0040309).
- 489 **Raser JM**, O'Shea EK. Control of Stochasticity in Eukaryotic Gene Expression. *Science*. 2004; 304(5678):1811–
490 1814. doi: [10.1126/science.1098641](https://doi.org/10.1126/science.1098641).
- 491 **Schreiber F**, Littmann S, Lavik G, Escrig S, Meibom A, Kuypers MMM, Ackermann M. Phenotypic Heterogeneity
492 Driven by Nutrient Limitation Promotes Growth in Fluctuating Environments. *Nat Microbiol*. 2016; 1(6):16055.
493 doi: [10.1038/nmicrobiol.2016.55](https://doi.org/10.1038/nmicrobiol.2016.55).
- 494 **Schumacher J**, Behrends V, Pan Z, Brown DR, Heydenreich F, Lewis MR, Bennett MH, Razzaghi B, Komorowski M,
495 Barahona M, Stumpf MPH, Wigneshweraraj S, Bundy JG, Buck M. Nitrogen and Carbon Status Are Integrated
496 at the Transcriptional Level by the Nitrogen Regulator NtrC In Vivo. *mBio*. 2013; 4(6):e00881–13–e00881–13.
497 doi: [10.1128/mBio.00881-13](https://doi.org/10.1128/mBio.00881-13).
- 498 **Schumacher J**, Joly N, Rappas M, Zhang X, Buck M. Structures and Organisation of AAA+ Enhancer Bind-
499 ing Proteins in Transcriptional Activation. *Journal of Structural Biology*. 2006; 156(1):190–199. doi:
500 [10.1016/j.jsb.2006.01.006](https://doi.org/10.1016/j.jsb.2006.01.006).
- 501 **Sepúlveda LA**, Xu H, Zhang J, Wang M, Golding I. Measurement of Gene Regulation in Individual Cells Re-
502 veals Rapid Switching between Promoter States. *Science*. 2016; 351(6278):1218–1222. doi: [10.1126/sci-](https://doi.org/10.1126/sci-)
503 [ence.aad0635](https://doi.org/10.1126/science.aad0635).
- 504 **Sevier SA**, Kessler DA, Levine H. Mechanical Bounds to Transcriptional Noise. *Proc Natl Acad Sci*. 2016;
505 113(49):13983–13988. doi: [10.1073/pnas.1612651113](https://doi.org/10.1073/pnas.1612651113).
- 506 **Shah VK**, Brill WJ. Nitrogenase IV. Simple Method of Purification to Homogeneity of Nitrogenase Components
507 from *Azotobacter Vinelandii*. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. 1973; 305(2):445–454. doi:
508 [10.1016/0005-2728\(73\)90190-4](https://doi.org/10.1016/0005-2728(73)90190-4).
- 509 **Sherman MS**, Lorenz K, Lanier MH, Cohen BA. Cell-to-Cell Variability in the Propensity to Transcribe Explains
510 Correlated Fluctuations in Gene Expression. *Cell Systems*. 2015; 1(5):315–325. doi: [10.1016/j.cels.2015.10.011](https://doi.org/10.1016/j.cels.2015.10.011).
- 511 **Simons AM**. Modes of Response to Environmental Change and the Elusive Empirical Evidence for Bet Hedging.
512 *Proc R Soc B*. 2011; 278(1712):1601–1609. doi: [10.1098/rspb.2011.0176](https://doi.org/10.1098/rspb.2011.0176).
- 513 **Skinner SO**, Sepúlveda LA, Xu H, Golding I. Measuring mRNA Copy Number in Individual *Escherichia Coli*
514 Cells Using Single-Molecule Fluorescent in Situ Hybridization. *Nat Protoc*. 2013; 8(6):1100–1113. doi:
515 [10.1038/nprot.2013.066](https://doi.org/10.1038/nprot.2013.066).
- 516 **So Lh**, Ghosh A, Zong C, Sepúlveda LA, Segev R, Golding I. General Properties of Transcriptional Time Series in
517 *Escherichia Coli*. *Nat Genet*. 2011; 43(6):554–560. doi: [10.1038/ng.821](https://doi.org/10.1038/ng.821).
- 518 **Suter DM**, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. Mammalian Genes Are Transcribed with Widely
519 Different Bursting Kinetics. *Science*. 2011; 332(6028):472–474. doi: [10.1126/science.1198817](https://doi.org/10.1126/science.1198817).
- 520 **Swain PS**, Elowitz MB, Siggia ED. Intrinsic and Extrinsic Contributions to Stochasticity in Gene Expression. *Proc*
521 *Natl Acad Sci*. 2002; 99(20):12795–12800. doi: [10.1073/pnas.162041399](https://doi.org/10.1073/pnas.162041399).
- 522 **Taniguchi Y**, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, Xie XS. Quantifying *E. Coli* Proteome and
523 Transcriptome with Single-Molecule Sensitivity in Single Cells. . 2010; 329:8.
- 524 **Yan J**, Hilfinger A, Vinnicombe G, Paulsson J. Kinetic Uncertainty Relations for the Control of Stochastic Reaction
525 Networks. *Phys Rev Lett*. 2019; 123(10):108101. doi: [10.1103/PhysRevLett.123.108101](https://doi.org/10.1103/PhysRevLett.123.108101).
- 526 **Yu Z**, Li S, Li Y, Jiang Z, Zhou J, An Q. Complete Genome Sequence of N 2 -Fixing Model Strain *Klebsiella Sp. Nov.*
527 *M5a1*, Which Produces Plant Cell Wall-Degrading Enzymes and Siderophores. *Biotechnology Reports*. 2018;
528 17:6–9. doi: [10.1016/j.btre.2017.11.006](https://doi.org/10.1016/j.btre.2017.11.006).

Figure 1-Figure supplement 1. Acetylene reduction as a percentage of initial acetylene concentration, measured at 9.5 hours.

Figure 2-Figure supplement 1. Raw images demonstrating the use of dual FISH probes. False colour luminescence from *glnK* and *nifHDK* probes (left) and DAPI stained cells (right).

Figure 3-Figure supplement 1. The MCMC chain is displayed in the upper figures for each of λ/δ , r , σ/δ and K/v . Lower figures display the posterior distributions along with the MAP estimate and the 95% credible intervals.