

## The interplay between growth rate and nutrient quality defines gene expression capacity

by

Juhyun Kim <sup>1,†</sup>, Alexander P.S. Darlington <sup>2,†</sup>, Declan G. Bates <sup>2,\*</sup>  
and Jose I. Jimenez <sup>3,\*</sup>

<sup>1</sup> Department of Chemical Engineering, Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom

<sup>2</sup> School of Engineering, University of Warwick, Coventry, CV4 7AL, United Kingdom

<sup>3</sup> Department of Life Sciences, Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom

<sup>†</sup>These authors contributed equally to this work

\* Corresponding authors: Declan G. Bates ([d.bates@warwick.ac.uk](mailto:d.bates@warwick.ac.uk)) and Jose I. Jimenez ([j.jimenez@imperial.ac.uk](mailto:j.jimenez@imperial.ac.uk))

**Running Title:** Growth rate and nutrient capacity define expression capacity

**Keywords:** ribosomes, cellular economics, metabolic constraints, (p)ppGpp

## Abstract

The gene expression capacity of bacterial cells depends on the interplay between growth and the availability of the transcriptional and translational machinery. Growth rate is widely accepted as the global physiological parameter controlling the allocation of cell resources. This allocation has an impact on the ability of the cell to produce both host and heterologous proteins required for synthetic circuits and pathways. Understanding the relationship between growth and resources is key for the efficient design of artificial genetic constructs, however, it is obscured by the mutual dependence of growth and gene expression on each other. In this work, we investigate the individual contributions of molecular factors, growth rate and metabolism to gene expression by investigating the behaviour of bacterial cells growing in chemostats in growth-limited conditions. We develop a model of the whole cell that captures trade-offs in gene expression arising from the individual contributions of different factors, and validate it by analysing gene couplings which emerge from competition for the gene expression machinery. Our results show that while growth rate and molecular factors, such as the number of rRNA operons, set the abundance of transcriptional and translational machinery available, it is metabolism that governs the usage of those resources by tuning elongation rates. We show that synthetic gene expression capacity can be maximised by using low growth in a high-quality medium. These findings provide valuable insights into fundamental trade-offs in microbial physiology that will inform future strain and bioprocesses optimisation.

## 20 Introduction

The ultimate goal of synthetic biology is to engineer reliable and robust genetic circuits. These circuits have a range of potential applications in biomedicine, industry and environmental science<sup>1</sup>, but ensuring their reliable performance requires precise control over gene expression. To this end, it is crucial to understand the inherent constraints on gene expression created by the  
25 limited cellular economy, in order to allow optimal allocation of resources between host and circuit genes. Recently, researchers have observed that the expression of a particular gene can affect the activity of another seemingly unconnected gene due to sharing of the cell's resources. Such resource-coupled gene expression is mainly caused by limitations in the numbers of ribosomes<sup>2-8</sup>. Several approaches have been developed to reduce this coupling<sup>4,6,8,9</sup>, for instance,  
30 we successfully mitigated competition for ribosomes in co-expressed genes using quasi-orthogonal ribosomes combined with a resource allocation controller<sup>8</sup>.

The association between cellular growth and ribosome abundance has previously been investigated in order to understand how cells allocate their protein synthesis capacity. As the  
35 synthesis of ribosomes is energetically costly, cells need to maintain an optimal concentration of these macromolecules to maximise cellular fitness<sup>10,11</sup>. In bacteria, evolutionary trade-offs have shaped the number of rRNA operons to achieve optimal ribosomal biosynthesis. For instance, wild type *Escherichia coli* carry seven copies of the operons and adapt to nutritional perturbations more quickly than other isogenic strains with different copy numbers of the rRNA operons<sup>12</sup>.  
40 The cellular concentration of ribosomes is also controlled by the nutritional stress-induced alarmone (p)ppGpp<sup>13-17</sup>. As a response to amino acid starvation (the stringent response), cells accumulate (p)ppGpp which represses the expression of both rRNAs and ribosomal-protein coding genes through direct interaction with the RNA polymerase<sup>18</sup>.

45 Bacterial growth rate acts as a global regulator of ribosomal concentrations by coordinating the distribution of the cellular proteome. This allocation has been described by partitioning the *E. coli* proteome into three fractions: the R fraction (ribosomes/ribosome-associated), the E fraction (containing metabolic enzymes), and a growth independent fraction<sup>19-21</sup>. Cells allocate more of their proteome to ribosome-associated fractions under nutrient rich conditions to support maximum growth, and thus lower amounts of metabolic proteins are produced. In  
50 contrast, under poor nutrient conditions, cells synthesize more metabolic proteins to produce ATP at the expense of ribosome production, which leads to slower growth<sup>19-21</sup>. Consequently,

the synthesis of ribosome-associated proteins competes with the production of non-ribosomal proteins.

55

To investigate such cellular fitness strategies, growth rate (and hence ribosome biosynthesis) has been manipulated by altering nutrient quality<sup>22-24</sup>. However, this can potentially produce confounding results, due to the impact of different metabolic effects on the resource allocation. For example, different carbon sources could lead to distinct profiles of amino acid synthesis which in turn could affect translation beyond the availability of ribosomes. In addition, the impact of growth rate on such resource-mediated coupling has so far not been explored, and this question is crucial to ensuring that multiple genes are expressed at the right combination and concentration under dynamic resource allocation regimes.

60

65

In this work, we investigate resource allocation in the cellular proteome by using a continuous culturing system that enables a systematic investigation of different growth scenarios for the same metabolic conditions. By combining theory and experiments, we obtain novel insights into the interplay between transcription and translation and the role of rRNA operons. These insights allow us to untangle the relationship between growth and metabolism to identify the impact on gene couplings of both carbon sources and the (p)ppGpp network. These results shed new light on fundamental trade-offs in microbial physiology and will enable optimised cultivation of microbial cell factories for diverse biotechnological applications.

70

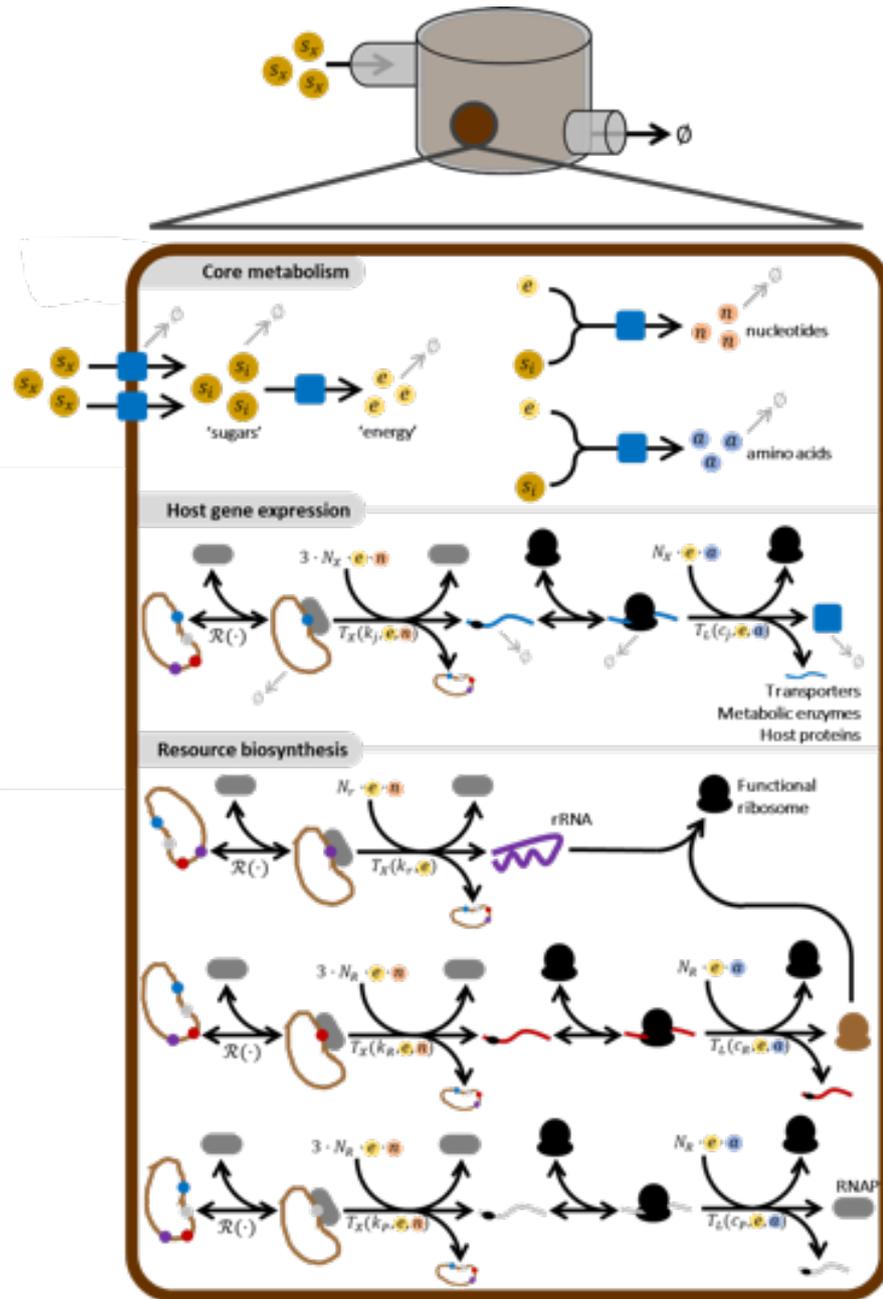
## Results

75

### **A dynamic model of metabolism, resource allocation and cellular growth**

To investigate the above questions, we developed a simplified model of the whole cell that accounts for multiple trade-offs in gene expression. By combining mechanistic models of gene expression and natural feedback with phenomenological models of growth and a simple representation of metabolism we produced a tractable model that is capable of recapitulating the different distributions of cellular proteins. This model goes beyond existing ones<sup>9,25</sup> by accounting for the following key universal constraints: (i) finite biosynthetic capacity, (ii) competition for total number of proteins, (iii) competition for RNA polymerases by mRNA and rRNA promoters, (iv) competition for ribosomes by mRNAs, (v) resource biosynthesis (both production of RNA polymerase and ribosomes, including their autocatalysis) and (vi) dynamic growth-based feedback.

85



**Figure 1. Model schematic.** The ordinary differential equation model describes a simple cell with a four-state metabolism, gene expression and resource biosynthesis. The external nutrient ( $s_x$ ) is imported by transporter proteins ( $p_T$ ) to become the internal substrate ( $s_i$ ). This is converted to an energy source ( $e$ ) by enzymes ( $p_E$ ). These enzymes also convert the internal metabolite to nucleotides ( $n$ ) and amino acids ( $a$ ) in reactions which also utilise the energy substrate. All host genes (such as transporters, enzymes and other host proteins) are expressed in a three-step process: each gene ( $j$ ) is bound by an RNA polymerase to form a transcription complex ( $k_j$ ) which undergoes transcription ( $T_X$ ), consuming energy and nucleotides in the process, to produce an mRNA ( $m_j$ ). The mRNA is then bound by a ribosome to form translation complex ( $c_j$ ) which undergoes translation ( $T_L$ ), consuming energy and amino acids in the process, to produce a protein ( $p_j$ ). The model explicitly considers resource biosynthesis. The RNA polymerase is produced like other host proteins. Ribosomes are produced in a three-step process. First, ribosomal proteins are produced in the same manner as other host proteins. Secondly, rRNAs are produced by transcription utilising RNA polymerase, energy and nucleotides in the same manner as other host RNAs are produced. The ribosomal proteins then bind the rRNA to form functional ribosomes which are capable of translation. *N.B.* All reactions are described in detailed in the Supplementary Material. Not all reactions are shown to simplify the schematic.

A schematic of the full model is shown in Fig. 1 and full mathematical details are described in the Supplementary Material. Here we summarise the main features and species within the model.

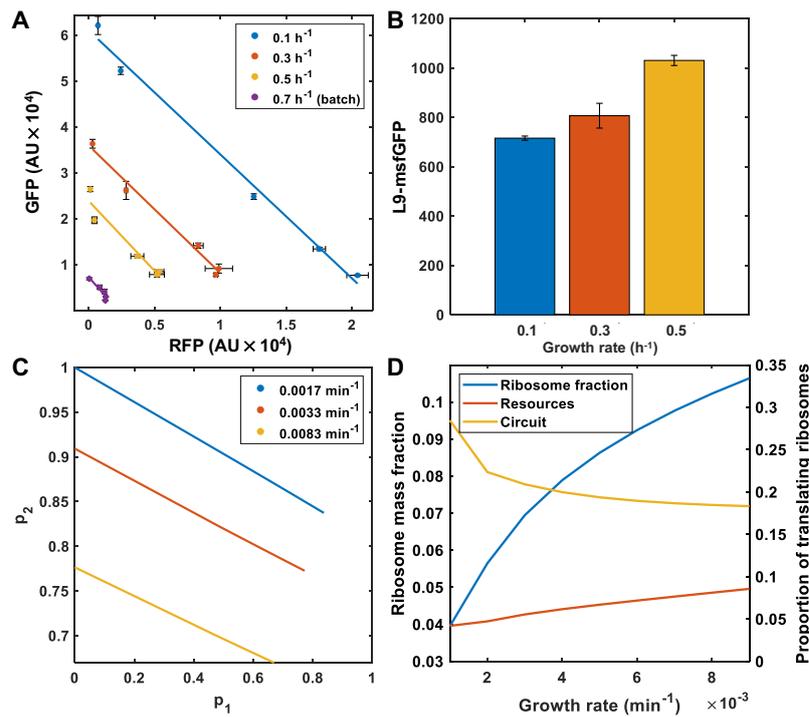
90 The model considers the growth of cells in the presence of a single external substrate ( $s_x$ ). This substrate is imported into the cell by transporter proteins ( $p_T$ ) to produce  $s_i$ , which itself is converted by enzymes ( $p_E$ ) into an expanded metabolism consisting of a proxy ‘energy’ species ( $e$ ), amino acids ( $a$ ) or nucleotides ( $n$ ). The production of  $a$  and  $n$  are dependent upon both the internal substrate  $s_i$  and energy  $e$ . Gene expression dynamics are captured by a simple four state  
95 model for each gene. Firstly, the RNA polymerase ( $P$ ) reversibly binds to free promoters to form transcription complexes ( $k_x$ , where  $x$  is the gene of interest). These are ‘consumed’ by transcription  $T_X(\cdot)$  to produce messenger RNAs ( $m_x$ ). The mRNA are reversibly bound by ribosomes ( $R$ ) to produce the translation complex ( $c_x$ ). The mRNA are released upon termination of translation at rate  $T_L(\cdot)$  which produces new proteins ( $p_x$ ). The transcription  
100 [translation] process is dependent upon both the energy species and nucleotides [amino acids]. We modeled gene expression regulation by scaling the RNA polymerase-promoter association rate as detailed in the Supplementary Material. All genes are regulated by the cell’s internal energy status. The dynamics of resource biosynthesis mirrors those of other genes. We modelled ribosome biosynthesis as a two-step process consisting of transcription of ribosomal RNAs and  
105 transcription/translation of ribosomal proteins. These bind to production functional ribosomes. The model was fit as described in the Methods to the resource economy data<sup>26</sup> and successfully recapitulates the key associations between growth and resource production (Figure S1A) as well as predicting peptide elongations (Figure S1B)

## 110 **Growth rates set gene expression trade-offs in carbon-rich conditions**

It is well established that changing nutrients within a growth media impacts both cell growth rate and ribosome content, with both being correlated<sup>27,22,28</sup>. Moreover, the interplay between gene expression and growth on a range of media has previously been established<sup>22-24</sup>. However, in these experiments the changes to the cell’s metabolism (such as changing carbon source) cause  
115 concurrent changes to growth rate, resource biosynthesis and gene expression. To dissect these relationships independently we utilized a chemostat to allow for the tight control of bacterial growth in glucose rich conditions. This allows the manipulation of bacterial growth rates and the quantification of gene expression levels under identical external nutrient conditions. To investigate resource-mediated gene coupling we utilized an *E. coli* strain that carries a plasmid  
120 encoding a constitutively expressed GFP and an inducible RFP under the control of LuxR (activated by acyl-homoserin lactone; AHL). This circuit enables a quantitative assessment of

gene coupling and resource utilization as described previously<sup>9</sup>. Cells were cultured at the following growth rates: 0.1, 0.3, and 0.5 h<sup>-1</sup>, as determined by the dilution rate of the minichemostat and in batch culture, using a range of AHL concentrations in each of those conditions. After five generations of growth, resulting in a physiologically steady state of the culture<sup>29</sup>, the production of the reporters at each given growth rate was measured by using flow cytometry.

When the chemostat was set at a 0.1 dilution rate, cells in the culture displayed a higher synthetic gene expression activity, compared to those in cultures with higher growth rates (Fig. 2A). The expression levels of GFP and RFP were three- and four-fold higher, respectively, in cells with a 0.1 h<sup>-1</sup> growth rate, compared to those in the culture set at a 0.5 h<sup>-1</sup>. Cells grown in glucose-limited batch culture showed the highest growth rates but the lowest activity of the two reporters (Fig. 2A). We measured ribosome production by labelling the ribosomal L9 subunit with



**Figure 2. Growth-rate dependence of gene expression and resource allocation.** (A) Resource mediated gene expression couplings were examined for different growth rates and AHL concentrations in cells carrying a genetic circuit that allows co-expression of GFP and RFP, which were determined by flow cytometry. Cells were also grown in batch cultures to quantify the reporter expression at the maximum growth rate. (B) Ribosomal protein production was quantified by monitoring the expression, for different growth rates, of a genomic fusion of the ribosomal protein L9 to msfGFP. (C) Simulations of the steady state concentration of the two gene circuit, normalized by maximum protein production, for different growth rates. Isocost lines were simulated by maintaining varying  $u_1 = 0$  to 1 and maintaining  $u_2 = 1$ . Simulations were carried out as described in the Methods. (D) Simulations of the cell's internal resource economy expressing two circuit genes (with induction constants of  $u_1 = 0.1$ ,  $u_2 = 1$ ) over a range of growth rates.

135 monomeric superfolder GFP (msfGFP) and found increased fluorescence with increased growth  
rate, reflecting a higher ribosomal protein synthesis (Fig. 2B). Therefore, the reduced synthesis of  
the reporters at high growth rates is not due to a lower production of ribosomes. These results  
are in agreement with previous observations<sup>23</sup> and are supported by our resource allocation  
140 model (Fig. 2C), which shows that at higher growth rates cells invest more in cellular resources,  
with RNA polymerase and ribosomes preferentially producing their own components, and a  
concurrent decrease in circuit mRNA and protein synthesis (Figs. 2D and S2).

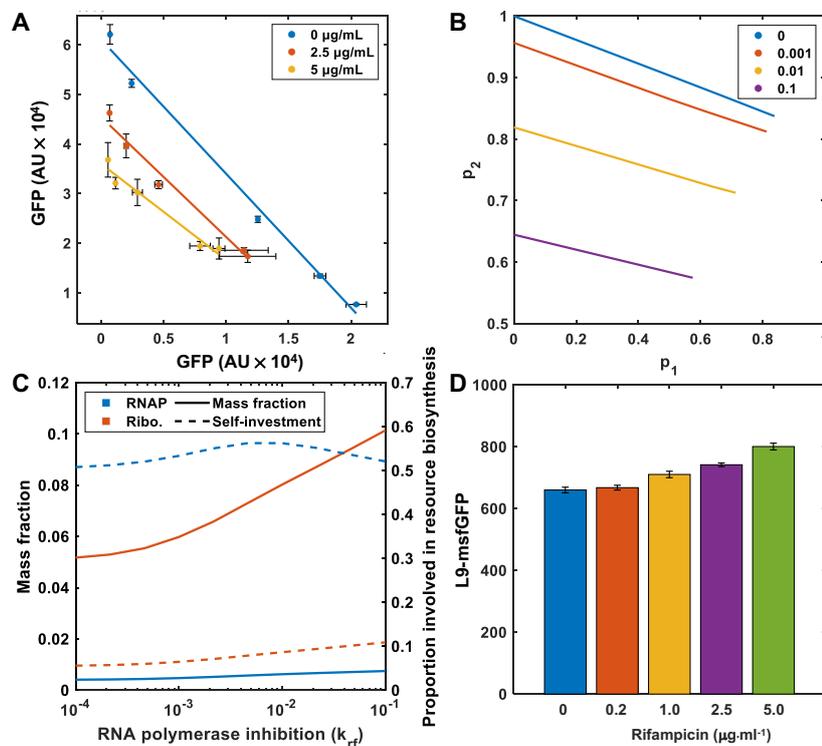
Our experimental system also enabled us to test whether the growth rate affects gene expression  
in resource-mediated competition. As shown in Fig. 2A, where the gene expression coupling  
145 occurred in continuous cultures; the expression level of GFP decreased when RFP accumulated  
in the cells regardless of the growth rate, and the concentration of the reporters is constrained by  
a linear manifold known as an ‘isocost line’<sup>5</sup>. The same level of coupling, defined as the ratio of  
the trade-off in the expression between the two genes, was maintained in all the conditions  
tested (Fig. 2A) suggesting that the cost of the protein production (as determined by the slope of  
150 the linear manifold – a measure of the resources required per protein produced<sup>5</sup>) is independent  
of the growth rate and is governed only by the abundance of ribosomes.

### **Effect of selective inhibition of transcription or translation on resource allocation**

Next, we sought to investigate the selective inhibition of components of the gene expression  
155 machinery to understand the effects of resource allocation at defined growth rates. To this end,  
sublethal concentrations of either rifampicin (Rif) or chloramphenicol (Cm) were added to the  
continuous culturing system for partial inhibition of, respectively, RNA polymerases or  
ribosomes. Again, the use of chemostats allows us to untangle the impact of resource inhibition  
whilst forcing cells to maintain a fixed growth rate (which is not possible in a batch culture  
160 system).

When cells were treated with Rif, lower intensities of both GFP and RFP were recorded, with  
expression levels that gradually decreased with increasing concentrations of the antibiotic (Fig.  
3A). We modelled the inhibition of RNA polymerase by sequestering RNA polymerase through  
165 varying  $k_{rf}$  within the model (Fig. 3B; see Supplementary Material for details). As  $k_{rf}$  is  
increased, the number of RNA polymerases and ribosomes dedicated to their own production  
increases (Fig. 3C). This results in a concurrent fall in production of other genes; i.e. the  
proportion of RNA polymerases [ribosomes] transcribing [translating] circuit genes falls (Fig.  
S3). This increase is required to maintain growth at the set rate (Fig. 3C). We confirmed this

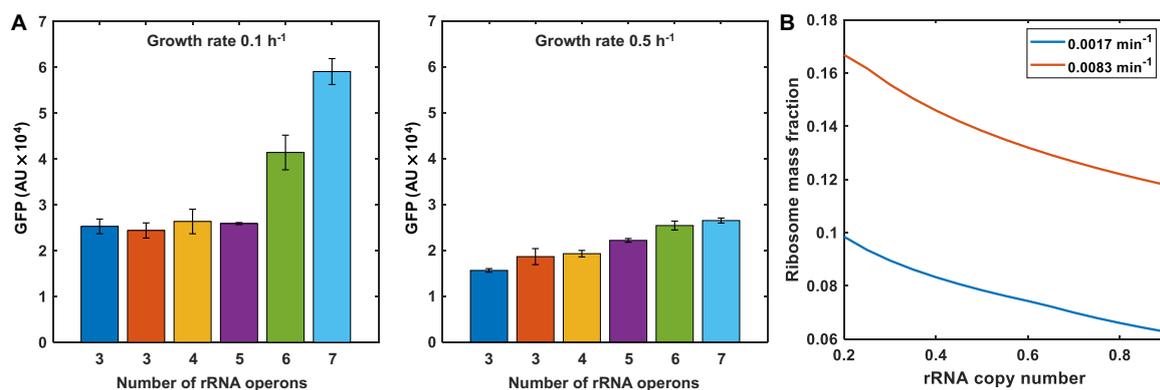
170 prediction by treating cells grown at  $0.1 \text{ h}^{-1}$  carrying the L9-msfGFP fusion with Rif. This  
 demonstrates increased synthesis of ribosomal proteins with increasing condition of Rif addition  
 (Fig. 3D). Similarly, the production of both reporters decreased when a sublethal concentration  
 of Cm was added to the culture (Fig. S4). Our model demonstrates that a simple natural  
 feedback emerging from resource autocatalysis (i.e. that resources are required to make  
 175 themselves) is sufficient to explain the fall in circuit expression which occurs during resource-  
 inhibition (Figs. S5 and S6). We also found that partial inhibition of both transcription or  
 translation at a fixed growth rate did not cause a significant change in coupling (i.e. the isocost  
 gradients are constant), rather, there was only a change in the amount of circuit protein  
 produced, indicating that reducing the transcription or translation activities does not change the  
 180 cellular resource allocation regime.



**Figure 3. Effect of the selective inhibition of transcription on resource availability.** (A) The strain carrying the GFP and RFP reporters was cultured in the minichemostat in the presence of sublethal concentrations of rifampicin and at different growth rates. (B) Simulations of the steady state concentration of the two reporter circuit normalized by maximum protein production, for different antibiotic transcriptional inhibition constant ranging from  $k_{rf}$  0 to 0.1. Isocost lines were simulated by maintaining varying  $u_1 = 0$  to 1 and maintaining  $u_2 = 1$ . Simulations were carried out as described in the Methods section. (C) Simulations of the cell's internal resource economy expressing the two-reporter circuit (with induction constants of  $u_1 = 0.1, u_2 = 1$ ) over a range of RNA polymerase inhibition constant  $k_{rf}$  varied on a log10 scale from -4 to -1. (D) The synthesis of the L9 ribosomal protein was examined in different concentrations of rifampicin at a fixed growth rate.

## The interplay between rRNA transcription, growth and gene expression

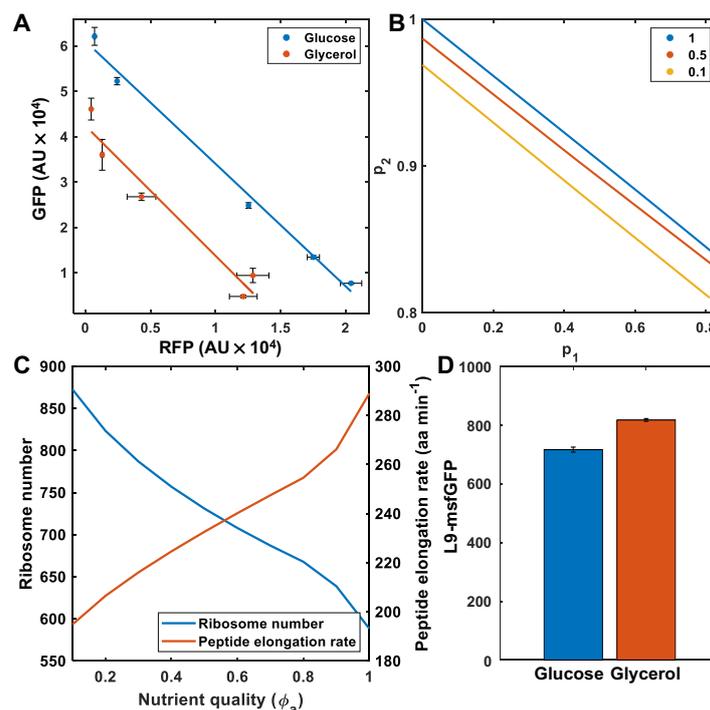
We investigated the effect of tuning the maximum size of the ribosomal pool on the protein production budget by using strains with successive deletions of the *rnm* operons<sup>30</sup>. Strains with different numbers of *rnm* operons were transformed with the resource-sensitive RFP-GFP circuit plasmid to measure the spare translation capacity and cultivated in the chemostat at either 0.1 or 0.5 h<sup>-1</sup> growth rates, representing slow and fast growth conditions, respectively. When the production of GFP was analysed, a higher accumulation of the protein was observed in the cells growing with the low dilution rate compared to the high dilution rate, regardless of their *rnm* copy number (Fig. 4A) – demonstrating again that growth and circuit gene expression are inversely correlated. The analysis also showed that the expression levels of GFP in fast-growing cells were weakly correlated with rRNA gene dosage. In contrast, the synthesis of GFP was strongly dependent on the copy number of the *rnm* operon in slow-growing cells (Fig. 4A). The synthesis of the fluorescent protein dropped by 30% and 60% in the strains containing one and two deleted *rnm* operons, respectively, compared to the wildtype strain growing at the same growth rate of 0.1 h<sup>-1</sup>. Nonetheless, only the two initial *rnm* deletions affected gene expression significantly, whereas the impact of further deletions was negligible (Fig. 4A). Our model predicts that this fall in expression is due to an increased self-investment effect; as the number of rRNA genes falls, rRNA transcription falls and so the number of functional ribosomes falls. Therefore, to maintain the desired growth rate, each functional ribosome spends more time translating ribosomal proteins. As the number of rRNA genes is scaled down from 100% to 0% of its original literature value the ribosomal mass fraction increases (Fig. 4B). This effect is predicted to be more pronounced at higher growth rates.



**Figure 4. Translational capacity dependent upon rRNA operon copy numbers.** (A) Constitutively expressed GFP levels were measured in strains carrying different number of *rnm* operons by growing them in the continuous culture at either low growth rate (0.1 h<sup>-1</sup>; left panel) or high growth rate (0.5 h<sup>-1</sup>; right panel). The strains containing ribosomal deletions used in this experiment were (from left to right) SQ78, SQ52, SQ49, SQ40, SQ37 from <sup>30</sup> and the wildtype. (B) Simulations of the cell's ribosomal mass fraction of cells expressing the one gene circuit where the rRNA copy number ( $g_{r,0}$ ) is scaled between 0 and 1 of its wild-type value. Simulations were carried out as described in the Material and Methods section.

## Metabolism plays a key role in the allocation of cellular resources

205 Previous work has investigated nutrient limitations as one of the key factors controlling  
 ribosome synthesis<sup>31,32</sup>, although in these batch growth studies the contribution of nutrients is  
 not isolated from the effect of changes in growth rate. To dissect the distinct contributions of  
 growth rate and metabolism on gene expression, we grew the RFP-GFP reporter carrying strain  
 at a set growth rate with different nutrient conditions. Samples were taken and circuit protein  
 210 expression assessed by flow cytometry. At constant growth rate in glucose cultures, we found  
 that increasing the nutrient richness by supplementing with casamino acids did not affect the  
 expression of circuit genes (Fig. S7). However, reducing nutrient quality by cultivating cells in  
 glycerol, instead of glucose, led to a 30% reduction in the reporter activity, compared to that in  
 the cells using glucose as sole carbon source (Fig. 5A). To identify the mechanistic cause of this  
 215 fall in circuit expression we modelled the impact of changing nutrient quality by varying  $\phi_e$   
 (which governs the  $s_i$  to  $e$  conversion rate),  $\phi_a$  (which governs the  $s_i$  and  $e$  to  $a$  conversion  
 rate) and  $\phi_n$  (which governs the  $s_i$  and  $e$  to  $n$  conversion rate) both independently and in



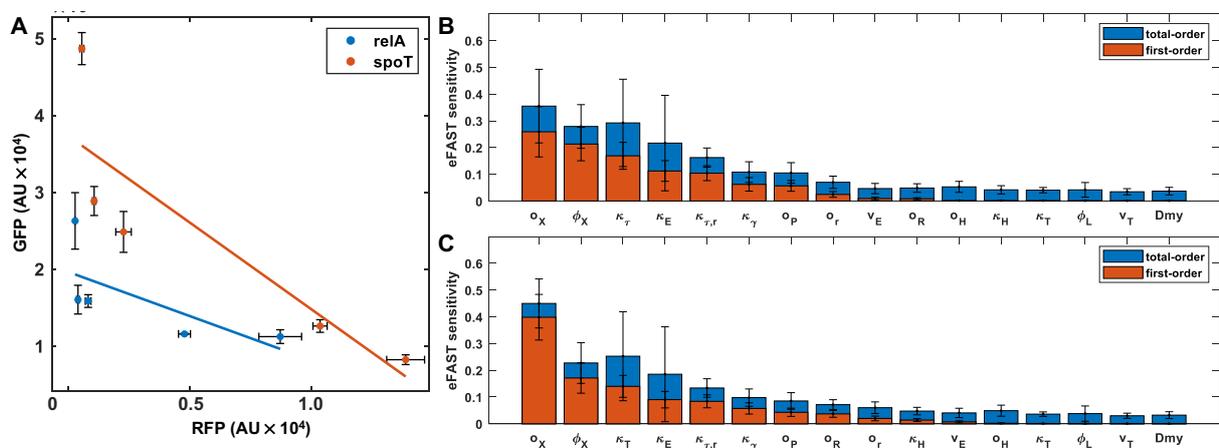
**Figure 5. Nutrient quality effect on the allocation of cellular resources.** (A) The expression capacity of the circuit genes was analysed by comparing the data of the fluorescent intensity obtained from the dual fluorescent reporter strain grown either in glucose (Figure 1A) or in glycerol as the carbon source. (B) Simulations of the steady state concentration of the two-reporter circuit, normalized by maximum protein production, for different nutrient quality. Different nutrient qualities were simulated by scaling  $\phi_a$  from 0.1 to 1 of its nominal value. Isocost lines were simulated by maintaining varying  $u_1 = 0$  to 1 and maintaining  $u_2 = 1$ . Simulations were carried out as described in the Material and Methods section. (C) Simulations of the cell's internal resource economy and peptide elongation rate of a cell expressing the two gene circuit (with induction constants of  $u_1 = 0.1, u_2 = 1$ ). (D) The expression level of the L9 ribosomal protein was measured in the same culture condition as described in (A).

combination. We found that varying  $\phi_a$  alone is sufficient to recapitulate the experimental result, suggesting that the impact of reducing the carbon source from glucose to glycerol reduces the availability of translational substrates, such as amino acids (Fig. 5B). Assessing the model's internal cellular economy and elongation rates showed that when  $\phi_a$  falls there is a concurrent fall in elongation rate, which is both energy and amino acid dependent (Fig. 5C). To maintain growth rate, the mass fraction of ribosomes rises due to increased resource investment in resource production (Fig. 5C, Fig. S8E and F). To corroborate this prediction, we measured the production of the L9 ribosomal protein and, as expected, glycerol-grown cells exhibited higher levels of fluorescent intensity relative to those in glucose-grown cells (Fig. 5D). This result suggests that the lower expression of circuit genes observed in the cells grown in glycerol is due to a subtle interaction between elongation rate and number of ribosomes, and that growth rate does not set resource levels *per se*. Rather, growth rate sets the global translation rate, which is the product of the elongation rate and the number of translating ribosomes (Eq. 25 in the Supplementary Materials). As the elongation rate falls (e.g. due to the poorer quality substrate, the number of translating ribosomes must rise for growth to be maintained.

### **(p)ppGpp affects the global distribution of the proteome**

Our findings show that the allocation of ribosomes to expendable proteins is inversely correlated with the growth rate, in agreement with the growth laws reported previously<sup>19-21,23,24</sup>. A salient question is what are the effects of physiological mechanisms that are known to play a role in the allocation of cellular resources on this trade-off. Ribosome synthesis is mainly controlled by the alarmone, (p)ppGpp<sup>15,17,18</sup>. (p)ppGpp accumulates at high levels when nutrients are limited, which results in the repression of both ribosomal proteins and rRNAs, and a decrease in growth rates<sup>15,17,18</sup>. In *E. coli*, (p)ppGpp is synthesised by RelA and degraded by the hydrolase SpoT<sup>14,18</sup>. We studied the role of the alarmone on resource allocation for set growth rates, monitoring the activity of the RFP-GFP circuit in either *relA* or *spoT* deleted strains, exhibiting low or high (p)ppGpp levels, respectively<sup>33,34</sup>. When the mutant strains, carrying the reporter genes, were cultured in the chemostat at a growth rate of 0.1 h<sup>-1</sup>, a two-fold decrease in fluorescent intensity was recorded for both GFP and RFP in the *relA* deleted strain compared to its counterpart *spoT* mutant strain (Fig. 6). This result suggests that the accumulation of the alarmone due to the mutation in the degradation activity encoded by *spoT* results in more resources being available for the circuit genes, due to a lower investment in ribosome synthesis. Interestingly, the RelA deficient strain showed a reduced resource-mediated coupling, which is associated with changes in the ratio between mRNAs and ribosome abundance (Fig. 6)<sup>5</sup>. This indicates that the alarmone

is not only involved in controlling ribosome concentration, but also in global transcriptional suppression<sup>17,35</sup> and thus the mutant strain exhibits a different cost of protein synthesis<sup>34</sup>. To identify the cellular processes that could have the largest effect on the gradient of the isocost line, we carried out a global sensitivity analysis of key metabolic parameters within the model<sup>36</sup>. We found that both the gradient and y-intercept of the isocost line are most sensitive to the  $\phi_X$  parameter (Figure 6B and C, respectively). This parameter represents the transcriptional energy threshold of the enzymes – i.e. the level of the internal energy species at which the transporter and enzyme production regulation is half maximal. We varied this parameter by +/-50% to assess the impact on the isocost line and found that decreasing  $\phi_X$  (i.e. making the enzymes more sensitive to the cell's internal energy supply) results in a decrease in the circuit output (Figure S9). We also find the isocost line is also sensitive to the energy threshold of the RNA polymerase ( $\phi_P$ ). Our global sensitivity analysis also identifies the elongation rate thresholds ( $\kappa_X$ ,  $\kappa_T$  and  $\kappa_Y$ ) as key determinants of the isocost line features. This suggests that the effect of the *relA* and *spoT* mutations on the isocost line is through a global reallocation of transcription and potential changes in the elongation process.



**Figure 6. (p)ppGpp mediated allocation of resources for gene expression. (A)** Resource mediated gene expression coupling was examined in mutants of the (p)ppGpp alarmone network. A *relA* mutant produces lower levels of (p)ppGpp than the wildtype strain while the *spoT* mutant accumulates more (p)ppGpp. Strains were cultured at a fixed growth rate in the minichemostat. **(B)** The global sensitivity of the isocost line gradient to each parameter was calculated using the eFAST method<sup>36</sup> using a parameter perturbation of up to 10%. **(C)** The global sensitivity of the isocost line y-intercept to each parameter was calculated as described in the main text.

## Discussion

In this study, we investigated how cellular resources are allocated to circuit gene expression under different cellular economic regimes. In artificial systems where ribosomal synthesis is not coupled to growth – e.g. cell-free extracts or cells endowed with orthogonal ribosomes –

translational capacity for circuit genes is directly proportional to the size of the ribosomal pool. However, in natural systems, bacterial ribosomes preferentially synthesize more ribosomes to obtain rapid growth. This is at the expense of the production of other proteins, including those of synthetic circuits or pathways. This strategy is defined by bacterial growth laws which address the partitioning of the cellular proteome, as first proposed by Hwa and colleagues<sup>19-21</sup>. In their original work, they cultivated *E. coli* cells in media containing varied carbon sources, resulting in different growth rates to obtain insights into the relation between the growth rate and the proteome distribution.

In this work we analysed the specific contributions of metabolism and growth to the allocation of resources for gene expression. To this end, we exploited a continuous culture system that allowed us to manipulate the growth rate while maintaining the carbon source, to enable the investigation of *both* growth-dependent and growth-independent effects on resource allocation. Under growth-limited (i.e. growth controlled) conditions, we found lower expression levels of circuit genes in fast-growing cells relative to those in slow-growing cells, which is in line with previous observations from carbon-limited conditions. Our model analysis showed that this is the result of a greater allocation of cellular resources to the synthesis of gene expression machineries, which is required for cells to achieve maximum growth. This strategy enables cells to maximise reproductive fitness. It is noteworthy that the cost of protein synthesis (i.e. the resources required per protein produced) was barely changed when cells redistributed resources, and thereby the level of resource mediated gene expression coupling was maintained.

This resource allocation strategy allows microbial cells to maintain cellular fitness in harsh conditions. For example, when transcription was partially inhibited with rifampicin, ribosomal protein synthesis increased, while the expression level of the circuit genes decreased. The model showed that this is the result of the self-investment of the gene expression machinery in its own synthesis to support the set growth rate. Similarly, we observed lower synthesis of the reporters in the presence of chloramphenicol-mediated translational inhibition. It is worth noting that although phenomenologically the cells behave in the same way, the specific regulatory mechanisms of resource (re-)allocation may vary depending on the point of inhibition. When treated with rifampicin, cells produce more ribosomes at the expense of the synthesis of other cellular RNAs, including circuit mRNAs, which leads to a reduced allocation of the proteome to heterologous genes. On the other hand, the inhibition of translation results in a reduced number of active ribosomes as evidenced in previous reports conducting polysome profiling<sup>22,37</sup>. Those

studies show that when the cells are cultured in the presence of Cm, they accumulate 70S monosomes, which are not significantly involved in protein synthesis.

310 Our results show that this relationship between growth rate and resource allocation is also affected by other factors. We found that as the *rrn* operon copy number increased, circuit gene expression increases in both slow and fast growth cells, but this effect is more pronounced in slow growth (where cells are investing less in resource production). The alarmone (p)ppGpp also plays an important role by setting the number of resources available for the synthesis of circuit  
315 proteins. We determined that when cultured at the same growth rate, a *relA* deficient strain, containing low (p)ppGpp levels, synthesized fewer heterologous proteins compared to a *spoT* mutant. Impairing RelA activity leads to reduced (p)ppGpp production, potentially raising ribosome synthesis while impairing SpoT, leading to (p)ppGpp accumulation and inhibition of ribosome synthesis<sup>14,15,18</sup>. As ribosome synthesis falls, circuit gene expression rises, thus  
320 explaining the different behaviour of these mutants.

We dissected the different contributions of growth rate and carbon source to the allocation of the cellular budget for gene expression. We determined that gene expression is not exclusively tuned by the number of ribosomes, but also depends on the peptide elongation rate. We found  
325 that increasing the quality of an already rich medium through amino acid supplementation did not change the translational capacity of the cell to produce the reporter proteins. On the other hand, when the cells metabolized glycerol, a poor-quality substrate, a lower expression of the circuit genes was observed compared to that in glucose (a high-quality carbon source *for E. coli*). These results are not due to a reduced availability of ribosomes in glycerol, and on the contrary,  
330 ribosomal protein synthesis was increased in glycerol. Instead, our model suggests that this is the result of a decrease in the translation elongation rates in the poor nutrient condition. In poor quality nutrients, such as glycerol, cells produce fewer translational building blocks – such as tRNA, elongation factor Tu, and guanosine triphosphate (GTP)<sup>22</sup>. Our model predicts that to compensate for the lower elongation rate and to maintain the externally set growth rate, cells  
335 synthesize more ribosomes in order to maintain their total translational capacity. This results in a reduction in spare capacity and so there is a reduction in the synthesis of non-ribosomal proteins.

The impact of host circuit interactions due to finite cellular resources has been blamed on the  
340 failure of synthetic gene circuits since the field's inception. Recently these interactions (notably in

terms of ribosome-mediated coupling between co-expressed genes) have begun to be rigorously quantified. However, our results show that in addition to ribosomal competition, metabolism and (p)ppGpp are key drivers in setting the cell's synthetic budget. These results have significant implications for the design of synthetic circuits and biotechnological pathways and imply that recombinant protein production can be increased by simple strategies such as identifying the carbon sources enabling the best elongation rates and artificially reducing the growth rates. The conceptual framework and results described here represent a step towards a complete understanding of the impact of this hitherto neglected layer of host-circuit interactions.

345

## 350 **Materials and methods**

**Minichemostat set up.** The custom-made mini continuous-culture system consisted of 15 reactors was set following the protocol<sup>29</sup> with some modifications. We used 80 ml glass bottles (VWR, Radnor, PA, USA) for the 50 ml cultures. We used silicone tubing (ID, 4 mm; OD 6 mm), 90 degree plastic tubes (ID, 4 mm; OD 6 mm), Norprene pump tubes (ID, 1.6 mm; OD, 5 mm, Sigma-Aldrich, St. Louis, MO, USA), plastic connectors, luer fittings, and syringe needles (14G and 19G) for the inflow and outflow of liquid cultures through a multi-channel peristaltic pump (Cole-Parmer, St. Neots, UK). A magnetic bar was used in each reactor for continuous culture agitation with a microplate stirrer (2mag, Munich, Germany). The pump flow rate was set to 80, 230 and 400  $\mu\text{l}\cdot\text{ml}^{-1}$  to generate growth rates of, respectively, 0.1, 0.3 and 0.5  $\text{h}^{-1}$ .

355

360

**Cultivation of bacterial cells.** Experiments in this study were conducted with *E. coli* MG1655<sup>38</sup> and derivatives (Table S1). Strains were grown in batch or continuous cultures using M9 minimal medium (6  $\text{g}\cdot\text{l}^{-1}$   $\text{Na}_2\text{HPO}_4$ , 3  $\text{g}\cdot\text{l}^{-1}$   $\text{KH}_2\text{PO}_4$ , 1.4  $\text{g}\cdot\text{l}^{-1}$   $(\text{NH}_4)_2\text{SO}_4$ , 0.5  $\text{g}\cdot\text{l}^{-1}$   $\text{NaCl}$ , 0.2  $\text{g}\cdot\text{l}^{-1}$   $\text{MgSO}_4\cdot 7\text{H}_2\text{O}$ ) supplemented with glucose (0.2%) and casamino acids (0.2%) at 30°C, unless otherwise indicated. The antibiotics kanamycin (50  $\mu\text{g}\cdot\text{ml}^{-1}$ ), ampicillin (150  $\mu\text{g}\cdot\text{ml}^{-1}$ ) and gentamicin (20  $\mu\text{g}\cdot\text{ml}^{-1}$ ) were added when necessary.

365

Chemostat cultures carrying the pSEVA63-Dual plasmid were initiated from overnight cultures in batch. Right after inoculation cultures were induced with N-acyl homoserine lactone (AHL, Sigma-Aldrich, St. Louis, MO, USA, final concentrations of 0.5, 1, 5, 10 nM) and/ or sublethal concentration of antibiotics such as rifampicin and chloramphenicol were added. After five generations of growth, when the population of cells in the chemostat reached a steady state, samples were collected and to determine fluorescent reporter expression levels (see below). *E.*

370

375 *coli* MG1655 L9-*msf*/GFP harbouring the L9 ribosomal protein labelled with GFP was cultured in the minichemostat without inducers.

**Cloning procedures and construction of reporter strains.** The characteristics of the bacteria, plasmids and primers used in this study are described in Table S1. DNA manipulation was carried out following standard protocol<sup>39</sup>. Plasmid DNA was isolated from bacterial cells using a  
380 commercial QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany). Restriction endonucleases were purchased from New England Biolabs (NEB, Ipswich, MA, USA).

The *E. coli* MG1655-derived strains containing the *rplI(L9)-msfGFP* fusion was constructed with a seamless allelic replacement method described previously<sup>40,41</sup>. The delivery plasmid pEMG<sup>41</sup> was  
385 used for the genomic exchange *rplI* x *rplI-msfGFP* and was built as follows: First, the entire *rplI* gene (TS1<sup>*rplI*</sup>, ~0.5 kb) and downstream (TS2<sup>*rplI*</sup>, ~0.5 kb) regions of the 3'-end of the gene were amplified with primer pairs *rplI*-TS1F/R and *rplI*-TS2F/R respectively (Table S1). Another DNA segment bearing the *msfGFP* sequence flanked by overlapping regions with those of the TS1<sup>*rplI*</sup> and TS2<sup>*rplI*</sup> fragments was prepared by amplification of plasmid pBG with primers *rplI*-  
390 *msfGFP*-F/R. The resulting TS1<sup>*rplI*</sup>, *msfGFP* and TS2<sup>*rplI*</sup> were then joined by isothermal assembly<sup>42</sup> and cloned in pEMG (Table S1) previously digested with *EcoRI* and *BamHI*, to yield pEMG-*rplI-msfGFP*, transformed into *E. coli* DH5 $\alpha$   $\lambda$ *pir*. This plasmid was transferred to *E. coli* MG1655 where it can not replicate and was co-integrated in the chromosome. Next, the pSW plasmid that expresses I-SceI endonuclease under the *Pm* promoter<sup>43</sup> was introduced by electroporation into  
395 the pEMG-*rplI-msfGFP* cointegrated. Selected clones were grown in LB medium with Ap and 3MBz (5 mM) to activate the *Pm* promoter, allowing I-SceI expression, which in turns forces a second recombination event by cleaving the co-integrate. Cells were plated on LB agar and the generation of a chromosomal *rplI-msf GFP* fusion was confirmed by testing the loss of the pEMG-*rplI-msf GFP*-encoded Km resistance gene. Km-sensitive clones were selected and  
400 chromosomal replacement further confirmed by PCR with oligonucleotides *rplI*-TS1F/TS2R (Supplementary Table S1). pSW was finally cured by serial dilutions in the absence of antibiotics, yielding the strain *E. coli* MG1655 *rplI-msf GFP*.

To inactivate the SpoT activity in the BW25113 strain, the target genomic region was replaced  
405 with a kanamycin antibiotic cassette (pKD4) using primers *spoT* KO F/-R, following a method described previously<sup>44,45</sup> and the genomic deletions were further confirmed by PCR.

**Fluorescence measurements.** Intracellular abundance of fluorescence proteins was determined with an Attune NxT Flow Cytometer (ThermoFisher, Waltham, MA, USA) analyzing GFP and RFP expression using blue (excitation 488 nm; emission 530/30 nm) and yellow (excitation 561 nm; emission 620/15 nm) lasers respectively. 100  $\mu$ l of the sample was taken from each culture and mixed with 200  $\mu$ l of PBS prior to injecting into the flow cytometer. 30,000 events of each sample were analysed in order to determine population means using the default software of the instrument.

415

**Mathematical modelling.** The full model is described in detail in the Supplementary Material. The system of ordinary differential equations was implemented in MATLAB2019b (The MathWorks Inc., Natick, MA, USA) and its behaviour simulated using the in-built stiff solver *ode15s*. All simulations were initiated with 10 molecules of each protein species, including ribosomes, and  $10^3$  molecules of energy. All simulations were run to steady state by increasing the simulation time span until the maximal absolute value of the derivative was very small ( $< 10^{-3}$ ). Isocost lines were simulated by incorporating two additional genes as described in the Supplementary Material.

**Model fitting.** The model was parameterized using MATLAB's in-built genetic algorithm function *ga*, with a large population of 320, and fitting ceasing when 100 generations pass without changing the cost function output significantly. The model was fit in the exponential growth regime and growth rate was varied by simulating the response of the model to a range of  $\phi_e$  values. The cost function was made of up three sums of squared errors relating to the growth rate  $\sigma_\lambda$  and mass fraction of RNA polymerase  $\sigma_P$  and ribosomes  $\sigma_R$ .  $\sigma_\lambda$  was defined as the sum of the squared error between the simulated growth rate  $\lambda_{sim}$  and the experimental growth rate  $\lambda_{exp}$ :

430

$$\sigma_\lambda = \sum \left( (\lambda_{sim} - \lambda_{exp})^2 \right)$$

435

Given that the mass fraction of the RNA polymerase  $\Phi_P$  and ribosomes  $\Phi_R$  are on different scales of magnitude, the sum of squared error was normalized by the sum of the respective experimental data squared:

$$\sigma_P = \frac{\sum \left( (\Phi_{P,sim} - \Phi_{P,exp})^2 \right)}{\sum (\Phi_{P,exp}^2)}$$

440

$$\sigma_R = \frac{\sum \left( (\Phi_{R,sim} - \Phi_{R,exp})^2 \right)}{\sum (\Phi_{R,exp}^2)}$$

The cost function used for data fitting was:

$$\text{cost} = \sigma_\lambda \cdot (\sigma_P + \sigma_R)$$

445 The parameters fit and their bounds are shown in the Supplementary Materials.

### **Conflict of interests**

The authors declare that they do not have any conflict of interests.

### **450 Acknowledgements**

The authors would like to thank Dr Manuel Salvador for the support developing the chemostat and the insightful discussions on the liquid handling system. The authors are indebted to Dr Helen King, Dr Mandy Fivian-Hughes, Fiona Watt and Anita Sicilia for their technical assistance. JK and JIJ acknowledge the support received from the Biology and Biotechnology  
455 Research Council (grants BB/M009769/1 and BB/T011289/1 from the ERA-Cobitech programme of the EU). APSD and DGB acknowledge the support of The Leverhulme Trust (grant RPG-2017-284).

## References

- 460 1 El Karoui, M., Hoyos-Flight, M. & Fletcher, L. Future trends in Synthetic Biology-A report. *Front. Bioeng. Biotechnol.* **7**, 175, (2019).
- 2 Borkowski, O. *et al.* Cell-free prediction of protein expression costs for growing cells. *Nat. Commun.* **9**, 1457, (2018).
- 3 Ceroni, F., Algar, R., Stan, G. B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* **12**, 415-418, (2015).
- 465 4 Ceroni, F. *et al.* Burden-driven feedback control of gene expression. *Nat. Methods* **15**, 387-393, (2018).
- 5 Gyorgy, A. *et al.* Isocost lines describe the cellular economy of genetic circuits. *Biophys. J.* **109**, 639-646, (2015).
- 470 6 Shopera, T., He, L., Oyetunde, T., Tang, Y. J. & Moon, T. S. Decoupling resource-coupled gene expression in living cells. *ACS Synth. Biol.* **6**, 1596-1604, (2017).
- 7 Carbonell-Ballester, M., Garcia-Ramallo, E., Montanez, R., Rodriguez-Caso, C. & Macia, J. Dealing with the genetic load in bacterial synthetic biology circuits: convergences with the Ohm's law. *Nucleic Acids Res.* **44**, 496-507, (2016).
- 475 8 Darlington, A. P. S., Kim, J., Jimenez, J. I. & Bates, D. G. Engineering translational resource allocation controllers: Mechanistic models, design guidelines, and potential biological implementations. *ACS Synth. Biol.* **7**, 2485-2496, (2018).
- 9 Darlington, A. P. S., Kim, J., Jimenez, J. I. & Bates, D. G. Dynamic allocation of orthogonal ribosomes facilitates uncoupling of co-expressed genes. *Nat. Commun.* **9**, 695, (2018).
- 480 10 Baracchini, E. & Bremer, H. Control of rRNA synthesis in *Escherichia coli* at increased *rrn* gene dosage. Role of guanosine tetraphosphate and ribosome feedback. *J. Biol. Chem.* **266**, 11753-11760 (1991).
- 11 Stevenson, B. S. & Schmidt, T. M. Life history implications of rRNA gene copy number in *Escherichia coli*. *Appl. Environ. Microbiol.* **70**, 6670-6677, (2004).
- 485 12 Gyorfy, Z. *et al.* Engineered ribosomal RNA operon copy-number variants of *E. coli* reveal the evolutionary trade-offs shaping rRNA operon number. *Nucleic Acids Res.* **43**, 1783-1794, (2015).
- 13 Corrigan, R. M., Bellows, L. E., Wood, A. & Grundling, A. ppGpp negatively impacts ribosome assembly affecting growth and antimicrobial tolerance in Gram-positive bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E1710-1719, (2016).
- 490 14 Lemke, J. J. *et al.* Direct regulation of *Escherichia coli* ribosomal protein promoters by the transcription factors ppGpp and DksA. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 5712-5717, (2011).
- 495 15 Magnusson, L. U., Farewell, A. & Nystrom, T. ppGpp: a global regulator in *Escherichia coli*. *Trends Microbiol.* **13**, 236-242, (2005).
- 16 Sun, D. *et al.* A metazoan ortholog of SpoT hydrolyzes ppGpp and functions in starvation responses. *Nat. Struct. Mol. Biol.* **17**, 1188-1194, (2010).
- 17 Zhu, M. & Dai, X. Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in *Escherichia coli*. *Nucleic Acids Res.* (2019).
- 500 18 Ross, W. *et al.* ppGpp Binding to a site at the RNAP-DksA interface accounts for its dramatic effects on transcription initiation during the stringent response. *Mol. Cell* **62**, 811-823, (2016).
- 19 Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099-1102, (2010).
- 505 20 Scott, M. & Hwa, T. Bacterial growth laws and their applications. *Curr. Opin. Biotechnol.* **22**, 559-565, (2011).

- 21 Scott, M., Klumpp, S., Mateescu, E. M. & Hwa, T. Emergence of robust growth laws  
510 from optimal regulation of ribosome synthesis. *Mol. Syst. Biol.* **10**, 747, (2014).
- 22 Dai, X. *et al.* Reduction of translating ribosomes enables *Escherichia coli* to maintain  
elongation rates during slow growth. *Nat. Microbiol.* **2**, 16231, (2016).
- 23 Klumpp, S., Zhang, Z. & Hwa, T. Growth rate-dependent global effects on gene  
expression in bacteria. *Cell* **139**, 1366-1375, (2009).
- 515 24 Mori, M., Schink, S., Erickson, D. W., Gerland, U. & Hwa, T. Quantifying the benefit of  
a proteome reserve in fluctuating environments. *Nat. Commun.* **8**, 1225, (2017).
- 25 Weisse, A. Y., Oyarzun, D. A., Danos, V. & Swain, P. S. Mechanistic links between  
cellular trade-offs, gene expression, and growth. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E1038-  
1047, (2015).
- 520 26 Bremer, H. & Dennis, P. P. *Modulation of chemical composition and other parameters of the cell by  
growth rate. Escherichia coli and Salmonella*, F. C. Neidhardt, ed. 2nd edn, 1553-1569 (ASM  
Press, 1996).
- 27 Li, G. W., Burkhardt, D., Gross, C. & Weissman, J. S. Quantifying absolute protein  
synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157**, 624-  
635, (2014).
- 525 28 Klumpp, S., Scott, M., Pedersen, S. & Hwa, T. Molecular crowding limits translation and  
cell growth. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 16754-16759, (2013).
- 29 Miller, A. W., Befort, C., Kerr, E. O. & Dunham, M. J. Design and use of multiplexed  
chemostat arrays. *J. Vis. Exp* **72**, e50262, (2013).
- 530 30 Quan, S., Skovgaard, O., McLaughlin, R. E., Buurman, E. T. & Squires, C. L. Markerless  
*Escherichia coli rrn* deletion strains for genetic determination of ribosomal binding sites. *G3  
(Bethesda)* **5**, 2555-2557, (2015).
- 31 Li, S. H. *et al.* *Escherichia coli* translation strategies differ across carbon, nitrogen and  
phosphorus limitation conditions. *Nat. Microbiol.* **3**, 939-947, (2018).
- 535 32 Dai, X. & Zhu, M. Coupling of ribosome synthesis and translational capacity with cell  
growth. *Trends Biochem. Sci.*, **45**, 681-692 (2020).
- 33 Bremer, H. & Dennis, P. Feedback control of ribosome function in *Escherichia coli*.  
*Biochimie* **90**, 493-499, (2008).
- 540 34 Shachrai, I., Zaslaver, A., Alon, U. & Dekel, E. Cost of unneeded proteins in *E. coli* is  
reduced after several generations in exponential growth. *Mol. Cell* **38**, 758-767, (2010).
- 35 Traxler, M. F. *et al.* The global, ppGpp-mediated stringent response to amino acid  
starvation in *Escherichia coli*. *Mol. Microbiol.* **68**, 1128-1148, (2008).
- 36 Marino, S., Hogue, I. B., Ray, C. J. & Kirschner, D. E. A methodology for performing  
global uncertainty and sensitivity analysis in systems biology. *Journal Theor. Biol.* **254**, 178-  
196, (2008).
- 545 37 Harvey, R. J. & Koch, A. L. How partially inhibitory concentrations of chloramphenicol  
affect the growth of *Escherichia coli*. *Antimicrob. Agents Chemother.* **18**, 323-337, (1980).
- 38 Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277**,  
1453-1462 (1997).
- 550 39 Sambrook, J. *Molecular cloning : a laboratory manual*. (Third edition. Cold Spring Harbor,  
N.Y. : Cold Spring Harbor Laboratory Press, 2001).
- 40 Nelson, K. E. *et al.* Complete genome sequence and comparative analysis of the  
metabolically versatile *Pseudomonas putida* KT2440. *Environ. Microbiol.* **4**, 799-808 (2002).
- 555 41 Martinez-Garcia, E. & de Lorenzo, V. Engineering multiple genomic deletions in Gram-  
negative bacteria: analysis of the multi-resistant antibiotic profile of *Pseudomonas putida*  
KT2440. *Environ. Microbiol.* **13**, 2702-2716, (2011).
- 42 Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred  
kilobases. *Nat. Methods* **6**, 343-345, (2009).

- 560 43 Wong, S. M. & Mekalanos, J. J. Genetic footprinting with mariner-based transposition in  
*Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 10191-10196 (2000).
- 44 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in  
*Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6640-6645,  
(2000).
- 565 45 Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout  
mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006:0008, (2006).