**On the quantitative inter-dependence between**

**gene regulation, protein levels and growth rate**

**Abstract**

In many microorganisms, the expression of a large fraction of the genome appears to be coordinately regulated with growth rate. However, although cellular growth rates, gene expression levels and gene regulation have been at the center of biological research for decades, recently there has been renewed interest in their quantitative interdependence. Here, we review recently published theoretical and experimental work in the field. We suggest a simplistic model, originating from the classic work of the Copenhagen school, that is useful as a quantitative benchmark on the interdependence of gene regulation, protein levels and growth rate. Based on an accumulated body of work we suggest that many of the gene expression changes observed across different environmental conditions are the result of passive redistribution of limited cellular resources, and are quantitatively reflected by the growth rate. This work emphasizes the extent to which gene expression can be modulated, even without employing specific regulation.

**Introduction**

A fundamental system-level challenge for cell physiology is the achievement of proper function in the face of fluctuating environments. It has been established for many years that in different environments cells differ in many properties, including their shape, size, growth rate and macromolecular composition 1–6, with strong interdependence between these parameters. Early on it was found that the expression of some genes is coordinated with growth rate. Classical experiments in bacteria, by researchers from what became known as the Copenhagen school, have shown that ribosome concentration increases in proportion to growth rate2. This increase is primarily the result of changes in protein production rate, with ribosome production increasing as the square of the growth rate1,7. The search for mechanisms in *E. coli* that underlie this observation yielded several candidates. Specifically, coordination between ribosome production and growth rate was attributed both to the pools of purine nucleotides 7,8, and the tRNA pools through the stringent response 9,10. Interestingly, these mechanisms were not identified in the model eukaryote *S. cerevisiae*, even though it displays the same quantitative coordination between ribosome production and growth rate 11–14, suggesting either the existence of different mechanisms in different organisms or the existence of additional mechanisms that underlie this basic physiological phenomenon across species. With time, the coordination between the expression level of other groups of genes with growth rate was observed; for example, the catabolic and anabolic genes in *E. coli*, a process mediated by cAMP 15.

In the last two decades, with the development of the ability to measure genome-wide expression levels, it was shown that coordination of expression and growth rate is not limited to ribosomal genes, but is actually much more wide-spread. In *S. cerevisiae*, it was shown that a surprisingly large fraction of the genome changes its expression levels in response to environmental conditions in a manner strongly correlated with growth rate 12,14,16–19. These studies raised many fundamental questions regarding the basic nature of gene regulation that are still open. For example, what is the degree of interconnection between gene expression and growth rate? What are the mechanisms underlying this connection? Is it mostly due to specific mechanisms, each affecting a distinct group of genes (such as the mechanisms detailed above) or is it a more global phenomenon shared across most genes in the genome?

In recent years, these questions have regained interest in the scientific community. Several studies examining the interplay between global and specific modes of regulation suggested that global factors play a major role in determining the expression levels of genes 14,20,21. In *E. coli*, this was mechanistically attributed to changes in the pool of RNA polymerase core and sigma factors 22. In *S. cerevisiae,* it was suggested that differences in histone modifications around the replication origins 23 or translation rates 12 across conditions may underlie the same phenomenon. Important advancements in understanding this process in *E. coli* were achieved by coupling measurements of fluorescent reporters to a model of expression built upon the empirical scaling of different cell parameters (such as gene dosage, transcription rate and cell size) with growth rate 21,24–26. These studies suggested that the expression of all genes change with growth rate, with different architectures of regulatory networks yielding differences in the direction and magnitude of these changes.

Despite these advancements, many gaps remain in our understanding of the connection between gene expression and growth rate. Importantly, whereas many of these studies depict gene expression as a function of growth rate, other studies suggest that the changes in expression temporally precede the changes in growth rate27. In addition, mechanistic insight and models for organisms other than *E. coli* are currently still missing. As such, a need remains for a quantitative model relating gene expression and growth rate which can provide a baseline for the behavior of endogenous genes in conditions in which they are not differentially regulated. Such a model can provide a basis on top of which different regulatory aspects can be added. In this work we review some of the basic notions in the field and, based on their formulations and concepts of “passive” regulation originating from the work of researchers in the Copenhagen school1, suggest a baseline model to quantify the relationship between gene regulation, protein abundance and growth rate. We examine this model alongside existing models and recent experimental results.

**Most proteins in *S. cerevisiae* and *E.coli* increase in concentration with increasing growth rate**

Starting with the development of 2-D gels, various high-throughput technologies have enabled examination of genome-wide protein concentrations in different conditions. Reanalysis of several such datasets, collected for either *E. coli* or *S. cerevisiae* in different growth rates, suggests that most proteins increase in concentration with increasing growth rate (Figure 1).

**Decoupling between the intrinsic strengths and final expression levels of genes**

Before moving to the full mathematical description presented below, we will give an intuitive overview of the components underlying it. Regulation of gene expression is a combination of global and gene-specific (“local”) modes of control. Genes and their associated regulatory sequences (such as promoters, 5'-UTRs etc.) have specific sequence features that drive the binding of specific regulatory factors in specific configurations, which lead to their unique expression levels. However, the resulting expression of a gene depends not only on these gene-specific control mechanisms, but also on global cellular factors, such as the availability of RNA polymerases and ribosomes and the metabolites pools within the cells such as amino acids and nucleotides. In recent years, there has been a growing appreciation of the important, at times dominant, role that global factors play in determining the expression levels of genes 14,20,21. To decouple these two modes of control we distinguish between the **intrinsic strength** of a gene and its final **expression level**.

The concept of intrinsic strength was first introduced by Maaloe1 stating that genes have some intrinsic ‘affinity for expression’. This affinity for expression is dictated both by the gene's sequence-encoded properties, which determine the identity of its regulators and its affinity to these regulators, and by the regulation state of the cell, which dictates the availability of specific regulators at a given time. For example, the genes belonging to the galactose-utilization pathway in *S. cerevisiae* are regulated by the transcription factor *GAL4*. Their intrinsic strength at any given time will be the result of both the regulation state of the cell (*e.g.*, *GAL4* concentration) and of their sequence (*e.g.*, the number of binding sites for *GAL4*, their location, affinity, etc.). Together, these determine the affinity of RNA polymerase to the promoters of these genes. Similarly, the sequence of the ribosome binding site in prokaryotes or the secondary structure of the 5'-UTR in eukaryotes, will determine the affinity of the ribosomes to these genes.

However, the final expression level of a gene depends on further modulation of the intrinsic strength by the availability of cellular resources shared among all genes. For example, the final rate of transcription initiation, which will largely contribute to the final expression level of a gene, will depend on the abundance and availability of RNA polymerases. The final elongation rate of the polymerase will depend on the nucleotide pool. To underscore the fact that many processes contribute to the final expression level of the gene (among them transcription initiation, elongation and termination, mRNA processing, export and degradation, translation initiation, elongation and termination), each with its gene-specific and shared global factors; we will deviate from the Copenhagen's school original definition of 'affinity for expression' and refer to the combined affinities of a gene to all complexes that perform these tasks as its intrinsic strength.

Gene-specific regulation acts to alter the intrinsic strengths of genes. For example, up or down regulation of a transcription factor (e.g. *GAL4*) will change the affinity for expression of all its downstream targets (e.g. galactose assimilation pathway). **A regulated gene will have the same intrinsic strength at any growth condition where its regulation state remains unchanged.** This definition also holds for the ubiquitous, though loosely defined, ‘constitutively expressed genes’. Here, we suggest that constitutively expressed genes do not have different regulation states and can therefore be more concretely defined as having the same intrinsic strength under all growth conditions. We note that differential availability of the shared cellular resources, such as RNA polymerases and ribosomes, under different growth conditions can result in different expression levels of a gene, even if its regulation state was not actively changed. As such, constitutively expressed genes, which according to our definition have a constant intrinsic strength across conditions, may actually display different final expression levels under different conditions.

Although this separation between 'specific' and 'global' determinants of expression has been proposed decades ago 1, in recent years it has received increasing attention due to new experimental measurement techniques. First, population measurements of gene expression and promoter activities across conditions have identified that many, if not most, expression changes across conditions result from global effects and not specific regulation 12,14,20,21. Independently, investigations in the field of cell-to-cell heterogeneity in gene expression (often referred to as 'gene expression noise') have identified that a major source of differential expression between genetically identical cells is extrinsic noise, which is caused by global differences in cellular environment and is shared across different genes 28. In support of the important role global factors play in expression, in several organisms (including the budding yeast *S. cerevisiae*) it was shown that most of the noise in gene expression is actually extrinsic 29–34. Together, these studies indicate that both the intrinsic strength of a gene and global cellular factors contribute to its resulting final expression level.

**What happens upon an environmental change?**

Two main strategies have been proposed to account for the coordination of gene expression with cellular growth rate under different environmental conditions 27. In the first strategy, the environment determines the growth rate, which in turn regulates gene expression (**Fig. 1A**). This hypothesis has led to mathematical models describing gene expression across conditions in which different cellular parameters (such as polymerase and ribosome content) are described as a function of growth rate 21,24,25. Other studies have proposed different mechanisms by which growth rate affects gene expression for different groups of genes; for example, through histone modifications around the replication origins 23, translation rates 12, the stringent response 9 or cAMP 15,20.

In the second strategy, the environment does not directly dictate the growth rate, but rather dictates gene expression levels, which in turn determine the growth rate (**Fig. 1B**). According to this model the cell senses its growth environment and responds by changing the expression of most genes in the genome in a manner that will match the final growth rate in this environment. This model has been supported by experiments in yeast which have examined the temporal order of changes in both gene expression and growth rate in response to an environmental change 27. However, despite the evidence that growth rate may actually follow gene expression, rather than the other way around, it remains unclear how this process is achieved and how signaling pathways would be fine-tuned to properly predict growth rate at each environment. One can also hypothesize intermediate strategies in which gene expression and growth rate continually modulate each other, in a feedback process, until reaching a new steady state.

Here we suggest the following simplistic model, extending the second strategy, which does not require fine-tuning of expression patterns in anticipation for a specific growth rate. According to this model the cell senses its growth environment and responds by changing the intrinsic strengths of a relatively limited set of regulators, which, in turn, modulate the intrinsic strengths of a limited set of target genes. The outcome of this process can be thought of as a “task-list" of genes to express (a kind of cellular look-up table), where each gene has its associated intrinsic strength. Schematically, this list takes the form of active genes, which are expressed by the transcription-translation machinery of the cell. While we realize that additional feedback between gene expression and growth rate may (and probably does) occur during the adjustment to a new condition, we will claim in the upcoming sections that the growth rate is dictated by the proteome composition. Specifically, we will discuss and derive how a change in the intrinsic strengths of a limited number of genes will affect the levels of all other genes in the genome, which will in turn determine the growth rate.

**The cell has limited, yet dynamic, resources**

How can a change in the intrinsic strength of a limited set of genes, as discussed above, affect the levels of other genes which did not have their intrinsic level of activation changed? To examine this question we consider a simple scenario of a cell growing in either of two environmental conditions: a favorable nutrient-rich environment and a harsh, nutrient-poor environment. This hypothetical cell contains two groups of genes: 1) Environmentally dependent genes. This group comprises genes which are differentially regulated between the two conditions. For example, it may include a specific metabolic pathway such as the galactose-utilization pathway in *S. cerevisiae*. For the cell to grow in the harsh environment it must actively up-regulate its environmental genes. 2) The rest of the genes. This group comprises all genes that are not differentially regulated in the two conditions. It includes genes required for growth in both conditions, such as genes involved in transcription, translation etc. (**Fig. 2**). Further discussion of the connection between harsh conditions and upregulation of environmental genes is available in the supplementary text.

The question of whether and how upregulation of the environmental genes will affect the rest of the genes depends on the exact property being measured and is thoroughly discussed in the supplementary text. Here, we focus on the fraction of the culture's proteome occupied by each group in each condition. Notably, upregulation of the environmental genes necessarily implies that the rest of the genes will occupy a smaller fraction of the culture's proteome. For example, upregulation of the environmental genes from 20% to 40% of the proteome will be accompanied by a decrease of the rest of the genes from 80% to 60%. Thus, the fraction of the not differentially regulated genes out of the overall protein mass is lower in the harsh condition than in the rich condition, and the implications of this will be discussed in the following section.

**Cells are self-replicators**

The last component in our model defines a relationship between gene expression and growth rate, and is based on the fact that cells are self-replicators. To this end we make the following assumptions: (A) The biosynthesis machinery operates at a constant rate per unit under the conditions investigated (i.e. the transcription and translation rates are invariant under different growth conditions, as was previously suggested6,35). (B) There is a set of genes that compose the biosynthesis machinery (e.g. genes of the transcriptional and translational machineries) which are not differentially regulated under the conditions investigated. These assumptions can be relaxed as is discussed in the supplementary text.

The first assumption entails that there is a tight connection between the fraction of the biosynthesis machinery out of the proteome and the time it takes a cell to double (**Fig. 3**). This can be intuitively understood for ribosomes. When a cell doubles itself, the ribosomes need to translate the entire proteome. If we assume that across conditions the translation rate is relatively constant 35, then the lower the fraction of ribosomes out of the proteome, the longer it will take them to translate it. The second assumption entails that because the genes of the biosynthesis machinery are non-differentially regulated, their fraction out of the proteome will be reduced under environmental conditions that require more expression of condition-specific genes. Altogether, these suggest that environmental conditions that require more expression of condition-specific genes, will display slower growth rates.

**Deriving a quantitative relationship between gene expression and growth rate**

In this section we incorporate the notions of intrinsic expression strengths, limited resources and self-replication detailed above, to arrive at a quantitative benchmark relationship between the expression of a non-differentially regulated gene and the growth rate. We show that due to these principles, once a cell changes the intrinsic strength of a limited group of genes, both the expression levels of all genes and the growth rate follow, with a concrete and predictive quantitative relationship between them.

Specifically, we will derive the interconnection between growth rate and A) the protein's fraction out of the proteome B) the protein's concentration out of the culture's biomass C) the protein production rate per biomass. The first is useful as it is unitless, and thus can be easily applied to any dataset irrespective of whether the actual quantities were measured per cell, as concentrations or as total amount in the culture. The other two are useful as they are independent of the cell volume, which changes across conditions according to constraints and cellular trade-offs which are beyond the scope of this work, and they can be readily correlated to the growth rate of an exponentially growing culture. Further discussion on these properties is available in the supplementary text.

**Relationship between growth rate and protein concentration**

We define the intrinsic strength of protein in condition as .

According to our model the biosynthetic resources of the cell are distributed among the different genes according to their intrinsic strengths, such that the fraction of protein , , out of the total proteome, , in condition ,denoted by , is equal to its intrinsic strength divided by the sum of intrinsic strengths of all of the genes under that condition.

In other words, is proportional to and sums to 1.

We note that a protein’s fraction out of the proteome, , can vary across conditions as a result of two mechanisms, as discussed above and before 1,14,24: (a) Active regulation by changing the intrinsic strength . (b) Passive change due to changes in the intrinsic strengths of other genes, i.e. a change in the sum over in the denominator.

To formally model the self-replicating property of cells we make the following assumptions:

1. There is a set of genes belonging to the biosynthesis machinery GB, which are not differentially regulated under the conditions investigated (e.g. genes of the transcriptional and translational machineries). Thus, their intrinsic strengths are constant for all investigated conditions:
2. The biosynthesis machinery operates at a constant rate per unit under the conditions investigated (i.e. the transcription and translation rates are invariant under different growth conditions, as was previously suggested6,35). This assumption was observed not to hold under slow growth rates and can be relaxed as discussed in the supplementary text.

Under these assumptions there exists a constant time, , that it takes a biosynthesis generating unit to replicate itself. For concreteness, one can think about the time it takes a single ribosome to translate the equivalent of its own proteins. We note that the time it takes the entire biosynthesis machinery to replicate itself is therefore also . For convenience, we denote the sum of intrinsic strengths of the biosynthesis machinery genes by:

As the intrinsic strength of these genes is constant (assumption 1), this sum includes no condition-specific terms and remains constant across different growth conditions.

When a cell doubles itself, every biosynthesis unit needs to synthesize the equivalent of its own proteins plus some portion of the rest of the proteome. As shown below, the total doubling time will be inversely proportional to the fraction of the biosynthesis machinery out of the total proteome. Based on equation 1, this fraction can be expressed as:

Therefore, the total time it takes the biosynthesis machinery to produce a desired proteome under condition c is:

is the doubling time of the proteome, and is equal to the doubling time of the biomass, assuming the proteome fraction does not vary significantly.

If the entire proteome is composed of biosynthesis machinery units, then the doubling time, is equal to . Each additional protein that does not belong to the biosynthesis machinery will increase the doubling time.

Since by definition *g*, the growth rate is then:

Substituting this term in the expression for the protein fraction of a protein out of the proteome (equation 1) we get that:

So the observed protein fraction of protein out of the proteome in condition depends linearly on the growth rate , on some condition-independent constants (and ) and its condition specific intrinsic strength .

We can group all constants to yield:

To get at the protein concentration per biomass, it is required to multiply by the ratio of the total protein to biomass, :

Thus, assuming that the ratio of protein to biomass does not change across conditions as suggested 4,6, our model suggests that for any gene whose regulation state is unaltered (i.e no change in ) the protein concentration will scale proportionally with the growth rate.

**Relationship between growth rate and protein production rate**

We now turn our attention to the connection between the protein production rate, , the protein fraction in the proteome and the growth rate .

The production rate of protein in condition is defined as:

where is the amount of protein in the culture and is the culture's biomass.

We note that in balanced growth:

Furthermore, under balanced growth and are constant36. Therefore:

Plugging this in equation 4 we get that:

That is, the production rate of protein depends both on the growth rate and on the protein’s concentration, as observed experimentally 14.

Substituting we get a more intuitive notion that protein 's concentration is proportional to the production rate multiplied by the doubling time **(Fig 4A)**:

Substituting for the expression for (equation 3) we get:

So the observed production rate for protein under condition depends linearly on the square of the growth rate, the protein/biomass ratio , some condition-independent constants () and the condition specific intrinsic strength of the relevant gene .

If we assume that the protein/biomass ratio is constant across conditions, then the model suggests that the protein production rate per biomass will scale like the square of the growth rate for any gene in conditions in which its regulation state in unaltered (**Fig 4B**).

**Model summary**

We have developed a model that offers a baseline prediction for the protein concentration and protein production rate per biomass as a function of growth rate for any gene across conditions in which it does not alter its regulation state. Our model suggests that for any gene, under conditions in which its regulation state in unaltered, protein concentration will scale like the growth rate, whereas protein production rate per biomass will scale like the growth rate squared (equations 3 and 5, **Fig. 5**). In other words, the expression of any gene should have a baseline response of scaling proportionally with growth rate, on top of which specific regulation will operate. These relations may require corrections if the protein ratio to biomass or the biosynthesis rates (e.g. transcription and translation rates) change between conditions. In addition, the model can be extended to include other organism-specific or condition-specific parameters. However, our model can serve as a basic guideline to assess whether an observed change in expression can be understood as a passive response without a need for specific change in regulation or indeed is a result of a change in regulation.

Notably, the final equations present protein concentration and protein production rate per unit biomass as a function of growth rate. This way of presentation is advantageous since growth rate is a variable that is easy to measure and compare across various laboratories and experimental setups. However, we note that in our model growth rate is not predetermined. It is an outcome of the fraction of proteins belonging to the biosynthetic machinery in that condition, which is in turn determined by the partitioning of resources between genes with different intrinsic strengths. This view of cellular physiology is complementary to many studies so far, which have generally regarded growth rate as the independent variable, and gene expression as the dependent variable 12,20–22,24,25,37.

Importantly, our model is not limited to a specific group of genes, such as ribosomes, constitutive genes or catabolic genes. It predicts the quantitative interdependence between protein concentration and growth rate for any gene across conditions in which it does not alter its regulation state. As such it can be applied also to condition-specific genes, between all conditions in which their regulatory state, and thus their intrinsic strength remains unaltered. For example the galactose genes under two conditions where they are repressed or two conditions in which they are activated. We suggest that it can serve as a baseline on top of which additional parameters, as well as specific regulation, can be incorporated.

**Experimental evidence and previous studies**

1. Ribosome fraction of proteome scales like the growth rate

One of the best established examples for a quantitative relationship between gene expression and growth rate is the case of ribosomal proteins. It was shown that ribosome concentration increases in proportion to growth rate2,19 and ribosome production increases as the square of the growth rate1,7. In *E. coli* this coordination was attributed to several mechanisms, primarily to the pools of purine nucleotides 7,8, and the tRNA pools through the stringent response 9. Interestingly, these mechanisms were not identified in the model eukaryote *S. cerevisiae*, even though it displays the same quantitative coordination between ribosome production and growth rate 12–14. Here, we suggest that the observed dependence between ribosome concentration and growth rate may be largely a passive result of limited resources without need for active regulation mechanisms, in line with models by Maaloe and colleagues 1. Additional mechanisms, such as the stringent response, may serve on top of the baseline behavior to allow for more precise regulation or faster response times. We find this model appealing due to its simplicity and parsimony, yet are aware that it is a simplification and that the range of its validity requires more experimental measurements.

1. The expression of genes not differentially regulated across two conditions should not stay constant but scale with the growth rate

Extending on both classic and recent observations and models, here we suggest that the expression of any gene (ribosomal or not) should have a baseline response of scaling proportionally with growth rate between conditions in which it is not differentially regulated. On top of this baseline response, specific regulation (for example the activation of specific transcription factors) and network architecture (for example, the existence of negative or positive feedback) will dictate the final relationship between growth rate and the expression level of a gene, as previously shown 24. This invokes the question of how prevalent will it be to observe the baseline response in actual expression data? Is the baseline response the norm across most genes, or are other forms of regulation dominating to the extent that this behavior is rarely observed?

Recent advances in technology, enabling genome-wide expression measurements, suggest that the proposed baseline behavior of expression scaling proportionally with growth rate is quite prevalent, and the expression of many genes is highly correlated with growth 3,5,10,12,16,17. In a study that examined yeast growing in 36 conditions it was found that close to 50% of the variation in expression in the entire dataset is attributed to genes scaling proportionally with growth rate 12. Another recent study showed that across conditions the promoter activity of most genes (70%) preserve proportionality and scale by a scaling factor dependent on both growth rate and magnitude of specific response, both in yeast and in *E. coli* 12. Here we show a good agreement between these reported scaling factors and the square of the culture’s growth rate, as predicted by our model (**Fig. 5b**). Altogether, it seems that for yeast there is strong evidence that the baseline behavior proposed here plays a dominant role in determining the final expression profile. It remains to be explored whether this is the case for other organisms.

**Implications**

The model presented here, together with the work discussed, have several important implications, both for our understanding of how cells function and for biological research. Primarily, it suggests that many of the previously-reported changes in expression across conditions can be understood as a passive response due to a redistribution of biosynthetic resources. Thus, a change in expression between conditions may not indicate active regulation. This suggests that global factors should be carefully taken into consideration when designing and analyzing studies that aim to understand gene expression regulation. Such a practice may lead to completely different interpretations of experimental data, as was demonstrated recently 37. Specifically for exponentially growing cells (microorganisms, cell lines etc.), it is critical to monitor the growth rate, report it, and analyze gene expression patterns accordingly.

It is intriguing to consider that phenotypes may be caused by either actively regulated genes or passively changing genes. Passively changing genes are commonly reported to be unchanging using standard normalization techniques, such as normalization to a housekeeping gene. However, a change in concentration of these genes, even if caused passively, may actually affect cellular behavior. Thus, interesting biology may lie in both actively regulated and passively responding genes. We suggest that the modelling framework presented here may serve as a benchmark for how the passive response of gene expression behaves and thus allow the decoupling between active and passive regulation.

**Figure legends**

**Figure 1: Models for coordination of gene expression and growth rate**

(A) A change in environment leads to a change in growth rate. This change affects the expression of all genes through the dependencies of the biosynthesis machinery (for example ribosomes and polymerases) on growth rate. (B) A change in environment leads to a coordinated change in gene expression. Signaling pathways must be fine-tuned to predict growth rate at each specific environment. (C) A change in environment leads to a change in expression of a limited number of genes. This affects the concentration of all other proteins due to limited cellular resources. Change in growth rate follows from the new concentrations of the biosynthesis machinery proteins. (Adapted from 27)

**Figure 2: Models for changes in protein concentrations across conditions**

Shown is a schematic of a proteome that is composed of 2 groups of proteins: environmental proteins (red) and the rest of the proteome (blue). The cell can grow in two conditions, with the second requiring strong upregulation of the environmetal genes. Upregulation of the environmental genes to a larger fraction of the proteome dictates a complimentary decrease in the fraction of all other proteins.

**Figure 3:** **The relationship between fraction of biosynthesis genes and growth rate**

A schematic example demonstrating that the fraction of biomass generating genes out of the proteome should be correlated to growth rate. A culture is growing in either condition I (A) or in condition II (B). The composition of the proteome differs in both conditions, as does the doubling time. In condition I the biosynthesis proteins (blue) comprise a quarter of the proteome, whereas in condition II they comprise half the proteome. Accordingly, the doubling time in condition I is twice longer than in condition II.

**Figure 4: Relationship between protein production rate, protein mass per cell and doubling time**

(A) Shown is the dependence of the protein mass per cell of a single protein (blue surface) on both cellular doubling time (x-axis) and production rate (y-axis). Shown are two conditions in which the same amount of protein is produced, either by high expression during a short doubling period (condition I), or by lower expression during a longer doubling period (condition II). (B) Shown is the proteome composition of a culture in either of two conditions, as in figure 2. The proteome is composed of 2 groups of proteins: environmental proteins (red) and the rest of the proteome, which includes the biosynthesis machinery (blue) . The cell can grow in two conditions, with the second requiring a twofold upregulation of the environmetal genes. This leads to a twofold decrease in the fraction of the biosynthesis genes in the proteome (e.g. from 2/3 to 1/3), and, consequently, to a twofold reduction in doubling time (equation 2). Altogether, a four-fold reduction is expected in the production rate of the biosynthesis genes in condition II (equation 5).

**Figure 5: Model summary and experimental support**

(A) Shown are over 200 measurements of the promoter activity (y-axis) of the yeast ribosomal gene *RPL3* as a function of growth rate. Growth rate was modulated by growing the cells in 6 environmental conditions (methods), and exploiting the natural variability in growth rate in these conditions between different wells in the plate. (B) Shown are the global scaling factors, by which most promoters change their activity between conditions (y-axis) and the square of the growth rate of each condition (x-axis) (data from Keren et al.).

**References**

1. Maaloe, O. in *Dev. Biol.* 33–58 (Elsevier, 1969).

2. Schaechter, M. Dependency on medium and temperature of cell size and chemical composition during balanced growth of Salmonella typhimurium. *J. Gen. …* **19,** 592–606 (1958).

3. Pedersen, S., Bloch, P. L., Reeh, S. & Neidhardt, F. C. Patterns of protein synthesis in E. coli: a catalog of the amount of 140 individual proteins at different growth rates. *Cell* **14,** 179–90 (1978).

4. Bremer, H. & Dennis, P. Modulation of chemical composition and other parameters of the cell by growth rate. *Escherichia coli Salmonella …* (1987). at <http://ctbp.ucsd.edu/qbio/beemer96.pdf>

5. Churchward, G., Bremer, H. & Young, R. Macromolecular composition of bacteria. *J. Theor. Biol.* **94,** 651–70 (1982).

6. John L. Ingraham, Ole Maaløe, F. C. N. *Growth of the bacterial cell*. 435 (. Sinauer Assoc., Sunderland, MA, 1983).

7. Gourse, R. L., Gaal, T., Bartlett, M. S., Appleman, J. A. & Ross, W. rRNA transcription and growth rate-dependent regulation of ribosome synthesis in Escherichia coli. *Annu. Rev. Microbiol.* **50,** 645–77 (1996).

8. Gaal, T., Bartlett, M. S., Ross, W., Turnbough, C. L. & Gourse, R. L. Transcription regulation by initiating NTP concentration: rRNA synthesis in bacteria. *Science* **278,** 2092–7 (1997).

9. Chatterji, D. & Ojha, A. K. Revisiting the stringent response, ppGpp and starvation signaling. *Curr. Opin. Microbiol.* **4,** 160–5 (2001).

10. Brauer, M. J. *et al.* Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast. *Mol. Biol. Cell* **19,** 352–67 (2008).

11. Zeevi, D. *et al.* Compensation for differences in gene copy number among yeast ribosomal proteins is encoded within their promoters. *Genome Res.* **21,** 2114–28 (2011).

12. Keren, L. *et al.* Promoters maintain their relative activity levels under different growth conditions. *Mol. Syst. Biol.* **9,** (2013).

13. You, C. *et al.* Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature* **advance on,** (2013).

14. Gasch, A. P. *et al.* Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* **11,** 4241–57 (2000).

15. Saldanha, A. J., Brauer, M. J. & Botstein, D. Nutritional homeostasis in batch and steady-state culture of yeast. *Mol. Biol. Cell* **15,** 4089–104 (2004).

16. Castrillo, J. I. *et al.* Growth control of the eukaryote cell: a systems biology study in yeast. *J. Biol.* **6,** 4 (2007).

17. Zaslaver, A. *et al.* Invariant distribution of promoter activities in Escherichia coli. *PLoS Comput. Biol.* **5,** e1000545 (2009).

18. Berthoumieux, S. *et al.* Shared control of gene expression in bacteria by transcription factors and global physiology of the cell. *Mol. Syst. Biol.* **9,** (2013).

19. Gerosa, L., Kochanowski, K., Heinemann, M. & Sauer, U. Dissecting specific and global transcriptional regulation of bacterial gene expression. *Mol. Syst. Biol.* **9,** 658 (2013).

20. Klumpp, S. & Hwa, T. Growth-rate-dependent partitioning of RNA polymerases in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 20245–50 (2008).

21. Regenberg, B. *et al.* Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in Saccharomyces cerevisiae. *Genome Biol.* **7,** R107 (2006).

22. Klumpp, S., Zhang, Z. & Hwa, T. Growth rate-dependent global effects on gene expression in bacteria. *Cell* **139,** 1366–75 (2009).

23. Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330,** 1099–102 (2010).

24. Levy, S. & Barkai, N. Coordination of gene expression with growth rate: a feedback or a feed-forward strategy? *FEBS Lett.* **583,** 3974–8 (2009).

25. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297,** 1183–6 (2002).

26. Acar, M., Becskei, A. & van Oudenaarden, A. Enhancement of cellular memory by reducing stochastic transitions. *Nature* **435,** 228–32 (2005).

27. Colman-Lerner, A. *et al.* Regulated cell-to-cell variation in a cell-fate decision system. *Nature* **437,** 699–706 (2005).

28. Raser, J. M. & O’Shea, E. K. Control of stochasticity in eukaryotic gene expression. *Science* **304,** 1811–4 (2004).

29. Volfson, D. *et al.* Origins of extrinsic variability in eukaryotic gene expression. *Nature* **439,** 861–4 (2006).

30. Newman, J. R. S. *et al.* Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441,** 840–6 (2006).

31. Raj, A. & van Oudenaarden, A. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell* **135,** 216–226 (2008).

32. Ehrenberg, M. & Kurland, C. G. Costs of accuracy determined by a maximal growth rate constraint. *Q. Rev. Biophys.* **17,** 45–82 (1984).

33. Costenoble, R. *et al.* Comprehensive quantitative analysis of central carbon and amino-acid metabolism in Saccharomyces cerevisiae under multiple conditions by targeted proteomics. *Mol. Syst. Biol.* **7,** 464 (2011).

34. Molenaar, D., van Berlo, R., de Ridder, D. & Teusink, B. Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Syst. Biol.* **5,** 323 (2009).

35. Ibarra, R. U., Edwards, J. S. & Palsson, B. O. Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* **420,** 186–9 (2002).

36. Barrick, J. E. & Lenski, R. E. Genome dynamics during experimental evolution. *Nat. Rev. Genet.* **14,** 827–839 (2013).

37. Lovén, J. *et al.* Revisiting global gene expression analysis. *Cell* **151,** 476–82 (2012).

38. Lubliner, S., Keren, L. & Segal, E. Sequence features of yeast and human core promoters that are predictive of maximal promoter activity. *Nucleic Acids Res.* **41,** 5569–81 (2013).

39. Wang, M. *et al.* PaxDb, a database of protein abundance averages across all three domains of life. *Mol. Cell. Proteomics* **11,** 492–500 (2012).

40. Setty, Y., Mayo, A. E., Surette, M. G. & Alon, U. Detailed map of a cis-regulatory input function. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 7702–7 (2003).

41. Field, Y. *et al.* Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput. Biol.* **4,** e1000216 (2008).

42. Tirosh, I. & Barkai, N. Two strategies for gene regulation by promoter nucleosomes. *Genome Res.* **18,** 1084–91 (2008).

43. Peccoud, J. & Ycart, B. Markovian Modeling of Gene-Product Synthesis. *Theor. Popul. Biol.* **48,** 222–234 (1995).