Signatures of nitrogen limitation in the elemental composition of the proteins involved in the metabolic apparatus

Claudia Acquisti^{1,2,*}, Sudhir Kumar^{1,2} and James J. Elser²

¹*Biodesign Institute, Center for Evolutionary Functional Genomics, Arizona State University, Tempe, AZ 85287-5301, USA*

²*School of Life Sciences, Arizona State University, Tempe, AZ 85287-4501, USA*

Nitrogen (N) is a fundamental component of nucleotides and amino acids and is often a limiting nutrient in natural ecosystems. Thus, study of the N content of biomolecules may establish important connections between ecology and genomics. However, while significant differences in the elemental composition of whole organisms are well documented, how the flux of nutrients in the cell has shaped the evolution of different cellular processes remains poorly understood. By examining the elemental composition of major functional classes of proteins in four multicellular eukaryotic model organisms, we find that the catabolic machinery shows substantially lower N content than the anabolic machinery and the rest of the proteome. This pattern suggests that ecological selection for N conservation specifically targets cellular components that are highly expressed in response to nutrient limitation. We propose that the RNA component of the anabolic machineries is the mechanistic force driving the elemental imbalance we found, and that RNA functions as an intracellular nutrient reservoir that is degraded and recycled during starvation periods. A comparison of the elemental composition of the anabolic and catabolic machineries in species that have experienced different levels of N limitation in their evolutionary history (animals versus plants) suggests that selection for N conservation has preferentially targeted the catabolic machineries of plants, resulting in a lower N content of the proteins involved in their catabolic processes. These findings link the composition of major cellular components to the environmental factors that trigger the activation of those components, suggesting that resource availability has constrained the atomic composition and the molecular architecture of the biotic processes that enable cells to respond to reduced nutrient availability.

Keywords: molecular evolution; cellular components; elemental composition; metabolic apparatus; nutrient limitation; biological stoichiometry

1. INTRODUCTION

A major challenge in biology is to integrate ecosystem and evolutionary perspectives to better understand patterns of adaptive variation in the biota. One such link involves the costs of resource acquisition from the environment and the consequences of environmental limitations on the biochemical composition of different organisms. Adaptation to environmental challenges such as limitation by key nutrient elements (e.g. nitrogen, N) may often manifest at the level of metabolic pathways, because the biochemical make-up of an organism has functional consequences that influence its fitness. Recent reports have begun to reveal some of the adaptive genetic mechanisms underlying stoichiometric evolution on different time scales (Baudouin-Cornu *et al.* 2001, 2004; Bragg & Hyder 2004; Bragg *et al.* 2006; Elser *et al.* 2006; Acquisti *et al.* 2007, 2009; Bragg & Wagner 2007). However, the potential for adaptive allocation of specific elements at the subcellular level of metabolic pathways and cellular components is still poorly understood. A deeper knowledge of the stoichiometric architecture of the subcellular apparatus, and its evolutionary dynamics

under varying environmental conditions, is paramount to understanding the mechanisms of adaptation to nutrient limitation.

The elemental composition of proteins and genes differs substantially among and within taxonomic groups, and the cost of nutrient acquisition appears to be one of the factors that influence the molecular evolution of proteomes and genomes. For example, Baudouin-Cornu *et al.* (2001) documented a significant correlation between atomic composition and metabolic function in a specific set of assimilatory enzymes in micro-organisms. They suggested that, in response to nutrient limitation, natural selection may have shaped the elemental composition of proteins with specific metabolic functions (Mazel & Marlière 1989; Baudouin-Cornu *et al.* 2001). However, whether selective mechanisms have shaped the stoichiometry of biotic apparatus involved in other metabolic processes and in multicellular organisms is still unclear.

Complex adaptive mechanisms enable organisms to survive periods of nutrient deprivation via the inhibition of growth and via the parallel mobilization of intracellular reservoirs of nutrients (Kraft *et al.* 2008). Indeed, a highly regulated equilibrium between anabolism and catabolism allows organisms to grow, reproduce, maintain their structures and respond to their variable environment.

* Author and address for correspondence: Biodesign Institute, Center for Evolutionary Functional Genomics, Arizona State University, Tempe, AZ 85287-5301, USA (claudia.acquisti.1@asu.edu).

Anabolic processes use energy to construct the molecular components of cells, whereas catabolic processes break down molecules to regulate their turnover and to harvest energy and recycle nutrients. During favourable environmental conditions (e.g. nutrient abundance), metabolic regulation favours anabolism over catabolism, resulting in net growth. By contrast, nutrient stress triggers the activation of catabolism and the attenuation of anabolism.

It would logically follow that, to optimize nutrient allocation, the catabolic machinery should have a lower demand for the nutrients that are most often limiting the biota. Growth and production of plants are often limited by N in many ecosystems (Elser *et al.* 2007). Such limitations are not surprising, as N is an essential component of nucleotides and amino acids, but external supplies are often insufficient for demand, thus establishing selection pressure for N conservation in the construction of biomass. Interestingly, the number of N atoms per monomer ranges between one and four in amino acids and between two and five in nucleotides (resulting in carbon : nitrogen ratios between 1.5 and 9 in amino acids and between 1 and 2.5 in nucleotides). Thus, nutrient limitation may in principle influence the molecular composition of genes and proteins by favouring monomer usage biases that conserve the limiting elements. Indeed, a number of recent studies have shown that environmental N limitation, in addition to structural and functional constraints, can directly affect the average composition of proteomes and genomes in different taxa (Bragg & Hyder 2004; Elser *et al.* 2006; Acquisti *et al.* 2009). Here, we investigate whether selection for nutrient conservation differentially impacts the stoichiometry of the metabolic apparatus that is centrally involved in cellular response to nutritional stress (catabolism) and to nutrient abundance (anabolism).

To test the possibility that the elemental composition of proteins involved in major metabolic pathways and cellular components is linked to the environmental factors that trigger their activation, we analysed the N content of the proteins involved in the anabolic and catabolic apparatus in four multicellular eukaryotic model organisms. We further investigated the observed N-usage patterns in the context of cellular location and the specific metabolic role of the proteins involved in the anabolic and catabolic apparatus. We distinguished two main classes of functions: a wide group of enzymes that catalyse specific metabolic reactions; and the intracellular structures in which broader metabolic activities take place ('metabolic machinery'), such as protein synthesis and protein degradation. Metabolic processes involve extensive co-operation between different types of enzymes and anabolic and catabolic machineries. However, due to their larger size and high copy number, the metabolic machinery typically represents a much larger fraction of the cellular mass. Therefore, if evolutionary pressures for strategic elemental investment have acted to conserve N in the catabolic apparatus, the strongest biases towards low-N monomers are expected in the catabolic machinery. We explored this hypothesis by comparing different classes of enzymes involved in the catabolism and anabolism of amino acids, nucleotides and lipids, and larger molecular complexes that act as anabolic and catabolic machineries (anabolic: ribosome and spliceosome; catabolic: proteosome, lysosome and vacuole). We further tested our

hypothesis by comparing the elemental composition of the anabolic and catabolic machineries with that of the rest of the proteome, in taxa with different levels of N limitation in their evolutionary history, contrasting plant and animal model organisms. Overall, our results suggest that selection for nutrient conservation has altered the elemental composition of major cellular components in response to their metabolic role in the flux of nutrient and energy in the cell. This suggests a fundamental role of nutrient availability in the evolution of the molecular architecture of metabolic pathways and cellular components.

2. MATERIAL AND METHODS

Protein sequences were obtained for four multicellular eukaryotic model organisms *Arabidopsis thaliana*, *Drosophila melanogaster*, *Homo sapiens* and *Mus musculus* from the Gene Ontology database (release go_200806 <http://www.geneontology.org>) for the following functional groups of proteins: cellular catabolism (GO:0044248); cellular anabolism (GO:0044249); anabolic machinery (ribosome GO:0005840; spliceosome GO:0005681); catabolic machinery (proteasome GO:0000502; vacuole GO:0005773; lysosome GO:0005764); catabolic pathways (RNA catabolism GO:0006401; amino acid catabolism GO:0009063; lipid catabolism GO:0016042); and anabolic pathways (RNA anabolism GO:0032774; amino acid anabolism GO:0008652; lipid anabolism GO:0008610). The species were chosen based on the amount of data available in the database, and *A. thaliana* was the only plant species for which none of the categories listed was empty. Different functional groups contain overlap, as some proteins may participate in more than one functional class. Paralogs and alternatively spliced forms annotated in the same functional GO group were included in the dataset. The size of each dataset is given in the caption of figure 1.

Full proteomes were obtained for *A. thaliana* (www.tair.org), *H. sapiens*, *M. musculus* and *D. melanogaster* (<http://genome.ucsc.edu/>).

The N content of each protein was estimated as follows:

$$N_c = \frac{\sum n_i \times p_i}{L},$$

where n_i is the number of N atoms of each i th amino acid side chain ($n=1$ for asparagine, glutamine, lysine, and tryptophan; $n=2$ for histidine; $n=3$ for arginine; and $n=0$ for the rest); p_i is the count of the i th amino acid in the protein analysed, and L is the protein length.

3. RESULTS

(a) N content of the anabolic and catabolic apparatus

Consistent with the expectations developed above, on a proteomic scale comparison of proteins involved in anabolic and catabolic cellular processes (GO:0044249 and GO:0044248, respectively), in four eukaryotic model organisms (*A. thaliana*, *D. melanogaster*, *H. sapiens* and *M. musculus*), we found that N content per amino acid side chain was 40 per cent higher in the anabolic than in the catabolic processes in each organism analysed (Student's t -test $p \ll 0.001$; data not shown).

If natural selection has constrained the elemental composition of specific metabolic pathways in a regime of efficient nutrient usage, we expect to find the lowest nutritional contents in the most abundant cellular

Figure 1. Quantile plot of N content of the proteins involved in anabolic and catabolic cellular machineries and enzymatic pathways (see §2 for data assembly) in (a,e) *A. thaliana* (anabolic machinery: mean=0.4817, s.d.=0.1640, $n=456$; catabolic machinery: mean=0.3129, s.d.=0.0606, $n=187$; anabolic enzymes: mean=0.3871, s.d.=0.0608, $n=1680$; catabolic enzymes: mean=0.3461, s.d.=0.0440, $n=103$), (b,f) *D. melanogaster* (anabolic machinery: mean=0.4745, s.d.=0.1323, $n=406$; catabolic machinery: mean=0.34552, s.d.=0.0644, $n=196$; anabolic enzymes: mean=0.3935, s.d.=0.0693, $n=1352$; catabolic enzymes: mean=0.3785, s.d.=0.0631, $n=162$), (c,g) *H. sapiens* (anabolic machinery: mean=0.4788, s.d.=0.1281, $n=542$; catabolic machinery: mean=0.3498, s.d.=0.0602, $n=464$; anabolic enzymes: mean=0.4068, s.d.=0.0801, $n=1335$; catabolic enzymes: mean=0.3682, s.d.=0.0601, $n=272$), (d,h) *M. musculus* (anabolic machinery: mean=0.443, s.d.=0.1274, $n=436$; catabolic machinery: mean=0.3269, s.d.=0.0666, $n=277$; anabolic enzymes: mean=0.3727, s.d.=0.0587, $n=952$; catabolic enzymes: mean=0.3549, s.d.=0.0592, $n=159$). s.d. is the standard deviation and n is the number of protein sequences analysed.

Figure 2. Scatter plot of the number of atoms in the side chain and pI for each of the 20 amino acids (Pearson's correlation coefficient $R=0.71$, $p=0.0004$).

components that respond to major changes in the flux of nutrients and energy in the cell. While enzymes and pathways are expressed in variable amounts in specific metabolic contexts, the anabolic and catabolic machineries represent the most abundant hardware ubiquitously involved in the synthesis and degradation processes, respectively. Since severe nutrient limitation reduces anabolic processes and promotes catabolic breakdown of cellular constituents, more efficient element use should be associated with the catabolic machinery. This is what we observed. The N content of catabolic enzymes is

significantly lower than that of anabolic enzymes in all the organisms studied (figure 1a–d; t -test $p \ll 0.01$). The highest differences are seen in the comparison of the anabolic and catabolic machineries (figure 1e–h), where the N content per amino acid side chain is over 50 per cent lower in the catabolic than in the anabolic apparatus (t -test $p \ll 0.001$). These results point to a more parsimonious element usage in the construction of the molecular apparatus of catabolism.

(b) Effects of structural constraints on the N content of the anabolic and catabolic machineries

Components of the anabolic machinery (spliceosome and ribosome) involved in different biological processes (transcription and translation) and localized in different cellular regions (nuclear and cytoplasmic) show very similar N content to each other, but were different from all the catabolic machineries analysed (figure 1). This raises the question of the mechanistic underpinnings of higher N use in the molecular complexes involved in anabolism and N conservation in those involved in catabolism.

We propose here that molecular interactions involving the riboprotein components of the anabolic machinery are a major driving force behind this atomic composition imbalance. The core structure of riboproteic complexes is substantially stabilized by electrostatic interactions between RNA and proteins. The negative charge of

Figure 3. N content of different functional groups of proteins. Mean values and standard deviations are shown for different functional classes of nucleic acid binding proteins, enzymes and membrane proteins in *H. sapiens*. Number of proteins analysed for each dataset: ribosome, GO:0005840, $n=409$; spliceosome, GO:0005681, $n=133$; nucleic acid-binding proteins, GO:0003676, $n=4867$; transcription, GO:006350, $n=3252$; translation, GO:0006412, $n=610$; hydrolase activity, GO:0016787, $n=3169$; isomerase activity, GO:0016853, $n=237$; ligase activity, GO:0016874, $n=517$; lyase activity, GO:0016829, $n=224$; oxidoreductase activity, GO:0016491, $n=1126$; transferase activity, GO:0016740, $n=2665$; and transporter activity, GO:0022857, $n=1180$.

Figure 4. N content of the metabolic apparatus within species (a,b) and between species (c,d). For sample sizes and mean values, see caption of figure 1.

RNA molecules constrains the protein component to have a high density of basic amino acids with positive charge and a high isoelectric point (pI). Amino acids that bind to nucleic acids have a higher N content than most other amino acids, because the pI correlates positively with the N content of the amino acids (Pearson's correlation coefficient $R=0.71$, $p=0.0004$; figure 2). In particular, in the range of pH close to neutrality that exists in intracellular environments (Karagiannis & Young 2001), all the amino acids with positive charge contain extra N atoms in their side chain (figure 2). Furthermore, compared with other intracellular compounds, nucleotides are highly nutrient rich with a carbon : nitrogen ratio (C : N) between 1 and 2.5 (Stern & Elser 2002). Therefore, riboproteic complexes are an extremely nutrient-rich pool in the cell due both to the selective

usage of N-rich amino acids in the protein component and to the high N and P content of the RNA molecules.

To test the hypothesis that the ability to bind to RNA molecules is a factor determining the N content of proteins, we compared nucleic acid-associated proteins (e.g. proteins involved in transcription and translation) with proteins with different structural requirements (e.g. membrane proteins and enzymes). While different functional classes of enzymes have N contents similar to that of the rest of the proteome (figure 3), all the functionally diverse classes of proteins that bind nucleic acids show a significantly higher N content than the rest of the proteome (t -test: $p \ll 0.001$). This indicates that the physical properties that enable proteins to bind to nucleic acid have a major impact on their biosynthetic costs in terms of N.

(c) N content of the catabolic apparatus in N-limited and N-sufficient species

Previous studies have suggested that ecophysiological selection has operated to conserve N in transcripts and proteins of plants compared with animals, reflecting more severe N limitation in the evolutionary history of plants (Elser *et al.* 2006; Acquisti *et al.* 2009). On this basis, one may also expect that the difference in N content between the catabolic machinery and the full proteome should be greater in *A. thaliana* than in *D. melanogaster*, *H. sapiens* and *M. musculus*. Also, the differences in N content between plant and animal proteins should be greater in the catabolic than in the anabolic machineries. Both these predictions are fulfilled. In animals, the N content of the catabolic machineries is very similar to that of the rest of the proteome (figure 4b). By contrast, the N content of the catabolic machineries was more than 10 per cent lower than the rest of the proteome in *A. thaliana* (figure 4a; $p \ll 0.001$ *t*-test). The N content of catabolic components is more than 10 per cent lower in *A. thaliana* than in animals (figure 4c), which is 10 times higher than the difference between the N content of the anabolic machineries of *A. thaliana* and animals (figure 4d). All together these results give further support to the hypothesis that selection pressures for N conservation constrained the composition of plant proteins (Elser *et al.* 2006; Acquisti *et al.* 2009), showing that in plants selection pressure for N conservation is stronger in proteins highly expressed during nutrient limitation, while it allows for higher N content of the anabolic apparatus involved in fast growth.

4. DISCUSSION

Our findings suggest that evolutionary pressures for strategic nutrient investment are among the forces that have fundamentally constrained the molecular architecture of metabolic pathways and cellular components in multicellular eukaryotes. These results extend previous studies on the nutritional costs of assimilatory enzymes in microbes (Baudouin-Cornu *et al.* 2001) to a wider cellular scale, with a major focus on the cellular components that are most abundantly expressed during different levels of nutrient availability. As expected under a strategy of efficient elemental allocation, we found that the signature of nutrient conservation is higher in cellular components that contribute the most to the cellular mass (catabolic machinery) rather than in specific enzymes.

Our data on the biochemical and structural factors leading to the different elemental costs of the anabolic and catabolic apparatus suggest that the RNA component of the anabolic machineries is a major factor underpinning the stoichiometric differences between the two apparatus. Owing to physical-chemical properties, the charge distribution of nucleotides requires a high N content in proteins that have a close physical interaction with nucleotides, because N-rich amino acids tend to have positive charge. This purely functional point of view suggests a high nutritional cost of the processes (such as transcription, translation and genome replication) that operate during fast growth and nutrient sufficiency, reinforcing the idea that resource availability and the optimization of nutrient allocation are among the factors shaping the molecular architecture of cellular structures.

We propose that the high N content of structural anabolic cellular components has an additional benefit in enhancing the reservoirs of N that are used during starvation when anabolic pathways are suppressed. This points to an effect of selection for robust nutrient allocation during the evolution of metabolic pathways. In the absence of nutrient limitation, growth and synthesis are promoted via increased transcription and translation rates that lead to an increased amount of ribosomes (Gourse *et al.* 1996) and most probably other anabolic machinery, such as spliceosomes. By contrast, in times of nutrient limitation, growth is inhibited and the anabolic machinery becomes redundant. In this scenario, the nutrients stored in the anabolic machinery can be released to the cell. Indeed, it has been shown that some types of autophagy, once thought to be a purely non-selective process, involve selective targeting of anabolic structural components (e.g. ribosomes) for degradation and recycling during starvation (Kraft *et al.* 2008). Such mechanisms of response to nutrient stress are especially efficient if the nutrient content of the anabolic components is higher than the N content of the catabolic apparatus. Thus, because of the positive anabolic N bias, the shift from an anabolic- to a catabolic-dominated metabolism leads to a larger liberation of key nutrients than would otherwise occur in the absence of such N-rich usage biases. These newly released nutrients then are used to efficiently construct the low-N machinery of catabolism. Recent results have shown that the fitness costs of protein synthesis primarily derive from the costs of occupying the machinery of expression, and not from the costs of protein production *per se* (Stoebel *et al.* 2008). This suggests that the number of copies of the transcriptional and translational apparatus is tightly regulated, further supporting our interpretation of the high nutritional costs of the anabolic machinery.

An alternative possibility is that the RNA component of the anabolic machinery is primarily dictated by structural interactions with nucleic acids than by nutritional constraints, as both the ribosome and the spliceosome interact with transcripts (while the proteasome, the vacuole and the lysosome do not). This possibility seems unlikely, as the catabolism of RNA molecules occurs in a multi-protein complex that lacks any RNA component (exosome). This is compatible with the hypothesis that the association between RNA and proteins is avoided in the catabolic apparatus due to evolutionary pressures for strategic resource allocation.

Comparative analysis of the N cost of the anabolic and catabolic apparatus in plants and animals represents a further test of our hypothesis. Consistent with the predictions of selection for N conservation in the catabolic apparatus in more chronically N-limited biota (plants), we found that the N content of the catabolic machineries is 10 per cent lower in *Arabidopsis* than in animals, while the N content of the anabolic apparatus did not differ significantly. Our results refine previous observations that highly expressed proteins have an especially lower N content in plants than in animals (Elser *et al.* 2006). We show here that this pattern preferentially reflects the structural properties and the elemental composition of the catabolic apparatus, components of which are highly expressed during nutrient limitation.

Our findings connect nutrient limitation to molecular evolution in the conceptual framework of biological stoichiometry (Elser & Hamilton 2007), and link the role of environmental growth limitation to the evolution of the subcellular components involved in the metabolic responses of organisms to their natural environments. The integration of the broad role of nutrient limitation into the study of the evolution of metabolic apparatus represents a further step in improving our understanding of how environmental factors have influenced molecular evolution, promising to shed new light on the evolution of metabolic pathways.

We thank Thomas Wiehe, Marcia Kyle, Alan Filipinski, Bill Fagan, James Gilbert, Antonio Marco, Bernhard Haubold and Fabia Battistuzzi for their valuable scientific discussions and Bernard Van Emden and Revak Raj Tyagi for their technical support. This work is supported in part by a research grant from the National Science Foundation (J.J.E. and S.K.).

REFERENCES

- Acquisti, C., Kleffé, J. & Collins, S. 2007 Oxygen content of transmembrane proteins over macroevolutionary time scales. *Nature* **445**, 47–52. (doi:10.1038/nature05450)
- Acquisti, C., Elser, J. J. & Kumar, S. 2009 Ecological nitrogen limitation shapes the DNA composition of plant genomes. *Mol. Biol. Evol.* **26**, 953–956. (doi:10.1093/molbev/msp038)
- Baudouin-Cornu, P., Surdin-Kerjan, Y., Marliere, P. & Thomas, D. 2001 Molecular evolution of protein atomic composition. *Science* **293**, 297–300. (doi:10.1126/science.1061052)
- Baudouin-Cornu, P., Schuerer, K., Marliere, P. & Thomas, D. 2004 Intimate evolution of proteins. Proteome atomic content correlates with genome base composition. *J. Biol. Chem.* **279**, 5421–5428. (doi:10.1074/jbc.M306415200)
- Bragg, J. G. & Hyder, H. C. 2004 Nitrogen versus carbon use in prokaryotic genomes and proteomes. *Proc. R. Soc. Lond. B* **271**, S374–S377. (doi:10.1098/rsbl.2004.0193)
- Bragg, J. G. & Wagner, A. 2007 Protein carbon content evolves in response to carbon availability and may influence the fate of duplicated genes. *Proc. R. Soc. B* **274**, 1293–1300. (doi:10.1098/rspb.2006.0290)
- Bragg, J. G., Thomas, D. & Baudouin-Cornu, P. 2006 Variation among species in proteomic sulphur content is related to environmental conditions. *Proc. R. Soc. B* **273**, 1063–1070. (doi:10.1098/rspb.2005.3441)
- Elser, J. J. & Hamilton, A. L. 2007 Stoichiometry and the new biology: the future is now. *PLoS Biol.* **5**, e181. (doi:10.1371/journal.pbio.0050181)
- Elser, J. J., Fagan, W. F., Subramanian, S. & Kumar, S. 2006 Signature of ecological resource availability in the animal and plant proteomes. *Mol. Biol. Evol.* **23**, 1946–1951. (doi:10.1093/molbev/msl068)
- Elser, J. J. *et al.* 2007 Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecol. Lett.* **10**, 1135–1142. (doi:10.1111/j.1461-0248.2007.01113.x)
- Gourse, R. L., Gaal, T., Bartlett, M. S., Appleman, J. A. & Ross, W. 1996 rRNA transcription and growth rate-dependent regulation of ribosomes synthesis in *E. coli*. *Annu. Rev. Microbiol.* **50**, 645–677. (doi:10.1146/annurev.micro.50.1.645)
- Karagiannis, J. & Young, P. G. 2001 Intracellular pH homeostasis during cell-cycle progression and growth state transition in *Schizosaccharomyces pombe*. *J. Cell Sci.* **114**, 2929–2941.
- Kraft, C., Deplazes, A., Sohrmann, M. & Peter, M. 2008 Mature ribosomes are selectively degraded upon starvation by an autophagy pathway requiring the Ubp3p/Bre5p ubiquitin protease. *Nat. Cell Biol.* **10**, 602–610. (doi:10.1038/ncb1723)
- Mazel, D. & Marlière, P. 1989 Adaptive eradication of methionine and cysteine from cyanobacterial light-harvesting proteins. *Nature* **341**, 245–248. (doi:10.1038/341245a0)
- Sterner, R. W. & Elser, J. J. 2002 *Ecological stoichiometry*. Princeton, NJ: Princeton University Press.
- Stoebel, D. M., Dean, A. M. & Dykhuizen, D. E. 2008 The cost of expression of *Escherichia coli* lac operon proteins is in the process, not in the products. *Genetics* **178**, 1653–1660. (doi:10.1534/genetics.107.085399)