Trends in
Microbiology

Opinion

The energetic costs of cellular complexity in evolution

Sergio A. Muñoz-Gómez [1,*]

The evolutionary history of cells has been marked by drastic increases in complexity. Some hypothesize that such cellular complexification requires a massive energy flux as the origin of new features is hypothetically more energetically costly than their evolutionary maintenance. However, it remains unclear how increases in cellular complexity demand more energy. I propose that the early evolution of new genes with weak functions imposes higher energetic costs by overexpression before their functions are evolutionarily refined. In the long term, the accumulation of new genes deviates resources away from growth and reproduction. Accrued cellular complexity further requires additional infrastructure for its maintenance. Altogether, this suggests that larger and more complex cells are defined by increased survival but lower reproductive capacity.

Highlights

The evolutionary relationship between energy and cellular complexity remains puzzling.

Proteome allocation theory provides a framework to understand the energetic costs of new genes and organelles.

New genes often emerge with weak functions that require overexpression, and this deviates greater resources prior to evolutionary refinement.

The evolutionary trajectory of new genes can be defined by early high relative energetic costs and low functional performance and is similar to that of complex cellular features that increase in functional specialization.

Large increases in cellular complexity decrease the proteome fraction that can be devoted to protein synthesis and thus directly reduce growth rate.

New complex features are often accompanied by infrastructure that supports their assembly and maintenance. This deviates further resources away from reproduction.

Cells with more complex proteomes inevitably allocate fewer resources to protein synthesis but arguably have a higher survival capacity.

Complexity and energy demands

The origin of eukaryotic cells (i.e., eukaryogenesis) represents the major structural transformation and leap of complexity in cell evolution [1,2]. It is estimated that ~6000 new genes, which underlie the structural and functional complexity of eukaryotes, evolved during this transition [3]. To explain the complexity gap at the prokaryote–eukaryote divide, some authors have argued that the drastic cellular complexification that accompanied the origin of eukaryotes required a major energy flux that could only have been provided by ATP-producing mitochondria [4]. In this view, mitochondria provided a twofold advantage by internalizing respiratory membranes (thus releasing them from following a sublinear scaling with cell volume), and supposedly creating an asymmetric genome architecture (i.e., reduced mitochondrial genomes supporting an expanded nuclear genome) that was biosynthetically much less costly [5,6].

In support of such a view, it has also been stated that the origin of complexity (e.g., complex features or new genes) requires a higher energetic investment than its subsequent evolutionary maintenance. For instance, Lane and Martin stated that: '…the energetic cost for the *de novo* 'invention' of complex traits like phagocytosis must far exceed the costs of simply inheriting a functional system' [7], and 'Once the inventing is over maintaining those traits is energetically moot. By analogy, it takes far more energy to build a suspension bridge than it does to maintain it, once finished.' [8].

These arguments have, in part, been made to counter the fact that anaerobic eukaryotes that exclusively rely on fermentation, a much lower energy-yielding metabolism, can be just as large and structurally complex as their close aerobic relatives [9]. The above statements have confused others, who have claimed that: 'There is no reason to believe that evolving a new trait ('expressing novel protein families') takes more cellular energy than maintaining it.' [10], and 'There is no theoretical or comparative evidence to support the imagination of such "exuberant evolutionary scaffolding" that would require a transient appearance of a huge number of genes exceeding the final count by up to an order of magnitude.' [11].

[1]Department of Biological Sciences, Purdue University, West Lafayette, IN, USA

*Correspondence:
samunozg@purdue.edu
(S.A. Muñoz-Gómez).

© 2024 Elsevier Ltd. All rights reserved.

The notion that increased complexity demands a higher energy flux is intuitive to many (e.g., [4] has been cited >1200 times) and appears to stem from analogies with human-made artefacts. For instance, smartphones (e.g., iPhone 14) have considerably many more parts and functions than older mobile phones (e.g., Nokia 3310), but also have a much higher power consumption. And the notion that new features demand more energy at their origin also applies to human-made artefacts that are relatively stable after an expensive manufacturing or construction process (e.g., bridges; [8]). However, living cells differ fundamentally from inert objects, as they grow, reproduce, and evolve.

The energy demands of a cell are expected to be a consequence of its volume and growth rate [6,9,12] (Box 1 and Figure I in Box 1). This is supported by a linear and continuous scaling of

Box 1. The evolutionary relationship between cellular complexity and size

There exists an indirect relationship between energy and cellular complexity. Energy demands are directly determined by cell volume and growth rate (see main text and Figure I). Thus, larger cell volumes impose higher absolute energy demands (Figure I). Larger cell volumes are, in turn, loosely correlated with increased cellular complexity – this correlation, however, only becomes apparent over a massive range of eight orders of magnitude in cell volume and across phylogenetically distant cell lineages. This complexity–size correlation furthermore emerges from the interrelated facts that (i) larger cell volumes become possible by complex adaptations that allow cells to overcome physiological or biophysical constraints, and (ii) larger cell volumes themselves impose selective pressures for the evolution of complex cellular features [50,51]. For example, mitochondria and the endomembrane system allowed for the expansion of respiratory and nutritional membranes, respectively, thus overcoming the surface area–volume constraints, and a motor-driven cytoskeleton allowed for active cytoplasmic transport over long distances, thus overcoming diffusion constraints.

Another indirect evolutionary relationship between energy and cellular complexity is possible. Koonin argued that larger eukaryotic cell volumes are energetically permitted by the internalized respiratory membranes within mitochondria that allow respiratory rates to scale linearly with cell volume (as opposed to as if they were localized at the cytoplasmic membrane/cell's surface) [4,52] – this has recently been quantitatively shown in relation to growth rate [6]. Larger cell volumes, Koonin argues, enable passive genome expansion because larger cells (i) progressively invest fewer energetic resources into DNA relative to other cellular components, and (ii) arguably make smaller populations where random genetic drift exerts a stronger power relative to natural selection. DNA can thus more easily accumulate in larger-celled species because of its decreased relative energetic costs and the higher probability of becoming evolutionarily fixed by stochastic processes [52]. An expanded genome may be permissive to proteome complexification. However, it is important to note that genome expansion in eukaryotic cells is primarily driven by non-coding DNA [53,54]. The above view furthermore leaves unaddressed the relative costs of proteome complexity, and how they change throughout evolutionary history (see main text).

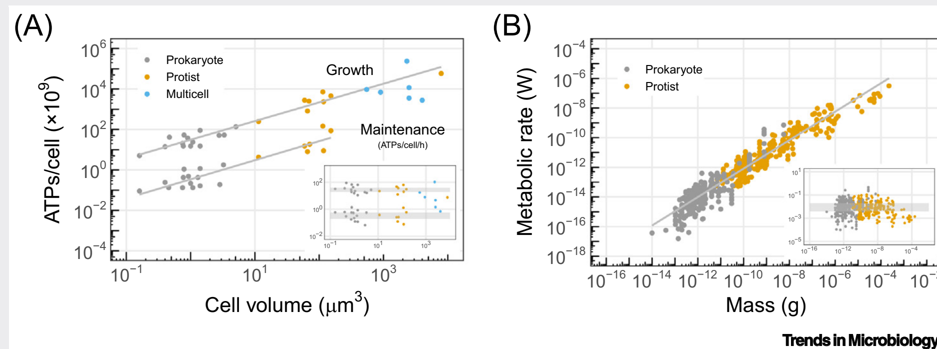


Figure I. The energy demands of a cell are primarily dictated by their volume regardless of complexity. (A) Lifetime ATP requirements of prokaryotic and eukaryotic cells scale linearly with volume. The regression equation for the ATP requirements for growth (in units of ATP/cell) and maintenance (in units of ATP/cell/h) are $y = 30.16x^{0.93}$ (excludes multicellular species) and $y = 0.40x^{0.88}$, respectively. The inset shows the ATP requirements per unit volume. The shaded area shows interquartile range. Adapted from [12]. (B) Mass-specific metabolic rate of prokaryotic and eukaryotic cells is approximately constant across a wide range in volume. The regression equation is $y = 0.0033x^{0.96}$. The inset shows the mass-specific metabolic rate. The shaded area shows interquartile range. Data from [13].

Glossary

Evolutionary trajectory: historical pattern of changes that have occurred in a gene, cellular feature, or cell lineage.
Gene amplification: the successive duplication of a gene by several mechanisms such as rolling circle amplification or nonequal crossing over between sister chromatids. By increasing gene dosage, gene amplification allows for much faster adaptive evolutionary response to selective pressures imposed by changes in both the internal and external environment.
Growth law: an informal term that refers to a series of empirical relationships observed in prokaryotes between growth rate, cellular composition (e.g., RNA: protein ratio), and cell size.
Lifetime ATP requirements: the number of ATP molecules or equivalents that are required for cellular growth and maintenance (or construction and operation) throughout the life cycle of a cell, that is, between cell divisions. This measure of energy demands was introduced by Lynch and Marinov [12].
Metabolic rate: the amount of energy per unit time that a cell requires to carry out its metabolic processes. It is often measured on growing and feeding (i.e., active metabolic rate) or resting or starved cells (inactive, endogenous, or basal metabolic rate). It is most often measured as the rate of oxygen consumption by a cell and expressed in units of nl O_2 per cell and thus refers to the rate at which aerobic respiration occurs.
Physiological constraint: limitations imposed by homeostatic properties of cells that need to be maintained within narrow boundaries for proper cellular function (e.g., biomass density) or biophysical factors that cannot be overcome by biological adaptations.
Proteomap: a Voronoi tree graph that summarizes the quantitative proteome of a cell by showing proteins as polygons whose sizes indicate abundance. Proteins involved in similar cellular functions are arranged in adjacent locations, creating regions whose areas give insight into the relative investment in each functional class. This visualization was introduced by Liebermeister *et al.* [37].
Proteome allocation: the relative fraction of proteins, in number or mass, dedicated to a particular cellular process or function in a cell.

CellPress

the **lifetime ATP requirements** (see Glossary) [12] (see Figure IA in Box 1), and **metabolic rates** [13–15] (see Figure IB in Box 1), with cell volume across prokaryotes and eukaryotes. This suggests that energy demands per unit volume are approximately constant across a 10^5- or 10^{10}-fold range in volume and mass, respectively, that spans phylogenetically and structurally diverse cells [16,17] (see Figure I in Box 1) – two cells that differ in their degree of complexity (e.g., number of protein types) demand the same amount of energy as long as they have the same volumes and growth rates. Cellular complexity, therefore, does not directly impinge on energy demands.

Proteome complexity: refers to the number of different proteins synthesized by a cell and is expected to be proportional to the number of genes encoded by a cell's genome. It may also be quantified as the number of protein domains or folds or unique combinations of these.

Unnecessary protein: a protein that is expressed by the cell but whose function does not provide any benefit under the current environmental conditions, for example, an enzyme in the absence of its substrate.

Global physiological constraints on cells

Are there energetic costs associated with increases in cellular complexity? The arguments provided above suggest that there is no direct relationship between energy and cellular complexity. However, a consideration of the global **physiological constraints** that act on cells may shed light on this question. One of the most fundamental physiological constraints on cells is the near-constancy of biomass and protein density (proteins constitute ~40–60% of a cell's dry mass [15]). This has been shown to hold within and across species that span more than ten orders of magnitude in cell volume [15,18] (Figure 1A,B). The observation that energy demands per unit volume are constant across phylogenetically disparate cells (see above; see Figure I in Box 1) follows naturally from this physiological constraint.

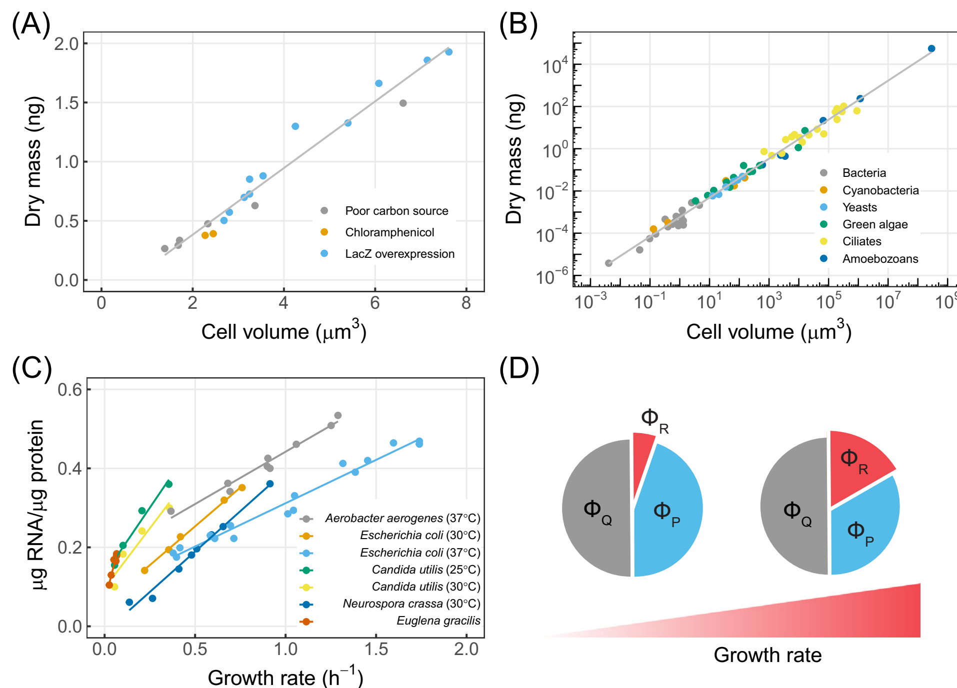


Figure 1. The constancy of cell density and the linear relationship between growth rate and ribosome concentration (i.e., the growth law) give rise to a proteome allocation constraint. (A) Constancy of protein density in *Escherichia coli*. The regression equation is $y = 0.28x - 0.18$. Adapted from [55]. (B) Approximately constant biomass (a proxy for protein) density across phylogenetically disparate cells. The regression equation is $y = 0.00058x^{0.92}$. Adapted from [15]. (C) Linear scaling of ribosome concentration with growth rate in different prokaryotic and eukaryotic cells. Adapted from [18]. (D) Resource allocation into three main proteome sectors in fast- and slower-growing cells. Proteome sector Φ_Q is constant, Φ_M is for biosynthetic enzymes, and Φ_R is for ribosome and associated translation factors (see [18]).

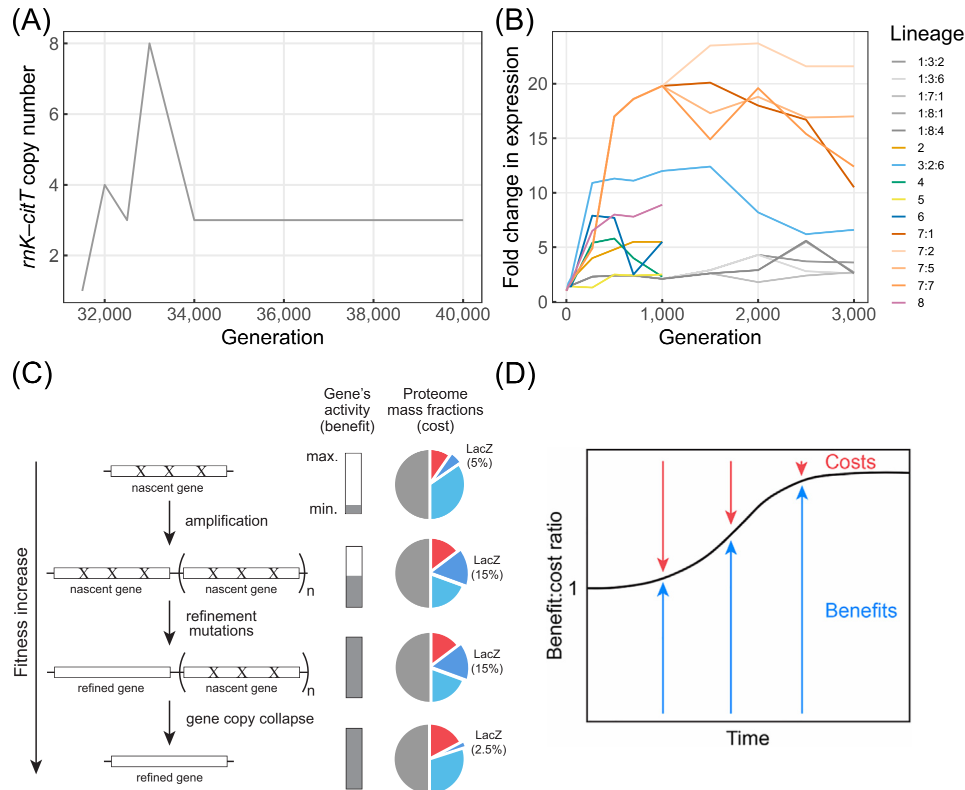
Another physiological constraint is that in which growth rate is proximally determined by the concentration of ribosomes in the cell. In the bacterium *Escherichia coli*, the number or concentration of ribosomes increases proportionally (linearly) with growth rate as the bacterium is grown in progressively nutrient-richer media (i.e., the **growth law**) [18] (Figure 1C). The same holds true for other heterotrophic prokaryotes and eukaryotes [18,19] (Figure 1C), including the yeast *Saccharomyces cerevisiae* [20], and even for photoautotrophic prokaryotes [19] and eukaryotes [21] (e.g., *Euglena gracilis*; Figure 1C). Such a correlation emerges because, as translation rate is approximately constant across cells, the only way to increase the rate of protein synthesis is by increasing the number of ribosomes. Thus, if a cell grows twice as fast, it needs to double its ribosome number.

Given that protein density is constant (protein density constraint), and that growth rate is directly determined by the concentration of ribosomes (growth rate–ribosome linear correlation constraint), a third, resource allocation, constraint emerges that imposes a limit on the non-ribosomal protein fraction of the cell [22]. There is thus a 'tug of war' within the proteome among sectors of functionally related proteins that are similarly regulated. As one proteome sector increases in size, the other necessarily decreases (Figure 1D). This immediately suggests that the expansion of a proteome sector, for example, by additions of newly expressed proteins, does not directly impose higher energetic demands but reallocates or redistributes resources (a fraction of the energy budget) from one proteome sector to another. This **proteome allocation** framework has been successfully used to explain whole-cell behavior or system-level properties such as overflow metabolism (also called the Crabtree effect in yeast), carbon catabolite repression (CCR), and the diauxic shift in both *E. coli* and *S. cerevisiae* [22–24].

The origin of new functional genes and their energetic costs

The origin of new functions has been observed in experimentally evolved populations. The most iconic, and perhaps best-studied, example is that of the evolution of the capacity to aerobically use citrate as a carbon and energy source by *E. coli* in the Long-Term Evolution Experiment (LTEE; [25]). The evolution of this new function has been divided into three stages: potentiation, which makes the new function possible; actualization, which creates the new function; and refinement, which improves the newly evolved function. In one of the 12 LTEE populations, the actualization of the new function occurred after >31 500 generations (~14 years) when the gene for an anaerobically expressed citrate transporter (citrate–succinate antiporter; *citT*) was placed under the control of the *rnk* promoter that allows for aerobic expression. This new gene arrangement provided a slight advantage in the form of an extremely weak Cit^+ phenotype or growth on citrate. This weak function was subsequently improved or refined by up to eight tandem duplications of the *rnk–citT* hybrid gene (Figure 2A). Further refinement of this new function included promoter mutations and amplification of the *dctA* gene, a succinate transporter, that allows for the recovery of the succinate exported in exchange for citrate by CitT [26]. A decrease in the *rnk–citT* gene copy number was observed at generation 34 000 (Figure 2A).

A similar refinement pattern of early **gene amplification** has been documented in the experimental test for the Innovation-Amplification-Divergence (IAD) model for the evolution of new genes [27]. In this study, a bifunctional gene with weak activities (low catalytic rates) for both histidine and tryptophan biosynthesis (HisA and TrpF activities, respectively) was transformed into a *Salmonella enterica* strain lacking the *hisA* and *trpF* genes. Within a few hundred generations of growth in medium without both amino acids, doubling time was decreased by half (from 5 to 1.9–2.5 h) by amplification of the weakly bifunctional gene to up to 20 copies in some lineages (Figure 2B). After ~3000 generations, doubling time was further reduced by beneficial mutations in some of the gene copies that led to the divergence of the duplicates/paralogs into specialized genes with



Trends in Microbiology

Figure 2. The early evolutionary history of new genes is accompanied by high relative energetic costs and low functional performance prior to evolutionary refinement. (A) Predicted *rnk–citT* module copy number in population 12 of the Long-Term Evolution Experiment (LTEE) across ~9000 generations after the inception of the Cit^{+} phenotype. Adapted from [25]. (B) Fold change in fluorescent signal of the bifunctional enzyme–YFP fusion protein across 3000 generations in the Innovation-Amplification-Divergence (IAD) evolution experiment [27]. Both (A) and (B) are consistent with early amplification and subsequent gene copy number collapse after evolutionary refinement. (C) Early gene amplification compensates for poor functional performance (e.g., low enzymatic activity) and this incurs higher associated proteome costs. (D) The evolutionary trajectory of new genes may be described by a progressive increase in functional benefits and a decrease in energetic costs under a selective regime where fitness improves.

either higher HisA or TrpF activities. A subsequent collapse into lower gene copy number was observed in most lineages (Figure 2B).

The examples above demonstrate that new genes often arise with weak or inefficient functions that need to be overexpressed to provide a strong enough selective advantage (Figure 2C). This is analogous to the inhibition of enzymatic activity which leads to enzyme overproduction to compensate for the decreased activity (e.g., translational inhibition by antibiotics leads to higher ribosome concentration; [18]). This overexpression results from, and is evidenced by, gene amplification early in the evolutionary history of new functions. Gene amplification constitutes one of the most mutationally accessible ways to improve function as duplication rates (e.g., 10^{-5}–10^{-2} per cell per generation) are much higher than those of beneficial point mutations (e.g., ~10^{-9} per cell per generation) [28,29]. Furthermore, gene amplification not only increases gene dosage but also provides a larger mutational target for beneficial mutations and a buffering capacity against harmful mutations. Several other examples have been documented in which gene amplification sustains early adaptation to environmental or intracellular selective pressures [28,30,31].

The overexpression of a weak function, however, comes at a cost. Because of proteome allocation constraints, the overexpression of a weakly functional gene deviates resources from other cellular functions. This is best illustrated by the effect that the synthetic overexpression of functionally neutral **'unnecessary proteins'** or 'unused proteins' has on growth rate [18,32]. For example, the overexpression of β-galactosidase in the absence of environmental lactose leads to a linearly proportional decrease in growth rate as a consequence of a reduced proteome fraction allocated to protein synthesis (i.e., ribosomes and associated proteins). Full induction of the *lac* operon (<0.5% of the proteome; [33]) in the absence of its conferred advantage reduces growth rate by ~4.5% [34]. This suggests that, although the new function provides a selective advantage (e.g., increased survival), its high expression level decreases growth capacity by deviating resources away from protein synthesis. Naturally, for any new function evolving under a selective regime, the fitness gain has to offset the associated energetic costs [35] (Figure 2D). After the refinement or optimization of a new function, net fitness can increase further by decreasing the energetic costs that accompany its early higher expression level (Figure 2D). This is evidenced by the collapse of genetically unstable tandem gene duplications in the examples discussed above [25,27] (Figure 2D).

The cost of new symbiotic organelles

The logic laid out above for the early evolution of new functional genes can also be applied to the macroevolution of more complex cellular features such as symbiotic organelles whose evolutionary histories are characterized by increased functional specialization. The **evolutionary trajectory** of mitochondria provides an example of increased specialization toward aerobic respiration. The ancestor of mitochondria is estimated to have had ~4000 genes, similar to *E. coli*. By contrast, modern mitochondrial genomes only have three to 67 protein-coding genes that are primarily involved in aerobic respiration (oxidative phosphorylation) and protein synthesis. At the proteome level, the largest fraction of proteins in yeast mitochondria, during non-fermentative growth on ethanol, is devoted to aerobic respiration (~33%), followed by transporters (~14%), amino acid biosynthesis (~7.5%), chaperones and proteases (~7.5%), and genome maintenance (~6.25); ribosomes constituted only ~2.0% of the mitochondrial proteome [36]. Conversely, the proteome fraction allocated to aerobic respiration in the bacterium *E. coli*, at a growth rate of ~0.75 h^{-1} (prior to the switch to overflow metabolism), is ~10% [23]. These numbers show that throughout the evolutionary specialization of mitochondria, there has been a proteome redistribution where a larger fraction has been allocated to aerobic respiration.

At the origin of symbiotic organelles, the selective advantage is arguably relatively small, and the energetic costs are high. A nascent symbiosis can be conceived of as being relatively inefficient because the newly acquired symbiont has a diverse proteome (and large gene repertoire) that allocates a large fraction to many different functions that contribute little to nothing to the main function the symbiont is being selected for. Thus, a larger number of symbionts or symbiont volume is initially required to provide a considerable benefit. As the symbiont loses genes that become unnecessary in the new nutrient-rich intracellular environment (through both random drift and selection for economy), its proteome is simplified and becomes progressively more specialized. This, together with new functional adaptations (e.g., crista developmental regulation or larger respiratory supercomplexes), increases the benefit that the symbiont provides to its host. An increased efficiency or functional performance subsequently allows for a decrease in symbiont number or volume, and this reduces energetic costs – the total volume taken up by symbionts is a proxy for their energetic cost [6]. Similar evolutionary trajectories are conceivable for other complex cellular features whose histories are characterized by increased functional specialization.

The macroevolutionary consequences of cellular complexity

The final proteome footprint of an evolutionarily refined functional gene is in most cases negligible. However, the macroevolutionary accruement of hundreds or thousands of novel genes during a major evolutionary transition, such as the origin of eukaryotes, is expected to take up a sizeable fraction of the proteome. The constitutive expression of much of the proteomic complexity of eukaryotes can be conceived of as expanding the proteome fraction that does not directly contribute to growth rate (often referred to as proteome sector Q; see [18]). One may thus predict that a drastic increase in **proteome complexity** correspondingly results in a marked decrease of the proteome fraction allocated to ribosomes. This prediction appears to be supported by the proteome allocation patterns (**proteomaps**; [37]) observed in the few model species for which wide-coverage mass-spectrometry data are available. In *E. coli*, for example, 33.5% of the proteome is devoted to the translation machinery (i.e., ribosomes, tRNA synthetases, initiation and elongation factors, etc.) (Figure 3A). In *S. cerevisiae* and *Homo sapiens* HeLa cells, the translation machinery represents 24.4% and 8.1% of the proteome, respectively (Figure 3B,C). An analogous pattern is observed when heterotrophs and photoautotrophs are compared, as the latter constitutively express

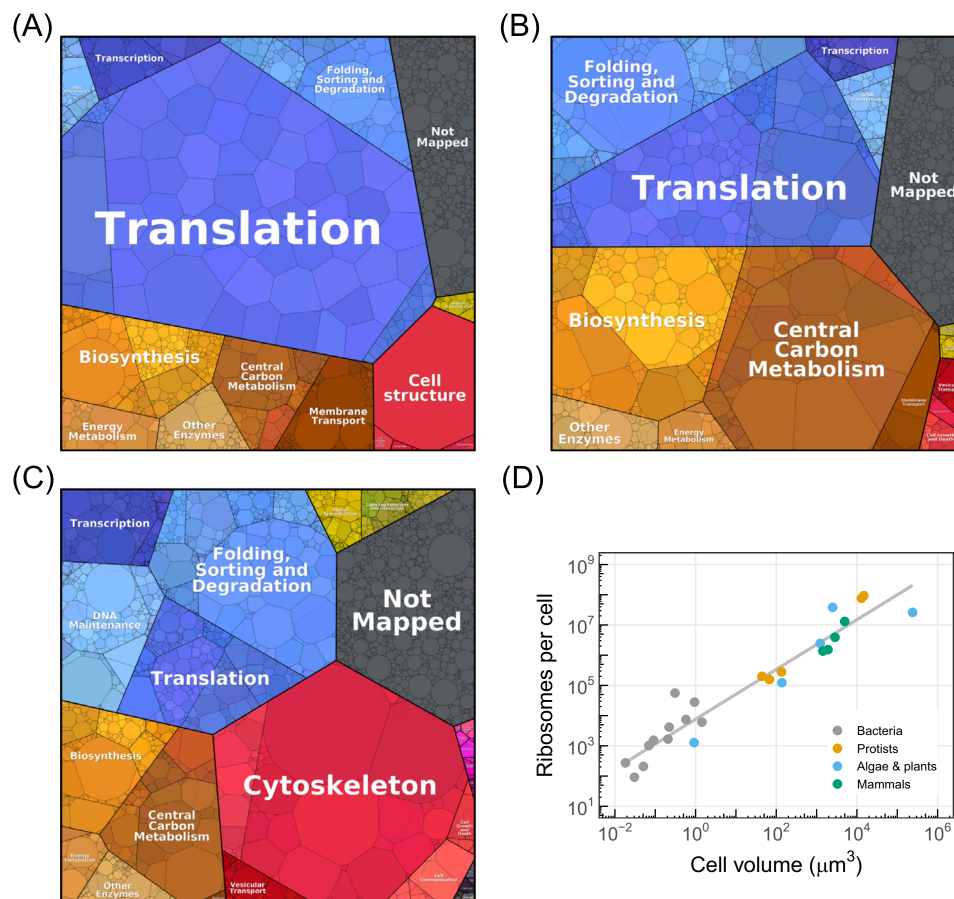


Figure 3. Cells with more complex proteomes allocate a smaller proteome mass fraction to the translation machinery. (A) A Voronoi treemap of the quantitative proteome (i.e., proteomap; [37]) of the bacterium *Escherichia coli* under complete medium with amino acids. (B) Proteomap of *Saccharomyces cerevisiae* grown in rich medium. (C). Proteomap of a *Homo sapiens* HeLa cell. (D). The sublinear scaling of ribosome number per cell with cell volume across prokaryotic and eukaryotic cells. The regression equation is $y = 7634.45x^{0.82}$. Adapted from [42].

a much more complex metabolic proteome; a large proteome investment into the photosynthetic machinery considerably reduces the amount of proteome resources that can be devoted to translation (i.e., 33.5% in *E. coli* vs. 10.4% in *Synechocystis* sp. 6803; see https://www.proteomaps.net). Conversely, a decrease in proteome complexity is expected to release proteome resources for translation. In agreement with this, the deletion of non-essential genes, such as those for flagellar proteins [38] or the biosynthesis of amino acids, vitamins, or nucleobases in the presence of external supplementation [39], leads to increases in growth rate and yield [40]. The examples above for *E. coli*, *S. cerevisiae*, and *H. sapiens* are consistent with the sublinear scaling of ribosome number with cell volume across prokaryotic and eukaryotic cells [41,42] (Figure 3D), which implies that larger cells devote a smaller proteome fraction to ribosomes – across such vast phylogenetic distances, cell volume roughly correlates with cellular complexity. Such a reduction in ribosome concentration ultimately reduces reproductive capacity and growth rate.

An evolutionary increase in cellular complexity is also associated with 'maintenance infrastructure'. This is most clearly illustrated by symbiotic organelles. As argued in the previous section, mitochondria specialized in ATP synthesis by aerobic respiration from a functionally highly versatile bacterial ancestor. Unlike prokaryotes, which place their respiratory chain at the cytoplasmic membrane, eukaryotes place it at the mitochondrial inner membrane. Such internalized respiratory membranes arguably provide macroevolutionary advantages by allowing the proportional increase of respiration with larger cell volumes without considerably compromising growth rates [6]. However, they also come with additional cellular complexity required for their maintenance, which is otherwise not present in a respiring prokaryote. This encompasses an additional outer membrane (of endosymbiotic origin), novel transporters (which represent up to 14% of the mitochondrial proteome [36]), protein-import machinery, dedicated genomes and associated proteins, etc. In addition, the use of oxygen as a terminal electron acceptor allows for a more efficient energy metabolism but also requires the expression of an increasing amount of oxygen-detoxifying enzymes (e.g., superoxide dismutase, catalase) to neutralize oxygen radicals in cells constantly exposed to oxygen [43]. This maintenance infrastructure deviates additional resources from reproduction, and thus ultimately lowers growth rate.

Further evidence for the notion that 'complexity begets complexity' is seen in the increasingly larger proteome fraction that more complex cells devote to chaperones and co-chaperones. Whereas prokaryotes allocate ~1–2%, eukaryotes devote ~3–10% of their proteomes to these folding-assisting proteins [44,45]. The increase in abundance of chaperones and co-chaperones is likely an adaptive evolutionary response to both adaptive and non-adaptive causes. For example, an adaptive increase in proteome complexity (measured as the number of different proteins, folds, or fold combinations [44]) demands more chaperones and new types of co-chaperones to assist the folding of new types of proteins, and non-adaptive protein divergence (due to a decreased power of natural selection in more complex and larger cells with smaller effective population sizes) may similarly require more (co-)chaperones to buffer passively accumulated deleterious mutations (e.g., as seen in intracellular symbionts; [46]). It is unlikely, however, that a constant proteome fraction of chaperones would have been able to cope with the drastic increase in complexity at the origin of eukaryotes, some of which is undeniably adaptive. The higher cellular complexity of eukaryotes, and associated maintenance infrastructure, arguably also contribute to the larger amounts of energy that eukaryotes allocate to maintenance processes, and ultimately to sublinear scaling of growth rate with cell volume observed in eukaryotes [47].

Concluding remarks

The relationship between energy and the origins of cellular complexity has been the subject of much disagreement in the field of cell evolution [4,6–8,10,11,48]. Some progress was achieved by closer

Outstanding questions

What are the proteome allocation programs of slow-growing cells in nutrient-depleted environments?

How does the linear relationship between ribosome concentration and growth rate (i.e., the growth law) vary across the tree of life?

What functional categories of genes incur higher energetic costs at their origin?

How does cellular complexity increase in cells that do not optimize proteome allocation?

inspection of the phylogenetic scaling of the energy requirements of cells across phylogenetic boundaries [12,42]. Here, I drew observations from experimental evolution, systems biology, microbial physiology, and energetics in an attempt to add conceptual clarity to some of the long-standing issues. Specifically, a resource allocation framework allows us to re-interpret some of the ideas previously posed and deemed controversial. I thus suggest that cellular complexity imposes energetic costs both at micro- and macro-evolutionary timescales. In the short-term, new genes with weak functions impose higher energetic demands as they divest a larger amount of proteome resources prior to evolutionary refinement. This may constitute a general route by which both new functional genes and more complex cellular features evolve. In the long-term, the accumulation of thousands of new genes inevitably takes up a sizeable fraction of the proteome that constrains the number of resources that can be allocated to growth and reproduction. New genes and cellular features thus do not impose higher absolute energetic demands but deviate resources away from pre-existing cellular processes. This notion is consistent with, and helps to explain, the phylogenetic scaling laws for the decrease of ribosome concentration and productivity across major cell size and complexity ranges [47]. These observations and arguments further reinforce the view that eukaryogenesis (i.e., a substantial increase in cellular complexity) moved eukaryotes into a fundamentally distinct adaptive zone where cellular resources are primarily channeled toward survival over reproduction (i.e., k-strategies dominate) [49]. Although these ideas are yet to be rigorously tested experimentally and comparatively, they promise to stimulate new research avenues at the intersection of systems biology and cell evolution (see Outstanding questions).

Acknowledgments

I thank Paul Schavemaker (Arizona State University) for discussions on cell evolution and energetics that stimulated some of the ideas developed in this essay.

Declaration of interests

No interests are declared.

References

1. Stanier, R.Y. *et al.* (1963) *The Microbial World* (2nd edn), Prentice-Hall
2. Cavalier-Smith, T. (2014) The Neomuran Revolution and phagotrophic origin of eukaryotes and cilia in the light of intracellular coevolution and a revised Tree of Life. *Cold Spring Harb. Perspect. Biol.* 6, a016006
3. Vosseberg, J. *et al.* (2021) Timing the origin of eukaryotic cellular complexity with ancient duplications. *Nat. Ecol. Evol.* 5, 92–100
4. Lane, N. and Martin, W. (2010) The energetics of genome complexity. *Nature* 467, 929–934
5. Lane, N. (2020) How energy flow shapes cell evolution. *Curr. Biol.* 30, R471–R476
6. Schavemaker, P.E. and Muñoz-Gómez, S.A. (2022) The role of mitochondrial energetics in the origin and diversification of eukaryotes. *Nat. Ecol. Evol.* 6, 1307–1317
7. Lane, N. (2011) Energetics and genetics across the prokaryote-eukaryote divide. *Biol. Direct* 6, 35
8. Lane, N. and Martin, W.F. (2015) Eukaryotes really are special, and mitochondria are why. *Proc. Natl. Acad. Sci. U. S. A.* 112, E4823
9. Muñoz-Gómez, S.A. (2023) Energetics and evolution of anaerobic microbial eukaryotes. *Nat. Microbiol.* 8, 197–203
10. Booth, A. and Doolittle, W.F. (2015) Eukaryogenesis, how special really? *Proc. Natl. Acad. Sci. U. S. A.* 112, 10278–10285
11. Szathmáry, E. (2015) Toward major evolutionary transitions theory 2.0. *Proc. Natl. Acad. Sci. U. S. A.* 112, 10104–10111
12. Lynch, M. and Marinov, G.K. (2015) The bioenergetic costs of a gene. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15690–15695
13. Hatton, I.A. *et al.* (2019) Linking scaling laws across eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 116, 21616–21622
14. Chiyomaru, K. and Takemoto, K. (2020) Revisiting the hypothesis of an energetic barrier to genome complexity between eukaryotes and prokaryotes. *R. Soc. Open Sci.* 7, 191859
15. Lynch, M.R. (2024) *Evolutionary Cell Biology: The Origins of Cellular Architecture*, Oxford University Press
16. Makarieva, A.M. *et al.* (2005) Energetics of the smallest: do bacteria breathe at the same rate as whales? *Proc. Biol. Sci.* 272, 2219–2224
17. Makarieva, A.M. *et al.* (2008) Mean mass-specific metabolic rates are strikingly similar across life's major domains: evidence for life's metabolic optimum. *Proc. Natl. Acad. Sci. U. S. A.* 105, 16994–16999
18. Scott, M. *et al.* (2010) Interdependence of cell growth and gene expression: origins and consequences. *Science* 330, 1099–1102
19. Jahn, M. *et al.* (2018) Growth of cyanobacteria is constrained by the abundance of light and carbon assimilation proteins. *Cell Rep.* 25, 478–486.e8
20. Metzl-Raz, E. *et al.* (2017) Principles of cellular resource allocation revealed by condition-dependent proteome profiling. *eLife* 6, e28034
21. Cook, J.R. (1963) Adaptations in growth and division in euglena effected by energy supply*. *J. Protozool.* 10, 436–444
22. Scott, M. and Hwa, T. (2023) Shaping bacterial gene expression by physiological and proteome allocation constraints. *Nat. Rev. Microbiol.* 21, 327–342
23. Basan, M. *et al.* (2015) Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature* 528, 99–104
24. Erickson, D.W. *et al.* (2017) A global resource allocation strategy governs growth transition kinetics of *Escherichia coli*. *Nature* 551, 119–123
25. Blount, Z.D. *et al.* (2012) Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* 489, 513–518
26. Quandt, E.M. *et al.* (2014) Recursive genomewide recombination and sequencing reveals a key refinement step in the evolution of a metabolic innovation in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 111, 2217–2222

27. Näsvall, J. *et al.* (2012) Real-time evolution of new genes by innovation, amplification, and divergence. *Science* 338, 384–387
28. Andersson, D.I. and Hughes, D. (2009) Gene amplification and adaptive evolution in bacteria. *Annu. Rev. Genet.* 43, 167–195
29. Katju, V. and Bergthorsson, U. (2013) Copy-number changes in evolution: rates, fitness effects and adaptive significance. *Front. Genet.* 4, 273
30. Andersson, D.I. *et al.* (1998) Evidence that gene amplification underlies adaptive mutability of the bacterial lac operon. *Science* 282, 1133–1135
31. Sandegren, L. and Andersson, D.I. (2009) Bacterial gene amplification: implications for the evolution of antibiotic resistance. *Nat. Rev. Microbiol.* 7, 578–588
32. Bienick, M.S. *et al.* (2014) The interrelationship between promoter strength, gene expression, and growth rate. *PLoS One* 9, e109105
33. Scott, M. and Hwa, T. (2011) Bacterial growth laws and their applications. *Curr. Opin. Biotechnol.* 22, 559–565
34. Dekel, E. and Alon, U. (2005) Optimality and evolutionary tuning of the expression level of a protein. *Nature* 436, 588–592
35. Lynch, M. and Trickovic, B. (2020) A theoretical framework for evolutionary cell biology. *J. Mol. Biol.* 432, 1861–1879
36. Di Bartolomeo, F. *et al.* (2020) Absolute yeast mitochondrial proteome quantification reveals trade-off between biosynthesis and energy generation during diauxic shift. *Proc. Natl. Acad. Sci. U. S. A.* 117, 7524–7535
37. Liebermeister, W. *et al.* (2014) Visual account of protein investment in cellular functions. *Proc. Natl. Acad. Sci. U. S. A.* 111, 8488–8493
38. Muntel, J. *et al.* (2014) Comprehensive absolute quantification of the cytosolic proteome of *Bacillus subtilis* by data independent, parallel fragmentation in liquid chromatography/mass spectrometry (LC/MSE). *Mol. Cell. Proteomics* 13, 1008–1019
39. D'Souza, G. *et al.* (2014) Less is more: selective advantages can explain the prevalent loss of biosynthetic genes in bacteria. *Evolution* 68, 2559–2570
40. Valgepea, K. *et al.* (2015) Lean-proteome strains – next step in metabolic engineering. *Front. Bioeng. Biotechnol.* 3, 11
41. Lynch, M. and Marinov, G.K. (2017) Membranes, energetics, and evolution across the prokaryote-eukaryote divide. *eLife* 6, e20437
42. Lynch, M. and Marinov, G. (2018) Correction: membranes, energetics, and evolution across the prokaryote-eukaryote divide. *eLife* 7, e35006
43. Speijer, D. *et al.* (2020) Comparing early eukaryotic integration of mitochondria and chloroplasts in the light of internal ROS challenges: timing is of the essence. *mBio* 11, 1, e00955–20
44. Rebeaud, M.E. *et al.* (2021) On the evolution of chaperones and cochaperones and the expansion of proteomes across the Tree of Life. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2020885118
45. Finka, A. and Goloubinoff, P. (2013) Proteomic data from human cell cultures refine mechanisms of chaperone-mediated protein homeostasis. *Cell Stress Chaperones* 18, 591–605
46. Poliakov, A. *et al.* (2011) Large-scale label-free quantitative proteomics of the pea aphid-*Buchnera* symbiosis. *Mol. Cell. Proteomics* 10, M110.007039
47. Lynch, M. *et al.* (2022) Evolutionary scaling of maximum growth rate with organism size. *Sci. Rep.* 12, 22586
48. Hampl, V. *et al.* (2019) Was the mitochondrion necessary to start eukaryogenesis? *Trends Microbiol.* 27, 96–104
49. Carlile, M. (1982) Prokaryotes and eukaryotes: strategies and successes. *Trends Biochem. Sci.* 7, 128–130
50. Haldane, J.B.S. (1926) *On Being the Right Size*, Harper's Magazine
51. Bonner, J.T. (2006) *Why Size Matters: From Bacteria to Blue Whales*, Princeton University Press
52. Koonin, E.V. (2015) Energetics and population genetics at the root of eukaryotic cellular and genomic complexity. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15777–15778
53. Cavalier-Smith, T. (1985) *The Evolution of Genome Size*, Wiley
54. Lynch, M. and Conery, J.S. (2003) The origins of genome complexity. *Science* 302, 1401–1404
55. Basan, M. *et al.* (2015) Inflating bacterial cells by increased protein synthesis. *Mol. Syst. Biol.* 11, 836