Trends in Microbiology

CellPress
REVIEWS

Review

Archaeal Histone Contributions to the Origin of Eukaryotes

Clifford F. Brunk[1] and William F. Martin[2,*]

The eukaryotic lineage arose from bacterial and archaeal cells that underwent a symbiotic merger. At the origin of the eukaryote lineage, the bacterial partner contributed genes, metabolic energy, and the building blocks of the endomembrane system. What did the archaeal partner donate that made the eukaryotic experiment a success? The archaeal partner provided the potential for complex information processing. Archaeal histones were crucial in that regard by providing the basic functional unit with which eukaryotes organize DNA into nucleosomes, exert epigenetic control of gene expression, transcribe genes with CCAAT-box promoters, and a manifest cell cycle with condensed chromosomes. While mitochondrial energy lifted energetic constraints on eukaryotic protein production, histone-based chromatin organization paved the path to eukaryotic genome complexity, a critical hurdle en route to the evolution of complex cells.

Highlights

The last common ancestor of eukaryotes had mitochondria, pointing to host–symbiont interactions at eukaryote origin. Mitochondria contributed energy, genes, and membranes to the eukaryotic lineage. What did the archaeal host contribute?

Recent metagenomic studies propose that close relatives of the archaeal host exist that are 'complex' (phagocytosing) and that the archaeal host brought that complexity to eukaryotes.

Yet complex archaea have not been found, and there are doubts that the metagenomic archaeal data represent truly complex archaea.

Alternatively, histones could have been the key archaeal contribution to eukaryote complexity.

In eukaryotes, histone modifications link gene expression to the physiological state of the cell via carbon-, energy-, and nitrogen-sensing through regulators such as AMPK, GCN2, and TOR.

Eukaryotes from Prokaryotes: Symbiotic Contributions

The origin of eukaryotes remains one of evolution's more pressing unresolved questions; however, it is beginning to yield some of its secrets. Over the past several decades, perspectives on eukaryogenesis have undergone a significant transformation [1]. Mitochondria are now thought to have been present in the eukaryote common ancestor [2,3], and the broad outlines of eukaryote lineage origin have increasingly come to include the concept of symbiosis [4–7]. Both the host lineage and the mitochondrial endosymbiont contributed to the eukaryote origin.

The mitochondrial contribution to the origin of eukaryotes encompassed energy [8], genes [9–11], and lipids [12]. A hallmark of eukaryotic cells is their endomembrane system, which comprises the endoplasmic reticulum (ER), the nucleus, and vesicular membrane traffic. Although eukaryotes are typically seen as being descendant from archaea [13], they possess bacterial lipids. Prokaryotes synthesize lipids at the plasma membrane, eukaryotes synthesize lipids in the ER and in mitochondria [14]. Even the origin of the eukaryotic endomembrane system, which has been a longstanding puzzle in cell evolution, now connects to mitochondria [14–16]. Mitochondria secrete single-membrane-bounded vesicles in the cytosol called mitochondrial-derived vesicles (MDVs) [14–16]. They are homologous to the outer-membrane vesicles (OMVs) of Gram-negative bacteria [17]. OMVs produced by the mitochondrial endosymbiont at the eukaryote origin are a promising candidate for the biochemical and physical source of the endomembrane system. MDVs produced by the ancestral mitochondrion in the cytosol of an archaeal host generate a membrane topology identical to that of the ER [15]. They also provide an ancestrally outward-directed vectoral membrane flux to the host's plasma membrane and thus a simple mechanism by which the archaeal lipids of the host could be replaced by the bacterial lipids of the mitochondrial endosymbiont. With energy, genes, and membranes coming from mitochondria (Figure 1), what did the archaeal host contribute to the eukaryotic lineage?

[1]Department of Ecology and Evolutionary Biology and Molecular Biology Institute University of California Los Angeles, Los Angeles, USA
[2]Institute of Molecular Evolution Heinrich-Heine-Universitaet Duesseldorf, Dusseldorf, Germany

*Correspondence:
bill@hhu.de (W.F. Martin).

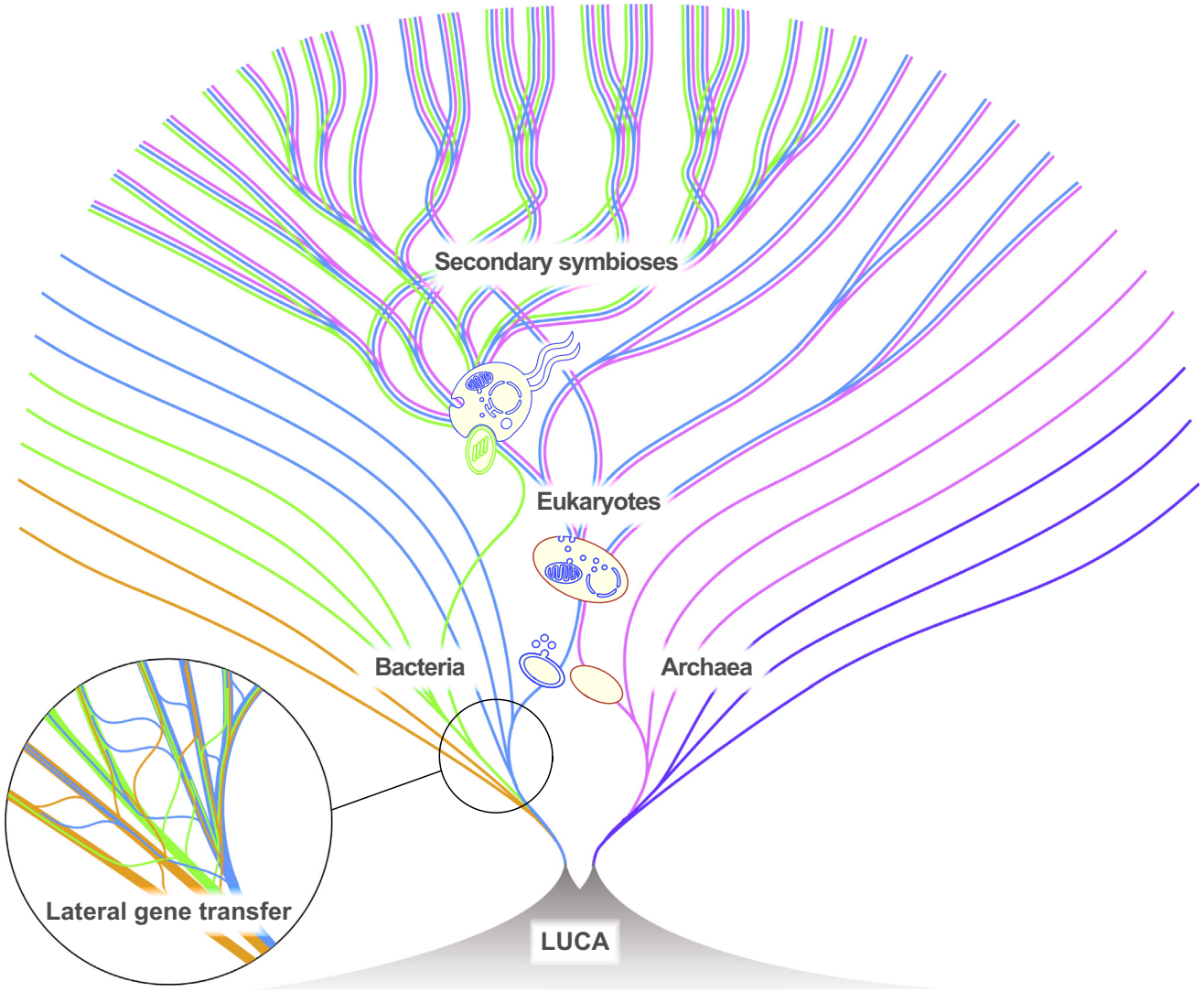

Figure 1. Schematic Diagram of Symbiosis in the Tree of Life and the 'Union' of the Bacterial and Archaeal Domains to Form the Eukaryotic Domain, Taking Lateral Gene Transfer (LGT) into Account, Redrawn from [106] and Earlier Renderings. Outer-membrane vesicles at the surface of the endosymbiont and their implicated role in the origin of bacterial lipids in eukaryotes and the eukaryote endomembrane system [12,14–16] are indicated. The physiological and metabolic context of host–symbiont interactions at eukaryote and mitochondrial origin has been reviewed elsewhere [29]. LUCA: last universal common ancestor. The figure is deliberately explicit about endosymbiosis being a mechanism at major transitions in eukaryote evolution. In the 50 years since the revival of endosymbiotic theory by Margulis (then named Sagan) [107], there has been no debate that symbiosis with a cyanobacterium gave rise to the plant lineage or that secondary symbioses gave rise to eukaryotic algae with plastids surrounded by three or four membranes [108], but there has always been resistance to the idea that either symbiosis or mitochondria gave rise to the eukaryotic lineage [31,32,109]. The time between the origin of mitochondria and the origin of plastids in the archaeplastidal lineage might be quite short [2]; the figure does not imply a specific time with regard to the relative timing of those events.

The Archaeal Contribution

Since their discovery, archaea have been implicated as relatives of the host lineage at the origin of eukaryogenesis and mitochondria [18,19]. There is now much discussion about the possibility that complex, phagocytosing archaea might be yet discovered in nature, based on metagenomic data [20–22]. In that view, the contribution of archaea is unequivocal: the archaeal host brought preformed complexity to the eukaryotic lineage. Indeed, the possibility that some archaea or archaeal relatives might possess a primitive cytoskeleton has been discussed for decades [23–26].

The prokaryotic precursors of actin and tubulin are present in many archaea as well as in bacteria [27,28]. So why did prokaryotes never utilize that starting material to evolve eukaryotic cellular complexity? From the bioenergetic standpoint, the simplest answer is that mitochondrial power was required for the evolution of eukaryotic complexity [8,29,30] and that prokaryotes lack true complexity because they lack mitochondria.

However, some researchers doubt that mitochondria had significance for eukaryote origin, arguing that the presence of mitochondria in the eukaryote common ancestor is pure coincidence, with phagotrophy (eating other cells as a form of obtaining carbon and energy) being the key innovation that allowed eukaryotes to become complex [31,32]. The insurmountable problem with that view, however, is that phagotrophy only provides physiological benefit to cells that already possess mitochondria (reviewed in [33]), which is very likely the reason why no phagocytosing prokaryotes have ever been found.

Doubts of a different nature are coming to bear upon the issue, however. That is, some researchers now doubt that the metagenomic data currently being interpreted as indicating complex archaea really reflect the existence of complex archaea [14,30,33–38]. We know that eukaryotes are complex because we can observe their complexity both under the microscope and in everyday life. The complex archaea about which much is being written [20–22] have yet to be seen. Until direct evidence comes forth for archaea with eukaryotic complexity, it remains possible, if not probable, that the archaeal contribution to the symbiotic merger was something other than preformed complexity.

For the sake of reasoning, let us assume for a moment that doubts about the existence of complex archaea [14,30,33–38] are justified. We would then ask the following question. What was the archaeal trait that made the archaeal contribution to the symbiosis a success? We have known for decades that archaea contributed the information-processing system, or informational genes, to the eukaryotic lineage [9,10,39,40]. The genes in the eukaryotic lineage that have remained most conserved in function relative to archaeal counterparts are involved in information processing: the RNA polymerase [4,41], the ribosome [42,43], and aminoacyl tRNA synthetases [44]. But the machinery of information processing itself does not set eukaryotes apart from prokaryotes. The information-processing machineries of eukaryotes and archaea are largely the same. Rather it is the amount of information that eukaryotes maintain, process, and inherit that sets them apart from prokaryotes.

Eukaryotic DNA Content

Although the increase in metabolic capacity of eukaryotic cells greatly exceeds that of their prokaryotic progenitors, the increase in DNA content of eukaryotic cells over their prokaryotic progenitors is even greater. A slight historical diversion into DNA content may be helpful in appreciating a critical capacity afforded by an increased amount of DNA. When the study of molecular biology commenced, bacteria and their viruses were the primary topics of investigation. The DNA content of these organisms seemed rational, the more complex the organism the larger the amount of DNA [45]. As eukaryotic cells came under scrutiny, this rational system seemed to break down, by and large eukaryotic cells had a great deal more DNA than expected [46]. Organisms in the same species had identical amounts of DNA, but very similar species often had vastly different amounts of DNA [47,48]. This phenomenon was so prevalent that it became known as the ‘C value paradox’, C value being the haploid amount of DNA associated with a species [49]. Eukaryotic cells have an inordinately large amount of DNA relative to prokaryotic cells, and the differences in DNA content across eukaryotes was not correlated with increased complexity, sophistication, or even gene number [45,50].

The differences between prokaryotic and eukaryotic genome size are accounted for primarily by mobile DNA because it can create new copies of itself, but without a positive contribution to the phenotype of the organism [51]. Mobile DNA is also called 'selfish DNA' [52] and makes up a substantial portion of the eukaryotic genome, well over 90% in *Homo sapiens* [53]. The presence of selfish DNA in the genomes of eukaryotes clearly indicates that the constraint on the amount of DNA in prokaryotic cells is not a limitation in eukaryotic cells, which can exhibit a super abundance of DNA. The ability to accommodate a vast amount of DNA produced by mobile DNA may have led to the origin of introns in the eukaryotic lineage and selective pressures that required the evolution of a nucleus in the eukaryote common ancestor [54]. The necessity of intron excision likely precipitated the decoupling of transcription from translation in the eukaryote ancestor, a major feature distinguishing eukaryotes from their prokaryotic predecessors.

The amount of DNA found in eukaryotes is so vastly greater than in prokaryotes that there is a minimal overlap of the distributions (Figure 2). Eukaryotes have DNA contents with a range spanning over 200 000-fold [48], while the range of expressed genes (proteomes) is much more limited, probably spanning about 50-fold. Early in the development of molecular biology, the replication of DNA was viewed as a major expense in the cell's energy economy and thus the expansion of DNA in eukaryotes was thought to be very expensive. A more detailed analysis of cellular energy expenditures indicates that the synthesis of proteins is far and away the major consumer of cellular energy, about 75% of the cell's energy budget, while DNA synthesis is relatively inexpensive, about 3% of the energy budget [8,55]. The limitation on prokaryotic genome size is related to the ability to manipulate and control the expression of large amounts of DNA, rather than the expense of synthesizing the DNA. Why are eukaryotic cells capable of manipulating vastly larger amounts of DNA than their prokaryotic progenitors? The answer would appear to lie in the organization of the DNA into nucleosomes.

Archaeal Histones

It would be an oversimplification to view the DNA in a cell independently of proteins bound to the DNA. All organisms have proteins bound to their DNA in order to prevent collapse of DNA at high

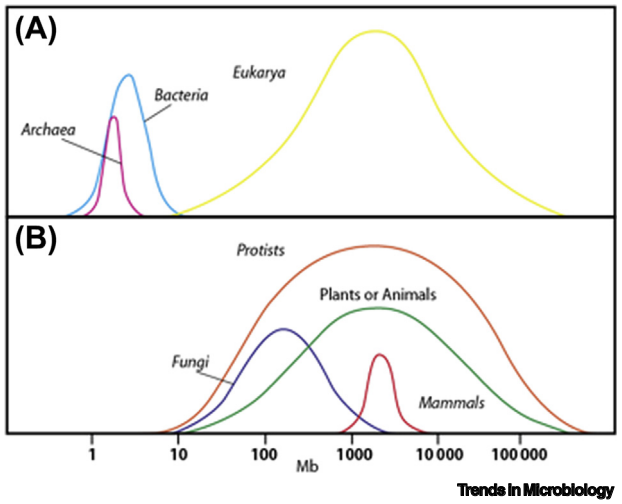

Figure 2. The Ranges of Genome Sizes for Various Groups. The ranges of genome sizes are displayed on a logarithmic scale. The vertical dimension indicates, but is not strictly proportional to, the number of taxa. The specific shape of the distributions is not the main point here, the range is accurately portrayed. (A) The comparative ranges of genome sizes of archaea, bacteria, and eukarya. (B) The ranges of genome sizes for various groups of eukaryotes. The range of genome sizes for plants is somewhat greater than for animals, but this difference is not evident on a logarithmic scale (www.genomesize.com) [50].

concentration [56]. In bacteria there are a number of DNA-binding proteins that fulfill this role [57]. Archaea in particular though, have several proteins that fulfill this role, including proteins with the histone-fold, which are widespread among the euryarchaea and nanoarchaea [58,59]. In these archaea, DNA binds to histone-fold proteins, which form homodimers and appear to produce extended fibers, as recent crystal structures for archaeal histones reveal [60]. The diversity of archaeal histones is not fully charted [61] but neither eukaryotic type nucleosomes nor eukaryotic type histone modifications have been found in archaea so far. Nuclear eukaryotic DNA, in contrast, is organized into nucleosomes, slightly less than 150 base pairs of DNA wrapped around an octamer histone core consisting of two H2A/H2B dimers and a H3/H4 tetramer [62]. The filament nature of archaeal histone polymerization contrasts to the 'beads on a string' nature of eukaryotic histones (Figure 3).

Histones make up a major portion of a nucleosome, ~ 55% of the mass, and their special orientation creates the nucleosome core. The nucleosomes are connected to each other by short 'linker' segments of DNA and comprise the primary structure of chromatin [62,63]. Organization of DNA into nucleosomes leads to a compaction of 30- to 40-fold [64].

Recent advances in chromatin imaging reveal 5 nm and 24–30 nm fibers of chromatin organization in living cells [65], underscoring the central role of nucleosomes in the primary organization of eukaryotic DNA. The nucleosome organization appears to permit a high degree of flexibility and removes a major constraint on manipulating large amounts of DNA [65,66]. Compaction and flexibility permit large amounts of DNA to be manipulated and the cellular genome size is free to expand, almost without limit [64] (Figure 3).

Histones Control Gene Expression

A dramatic increase in the DNA content of cells requires not only improved mechanisms for replication and maintenance but also enhanced control of gene expression [67,68]. The evolution and development of a nucleosome-based chromatin structure in eukaryotes addresses both of these requirements. Initially, nucleosomes were viewed as static structural elements facilitating the manipulation of the DNA; the role of histones in organizing eukaryotic DNA into nucleosomes provides a basis for enhanced control of gene expression. Having a greatly expanded DNA content dramatically increases the necessity for control of genetic expression in eukaryotes well beyond the prokaryotic mechanisms for regulation of gene expression. Within the nucleosome structure, particularly in the polypeptide extensions – the N terminal and C terminal 'tails' – beyond the core histone-fold protein, histone modification provides a powerful element for the control of eukaryotic gene expression [67–72].

Central to the role of histones as a mechanism for the control of gene expression are histone acetyltransferases (HATs) [70,71]. Histone modification plays the central role in the massively expanded epigenetic control of gene expression in eukaryotes relative to prokaryotes. The 'histone barcode' concept has it that specific histone modification, or combinations thereof, can affect distinct downstream cellular events by altering the structure of chromatin or by generating a binding platform for effector proteins [68,70,73]. The role of covalent modification of histones in the regulation of gene expression has become a central theme in eukaryote gene regulation [74].

The central portion of the eukaryotic histones, dominated by the histone-fold, is the core of the nucleosome; however, extensions of this core – particularly at the amino terminus – provide the substrate for modifications that control gene expression [75]. Post-translational acetylation of histone termini, particularly of lysine residues, as well as phosphorylation and methylation, dramatically affects their charge and function. Histone modification affords the elaborate control of gene

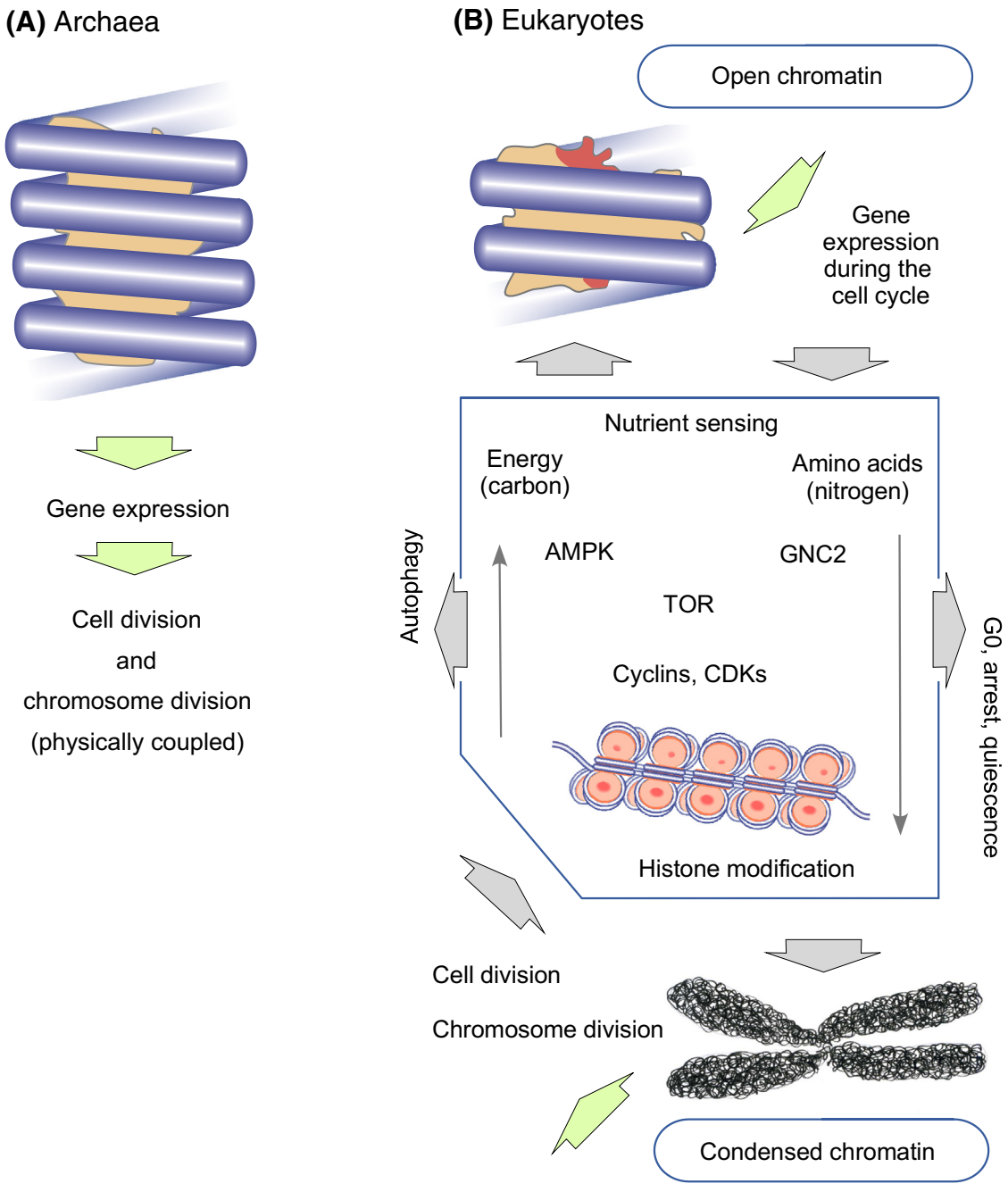

Figure 3. Histones in Archaea and Eukaryotes. (A) The crystal structure of DNA bound to archaeal histones was recently determined to 4 Å for *Methanothermus fervidus* [60]. The schematic representation of the structure shown here indicates DNA wrapped four times around the central protein superhelix of nine archaeal histone dimers. The archaeal structure forms a fiber. (B) Eukaryotic histones form the familiar nucleosome structure that is the basic structural and functional unit of eukaryotic chromatin. The red surfaces indicate the N terminal (top) and C terminal extensions of eukaryotic histones. These extensions contain the main sites of histone modification that modulate chromatin condensation/decondensation and that govern gene expression. These functional extensions are crucial to eukaryotic nucleosome organization, but are lacking in archaeal histones [60], which do not form nucleosomes [60,61]. In contrast to archaea, eukaryotes have a bona fide cyclin-dependent cell cycle in which metaphase chromosomes condense for chromosome segregation and cell division (lower panel). Regulation of the cell cycle and decisions to exit into G0, arrest, or autophagy involve histone modifications and are governed by nutrient sensing (Box 1). Note that, in eukaryotes with a syncytial (coenocytic) lifestyle, chromosome division and cell division are not coupled [94], in contrast to archaea and bacteria. AMPK, AMP-activated protein kinase; CDKs, cyclin-dependent kinases; TOR, target of rapamycin.

Box 1. Eukaryotic Histones: An Archaeal Tool Put to Greater Use

Mattiroli *et al*. [60] recently reported the structure of histone-bound DNA from archaea. As sketched in Figure 3A, it has more the form of a continuous cylinder than the familiar beads-on-a-string model found in eukaryotes (Figure 3B). Archaeal histone mutants that cannot polymerize show effects on gene expression [60], but the classical condensation of DNA into chromatin of the type seen in eukaryotic metaphase chromosomes has not been observed in archaea. The extensive role of histone modification – acetylation, methylation, phosphorylation – in chromatin organization and gene expression typical of eukaryotes [83,85,102] is also so far unknown for archaea. One of the most fundamental differences between prokaryotes and eukaryotes is that eukaryotes condense their DNA for microtubule-dependent chromosome separation at every cell division [94] as an integral component of the cell cycle. This condensation is histone-dependent and is dependent upon their modification state. In general, phosphorylation of histone H1 promotes decondensation [85], while acetylation of histones H3 and H4 promotes decondensation [83]. Histone methylation can either promote or counteract gene activity [102]. The effects of individual modifications are not strictly conserved across eukaryotes, but the principle of histone modification and histone-dependent condensation is.

Histones have a central role in the eukaryotic cell cycle (Figure 3). The decision of when to condense chromosomes is made by complex regulatory networks of signal transduction in which cyclins and cyclin-dependent kinases (CDKs) have the last word. The DNA-binding domain of cyclins traces to archaeal information processing: archaeal transcription factor B, from which eukaryotic transcription factor IIB, and cyclins, in turn, are evolutionarily derived [94,103]. Cyclins receive their signals from a number of pathways and regulators, perhaps the most important being TOR (target of rapamycin). TOR is a master regulator of cell division [87]; it takes its signals from nutrient sensing [90]. The term nutrients here refers to basic bulk nutrition – carbon, energy, and nitrogen (amino acids) – required for cell division and growth. An important sensor of nitrogen availability in eukaryotes, in terms of hierarchy and conservation across lineages, is GCN2 (general control nonderepressable), while the most central energy sensor is AMPK (AMP-activated protein kinase) [90]. GCN2 binds uncharged tRNAs as a proxy for cytosolic nitrogen availability. If uncharged tRNAs are abundant, genes for amino acid synthesis are switched on. Persistent amino acid starvation leads to signal transmission between GNC2 and TOR that, in yeast, instructs the cell to exit the cell cycle into G0, or in *Caenorhabditis* induces an inactive survival state called *Dauerlarvae*. AMPK senses ATP levels and carbon availability. AMPK senses the ratio of ATP to ADP and AMP and thereby acts as a fuel gauge for the cell [90], conveying to TOR signals required for informed decisions as to whether to sustain growth (continue the cell cycle) or to exit the cell cycle and enter G0 or quiescence. Starvation sensing via AMPK and the TOR family can also send signals that lead to autophagy [88,104], a self-digestion process in which cytoplasmic organelles, including mitochondria, are enzymatically degraded in vacuoles to provide basic nutrients for survival [93]. The signals sent by AMPK and TOR (GCN2 via TOR) elicit changes in histone modification [88,89], forging links between histone-dependent regulation of genes and the cell cycle on the one hand, with the most basic, vital needs of the cell (nutrition) on the other. In humans, nutrition limitation in early life can elicit epigenetic effects [105].

expression required by the vastly expanded cellular DNA complement of eukaryotes (Figure 3, Box 1). This epigenetic control can persist over many cell generations, allowing the differentiation of stable gene-expression patterns in various cell types [76]. Such gene control is essential in multicellular organisms in which different portions of the genome are expressed in different tissues. This type of control permits 'silencing' of large portions of the genome, making the expansion of eukaryotic genomes by 'selfish DNA' feasible [77]. These developmental patterns of gene expression are fundamental to the evolution of multicellular organisms, and a hallmark of eukaryotes.

Energy and Information: A Good Recipe for Complexity

The origin of eukaryotic histones traces to archaea. Histones are required for genomic complexity and the eukaryotic cell cycle, both of which are integral to eukaryotic cellular complexity (Figure 3). Dennis Searcy was investigating histone-like proteins in the archaea even before the archaea were recognized as a domain [78]. John Reeve and colleagues found bona fide histones in methanogens (euryarchaeotes) [58–60] and other archaea [79], emphasizing over decades the evolutionary significance of archaeal histones with regard to the eukaryote origin. It is possible that the role of histones in the process of eukaryote origin has been underappreciated. Perhaps histones were key.

The mitochondrion contributed genes, energy, and membranes to the origin of the eukaryote lineage. What did the archaeal partner contribute? That is, what archaeal trait allows eukaryotic cells

to accumulate, inherit, and express the large complex genomes that underpin eukaryote complexity? The answer might be simple: with the energetic configuration that mitochondria bestowed upon eukaryotes, the histone-dependent organization of DNA into chromatin allowed eukaryotes to take the first hurdle towards true complexity by regulating and managing large amounts of DNA [66–68]. The archaeal partner brings proteins with the histone-fold, from which eukaryotic histones, nucleosomes, and the substrate for epigenetic control of gene expression evolved [58–60]. The multiple-origin nature of archaeal DNA replication is also essential in the propagation of large amounts of DNA, which was necessary to replicate chromosomes of eukaryotic size [80,81].

Unlike the metabolic contribution of the bacterial partner, which is fully developed at the time of the symbiotic merger, the archaeal partner's contribution is only the 'potential' for manipulating large amounts of DNA in the form of proteins with the histone-fold. These proteins evolve into the eukaryotic histones, which are at the heart of the eukaryotic information management [70,75]. Stated another way, the union of a bacterium and an archaeon set the stage for the development of the eukaryotic lineage whereby the archaeal host compartment underwent a great deal more evolutionary modification than the mitochondrion did – the archaeal host was transformed from within, via gene transfer from the endosymbiont.

Histones, Decisions, and Signaling in the Cell Cycle

Energy powers the evolutionary process because energy powers life – without energy, no life, much less evolution. There are very interesting and important connections between histone modification and the energetic status of the eukaryotic cell. The essence of histone acetylation is that acetylation preferentially takes place at lysine residues, reducing their positive charge and hence their affinity to DNA, opening up chromatin for transcription. Mitochondrial proteins tend to undergo nonenzymatic (spontaneous) acylation by acetyl-CoA, an energy-rich thioester intermediate of core energy metabolism [82]. Might such a simple mechanism have been the seed of eukaryotic gene regulation through histone acetylation (see Outstanding Questions)? At eukaryote origin, sensing the nutritional status of the cell in the thioester currency of acetyl-CoA levels (histone acetylation) could have transferred information about the physiological state of the cell directly to chromosome activity in a meaningful manner. Acetylation of histones H3 and H4 promotes chromatin decondensation [83], removal of histone acetylation signals low acetyl-CoA levels (low nutrients) and leads to condensed inactive chromosomes. In an unrefined initial process, high nutrient levels would evoke decondensed active chromosomes, while low nutrient levels would lead to condensed, transcriptionally inactive chromatin.

Today, histone modification is highly regulated and tightly linked to (i) the energetic status of the cell [84], (ii) chromatin condensation [85], and (iii) decisions about the eukaryotic cell cycle [86]. Eukaryotes sense the energetic state of the cell with the help of the sensor proteins AMP-activated protein kinase (AMPK) and target of rapamycin (TOR) [87–89], and they sense the nutritional status of the cell (carbon, energy, and amino acid availability) via TOR and GCN2 [90]. By dry weight, nonphotosynthetic cells are about 50% protein and 10% nitrogen [91]. The three main energy and nutrient signaling pathways of eukaryotes – AMPK, TOR, and GCN2 – are conserved across eukaryotic groups [90]. All three master regulators transmit signals to chromatin via histone modification [88,89] (Box 1). By sensing the nutrient state of the cell they connect mitochondrial energy supply – and mitochondrial quality-control via autophagy [92,93] – with the eukaryotic cell cycle. The cell cycle is the basic and indivisible quantum unit of eukaryotic cell survival [94] (Figure 3). Of course, many eukaryotes have reduced the bioenergetic functions of their mitochondria or have transferred ATP synthesis to the cytosol [5,32], but in evolutionary terms the cell cycle was built upon archaeal histones in a cell that had mitochondria.

Even the CCAAT Box Traces Back to Histones

Archaeal histones were even instrumental in the early evolution of basic eukaryotic transcriptional regulation. How so? The CCAAT-box is found in 30% of eukaryotic promoters [95]. It binds the CCAAT-box binding complex, CBC, which promotes the recruitment of RNA polymerase II to eukaryotic DNA for transcription. Because it has undergone many duplications and functional specialization for regulation of various genes, CBC has several names in the literature, the most common being nuclear factor Y, or NF-Y [96]. The core subunits of CBC are evolutionarily derived from eukaryotic and archaeal histones [97,98]. The structural basis for CBC binding was resolved for the protein from *Aspergillus nidulans* [99]. *Aspergillus* CBC has the crystal structure of a histone H2A/H2B heterodimer [99]. It bends DNA, it can interact with the histone pair H3/H4, and it possesses a specific structural motif, the Nα helix of HapB, which binds directly to the CCAAT box and is strictly conserved throughout all eukaryotes [99]. Though lacking in archaea, the CCAAT box is ubiquitous among eukaryotic genomes, hence it was present in the eukaryotic ancestor. Its binding protein, CBC, is derived from archaeal histones, it is modified in the same manner as histones are [100], it is recruited to promoters specific to the cell cycle [101], it has retained the DNA-bending structure of histones [99], and it helps to activate about one in three eukaryotic genes [95]. Clearly, archaeal histones were important for the establishment of gene regulation that led to chromatin condensation, the establishment of a cell cycle (Figure 3), transcriptional response to nutrient availability, and complexity in eukaryotes that involves CCAAT box-dependent gene regulation. Taken together, that is a substantial contribution from archaeal histones.

Concluding Remarks

At eukaryotic origin, the bacterial partner contributed genes, energy, and membranes to the union, resulting in the eukaryotic lineage, while the host contributed information-processing ability

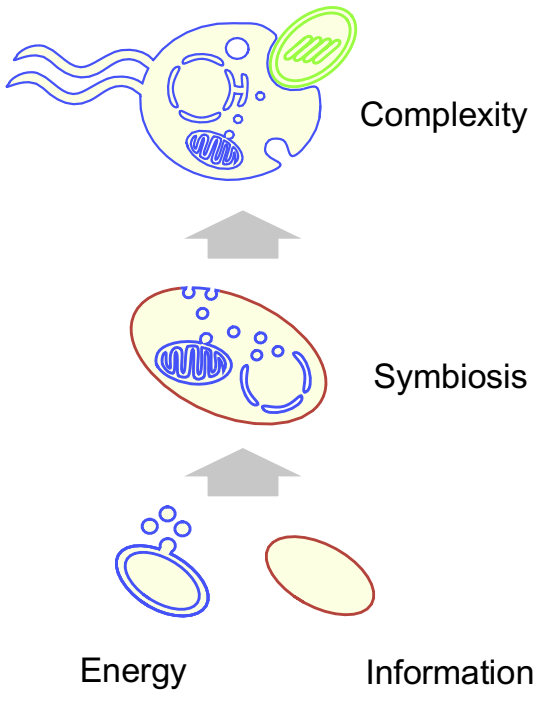

Figure 4. A Suggestion for the Role of the Archaeal Host at the Eukaryote Origin. Histones are a crucial component of information inheritance and processing in eukaryotes and furthermore reside at the heart of the eukaryotic cell cycle (see text). Outer-membrane vesicles at the surface of the endosymbiont and their implicated role in the origin of bacterial lipids in eukaryotes and the eukaryote endomembrane system [12,14–16] are indicated.

Outstanding Questions

If the host that acquired the mitochondrion was a complex phagocytosing cell, as some popular views currently have it, why has no one seen such cells, and why is it that only the cells that are descended from mitochondrial origin are truly complex? Were mitochondria and mitochondrial ATP supply irrelevant to the evolutionary process, as proponents of phagotrophic eukaryote origins contend?

Proteins with the histone-fold arose in archaea and are common in archaea where they function in DNA packaging. Why do archaea not have eukaryote-sized genomes with euchromatin, heterochromatin, epigenetics, and complex gene regulation founded in histone modification? Is it because mitochondria were required for the evolution of eukaryote genome complexity, or is it mere coincidence?

Histone modification is a chemical reaction involving the energy-rich donors acetyl-CoA (acetylation), ATP (phosphorylation), and *S*-adenosylmethionine (methylation), which have free energies of hydrolysis of –43, –31, and –26 kJ/mol, respectively. Is histone modification an ancient mechanism to directly sense the energetic state of the eukaryotic cell by translating metabolic information into gene regulation via histones?

The eukaryotic cell cycle involves dense condensation of chromosomes to the metaphase state. By the measure of evolutionary conservation, this condensation occurs as a prelude to every single cell division that has ever occurred within the eukaryotic kingdom during its 1.6 billion year history. Histones are the key to chromosome condensation and thus sit at the center of the cell cycle. Does the essential role of histones in the eukaryotic cell cycle reflect a key series of events at eukaryote origin?

(Figure 4). A case can be made that the major contribution of the archaeal partner to the union was the provision of proteins with the histone-fold that evolved into *bona fide* eukaryotic histones, which became the core of both eukaryotic nucleosomes [59] and the cell cycle (Figure 3).

The compaction of eukaryotic DNA into nucleosomes allowed the nascent eukaryotic lineage to store and manage vastly more information than was possible in prokaryotes, and thus to express a dramatically larger number of different proteins than were expressed in prokaryotes. At the same time, the multiple replication origins germane to archaeal DNA [80,81] allowed eukaryotic chromosomes to expand to the size range now characteristic of eukaryotic cells. Through acetylation and methylation, histones also provided a simple mechanism of information transfer about the nutrient and energetic state of the nascent eukaryotic cell into chromatin condensation and activity.

With the combination of increased energy available per gene, provided by multiple mitochondria, and virtually unlimited genetic capacity, provided by eukaryotic chromatin structure, the eukaryotic lineage embarked on the exploration of evolutionary innovations that came to include meiosis and the cell cycle, multicellular organisms with different cell types, and complex cellular interactions within an organism. Increased genetic capacity allowed for the elucidation of complex developmental programs, many underpinned by CCAAT box-dependent gene regulation, as well as innovation in protein structure. The characteristics we associate with the eukaryotic lineage are the direct result of the increased energy provided by the bacterial partner (mitochondria) coupled with an increased capacity to handle DNA afforded by organization of eukaryotic DNA into nucleosomes contributed by the histones from the archaeal partner. This combination of traits endowed the first eukaryote with evolutionary options, particularly in the realm of cell structures and multicellularity, unavailable to any prokaryotic lineage. The cell cycle is the cornerstone of eukaryotic cell division and growth. At its heart reside histones, both in a regulatory and a structural role, giving them a special place in the origin and evolution of both the eukaryotic cell and the eukaryotic cell cycle.

Acknowledgments

W.F.M. thanks the European Research Council (Grant 666053) and the Volkswagenstiftung, Germany (Grant 93 046) for financial support. We thank Sriram Garg, Sven Gould, and Michael Grunstein for discussions and comments on the paper, and Verena Zimorski for help in preparing the manuscript.

References

1. McInerney, J.O. *et al.* (2014) The hybrid nature of the Eukaryota and a consilient view of life on Earth. *Nat. Rev. Microbiol.* 12, 449–455
2. Betts, H.C. *et al.* (2018) Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nat. Ecol. Evol.* 2, 1556–1562
3. Judson, O.P. (2017) The energy expansions of evolution. *Nat. Ecol. Evol.* 1, 0138
4. Zillig, W. *et al.* (1989) Did eukaryotes originate by a fusion event? *Endocyt. Cell Res.* 6, 1–25
5. Martin, W. and Müller, M. (1998) The hydrogen hypothesis for the first eukaryote. *Nature* 392, 37–41
6. Rivera, M.C. and Lake, J.A. (2004) The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature* 431, 152–155
7. Blackstone, N.W. (2016) An evolutionary framework for understanding the origin of eukaryotes. *Biology* 5, 18
8. Lane, N. and Martin, W.F. (2010) The energetics of genome complexity. *Nature* 467, 929–934
9. Pisani, D. *et al.* (2007) Supertrees disentangle the chimerical origin of eukaryotic genomes. *Mol. Biol. Evol.* 24, 1752–1760
10. Cotton, J.A. and McInerney, J.O. (2010) Eukaryotic genes of archaebacterial origin are more important than the more numerous eubacterial genes, irrespective of function. *Proc. Natl. Acad. Sci. U. S. A.* 107, 17252–17255
11. Ku, C. *et al.* (2015) Endosymbiotic origin and differential loss of eukaryotic genes. *Nature* 524, 427–432
12. Gould, S.B. (2018) Membranes and evolution. *Curr. Biol.* 28, R381–R385
13. Williams, T.A. *et al.* (2013) An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* 504, 231–236
14. Gould, S.B. *et al.* (2016) Bacterial vesicle secretion and the evolutionary origin of the eukaryotic endomembrane system. *Trends Microbiol.* 24, 525–534
15. McBride, H.M. (2018) Mitochondria and endomembrane origins. *Curr. Biol.* 28, R367–R372
16. Sugiura, A. *et al.* (2017) Newly born peroxisomes are a hybrid of mitochondrial and ER-derived pre-peroxisomes. *Nature* 542, 251–254
17. Deatherage, B.L. and Cookson, B.T. (2012) Membrane vesicle release in bacteria, eukaryotes, and archaea: a conserved yet underappreciated aspect of microbial life. *Infect. Immun.* 80, 1948–1957
18. Fox, G. *et al.* (1980) The phylogeny of prokaryotes. *Science* 209, 457–463

19. van Valen, L.M. and Maiorana, V.C. (1980) The Archaebacteria and eukaryotic origins. *Nature* 287, 248–250
20. Spang, A. *et al.* (2015) Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521, 173–179
21. Embley, T.M. and Williams, T.A. (2015) Evolution: steps on the road to eukaryotes. *Nature* 521, 169–170
22. Zaremba-Niedzwiedzka, K. *et al.* (2017) Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 541, 353–358
23. Searcy, D.G. *et al.* (1981) A mycoplasma-like archaebacterium possibly related to the nucleus and cytoplasm of eukaryotic cells. *Ann. N. Y. Acad. Sci.* 361, 312–324
24. Cavalier-Smith, T. (1989) Archaebacteria and archezoa. *Nature* 339, 100–101
25. Carballido-Lopez, R. (2006) The bacterial actin-like cytoskeleton. *Microbiol. Mol. Biol. Rev.* 70, 888–909
26. Hara, F. *et al.* (2007) An actin homolog of the archaeon *Thermoplasma acidophilum* that retains the ancient characteristics of eukaryotic actin. *J. Bacteriol.* 189, 2039–2045
27. Erickson, H.P. (2007) Evolution of the cytoskeleton. *BioEssays* 29, 668–677
28. Löwe, J. and Amos, L.A. (2009) Evolution of cytomotive filaments: the cytoskeleton from prokaryotes to eukaryotes. *Int. J. Biochem. Cell Biol.* 41, 323–329
29. Martin, W.F. *et al.* (2015) Endosymbiotic theories for eukaryote origin. *Philos. Trans. R. Soc. Lond. B* 370, 20140330
30. Lane, N. (2017) Serial endosymbiosis or singular event at the origin of eukaryotes. *J. Theor. Biol.* 434, 58–67
31. Booth, A. and Doolittle, W.F. (2015) Eukaryogenesis, how special really? *Proc. Natl. Acad. Sci. U. S. A.* 112, 10278–10285
32. Hampl, V. *et al.* (2019) Was the mitochondrion necessary to start eukaryogenesis? *Trends Microbiol.* 27, 96–104
33. Martin, W.F. *et al.* (2017) The physiology of phagocytosis in the context of mitochondrial origin. *Microbiol. Mol. Biol. Rev.* 81, e00008–e00017
34. Surkont, J. and Pereira-Leal, J.B. (2016) Are there Rab GTPases in archaea? *Mol. Biol. Evol.* 33, 1833–1842
35. Dey, G. *et al.* (2016) On the archaeal origins of eukaryotes and the challenges of inferring phenotype from genotype. *Trends Cell Biol.* 26, 476–485
36. McInerney, J.O. and O'Connell, M.J. (2017) Mind the gaps in cellular evolution. *Nature* 541, 297–299
37. Da Cunha, V. (2017) Lokiarchaea are close relatives of Euryarchaeota, not bridging the gap between prokaryotes and eukaryotes. *PLoS Genet.* 13, e1006810
38. Burns, J.A. *et al.* (2018) Gene-based predictive models of trophic modes suggest Asgard archaea are not phagocytotic. *Nat. Ecol. Evol.* 2, 697–704
39. Esser, C. *et al.* (2004) A genome phylogeny for mitochondria among α-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol. Biol. Evol.* 21, 1643–1660
40. Rivera, M.C. *et al.* (1998) Genomic evidence for two functionally distinct gene classes. *Proc. Natl. Acad. Sci. U. S. A.* 95, 6239–6244
41. Zillig, W. *et al.* (1985) Archaebacteria and the origin of the eukaryotic cytoplasm. *Curr. Top. Microbiol. Immunol.* 114, 1–18
42. Woese, C.R. (2002) On the evolution of cells. *Proc. Natl. Acad. Sci. U. S. A.* 99, 8742–8747
43. Lake, J.A. *et al.* (1984) Eocytes: a new ribosome structure indicates a kingdom with a close relationship to eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 81, 3786–3790
44. Woese, C.R. *et al.* (2000) Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. Mol. Biol. Rev.* 64, 202–236
45. Gregory, T.R. (2001) Coincidence, coevolution, or causation DNA content, cell size, and the C-value enigma. *Biol. Rev.* 76, 65–101
46. Elliott, T.A. and Gregory, T.R. (2015) What's in a genome? The C-value enigma and the evolution of eukaryotic genome content. *Philos. Trans. R. Soc. Lond. B* 370, 20140331
47. Mirsky, A.E. and Ris, H. (1951) The desoxyribonucleic acid content of animal cells and its evolutionary significance. *J. Gen. Physiol.* 34, 451–462
48. Graur, D. (2016) Molecular and Genome Evolution, Sinauer
49. Thomas, C.A. (1971) The genetic organization of chromosomes. *Annu. Rev. Genet.* 5, 237–256
50. Cavalier-Smith, T. (1982) Skeletal DNA and the evolution of genome size. *Annu. Rev. Biophys. Bioeng.* 11, 273–302
51. Nevers, P. and Saedler, H. (1977) Transposable genetic elements as agents of gene instability and chromosomal rearrangements. *Nature* 268, 109–115
52. Orgel, L.E. and Crick, F.H.C. (1980) Selfish DNA: the ultimate parasite. *Nature* 284, 604–607
53. Graur, D. *et al.* (2013) On the immortality of television sets: 'function' in the human genome according to the evolution-free gospel of ENCODE. *Genome Biol. Evol.* 5, 578–590
54. Martin, W. and Koonin, E.V. (2006) Introns and the origin of nucleus–cytosol compartmentalization. *Nature* 440, 41–45
55. Harold, F.M. (1986) *The Vital Force: A Study of Bioenergetics*, W.H. Freeman
56. Post, C.B. and Zimm, B.H. (1982) Theory of DNA condensation: collapse versus aggregation. *Biopolymers* 21, 2123–2137
57. Sandman, K. *et al.* (1998) Diversity of prokaryotic chromosomal proteins and the origin of the nucleosome. *Cell. Mol. Life Sci.* 54, 1350–1364
58. Sandman, K. and Reeve, J.N. (1998) Origin of the eukaryotic nucleus. *Science* 280, 501–503
59. Sandman, K. and Reeve, J.N. (2006) Archaeal histones and the origin of the histone fold. *Curr. Opin. Microbiol.* 9, 520–525
60. Mattiroli, F. *et al.* (2017) Structure of histone-based chromatin in Archaea. *Science* 357, 609–612
61. Henneman, B. *et al.* (2018) Structure and function of archaeal histones. *PLoS Genet.* 14, e1007582
62. Luger, K. *et al.* (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? *Nat. Rev. Mol. Cell Biol.* 13, 437–447
63. Tremethick, D.J. (2007) Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell* 128, 651–654
64. Minsky, A. *et al.* (1997) Nucleosomes: a solution to a crowded intracellular environment. *J. Theor. Biol.* 188, 379–385
65. Ou, H.D. *et al.* (2017) ChromEMT: visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* 357, eaag0025
66. Vafabakhsh, R. and Ha, T. (2012) Extreme bendability of DNA less than 100 base pairs long revealed by single-molecule cyclization. *Science* 337, 1097–1101
67. Grunstein, M. (1997) Histone acetylation in chromatin structure and transcription. *Nature* 389, 349–352
68. Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature* 403, 42–45
69. Jenuwein, T. and Allis, C.D. (2001) Translating the histone code. *Science* 293, 1074–1080
70. Kurdistani, S.K. and Grunstein, M. (2003) Histone acetylation and deacetylation in yeast. *Nat. Rev. Mol. Cell Biol.* 4, 276–284
71. Roth, S.Y. *et al.* (2001) Histone acetyltransferases. *Annu. Rev. Biochem.* 70, 81–120
72. Sperling, A. and Grunstein, M. (2009) Histone H3 N-terminus regulates higher order structure of yeast heterochromatin. *Proc. Natl. Acad. Sci. U. S. A.* 106, 13153–13159
73. Hake, S.B. and Allis, C.D. (2006) Histone H3 variants and their potential role in indexing mammalian genomes: the 'H3 barcode hypothesis'. *Proc. Natl. Acad. Sci. U. S. A.* 103, 6428–6435
74. Peeters, J.G.C. *et al.* (2019) Transcriptional and epigenetic profiling of nutrient-deprived cells to identify novel regulators of autophagy. *Autophagy* 15, 98–112
75. Badeaux, A.I. and Shi, Y. (2013) Emerging roles for chromatin as a signal integration and storage platform. *Nat. Rev. Mol. Cell Biol.* 14, 211–224
76. Butler, J.S. *et al.* (2012) Histone-modifying enzymes regulators of developmental decisions and drivers of human disease. *Epigenomics* 4, 163–177
77. Soshnev, A.A. *et al.* (2016) Greater than the sum of parts: complexity of the dynamic epigenome. *Mol. Cell* 62, 681–694
78. Searcy, D.G. (1975) Histone-like protein in the prokaryote *Thermoplasma acidophilum*. *Biochim. Biophys. Acta* 395, 535–547
79. Cubonova, L. *et al.* (2005) Histones in crenarchaea. *J. Bacteriol.* 187, 5482–5485

80. Barry, E.R. and Bell, S.D. (2006) DNA replication in the archaea. *Microbiol. Mol. Biol. Rev.* 70, 876–887
81. Sarmiento, F. *et al.* (2014) Diversity of the DNA replication system in the archaea domain. *Archaea* 2014, 675946
82. Trub, A.G. and Hirschey, M.D. (2018) Reactive acyl-CoA species modify proteins and induce carbon stress. *Trends Biochem. Sci.* 43, 369–379
83. Eberharter, A. and Becker, P.B. (2002) Histone acetylation: a switch between repressive and permissive chromatin. *EMBO Rep.* 3, 224–229
84. Fan, J. *et al.* (2015) Metabolic regulation of histone post-translational modifications. *ACS Chem. Biol.* 10, 95–108
85. Fyodorov, D.V. *et al.* (2018) Emerging roles of linker histones in regulating chromatin structure and function. *Nat. Rev. Mol. Cell Biol.* 19, 192–206
86. Ma, Y. *et al.* (2015) How the cell cycle impacts chromatin architecture and influences cell fate. *Front. Genet.* 6, 19
87. Wang, X. and Proud, C.G. (2009) Nutrient control of TORC1, a cell-cycle regulator. *Trends Cell Biol.* 19, 260–267
88. Vancura, A. *et al.* (2018) Reciprocal regulation of AMPK/SNF1 and protein acetylation. *Int. J. Mol. Sci.* 19, 3314
89. Beckwith, S.L. *et al.* (2018) The INO80 chromatin remodeler sustains metabolic stability by promoting TOR signaling and regulating histone acetylation. *PLoS Genet.* 14, e1007216
90. Chantranupong, L. *et al.* (2015) Nutrient sensing mechanisms across evolution. *Cell* 161, 67–83
91. Fagerbakke, K.M. *et al.* (1996) Content of carbon, nitrogen, oxygen, sulfur and phosphorus in native aquatic and cultured bacteria. *Aquat. Microb. Ecol.* 10, 15–27
92. Jung, C.H. *et al.* (2010) mTOR regulation of autophagy. *FEBS Lett.* 584, 1287–1295
93. He, L. *et al.* (2018) Autophagy: The last defense against cellular nutritional stress. *Adv. Nutr.* 9, 493–504
94. Garg, S.G. and Martin, W.F. (2016) Mitochondria, the cell cycle, and the origin of sex via a syncytial eukaryote common ancestor. *Genome Biol. Evol.* 8, 1950–1970
95. Bucher, P. (1990) Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol.* 212, 563–578
96. Mantovani, R. (1999) The molecular biology of the CCAAT-binding factor NF-Y. *Gene* 239, 15–27
97. Romier, C. *et al.* (2003) The NF-YB/ NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. *J. Biol. Chem.* 278, 1336–1345
98. Sandman, K. *et al.* (1990) HMf, a DNA-binding protein isolated from the hyperthermophilic archaeon *Methanothermus fervidus*, is most closely related to histones. *Proc. Natl. Acad. Sci. U. S. A.* 87, 5788–5791
99. Huber, E.M. *et al.* (2012) DNA minor groove sensing and widening by the CCAAT-binding complex. *Structure* 20, 1757–1768
100. Peng, Y. and Jahroudi, N. (2003) The NFY transcription factor inhibits von Willebrand factor promoter activation in non-endothelial cells through recruitment of histone deacetylases. *J. Biol. Chem.* 278, 8385–8394
101. Caretti, G. *et al.* (2003) Dynamic recruitment of NF-Y and histone acetyltransferases on cell-cycle promoters. *J. Biol. Chem.* 278, 30435–30440
102. Torres, I.O. and Fujimori, D.G. (2015) Functional coupling between writers, erasers and readers of histone and DNA methylation. *Curr. Opin. Struct. Biol.* 35, 68–75
103. Gibson, T.J. *et al.* (1994) Evidence for a protein domain superfamily shared by the cyclins, TFIIB and RB/p107. *Nucleic Acids Res.* 22, 946–952
104. Mihaylova, M.M. and Shaw, R.J. (2011) The AMPK signalling pathway coordinates cell growth, autophagy and metabolism. *Nat. Cell Biol.* 13, 1016–1023
105. Canani, R. *et al.* (2011) Epigenetic mechanisms elicited by nutrition in early life. *Nutr. Res. Rev.* 24, 198–403
106. Martin, W.F. (2017) Symbiogenesis, gradualism, and mitochondrial energy in eukaryote evolution. *Period. Biol.* 119, 141–158
107. Sagan, L. (1967) On the origin of mitosing cells. *J. Theor. Biol.* 14, 225–274
108. Gould, S.B. *et al.* (2008) Plastid evolution. *Annu. Rev. Plant Biol.* 59, 491–517
109. Harish, A. and Kurland, C.G. (2017) Mitochondria are not captive bacteria. *J. Theor. Biol.* 434, 88–98