

Life With a Few Genes: A Survey on Naturally Evolved Reduced Genomes

Luis Delaey^{*1,2}, Rosario Gil^{1,3,4}, Juli Peretó^{1,4,5}, Amparo Latorre^{1,3,4} and Andrés Moya^{1,3,4}

¹Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, València, Spain

²Facultad de Ciencias, Universidad Nacional Autónoma de México, México

³Departament de Genètica, Universitat de València, Spain

⁴CIBER de Epidemiología y Salud Pública, Instituto de Salud Carlos III

⁵Departament de Bioquímica i Biologia Molecular, Universitat de València, Spain

Abstract: Engineering a cell with the minimum number of genes, functions or components required for living in a defined environment has become one of the promises of system biology. Theoretical as well as experimental approaches have been undertaken in order to identify essential genes and functions for different model organisms. These experiments have shown that, despite the well-known group of universally conserved genes mainly related to transcription and translation, there is a fraction of essential genes which are particular to each one of the species analyzed, thus defining a diversity of possible minimal genomes. Comparative analysis of naturally evolved near-minimal genomes have the potential to uncover how nature has solved in different ways the accomplishment of basic functions to sustain a cell. Here we review the state of the art of reduced sequenced genomes and discuss the concept of a minimal cell at the light of this biodiversity.

Keywords: Minimal cell, reduced genomes, endosymbionts.

THE MINIMAL CELL CONCEPT

The quest for the minimal cell can be posed as the search for the necessary and sufficient features of life. This question, with a long tradition in theoretical biology [1], has become amenable to experimentation thanks to the spectacular advances in molecular biology techniques. The availability of complete genome sequences for several (often closely related) species, and the development of *in vitro*, as well as *in silico* quantitative analytical techniques designated to study collectively the molecules that constitute the intricate machinery of cells, have the potential to burst our present day understanding of biological systems. However, as discussed here, an evolutionary perspective of cells provides the framework to better understand the properties of minimal living systems. In particular, attention must be given to the diversity of naturally evolved reduced genomes.

Approaches to identify minimal components of life can be conceptually divided in two, namely the *bottom-up* and the *top-down* [2] (Fig. 1). The *bottom-up* approach aims at constructing artificial chemical systems capable of replication and evolution starting from simpler components (i.e., building up cells from scratch). By recreating some of the steps that pre-cellular systems may have followed from chemistry to biology under primitive Earth conditions, this research program is motivated by the possibility to better understand the origin of life on Earth [2]. Although no such a system has been developed yet, extremely simple systems might eventually be constructed from this approach showing

some typical biological properties. Thus, subsystems considered as supra-chemical or infra-biological [3] might be composed of a combination of three basic properties of cells, namely: Metabolism (M), Template based replication of hereditary material (T) and Boundary of the system (B) [4]. Progress in the combinatorial of some of these subsystems includes the vesicle-encapsulated, template-directed, enzyme-free synthesis of a genetic polymer from externally provided monomers [5]. It is worth noting that those efforts are in context of the origin-of-life research program as originally formulated by Oparin [6], i.e., to discover the articulation of chemical subsystems exhibiting life-like properties under plausible early Earth conditions.

The *top-down* approach starts from existing organisms with the objective of simplifying their genome, leading to a minimal (or at least reduced) gene set. The rationale behind is common ground. Since the cell's hereditary information needed for structure and function is coded in its genes, cell simplification can be reduced to the task of engineering the genome [7]. Essential genes are identified through deletion mutant experiments or by *in silico* comparative genome analysis. The difference between computational and experimental approaches is that the former identifies a set of essential genes that is shared among diverse taxa, whereas the latter searches for individual genes that are essential for growth in a single species and in the conditions under study [8]. Then, a theoretical minimal genome is proposed comprising the set of essential genes identified. For instance, systematic attempts to identify essential genes through deletion mutant experiments includes *Escherichia coli* [9], *Bacillus subtilis* [10] and *Mycoplasma genitalium* [11] where 620 (out of 3746), 271 (out of ~4100) and 382 (out of 482) of the genes were shown to be indispensable respectively when

*Address correspondence to this author at the Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Universitat de València, València, Spain, Tel: + 34 96 354 38; Fax: 34 963543670; E-mail: luis.delaey@uv.es

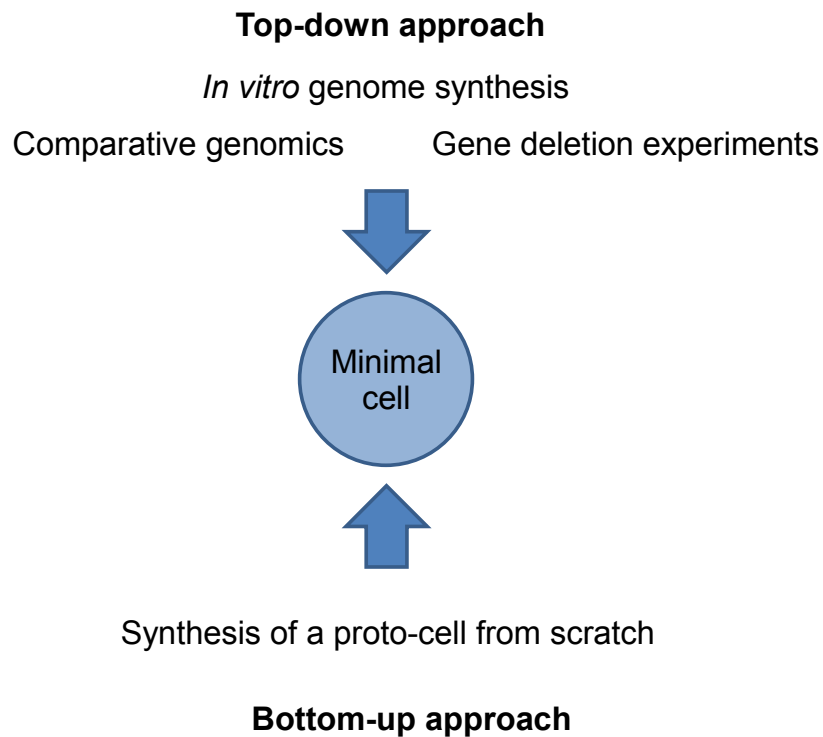


Fig. (1). Bottom-up and top-down approaches to the minimal cell.

singly deleted under the experimental conditions tested. Examples of identification of non essential genes through comparative analysis of several strains has been applied to *E. coli* where it has been possible to delete ~15% of its genome, without comprising its metabolic efficiency and unexpected increase in its electroporation efficiency [12]. The synthesis of the complete genome sequence of *M. genitalium* [13] is also an effort towards the engineering of simpler biological systems starting from present day cells.

Bottom-up and top-down approaches will ultimately provide models for better understanding the early history of life on Earth and will help to untangle the intricate molecular machinery of modern cells. In particular, top-down approaches can enormously benefit from the study of present day organisms with small genomes by showing different ways in which relatively simple biological systems might operate. In this sense, genomes from endosymbionts, parasites and free-living organisms with reduced genomes are *par excellence* examples of naturally evolved minimal gene sets. Because small genomes preferentially retain basic functions for sustaining a cell, a comparative analysis of their diverse biology might help to better understand the properties of a hypothetical minimal cell [14]. Here, we review the state of the art of sequenced naturally evolved reduced genomes and discuss some aspects of their evolution in relation to the minimal cell concept.

NATURALLY EVOLVED REDUCED GENOMES

Morowitz [15] made a distinction between minimal cells requiring precise and chemically complex environments for heterotrophic growth, and those demanding only minimal components to the environment to autotrophically self construct their biomolecules. Paradigmatic cases for the above

are *M. genitalium* on the heterotrophic side and cyanobacteria on the autotrophic one. Not unexpectedly, there is an inverse relationship between the number of genes a reduced genome harbors, and the chemical complexity of the environment the organism requires for growth. As shown in Fig. (2), obligate host associated prokaryotes have the smallest genomes, being endosymbiotic ones those endowed with the smallest genomes.

Nowadays, the sequenced genome with the smallest number of genes from a free-living organism belongs to an uncultured but abundant ocean beta-proteobacterium from the clade OM43, strain HTCC2181 with 1377 genes [16] (Fig. 2). This genome is followed closely by others like that of the cosmopolitan oceanic bacterium *Candidatus Pelagibacter ubique* HTCC1062 (one of the most successful clade of prokaryotes), the dehalorespirant *Dehalococcoides* sp. BAV1 and the hyperthermophilic crenarchaeon *Ignicoccus hospitalis* KIN4/I, with 1394, 1436 and 1494 genes respectively [17-19].

The diversity of lineages, nutritional strategies and ecological niches occupied by these free-living organisms with small genomes is noteworthy (Table 1). Among the bacterial domain, the proteobacterium HTCC2181 is an obligate methylotroph marine bacterium that specialize on C1 compounds like methanol and formaldehyde as sources of carbon and energy [16]. *Candidatus Pelagibacter ubique* HTCC1062 is a heterotroph that grow by assimilating organic compounds from the ocean's dissolved organic carbon reservoir and can generate energy by a light-driven proteorhodopsin proton pump or by respiration [17]. The *Dehalococcoides* sp. BAV1 (Chloroflexi phylum) is a dehalorespiring heterotrophic strictly anaerobic bacterium capable of derive all its energy required for growth from the reduction of vinyl chlo-

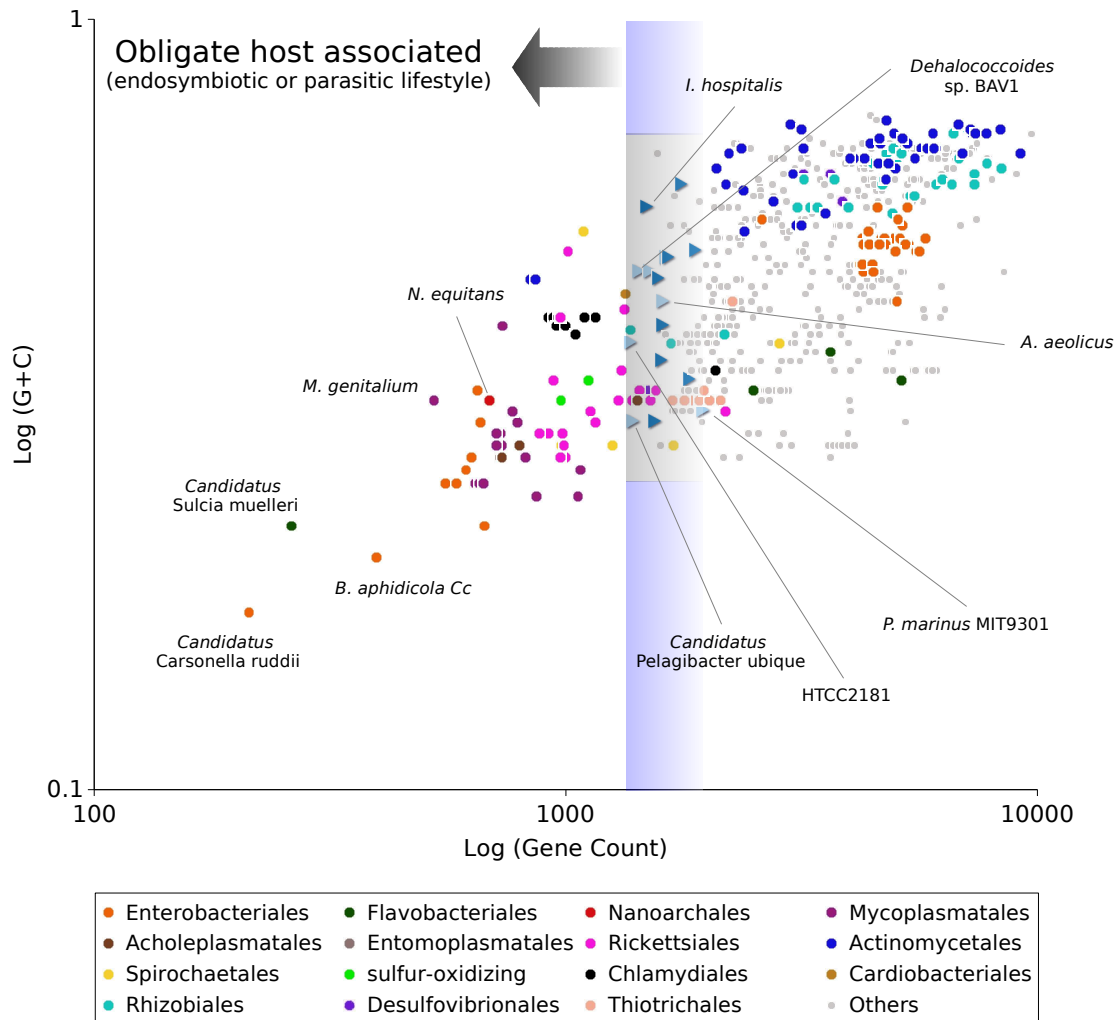


Fig. (2). Diversity of naturally evolved reduced prokaryotic genomes. Vertical shaded line indicates extant limit of free-living sequenced genomes. Dark-blue triangles indicate free-living archaea; Pale-blue triangles indicate free-living bacteria. Gene count and G+C content according to [69].

ride (VC) to ethene [18]. The smallest photosynthetic cells, represented by *Prochlorococcus marinus* have slightly larger genomes (in the order of 1.7 Mbp) than the previous prokaryotes [20]. The hyperthermophilic bacteria *Aquifex aeolicus*, a chemolithoautotroph, is capable of growing on hydrogen, oxygen, carbon dioxide, and mineral salts [21]. Among archaeal prokaryotes, *I. hospitalis* is another chemolithoautotroph that couples CO₂ fixation with sulfur respiration using molecular hydrogen in high temperature hydrothermal vent systems [19]. *Methanococcus aeolicus* Nakai-3 is a CO₂ reducing methanogenic marine euryarchaeota [22]. Another euryarchaeota, the strictly aerobic extremophile *Picrophilus torridus* DSM 9790 is a heterotroph that can use many type of sugars as their energy source, as well as propionate, and is also able to grow around pH 0 at up to 65 degrees C, thus they represent one of the most thermoacidophilic organisms known [23]. Last but not least, the aerobic hyperthermophilic korarchaea *Candidatus* Korarchaeum cryptofilum OPF8 grows heterotrophically, using a variety of peptide and amino acid degradation pathways [26].

As shown in Table 1, several species of free-living prokaryotes belonging to different orders have evolved small genomes with a similar number of genes, indicating that extant cells requires around 1400 to 1500 genes to sustain a free-living lifestyle under present-day natural conditions. A similar conclusion was achieved by a study on published genome sizes determined by pulse-field gel electrophoresis (PFGE) on 641 prokaryotic genomes [28]. These observations, although limited by the sample of available data, are nevertheless supported by the diversity of lineages as well as the different nutritional strategies and ecological niches represented by them. They also indicate that the minimal genome for a free-living cell may be in the order of 1Mbp [19].

Genomes with fewer genes than the smallest free-living prokaryote belong to parasitic or endosymbiotic organisms and are found in fifteen different orders among extant sequenced genomes (Table 2). Evolved from free-living relatives by a process of massive genome reduction, they have been traditionally studied as natural models of minimal cells [8, 14]. Perhaps one of the most interesting surprises found

Table 1. Naturally Evolved Free-Living Prokaryotes with Small Genomes

Taxonomy	Species	Gene count	G+C content	Genome size (Mbp)	Comment on nutritional strategy
Aquificae					
Aquificae	<i>Aquifex aeolicus</i> VF5	1613	0.43	1.59	Chemolithoautotroph [20]
Cyanobacteria					
Cyanobacteria	<i>Prochlorococcus marinus</i> (12 strains)	1765 – 3127	0.31 – 0.51	1.66 – 2.68	Photoautotroph [21]
Chloroflexi					
Dehalococcoidetes	<i>Dehalococcoides</i> sp. (2 strains)	1436 – 1641	0.47	1.34 – 1.40	Dehalorespiring heterotroph [18]
Proteobacteria					
Alpha-Proteobacteria					
	<i>Candidatus Pelagibacter ubique</i>	1394	0.3	1.31	Heterotroph [17]
Beta-Proteobacteria					
	Methylophilales bacterium HTCC2181	1377	0.38	1.3	Methylotroph [16]
Crenarchaeota					
Thermoprotei	<i>Ignicoccus hospitalis</i> KIN4/I	1494	0.57	1.28	Chemolithoautotroph [19]
Euryarchaeota					
Thermoplasmata	<i>Picrophilus torridus</i> DSM 9790	1605	0.36	1.54	Heterotroph, aerobic [22]
Methanococci	<i>Methanococcus aeolicus</i> Nankai-3	1555	0.3	1.57	Chemoautotroph [23]
Methanopyri	<i>Methanopyrus kandleri</i> AV19	1768	0.61	1.69	Chemolithoautotroph [24]
Methanobacteria	<i>Methanothermobacter thermoautotrophicus</i> Delta H	1893	0.5	1.75	Lithoautotroph [25]
Korarchaeota					
Korarchaea	<i>Candidatus Korarchaeum cryptofilum</i>	1660	0.49	1.59	Heterotroph [26]
Thaumarchaeota					
Cenarchaea	<i>Nitrosopumilus maritimus</i> SCM1	1840	0.34	1.64	Autotroph [27]

in these reduced genomes is the lack of genes previously thought to be essential for life. For example, in *Candidatus Phytoplasma asteris* OY strain as well as in *Buchnera aphidicola* BCc genomes, there is a lack for the genes coding for the ATP-synthase subunits [29, 30] and *Ureaplasma urealyticum* lacks the heat shock protein/chaperonins GroEL and GroES and the cell-division protein FtsZ [31]. Examples of simpler biochemical pathways are also found in *U. urealyticum*. This bacterium generates 95% of its ATP through the hydrolysis of urea by urease [32]. In terms of the number of genes involved, ATP production by urea hydrolysis is a simpler process than carbohydrate metabolism is [31]. A comparison of the gene set of the minimal genome derived from the comparison of *M. genitalium* and *Haemophilus influenzae* [33] to the *Mycoplasma mycoides* subsp. *mycoides* SC gene set showed that 11 out of 254 genes of the minimal genome are absent in this strain [34]. In addition, several genes shown to be essential in *E. coli* are absent in the related gamma-proteobacterium symbiont of the clam *Calyp-*

togena okutanii [35], including essential genes for cytokinesis (*ftsZ* and related genes) among others.

EVOLUTIONARY MECHANISMS OF GENOME REDUCTION

Reduced genomes from endosymbiotic prokaryotes share some general trends that together conform what has been called a reductive genome syndrome [36]. These trends include low G+C content, acceleration of nucleotide substitution rate, and the loss of adaptive codon bias, among others [37, 38]. The reductive genome process in these bacteria seems to be the result of relaxation of natural selection for maintaining genes used in the biosynthesis of compounds that can be imported from the host. This process, combined with an increase of genetic drift, leading to the accumulation of slightly deleterious mutations in a process known as Muller's ratchet, might eventually lead to the functional inactivation and loss of non-essential genes [38].

Table 2. Naturally Evolved Symbiotic or Parasites Prokaryotes with Reduced Genomes

Taxonomy	Species	Gene count	G+C content	Genome size (Mbp)	Life-style*
Gamma-Proteobacteria					
Enterobacteriales					
	<i>Candidatus Carsonella ruddii</i> PV	213	0.17	0.16	OE
	<i>Buchnera aphidicola</i> (Cc, Bp, Sg and APS strains)	397 – 614	0.2 – 0.26	0.42 – 0.65	OE
	<i>Baumannia cicadellinicola</i> Hc	650	0.33	0.69	OE
	<i>Wigglesworthia glossinidia</i> end. of <i>Glossina brevipalpis</i>	671	0.22	0.70	OE
	<i>Candidatus Blochmannia floridanus</i>	632	0.27	0.71	OE
	<i>Candidatus Blochmannia pennsylvanicus</i> BPEN	660	0.3	0.79	OE
Sulfur-oxidizing					
	<i>Candidatus Vesicomysocius okutanii</i> HA	977	0.32	1.02	OE
Unclassified					
	<i>Candidatus Ruthia magnifica</i> Cm	1117	0.34	1.16	OE
Cardiobacteriales					
	<i>Dichelobacter nodosus</i> VCS1703A	1341	0.44	1.39	OP
Thiotrichales					
	<i>Francisella tularensis</i>	1.4 – 2.1 ⁿ	0.32	~1.9	OIP
Cardiobacteriales					
	<i>Dichelobacter nodosus</i> VCS1703A				
Alpha-Proteobacteria					
Rickettsiales					
	<i>Neorickettsia sennetsu</i> Miyayama	973	0.41	0.86	OIP
	<i>Rickettsia typhi</i> Wilmington	920	0.29	1.11	OIP
	<i>Rickettsia prowazekii</i> Madrid E	881	0.29	1.11	OIP
	<i>Rickettsia canadensis</i> McKiel	1129	0.31	1.16	OIP
	<i>Rickettsia akari</i> Hartford	1294	0.32	1.23	OIP
	<i>Rickettsia rickettsii</i> (2 strains)	1.3 – 1.4 ⁿ	0.32	1.23 – 1.27	OIP
	<i>Rickettsia conorii</i> Malish 7	1416	0.32	1.27	OIP
	<i>Rickettsia massiliae</i> MTU5	1435	0.33	1.38	OIP
	<i>Rickettsia bellii</i> (2 strains)	1.4 – 1.5 ⁿ	0.32	~1.52	OIP
	<i>Rickettsia felis</i> URRWXCAl2	1552	0.33	1.59	OIP
	<i>Anaplasma marginale</i> St. Maries	1011	0.5	1.20	OIP
	<i>Anaplasma phagocytophilum</i> HZ	1333	0.42	1.47	OIP
	<i>Ehrlichia canis</i> Jake	985	0.29	1.31	OIP
	<i>Ehrlichia ruminantium</i> (3 strains)	974 – 991	0.27– 0.28	~1.50	OIP
	<i>Ehrlichia chaffeensis</i> Arkansas	1157	0.3	1.18	OIP
	<i>Wolbachia endosymbiont</i> TRS of <i>Brugia malayi</i>	940	0.34	1.08	OE
	<i>Wolbachia endosymbiont</i> of <i>Drosophila melanogaster</i>	1312	0.35	1.27	OIP

(Table 2). Contd.....

Taxonomy	Species	Gene count	G+C content	Genome size (Mbp)	Life-style*
Rhizobiales					
	<i>Bartonella bacilliformis</i> KC583	1373	0.38	1.44	OIP
	<i>Bartonella quintana</i> Toulouse	1358	0.39	1.58	OIP
	<i>Bartonella henselae</i> Houston-1	1670	0.38	1.93	OIP
Delta-Proteobacteria					
Desulfovibrionales					
	<i>Lawsonia intracellularis</i> PHE/MN1-00	1493	0.33	1.72	OIP
Bacteroidetes					
Flavobacteriales					
	<i>Candidatus Sulcia muelleri</i> GWSS	262	0.22	0.24	OE
Mollicutes					
Mycoplasmatales					
	<i>Mycoplasma genitalium</i> G37	525	0.32	0.58	OP
	<i>Ureaplasma parvum</i> (2 strains)	643 – 655	0.25	0.75 – 0.75	OP
	<i>Mycoplasma mobile</i> 163K	669	0.25	0.78	OP
	<i>Mycoplasma synoviae</i> 53	732	0.28	0.80	OP
	<i>Mycoplasma pneumoniae</i> M129	733	0.4	0.81	OP
	<i>Mycoplasma agalactiae</i> PG2	791	0.3	0.88	OP
	<i>Mycoplasma hyopneumoniae</i> (3 strains)	714 – 728	0.28–0.29	0.89 – 0.92	OP
	<i>Mycoplasma pulmonis</i> UAB CTIP	820	0.27	0.96	OP
	<i>Mycoplasma gallisepticum</i> R	769	0.31	1.00	OP
	<i>Mycoplasma capricolum capricolum</i> ATCC 27343	866	0.24	1.01	OP
	<i>Mycoplasma mycoides mycoides</i> SC PG1	1061	0.24	1.21	OP
	<i>Mycoplasma penetrans</i> HF-2	1074	0.26	1.36	OIP
Acholeplasmatales					
	Aster yellows witches-broom phytoplasma AYWB	732	0.27	0.72	OIP
	Onion yellows phytoplasma OY-M	796	0.28	0.86	OIP
Entomoplasmatales					
	<i>Mesoplasma florum</i> L1	725	0.27	0.79	HA
Actinobacteria					
Actinomycetales					
	<i>Tropheryma whipplei</i> (2 strains)	840 -864	0.46	~0.93	OP
Spirochaetes					
Spirochaetales					
	<i>Borrelia garinii</i> PBi	979	0.28	0.99	OP
	<i>Treponema pallidum pallidum</i> Nichols	1088	0.53	1.14	OP
	<i>Borrelia afzelii</i> PKo	1256	0.28	1.20	OP

(Table 2). Contd.....

Taxonomy	Species	Gene count	G+C content	Genome size (Mbp)	Life-style*
Chlamydiae					
	Chlamydiales				
	<i>Chlamydia trachomatis</i> (4 strains)	917 – 965	0.41	1.04 – 1.05	OIP
	<i>Chlamydia muridarum</i> Nigg	957	0.4	1.08	OIP
	<i>Chlamydomphila abortus</i> S26/3	999	0.4	1.14	OIP
	<i>Chlamydomphila felis</i> Fe/C-56	1054	0.39	1.17	OIP
	<i>Chlamydomphila caviae</i> GPIC	1049	0.39	1.18	OIP
	<i>Chlamydomphila pneumoniae</i> (4 strains)	1.0 – 1.1 ⁿ	0.41	1.23 – 1.26	OIP
Nanoarchaea					
	Nanoarchaeales				
	<i>Nanoarchaeum equitans</i> Kin4-M	690	0.32	0.49	OS

*OS, obligate symbiont; OE, obligate endosymbiont; OP, obligate parasite; OIP, obligate intracellular parasite; HA, host associated. ⁿIn mega-bases. Life-styles according to [69].

Reduced genomes from free-living prokaryotes share several trends associated to genomes from host-associated bacteria, like low G+C content and accelerated substitution rate. However, the reductive genome process in these cases seems to be the result of streamlining selection for the efficient utilization of nutrient resources [16, 17, 39]. This hypothesis suggests that, in free-living species with very large population sizes, selection acts to reduce genome size because of the metabolic burden of replicating DNA with no adaptive value [17]. This process has been invoked to explain the genome reduction and metabolic specialization in *Prochlorococcus marinus* MED4 and SS120 strains, *Candidatus Pelagibacter ubique* and in the beta-proteobacterium HTCC2181 [16, 17, 39]. In agreement with this hypothesis, in a recent analysis on the genome of *P. marinus* no evidence of strong genetic drift has been detected [40]. It is too early to know if genome reduction in free-living prokaryotes is always mediated by this proposed process of streamlining selection. For the sake of comparison, a different hypothesis has been used to explain genome reduction in the *Nanoarchaeum-Igniococcus* system, where a very active genome recombination process, combined with the peculiar relationship between both organisms (and widely differing in population size to pelagic bacteria), might counteract genome expansion [19].

Genomes from free-living prokaryotes often contain redundant metabolic capacities. Simpler and streamlined genomes might evolve by retaining alternatively redundant metabolic pathways. By modeling the evolution of reduced genomes of endosymbiotic bacteria and controlling for environment, Pál and co-workers [41] were able to obtain several different minimal gene sets due to contingency-dependent loss of alternative pathways. These diverse minimal gene sets were enriched in genes present in naturally evolved endosymbionts. Among sequenced reduced genomes, there are examples that alternative loss of metabolic pathways has indeed taken place. As shown by [42] the one-carbon pool tetrahydrofolate (C₁-THF) metabolism in mycoplasma genomes clearly illustrates this process, where different species

have differentially retained the metabolically redundant glycine hydroxymethyltransferase (GlyA) and formate-dihydrofolate ligase (FthS) enzymes. Therefore, the evolution of reduced genomes can be thought as environmental as well as taxonomically constrained, but depending also on contingent dependent loss of alternative metabolic pathways. The outcome is that, for a given environment and a given species, several different equally-likely minimal genomes are possible.

THE MINIMAL GENE SET AND THE BOUNDARY BETWEEN CELL AND ORGANELLE

How small a genome of an intracellular prokaryote can become without losing the status of a cell and becoming an organelle? The boundary between cells and organelles might appear sometimes fuzzy, as it has been suggested for the photosynthetic body in the filose thecamoeba *Paulinella chromatophora*, where there is a high level of cell-cycle integration between the chromatophore and the amoebae, but the genes retained in the chromatophore genome suggest that this cellular entity is dependent on its host for consumables, although perhaps not for organelle-specific information [43, 44]. However, a good criterion to distinguish among endosymbionts and organelles seems to be whether or not all the functional proteins in the cytosol of the intracellular entity are encoded in its own genome [45]. Incidentally, this criterion was also used when comparing the genomes of *Haemophilus influenzae* and *M. genitalium* to derive a theoretical minimal gene set to sustain life: “however small, a cellular gene set has to be self-sufficient in the sense that cells generally import metabolites and not functional proteins” [33].

Due to the lack of several important functions (like an incomplete set of tRNA synthetases) in the genome of *Candidatus Sulcia muelleri* GWSS [46], and in *Candidatus Carsonella ruddi* PV [47], it has been suggested that some of these functions might be encoded in the nucleus of the host (thus reaching the status of an organelle). This is a possibility that might also be true for the peculiar *Nanoarchaeum-Igniococcus* system [19]. For instance, a detailed analysis of

the genes coded in the genome of *Candidatus Carsonella ruddi* PV showed that the extensive degradation of the genome is not compatible with its consideration as a mutualistic endosymbiont and, even more, as a living organism [48]. In this cellular entity, the genes for the synthesis of three essential amino acids required for the host have been completely lost, as well as several genes to replicate, transcribe and synthesize proteins, thus questioning its consideration as an independent living entity [48]. It remains to be seen to what extent these cells have reached so small genomes at the expenses of transferring some of the genes to the host genome (coupled with targeting back the translated proteins to the endosymbiont), or by gene loss and modification of the remaining genes to accomplish more than one function [47, 46]. While this question remains unclear, using these cells as models of minimal gene sets should take this into consideration.

THE MINIMAL CELL CONCEPT AND THE LAST COMMON ANCESTOR

It has been suggested that the idea behind the minimal cell is that “some set of basic parameters and principles are common to all cellular life on this planet, and the difficulty lies in finding the proper way to strip away the outer layers of complexity in order to uncover what is truly general and therefore satisfactorily describes all cells” [49]. This is, research on minimal cells seeks ultimately for principles common to all life forms. Since all life on Earth is related by having descended from a last common ancestor (LCA), a discussion of the minimal cell concept in relation to the nature of the LCA is unavoidable.

It is reasonable to assume that, at some point in evolution, extant cells were preceded by simpler cellular systems [50]. For instance, it has been suggested that hypothetical protocells might have been composed of an RNA genome coding for a self-replicating replicase and a lipid-synthesizing ribozyme inside a lipid membrane vesicle [51]. Such systems, if ever existed, would have been minimal in the sense that the lack of any of its components would have destroyed the system.

However, as indicated by [28] the LCA was by no means a minimal cell. On the one hand, it should not be confused the experimental search of an engineered minimal cell/genome coding for the sufficient features of life which is taxonomic as well as environmental dependent, with the reconstruction of the genome of the LCA. As suggested by [49], the minimal cell concept is mostly a “*useful tool for organizing the thoughts about gene function in a particular organism in a particular environment*”. Therefore, there can be a plethora of different minimal cells [52] (as illustrated by the 42 genes from the mutation deletion experiment of *M. genitalium* which are unique to the mycoplasmas [53]), but only one particular biology for the LCA. On the other hand, the universality of the genetic code and the high level of conservation of transcription and translation machinery, strongly suggest that the LCA had already evolved the capacity for ribosome mediated biosynthesis of proteins [54] and there is no reason to think that the genome of the LCA coded just for the essential set of proteins for survival (i.e., the lack of any one of its components would not have destroyed the entity).

Nevertheless, there is a clear relationship between the essentialness of a gene and its degree of evolutionary conservation, as shown by the phylogenetic distribution of essential genes identified in the genome of *B. subtilis* [10]. Therefore, some of the genes identified as essential in extant cells, were likely to be present in the genome of the LCA. This incidentally suggests that the complexity, measured in number of genes, of extant naturally-evolved reduced genomes is likely to be internally constrained (due to epistatic interactions) by having descended from such complex universal common ancestor, and externally, by the selective pressure exerted by extant competitors with similar evolved refinements.

Despite the set of highly conserved genes related to transcription and translation, comparative genome analysis performed to identify universally conserved proteins inherited from the LCA of all life, vary in their results depending on the sampled genomes and the methodology used for the back-track characterization [54, 55]. These datasets are, in some occasions, partially overlapping. The set of inferred genes can be as small as 80 universally conserved protein families (50 of them showing three domain phylogenies i.e., archaea, bacteria and eukarya branching separately) [56], or as large as ~1000 genes [57]. By using an algorithm that reconstructs the gene content of the LCA (penalizing horizontal gene transfers and secondary gene losses under parsimony criteria) [58] suggested a LCA with ~600 genes.

Whatever sophisticated the methods to characterize the LCA genome, it is unlikely that a functional cell can be reconstructed solely by defining the set of inferred ancestral genes. Clearly essential functions, like the main polymerase involved in DNA replication, are lacking when looking for universally conserved homologous genes [59, 60]. Non-orthologous gene displacement [61] might be one of the reasons for the lack of conserved genes otherwise expected to be present among a set of universal and conserved genes related to the basic functions of cells. How far the process of non-orthologous gene displacement has gone in hindering early stages of cellular evolution is still an unanswered question. However, it is likely that the same process that hides the past could be used to gain understanding on the basic functions genes have to perform in cells in order to keep them alive. For example, most bacteria possess a cell wall which provides structural stability and protects against osmotic stress. The wall-less *M. genitalium* and *M. pneumoniae* do not code for a single gene involved in cell wall formation, but they possess a cytoskeleton-like as a substitute [62]. Mycoplasmatales very likely evolved from Gram-positive walled bacteria [63]; therefore, the substitution of a cell wall by a cytoskeleton indicates the relevance of the function that is accomplished by two different non-homologous cellular subsystems.

THE MINIMAL CELL CONCEPT AND A DEFINITION OF LIFE

It has been suggested that a definition of a minimal set of genes and/or their associated functions required to sustain cellular life involves ultimately an agreed-upon definition of what is living and what is not [49]. However, organisms are the product of evolution, and their history plays a central role on its phenomenology [64]. As argued by Moreno and co-

workers [65, 66], when defining life (and searching for its necessary and sufficient features), it is imperative to take into account the historical-collective and evolutionary perspective, because evolutionary change due to natural selection acting upon a system replicating with variation is a unique feature of living systems [65, 67]. This notion of life has remarkable implications for minimal cells. As mentioned above, three basic sub-systems of cells are metabolism (M), template based replication (T) and boundary (B) [4]. Under this view of life, a system resulting from the combinatorial of those sub-systems, in order to be considered fully alive, must show an open-ended evolutionary capacity.

CONCLUSIONS

A survey on the *Aristotelian* reality represented by the diversity of cells with small and reduced genomes evolved in nature can help to better delineate the *Platonic* idea of a hypothetical minimal cell. Reduced genomes from free-living organisms should be included, together with those of endosymbionts and parasites, as models of minimal cells. Extant biodiversity suggest that the minimal genome for a free-living cell may contain approximately 10^3 genes. Differing from endosymbionts, streamlining selection might explain the reductive process in free-living prokaryotes. Ultimately, minimal cells must be endowed with the capacity to evolve.

NOTE ADDED AT PROOF

While this manuscript was been reviewed the genome of an α -Proteobacteria named *Candidatus* *Hodgkinia cicadicola* (endosymbiont of the cicada *Diceroprocta semicincta*) was published [68]. With 169 protein coding genes and a surprisingly G+C content of 58% is the endosymbiont with the smallest genome.

ACKNOWLEDGEMENT

This work was supported by a grant from the European Union, Seven Framework Program, project TARPOL number 212894, to AM; grant BFU2006-06003/BMC (Ministerio de Educación y Ciencia, Spain) to AL, and GV/2007/050 (Generalitat Valenciana, Spain) to RG.

REFERENCES

- [1] Szathmáry, E. In search of the simplest cell. *Nature*, **2005**, *433*, 469-470.
- [2] Peretó, J.; Catalá, J. The renaissance of synthetic biology. *Biol. Theor.*, **2007**, *2*, 128-130.
- [3] Szathmáry, E. Coevolution of metabolic networks and membranes: the scenario of progressive sequestration. *Phil. Trans. R. Soc. B*, **2007**, *362*, 1781-1787.
- [4] Fernando, C.; Santos, M.; Szathmáry, E. Evolutionary potential and requirements for minimal protocells. *Top. Curr. Chem.*, **2005**, *259*, 167-211.
- [5] Mansy, S.S.; Schrum, J.P.; Krishnamurthy, M.; Tobé, S.; Treco, D.A.; Szostak, J.W. Template-directed synthesis of a genetic polymer in a model protocell. *Nature*, **2008**, *454*, 122-125.
- [6] Oparin, A.I. (1924) *Proiskhozhedenie Zhizni*. Moscow: Moscovskii Rabotchii. In *The Origin of Life*; Deamer, D.W.; Fleischaker, G.R. Eds.; The Central Concepts. Bartlett and Jones: Boston, **1994**.
- [7] Fehér, T.; Papp, B.; Pál, C.; Pósfai, G. Systematic genome reductions: theoretical and experimental approaches. *Chem. Rev.*, **2007**, *107*, 3498-3513.
- [8] Klasson, L.; Andersson, S.G. Evolution of minimal-gene-sets in host-dependent bacteria. *Trends Microbiol.*, **2004**, *12*, 37-43.
- [9] Gerdes, S.Y.; Scholle, M.D.; Campbell, J.W.; Balázsi, G.; Ravasz, E.; Daugherty, M.D.; Somera, A.L.; Kyrpides, N.C.; Anderson, I.; Gelfand, M.S.; Bhattacharya, A.; Kapatral, V.; D'Souza, M.; Baev, M.V.; Grechkin, Y.; Meeh, F.; Fonstein, M.Y.; Overbeek, R.; Ba-
- [10] rabási, A.L.; Oltvai, Z.N.; Osterman, A.L. Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655. *J. Bacteriol.*, **2003**, *185*, 5673-5684.
- [11] Kobayashi, K.; Ehrlich, S.D.; Albertini, A.; Amati, G.; Andersen, K.K.; Arnaud, M.; Asai, K.; Ashikaga, S.; Aymerich, S.; Bessieres, P.; Boland, F.; Brignell, S.C.; Bron, S.; Bunai, K.; Chapuis, J.; Christiansen, L.C.; Danchin, A.; Débarbouille, M.; Dervyn, E.; Deuerling, E.; Devine, K.; Devine, S.K.; Dreesen, O.; Errington, J.; Fillinger, S.; Foster, S.J.; Fujita, Y.; Galizzi, A.; Gardan, R.; Eschevins, C.; Fukushima, T.; Haga, K.; Harwood, C.R.; Hecker, M.; Hosoya, D.; Hullo, M.F.; Kakeshita, H.; Karamata, D.; Kasahara, Y.; Kawamura, F.; Koga, K.; Koski, P.; Kuwana, R.; Imamura, D.; Ishimaru, M.; Ishikawa, S.; Ishio, I.; Le Coq, D.; Masson, A.; Mauël, C.; Meima, R.; Mellado, R.P.; Moir, A.; Moriya, S.; Nagakawa, E.; Nanamiya, H.; Nakai, S.; Nygaard, P.; Ogura, M.; Ohanan, T.; O'Reilly, M.; O'Rourke, M.; Pragai, Z.; Pooley, H.M.; Rapoport, G.; Rawlins, J.P.; Rivas, L.A.; Rivolta, C.; Sadaie, A.; Sadaie, Y.; Sarvas, M.; Sato, T.; Saxild, H.H.; Scanlan, E.; Schumann, W.; Seegers, J.F.; Sekiguchi, J.; Sekowska, A.; Séror, S.J.; Simon, M.; Stragier, P.; Studer, R.; Takamatsu, H.; Tanaka, T.; Takeuchi, M.; Thomaidis, H.B.; Vagner, V.; van Dijl, J.M.; Watabe, K.; Wipat, A.; Yamamoto, H.; Yamamoto, M.; Yamamoto, Y.; Yamane, K.; Yata, K.; Yoshida, K.; Yoshikawa, H.; Zuber, U.; Ogasawara, N. Essential *Bacillus subtilis* genes. *Proc. Natl. Acad. Sci. USA*, **2003**, *100*, 4678-4683.
- [12] Glass, J.I.; Assad-Garcia, N.; Alperovich, N.; Yooseph, S.; Lewis, M.R.; Maruf, M.; Hutchison, C.A. 3rd; Smith, H.O.; Venter, J.C. Essential genes of a minimal bacterium. *Proc. Natl. Acad. Sci. USA*, **2006**, *103*, 425-430.
- [13] Pósfai, G.; Plunkett, G.3rd; Fehér, T.; Frisch, D.; Keil, G.M.; Umenhoffer, K.; Kolisnichenko, V.; Stahl, B.; Sharma, S.S.; de Arruda, M.; Burland, V.; Harcum, S.W.; Blattner, F.R. Emergent properties of reduced-genome *Escherichia coli*. *Science*, **2006**, *312*, 1044-1046.
- [14] Gibson, D.G.; Benders, G.A.; Andrews-Pfannkoch, C.; Denisova, E.A.; Baden-Tillson, H.; Zaveri, J.; Stockwell, T.B.; Brownley, A.; Thomas, D.W.; Algire, M.A.; Merryman, C.; Young, L.; Noskov, V.N.; Glass, J.I.; Venter, J.C.; Hutchison, C.A. 3rd; Smith, H.O. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science*, **2008**, *319*, 1215-1220.
- [15] Gil, R.; Silva, F.J.; Peretó J.; Moya, A. Determination of the core of a minimal bacterial gene set. *Microbiol. Mol. Biol. Rev.*, **2004**, *68*, 518-537.
- [16] Morowitz, H. *Beginnings of Cellular Life*, Yale University Press: New Haven and London, **1992**.
- [17] Giovannoni, S.J.; Hayakawa, D.H.; Tripp, H.J.; Stingl, U.; Givan, S.A.; Cho, J.C.; Oh, H.M.; Kitner, J.B.; Vergin, K.L.; Rappé, M.S. The small genome of an abundant coastal ocean methylotroph. *Environ. Microbiol.*, **2008**, *10*, 1771-1782.
- [18] Giovannoni, S.J.; Tripp, H.J.; Givan, S.; Podar, M.; Vergin, K.L.; Baptista, D.; Bibbs, L.; Eads, J.; Richardson, T.H.; Noordewier, M.; Rappé, M.S.; Short, J.M.; Carrington, J.C.; Mathur, E.J. Genome streamlining in a cosmopolitan oceanic bacterium. *Science*, **2005**, *309*, 1242-1245.
- [19] He, J.; Ritalahti, K.M.; Yang, K.L.; Koenigsberg, S.S.; Löffler, F.E. Detoxification of vinyl chloride to ethene coupled to growth of an anaerobic bacterium. *Nature*, **2003**, *424*, 62-65.
- [20] Podar, M.; Anderson, I.; Makarova, K.S.; Elkins, J.G.; Ivanova, N.; Wall, M.A.; Lykidis, A.; Mavromatis, K.; Sun, H.; Hudson, M.E.; Chen, W.; Deciu, C.; Hutchison, D.; Eads, J.R.; Anderson, A.; Fernandes, F.; Szeto, E.; Lapidus, A.; Kyrpides, N.C.; Saier, M.H.Jr.; Richardson, P.M.; Rachel, R.; Huber, H.; Eisen, J.A.; Koonin, E.V.; Keller, M.; Stetter, K.O. A genomic analysis of the archaeal system *Ignicoccus hospitalis*-*Nanoarchaeum equitans*. *Genome Biol.*, **2008**, *9*, R158.
- [21] Ketter, G.C.; Martiny, A.C.; Huang, K.; Zucker, J.; Coleman, M.L.; Rodrigue, S.; Chen, F.; Lapidus, A.; Ferriera, S.; Johnson, J.; Steglich, C.; Church, G.M.; Richardson, P.; Chisholm, S.W.; Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*. *PLoS Genet.*, **2007**, *3*, e231.
- [22] Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Overbeek, R.; Snead, M.A.; Keller, M.; Au-jay, M.; Huber, R.; Feldman, R.A.; Short, J.M.; Olsen, G.J.; Swanson, R.V. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature*, **1998**, *392*, 353-358.

- [22] Kendall, M.M.; Liu, Y.; Sieprawska-Lupa, M.; Stetter, K.O.; Whitman, W.B.; Boone, D.R. *Methanococcus aeolicus* sp. nov., a mesophilic, methanogenic archaeon from shallow and deep marine sediments. *Int. J. Syst. Evol. Microbiol.*, **2006**, *56*, 1525-1529.
- [23] Fütterer, O.; Angelov, A.; Liesegang, H.; Gottschalk, G.; Schleper, C.; Schepers, B.; Dock, C.; Antranikian, G.; Liebl, W. Genome sequence of *Picrophilus torridus* and its implications for life around pH 0. *Proc. Natl. Acad. Sci. USA*, **2004**, *101*, 9091-9096.
- [24] Slesarev, A.I.; Mezhevaya, K.V.; Makarova, K.S.; Polushin, N.N.; Shcherbinina, O.V.; Shakhova, V.V.; Belova, G.I.; Aravind, L.; Natale, D.A.; Rogozin, I.B.; Tatusov, R.L.; Wolf, Y.I.; Stetter, K.O.; Malykh, A.G.; Koonin, E.V.; Kozyavkin, S.A. The complete genome of hyperthermophile *Methanopyrus kandleri* AV19 and monophyly of archaeal methanogens. *Proc. Natl. Acad. Sci. USA*, **2002**, *99*, 4644-4649.
- [25] Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.; Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.; Lumm, W.; Pothier, B.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Caruso, A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.; McDougall, S.; Shimer, G.; Goyal, A.; Pietrkowski, S.; Church, G.M.; Daniels, C.H.J.; Mao, J.-I.; Rice, P.; Nölling, J.; Reeve, J.N. Complete genome sequence of *Methanobacterium thermoautotrophicum* deltaH: functional analysis and comparative genomics. *J. Bacteriol.*, **1997**, *179*, 7135-7155.
- [26] Elkins, J.G.; Podar, M.; Graham, D.E.; Makarova, K.S.; Wolf, Y.; Randau, L.; Hedlund, B.P.; Brochier-Armanet, C.; Kunin, V.; Anderson, I.; Lapidus, A.; Goltsman, E.; Barry, K.; Koonin, E.V.; Hugenholtz, P.; Kyrpides, N.; Wanner, G.; Richardson, P.; Keller, M.; Stetter, K.O. A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc. Natl. Acad. Sci. USA*, **2008**, *105*, 8102-8107.
- [27] Könneke, M.; Bernhard, A.E.; de la Torre, J.R.; Walker, C.B.; Waterbury, J.B.; Sthal, D.A. Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature*, **2005**, *437*, 543-546.
- [28] Islas, S.; Becerra, A.; Luisi, P.L.; Lazcano, A. Comparative genomics and the gene complement of a minimal cell. *Orig. Life Evol. Biosph.*, **2004**, *34*, 243-256.
- [29] Oshima, K.; Kakizawa, S.; Nishigawa, H.; Jung, H.Y.; Wei, W.; Suzuki, S.; Arashida, R.; Nakata, D.; Miyata, S.; Ugaki, M.; Namba, S. Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. *Nat. Genet.*, **2004**, *36*, 27-29.
- [30] Perez-Brocail, V.; Gil, R.; Ramos, S.; Lamelas, A.; Postigo, M.; Michelena, J.M.; Silva, F.J.; Moya, A.; Latorre, A. A small microbial genome: the end of a long symbiotic relationship? *Science*, **2006**, *314*, 312-313.
- [31] Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H. The complete sequence of the mucosal pathogen *Ureaplasma urealyticum*. *Nature*, **2000**, *407*, 757-762.
- [32] Smith, D.G.; Russell, W.C.; Ingledew, W.J.; Thirkell, D. Hydrolysis of urea by *Ureaplasma urealyticum* generates a transmembrane potential with resultant ATP synthesis. *J. Bacteriol.*, **1993**, *175*, 3253-3258.
- [33] Mushegian, A.R.; Koonin, E.V. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 10268-10273.
- [34] Westberg, J.; Persson, A.; Holmberg, A.; Goesmann, A.; Lundeberg, J.; Johansson, K.E.; Pettersson, B.; Uhlén, M. The genome sequence of *Mycoplasma mycoides* subsp. *mycoides* SC type strain PG1T, the causative agent of contagious bovine pleuropneumonia (CBPP). *Genome Res.*, **2004**, *14*, 221-227.
- [35] Kuwahara, H.; Yoshida, T.; Takaki, Y.; Shimamura, S.; Nishi, S.; Harada, M.; Matsuyama, K.; Takishita, K.; Kawato, M.; Uematsu, K.; Fujiwara, Y.; Sato, T.; Kato, C.; Kitagawa, M.; Kato, I.; Maruyama, T. Reduced genome of the thioautotrophic intracellular symbiont in a deep-sea clam, *Calyptogena okutanii*. *Curr. Biol.*, **2007**, *17*, 881-886.
- [36] Moya, A.; Peretó, J.; Gil, R.; Latorre, A. Learning how to live together: genomic insights into prokaryote-animal symbioses. *Nat. Rev. Genet.*, **2008**, *9*, 218-229.
- [37] Moran, N.A. Microbial minimalism: genome reduction in bacterial pathogens. *Cell*, **2002**, *108*, 583-586.
- [38] Delmotte, F.; Rispé, C.; Schaber, J.; Silva, F.J.; Moya, A. Tempo and mode of early gene loss in endosymbiotic bacteria from insects. *BMC Evol. Biol.*, **2006**, *6*, 56.
- [39] Dufresne, A.; Ostrowski, M.; Scanlan, D.J.; Garczarek, L.; Mazard, S.; Palenik, B.P.; Paulsen, I.T.; de Marsac, N.T.; Wincker, P.; Dossat, C.; Ferreira, S.; Johnson, J.; Post, A.F.; Hess, W.R.; Partensky, F. Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol.*, **2008**, *9*(5), R90.
- [40] Hu, J.; Blanchard, J.L. Environmental sequence data from the Sargasso Sea reveal that the characteristics of genome reduction in *Prochlorococcus* are not a harbinger for an escalation in genetic drift. *Mol. Biol. Evol.*, **2009**, *26*, 5-13.
- [41] Pál, C.; Papp, B.; Lercher, M.J.; Csermely, P.; Oliver, S.G.; Hurst, L.D. Chance and necessity in the evolution of minimal metabolic networks. *Nature*, **2006**, *440*, 667-670.
- [42] Vasconcelos, A.T.; Ferreira, H.B.; Bizarro, C.V.; Bonatto, S.L.; Carvalho, M.O.; Pinto, P.M.; Almeida, D.F.; Almeida, L.G.; Almeida, R.; Alves-Filho, L.; Assunção, E.N.; Azevedo, V.A.; Bogo, M.R.; Brigido, M.M.; Brocchi, M.; Burity, H.A.; Camargo, A.A.; Camargo, S.S.; Carepo, M.S.; Carraro, D.M.; de Mattos Cascardo, J.C.; Castro, L.A.; Cavalcanti, G.; Chemale, G.; Collevatti, R.G.; Cunha, C.W.; Dallagiovanna, B.; Dambrós, B.P.; Dellagostin, O.A.; Falcão, C.; Fantinatti-Garborggini, F.; Felipe, M.S.; Fiorentin, L.; Franco, G.R.; Freitas, N.S.; Frías, D.; Grangeiro, T.B.; Grisard, E.C.; Guimarães, C.T.; Hungria, M.; Jardim, S.N.; Krieger, M.A.; Laurino, J.P.; Lima, L.F.; Lopes, M.I.; Loreto, E.L.; Madeira, H.M.; Manfio, G.P.; Maranhão, A.Q.; Martinkovics, C.T.; Medeiros, S.R.; Moreira, M.A.; Neiva, M.; Ramalho-Neto, C.E.; Nicolás, M.F.; Oliveira, S.C.; Paixão, R.F.; Pedrosa, F.O.; Pena, S.D.; Pereira, M.; Pereira-Ferrari, L.; Piffer, I.; Pinto, L.S.; Potrich, D.P.; Salim, A.C.; Santos, F.R.; Schmitt, R.; Schneider, M.P.; Schrank, A.; Schrank, I.S.; Schuck, A.F.; Seunaez, H.N.; Silva, D.W.; Silva, R.; Silva, S.C.; Soares, C.M.; Souza, K.R.; Souza, R.C.; Staats, C.C.; Steffens, M.B.; Teixeira, S.M.; Urmenyi, T.P.; Vainstein, M.H.; Zuccherato, L.W.; Simpson, A.J.; Zaha, A. Swine and poultry pathogens: the complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae*. *J. Bacteriol.*, **2005**, *187*, 5568-5577.
- [43] Keeling, P.J.; Archibald, J.M. Organelle evolution: what's in a name? *Curr. Biol.*, **2008**, *18*, R345-247.
- [44] Nowack, E.C.; Melkonian, M.; Glöckner, G. Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Curr. Biol.*, **2008**, *18*, 410-418.
- [45] Theissen, U.; Martin, W. The difference between organelles and endosymbionts. *Curr. Biol.*, **2006**, *16*, R1016-R1017.
- [46] McCutcheon, J.P.; Moran, N.A. Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. *Proc. Natl. Acad. Sci. USA*, **2007**, *104*, 19392-19397.
- [47] Nakabachi, A.; Yamashita, A.; Toh, H.; Ishikawa, H.; Dunbar, H.E.; Moran, N.A.; Hattori, M. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. *Science*, **2006**, *314*, 267.
- [48] Tamames, J.; Gil, R.; Latorre, A.; Peretó, J.; Silva, F.J.; Moya, A. The frontier between a cell and an organelle: genome analysis of *Candidatus Carsonella ruddi*. *BMC Evol. Biol.*, **2007**, *7*, 181.
- [49] Peterson, S.N.; Fraser, C.M. The complexity of simplicity. *Genome Biol.*, **2001**, *2*, comment 2002.1-2002.8.
- [50] Joyce, G.F.; The antiquity of RNA-based evolution. *Nature*, **2002**, *418*, 214-221.
- [51] Szostak, J.W.; Bartel, D.P.; Luisi, P.L. Synthesizing life. *Nature*, **2001**, *409*, 387-390.
- [52] Huynen, M. Constructing a minimal genome. *TIG*, **2000**, *16*, 116.
- [53] Hutchison, C.A.; Peterson, S.N.; Gill, S.R.; Cline, R.T.; White, O.; Fraser, C.M.; Smith, H.O.; Venter, J.C. Global transposon mutagenesis and a minimal *Mycoplasma* genome. *Science*, **1999**, *286*, 2165-2169.
- [54] Becerra, A.; Delaye, L.; Islas, S.; Lazcano, A. The very early stages of biological evolution and the nature of the last common ancestor of the three major cell domains. *Annu. Rev. Ecol. Evol. Syst.*, **2007**, *38*, 361-379.
- [55] Koonin, E.V. Comparative genomics, minimal gene-sets and the last universal common ancestor. *Nat. Rev. Microbiol.*, **2003**, *1*, 127-136.
- [56] Harris, J.K.; Kelley, S.T.; Spiegelman, G.B.; Pace, N.R. The genetic core of the universal ancestor. *Genome Res.*, **2003**, *13*, 407-412.
- [57] Ouzounis, C.A.; Kunin, V.; Darzentas, N.; Goldovsky, L. A minimal estimate for the gene content of the last universal common ancestor--exobiology from a terrestrial perspective. *Res. Microbiol.*, **2006**, *157*, 57-68.

- [58] Mirkin, B.G.; Fenner, T.I.; Galperin, M.Y.; Koonin, E.V. Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes. *BMC Evol. Biol.*, **2003**, *6*, 3:2.
- [59] Leipe, D.D.; Aravind, L.; Koonin, E.V. Did DNA replication evolve twice independently? *Nucl. Acids Res.*, **1999**, *27*, 3389-3401.
- [60] Delaye, L.; Becerra, A.; Lazcano, A. The nature of the last common ancestor. In *The genetic code and the origin of life.*; Lluís Ribas de Pouplana, Ed.; Landes Bioscience and Kluwer Academic: USA, **2004**; pp. 34-47.
- [61] Koonin, E.V.; Mushegian, A.R.; Bork, P. Non-orthologous gene displacement. *TIG*, **1996**, *12*, 334-336.
- [62] Himmelreich, R.; Plagens, H.; Hilbert, H.; Reiner, B.; Herrmann, R. Comparative analysis of the genomes of the bacteria *Mycoplasma pneumoniae* and *Mycoplasma genitalium*. *Nucl. Acids Res.*, **1997**, *25*:701-712.
- [63] Woese, C.R.; Maniloff, J.; Zablen, L.B. Phylogenetic analysis of the mycoplasmas. *Proc. Natl. Acad. Sci. USA*, **1980**, *77*, 494-498.
- [64] Gould, S.J. "*Wonderful Life: The Burgess Shale and the Nature of History*", W.W. Norton & Co.: USA, **1989**.
- [65] Ruiz-Mirazo, K.; Peretó, J.; Moreno, A. A universal definition of life: autonomy and open-ended evolution. *Orig. Life Evol. Biosph.*, **2004**, *34*, 323-346.
- [66] Ruiz-Mirazo, K.; Umerez, J.; Moreno, A. Enabling conditions for open-ended evolution. *Biol. Philos.*, **2008**, *23*, 67-85.
- [67] Lazcano, A. What is life? a brief historical overview. *Chem. Biodivers.*, **2008**, *5*, 1-15.
- [68] McCutcheon, J.P.; McDonald, B.R.; and Moran, N.A. Origin of an alternative genetic code in the extremely small and GC-rich genome of a bacterial symbiont. *PLoS Genet.*, **2009**, *5*(7), e1000565.
- [69] Markowitz, V.M.; Szeto, E.; Palaniappan, K.; Grechkin, Y.; Chu, K.; Chen, I.M.; Dubchak, I.; Anderson, I.; Lykidis, A.; Mavromatis, K.; Ivanova, N.N.; Kyrpides, N.C. The integrated microbial genomes (IMG) system in 2007: data content and analysis tool extensions. *Nucl. Acids Res.*, **2008**, *36*, D528-533.

Received: November 30, 2009

Revised: December 28, 2009

Accepted: January 24, 2010

© Delaye et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.