Reductive evolution of resident genomes

Siv G.E. Andersson and Charles G. Kurland

The genomes of obligate parasites, endosymbionts and cellular organelles evolve under conditions that are radically different from those of free-living organisms. One distinguishing factor is that they are residents of an environment that is conditioned by a host genome. For example, it is possible to imagine an evolutionary sequence in which a bacterium begins as a facultative resident within a cellular domain. During the course of adaptation to the intracellular environment, the bacterium can take one of two alternative evolutionary routes. On the one hand, the host cell can become dependent on products provided by the activities of the bacterial genome. In this case, the fates of the cell and the resident are linked through a symbiotic dependence that is progressively strengthened by the tendency to lose gene functions from one or other genome as a consequence of the redundancy of their overlapping genomic functions. In the extreme, the two genomes can evolve a relationship as profound as that between the eukaryote nucleus and the mitochondria. Here, the symbiosis initiated between the primitive eukaryote nucleus and an aerobic bacterium has enabled the eukaryote to evolve an aerobic lifestyle. Gene content is markedly reduced in mtDNAs, compared with those of their eubacterial ancestors, and some mitochondrial genomes have undergone so much reduction in size that it is progressive in asexual viral populations if the rate of deleterious substitutions and deletions in an irreversible manner, which in the long-term will lead to mutational meltdown and genome decay. Here, we discuss the influence of reductive processes on the evolution of genomes that replicate within the domain of a host genome.

Small, asexual populations are expected to accumulate deleterious substitutions and deletions in an irreversible manner, which is progressively strengthened by the tendency to lose gene functions from one or other genome as a consequence of the redundancy of their overlapping genomic functions. In the extreme, the two genomes can evolve a relationship as profound as that between the eukaryote nucleus and the mitochondria. Here, the symbiosis initiated between the primitive eukaryote nucleus and an aerobic bacterium has enabled the eukaryote to evolve an aerobic lifestyle. Gene content is markedly reduced in mtDNAs, compared with those of their eubacterial ancestors, and some mitochondrial genomes have undergone so much reduction in size that it is progressively strengthened by the tendency to lose gene functions from one or other genome as a consequence of the redundancy of their overlapping genomic functions. In the extreme, the two genomes can evolve a relationship as profound as that between the eukaryote nucleus and the mitochondria. Here, the symbiosis initiated between the primitive eukaryote nucleus and an aerobic bacterium has enabled the eukaryote to evolve an aerobic lifestyle. Gene content is markedly reduced in mtDNAs, compared with those of their eubacterial ancestors, and some mitochondrial genomes have undergone so much reduction in size that it is progressively strengthened by the tendency to lose gene functions from one or other genome as a consequence of the redundancy of their overlapping genomic functions.

Despite the differences in evolutionary dynamics, the transmission of both types of resident genomes from one host to the next frequently involves bottlenecks and replication in small populations, with little opportunity for recombination between variants. For this reason, the resident genomes accumulate deleterious mutations at a rate that is higher than that for free-living organisms. Some of these deleterious mutations lead to the loss of coding sequences, while others lead to a marked variability of resident genome architectures. This tendency of small asexual populations to accumulate deleterious mutations is referred to as Muller’s ratchet (Box 1).

The validity of Muller’s ratchet has been investigated in RNA viruses, which have extraordinarily high mutation rates and are subjected to recurrent bottlenecks. Co-infection provides opportunities for genetic exchange through recombination or segmentation in some viral lineages, while others reproduce in a strictly asexual manner. When populations of asexual and sexual viruses are forced through a series of bottlenecks, deleterious mutations accumulate by genetic drift, resulting in a loss of fitness in both populations. However, while this loss can be reversed in sexual viruses by recombination and/or segmentation during intervening periods of mass selection in large populations, loss of fitness is effectively irreversible in asexual viral populations if the rate of compensatory back-mutations is low.

Because of their obligate intracellular lifestyles, there are few or no opportunities for resident genomes of bacteria and organelles to recover from the fitness decline caused by genetic drift during transmission from one host to the other. Indeed, mitochondrial genomes with low recombination frequencies undergo higher fixation rates for new mutations than their sexually reproducing host genomes. Thus, we may wonder if some lineages of resident genomes, such as those of obligate intracellular parasites, are driven to extinction because of the effects of Muller’s ratchet.

Mitochondria and their codes

The most extreme examples of reductively evolved resident genomes are those of the cellular organelles. Phylogenetic reconstructions suggest that mitochondrial genomes
REVIEWS

Box 1. Muller’s ratchet

Identifying the evolutionary forces that drive the fixation rates for mutations is an important task of evolutionary biology. Nonsynonymous substitutions consist of synonymous substitutions that cause amino acid replacements. It is anticipated that some nonsynonymous substitutions will have a deleterious or slightly deleterious effect on fitness. A gradual accumulation of deleterious mutations can induce an irreversible loss of the most-fit class of organisms from the population1–10. This loss can be triggered by high mutation rates, lack of recombination and/or small population sizes. Under these conditions, genetic drift can result in the successive loss of the most-fit class, in a ratchet-like manner. Such a gradual decline in the fitness of a population is referred to as Muller’s ratchet11,12.

However, most nonsynonymous mutations are unlikely to be fixed in the population. Indeed, Muller’s ratchet can be opposed by high rates of compensatory mutations, sexual reproduction and/or by large population sizes driven by purifying selection. Because resident genomes are effectively asexual and often subjected to bottlenecks during each transmission from one host to the other, the loss of fitness induced by genetic drift will not be slowed down to the same extent as in large, free-living bacterial populations. Thus, abnormally higher fixation rates for nonsynonymous substitutions, as observed for intra-cellularly replicating bacteria, are an indicator that Muller’s ratchet is operative in these lineages.

genomes derive from much larger genomes. In particular, they appear to be most closely related to the α-proteobacteria and more specifically to an ancestor of the Rickettsiaceae10–13. Some mitochondrial genomes have retained <1% of the gene repertoire of modern bacteria. This marked reduction in genome size can be partly explained by a massive transfer of genes from the mitochondrial genome to the nuclear genome14, a transfer that began long as 1–2 billion years ago when the ancestral bacterium began to establish its symbiotic relationship with the eukaryotes15. Nevertheless, examples of transfer events from mitochondria to nuclear genomes as recently as within the past 200 million years have been detected16,17, suggesting that the process might be continuing.

A eukaryotic cell can contain hundreds of mitochondria and thousands of mitochondrial genomes. These genomes are maternally inherited through the cytoplasm of the egg with little or no paternal contribution. This mode of transmission effectively restricts the exchange of genetic material between mitochondrial genomes; in effect, the egg is a bottleneck for mitochondrial genomes. Consequently, mitochondria from multicellular organisms display characteristics that are generally associated with small, asexual populations of resident genomes18,19. They are the victims of drift and Muller’s ratchet, as well as the neutralizing effect of the host genome on redundant genes.

Animal mitochondria have diverged so far from their free-living bacterial ancestors that they have evolved new codes and a correspondingly specialized transfer RNA (tRNA) ensemble. If the universal genetic code were to be translated unambiguously into the canonical 20 amino acids, an absolute minimal set of 24 tRNA species would be required to translate the code. One tRNA species is required for each amino acid that is coded by a single codon or by a two- or four-codon box (16 in total). Two tRNA species are required for each amino acid encoded by the three-codon box for isoleucine and by the six-codon boxes for leucine, serine and arginine. In this minimal set, the tRNA species would have a greater redundancy than is normally seen; for example, each of the four codon readers would be indifferent to the nucleotide in codon position three. In some animal mitochondria, the number of tRNA genes has decreased below the canonical minimum of 24. Thus, there may be as few as 22 tRNA species20. The two missing tRNAs in these systems correspond to an isoleucine-tRNA normally cognate to the codon AAA, and an arginine-tRNA normally cognate to the codons AGA and AGG. In addition, AAA is read by a methionine-tRNA that also translates AUG. The codons AGA and AGG now appear as termination codons or are translated by a serine-tRNA that reads the AGN box. The codon reassignments in these systems correspond to an isoleucine-tRNA normally cognate to the codons AGA and AGG.
Positions (S.G.E. Andersson, unpublished). Indeed, for selective constraints on synonymous third-codon variety of genes from gene evolves at a rate of 1.3–2.5 for five species of coding for elongation factor Tu has been carried out the endosymbionts. For example, analysis of the gene can, in turn, be used to calculate rates of evolution for ferred to the bacterial subtree. These dated nodes can the dated nodes in the aphid subtree can be trans- aphids. By assuming that the symbionts are co-specific, the hosts have diverged from their respective ancestors at support for the idea that the endosymbionts and their constructions based on rRNA genes provide strong the symbiotic relationship is mutual. Phylogenetic re- bionts outside their aphid hosts, which suggests that it has not been possible to cultivate the endosym- with vital amino acids. Therefore, antibiotic treat- mission of mitochondria from one host generation tomitted via the egg in a manner resembling the trans- cells called bacteriocytes that are maternally trans- the same time and have co-speciated22. Indeed, the hosts22–25. These bacteria grow inside specialized which live as symbionts of Buchnera, which live as symbionts of E. coli. Buchnera provide no evidence Buchnera provides no evidence for selective constraints on synonymous third-codon positions (S.G.E. Andersson, unpublished). Indeed, synonymous substitution rates appear to be roughly similar among genes with different expression levels, in support of their neutrality26,27. Similar results have been obtained from the obligate intracellular parasite Rickettsia prowazekii, where selection does not appear to influence base composition patterns at synonymous third-codon positions26. Also, synonymous substitution rates appear to be similar among genes in this species. This suggests that the genomes of Buchnera and Rickettsia are under a randomizing mutational pressure from Muller’s ratchet, which overwhelms the weak selection pressure for preferred synonymous codons. Thus, for both of these species, synonymous rates should reflect the neutral, intrinsic mutation rate.

By contrast, measurable selective pressures for sub- sets of synonymous codons characterize free-living organisms, such as Escherichia coli and Salmonella typhimurium, and these are correlated to the abundance of the corresponding isoacceptor rRNAs26,27. In these species, highly expressed genes have stronger codon-usage biases and lower synonymous substitution rates than less-expressed genes. However, the mutual divergence between E. coli and S. typhimurium has been calculated26, and the resulting estimates are remarkably similar to the observed synonymous substitution rate in Buchnera26. This suggests that the intrinsic mutation rate may be very similar in Buchnera, E. coli, and S. typhimurium. By contrast, the substitution rates at nonsynonymous sites differ by more than a factor of ten in these species26,27. Consequently, amino acid replacements occur much faster in Buchnera than in E. coli and S. typhimurium (Table 1). These findings suggest that resident genomes, such as those of organelles, endosymbionts and obligate intracellular parasites, endure an enhanced rate of deleterious substitution26,27. This is a verification of the influence of drift and Muller’s ratchet on their evolution.

### Small, scrambled genomes

Obligate intracellular parasites, such as Rickettsia, Chlamydia and Coxiella, have genome sizes in the 1-Mbp range26. The free-living bacteria from which these parasites have evolved probably had genomes that were 4–5 times larger19,20. Such genomic shrinkage is not a random process, in the sense that any and all genes may be deleted from the genome: it is potentially less damaging to a parasite to lose a gene that is

---

**Table 1. Mutation and substitution rates in Escherichia coli, Salmonella typhimurium and Buchnera sp.**

<table>
<thead>
<tr>
<th>Species</th>
<th>Lifestyle</th>
<th>Tuf Estimated mutation rate (× 10⁻⁹)</th>
<th>Tuf Synonymous rate (× 10⁻⁹)</th>
<th>Tuf Nonsynonymous rate (× 10⁻⁹)</th>
<th>TopB Estimated mutation rate (× 10⁻⁹)</th>
<th>TopB Synonymous rate (× 10⁻⁹)</th>
<th>TopB Nonsynonymous rate (× 10⁻⁹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buchnera sp.</td>
<td>Intracellular</td>
<td>4–8</td>
<td>4–8</td>
<td>0.1–0.2</td>
<td>6–12</td>
<td>0.1–0.2</td>
<td>6–12</td>
</tr>
<tr>
<td>E. coli/S. typhimurium</td>
<td>Free-living</td>
<td>4–5</td>
<td>2–0.3</td>
<td>&lt;0.02</td>
<td>9–12</td>
<td>3–4</td>
<td>0.04–0.05</td>
</tr>
</tbody>
</table>

*Data taken from Ref. 26.

*Number of mutations per position per year.

*Number of substitutions per position per year.
also present in the host genome. Therefore, it is perhaps not surprising that many genes coding for proteins involved in amino acid biosynthesis have been discarded from the genomes of R. prowazekii and Mycoplasma genitalium\(^{33,34}\). Not unexpectedly, these bacterial genomes also have a very limited set of regulatory genes\(^{35}\). The finding that a gene coding for 3-adenosylmethionine synthetase is in different stages of mutational meltdown in a number of Rickettsia species suggests that elimination of biosynthetic genes is an ongoing process in Rickettsia (J.O. Andersson and S.G.E. Andersson, unpublished). A consequence of these reductions in the metabolic repertoire of intracellular and cell-surface parasites is that a relatively larger fraction of the genome is devoted to basic genetic functions, such as replication, transcription and translation.

Why should the depletions of Muller’s ratchet generate smaller genomes? The principle mechanism for excision and insertion of genomic sequences in bacteria seems to involve intrachromosomal recombination at repeated sequences\(^{19,41}\). Deviations from the ancestral gene-order structure are associated with bacterial genomes in the 1-Mbp range, having only a single or two copies of each of the rRNA genes\(^{38,39}\). Because there is a very small probability that 30–40 genes have become linked independently of each other, it is generally believed that these shuffled structures represent an ancient conserved genomic motif\(^{35}\). Indeed, remnants of the super-ribosomal protein gene operon have been observed in plastid genomes, as well as in the mitochondrial genome of the freshwater protozoan Reclinomonas americana\(^{36}\), suggesting that these genes were clustered in an ancestral bacterial genome prior to the divergence of mitochondria and chloroplasts from their bacterial ancestors.

Despite the unusually high degree of gene-order conservation in the ribosomal protein gene operons, we have identified both inversion and deletion events in this area of the R. prowazekii genome\(^{40,41}\). We suggest that an intrachromosomal recombination event between the two ancestral tuf genes resulted in a major inversion of the genome, after which one of the two copies was deleted\(^{40,41}\). Deletions from the ancient organization of these genes have also been detected in the tiny genome of M. genitalium. For this species, numerous rearrangement events have to be inferred to explain the distribution of this set of genes into the six different gene blocks observed\(^{37}\).

Additional examples of scrambled gene-order structures in small parasitic genomes are provided by the tRNA genes. Most of the deviations from the conventional gene-order structure (16S–23S–5S) are associated with bacterial genomes in the 1-Mbp range, having only one or two copies of each of the tRNA genes\(^{38,39}\). A general survey of gene order conservation in the R. prowazekii genome is consistent with the idea that this genome is a highly derived, rearranged minimal genome\(^{35}\). Thus, small sizes and scrambled gene organization patterns seem to be a characteristic feature of the genomes of resident bacteria.
Genomic extinction

Transitions to new environmental niches can induce the loss of metabolic functions previously supplied by a long-term endosymbiont. For example, photosynthesis has been secondarily lost in some plants, algae and protists that live as parasites of other plants. Not surprisingly, the chloroplast genomes of these species are markedly reduced and the photosynthetic genes have either been eliminated or are present as pseudogenes. The presence of a residual plastid genome is indicative of a role for chloroplast gene products in metabolic processes other than photosynthesis. The most extreme examples of such reductive processes are the complete loss of mitochondrial genomes in eukaryotes that have adapted to anaerobic or obligate intracellular lifestyles, their past being evident only from the presence of the mitochondrial heat-shock proteins Hsp10, Hsp60 and Hsp70 (Refs 48–50).

Another interesting case of extreme reductive evolution is the incorporation of a green algae by a group of marine protists called chlorarchinophytes, in which the remnants of the invading nuclear genome, the nucleomorph, have been reduced to <1 Mbp (Refs 51, 52).

Fig. 3. Genomic rearrangements. (a) It is believed that an ancestor of Bacteria and Archaea had a set of transcriptional and translational genes located in close proximity to each other in the following order: nusG, L11 (rpsLA), L10 (rpsKJ), rif (rpoBC), str (rpsLG–tuf–fus), S10 (rpsJ–rplCDWB–rpsS–rplV–rpsC–rplP–rpmC–rpsQ), spc (rplNXE–rpsNH–rplFR–rpsE–rpmD–rplO–prlA–rpmJ) and α (rpsMKD–rpoA–rplQ). These segments are conservedly arranged in the genomes of (a) Bacillus subtilis and (b) Escherichia coli, but are scattered around the genomes of (d) Haemophilus influenzae, (c) Mycoplasma genitalium and (f) Rickettsia prowazekii. The organization of these genes in the R. prowazekii genome is consistent with an intrachromosomal recombination event at the tuf genes (41). The circumferences of the circles are directly proportional to the genome sizes. Data taken from Ref. 39.
REVIEWS

How small must sexual populations be to behave as asexual populations (i.e. populations where there is little or no recombina-

Are bottle-necked subpopulations, which start growth cycles from a limited number of individuals, a valuable source of mutational variation for large, sexual populations?

Does the successful incorporation of resident genomes also serve as a bottleneck for their hosts, by eliminating a large num-

structures provide support for a similar kind of sec-

Questions for future research

• How small must sexual populations be to behave as asexual

• Are bottle-necked subpopulations, which start growth cycles from a limited number of individuals, a valuable source of mutational variation for large, sexual populations?

• Does the successful incorporation of resident genomes also serve as a bottleneck for their hosts, by eliminating a large number of infected hosts, in the case of pathogens, or by boosting the population with infected offspring, in the case of symbionts?

Conclusions

Duplication events can, in principle, be reversed by compensatory deletion events. By contrast, deleted se-

References

ternal events and the gradual disruptions or loss of genes in the resident genomes are, therefore, ex-


7 Chao, L. (1997) Genet. 201, 301–308


33 Fraser, C.M. et al. (1999) Science 283, 1297–403


41 Syvanen, A.C. et al. (1996b) J. Bacteriol. 178, 6192–6199


55 Kohler, S. et al. (1997) Science 275, 1485–1489