Symbiosis, 34 (2003) 301–316
Balaban, Philadelphia/Rehovot

Review article

Tempo and Mode of Genomic Evolution in Endosymbiotic Bacteria of Insects: The Case of *Buchnera aphidicola*

AMPARO LATORRE, ROSARIO GIL, FRANCISCO J. SILVA, DAVID MARTÍNEZ-TORRES, and ANDRÉS MOYA*
Instituto Cavanilles de Biodiversidad y Biología Evolutiva and Departament de Genètica, Universitat de València, Apartado Postal 22085, 46071 Valencia, Spain, Tel. +34-96-3543480, Fax. +34-96-3543670, Email. andres.moya@uv.es

Received March 10, 2003; Accepted April 24, 2003

Abstract

Many bacteria, especially proteobacteria, have established a symbiotic association with insects. Insects confer to their partners a stable environment, while the bacteria provide the host with new metabolic capabilities, complementing the insect nutritional requirements and contributing to the exploitation of new ecological niches. The analyses of the complete sequence of five genomes of endosymbionts of insects has shown that both, at gene and genomic levels, endosymbiotic bacteria have experienced drastic transformations, when compared with their free-living relatives. A massive genome reduction occurred during the establishment of symbiosis. The loss of genes involved in DNA repair and recombination, probably some of the first gene losses that occurred in the reductive process, had important posterior consequences. The remnant gene repertoire can be classified in two groups: a set of genes shared by all endosymbionts and another set of genes having specific role in the insect-bacteria association. In this review we take advantage of the recent genome sequencing of three *Buchnera aphidicola* strains

*The author to whom correspondence should be sent.

0334-5114/2003/$05.50

living on three different aphid hosts, to get some light into the factors that are responsible for the genome reduction syndrome and its associated features, such as the accumulation of deleterious mutations. We also discuss about the tempo and mode of genomic evolution in endosymbionts.

Keywords: Aphids, endosymbionts, *Buchnera aphidicola*, genome reduction, evolution

1. Introduction

The term symbiosis (from the Greek *simbios* or living together) was first introduced by Anton de Bary in 1879. This author explicitly included parasitism as a type of symbiosis and excluded short-term associations. Some authors do not accept Bary's definition and hold that short-term interactions should be considered true symbiosis and parasitic associations must be rejected as such. Most researchers consider, in a more restrictive meaning, that there is symbiosis only when both partners benefit from the association. Douglas (1994) narrows the term even further and confines symbiosis to associations in which at least one partner bestows the other with some sort of novel metabolic capability, as it is the case of most of the well known associations.

Contrary to most prokaryotes, eukaryotes have rather limited metabolic capabilities and, hence, symbiosis has provided an evolutionary strategy through which eukaryotes gain access to a wider range of metabolic resources. Insects are particularly receptive to symbiotic processes, and it has been estimated that at least 15–20% of all insects live in such symbiotic relationships (Buchner, 1965), allowing them to explore a great variety of ecological niches (Douglas, 1989; Moran and Bauman, 2000). One special case of symbiosis is endosymbiosis, meaning that one partner, generally a prokaryote, is located inside the body of the other. In some cases, the prokaryote is literally sequestered within a eukaryotic cell. Insects that have established endosymbiotic processes with bacteria are characterized, in general, by feeding upon unbalanced diets, poor in essential nutrients such as amino acids, sterols or vitamins, which are provided by the symbionts (Bauman et al., 1995, 2000; Douglas, 1998).

Mutualistic and obligate insect-bacteria symbiosis have been intensively studied in the last years, particularly those involving *Buchnera aphidicola*, *Wigglesworthia glossinidia* and *Blochmannia floridanus*, primary endosymbionts of aphids, tsetse flies and carpenter ants, respectively.

Aphids are plant phloem-feeding insects that harbor the bacterium *B. aphidicola* as a primary endosymbiont (Munson et al., 1991). The phloem is rich in sugars, but poor in nitrogenous compounds. In particular, aphids diet is

deficient in essential amino acids, vitamins and several essential lipids. Three *B. aphidicola* genomes have been sequenced (Shigenobu et al., 2000; Tamas et al., 2002; van Ham et al., 2003), and their gene content revealed the presence of genes coding for the essential nutrients that are lacking in aphids diet, thus giving support to the previously proposed nutritional role for the symbiosis between aphids and *B. aphidicola*.

The tsetse fly, the vector of African trypanosomes, harbors two symbiotic bacteria: *W. glossinidia*, as an obligate primary symbiont, and *Sodalis glossinidia*, as a commensal secondary symbiont. The genome sequence of *W. glossinidia* has shown that genes involved in the biosynthesis of vitamin metabolites are retained, which also suggests a nutritional role for this symbiont, supplying the deficiencies in the tsetse fly diet, composed exclusively of blood from the vertebrate host (Akman et al., 2002).

Finally, social insects, such as ants, have developed numerous interactions with different species of animals, plants and microorganisms. The genome of *Bl. floridanus*, the primary endosymbiont of carpenter ants of the species *Camponotus floridanus*, has recently been sequenced by our group (Gil et al., 2003), revealing that this symbiosis also has a nutritional basis, in spite of the fact that these ants feed on a complex diet. In this case, *Bl. floridanus* supplies nitrogen and sulfur compounds to the host.

A very wide phylogenetic study has recently been carried out, based on the alignment of 61 concatenated conserved protein-coding genes involved in translation from 19 selected γ-proteobacteria, including the insect endosymbionts *Bl. floridanus*, *W. glossinidia* and three strains of *Bu. aphidicola* (Gil et al., 2003). The phylogenetic relationships were evaluated by means of maximum likelihood and Bayesian methods (Schmidt et al., 2002; Huelsenbeck and Ronquist, 2001). Briefly, the maximum likelihood tree was obtained by the quartet puzzling method. A matrix of amino acid substitution, along with a gamma model (α=0.99) for rate variation among sites, and a proportion p=0.14 for invariant sites, were used with 4,000 puzzling steps. The Bayesian analysis proceeded with a given model for amino acid substitution, complemented with a gamma + invariant model for rate heterogeneity among sites. Fig. 1 is part of the obtained phylogeny, showing the phylogenetic relationships for the five endosymbionts of insects plus four close-related free-living bacteria whose genomes have also been sequenced (*Escherichia coli* K12; *Salmonella enterica*, *Haemophilus influenzae* and *Vibrio cholerae*). The genome sizes and the number of genes of all of the compared bacteria are also included. As it can be observed, the three symbiotic bacterial lineages form a monophyletic cluster belonging to the γ-3 Proteobacteria subdivision. The analyses of the complete sequence of the five known genomes from bacterial endosymbionts of insects (three strains of *B. aphidicola*, *W. glossinidia* and *Bl. floridanus*) reveal a dramatic process of gene and genome reduction in the three

bacterial lineages, when compared to their closest free-living relatives such as *Escherichia coli*. An interesting feature, with remarkable consequences for the genome evolution of the endosymbionts, is the loss of most of the genes involved in DNA repair and recombination. Furthermore, in general no orphan genes are present, indicating that no new genes have been acquired by horizontal gene transfer (Silva et al., 2003). The remnant gene repertoire can be grouped into two major groups. One group corresponds to the genes that are shared by all analysed endosymbionts, which may well represent a minimum gene set necessary to sustain symbiotic life. The second one is formed by those genes having a specific role in the association with their corresponding hosts, thus being specific for each bacterial species.

For a better understanding of the genome reduction process it is necessary to have a deep knowledge of the biology of the symbiotic association, i.e., that of each of the partners and of the symbiotic interaction itself. This is the case of aphids and *B. aphidicola*, for which many studies have been carried out in the last years regarding the biology of aphids, *B. aphidicola*, and on the aphid-*B. aphidicola* association (reviewed in Bauman et al., 2000). In addition, *B. aphidicola* is the only endosymbiotic bacteria for which the genome of three strains obtained from three different aphid species has been reported (Shigenobu et al., 2000; Tamas et al., 2002; van Ham et al., 2003). Thus, we will focus most of the remaining part of this review analyzing *B. aphidicola* and its symbiotic relationship with aphids.

2. The Aphid-*B. aphidicola* Symbiotic Association

The genus *Buchnera* contains one species, *B. aphidicola*, being the type strain of the endosymbiont of the aphid *Schizaphis graminum*, as defined by Munson et al. (1991). Currently this species name designates the lineage of these bacteria in the different aphid species. Thus, we will refer to the primary endosymbionts of the different aphid species as *B. aphidicola* strains. *B. aphidicola* is confined into specialized polyploid cells, called bacteriocytes. These cells form aphid organ-like structures called bacteriomes, localized in the haemocoel of the insect, above the digestive tube. *B. aphidicola* is maternally inherited by means of infection of the eggs or embryos during the blastoderm stage (Buchner, 1965; Hinde, 1971; Houk and Griffiths, 1980). This symbiosis is an example of obligate mutualism. On one hand, the aphid cannot reproduce without *B. aphidicola*, since aposymbiotic aphids, obtained after treatment with antibiotics, have a slow growth and have no fertile offspring. On the other hand, the bacteria cannot survive outside the aphid, since *B. aphidicola* cannot be cultured in artificial media (Bauman et al., 1995; Douglas, 1998; Prosser and Douglas, 1991).

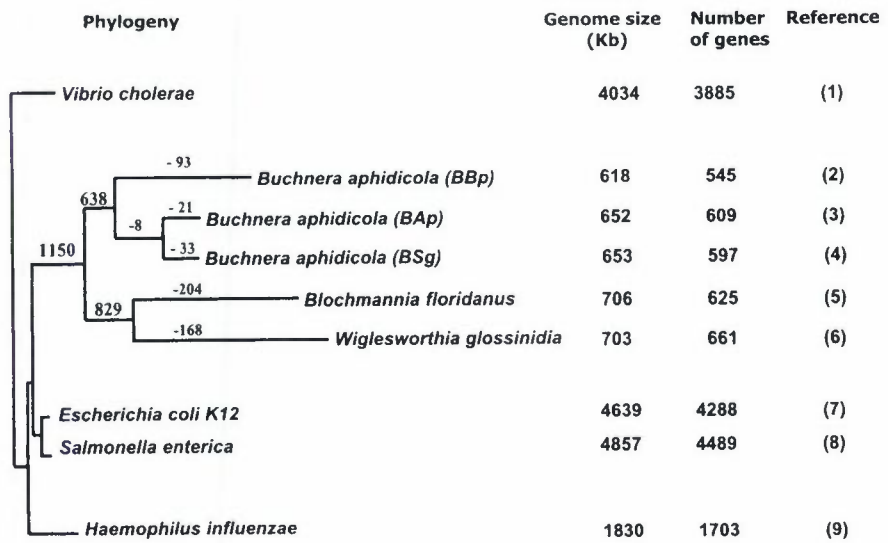
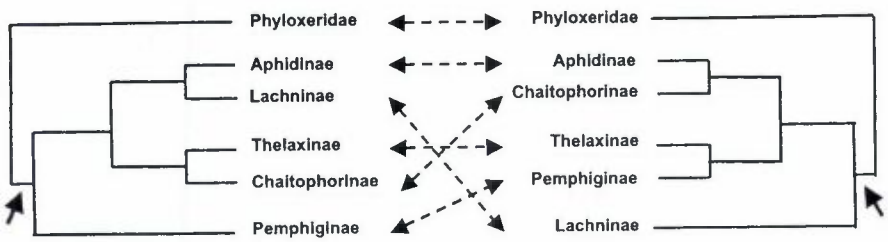


Figure 1. Phylogeny, genome size and number of genes of the five sequenced genomes of insect endosymbionts, and five free-living bacteria. For the methods used to infer the phylogenetic relationships and the statistical support, see text. All the nodes are highly supported statistically (between 85 to 100%). The positive numbers above the branches show the inferred minimum number of corresponding ancestral genomes. The negative numbers show the postulated genes losses occurred on each one of the bacterial symbionts. (1) Heidelberg et al. (2000); (2) van Ham et al. (2003); (3) Shigenobu et al. (2000); (4) Tamas et al. (2002); (5) Gil et al. (2003); (6) Akman et al. (2002); (7) Blattner et al. (1997); (8) Parkhill et al. (2001); (9) Fleischmann et al. (1995).



A) B)

Figure 2. Evolutionary relationships among the aphid subfamilies analyzed in this study according to A) Heie (1987), and B) Ortiz et al. (2003). The taxonomy used in this study is the one proposed by Remaudière and Remaudière (1997) that consider the range of family for the Aphidoidea (instead of superfamily as in the previous one) and, consequently of subfamily for the derived aphid lineages. Arrows represents the origin of the *B. aphidicola* infection and broken lines represent the non-congruence of several subfamilies between both trees (see text for more details).

Phylogenetic analysis based on 16S rDNA sequences of several *B. aphidicola* strains have shown that they belong to the γ-3 subdivision of the class Proteobacteria, having *E. coli* and related members of the Enterobacteriaceae as their closest known relatives (Moran et al., 1993; see also Fig. 1). These results also revealed that this symbiosis resulted from a single bacterial infection of the common ancestor of all modern aphids, approximately 100–250 million years ago, and followed by co-evolution of both partners.

Several molecular phylogenetic studies using different genes from *B. aphidicola* (Brynnel et al., 1998; Baumann et al., 1999; van Ham et al., 1999, 2000), were in good agreement with the morphology-based phylogeny of aphids proposed by Heie (1987) (Fig. 2a). However, in most of these phylogenetic studies, species sampling was limited to representatives of a few subfamilies. Van Ham et al. (1997) using also 16S rDNA sequences of *B. aphidicola* from aphids representative of additional subfamilies, put into question some of the traditionally accepted relationships between some aphid subfamilies.

Recently, by using mitochondrial ribosomal DNA sequences, von Dohlen and Moran (2001) carried out a broad phylogenetic study where aphid representatives of many subfamilies were chosen. They found, in general, a good agreement of external branches (i.e., at the level of tribes) with phylogenies based on morphological data, but the resolution of the internal nodes was very poor, therefore being unable to support or contradict the relationships among subfamilies put forward by morphology based phylogenies.

Finally, Martínez-Torres et al. (2001) and Ortiz-Rivas et al. (2003) carried out large studies using several genes from both *B. aphidicola* and aphid species representative of several subfamilies, including some belonging to the Lachninae. From their results, some of the traditionally accepted groupings, both in bacterial and aphid phylogenies, are put into question (Fig. 2b). In particular, the subfamilies Lachninae and Aphidinae do not seem to be as closely related as previously reported. Further knowledge about the relationships among the different subfamilies of the family Aphididae will provide important clues to understand the evolutionary process of the symbiosis in *B. aphidicola*, leading to features such as the variation of the genome size of different *B. aphidicola* strains, plasmidic or chromosomal location of genes involved in biosynthesis of amino acids, and accelerated evolutionary rate of some lineages.

3. Genome Size of *Buchnera aphidicola*

Charles and Ishikawa (1999) determined by Pulse-Field Gel Electrophoresis (PFGE) the chromosome size of *B. aphidicola* from the aphid *Acyrthosiphon*

pisum (BAp), a member of the subfamily Aphidinae, as 0.64 Mb. This size was confirmed with the complete genome sequence by Shigenobu et al. (2000). At that time, this was the second smallest bacterial genome reported, being the genome of *Mycoplasma genitalium*, the obligate parasite of epithelial cells, the smallest one (580 Kb) (Fraser et al., 1995).

These data, together with new genome size estimates for several endosymbionts and obligate intracellular pathogens, showed that symbionts and obligate pathogens maintain a highly compact genome with a minimum number of coding genes, due to a rapid loss rate of genes, in contrast to the free-living bacteria, that tend to increase their genome sizes by means of horizontal gene transfer and gene duplication (Silva et al., 2001, 2003). For example, the genome sizes of *Bl. floridanus*, and *W. glossinidia*, are 706 and 703 Kb, respectively (Akman et al., 2002; Gil et al., 2003). The genome size of *Rickettsia prowazekii*, the agent producing the epidemic typhus is 1.1 Mb (Andersson et al., 1998), and *Haemophilus influenzae*, another human pathogen, has also a reduced genome size of 1.8 Mb (Fleischmann et al., 1995).

The analysis by PFGE of the chromosome size of three *B. aphidicola* strains from aphids of the subfamilies Aphidinae and Pemphiginae (Wernegreen et al., 2000) showed a slightly smaller size for the latter species. Gil et al. (2002) analyzed the chromosome size of nine *B. aphidicola* strains belonging to five aphid subfamilies, including the same two subfamilies studied by Wernegreen et al. (2000). As it can be seen in Table 1, great genome size variation was found, ranging from 670 to 450 Kb. Thus, *B. aphidicola* strains from aphids of the subfamilies Chaitophorinae, Thelaxinae and Lachninae posses genome sizes smaller than *M. genitalium* (from approximately 450 to 550 Kb), being *B. aphidicola* strains from aphids of the genus *Cinara* (subfamily Lachninae) the ones with the smallest genomes. In fact, we can consider the genome of *B. aphidicola* from the aphid *Cinara cedri* (BCce) the smallest known bacterial genome reported so far (450 Kb) (see Table 1).

The comparison of the genome sizes of several *B. aphidicola* strains revealed great differences among aphid subfamilies. This is, for instance, the case of the three *B. aphidicola* from aphids of the subfamily Lachninae, that have 200 Kb less than the two *B. aphidicola* from aphids of the subfamily Aphidinae. However, the differences within subfamilies are much smaller, being of average of 20 Kb (see Table 1).

Although we can hypothesize that the great majority of genome shrinkage probably occurred in the Last Common Symbiotic Ancestor (LCSA) of *B. aphidicola* at the beginning of the symbiotic integration (see below), since then the different strains of *B. aphidicola* have undergone a reductive process in a way that correlates with their host, at least at the subfamily level, and they may well be evolving towards the minimal genome required for the survival into the corresponding aphid host.

Table 1. Chromosome and leucine and tryptophan plasmid sizes for several *B. aphidicola* species belonging to different aphid subfamilies.

Aphid subfamily	Tribe	*Buchnera* species	Chromosome (bp)	Leucine plasmid (bp)	Leucine cryptic plasmid (bp)	Tryptophan plasmid (bp)
Pemphiginae	Fordini	BBp*	615,980 [1]	No	2.399 [1]	No
	Eriosomatini	BTc	~565,000 [2]	No	1.740 [3]	3.000 [4]
Thelaxinae		BThs	~544,000 [2]	8,500 [3]	No	No
Chaitophorinae	Chaitophorini	BChp	~508,000 [2]	No	No	No
Lachninae	Cinarini	BCce	~448,000 [2]	~6,500 [u]	No	No
		BCcu	~476,000 [2]	~6,500 [u]	No	No
		BCt	~477,000 [2]	~6,500 [u]	No	No
Aphidinae	Macrosiphini	BAp*	640,681 [5]	7,805 [6]	No	~7.2585
		BMr	~669,000 [2]	~7,800 [u]	No	?
	Aphidini	BSg*	641,454 [7]	7,967 [8]	No	3.6 (4 t.r.) [9]

Abbreviation are as follows: B from *Buchnera aphidicola* followed by the aphid species. Bp, *Baizongea pistaceae;* Tc, *Tetraneura caerulescens;* Ths, *Thelaxes suberi;* Chp, *Chaitophorus populeti;* Cce, *Cinara cedri;* Ccu, *Cinara cupresi;* Ct, *Cinara tujafilina;* Ap, *Acyrtosiphon pisum;* Mr, *Macrosiphun rosae;* Sg, *Schizaphis graminum.* [1]van Ham et al. (2002); [2]Gil et al. (2002); [3]van Ham et al. (1997); [4]van Ham et al. (1999); [5]Shigenobu et al. (2000); [6]Soler et al. (2000); [7]Tamas et al. (2002); [8]Baumann et al. (1999); [9]Lai et al. (1994). *Sequenced genomes; t.r., tandem repeats. [u]unpublished results.

4. *B. aphidicola* Genome Analysis

The genome of *B. aphidicola* is formed by one main chromosome plus one of two plasmids (Bracho et al., 1995; Lai et al., 1994), depending on the *B. aphidicola* strain (see Table 1). A leucine plasmid, ranging in size from 6.3 to 8.2 Kb, is present in the subfamilies Aphidinae, Thelaxinae and Lachninae (Silva et al., 1998; van Ham et al., 1997, 2000). It contains the four structural genes for the synthesis of leucine (*leu*A, *leu*B, *leu*C and *leu*D) and one or two genes coding for a plasmid replicase (*rep*A). Some plasmids also contain the *ibp* gene, coding for a heat shock protein, and an ORF1 coding for a putative membrane protein (*yqh*A). Species from the subfamily Pemphiginae posses a cryptic leucine plasmid of 1.7–2.4 Kb (van Ham et al., 2000) that does not contain the leucine genes. *B. aphidicola* from aphids of the subfamily Aphidinae and some tribes of the subfamily Pemphiginae also contain tryptophan plasmids (Rouhbakhsh et al., 1996; van Ham et al., 1999), ranging in size from 3.0 to 12.8 Kb, that contain the two first genes of the tryptophan pathway (*trp*EG). The variability in size is mainly due to variability in the number of tandem repeats of these genes or pseudogenes. It is worth noticing that, with the only exception of the duplicate gene *grp*E in the main chromosome, genes and/or pseudogenes duplications are only found in the plasmids of *B. aphidicola*. Sabater-Muñoz et al. (2002) have postulated that a leucine plasmid was present in the *B. aphidicola* LCSA that preceded the diversification of all the endosymbionts, and that the chromosomal location of the leucine genes observed in some *B. aphidicola* strains arose by a transfer of such genes from a plasmid to the main chromosome. The reason why some lineages transferred back the plasmid to the main chromosome is an open and unsolved question, but due to the metabolic role of *B. aphidicola* in the symbiosis, it is likely to be related to the quality of phloem-sap from different host plants and/or the growth rate of the aphids.

The sequencing of the genome of *B. aphidicola* BAp (Shigenobu et al., 2000) allowed the comparison with *E. coli* (Blattner et al., 1997) and *Vibrio cholerae* (Heidelberg et al., 2000), the two closest free-living bacteria whose genomes have been sequenced. These comparisons have shown that the endosymbiont not only has experienced large genome reductions, but also large genome reorganizations (Silva et al., 2001).

Two new genomes of *B. aphidicola* from the aphids, *Schizaphis graminum* (BSg) and *Baizongia pistacea* (BBp), have been sequenced, providing new insights for understanding the genome evolution of these species. The comparison of the strains BAp and BSg coming from the same aphid subfamily (Aphidinae), but from different tribes (Macrosiphini and Aphidini, respectively) (see Table 1), with estimated divergence time of 50 to 70 million years, revealed an extreme conservation of the genome order with neither

chromosomal rearrangements (translocations, inversions or duplications), nor gene acquisition by horizontal gene transfer, thus being the most extreme case of genome stability to date (Tamas et al., 2002). The comparison with a third strain, *B. aphidicola* BBp, coming from aphids of the subfamily Pemphiginae, revealed a nearly perfect gene-order conservation, with only four minor rearrangements (two inversions and two translocations involving the leucine and tryptophan plasmid-contained genes) in the BBp strain relative to the BAp and BSg strains. Since the Aphidinae and Pemphiginae lineages diverged about 80 to 150 million years, van Ham et al. (2003) suggested that *B. aphidicola* can be considered as an enterobacterial "gene-order fossil" and that the onset of genomic stasis coincided with the establishment of the symbiosis with aphids, approximately 200 million years ago. However, the gene content is different in the three lineages, indicating that independent gene losses have occurred from the LCSA of *B. aphidicola*. The complete set of coding genes present in the three *B. aphidicola* strains, obtained as the sum of all shared and non-shared genes and pseudogenes, is 638. This represents the minimum number of genes that LCSA would contain (see Fig. 1). Thus, a minimum of 155 independent gene losses have occurred in the three *B. aphidicola* strains analyzed, 93 in the BBp lineage, 8 before the split of the two Aphidini, and 21 and 33 in the BAp and BSg lineages, respectively. The reason why some genes are lost (or retained) in some lineages is not clear. The current sequencing of the genome of *B. aphidicola* BCce (see Table 1) could provide new clues to address that question, since it possesses the smallest bacterial genome ever found, and could represent the minimum number of genes carried out by a symbiont before becoming extinct.

Using the published data of the *W. glossinidia* genome (Akman et al., 2002) and our sequence data for *Bl. floridanus* (583 protein-coding genes and 42 rRNA genes; Gil et al., 2003) we have estimated that the minimum number of genes of the LCSA between both species would be 829. Finally, taking into account all the data, we can estimate that the minimum number of genes of the LCSA for the five analyzed endosymbionts would be 1150 (Fig. 1). This last number is still very distant from the number of genes present in their closest free living relatives, even when genes acquired in these species by horizontal gene transfer are not considered.

The genomic alignment of orthologous gene position in pairwise comparisons between *B. aphidicola* BAp, *Bl. floridanus* and *E. coli* has revealed that many chromosomal rearrangements have occurred since the divergence of the *B. aphidicola* and *Bl. floridanus* lineages, with different rearrangements in the different species (Silva et al., 2001; Gil et al., 2003). These results let us postulate that the massive gene loss that took place in the process towards the LCSA of *B. aphidicola* and probably of *Bl. floridanus* was accompanied by many chromosomal rearrangements. So, recombination by means of unequal

crossing-over between repeat sequences, including transposable elements before their disappearance, would be the most probable mechanism for these rearrangements, and would indicate that the loss of most genes involved in recombination and gene repair, would have occurred in a similar moment for both species, i.e., at the beginning of the process towards the symbiotic life. We know that, since then, at least in *B. aphidicola*, the only important genome process has been reduction in gene content in the different lineages, with the exception of the back transfer of leucine and tryptophan plasmids to the main chromosome in some lineages.

In addition to its drastic genome reduction and the proliferation of plasmids involved in the biosynthesis of leucine and tryptophan, *B. aphidicola* has suffered many other important molecular and biochemical changes, when compared to their free-living relatives. These changes can be summarized in the following features: amplification of the number of copies of the main chromosome per cell, an almost total absence of recombination, increase in the rate of nucleotide substitution, high A+T content, accumulation of deleterious mutations by random genetic drift, and loss of codon bias (Baumann et al., 2000; Clark et al., 1999; Komaki and Ishikawa, 1999; Moran et al., 1995; Moran, 1996; Moya et al., 2002).

5. Tempo and Mode of Genome Evolution

Following the punctuated versus gradual debate on biological evolution, we can talk about two different models regarding genome reduction. Punctuated genome reduction involves a combination of short periods of large genome reduction and large periods of no genome reduction or evolutionary stasis (i.e., the tempo is not uniform). Although there is a general agreement for this "tempo" in *B. aphidicola*, whose genome reduction took place mainly at the beginning of the symbiotic process, it is still to be tested what is the "mode" of such process, that is, if it occurred by massive deletions of varying size (Moran and Mira, 2001) or by gene disintegration (Silva et al., 2001), and if this last process is still active. For instance, as mentioned above, two strains of *B. aphidicola* clonally evolving inside different aphid species of the subfamily Aphidinae (*A. pisum* and *S. graminum*) suffered genomic stasis during the last 50 MY of evolution (Tamas et al., 2002) as revealed by the almost perfect match in gene order when comparing both genomes, with no chromosome rearrangements or gene acquisitions, and with similar genome sizes. However, we have also shown that other strains evolving inside other aphid subfamilies showed a much smaller genome (Table 1). In other words, there is no stasis when considering genome size.

Another open question is whether the process of genome reduction plus the

accumulation of deleterious mutations will end with the extinction of these endosymbionts. Due to their maternal inheritance, only a small fraction of bacteria pass to the next generation, then experiencing continuous bottlenecking. A consequence of this mode of inheritance, whose effects have been studied in *B. aphidicola*, is the accumulation of deleterious mutations in an irreversible way, a phenomenon known as Muller's ratchet (Moran, 1996). The action of random drift promoting this accumulation compromises the effectiveness of purifying selection, which is also present purging less efficient bacterial genotypes. So, we can deduce that the non-lost genes are accumulating deleterious mutations, but we do not know whether the process of genome degradation (gene loss and accumulation of deleterious mutations) is a transient state towards their complete extinction or if there is any evidence of some process that may counter-balance the impact on the genome of the Muller's ratchet.

Recently, the simulation of the process of bottlenecking using experimental populations of *E. coli* has demonstrated the role played by GroEL. This heat-shock protein acts as a molecular chaperone, which is overproduced and constitutively over expressed in endosymbiotic but not in free-living bacteria (Fares et al., 2002b). It seems that GroEL assists in the folding of conformationally damaged proteins. Moreover, by performing phylogenetic analysis of GroEL, it has been shown that this gene has been positively selected in different lineages of *B. aphidicola* (Fares et al., 2002a). GroEL, then, may act by buffering the effect of the accumulation of amino acid replacements, therefore maintaining the functionality of the endosymbiont proteome. This and other compensatory mechanisms may help to solve the problem of the accumulation of deleterious mutations. The demonstration of such active mechanisms will be of particular relevance for understanding the evolution of symbiosis.

Acknowledgements

This research was funded by grant GV01-177 from Oficina de Ciència i Tecnología del Govern Valencià to A. Latorre, and grant BFM2000-1313 from Ministerio de Ciència y Tecnología (MCYT) to A. Moya.

REFERENCES

Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M., and Aksoy, S. 2002. Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nature Genetics* **32**: 402–407.

Andersson, S.G., Zomorodipou, A., Andersson, J.O., Sicheritz-Ponten, T., Alsmark, U.C., Podowski, R.M., Naslund, A.K., Eriksson, A.S., Winkler, H.H., and Kurland, C.G. 1998. The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396**: 133–140.

Baumann, P., Baumann, L., Lai, C.Y., Roubakhsh, D., Moran, N.A., and Clark, M.A. 1995. Genetics, physiology, and evolutionary relationships of the genus *Buchnera*: intracellular symbionts of aphids. *Annual Review of Microbiology* **49**: 55–94.

Baumann, L., Baumann, P., Moran, N.A., Sandström, J., and Thao, M.L. 1999. Genetic characterization of plasmids containing genes encoding enzymes of leucine biosynthesis in endosymbionts (*Buchnera*) of aphids. *Journal of Molecular Evolution* **48**: 77–85.

Baumann, P., Moran, N.A., and Baumann, L. 2000. Bacteriocyte-associated endosymbionts of insects. In: *The Prokaryotes*. M. Dworkin, ed. Springer-Verlag, New York, pp. 1–67. http://link.springer.de/link/service/books/10125/.

Blattner, F.R., Plunkett, G. 3rd., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., and Shao, Y. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* **227**: 1453–1474.

Bracho, A.M., Martínez-Torres, D., Moya, A., and Latorre, A. 1995. Discovery and molecular characterization of a plasmid localized in *Buchnera* sp. bacterial endosymbionts of the aphid *Rhopalosiphum padi*. *Journal of Molecular Evolution* **41**: 67–73.

Buchner, P. 1965. *Endosymbiosis of Animals with Plants and Microorganisms*. Interscience Publishers, Inc. New York.

Brynnel, J.K., Kurland, C.G., Moran, N.A., and Andersson, S.G.E. 1998. Evolutionary rates for *tuf* genes in endosymbionts of aphids. *Molecular Biology and Evolution* **15**: 574–582.

Charles, H. and Ishikawa, H. 1999. Physical and genetic map of the genome of *Buchnera*, the primary endosymbiont of the pea aphid *Acyrthosiphon pisum*. *Journal of Molecular Evolution* **48**: 142–150.

Clark, M., Moran, N.A., and Baumann, P. 1999. Sequence evolution in bacterial endosymbionts having extreme base composition. *Molecular Biology and Evolution* **16**: 1586–1598.

Douglas, A.E. 1989. Mycetocyte symbiosis in insects. *Biological Reviews of the Cambridge Philosophical Society* **64**: 409–434.

Douglas, A.E. 1994. *Symbiotic Interactions*. Oxford University Press. Oxford.

Douglas, A.E. 1998. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. *Annual Review of Entomology* **43**: 17–37.

Fares, M.A., Barrio, E., Sabater-Muñoz, B., and Moya, A. 2002a. The evolution of heat-shock protein GroEL from *Buchnera*, the primary endosymbiont of aphids, is governed by positive selection. *Molecular Biology and Evolution* **19**: 1162–1170.

Fares, M.A., Ruiz-Gonzalez, M.X., Moya, A., Elena, S.F., and Barrio, E. 2002b. GroEL buffers against deleterious mutations. *Nature* **417**: 398–398.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**: 496–512.

Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., Bult, C.J., Kerlavage, A.R., Sutton, G., Kelley, J.M., et al. 1995. The minimal gene complement of *Mycoplasma genitalium*. *Science* **270**: 397–402.

Gil, R., Sabater-Muñoz, B., Latorre, A., Silva, F.J., and Moya, A. 2002. Extreme genome reduction in *Buchnera* spp.: toward the minimal genome needed for symbiotic life. *Proceedings of the National Academy of Science USA* **99**: 4454–4458.

Gil, R., Silva, F.J., Zientz, E., Delmotte, F., González-Candelas, F., Latorre, A., Raussell, C., Kamerbeek, J., Gadau, J., Hölldobler, B., van Ham, R.C.H.J., Gross, R., and Moya A. 2003. The genome sequence of *Blochmannia floridanus*: Comparative analysis of reduced genomes. *Proceedings of the National Academy of Science, USA*. In press.

Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L., Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D., Ermolaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragoi, I., Sellers, P., McDonald, L., Utterback, T., Fleishmann, R.D., Nierman, W.C., and White, O. 2000. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406**: 477–483.

Heie, O.E. 1987. Paleontology and phylogeny. In: *Aphids. Their Biology, Natural Enemies and Control*. A.K. Minks and P. Harrewijn, eds. Elsevier, Amsterdam, pp. 367–391.

Hinde, R. 1971. Maintenance of aphids cells and the intracellular symbiotes of aphids *in vitro*. *Journal of Invertebrate Pathology* **17**: 333–338.

Houk, E.J. and Griffiths, G.W. 1980. Intracellular symbiotes of the Homoptera. *Annual Review of Entomology* **25**: 161–187.

Huelsenbeck, J.P. and Ronquist, F. 2001. MrBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.

Komaki, K. and Ishikawa, H. 1999. Intracellular bacterial symbionts of aphids posses many genomic copies per bacterium. *Journal of Molecular Evolution* **48**: 717–722.

Lai, C.Y., Baumann, L., and Baumann, P. 1994. Amplification of *trp*EG: adaptation of *Buchnera aphidicola* to an endosymbiotic association with aphids. *Proceedings of the National Academy of Science USA* **91**: 3819–3823.

Martínez-Torres, D., Buades, C., Latorre, A., and Moya, A. 2001. Molecular systematics of aphids and their primary endosymbionts. *Molecular Phylogenetics and Evolution* **20**: 437–449

Moran, N.A. 1996. Accelerated evolution and Muller's ratchet endosymbiotic bacteria. *Proceedings of the National Academy of Science USA* **93**: 2873–2878.

Moran, N.A. and Baumann, P. 2000. Bacterial endosymbionts in animals. *Current Opinions in Microbiology* **3**: 270–275.

Moran, N.A, Munson, M.A., Baumann, P., and Ishikawa, H. 1993. A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. *Proceedings of the Royal Society of London B* **253**: 167–171.

Moran, N.A. and Mira, A. 2001. The process of genome shrinkage in the obligate symbiont *Buchnera aphidicola*. *Genome Biology* **2**: RESEARCH 0054.

Moran, N.A., von Dohlen, C.D., and Baumann, P. 1995. Faster evolutionary rate in endosymbiotic bacteria than in cospeciating insect hosts. *Journal of Molecular Evolution* **41**: 727–731.

Moya, A., Latorre, A., Sabater, B., and Silva, F.J. 2002. Comparative molecular evolution of primary (*Buchnera*) and putative secondary endosymbionts of aphids based on two protein-coding genes. *Journal of Molecular Evolution* **54**: 127–137.

Munson, M.A., Baumann, P., and Kinsey, M.G. 1991. *Buchnera* gen. nov. and *Buchnera aphidicola* sp. nov., a taxon consisting of the mycetocyte-associated, primary endosymbionts of aphids. *International Journal of Systematic Bacteriology* **41**: 566–568.

Ortiz-Rivas, B., Moya, A., and Martínez-Torres, D. 2003. Molecular systematic of aphids (Homoptera: Aphididae): new insights from the long-wave length Opsina gene. *Molecular Phylogenetics and Evolution*. In press.

Parkhill, J., Dougan, G., James, K.D., Thomson, N.R., Pickard, D., Wain, J., Churcher, C., Mungall, K.L., Bentley, S.D., Holden, M.T., Sebaihia, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T., Connerton, P., Cronin, A., Davis, P., Davies, R.M., Dowd, L., White, N., Farrar, J., Feltwell, T., Hamlin, N., Haque, A., Hien, T.T., Holroyd, S., Jagels, K., Krogh, A., Larsen, T.S., Leather, S., Moule, S., O'Gaora, P., Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S., and Barrell, B.G. 2001. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature* **423**: 848–852.

Prosser, W.A. and Douglas, A.E. 1991. The aposymbiotic aphid: an analysis of chlortetracycline-treated pea aphid, *Acyrthosiphon pisum*. *Journal of Insect Physiology* **37**: 713–719.

Remaudière, G. and Remaudière, M. 1997. Catalogue des Aphididae du monde. Homoptera. Aphidoidea. INRA, Paris.

Rouhbakhsh, D., Lai, C.Y., von Dohlen, C., Clark, M.A., Baumann, L., Baumann, P., Moran, N.A., and Voegtlin, D.J. 1996. The tryptophan biosynthetic pathway of aphids endosymbionts (*Buchnera*): genetics and evolution of plasmid-associated anthranilate synthase (*trp*EG) within the Aphididae. *Journal of Molecular Evolution* **42**: 414–421.

Sabater-Muñoz, B., Gómez-Valero, L., van Ham, R.C.H.J., Silva, F.J., and Latorre, A. 2002. Molecular characterization of the leucine cluster in *Buchnera* sp. Strain PSY, a primary endosymbiont of the aphid *Pemphigus spyrothecae*. *Applied and Environmental Microbiology* **68**: 2572–2575.

Schmidt, H.A., Strimmer, K., Vingron, M., and von Haeseler, A. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**: 502–504.

Shigenobu, S., Watanabe, H., Hattori, M., and Ishikawa, H. 2000. Genome sequence of the endocellular bacterial symbionts of aphids *Buchnera* sp. APS. *Nature* **407**: 81–86.

Silva, F.J., van Ham, R.C.H.J., Sabater-Muñoz, B., and Latorre, A. 1998. Structure and evolution of the leucine plasmids carried by the endosymbionts (*Buchnera aphidicola*) from aphids of the family Aphididae. *FEMS Microbiology Letters* **168**: 43–49.

Silva, F.J., Latorre, A., and Moya, A. 2001. Genome size reduction through multiple events of gene disintegration in *Buchnera* APS. *Trends in Genetics* **17**: 615–618.

Silva, F.J., Latorre, A., and Moya, A. 2003. Why are the genomes of endosymbiotic bacteria so stable? *Trends in Genetics* **19**: 176–180.

Soler, T., Latorre, A., Sabater, B., and Silva, F.J. 2000. Molecular characterization of the leucine plasmid from *Buchnera aphidicola*, primary endosymbionts of the aphid *Acyrtosiphon pisum*. *Current Microbiology* **40**: 264–268.

Tamas, I., Klasson, L., Canback, B., Naslund, A.K., Eriksson, A.S., Wernegreen, J.J., Sandstrom, J.P., Moran, N.A., and Andersson, S.G. 2002. 50 million years of genomic stasis in endosymbiotic bacteria. *Science* **296**: 2376–2379.

van Ham, R.C.H.J., Gonzalez-Candelas, F., Silva, F.J., Sabater, B., Moya, A., and Latorre, A. 2000. Postsymbiotic plasmid acquisition and evolution of the repA1-replicon in *Buchnera aphidicola*. *Proceedings of the National Academy of Science USA* **97**: 10855–10860.

van Ham, R.C.H.J., Kamerbeek, J., Palacios, C., Rausell, C., Abascal, F., Bastolla, U., Fernandez, J.M., Jimenez, J.M., Postigo, L., Silva., F.J., Tamames, J., Viguera, E., Latorre, A., Valencia, A., Morán, F., and Moya, A., 2003. Reductive genome evolution in *Buchnera aphidicola*. *Proceedings of the National Academy of Science USA* **100**: 581–586.

van Ham, R.C.H.J., Martínez-Torres, D., Moya, A., and Latorre, A. 1999. Plasmid-encoded anthranilate synthase (trpEG) in *Buchnera aphidicola* from aphids of the family Pemphigidae. *Applied and Environmental Microbiology* **65**: 117–125.

van Ham, R.C.H.J., Moya, A., and Latorre, A. 1997. Putative evolutionary origin of plasmids carrying the genes involved in leucine biosynthesis in *Buchnera aphidicola* (endosymbiont of aphids). *Journal of Bacteriology* **179**: 4768–4777.

von Dohlen, C.D. and Moran, N.A. 2001. Mitochondrial ribosomal sequences and fossils support a rapid radiation of aphids in the Cretaceous and multiple origins of a complex life cycle. *Biological Journal of the Linnean Society* **71**: 689–717.

Wernegreen, J.J., Ochman, H., Jones, I.B., and Moran, N.A. 2000. Decoupling of genome size and sequence divergence in a symbiotic bacterium. *Journal of Bacteriology* **182**: 3867–3869.