MOLECULAR EVOLUTION OF APHIDS AND THEIR PRIMARY (*Buchnera* sp.) AND SECONDARY ENDOsyMBIONTS: IMPLICATIONS FOR THE ROLE OF SYMBIOSIS IN INSECT EVOLUTION

Beatriz Sabater, Roeland C.H.J. van Ham, David Martínez-Torres, Francisco Silva, Amparo Latorre and Andrés Moya

Beatriz Sabater Muñoz. Bachelor in Biology. University of Valencia, Spain. e-mail: beatriz.sabater@uv.es.

Roeland C.H.J. van Ham. Ph.D. in Biology, University of Utrecht, The Netherlands. Postdoctoral researcher, Centro de Astrobiología (INTA-CSIC), Spain. e-mail: roeland.vanham@uv.es

David Martínez Torres. Ph.D. in Biology. Assistant Professor, University of Valencia, Spain. e-mail: david.martinez@uv.es

Francisco Silva Moreno. Ph.D. in Biology. Associate Professor, University of Valencia, Spain. e-mail: francisco.silva@uv.es

Amparo Latorre Castillo. Ph.D. in Biology. Associate Professor, University of Valencia, Spain. e-mail: amparo.latorre@uv.es

Andrés Moya Simarro. Ph.D. in Biology, Ph.D. in Philosophy and Full Professor, University of Valencia, Spain. Address: Instituto Cavanilles de Biodiversidad y Biología Evolutiva, Departamento de Genética, Universidad de Valencia, Spain, Apartado de Correos 2085, 46071, Valencia, Spain. e-mail: andrés.moya@uv.es

Summary

*Aphids maintain an obligate, endosymbiotic association with Buchnera sp., a bacterium closely related to*
Escherichia coli. Bacteria are housed in specialized cells of organ-like structures called bacteriomes in the hemocoel of the aphid and are maternally transmitted. Phylogenetic studies have shown that the association had a single origin, dated about 200-250 million years ago, and that host and endosymbiont lineages have evolved in parallel since then. However, the pattern of deepest branching within the aphid family remains unsolved, which thereby hampers an appraisal of, for example, the role played by horizontal gene transfer in the early evolution of Buchnera.

The main role of Buchnera in this association is the biosynthesis and provisioning of essential amino acids to its aphid host. Physiological and metabolic studies have recently substantiated such nutritional role. In addition, genetic studies of Buchnera from several aphids have shown additional modifications, such as strong genome reduction, high A+T content compared to free-living bacteria, differential evolutionary rates, a relative increase in the number of non-synonymous substitutions, and gene amplification mediated by plasmids. Symbiosis is an active process in insect evolution as revealed by the intermediate values of the previous characteristics showed by secondary symbionts compared to free-living bacteria and Buchnera.

Os pulgões mantém uma associação endosimbiótica obrigatória com Buchnera sp., uma bactéria estreitamente relacionada com Escherichia coli. A bactéria se encontra alojada em estruturas celulares chamadas bacteriomas no hemocele dos pulgões, e são maternamente transmitidas. Os estudos filogenéticos indicam que a associação teve uma única origem há aproximadamente 200-250 milhões de anos, e desde então hospedeiro e endosimbionte têm evoluído em paralelo. Porém, o padrão de ramificação na família dos áfidos não foi resolvido, o qual impede a avaliação de, por exemplo, o papel da transferência horizontal de gens na evolução precoce de Buchnera.

O papel principal de Buchnera nesta associação é a biosíntese e aprovisionamento de aminoácidos essenciais a seu áfido hospedeiro. Esta afirmação está apoiada por estudos fisiológicos, metabólicos e recentemente por estudos genéticos. As transformações genéticas experimentadas por Buchnera em sua adaptação à vida endosimbiótica são variadas. Entre elas, a redução de seu genoma comparando com bactérias de vida livre, o incremento em A+T, taxas de evolução diferentes, incremento relativo das taxas de substituição não sinônimas e a amplificação gênica mediada por plásmidos. Mas a endosimbiose é um processo ativo nos insetos, como é comprovada pelo estado intermediário dos endosimbiontes secundários, os quais apresentam valores intermediários das características anteriormente citadas em comparação com as bactérias de vida livre e Buchnera.

KEYWORDS / Aphids / Buchnera sp. / Secondary Endosymbionts / Phylogenetic Analysis / Genome Reduction /


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
Molecular evolution of aphids and their primary 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. In most of the cases one member of the association is an eukaryote. Contrary to most prokaryotes, eukaryotes have rather limited metabolic capabilities and hence, symbiosis has provided and continues to provide an evolutionary strategy through which eukaryotes have access to a wider range of metabolic resources.

One special case of symbiosis is endosymbiosis, where the prokaryote is literally sequestered within an eukaryotic cell. The paradigm of endosymbiosis is the origin of the eukaryotic cell, in which different bacteria successively entered a protoeukaryotic cell giving rise to mitochondria first, and chloroplasts later (Margulis, 1970; Fisher, 1989; Gray, 1989). Insects are particularly receptive to endosymbiotic processes. It has been estimated than more than 10% of the extant insect species harbor endosymbionts at different stages of accommodation to their host. Some of them are recent and casual, but others are well established associations that are as relevant to the insect and its physiology as true organelles are for an eukaryotic cell (Buchner, 1965).

Empirical research on symbiosis, specially at the molecular level, is a growing and promising field. The impact of symbiosis on active speciation processes, on the genomic transformations of both partners, and on the evolutionary success of many taxonomic lineages is beginning to be unveiled.

With respect to the role played by symbiosis in evolution, there are two prevailing views. The first acknowledges that symbiosis is an important component in the biology of many organisms, but yet does not consider it to be a major evolutionary force per se, because many mutualistic associations have limited chances of promoting significant evolutionary innovation (Maynard Smith, 1989). Margulis (1991) holds a completely opposite view and broadly argues against the Neodarwinian theory as an incomplete theory of evolution. Her argument is based on the scarcity of cases of observed speciation and she maintains that the major taxonomic groups (i.e. phyla, kingdoms) have evolved through engagements in symbiotic associations. According to Charlesworth (1991) Margulis’ view is rather radical. Although it is possible to admit evolutionary accommodations after a symbiotic event between two species, there is not an a priori reason to discard compatibility between this fact and the presence of natural selection as a driving force promoting the evolution of the symbiotic association. Moreover, the relevance of symbiotic events determining the appearance of major evolutionary lineages or more precisely, the role played by symbiosis in traits that define higher taxa, should be evaluated.

**Endosymbiosis: A Classification**

In most, if not all, endosymbiotic associations one of the two partners is eukaryotic (Table I). Although they display great diversity and morphological complexity, eukaryotes show limited metabolic capabilities, are rarely anaerobic, lack photosynthesis (except for plants and some protists), cannot fix nitrogen, and some groups have lost the ability to synthesize essential amino-acids and coenzymes necessary for their basic metabolism. Some eukaryotes solve these limitations by means of a symbiotic association with another eukaryotic or prokaryotic organism. The examples shown in Table I should be considered as recent endosymbiotic events compared with respiration and photosynthesis carried out by
eukaryotic cells, where mitochondria and chloroplasts are endosymbionts derived from ancestral free living proteo- and cyanobacteria, respectively.

**TABLE I**

<table>
<thead>
<tr>
<th>Metabolic novelty</th>
<th>Symbiont</th>
<th>Host</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photosynthesis</td>
<td>Algae and Cyanobacteria</td>
<td>Several protists, invertebrates, lichenized fungi</td>
</tr>
<tr>
<td>Chimiosynthesis</td>
<td>Gamma-proteobacteria</td>
<td>Several invertebrates</td>
</tr>
<tr>
<td>N₂ fixation</td>
<td><em>Rhizobium, Frankia</em> and cyanobacteria</td>
<td>Several plants</td>
</tr>
<tr>
<td>Methanogenesis</td>
<td>Methanogenic bacteria</td>
<td>Some protists</td>
</tr>
<tr>
<td>Celulases</td>
<td>Bacteria, ciliated protists and phycomecetes fungi</td>
<td>Herbivore vertebrates and termites</td>
</tr>
<tr>
<td>Luminiscence</td>
<td><em>Vibrio and Photobacterium</em></td>
<td>Teleosts, cephalopods, and some insects, protists</td>
</tr>
<tr>
<td>Synthesis of essential nutrients</td>
<td><em>Bacteria</em></td>
<td>Some insects, protists</td>
</tr>
</tbody>
</table>

**Aphids and their Endosymbionts**

Within insects, the Homoptera are a highly specious group, with more than 30000 extant species (Remaudière and Remaudière, 1997). The most relevant superfamilies within this order are Phylloseroidea, Aphidoidea, Psylloidea, Aleurodoidea, and Coccoidea.

With more than 4400 extant species, the superfamily Aphidoidea is one of the most diversified and widely distributed groups of Homoptera (Blackman and Eastop, 1994). Representatives occur from Ecuador to the Polar Circle, with the Temperate Zones showing the highest diversity. Their evolutionary success might be attributed to their ability to suck phloem from plants or from the acquisition of bacterial endosymbionts. Both facts seem to be related because the phloem diet, poor in essential amino acids is compensated by essential nutrients supplied by endosymbionts (Buchner, 1965; Douglas, 1989; Fukatsu et al., 1994). Aphid populations are enormous in size, with densities reaching up to 5x10⁵ aphids/m². This reproductive potential is due primarily to their exceptional mode of reproduction. Beside the sometimes complex alternation of sexual and non-sexual phases, many species of aphids pass through a cycle of viviparous, parthenogenetic reproduction. In this cycle, the phenomenon of "telescopic generations" is often observed, in which embryogenesis starts in yet unborn daughters, resulting in the enclosure of up to three generations developing within an adult individual. These high reproductive rates make of aphids an important agricultural pest, promoting in their hosts the loss of nutrients, sheet ruffling, galls, and other forms of structural damage. In addition, they are vectors of important plant viral diseases.
The phloem which aphids feed upon is rich in sugars but poor in nitrogen compounds. In particular, the diet is deficient in essential amino acids, vitamins and several lipids that are also essential (Dixon, 1975; Houk and Griffiths, 1980; Raven, 1983; Sasaki et al., 1991; Douglas, 1993).

Species of the genus *Buchnera* are the primary endosymbionts of aphid species and they are required for its hosts’ normal development and fecundity. These bacteria are located in specialized, polyploid cells, called bacteriocytes that form an organ (bacteryome) localized in the hemocoel above the digestive tube. They are maternally inherited by means of infection of the embryos during the blastoderm stage (Hinde, 1971). The symbiosis is an example of obligatory mutualism, in the sense that the aphid cannot survive without *Buchnera* and *Buchnera* cannot survive outside the aphid. For example, aposymbiotic aphids (aphids treated with antibiotics) have a slow growth rate and have no offspring, and *Buchnera* spp. cannot be cultured in artificial media (revised in Bauman et al., 2000). Several experiments have led to the conclusion that *Buchnera* species are the main source of essential amino acids for aphids, and responsible in part of the high growth rate of their hosts (Prosser et al., 1992; Sasaki et al., 1991; Liadouze et al., 1995).

### Comparative Phylogenetics of *Buchnera* sp. and Aphids

**Analysis of Buchnera spp.** 16S rDNA sequences from different aphid species supported their monophyly within the γ3 subdivision of the class proteobacteria, having *Escherichia coli* and related members of the enterobacteriaceae as their closest known relatives (Moran et al., 1993). In addition, the 16S rDNA based phylogeny of representatives of *Buchnera* from four aphid families showed complete concordance with the morphology-based phylogeny of their aphid hosts (Moran and Baumann, 1994). These results pointed to a single original infection in a common ancestor of all modern Aphidoidea occurred approximately 100-250 MY followed by co-speciation of aphids and *Buchnera* (Moran et al., 1993).

Several molecular phylogenetic studies using different genes from *Buchnera* have also shown a congruence with the morphology-based phylogeny of aphids (Brynnel et al., 1998; Baumann et al., 1999; van Ham et al., 1999, 2000). However, in most of these phylogenetic studies taxonomic sampling was limited to representatives of the family Aphididae. Indeed, van Ham et al. (1997) proposed a topology for the phylogenetic relationship among 16S rDNA sequences of *Buchnera* from some aphid families which might question either the traditionally accepted phylogeny of aphids or the perfect parallelism between aphids and their primary endosymbiont phylogenies (see Figure 1). In fact, since most of the molecular phylogenetic reconstructions done with *Buchnera* were in good agreement with the accepted phylogeny of their host (Heie, 1980; 1987), the latter was never questioned before the report by van Ham et al. (1997). Moran et al. (1995) compared the evolutionary rates in bacteria and their host, but the study was restricted to three aphid families. More recently, von Dohlen and Moran (2001) have used mitocondrial ribosomal DNA sequences to reconstruct the phylogeny of aphids. They found, in general, good correlation with morphological data at the level of tribes. Finally, Martínez-Torres et al. (2001) have carried out a large study using genes from either *Buchnera* and aphid species representative from five aphid families, including Lachnidae. Although the results of Martínez-Torres et al. (2001) do not define completely evolutionary relationships among aphid families, some traditionally accepted groupings are questioned from both bacterial and insect data. In particular the Lachnidae and the Aphididae, which on the basis of morphological data are considered to be sister groups, do not seem to be as closely related as expected.
Molecular changes experienced by primary (*Buchnera* sp.) and secondary endosymbionts

*Buchnera* is present in all aphid species studied to date, with the exception of some species from the Hormaphididae family (Baumann *et al*., 1995). Molecular and genetic studies have shown that *Buchnera* overexpresses the essential amino acids tryptophan and leucine by means of gene amplification into plasmids (Lai *et al*., 1994; Bracho *et al*., 1995; Bauman *et al*., 1997; 1999; van Ham *et al*., 1997; 1999; 2000; Silva *et al*., 1998). The characteristic and evolutionary history of these plasmids are complex as not all the families of aphids carry plasmids and not all plasmids have the same gene content and/or gene order. In the case of tryptophan, the genes *trpEG*, the two first genes of the pathway, are carried in several copies, arranged as tandem repeats on a low copy-number plasmid from species of the family Aphididae. The remaining genes of the pathway are chromosomal (Lai *et al*., 1994). The amplification of the same genes has been shown by van Ham *et al*. (1999), but in a different replicon in *Pemphigus spyrothecae*, a member of the Pemphiginae, the most divergent family of aphids (Heie, 1980), whereas in *Schendalia chinensis*, a Pemphigidae belonging to a different tribe, *trpEG* was shown to be chromosomally encoded, though unlinked from the remaining genes from the pathway (Figure 1).

![Figure 1. Phylogenetic trees based on *Buchnera* 16S rDNA gene (van Ham *et al*., 1997) and morphological traits and geological data records from aphids (Heie, 1980; 1987).](http://www.scielo.org.ve/scielo.php?pid=S0378-184420010001000015&script=sci_arttext (6 of 14)1-9-2006 13:24:41)

In the case of leucine plasmids, only one single replicon, named *repAl* has been found, but the gene content and/or gene order is different in different lineages of the family Aphididae, indicating a great plasticity of the Leu plasmids during *Buchnera* evolution. The first leucine plasmid was described (Bracho *et al*., 1995) in *Buchnera* from *Rhopalosiphum padi* (pLeu-BRp) a member of the Aphidinae subfamily. It contains the genes *leuABCD*, for the biosynthesis of leucine, two copies of *repA* and ORF1, a putative integral membrane protein. The same gene content and in the same order was found in other species from the same subfamily (Bauman *et al*., 1999). However, a different gene arrangement has been
found in pLeu-BPp, the plasmid of *Pterocomma populeum*, a species belonging to Pterocommatini, the most divergent tribe of the Aphididae family (Silva et al., 1998). The plasmid of *Buchnera* of *Thelaxes suberi* (pLeu-BTs), a member of a different family (Thelaxidae) differed from the previous by the presence of a small heat shock gene (*ibp*) and in the order of the *leu* and *repA* genes (van Ham et al., 1997). Finally, the plasmid of *Buchnera* of the pemphgid *Tetraneura caerulescen* (pLeu-BTc) encodes only one *repA* copy plus ORF1, probably representing the ancestral replicon, related to the IncFII plasmids in which the other genes were relocated (van Ham et al. 1997) (see Figure 2).

**Figure 2.** Genetic organization of the leucine and tryptophan gene cluster in *Buchnera* from five aphid families. *E. coli* is used as reference of bacterial operon. Localization: chromosomal (c), plasmidic (p).

The complete genome of *Buchnera* from *Acyrthosiphon pisum* has shown the presence of a set of biosynthetic genes that are absent in intracellular pathogenic bacteria, suggesting that most of these genes may be involved in the provision of small molecules to the host (Shigenobu et al., 2000). This has been demonstrated at least in the case of the vitamin riboflavin (Nakabachi and Ishikawa, 1997; 1999).

Wernegreen and Moran (1999) have carried out a large study of the effects of selection and random genetic drift on the evolution of several loci of *Buchnera*, as compared to free-living bacteria. The major conclusion of their study, which also supports previous findings (Moran, 1996), is that random drift affects the potential effect of selection in the evolution of *Buchnera* in a crucial way. Due to its maternal inheritance, the effective population size of *Buchnera* is much lower than the effective population size of free-living bacteria. This has at least three important evolutionary consequences at the molecular level: a negative effect on the process of optimization of codon usage, an increase in the number of non-synonymous substitutions, and an acceleration of the evolutionary rate. According to Li (1987), the coefficient of selection needed to drive a population from a less to a more efficient codon usage (i.e.,
translational selection) should be larger than \(2/Ne\). Hence, those populations having great effective size \((Ne)\) will experience the effects of translational selection with weaker coefficients of selection than those others with less effective population size.

Besides *Buchnera* as primary symbiont, a number of aphids are known to contain other intracellular bacteria denoted as secondary symbionts (Buchner, 1965; Fukatsu and Ishikawa, 1993; 1998; Sandström *et al.*, 2001). The location, presence and pattern of inheritance of the secondary endosymbionts in aphids have been less studied and most of the current information has been obtained from those of *A. pisum* and *Macrosiphum rosae* (Unterman *et al.*, 1989; Chen and Purcell, 1997; Fukatsu *et al.*, 1998; 2000;). *In situ* hybridization and electron microscopy revealed that, at least in *A. pisum*, secondary endosymbionts are located in a different type of bacteriocytes adjacent to those formed by *Buchnera* (Fukatsu *et al.*, 2000).

Available data indicate that secondary endosymbionts have been acquired several times during aphid evolution and, contrary to the origin of *Buchnera*, they do not represent a monophyletic clade (Fukatsu and Ishikawa, 1993; Moran and Telang, 1998). Recently, three types of bacteria belonging to the Enterobacteriaceae have been identified in several species from the Macrosiphini tribe, named U-, R-, and T-type (Sandström *et al.*, 2001). The phylogenetic distribution of secondary endosymbionts indicates occurrence of repeated horizontal transmissions.

A+T content, phylogenetic relationships, codon usage, evolutionary rates, and ratio of synonymous versus non-synonymous substitutions have been studied in partial sequences of the *atpD* and *aroQ/pheA* genes of primary (*Buchnera*) and putative secondary endosymbionts of aphids and a set of selected non-symbiotic bacteria, belonging to the five subdivisions of the Proteobacteria. Compared to the homologous genes of the last group, both genes belonging to *Buchnera* behave in a similar way, showing a higher A+T content, forming a monophyletic group, a loss in codon bias, especially in third base position, an evolutionary acceleration and an increase in the number of non-synonymous substitutions. Some of the putative secondary endosymbionts show a high A+T content, an intermediate, unresolved phylogenetic position between *Buchnera* and the other \(\gamma\)-Proteobacteria, a loss in codon bias, and a significant evolutionary acceleration in the case of the three-*atpD* genes. These results lead us to conclude that some of them are true endosymbionts at different stages of symbiotic accommodation to the host (Moya *et al.*, unpublished results).

**Genome Reduction of Buchnera sp.**

The comparative analysis of *Buchnera* sp. APS genome (Shigenobu *et al.*, 2000), with respect to those from *E. coli* and *Vibrio cholerae*, the two closest free living bacteria whose genomes have been sequenced, reveal that the endosymbionts not only have experienced large genome reductions, but also large genome reorganizations (Silva *et al.*, 2001). This is inferred when considering that synteny is only preserved in a number of short length fragments, being the average number of genes per fragment larger when comparing *Buchnera* versus *E. coli* than when comparing with *V. cholerae*, as expected if we consider that *E. coli* is closer phylogenetically to *Buchnera* than to *V. cholerae*. In order to determine if the genome reduction involved large genome fragments, we divided the circular *E. coli* chromosome in ten sections, being the first center on the origin and the sixth on the replication terminus, and screened the *Buchnera* chromosome for orthologous protein coding genes. The number of genes of each section (*Figure 3*) showed that the actual *Buchnera* genes come from every section of the *E. coli* chromosome, except the region around the terminus. A more detailed examination shows that there is only 4
orthologous genes in the *E. coli* region between *ter site* C and A. This large absence could be due to a large deletion of a region higher than 300kb in the ancestor of *Buchnera*. However, the fact that a similar circumstance occurred in the comparison with *V. cholerae* chromosome ([Figure 3](#)), and that the region around *ter sites* is not well-conserved between the *E. coli* strains K-12 and O157:H7 (Perna *et al*., 2001) points to this region being very unstable and subjected to lateral transfer events.

![Figure 3. Number of orthologous genes that have been found in ten regions of the *Buchnera* sp APS genome (see text for more details) when compared with *E. coli* and *V. cholerae* genomes.](http://www.scielo.org.ve/scielo.php?pid=S0378-18442001001000015&script=sci_arttext)

**Discussion and Perspectives**

Although it is apparent that symbiosis has played a major role in the adaptation and evolution of insects, theoretical problems remain as to how the accommodation process of endosymbionts has taken place. Primary endosymbionts are maternally inherited; they have small effective population sizes and the continuous bottlenecking they experience drives the increment of the average number of deleterious mutations they carry. The effects of this process on endosymbiotic genomes, known as Muller’s ratchet, have been demonstrated in *Buchnera* (Moran, 1996), but also in secondary endosymbionts of aphids as well (Moya *et al*., unpublished results). Such process should finish with the extinction of the corresponding bacterial lineages and, concomitantly, with the insects harboring them. However, *Buchnera* and other primary endosymbionts of insects have survived for over 100 MY. From that we can hypothesize that some bacterial lineages have evolved one or more compensatory mechanisms to solve the problem of accumulation of deleterious mutations. Genome reduction or active presence of proteins able to compensate defective gene functions have been advanced by Wernegreen and Moran (1999). Demonstrating the active presence of these and other factors is of particular relevance in the future research in endosymbiont evolution.

An intriguing aspect of symbiont evolution in insects is the potential interplay between old-primary and recent-secondary endosymbionts, which also deserves research. As already mentioned the presence of secondary endosymbionts indicates that new symbiotic accommodation processes are taking place in insect evolution. The point then is in what respect secondary endosymbionts are replacing and/or complementing metabolic functions of the primary endosymbionts. In the case of *Buchnera*, for instance, it is of great relevance to clarify if the fate of *Buchnera* is to become replaced by a new generation of
symbionts. Although it is too preliminary, we have evidence (unpublished results) that the genome size of *Buchnera* is not uniform across aphid lineages (see however Wernegreen *et al.*, 2000), and that depending on the active metabolic presence of other symbionts *Buchnera* might evolve towards an increasing genome reduction and, eventually, go to extinction.

**ACKNOWLEDGEMENTS**

The authors thank the *Servicio Central de Soporte a la Investigación Científica* (Universitat de València) for bioinformatics and sequencing facilities. The work has been supported by grants GV-3216/95 from Generalitat Valenciana (Spain), BFM2000-1383 from MEC (Spain) and FEDER 2FD1997-1006 (Spain).

**REFERENCES**


