Looking for Darwin in Endosymbionts

Mario A. Fares*

Evolutionary Genetics and Bioinformatics Laboratory, Department of Genetics, University of Dublin, Trinity College, Dublin, Ireland

Abstract: Darwinian evolution has championed mechanistic models explaining the forces shaping species diversification. This principle however has been challenged by tantalizing observations in many organismal systems where selection seemed to have a minor role. One such a system is the case of endosymbiotic bacteria of insects that present evolutionary features only conceivable if stochastic variation over-weights selection. This is because these bacteria present proteomes with high rates of evolution that have driven the biological system to the edge of catastrophe error. In this review I discuss and demonstrate, based on recent emerging literature and other observations, that rather than a fact this stochasticity is only apparent. Most of the variation patterns can be explained if we introduce new variables in the equation that defines selection-drift balance such as translational robustness, functional divergence and epistasis. The main conclusion of this review is that, endosymbiotic bacterial proteomes have evolved an evolutionary complex proteome/interactome system robust to mutations.

INTRODUCTION

Darwinian view of evolution has revolutionized studies aimed at understanding adaptation and biological innovation. Ever since the publication of Darwin's work in the origin of Species (1957), a burgeoning research literature has examined a myriad of different biological systems to demonstrate the very basic principles of natural selection. In general terms, organisms evolve under strict and gradual purifying selection intermingled with bursts of punctually fixed advantageous mutations (Diversifying selection). Purifying selection removes functionally/structurally unstable mutations preserving therefore proteins' functions [1]. Darwinian diversifying selection on the other hand contributes to the emergence of new protein functions [2]. The neutral and nearly neutral theory of molecular evolution [3, 4] has complemented and provided the mathematical background to understand how mutations are fixed in proteins. An apparently counter-intuitive conclusion derived from these studies is that most of the mutations are neutrally fixed in proteins. The actual conception yielded by a plethora of analyses is that natural selection acts through three mechanisms whereby the majority of mutations are removed by selection because of their harmful effects on proteins. A very rare fraction of these mutations are fixed by adaptive evolution and the vast majority of these mutations are fixed neutrally by genetic drift. Much has been gained from these studies although challenges to Darwinian theories have led to heated debates about whether selection or drift is the responsible for the emerging biological complexity. The importance of natural selection as the central organizing principle of biology has been brought into question many times. For example, the evolutionary characteristics of the ubiquitous insect heritable symbiotic bacteria have posed a challenge to the universality of this principle. Endo-cellular symbiotic bacteria of insects that are housed in specialized cells formed during the ontogeny of the insect host [5] star this challenge. Due to their clonal and strict vertical transmission between host generations, these bacteria present small effective population sizes, reduced genomes, and AT positive bias [6-9]. The universality of these characteristics among endo-cellular symbiotic bacteria of insects, have converged into the symbiosis syndrome that is mostly characterized by observations that has challenged the Darwinian selection view during the last decade.

CAN WE FIND DARWIN AMONG SYMBIONTS?

As mentioned above, primary symbiotic bacteria of insects, such as the intra-cellular symbiont of the aphid insect *Buchnera aphidicola*, are characterized by small effective population sizes owing to their vertical clonal transmission among host generations. Reduced effective population sizes have the effect of increasing the rate of fixation of slightly deleterious mutations in populations as a result of genetic drift. This has been observed to irreversibly occur in primary symbiotic bacteria of insects. In fact, comparison of the evolutionary rates of protein–coding genes between *Buchnera aphidicola*, the primary symbiotic bacterium of aphids, and their free-living relatives (Fig. **1A**) shows greater accelerated evolution in endosymbionts (Fig. **1B**). The irreversible neutral fixation of mutations at these genomes has been interpreted to be an example of Muller's ratchet [7, 8, 10-14]. In theory, in small population sizes the chance to fix individuals with less mutations (for example, they reproduce less) is enhanced [15] and the mutations are then fixed in an irreversible manner [16, 17]. This effect has been a general trend in endosymbiotic bacteria of insects leading to an irreversible rampant increase of the genome mutational load. Increased fixation of slightly deleterious mutations in these bacteria has been shown to lead to low levels of intra-specific polymorphism [13, 18], although high levels of inter-specific variation (Fig. **2**), and low stability of RNAs [19]

*Address correspondence to this author at the Evolutionary Genetics and Bioinformatics Laboratory, Department of Genetics, University of Dublin, Trinity College, Dublin 2, Dublin, Ireland; Tel: +353 1 8963521; E-mail: faresm@tcd.ie
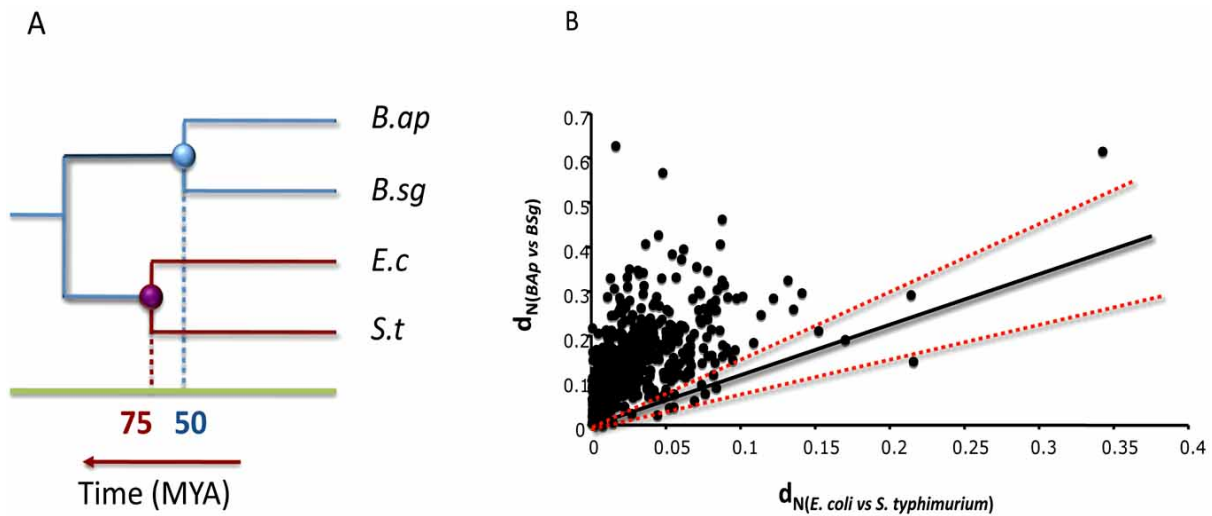
Fig. (1). Accelerated evolution in endosymbiotic bacteria of aphids. I compared protein evolutionary rates in *Buchnera aphidicola*, the primary symbiotic bacterium of aphids, between strains *Acyrthosiphon pisum* (*B.ap*) and *Schizaphis graminum* (*B.sg*), to those in their closest free-living cousins *Escherichia coli* (*E.c*) and *Salmonella typhimurium* (*S.t*) because they present similar divergence times [8, 81] (**A**). (**B**) The pairwise non-synonymous substitutions was estimated by maximum-likelihood using Goldman and Yang model implemented in CODEML program from the PAML package version 4.0 [82]. The black continuous line and the two red dashed lines represent the confidence area in which there is no evidence for the difference in the evolutionary rates between the endosymbionts and the free-living bacteria. This graph is based on the analysis of 510 genes from B.ap and B.sg that present orthologs in *Ec* and *St*. This figure is adapted from analyses performed in a previous thesis [83].
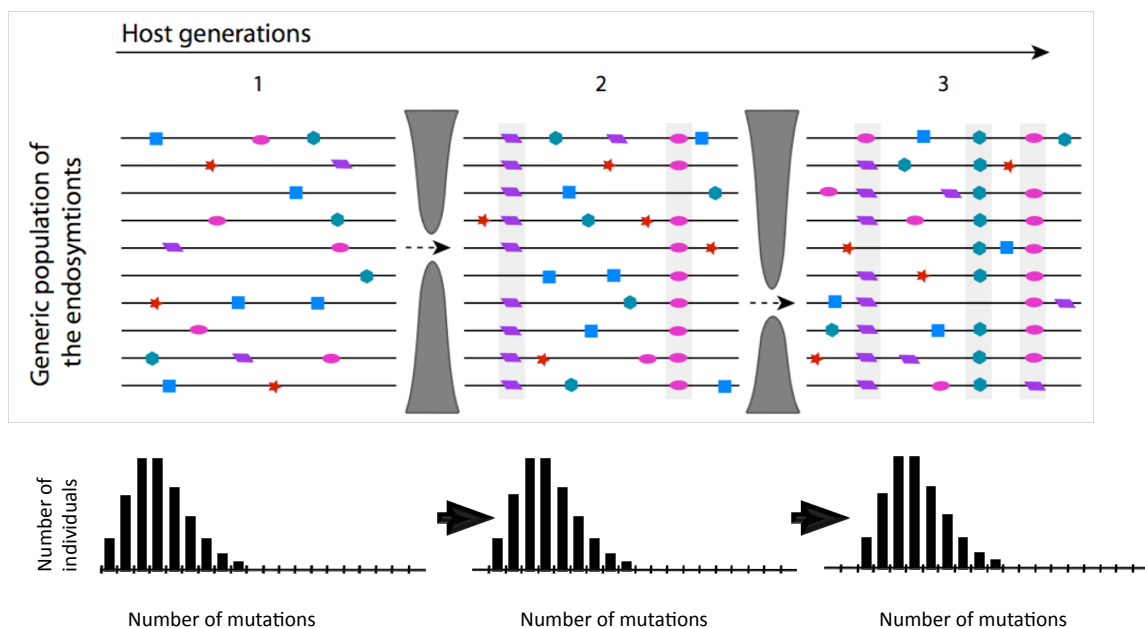
Fig. (2). Dynamics of Muller's ratchet in endosymbiotic bacteria. Slightly deleterious mutations accumulate in endosymbiotic bacteria of insects due to the bottlenecks to the effective population sizes during their intergenerational transmissions. This leads to an increased load of slightly deleterious mutations fixed by genetic drift while intra-specific polymorphism is drastically reduced.

and of proteins [20]. Given these evidence it seems reasonable to assume that genetic drift has apparently led the evolution of this system and that little role for selection is expected. However, a stochastically drifting system is condemned to its demise unless selection is imposed in the biological systems resulting from the combination of both the insect host and the bacterium. The question remaining therefore is whether genetic drift is the leading evolutionary force of these bacteria or whether such simple genomes hide a complex dynamic of selective constraints difficult to highlight given the amount of stochastic evolutionary noise at these genomes. In favor of the hypothesis that genetic drift is acting is the fact that such genomes show extremely accelerated rates of evolution and drastic genome downsizing resulting from an apparent random gene non-functionalization and subsequent disintegration [6, 17, 21-25]. Few lines of evidence however unearth Darwinian characteristics among the apparent stochastic evolutionary noise and argue hence

against random non-functionalization. For example, in spite of that many endosymbiotic genomes such as *Buchnera*, *Wigglesworthia* and *Baumannia* are reduced up to 20% of the ancestral free-living relative bacterium *Escherichia coli* [23, 26, 27], most of the disintegrated genes represent a redundant pool of functions for the bacterium and the host [28]. Conversely, endosymbionts such as *Buchnera* have conserved genes from metabolic pathways essential for the host [29]. Most of these genes seem to have suffered little to no modifications since their ancestral origin, something that was striking after the early discovery in *Buchnera* of rate-limiting biosynthetic genes for tryptophan (*trpEG*, [30]) and for leucine (*leuABCD*, [31]) and the subsequent confirmation of the evolutionary stasis of these genes [32-34]. Moreover, genomes of obligate symbionts of insects present examples of modifications of an ancestral genome driven by the co-adaptation with the insect host [35]. Although it becomes obvious that a balance exists between drift and selection, a mechanistic explanation is required in order to fully understand the detailed parametric properties of such balance.

GroEL: THE SENTINEL OF THE ENDOSYMBIOTIC PROTEOME

One of the most striking findings that accompanied the demonstration of the high evolutionary rates in *Buchnera* [7] was that the molecular chaperonin GroEL was over-expressed in most if not all of the intra-cellular symbiotic genomes [36]. GroEL is codified by the operon *groE* (*mopBA*) and is the best characterized chaperonin to date [37-39]. This chaperonin is specifically found in bacteria and eukaryotic organelles and presents great functional flexibility as it folds proteins with a wide range of sizes and intermediate folding stages [40, 41]. GroEL also has a remarkable ability to unfold and subsequently fold proteins into their native conformation [42]. This protein also constitutes up to 10% of the *Buchnera* proteome [36]. The fact that GroEL is able to suppress the destabilizing effects of temperature sensitive mutations [43] fuelled the hypothesis that GroEL may be the stabilizer of the endosymbionts proteomes. In an experimental simulation study where the free-living relative of many endosymbiotic bacteria, the proteobacterium *Escherichia coli*, was subjected to strong bottlenecks between generations, over-expression of GroEL permitted the almost total recovery of the fitness of highly compromised strains [44]. Rather than a particularity, this phenomenon seems to be the rule in bacteria presenting declined fitness due to protein-structure destabilizing mutations [45, 46]. Finally, the low number of slightly deleterious mutations fixed in GroEL [47], the strong evidence for purifying selection as well as the fixation of advantageous mutations at amino acid sites of this chaperonin involved in protein folding [48, 49] support the important role of this protein in endosymbiotic cells. Is GroEL a realistic solution for the persistence of endosymbiotic genomes? Are there other mechanisms involved in the successful survival of these bacteria?

ADAPTIVE EVOLUTION IN ENDOSYMBIOTIC GENOMES

As discussed earlier, endosymbiotic genomes are subjected to strong AT enrichment and possible saturation of synonymous sites. Identifying selection using conventional methods (for example, the ratio between synonymous and non-synonymous substitution rates ω) may become a challenging endeavor. This is mostly due to the fact that such methods are based upon the assumption of neutrality at synonymous sites and therefore violation of such assumption deems such approach inappropriate. In endosymbiotic bacteria the saturation in third codon positions due probably to nucleotide bias (for example, a high percentage of AT nucleotides) can lead to underestimates of synonymous nucleotide substitutions and hence to inflated ω. Synonymous sites can also be under selection [50, 51], Moreover, in a recent manuscript authors have demonstrated that varying synonymous sites in a library of synthetic peptides that encode the green fluorescent protein GFP can influence gene expression [52]. All these lines of arguments in addition to the limited sensitivity of methods using ω to detect selection makes it convincing that indirect evidence may more likely indicate whether selection is ongoing in endosymbiotic bacterial genomes. Is there justifiable hope for identifying selection signatures in endosymbiotic bacterial genomes? The signals of selection in endosymbiotic bacteria could be easily confounded by genetic drift when we are at the species level. For instance, strong genetic drift effects due to population bottlenecks will dramatically increase the number of fixed mutations with slight fitness effects increasing hence the inter-species divergence. The challenge therefore comes from the possibility of identifying slightly advantageous mutations from those that are slightly disadvantageous. One possibility is the use of approaches that identify interaction relationships between mutations (epistasis). In any case, what is clear is that the pattern of genetic drift should affect the entire genome of endosymbionts while those under selection may pop out from the background mutational noise. As for linked selection signatures (for example, "hitchhiking", [53]), these are only possible to identify if the neutral mutation rates varies because divergence levels between two species does not depend on the effective population size [54] (see also [55] for a comprehensive commentary on the subject). Given these difficulties, it may be interesting exploring other ways to identify signatures of selection. Preliminary analyses of proteins have recently provided the first clues although further analyses are a must to provide direct evidence of the quantitative action of selection in the evolution of complex biological systems.

SELECTION FOR TRANSLATIONAL ROBUSTNESS IN ENDOSYMBIOTIC PROTEOMES

Endo-cellular symbiotic bacteria of insects are characterized by high fixation rates of slightly deleterious mutations, which mainly have protein structure destabilizing effects [20]. Continuous fixation of slightly deleterious mutations would eventually lead to genome wide non-functionalization, erosion and subsequent extinction of the biological system [14, 56-58]. The fact that actual genomes have successfully survived for tens to hundreds of millions of years challenges this outcome [59]. Other factors, either extrinsic, intrinsic or both, may therefore be keeping a stable equilibrium between drift and selection in this system. For instance, resistance to AT enrichment [60] and purifying selection in *Buchnera aphidicola* [61] strongly point to a slow down of Muller's ratchet. Several researchers sought to identify signals of factors counteracting genetic drift effects. Drummond and colleagues contributed to the explanation of variations in

the rates of protein evolution by demonstrating that such variation is correlated with the robustness of protein structures to mistranslation errors [62]. They showed that highly expressed proteins evolve faster because their structures are more resistant to unfolding or misfolding caused by the accumulation of mistranslation errors. Using this hypothesis, Toft and Fares tested whether selection favored proteins robust to structure destabilizing mutations fixed by genetic drift in the four fully sequenced genomes of *Buchnera aphidicola* strains *Acyrthosiphon pisum* (*B.ap*), *Schizaphis graminum* (*B.sg*), *Baizongia pistaciae* (*B.bp*) and *Cinara cedri* (*B.cc*) [63]. Their main findings support the hypothesis that translational robustness was indeed a driving force of endosymbiotic proteome evolution. Further, selection for translational robustness seemed to be correlated with genome downsizing and with essential functional categories. This pinpointed the idea that, in contrast to the general thought, *Buchnera* genomes are highly stable or at least they evolve towards minimizing their genome sizes to contain robust protein structures essential for intra-cellular survival [63]. It is worth mentioning that, despite the tendency of endosymbiotic proteomes to evolve toward having protein structures highly robust to mistranslation errors, these proteins may always wander at the edge of structure stability because this property is essential for protein evolvability. This is specially true for globular proteins that have shown a tendency to be marginally stable [64]. This has been further addressed by other researchers showing that in fact such property can be adaptively convenient since it provides proteins by the plasticity to develop new functions [65], which would support a further explanation for the possibility of functional divergence in endosymbiotic proteomes [66].

Intra-cellular symbiosis involved a revolution of life on earth and has been regarded as the main fuel of biological innovation and species diversity [67-69]. This process is a complex evolutionary phenomenon because the metabolisms of two biological systems with differing complexities fuse into one. Complex and rapid changes at the genome, proteome and metabolome levels may have therefore accompanied the symbiosis of the eukaryotic insect host and their prokaryotic bacteria. What are the molecular changes occurred and what type of selective forces shape these changes remain a matter of discussion. Nonetheless, advances have been made that aid us in understanding the complexity underlying the endosymbiotic events.

## WIDE-GENOME FUNCTIONAL DIVERGENCE IN ENDOSYMBIOTIC BACTERIA OF INSECTS

In most evolutionary studies, it is assumed that selective constraints on proteins are invariable through evolution. However, there are several scenarios under which changes in selective constraints may occur. Gene duplication is the most prominent case [70-78], although many other cases may lead to such outcome, including a change in the lifestyle of organisms. Symbiosis of proteobacteria with insects is a striking case of a revolutionized change in the lifestyle of a free-living bacterium to an intra-cellular dependent one. This change is dramatic because the bacterium no longer depend on the environmental conditions to acquire its products and metabolites and consequently not all the genes for such endeavors are needed any longer in the enclosed stable cellular environment they inhabit. There are two well-characterized

genome scale changes in endosymbiotic bacteria resulting from the intra-cellular life, including loss of redundant genes and accelerated rates of evolution. It is however expected that many changes in protein functions may have accompanied symbiosis stemming from the need for coupling the metabolic requirements of the host and the bacterium. Changes in protein functions are possible through changes in the selective constraints operating on these proteins (Functional divergence). A recent study has shown that indeed functional divergence has been fundamental to ensure the metabolic crosstalk of the endosymbiotic bacteria of aphids and carpenter ants with their insect hosts [66]. This functional divergence seems to have only affected the metabolic genes strictly involved in supplementing the host's diet deficient in essential amino acids of the insect hosts. Interestingly, genes affected by functional divergence varied depending on the ecological requirements of the host further demonstrating the metabolic link between insect and bacterium.

A MODEL FOR THE SURVIVAL OF ENDOSYMBIOTIC BACTERIA OF INSECTS

In the previous sections and throughout this review I show that endosymbiotic bacteria of insects undergo fixation of slightly deleterious mutations by genetic drift. Most of these mutations are expected to have destabilizing effects on protein's structures due to the stochastic nature of mutagenesis. Several mechanisms have had to evolve in these bacteria that may have ablated the demise of endosymbiotic lineages. In this review I propose translational robustness as a main mechanism intrinsic to the endosymbiotic proteome that may slow the Muller's ratchet. Nonetheless other mechanisms may have been involved to ultimately maintain equilibrium between destabilizing mutations while keeping a minimum genomic information, making these bacteria always wonder at the edge of error catastrophe. An elegant and simple mechanism is the accumulation of intra-protein and inter-protein compensatory mutations (antagonistic epistatic mutations). We need two types of mutations, inter- and intra-protein mutations to stabilize the biological system. First, mutations at the surface of proteins may force correlated fixation of structure surface mutations in interacting protein partners, keeping thereby protein-protein interfaces stable. This inter-protein antagonistic epistasis should be more dramatic with genome downsizing due to the reduced gene (functional) redundancy. Indeed, a previous study demonstrated that epistasis correlates with genome complexity [79]. Second, mutations within the protein structure require compensatory mutations nearby in the structure to maintain this structure stable while providing the evolutionary opportunity for the fixation of novel functional mutations necessary to adapt the bacterium to the host. It has been shown that functional divergence may have been a key event in the metabolic communication between the insect host and the bacterium. This means that the endosymbiotic genome has accumulated mutations that led to the functional divergence of a set of proteins involved in the interaction between the two biological systems. However, functionally innovative mutations are usually structurally destabilizing and therefore compensatory intra-molecular mutations may have been paramount for the stability of the functionally divergent proteins.

This line of argument leads consequently to a simple model to explain the persistence of endosymbiotic bacteria despite the effects of Muller's ratchet (Fig. **3**). In this model there are two non-mutually exclusive steps for the adaptation and metabolic marriage of insects and their endosymbiotic bacteria. First, the parasitic infection of the insect by a bacterium may have derived into an advantage due to a change in the ecological conditions so that the cost of holding a bacterium may have been largely compensated by the ability acquired by the host as a result of the infection to colonize new environments. This direct transition may have entrenched the host into a direct dependence on the bacterium for its survival. Consequently, insects with specialized cells with rich and stable environments housing symbiotic bacteria may have been selectively advantageous. With the strict addiction of symbiotic bacteria to stable intra-cellular environments, many of the genes essential for extra-cellular life and metabolites search became redundant. The population genetics dynamics of the bacterium meant that slightly deleterious mutations became fixed in symbiotic genomes at higher rates in non-essential genes than in essential genes. Some of these mutations became positively selected because they provided advantage to the host and the bacterium. Redundant genes became rapidly non-functionalized and subsequently disintegrated leading to highly streamlined genomes. The deregulation of the chaperonin GroEL and the possible accumulation of compensatory mutations counteracted the fitness decline of these bacteria as a result of the progressive accumulation
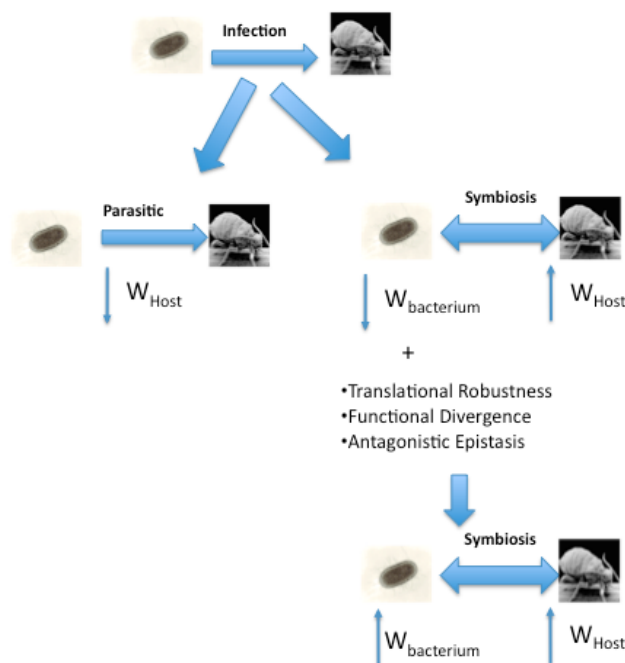
Fig. (3). Model of survival of endosymbiotic bacterial genomes. The initial infection of an insect by a bacterium may have repeatedly led to parasitic relationship that ended in extinction because of the compromised biological fitness of the host (W_{Host}). An subtle beneficial relationship between both these organisms may have led to the increase in W_{Host} but decreasing $W_{bacterium}$ owing to its population genetics dynamics. This may have been compensated by the evolution of proteins to acquire structures robust to mistranslation errors, protein functional divergence and antagonistic epistasis within and between proteins.

of mutations and loss of repair genes [29, 80]. Moreover, this compensatory event led to two main processes, including the functional divergence of genes and the emergence and fixation of mutations conferring structural robustness to proteins. In sum, the successful survival of endosymbiotic bacteria is a striking example of Darwinian selection hidden in complex evolutionary patterns resulting from the instability in the equilibrium between selection and drift.

ACKNOWLEDGMENTS

REFERENCES

[1] Messier, W.; Stewart, C.B. Episodic adaptive evolution of primate lysozymes. *Nature*, **1997**, *385*, 151-154.
[2] Gould, S.J.; Eldredge, N. Punctuated equilibrium comes of age. *Nature*. **1993**; 366, 223-227.
[3] Kimura, M. The neutral theory of molecular evolution. *Cambridge University Press, London*. **1983**.
[4] Ohta, T.; Kimura, M. On the constancy of the evolutionary rate of cistrons. *J. Mol. Evol.*, **1971**, *1*, 8.
[5] Allen, J.M.; Light, J.E.; Perotti, M.A.; Braig, H.R.; Reed, D.L. Mutational meltdown in primary endosymbionts: selection limits Muller's ratchet. *PLoS ONE.*, **2009**, *4*, e4969.
[6] Gil, R.; Latorre, A.; Moya, A. Bacterial endosymbionts of insects: insights from comparative genomics. *Environ. Microbiol.*, **2004**, *6*, 1109-1122.
[7] Moran, N.A. Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 2873-2878.
[8] Clark, M.A.; Moran, N.A.; Baumann, P. Sequence evolution in bacterial endosymbionts having extreme base compositions. *Mol. Biol. Evol.*, **1999**, *16*, 1586-1598.
[9] Lutzoni, F.; Pagel, M. Accelerated evolution as a consequence of transitions to mutualism. *Proc. Natl. Acad. Sci. USA*, **1997**, *94*, 11422-11427.
[10] Lynch, M. Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. *Mol. Biol. Evol.*, **1996**, *13*, 209-220.
[11] Lynch, M. Mutation accumulation in nuclear, organelle, and prokaryotic transfer RNA genes. *Mol. Biol. Evol.*, **1997**, *14*, 914-925.
[12] Brynnel, E.U.; Kurland, C.G.; Moran, N.A. Andersson, S.G. Evolutionary rates for tuf genes in endosymbionts of aphids. *Mol. Biol. Evol.*, **1998**, *15*, 574-582.
[13] Funk, D.J.; Wernegreen, J.J.; Moran, N.A. Intraspecific variation in symbiont genomes: bottlenecks and the aphid-buchnera association. *Genetics*, **2001**, *157*, 477-489.
[14] Rispe, C.a.M.; N.A. Accumulation of deleterious mutations in endosymbionts: Muller's ratchet with two levels of selection. *Am. Nat.*, **2000**, *156*, 425-41.
[15] Muller, H.J. The relation of recombination to mutational advance. *Mutat. Res.*, **1964**, *1*, 2-9.
[16] Felsenstein, J. The evolutionary advantage of recombination. *Genetics*, **1974**, *78*, 737-756.
[17] Lynch, M.; Gabriel, W. Mutation load and the survival of small populations. *Evolution*, **1990**, *44*, 1725-1737.
[18] Abbot, P.; Moran, N.A. Extremely low levels of genetic polymorphism in endosymbionts (Buchnera) of aphids (Pemphigus). *Mol. Ecol.*, **2002**, *11*, 2649-2660.
[19] Lambert, J.D.; Moran, N.A. Deleterious mutations destabilize ribosomal RNA in endosymbiotic bacteria. *Proc. Natl. Acad. Sci. USA*, **1998**, *95*, 4458-4462.
[20] van Ham, R.C.; Kamerbeek, J.; Palacios, C.; Rausell, C.; Abascal, F.; Bastolla, U.; Fernandez, J.M.; Jimenez, L.; Postigo, M.; Silva, F.J.; Tamames, J.; Viguera, E.; Latorre, A.; Valencia, A.; Moran, F. Moya, A. Reductive genome evolution in Buchnera aphidicola. *Proc. Natl. Acad. Sci. USA*, **2003**, *100*, 581-586.

[21] Gil, R.; Sabater-Munoz, B.; Latorre, A.; Silva, F.J.; Moya, A. Extreme genome reduction in Buchnera spp.: toward the minimal genome needed for symbiotic life. *Proc. Natl. Acad. Sci. USA*, **2002**, *99*, 4454-4458.

[22] Gil, R.; Silva, F.J.; Pereto, J.; Moya, A. Determination of the core of a minimal bacterial gene set. *Microbiol. Mol. Biol. Rev.*, **2004**, *68*, 518-537.

[23] Gil, R.; Silva, F.J.; Zientz, E.; Delmotte, F.; Gonzalez-Candelas, F.; Latorre, A.; Rausell, C.; Kamerbeek, J.; Gadau, J.; Holldobler, B.; van Ham, R.C.; Gross, R. Moya, A. The genome sequence of Blochmannia floridanus: comparative analysis of reduced genomes. *Proc. Natl. Acad. Sci. USA*, **2003**, *100*, 9388-9393.

[24] Gomez-Valero, L.; Latorre, A.; Silva, F.J. The evolutionary fate of nonfunctional DNA in the bacterial endosymbiont Buchnera aphidicola. *Mol. Biol. Evol.*, **2004**, *21*, 2172-2181.

[25] Perez-Brocal, 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.

[26] Moran, N.A.; Mira, A. The process of genome shrinkage in the obligate symbiont Buchnera aphidicola. *Genome Biol.*, **2001**; *2*, RESEARCH0054.

[27] Silva, F.J.; Latorre, A.; Moya, A. Genome size reduction through multiple events of gene disintegration in Buchnera APS. *Trends Genet.*, **2001**, *17*, 615-618.

[28] Toft, C.; Fares, M.A. The evolution of the flagellar assembly pathway in endosymbiotic bacterial genomes. *Mol. Biol. Evol.*, **2008**, *25*, 2069-2076.

[29] Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H. Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS. *Nature*, **2000**, *407*, 81-86.

[30] Lai, C.Y.; Baumann, L.; Baumann, P. Amplification of trpEG: adaptation of Buchnera aphidicola to an endosymbiotic association with aphids. *Proc. Natl. Acad. Sci. USA*, **1994**, *91*, 3819-3823.

[31] Bracho, A.M.; Martinez-Torres, D.; Moya, A.; Latorre, A. Discovery and molecular characterization of a plasmid localized in Buchnera sp. bacterial endosymbiont of the aphid Rhopalosiphum padi. *J. Mol. Evol.*, **1995**, *41*, 67-73.

[32] Baumann, L.; Baumann, P. Moran, N.A. The endosymbiont (Buchnera) of the aphid Diuraphis noxia contains all the genes of the tryptophan biosynthetic pathway. *Curr Microbiol.* **1998**; 37, 58-59.

[33] Rouhbakhsh, D.; Clark, M.A.; Baumann, L.; Moran, N.A.; Baumann, P. Evolution of the tryptophan biosynthetic pathway in Buchnera (aphid endosymbionts): studies of plasmid-associated trpEG within the genus Uroleucon. *Mol. Phylogenet. Evol.*, **1997**, *8*, 167-176.

[34] Van Ham, R.C.; Gonzalez-Candelas, F.; Silva, F.J.; Sabater, B.; Moya, A.; Latorre, A. Postsymbiotic plasmid acquisition and evolution of the repA1-replicon in Buchnera aphidicola. *Proc. Natl. Acad. Sci. USA*, **2000**, *97*, 10855-10860.

[35] Moran, N.A.; Plague, G.R.; Sandstrom, J.P.; Wilcox, J.L. A genomic perspective on nutrient provisioning by bacterial symbionts of insects. *Proc. Natl. Acad. Sci. USA*, **2003**, 100 Suppl 2, 14543-14548.

[36] Wilcox, J.L.; Dunbar, H.E.; Wolfinger, R.D.; Moran, N.A. Consequences of reductive evolution for gene expression in an obligate endosymbiont. *Mol. Microbiol.*, **2003**, *48*, 1491-1500.

[37] Sigler, P.B.; Xu, Z.; Rye, H.S.; Burston, S.G.; Fenton, W.A.; Horwich, A.L. Structure and function in GroEL-mediated protein folding. *Annu. Rev. Biochem.*, **1998**, *67*, 581-608.

[38] Thirumalai, D.; Lorimer, G.H. Chaperonin-mediated protein folding. *Annu. Rev. Biophys. Biomol. Struct.*, **2001**, *30*, 245-269.

[39] Grallert, H.; Buchner, J. Review: a structural view of the GroE chaperone cycle. *J. Struct. Biol.*, **2001**; *135*, 95-103.

[40] Lilie, H. Buchner, J. Interaction of GroEL with a highly structured folding intermediate: iterative binding cycles do not involve unfolding. *Proc. Natl. Acad. Sci. USA*, **1995**, *92*, 8100-8104.

[41] Clark, A.C.; Frieden, C. GroEL-mediated folding of structurally homologous dihydrofolate reductases. *J. Mol. Biol.*, **1997**, *268*, 512-525.

[42] Zahn, R.; Lindner, P.; Axmann, S.E. Pluckthun, A. Effect of single point mutations in citrate synthase on binding to GroEL. *FEBS Lett.*, **1996**, *380*, 152-156.

[43] Gordon, C.L.; Sather, S.K.; Casjens, S. King, J. Selective *in vivo* rescue by GroEL/ES of thermolabile folding intermediates to phage P22 structural proteins. *J. Biol. Chem.*, **1994**, *269*, 27941-27951.

[44] Fares, M.A.; Ruiz-Gonzalez, M.X.; Moya, A.; Elena, S.F.; Barrio, E. Endosymbiotic bacteria: groEL buffers against deleterious mutations. *Nature*, **2002**, *417*, 398.

[45] Maisnier-Patin, S.; Roth, J.R.; Fredriksson, A.; Nystrom, T.; Berg, O.G.; Andersson, D.I. Genomic buffering mitigates the effects of deleterious mutations in bacteria. *Nat. Genet.*, **2005**, *37*, 1376-1379.

[46] Tokuriki, N.; Tawfik, D.S. Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature*, **2009**, *459*, 668-673.

[47] Herbeck, J.T.; Funk, D.J.; Degnan, P.H.; Wernegreen, J.J. A conservative test of genetic drift in the endosymbiotic bacterium Buchnera: slightly deleterious mutations in the chaperonin groEL. *Genetics*, **2003**, *165*, 1651-1660.

[48] Fares, M.A.; Barrio, E.; Sabater-Munoz, B. Moya, A. The evolution of the heat-shock protein GroEL from Buchnera, the primary endosymbiont of aphids, is governed by positive selection. *Mol. Biol. Evol.*, **2002**, *19*, 1162-1170.

[49] Fares, M.A.; Moya, A.; Barrio, E. Adaptive evolution in GroEL from distantly related endosymbiotic bacteria of insects. *J. Evol. Biol.*, **2005**, *18*, 651-660.

[50] Chamary, J.V.; Parmley, J.L.; Hurst, L.D. Hearing silence: nonneutral evolution at synonymous sites in mammals. *Nat. Rev. Genet.*, **2006**, *7*, 98-108.

[51] Mayrose, I.; Doron-Faigenboim, A.; Bacharach, E.; Pupko, T. Towards realistic codon models: among site variability and dependency of synonymous and non-synonymous rates. *Bioinformatics*, **2007**, *23*, i319-327.

[52] Kudla, G.; Murray, A.W.; Tollervey, D.; Plotkin, J.B. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science*, **2009**, *324*, 255-258.

[53] Smith, J.M.; Haigh, J. The hitch-hiking effect of a favourable gene. *Genet. Res.*, **1974**, *23*, 23-35.

[54] Kimura, M. Evolutionary rate at the molecular level. *Nature*, **1968**, *217*, 624-626.

[55] Hahn, M.W. Toward a selection theory of molecular evolution. *Evolution*, **2008**, *62*, 255-265.

[56] Pamilo, P.; Nei, M.; Li, W.H. Accumulation of mutations in sexual and asexual populations. *Genet. Res.*, **1987**, *49*, 135-146.

[57] Kondrashov, A.S. Contamination of the genome by very slightly deleterious mutations: why have we not died 100 times over? *J. Theor. Biol.*, **1995**, *175*, 583-594.

[58] Moran, N.A.; McLaughlin, H.J.; Sorek, R. The dynamics and time scale of ongoing genomic erosion in symbiotic bacteria. *Science*, **2009**, *323*, 379-382.

[59] Ochman, H.; Elwyn, S.; Moran, N.A. Calibrating bacterial evolution. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 12638-12643.

[60] Rispe, C.; Delmotte, F.; van Ham, R.C.; Moya, A. Mutational and selective pressures on codon and amino acid usage in Buchnera, endosymbiotic bacteria of aphids. *Genome Res.*, **2004**, *14*, 44-53.

[61] Itoh, T.; Martin, W.; Nei, M. Acceleration of genomic evolution caused by enhanced mutation rate in endocellular symbionts. *Proc. Natl. Acad. Sci. USA*, **2002**, *99*, 12944-12948.

[62] Drummond, D.A.; Bloom, J.D.; Adami, C.; Wilke, C.O.; Arnold, F.H. Why highly expressed proteins evolve slowly. *Proc. Natl. Acad. Sci. USA*, **2005**; *102*, 14338-14343.

[63] Toft, C.; Fares, M.A. *Selection for Translational Robustness in Buchnera aphidicola, Endosymbiotic Bacteria of Aphids. Mol. Biol. Evol.*, **2009**. .

[64] Rasmussen, B.F.; Stock, A.M.; Ringe, D.; Petsko, G.A. Crystalline ribonuclease A loses function below the dynamical transition at 220 K. *Nature*, **1992**, *357*, 423-424.

[65] Taverna, D.M.; Goldstein, R.A. Why are proteins marginally stable? *Proteins*, **2002**, *46*, 105-109.

[66] Toft, C.; Williams, T.A.; Fares, M.A. Genome-wide functional divergence after the symbiosis of proteobacteria with insects unraveled through a novel computational approach. *PLoS Comput. Biol.*, **2009**, *5*, e1000344.

[67] Levin, D. Flowering phenology in relation to adaptive radiation. *Sust. Bot.*, **2006**, *31*, 239-246.

[68] Rundle, H.; Nosil, P. Ecological speciation. *Ecol. Lett.*, **2005**, *8*, 336-352.

[69]   Schluter, D. *The ecology of adaptive radiation*. Oxford, UK: Oxford University Press. **2000**; .

[70]   Fitch, W.M.; Markowitz, E. An improved method for determining codon variability in a gene and its application to the rate of fixation of mutations in evolution. *Biochem. Genet.*, **1970**, *4*, 579-593.

[71]   Li, W.H.; Gojobori, T. Rapid evolution of goat and sheep globin genes following gene duplication. *Mol. Biol. Evol.*, **1983**, *1*, 94-108.

[72]   Clark, A.G. Invasion and maintenance of a gene duplication. *Proc. Natl. Acad. Sci. USA*, **1994**, *91*, 2950-2954.

[73]   Hughes, A.L. The evolution of functionally novel proteins after gene duplication. *Proc. Biol. Sci.*, **1994**; *256*, 119-124.

[74]   Nei, M.; Gu, X.; Sitnikova, T. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc. Natl. Acad. Sci. USA*, **1997**, *94*, 7799-7806.

[75]   Force, A.; Lynch, M.; Pickett, F.B.; Amores, A.; Yan, Y.L. Postlethwait, J. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, **1999**, *151*, 1531-1545.

[76]   Gu, X. Evolution of duplicate genes versus genetic robustness against null mutations. *Trends Genet.*, **2003**, *19*, 354-356.

[77]   Ruano-Rubio, V.; Fares, M.A. Artifactual phylogenies caused by correlated distribution of substitution rates among sites and lineages: the good, the bad, and the ugly. *Syst. Biol.*, **2007**, *56*, 68-82.

[78]   Ruano-Rubio, V.; Fares, M.A. Testing the neutral fixation of hetero-oligomerism in the archaeal chaperonin CCT. *Mol. Biol. Evol.*, **2007**; *24*, 1384-1396.

[79]   Sanjuan, R. Elena, S.F. Epistasis correlates to genomic complexity. *Proc. Natl. Acad. Sci. USA*, **2006**, *103*, 14402-14405.

[80]   Moran, N.A.; McCutcheon, J.P.; Nakabachi, A. Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.*, **2008**, *42*, 165-190.

[81]   Ochman, H.; Wilson, A.C. Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *J. Mol. Evol.*, **1987**, *26*, 74-86.

[82]   Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.*, **1997**, *13*, 555-556.

[83]   Toft, C. *Complex Evolutionary Dynamics in Simple Genomes: The Paradoxical Survival of Intracellular Symbiotic Bacteria. Ph.D. Thesis.* **2008**.