



Editorial

## Biological evolution: Lessons to be learned from microbial population biology and genetics

Werner Arber

*Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland*

Received 19 January 2004; accepted 27 January 2004

Available online 19 March 2004

This issue is dedicated to the memory of Michel Blot, who died in a tragic accident in the French Alps in September 2002, bringing his fruitful scientific activities to a sudden end. After having completed his basic and doctoral education in population biology of eukaryotic organisms, Michel Blot joined my laboratory in 1989 with the intention of becoming familiar with microbial genetics and of combining the two fields of population biology and microbial genetics in studies on mechanisms of biological evolution. His contributions to this interdisciplinary approach rapidly attracted the attention of the scientific community, bringing him a well merited recognition of his scientific activities which he deployed for the last seven years as Professor and Group Leader at the University J. Fourier in Grenoble, France. The articles comprising this special issue have been written by colleagues and friends of Michel Blot and shall document the genetic plasticity of microbial genomes and its impact on microbial evolution. They complement related reports on bacteriophages, on integrons and on repetitive DNA sequences in microbial genomes published in recent special issues of this journal (Refs. [8–10]). The general relevance of studies on microbial genetics and microbial population biology for a deeper understanding of molecular evolution shall briefly be outlined in this editorial.

The Neo-Darwinian theory of biological evolution was elaborated before it became known that genetic information is contained in DNA molecules. While this theory postulates that it is genetic variation which drives the evolutionary process, it cannot explain the molecular nature of genetic variation. With the advent of microbial genetics identifying DNA as the carrier of genetic information, and shortly thereafter, with the description of the double-helical structure of DNA molecules, the door was open to investigating molecular mechanisms of genetic variation.

The most accessible and successful approaches for experimental investigations used, and still use, microorganisms, mainly bacteria and viruses, as study objects. In population studies it is possible to trace individual mutagenic events by comparing DNA sequences before and after the occurrence of a genetic variation. On the other hand, the comparison of DNA sequences of functional domains, genes, groups of genes and entire genomes from more or less related organisms can suggest which kinds of events must have brought about the actual gene and genome structures in their past evolutionary developments since their evolutionary separation.

In the following sections we will use the terms “genetic variation” and “mutation” as synonyms, and for their molecular genetic definition, i.e., regarding any alteration in the genomic DNA sequence as a mutation. This contrasts with the classical definition of a mutation as an inheritable change in the phenotypic properties of an organism. It is well known that not all DNA sequence alterations result in alterations of the phenotype.

Fig. 1 summarizes our present knowledge on genetic variation and puts it into the context of natural selection (favoring some of the genetic variants and rejecting many others) and of geographic and reproductive isolation. We note that there is a multitude of sources of genetic diversity. Behind each individual spontaneous mutational event there is a specific molecular mechanism, but these mechanisms can be quite different for different events. Fig. 1 classifies all possible sources of genetic variation into three major natural strategies. One of these strategies brings about small local changes in the DNA sequences, such as nucleotide substitution, the deletion or the additional insertion of one or a few nucleotides, or a scrambling of a few nucleotides. A second strategy is called DNA rearrangement and is largely based on recombinational events within the genome of the concerned organism. A third strategy, here called DNA acquisition, involves the horizontal transfer of a DNA segment from a donor to a recipient organism, often mediated by a

*E-mail address:* [werner.arber@unibas.ch](mailto:werner.arber@unibas.ch) (W. Arber).

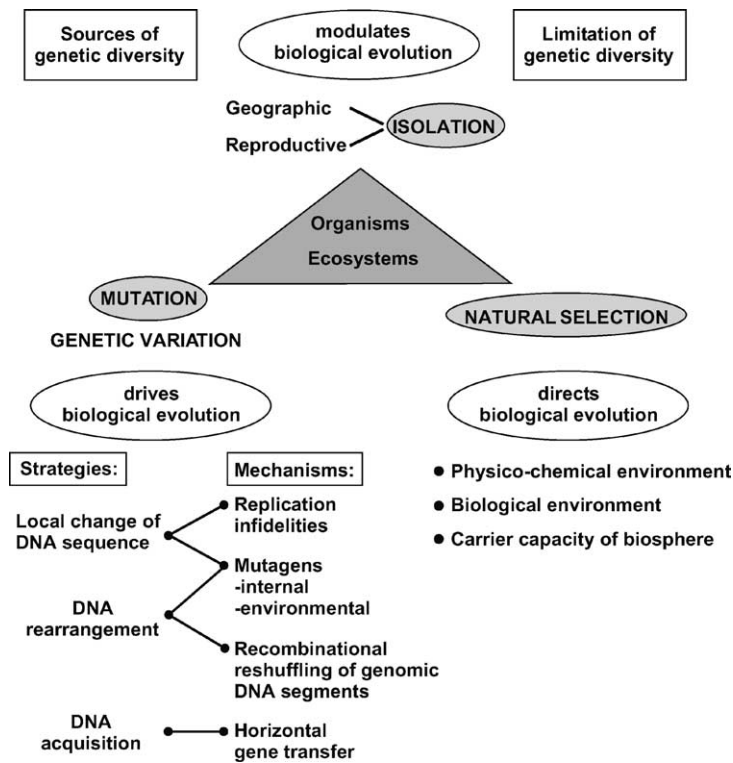


Fig. 1. Synoptical view of the main elements of biological evolution with particular attention to the natural strategies of genetic variation [5]. See text and Refs. [1–3,6] for explanation and background.

natural gene vector such as a virus or a conjugative plasmid, but in other instances also by transformation involving free DNA segments.

The three natural strategies of genetic variation differ in the quality of their contributions to biological evolution. Local sequence changes mainly contribute to a stepwise improvement and adaptation of existing gene activities. This strategy might occasionally also bring about a novel gene activity, but one has to expect that such a development will only become efficient once the gene product becomes a substrate for natural selection. The strategy of DNA rearrangement can bring about the fusion of two different functional domains which can by chance give rise to a new functional product. Similarly, this strategy can fuse an open reading frame with a hitherto unrelated regulatory element for gene expression. Alternatively, this strategy can also bring about gene duplication, offering the possibility of different evolutionary developments of the resulting copies. Finally, the strategy of DNA acquisition offers to the recipient organism a sharing in successful evolutionary developments made by other organisms. This strategy of acquiring functional genetic information is quite efficient per event. It has a good chance of success, particularly if it occurs in a small step, i.e., by the acquisition of a relatively short DNA segment so that the functional harmony of the recipient cell has only a small risk of being disturbed.

One can assume that the fitness for further evolution of a bacterium will depend on its capacity to use all three

described strategies of genetic variation, ideally by having available a few different specific mechanisms for each strategy. For example, quite different recombination mechanisms can contribute in different ways to DNA rearrangements. As we know, general recombination reassorts DNA at relatively extended segments of homologous sequences. This mechanism is well known to contribute efficiently to the genetic diversity of diploid organisms. But it is also active in prokaryotes, particularly by repairing damaged chromosomes and by recombining segments of homology, such as IS elements, located at different sites of the chromosome or on a plasmid. Bacteria often also possess systems for site-specific recombination. These use either specific or consensus sites for recombination with relatively high efficiency. Much less often they may also use secondary sites of crossing over, deviating considerably from consensus. Such rare events involving a secondary site of recombination are a good source of evolutionarily relevant DNA rearrangements. Last but not least, bacterial mobile genetic elements are well known to contribute to spontaneous mutagenesis by transpositional DNA rearrangements including, in addition to simple transposition, also deletion formation and DNA inversion as well.

In order to avoid misunderstandings, it should be mentioned that some specific mechanisms of mutagenesis contribute to more than one strategy of genetic variation. For example, the transposition producing a rearrangement of segments of DNA also causes, in the insertion target region, the duplication of a few nucleotides, hence a local sequence

change. Another example is the involvement of recombination systems in most processes of horizontal gene transfer.

Many of the specific mechanisms of genetic variation are mediated by specific enzymes. Transposases and recombinases are prominent examples. As far as we know, transposition is not one of the essential functions required in bacteria for their physiology and for exponential propagation. Transposable genetic elements can thus be seen as genetic systems serving as generators of genetic variations with obvious evolutionary relevance; we therefore classify them as evolution genes. Products of evolution genes are also involved in other mechanisms of DNA rearrangement, as well as in DNA acquisition and in at least some processes of local sequence changes. However, many specific mechanisms of genetic variation not only depend on products of evolution genes, but frequently on non-genetic factors. Such factors may be intrinsic properties of matter affecting, for example, the chemical stability of nucleotides or influencing the structural flexibility of nucleotides. Tautomerism of nucleotides is well known to influence base pairing in double-helical DNA molecules. Indeed, mispairings resulting from these effects would seriously affect a certain required genetic stability of genomes if a large fraction of the so-called replication infidelities were not repaired by special enzyme systems. In this case, the repair enzymes serve as modulators of the frequency of genetic variation, and we consider the relevant genes as evolution genes as well.

In line with the stringent requirement to maintain spontaneous mutation rates low in order to ensure a certain genetic stability within each microbial strain is the observation that genetic variation generators usually act quite inefficiently. Transposable genetic elements, for example, use a number of different strategies to keep transposition frequencies quite low.

One can assume that evolution genes serving either as generators of genetic variations or as modulators of the frequencies of genetic variation have their own past evolutionary history, in which they must have been fine-tuned for their specific functional activities. Selection for these activities must have been indirect at the level of populations. We call this process second-order selection.

The existence of evolution genes implies a duality in the genome of an organism. The products of many genes serve each individual for its normal life functions. These are housekeeping genes and genes for products required under particular living conditions, such as in the presence of antibiotics. In contrast, the products of evolution genes are not stringently required for each individual moving from one generation to the next. Rather, they serve to occasionally produce a genetic variation in one individual being a member of a larger population. For example, in propagating cultures of *Escherichia coli* bacteria, overall spontaneous genetic variation affects between 0.1 and 1% of the cells per generation. A very strict classification of microbial gene functions into those needed for each individual cell and those

serving to ensure genetic plasticity and thus to drive biological evolution would not correspond to the encountered reality. Indeed, the products of some genes are relevant for both the individual lives and the evolutionary development. Examples are DNA topoisomerases and DNA ligases. One can assume that genes serving different purposes have been fine-tuned in their evolutionary history to reliably accomplish their different tasks.

It is a general observation that only a minority of spontaneously occurring genetic variations are favorable and provide a selective advantage to the mutant organism. This indicates that genetic variation is in general not directed. This could be related to the multitude of molecular mechanisms generating genetic variations. Possibly, some very specific mechanisms may produce an increased proportion of favorable mutations under particular physiological and/or environmental conditions. This possibility will have to be explored case by case. It might have its relevance for cases of adaptive mutations.

While, in a crude view, overall genetic variation occurs more or less randomly along the genome, this may not be strictly so for some given specific mechanisms. Indeed, local hot targets and target regions are known to serve for some processes of genetic variation. Similarly, frequencies of spontaneous genetic variation can vary, e.g., in subclones having assumed a special genomic configuration of increased genetic instability. Dimeric forms of IS elements giving rise to bursts of transposition are a good example.

Another variable factor to take into account in biological evolution is natural selection. This depends on the living conditions encountered by the organism. It is clear that living conditions can vary both with space and with time. They not only depend on the physicochemical environment, but also on the biological environment, i.e., different kinds of organisms living in a particular ecological niche influence each other. A convenient way to study the impact of environmental factors on natural selection involves competition experiments with microorganisms of different genetic traits. Mixed populations of organisms and/or genetic variants are thereby grown under defined environmental conditions. This allows us to follow the relative frequencies of the input forms as a function of time.

In conclusion, experimental investigations in microbial genetics and in microbial population biology reveal a multitude of specific molecular mechanisms that produce genetic variants. These investigations can also deepen our knowledge of the effect of natural selection on any given strain of microorganism and its genetic variants. The lessons thereby learned are at the basis of the theory of molecular evolution or, in other words, Darwinism at the molecular level [4,5].

Unicellular microorganisms are thought to have populated our planet for more than three billion years, long before higher multicellular organisms showed up. It is most likely that evolution genes whose products act as generators of genetic variations and/or as modulators of the frequencies of genetic variation developed in the microbial world at these

early periods and that their activities later enabled the living world to experience division of labor in microbial populations, and finally also the development of multicellular organisms [7]. It is most likely that symbiosis with microorganisms was thereby involved together with the described strategies of genetic variation. Endosymbiosis must also offer good conditions for occasional horizontal gene transfer.

We are aware that some of the functions that had been developed by single-cellular organisms for the purpose of biological evolution are used by higher, multicellular organisms to the benefit of the individual. Examples are the repair of DNA damage in somatic cells and the developmental establishment of a functional immune system. However, one can assume that basic principles of mechanisms of genetic variation as outlined here for the microbial world will also serve in higher organisms for their further evolutionary development. Increasing evidence in support of this can be expected from genomics and, in particular, from DNA sequence comparison with the help of bioinformatic tools. Direct experimental investigations are still rather difficult because of the system-inherent inefficiency of spontaneous genetic variation and its non-reproducibility from case to case, as well as the multitude of specific mechanisms contributing to spontaneous mutagenesis. We are confident, however, that

appropriate strategies for investigations of molecular evolution will be found to strengthen and deepen our knowledge of the evolutionary development of the living world.

## References

- [1] W. Arber, Elements in microbial evolution, *J. Mol. Evol.* 33 (1991) 4–12.
- [2] W. Arber, The generation of variation in bacterial genomes, *J. Mol. Evol.* 40 (1995) 7–12.
- [3] W. Arber, Genetic variation: Molecular mechanisms and impact on microbial evolution, *FEMS Microbiol. Rev.* 24 (2000) 1–7.
- [4] W. Arber, Elements for a theory of molecular evolution, *Gene* 317 (2003) 3–11.
- [5] W. Arber, Genetic variation and molecular evolution, in: R.A. Meyers (Ed.), *Encyclopedia of Molecular Cell Biology and Molecular Medicine*, vol. 5, Wiley-VCH, Weinheim, 2004, pp. 331–352.
- [6] M. Blot, Transposable elements and adaptation of host bacteria, *Genetica* 93 (1994) 5–12.
- [7] L.H. Caporale (Ed.), *Molecular Strategies in Biological Evolution*, *Ann. NY Acad. Sci.*, vol. 870, 1999.
- [8] P. Forterre, H. Krish (Eds.), Viruses: Origin, evolution and biodiversity, *Res. Microbiol.* 154 (2003) 223–313.
- [9] D.A. Rowe-Magnus, A.M. Guérout, D. Mazel, Superintegrons, *Res. Microbiol.* 150 (1999) 641–651.
- [10] M. Hofnung, J.A. Shapiro (Eds.), Repetitive DNA sequences in microbial genomes, *Res. Microbiol.* 150 (1999) 577–799.