

Genetic variation: molecular mechanisms and impact on microbial evolution

Werner Arber *

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

Received 18 August 1999; received in revised form 29 October 1999; accepted 1 November 1999

Abstract

On the basis of established knowledge of microbial genetics one can distinguish three major natural strategies in the spontaneous generation of genetic variations in bacteria. These strategies are: (1) small local changes in the nucleotide sequence of the genome, (2) intragenomic reshuffling of segments of genomic sequences and (3) the acquisition of DNA sequences from another organism. The three general strategies differ in the quality of their contribution to microbial evolution. Besides a number of non-genetic factors, various specific gene products are involved in the generation of genetic variation and in the modulation of the frequency of genetic variation. The underlying genes are called evolution genes. They act for the benefit of the biological evolution of populations as opposed to the action of housekeeping genes and accessory genes which are for the benefit of individuals. Examples of evolution genes acting as variation generators are found in the transposition of mobile genetic elements and in so-called site-specific recombination systems. DNA repair systems and restriction-modification systems are examples of modulators of the frequency of genetic variation. The involvement of bacterial viruses and of plasmids in DNA reshuffling and in horizontal gene transfer is a hint for their evolutionary functions. Evolution genes are thought to undergo biological evolution themselves, but natural selection for their functions is indirect, at the level of populations, and is called second-order selection. In spite of an involvement of gene products in the generation of genetic variations, evolution genes do not programmatically direct evolution towards a specific goal. Rather, a steady interplay between natural selection and mixed populations of genetic variants gives microbial evolution its direction. © 2000 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.

Keywords: DNA acquisition; DNA rearrangement; Evolution gene; Genetic variation; Natural selection; Theory of molecular evolution

Contents

1. Historical introduction	2
2. Gene involvement in the generation of genetic variations	2
3. Integral view of natural strategies to generate genetic variations	2
3.1. Local sequence change	3
3.2. DNA rearrangement	3
3.3. DNA acquisition	4
4. Evidence for the occurrence of evolution genes	4
4.1. Site-specific recombination systems	4
4.2. Transposable elements	5
4.3. DNA repair systems	5
4.4. Restriction-modification systems	5
4.5. Bacterial viruses and plasmids	5
5. Involvement of non-genetic factors in the generation of genetic variants	5
6. Evolution of evolutionarily active elements	6
7. Synthesis between evolutionary and molecular biology	6
References	6

* Tel.: +41 (61) 267 21 30; Fax: +41 (61) 267 21 18; E-mail: werner.arber@unibas.ch

1. Historical introduction

The occurrence of spontaneously arising genetic variants of microorganisms and the experimental use of chemical and physical mutagens to induce genetic variations are at the basis of microbial genetics which has its origin in the 1940s. As compared to the diploid nature of most higher organisms, prokaryotic microorganisms, bacteria and their viruses, are usually haploid and they thus rapidly reveal phenotypic consequences of genetic variation. This and the fast propagation of these microorganisms facilitate genetic experimentation. Natural means to obtain temporary or more permanent partial diploidy in bacteria by transformation, conjugation or virus-mediated transduction have extensively been used to study both functional complementation and genetic recombination between genetic variants. In this experimental approach individual mutations can be assigned to specific genes and genetic maps can be established. The same applies to bacteriophage genetics upon mixed infection of host bacteria with different phage mutants. These classical microbial genetic studies were often carried out with mutants which had been induced with specific mutagens giving rise to nucleotide substitution and resulting in missense and nonsense mutations. This had been facilitated by available knowledge on specific mutagen action. In this context one has often assumed that spontaneous genetic variants were normally substitution mutants or more generally point mutants.

In the late 1960s microbial geneticists started to pay attention to a class of spontaneous mutations which often exerted polar effects and which turned out to be caused by the insertion of mobile genetic elements [1]. Subsequent studies on transposition of mobile elements made it clear that by far not all microbial recombination activities concerned homologous recombination occurring between identical DNA sequences. As a matter of fact, site-specific recombination had already been studied at that time, particularly the integration of proviral genomes into the bacterial host chromosome [2]. It was then revealed that site-specific recombination can also occur, although at lower frequencies, at DNA sequences deviating from the normal recombination site. Both DNA rearrangement activities mentioned here, transposition and site-specific recombination, can also give rise to deletion formation as well as DNA inversion. In addition, both of these activities can bring about an association of segments of chromosomal DNA with viral genomes and plasmids which thus act as natural gene vectors. Studies on specialized transduction, F' conjugative plasmids and antibiotic resistance plasmids were very instructive in this regard.

2. Gene involvement in the generation of genetic variations

Genetic variation is a prerequisite for biological evolu-

tion. Indeed, mixed populations of parental forms and different genetic variants are the substrate for natural selection which is exerted by the living conditions encountered by the organisms. It is still a widely spread opinion that genetic variants spontaneously result from errors and mistakes made during replication or accidents happening to the DNA resulting in changes in the base sequence. With a similar attitude mobile genetic elements are often considered selfish or parasitic, and rare, non-reproducible recombination events are called illegitimate. In the following considerations, I will defend a different view which I have already propagated for a number of years [3–6].

The thesis to be defended here postulates the existence of evolution genes, the products of which act for the benefit of the biological evolution of the population of organisms. According to this hypothesis there are principally two types of evolution genes. The products of one of these types are generators of genetic variations, while the products of the other type of genes are modulators of the frequency of genetic variation. Hence, evolution genes generally influence the kind and the frequency of genetic variation occurring in a population of organisms and sometimes they may ensure a certain degree of genetic isolation.

Bacteria must have evolved on our planet for about 3.8 billion years. The currently found diversity of bacterial life forms is impressive, particularly with regard to their differences in the capacity to live under extreme environmental conditions, such as chemical composition, pH, temperature and pressure. In addition, many bacterial species have become symbionts or pathogens and have adapted to living in close community with other organisms. If the latter are higher organisms, the evolutionary adaptation of today's pathogenic bacteria must have occurred relatively recently. We may therefore expect to find good examples of evolution genes at work in today's bacteria. As a matter of fact, it is likely that evolution genes have themselves undergone a long evolution and are now fine-tuned for their present functions.

The theory of molecular evolution postulates that the products of evolution genes work hand in hand with other factors such as structural and stability features of biological macromolecules, the action of chemical and physical mutagens, as well as the chance of random encounters of interactive components.

A synoptic view of genetic and non-genetic elements which the theory of molecular evolution postulates to contribute to the spontaneous generation of genetic variants is given in Fig. 1.

3. Integral view of natural strategies to generate genetic variations

In the following considerations we follow the molecular genetic custom of defining as a mutation any alteration in

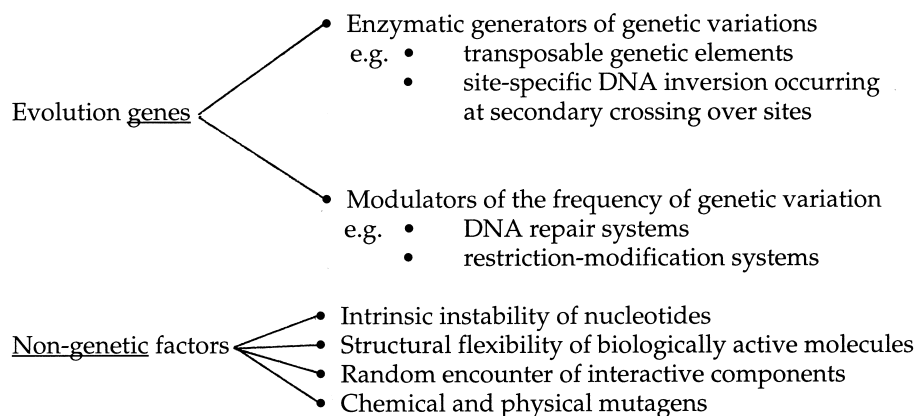


Fig. 1. Evolution genes and non-genetic factors contribute to the generation of genetic variations.

the nucleotide sequence of the genome without paying attention to whether such a change brings about a phenotypic change or not. This clearly differs from the classical genetic custom of identifying a mutation by a change in the phenotypic properties of an organism. However, as far as we can judge, molecular mechanisms producing DNA sequence alterations exert their interactions on DNA molecules without paying attention to whether such interaction causes a phenotypic change. Indeed, it has frequently been observed that quite many nucleotide changes are silent or neutral, i.e. without any immediate effect on the phenotype of the organism. It is also known that in general relatively few DNA sequence alterations bring a selective advantage to the organism in its normal habitat, while many more sequence alterations cause a selective disadvantage or, in the extreme case, lethality. This rule indicates that in general spontaneous genetic variations are not in response to a specific need of an organism in its environment.

According to the experience of microbial genetics three qualitatively different strategies contribute to the overall production of genetic variants (Fig. 2). The first of these strategies brings about local DNA sequence changes such as nucleotide substitutions, small deletions and small insertions of one or a few nucleotides or a limited scrambling of nucleotides. The second strategy is a segment-wise rearrangement of genomic DNA sequences by recombinational reshuffling. The third strategy is based on horizontal gene transfer and results in the acquisition of foreign DNA sequences.

These three major strategies to generate genetic variations differ from each other in the quality of their effects. With regard to their evolutionary impact, individual mechanisms belonging to the same strategy may substitute for one another, but not for a member of another type of strategy. Hence, evolutionarily well adaptable bacteria may be expected to be genetically equipped with at least one or – what would be even better – several systems for each of the three different evolutionary strategies.

3.1. Local sequence change

Local sequence changes might theoretically represent individual steps in the development of new biological functions. However, there is no general evidence for a specific directionality of such steps. The mutational generation of a completely novel biological function may thus be extremely rare as long as natural selection cannot exert its pressure on an at least primitive function displayed by the product of the DNA sequence implied. The situation may change in the stepwise improvement of already available biological functions on which natural selection steadily exerts its pressure. This can apply to functional domains, entire genes as well as to functional systems composed of different genes.

Evolutionary clocks used by the scientists to estimate evolutionary distances are based on local sequence changes often within a specific gene, for example a ribosomal RNA gene. Within functional genes lethal and heavily contraselective mutations will of course not be maintained. Also, local sequence changes are known to affect different DNA regions with different efficiencies. For these reasons the precision of evolutionary clocks is limited, but the method has its merits for the relative comparison of the evolutionary relations between different species. For this purpose strategies of genetic variation other than local change can be ignored except in the case of gene conversion upon DNA acquisition by horizontal gene transfer.

3.2. DNA rearrangement

Gene conversion can also result from the strategy of intragenomic reshuffling of DNA sequences, if related sequences carried in the genome undergo recombination. Another, often innovative and thus important effect of the DNA rearrangement strategy is the novel combination of available capacities by the fusion of different functional domains. Genomics reveals more and more that different functional genes often share related sequence motifs and

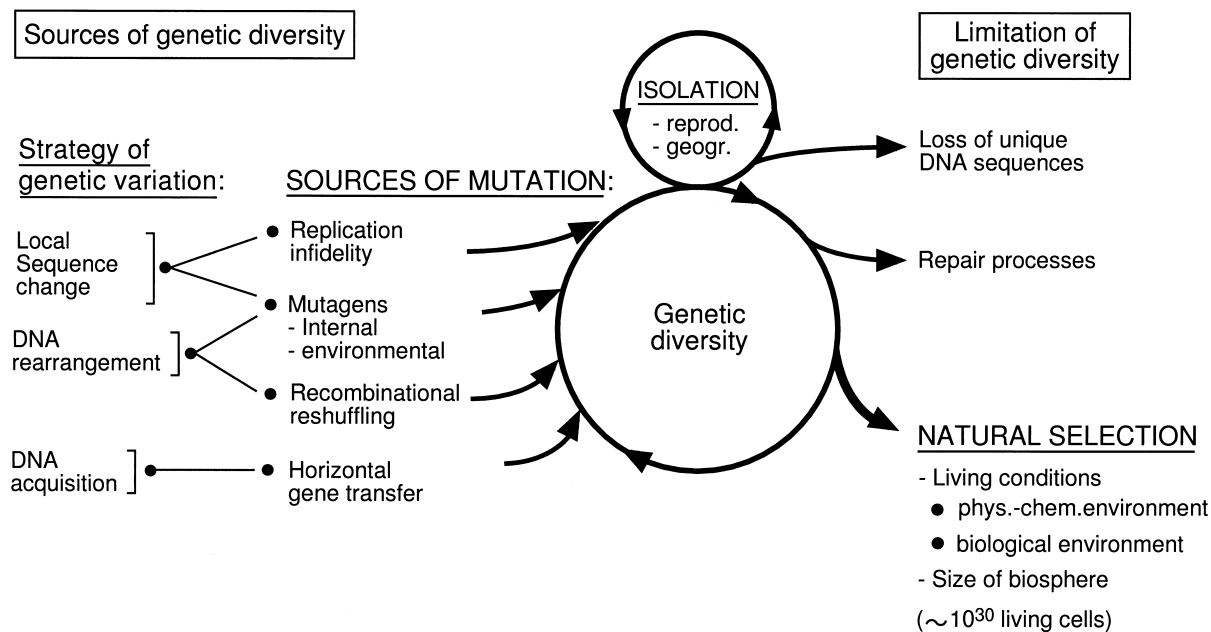


Fig. 2. Synoptic view of the elements of molecular evolution of prokaryotic microorganisms.

functional domains. Occasional DNA reshuffling is a good source for such fusion genes. The same holds true for the reassortment of expression control signals with different reading frames serving in protein production. DNA rearrangements can also bring about the duplication of DNA segments which may serve as substrates for further evolution. Similarly, deletion of DNA segments can result from DNA rearrangement processes and can help in conjunction with natural selection to discharge the genome from inessential sequences.

3.3. DNA acquisition

As was widely explored with antibiotic resistance genes, the strategy of DNA acquisition allows microorganisms to share the evolutionary success of others. This strategy is very efficient and can result in essential new capacities in a single step of acquisition. This is of particular relevance for accessory genes. Such genes may also open the possibility for a microorganism to populate new habitats. The evolutionary contribution of this strategy is thus considerable. As already mentioned, the acquisition strategy can also affect functions of essential housekeeping genes by conversion upon horizontal transfer of DNA segments.

4. Evidence for the occurrence of evolution genes

In the following section we discuss evidence for genes acting as variation generators or as modulators of the frequency of genetic variation and we will see that in some genetic systems both of these functions are exerted.

4.1. Site-specific recombination systems

Site-specific recombination systems are widespread in bacteria and several of them have been thoroughly studied in their molecular reactivities. By the intervention of specific gene products some site-specific recombination systems preferentially cause the inversion of DNA segments flanked by consensus sites of crossing over. Examples of this situation are found in flip-flop systems, which relatively quickly bring to an equilibrium the two or sometimes more possible DNA arrangements in the microbial population [7]. It was observed that the enzymes involved in DNA inversion can occasionally also use so-called secondary sites of crossing over, which may deviate considerably from the consensus [8,9]. Therefore DNA may contain a considerable number of such secondary recombination sites and each of them has its own characteristic frequency to serve in recombination. The result of such recombination is thus only statistically predictable, but not for an individual reaction. In addition, the recombination product is relatively stable, since the resulting hybrid sites also differ from the consensus and sometimes display additional local sequence variations. While DNA inversion involving a secondary site of recombination could be seen as an error of the enzyme system, it could, alternatively, also be seen as of evolutionary relevance. This depends on the attitude of the scientist interpreting the experimental observations. Carrying this kind of reasoning further, I see in site-specific DNA inversion primarily an evolutionary function providing at low frequencies novel DNA sequence fusions, some of which may be found useful upon natural selection. With this evolutionary function in mind, the more frequently, reproducible DNA inversion seen as flip-flop might be just a by-product

of the activity of the involved evolution gene. A DNA segment undergoing frequent reproducible inversion, i.e. flip-flop, is tolerable as long as none of its possible frequent DNA arrangements displays a strong selective disadvantage.

4.2. *Transposable elements*

An impressive number of prokaryotic mobile genetic elements have already been described, in particular the basic elements known as insertion sequences (IS). Their various DNA rearrangement activities largely depend on the target selection criteria characteristic for each element [10]. For some IS elements targets for insertion are more or less strictly respected consensus sequences, while other IS elements seem to follow more complex rules and may, for example, prefer particular DNA regions within which they alternatively insert into many different DNA sequences [11].

Transpositional DNA rearrangements do not immediately result in gene fusion, but a later excision of an IS element can bring about such an effect. However, IS elements also have a number of other means to influence genome evolution. For example, by providing sequence homologies at various sites on DNA molecules they can bring about DNA rearrangements which are mediated by homologous recombination. This, and transpositional DNA rearrangements, can be involved both in the association and in the subsequent dissociation of chromosomal DNA segments to and from natural gene vectors which is relevant for the gene acquisition strategy.

In general, IS elements themselves regulate their transposition activities by different means in order to provide DNA rearrangements only at low frequencies which are tolerable for the maintenance of a given microbial strain. However, it was observed with some IS elements that particular structural configurations in the location of IS elements in DNA molecules (dimeric IS structures) can temporarily result in very high frequencies of different kinds of DNA rearrangements [12]. This can manifest itself in a burst of transpositions in a subpopulation of the bacterial strain carrying the IS element in question [13].

4.3. *DNA repair systems*

While a majority of intragenomic DNA rearrangements and systems for horizontal DNA transfer must be mediated by genetic functions, the situation may be different in the strategy of genetic variation by local sequence change. A major cause for such changes is the intrinsic structural instability of nucleotides which may result either in direct mutagenicity or in infidelity upon DNA replication. Other causes for local sequence changes depend on the interaction of environmental mutagens with DNA. While the strategy of producing genetic variations by local sequence change may profit from these natural sources of mutation,

care must be taken by the organisms to limit the frequency of such changes to tolerable levels. This is brought about by DNA repair systems. Their main function is to ensure a certain degree of genetic stability and we thus classify them as evolution genes modulating the frequency of genetic variation.

4.4. *Restriction-modification systems*

Similarly, restriction enzymes considerably reduce the uptake of foreign DNA by bacteria. This seriously reduces the acquisition of foreign gene functions again to tolerable levels. Interestingly, restriction enzymes cut foreign DNA molecules upon their penetration into recipient bacteria into fragments. These DNA fragments are a substrate for exonucleolytic degradation. However, they are also recombinogenic and may be used for a permanent fixation into the recipient genome if an opportunity to do so is provided. In this sense, in bacterial restriction-modification systems we see evolution genes at work which, on the one hand, limit DNA acquisition to low levels and, on the other hand, stimulate DNA acquisition in small steps.

4.5. *Bacterial viruses and plasmids*

The question might be raised if one should also consider plasmids and viruses as elements with primarily evolutionary functions. Their role as natural gene vectors in microorganisms is uncontested and thus also their involvement in the DNA acquisition strategy. Together with DNA transformation they open the entire rich pool of genetic diversity to contribute by horizontal gene transfer to any future evolutionary development on any branch of the evolutionary tree.

5. *Involvement of non-genetic factors in the generation of genetic variants*

After having discussed examples of evolution genes, it is indicated to clearly state that nature also takes advantage of non-genetic factors for the production of genetic variants (Fig. 1). Both intrinsic structural features of biological macromolecules and external influences, such as those exerted by mutagens interacting with DNA, play their role. This is not limited to the strategy of local sequence change but can also affect the other strategies of variation generation. We postulate that in the infrequently occurring interaction of site-specific recombination enzymes with secondary recombination sites, rare, short-living structural variations of either the DNA substrate or the enzyme protein may be relevant. Similarly, in the transposition into variants of consensus target sequences structural instabilities might be involved. We realize that this hypothesis may be difficult to investigate experimentally by biophys-

ical or biochemical methods because of the inefficiency of this kind of reactions and in view of the often very short lifetime of structural variants of biological macromolecules.

Another non-genetic factor influencing the production of genetic variations is the chance of random encounter. This plays its role in the interaction of enzymes with their substrate such as a site of recombination on DNA molecules. Random encounter is also clearly involved in the DNA acquisition strategy. In transformation, DNA molecules must encounter recipient bacteria; in transduction, the transducing bacteriophage must encounter a host bacterium; and in conjugation, donor and recipient bacteria must encounter each other.

Repeated attempts to investigate if bacteria possess sensors to specifically identify evolutionary needs and to answer either by increased mutagenicity or more specifically by producing the most appropriate genetic variants, have not given conclusive answers. In general, the generation of genetic variations occurs independently of functional requirements related to particular living conditions. However, from case to case, it may be expected that exceptions to this rule may be found and, possibly, may also be explained causally according to the dictum 'no rule without exception'.

6. Evolution of evolutionarily active elements

Biological constituents of the process of biological evolution are likely to undergo biological evolution themselves. However, since the products of evolution genes do not provide a direct substrate for natural selection in each individual cell of microbial populations, selection for their evolutionary improvement must occur at the level of populations based on their mutagenic action on directly selectable genes. This has been called second order selection [14]. This is thought to have resulted in a fine tuning of the evolutionary functions in the course of time. Such fine tuning may concern various aspects such as the reactivity of recombination enzymes with variously structured target DNA sequences, or the intrinsic regulation of transposition activity of mobile genetic elements. It is also likely that structural elements on genomes have an evolutionary history, in particular nucleotide sequences affecting regional differences in the response to certain mechanisms of generation of genetic variations. Strategies and methodology of molecular genetics provide us with means to experimentally investigate this kind of questions and to measure the contribution of specific mechanisms of genetic variation relative to overall mutagenesis.

7. Synthesis between evolutionary and molecular biology

The interpretation given here for the generation of ge-

netic variations can be seen as an attempt to induce a synthesis between evolutionary biology on the one hand, and molecular biology and biochemistry on the other hand. It is true that evolutionary biology has paid relatively little attention either to microbial evolution or to interaction mechanisms at the level of biological macromolecules. It is also true that molecular biology has largely addressed its interests to both efficient and reproducible reactions such as those exerted by many products of housekeeping genes. The proposed variation generators clearly contrast with these conditions. They must function inefficiently as well as non-reproducibly, or at most, only statistically predictably. It will be a challenge for future studies to investigate the more general nature of such hitherto unexpected gene activities. This principle of not precisely programmed gene action might be of relevance also to fields of genetic influence other than biological evolution.

Another challenge for further studies will be to investigate if higher organisms also possess evolution genes with activities similar to those in bacteria. Genomics will give an excellent basis for such investigations. Since higher organisms must have evolved from microorganisms, we indeed expect the former to also take care of their own evolution, again with the rule seen with bacteria that evolution genes have no strict finality, i.e. no particular goal to be reached. The objective of evolution genes is rather thought to ensure a rich biodiversity which represents a good guarantee for a long-lasting maintenance and development of life on earth.

References

- [1] Shapiro, J.A. (1983) *Mobile Genetic Elements*. Academic Press, New York.
- [2] Campbell, A.M. (1962) Episomes. *Adv. Genet.* 11, 101–145.
- [3] Arber, W. (1991) Elements in microbial evolution. *J. Mol. Evol.* 33, 4–12.
- [4] Arber, W. (1993) Evolution of prokaryotic genomes. *Gene* 135, 49–56.
- [5] Arber, W. (1995) The generation of variation in bacterial genomes. *J. Mol. Evol.* 40, 7–12.
- [6] Arber, W. (1999) Involvement of gene products in bacterial evolution. In: *Molecular Strategies in Biological Evolution* (Caporale, L.H., Ed.), Vol. 870, pp. 36–44. *Annals of the New York Academy of Science*, New York.
- [7] Glasgow, A.C., Hughes, K.T. and Simon, M.I. (1989) Bacterial DNA inversion systems. In: *Mobile DNA* (Berg, D.E. and Howe, M.M., Eds.), pp. 637–659. *American Society of Microbiology*, Washington, DC.
- [8] Iida, S. and Hiestand-Nauer, R. (1986) Localized conversion at the crossover sequences in the site-specific DNA inversion system of bacteriophage P1. *Cell* 45, 71–79.
- [9] Iida, S. and Hiestand-Nauer, R. (1987) Role of the central dinucleotide at the crossover sites for the selection of quasi sites in DNA inversion mediated by the site-specific *Cin* recombinase of phage P1. *Mol. Gen. Genet.* 208, 464–468.
- [10] Craig, N.L. (1997) Target site selection in transposition. *Annu. Rev. Biochem.* 66, 437–474.

- [11] Sengstag, C. and Arber, W. (1983) IS2 insertion is a major cause of spontaneous mutagenesis of the bacteriophage P1: non-random distribution of target sites. *EMBO J.* 2, 67–71.
- [12] Olasz, F., Stalder, R. and Arber, W. (1993) Formation of the tandem repeat (IS30)₂ and its role in IS30-mediated transpositional DNA rearrangements. *Mol. Gen. Genet.* 239, 177–187.
- [13] Naas, T., Blot, M., Fitch, W.M. and Arber, W. (1994) Insertion sequence-related genetic variation in resting *Escherichia coli* K-12. *Genetics* 136, 721–730.
- [14] Weber, M. (1996) Evolutionary plasticity in prokaryotes: A panglossian view. *Biol. Phil.* 11, 67–88.