

CULTURAL ASPECTS OF THE THEORY OF MOLECULAR EVOLUTION

WERNER ARBER

Summary

Applications of scientific knowledge often refer to technological uses, but their impact on our world view can also be of great importance. Both of these kinds of applications have their cultural values. This view is here exemplified with recent developments in molecular genetics and evolution. Darwinian evolution resides on three pillars: genetic variation (or mutation), natural selection, and geographic and reproductive isolation. It is only since about 50 years that genetic information is known to be carried in DNA molecules. Molecular genetics and the theory of molecular evolution are the fruits of this knowledge. Molecular evolution is investigated by the comparison of genomic sequences and by the study of the molecular mechanisms generating genetic variations. Much of the available knowledge comes from microbial genetics. Genetic variation is brought about by a number of different specific mechanisms, which can be grouped into three different strategies, namely small local changes in the DNA sequences, rearrangement of DNA segments within the genome by recombinational processes, and the acquisition of foreign DNA by horizontal gene transfer. The three strategies to generate genetic variants have different qualities with regard to their contribution to the evolutionary progress. The available data clearly show the involvement both of gene products (acting as variation generators or as modulators of the frequency of genetic variation) and of non-genetic elements in the production of genetic variations. There is no real evidence that genetic variation would in general be a specific response to an identified need imposed by the environment that exerts natural selection. Rather, genetic variation is generally to some degree

aleatoric, and it is natural selection together with the availability of appropriate genetic variants which determines the direction(s) of biological evolution. In view of the activities of specific gene products to the benefit of biological evolution a dual nature of the genome becomes obvious. While many of its genes serve for the fulfillment of each individual life, others (the evolution genes) serve at the level of populations for the expansion of life, including the building up and replenishing of a rich biodiversity. The pertinence of this knowledge for our world view, as well as for the strategies of genetic research and its technological applications is discussed.

From fundamental scientific research to the application of its results

In this presentation I will focus attention to molecular genetics and more specifically to the mechanisms of molecular evolution. Upcoming knowledge on the genetic basis of biological activities and on their specific molecular mechanisms can serve us as guidance to apply that knowledge responsibly with the aim to facilitate the human life. This obviously enriches the patrimony of our civilization and represents thus cultural values. This kind of reflexion applies to many different fields of investigations in the natural sciences, so that the general conclusions regarding cultural values are of wide relevance.

Traditional research strategies in the biological sciences are largely observing and descriptive. In recent times experimental strategies of research are given increasing importance. They are often invasive, disturbing the system under study. For example, by knocking out the activity of a gene one can try to identify the biological function of that gene, by comparing the phenotype of an organism lacking the gene function with that of a genetically unaltered organism. Observing and invasive research strategies often differ in the kind and quality of their contributions to knowledge, they are largely complementary to each other. Applying different research strategies often involving trans- and interdisciplinary research is a good means to enrich our knowledge base. Since a knowledge base represents cultural values, these values increase with the increasing richness of the knowledge base.

Accumulated knowledge can lead to two kinds of applications. On the one hand, an application can be practical, often technological, and it is frequently invasive, causing some disturbance to the natural situation. Such practical applications may contribute to the shaping of the future, they may exert their influence on the longer-term development of things. On the

other hand, the knowledge base is also an important source of our world view; novel knowledge can bring about changes in the generally accepted world view. Intrinsically, the world view represents philosophical, thus cultural values. The responsibility assumed by human beings is largely based on their validated world view. The latter can indeed provide guidance to society in the shaping of the future. This represents an important feedback of the world view to the ways and the intensity by which practical, technological applications of scientific knowledge are made. Similar reflections can apply to policy decisions such as on legal regulations that are related to available scientific knowledge, to the search for novel knowledge and to the practical application of such knowledge. It becomes more and more obvious that the principle of sustainability should govern the influence exerted by human activities on the natural environmental conditions. Therefore, the world view aspects of scientific knowledge deserve as much attention as the immediate utility attributed to technologically based applications of the scientific knowledge base. These considerations shall be illustrated in the following sections by the relevance of a deepened knowledge on the process of biological evolution for genetic research and biotechnology.

Quest for molecular mechanisms of biological evolution

The theory of Darwinian evolution resides on three pillars: genetic variation, natural selection and isolation. Genetic variation is brought about by a number of different mechanisms causing alterations in the genetic information of an organism. Genetic variants (or mutants) represent the driving force of biological evolution. In contrast, natural selection together with the range of available genetic variants guides biological evolution into specific directions. Geographic and reproductive isolations modulate the evolutionary process.

Biological evolution is a relatively slow, but steady process, in which once in a while an individual member of a population of organisms is hit by a mutational event. In ecosystems, mixed populations of different organisms and different variants thereof are steadily submitted to natural selection. Thereby, those organisms that succeed to cope best with the encountered living conditions have a selective advantage, so that they will at longer-term overgrow their competitors. This largely depends on the genetic setup.

We know that genetic information is encoded by linear sequences of nucleotides in filamentous DNA molecules. These sequences contain genes and intergenic regions. A gene typically contains an open reading frame

that serves upon gene expression to instruct the manufacture of a specific gene product, which is often a protein with enzyme functions. The gene also contains expression control signals that serve to regulate the time and intensity of gene expression. The total genetic information present in each cell of an organism is called the genome.

In molecular genetics it has become a habit to call any alteration of the inherited nucleotide sequence a mutation. This contrasts with classical genetics in which the term mutation refers to an observed alteration of the phenotype, that results from the activities of the gene products. We will apply here the molecular genetic definition.

Spontaneous alterations in the nucleotide sequences are often attributed to errors upon DNA replication and to accidents occurring to the DNA. An alternative view, to be defended here, is to attribute mutagenesis to the common influence of particular gene products and of a number of non-genetic factors. Evidence for this interpretation can be expected from a deeper knowledge on the molecular mechanisms involved in the generation of genetic variants.

At present two approaches are available to explore the molecular mechanisms of genetic variation. One of these approaches is the systematic comparison of available nucleotide sequences of more or less closely related organisms. This strategy involving bioinformatic tools can be applied at the level of a gene for a specific function, at the level of a group of genes and also at the level of the genome. This can reveal single nucleotide changes, the reassortment of functional domains, as well as aspects of the genome organisation. Results obtained in such investigations can provide hints with regard to historical events that occurred to the ancestors of the compared organisms.

A more straightforward approach is the study of individual events generating genetic variants. Since these processes are both inefficient and generally not reproducible, their investigation is relatively difficult and has to be mostly indirect, by comparing the nucleotide sequences just before and after an event of mutagenesis.

Most of the data available so far on the generation of genetic variants at the molecular level come from microbial genetics, particularly from studies of bacterial and viral genomes. Bacteria are single-cellular organisms that propagate by cell division with typical generation times in the order of 30 minutes under optimal nutritional conditions. This facilitates population genetic approaches. Since the bacterial genome is haploid, spontaneously occurring mutations become phenotypically manifested rapidly.

These facts and the relatively small size of the microbial genome render studies of the molecular basis of genetic variation possible. From the available data it is clear that several different specific processes are at work, as will be discussed in more detail below.

It is a common observation made by many investigators that useful, beneficial mutations, are rather rare among the spontaneously generated DNA sequence alterations. More often, a mutation inhibits some life functions, thus providing a selective disadvantage (in extreme cases leading to lethality). Many other spontaneous DNA sequence alterations are without immediate influence on life processes. These are silent or neutral mutations. This situation is in line with the view that the spontaneous generation of DNA sequence alterations is in general not a specifically targeted answer to an identified need for adaptation of one or a few specific genes. In other words, there is no good evidence for a strict directedness of spontaneous mutagenesis.

Three major, natural strategies with different qualities contribute to the spontaneous formation of genetic variants

Referring to a more detailed outline given at the plenary session of our Academy in October 1996 (Arber, 1997), I can limit my presentation to an overview of the various molecular mechanisms that contribute each in its specific way to the generation of genetic variants. These conclusions are largely based on data obtained in microbial genetics, but they are likely to be generally valid also for higher organisms (Caporale, 1999).

Spontaneous genetic variation can be attributed to a number of mechanistically different events. These different mechanisms can be classified into three general strategies of genetic variation: local sequence change, DNA rearrangement within the genome and DNA acquisition. Each of these will be briefly characterized here.

The local sequence change brings about the substitution, deletion or insertion of a single or a few adjacent nucleotides. It can also result in a local scrambling of a few nucleotides. Several causes for such reactions have been identified, such as a limited chemical stability of nucleotides, a structural flexibility (tautomerism) of nucleotides implying alternative base pairing, other types of replication infidelities (e.g. replication slippage), as well as the effects of some chemical and physical, internal and environmental mutagens. The quality of local sequence changes resides primarily in their possibilities for a stepwise functional improvement of a gene and

for the potential adaptation to alternative living conditions. In the long term, a series of subsequent local sequence changes can, in principle, also result in a novel gene function, but one can assume that this comes only to bear once the product of the genetic information in question becomes a substrate for natural selection. To a large extent, local sequence changes are initiated by non-genetic factors, mainly intrinsic properties of matter and responses to interactions of matter. However, various enzymatic repair systems have been developed by living organisms to limit the frequencies of local sequence changes and their detrimental consequences to relatively low levels ensuring a certain degree of genetic stability but still allowing for a low, evolutionarily useful frequency of mutagenesis.

Intragenomic DNA rearrangements generally affect DNA segments of variable length. These can undergo deletion, inversion, translocation to another site in the genome, duplication and higher amplification. These genomic changes are sometimes accompanied by additional local sequence changes at the junction sites. DNA rearrangements are often, or perhaps as a rule, brought about by the recombinogenic activities of specific gene products. We call these gene products variation generators. Generally, they work inefficiently and act on the DNA molecules at one of many different possible sites, so that the results of their reactions are not strictly reproducible; they are at most statistically reproducible. DNA rearrangements can bring about novel gene fusions (the fusion of a part of one gene with a part of another gene). This can in some cases result in a novel genetic activity. Alternatively, a DNA rearrangement can also fuse a given reading frame of a gene with a hitherto unrelated signal for the control of gene expression. In the case of the duplication of functional sequences, a duplicate copy can later serve as a substrate for further evolutionarily relevant events while the other copy can continue to exert its normal function.

In diploid eukaryotic organisms general (homologous) recombination between the paternal and the maternal genomes, as well as the meiotic assortment of chromosomes, are other well known sources of genetic variations.

The third strategy to generate genetic variations, DNA acquisition, depends on the horizontal transfer of genetic information from a donor to a recipient organism. This process is well studied with bacteria where several different mechanisms contribute to the overall horizontal DNA transfer. The process occurs also in higher organisms where it is, however, less well explored than with bacteria. In the latter case, several factors have been identified to limit gene acquisition to low frequencies and to relative-

ly short segments of DNA. This strategy of gene acquisition in small steps represents a sharing in successful developments made by other kinds of organisms. The process is relatively effective: in a single event of alteration of the genome the recipient organism can gain a fully functional activity which may by chance satisfy an upcoming need for adaptation to changing living conditions, such as in the sudden presence of an antibiotic.

In view of the occasional horizontal flux of genome segments, the classical evolutionary tree should be drawn with randomly placed horizontal connectors (Arber, 1991). Remember that, in general, relatively short genome segments become horizontally transferred to another organism, while in the vertical transmission of the hereditary information from generation to generation, the entire genome becomes transmitted to the progeny and steadily represents a target for genetic alteration by local sequence changes as well as by intragenomic DNA rearrangements.

The theory of molecular evolution postulates the generation of genetic variations to depend on the coordinated action of the products of specific evolution genes and of non-genetic elements

There is no doubt that a number of non-genetic factors contribute each in a relatively specific way to the production of genetic variants. As was already mentioned, this mutagenesis often depends on properties of matter such as a certain degree of chemical instability and of structural flexibility of biological molecules. In addition, random encounter also plays its role, such as in the interaction of an enzyme with its substrate, in the random choice of a recipient organism for horizontal gene transfer, or when a DNA segment is hit by a mutagen.

On the other side, increasingly strong evidence supports the interpretation that the products of a number of specific genes act primarily for the benefit of biological evolution. Some of these so-called evolution genes act as generators of genetic variations, while others act as modulators of the frequency of genetic variation. Examples for the latter activities are found among the already cited repair systems. The transposition of mobile genetic elements is a good example of a variation generator. The theory of molecular evolution also postulates that those evolution genes that are encountered today in living organisms had been fine-tuned for their specific activities in their own evolutionary development involving second order selection, a selection process acting at the level of populations (Arber, 2003a).

Dual nature of the genomic information

The genome is usually thought to contain genes with specific tasks to be carried out for the benefit of each individual organism. These are the housekeeping genes and genes of use under particular life conditions. The developmental genes ensuring in higher organisms the development from a fertilized egg to the adult organism can also be counted to this large class of genes of general relevance for each living being. They serve for the fulfillment of each particular life.

Until recently, evolutionary developments were generally assumed to depend on errors and accidents occurring to the DNA molecules. In view of evidence for enzyme activities of unique relevance for biological evolution, but dispensable for the individual life span extending from one generation to the next, the view of the existence of evolution genes in the genomes obtains increasing support. If this view is correct, the genomic information must be of a dual nature with regard to its purpose. Clearly, as was already said, many genes act for the benefit of the individuals. In contrast, evolution genes act for the benefit of an evolutionary development of the population, by serving as generators of occasional DNA sequence variations and as modulators of the frequencies of such variations. By doing so, they sometimes cause harm to an essential life function, if a novel mutation happens to provide a selective disadvantage. This has to do with the fact that genetic variation is, in general, a largely random event rather than a precise, directed response to an identified need. The duality discussed here can be seen as a consequence of the engagement of nature to care not only for the fulfillment of individual lives but also for the evolutionary expansion of life and hence, for the evolutionary installment and replenishment of a high diversity of life forms on our planet. It should be added to this discussion that the products of some genes are used for both purposes: for the benefit of the individuals and for the evolutionary development. Genes of this kind may have been evolutionarily fine-tuned to carry out both of their tasks properly.

Cultural values of the knowledge on mechanisms of molecular evolution

The involvement of products of specific evolution genes for the driving of biological evolution that insures a rich biodiversity implies a widely unexpected and surprising modification of our world view. Nature cares actively for the evolutionary expansion of life. Properties of matter and genetically determined mechanistic capacities of life itself are identified as

coordinated driving forces of evolution. This represents an expansion of the Darwinian theory to the level of molecular processes and it strengthens the validity of this theory. The philosophical and hence cultural values of these conclusions are evident. Some of the aspects relating to an evolutionary, permanent creation were discussed in more detail elsewhere (Arber, 2003b), pleading for a reconciliation between, on the one hand, traditional wisdom such as the one transmitted in the Old Testament and, on the other hand, recently acquired scientific knowledge.

For several reasons the postulate of the presence in the genome of evolution genes and the knowledge on their ways of action also represent an enrichment for research on genomics and proteomics, as well as for the practical application of the results of these investigations in biotechnology. One aspect clearly illustrated by the enzymes generating genetic variations relates to the widely spread belief that genes encode strict programs for life processes. According to this view, primary gene products are thought to normally serve as enzymes in a sequence of events, the final output of which would be reproducible and therefore also predictable. In addition, the belief is quite widespread that enzymatic reactivities are always efficient. Evolutionarily relevant variation generators do not have these properties. Rather, they are inefficient and in the rare cases of their activities the output (which is a novel genetic variant) is not reproducible and not predictable from case to case. These aspects deserve due attention in the definition of the gene concept.

The knowledge on the three different basic natural strategies contributing to the generation of genetic variations forms a welcome basis for the evaluation of conjectural risks of genetic engineering, in particular in cases of deliberate release of a genetically modified organism (GMO). In genetic engineering, the genetic information is deliberately altered in a planned and thus *a priori* known way, e.g. by site-directed mutagenesis or by the horizontal transfer of a natural DNA sequence from one organism to another kind of organism. In all of these processes the investigator may apply, in principle, one or a combination of more than one of the described three natural strategies of genetic variation, i.e. local sequence change, intragenomic DNA rearrangement and horizontal gene transfer. Thereby, the quantity of involved base-pairs is, as a rule, in the same span as that observed in natural events of genetic variation. The use in genetic engineering of principally natural strategies which must have served in nature since a few billion years for promoting the evolutionary progress can at least suggest to us that conjectural, long-term risks of genetic engineering must be similar to those

related to the natural evolutionary process. These considerations have been outlined in more detail in a contribution to a workshop held by our Academy in February 2001 (Arber, 2002). This is a good illustration for how the scientifically based world view can have its feedback on technological applications of scientific knowledge. Such feedback provides means to responsibly carry out technological applications as contributions to the sustainable shaping of the future for the benefit of Mankind.

REFERENCES

- Arber, W. (1991) Elements in microbial evolution. *J. Mol. Evol.* 33, 4-12.
- Arber, W. (1997) *The influence of genetic and environmental factors on biological evolution*. In: Plenary Session on the Origin and Early Evolution of Life (Part I). The Pontifical Academy of Sciences, Commentarii Vol. IV, N3, pp. 81-100.
- Arber, W. (2002) *Molecular evolution: comparison of natural and engineered genetic variations*. In: The Challenges of Science. The Pontifical Academy of Sciences, Scripta Varia 103, pp. 90-101.
- Arber, W. (2003a) Elements for a theory of molecular evolution. *Gene* (in press).
- Arber, W. (2003b) *Traditional wisdom and recently acquired knowledge in biological evolution*. Proceedings of UNESCO Conference 'Science and the Quest for Meaning' (in press).
- Caporale, L.H. (Ed.) (1999) *Molecular strategies in biological evolution*. Annals of the New York Academy of Sciences, Vol. 870.

DISCUSSION ON THE PAPER BY ARBER

MALDAMÉ: Professor Arber, could I ask you three short questions? The first refers to your slide, the evolutionary tree. What is the importance of this horizontal transfer in evolution? The second question is: in our genome we carry a lot of viral, genomic and bacterium particles. Could they have a role in evolution? And then there is my third question, which is probably difficult and always comes up in discussions. You see, I share your view, but I would like to hear more arguments from you. We have often been told that this is an error mutation. You say that it is not an accident. You say that it's natural law. What are your reasons for saying this?

ARBER: I will begin with your last question. I think it is important that we scientists should not always base our views on textbook interpretations of the available data. We have to have some flexibility. I have become more and more convinced of the major importance of biological evolution. I can therefore not consider that biological evolution could be driven by errors and accidents. Rather, we should look for functions that actively generate genetic variations. Transposable genetic elements are a good example, they exert no other function, as far as we know, than to produce occasionally novel genetic variations. As to the replication infidelities, they largely depend on properties of matter, such as a certain degree of structural flexibility and of chemical instability of nucleotides. These are intrinsic properties, not accidents.

In your first and second questions you asked about horizontal gene transfer involving, among other elements, viruses and in higher organisms sometimes even bacteria. The evolutionary role of these elements is well studied with bacterial populations. The horizontal gene transfer is in general a rare event. But if an acquired gene provides to the recipient an advantage, natural selection will not only maintain it, but also amplify selectively the novel hybrid. In this regard we learned a lot by observing the wide spreading of antibiotic resistance genes in bacterial populations due to the

extensive use of highly selective antibiotics in human and veterinary medicine in the last sixty years. There is still less knowledge on the role played by horizontal gene transfer in higher organisms, where more relevant research is required. But it is known that some viruses can serve as natural gene vectors also in higher animals and thus in man. As to the general importance of horizontal gene transfer, this resides in its different quality as compared to the other natural strategies of generating genetic variants. In horizontal gene transfer an organism may have the chance to acquire from some other organism a genetic function which it had not possessed before. Such an acquisition is a one-step event. If these organisms would have to develop the same function itself, this would be a very laborious multistep process.

MALDAMÉ: I have a question. You reported at the end that genes' actions serve a purpose. What kind of purpose?

ARBER: I think the important purpose of many genes is to be seen in the fulfillment of individual lives. In contrast, evolution genes working at the level of populations provide means to produce and to replenish biodiversity.

By the way, the knowledge on molecular mechanisms of evolution can offer insights into the sense of life and the sense of death, although only in the context of the evolutionary development. We can compare naked DNA molecules with a closed library: nothing happens as long as there are no readers present. Any potential actions depend on the activities of readers. Reading of the genetic information on organisms is also the prerequisite for life manifestations. Remember that it is the life manifestations, not the DNA molecules, which are the substrate for natural selection. Active life has thus its clear evolutionary meaning. As to the evolutionary sense of death, remember that genetic variation is the driving force of biological evolution. Genetic variation depends on a steady renewal of the populations. Eternal lives could not satisfy this condition. In addition the space for living organisms in the biosphere is limited. The turnover of populations necessitates the death of individuals after having served for some time as substrates for natural selection.

SINGER: Prof. Arber, you gave us a very interesting and detailed description of molecular evolution where it can be studied. Would you agree with me that there has been no further evolution in human beings over the last 5,000 years because of a lack of selection pressure and because of a lack of inbreeding?

ARBER: I don't agree. I think we evolve steadily, like any other living organism, but in our short life span we cannot spot evolution so easily – this is not possible. Reading Genesis we can learn that the descendants of Adam and Eve (the products of creation) are not clones with identical properties. Each descendant has a specific character. I see in that description a traditional wisdom that within a species, all living beings are different from one another, and this is exactly the driving force of evolution. So I cannot say that we do not continue to evolve.

SINGER: Just a remark referring to your account of creation. Recently I talked to a Jewish priest who told me that in the original text of the Hebrew Old Testament there is a predecessor of Eve, Lilith. She was made out of clay at the same time as Adam, and they had equal rights. She was not part of man, she was no clone, but Adam could not cope with her because she was too self-determined. So he sent her away and she then lived with animals in the region of Jordan, but she had offspring.

JAKI: He should have told you that Lilith is in the Talmud and not in the Bible. That is the first thing. The second thing is that whenever a creationist claims that we have to take in a scientific, in a modern scientific sense that God created everything according to its kind, that is, each species separately, we should remind him that if you take one single phrase in the first chapter of Genesis in a scientific sense, then the basic rules of interpretation demand that you should take all the other statements in a scientific sense: then you have light coming before the sun, then you have plants coming before sunlight, and finally you may ask them: did the astronauts wear helmets to protect their heads when they went through the firmament?

ARBER: I agree with you that Genesis is not a strictly scientific text. I also became aware that plants were created before animals, and my interpretation is that this is in line with traditional knowledge. You know that animals eat plants, that's their food. You cannot create animals and then plants only the next day, because otherwise these created animals will die in the meantime. So, there is some logic in the sequence of events. And there is no mention of micro-organisms because human beings did not know at that time that there were microbes around.

VICUÑA: This may be a semantic problem, but I wonder if you can eliminate error as a source of evolution, because the enzymes that make

DNA, DNA polymerase, do make errors, otherwise they wouldn't have an additional activity that allows them to correct those errors. I mean, nucleotides are so similar to each other that I would imagine that an enzyme that is copying a template would make an error. They do the job so fast that they can make mistakes.

ARBER: The term error is a human interpretation of some unusual, often unexpected observation. In the case of DNA replication, we know several specific reasons for the incorporation of another nucleotide than the one expected from the sequence of the template strand. One of these reasons are tautomeric forms of nucleotides. These conformational variants are relatively rare and very shortlived. Most importantly for our discussion, they give rise to an altered base-pairing. This can result in a substitution mutation. To my mind, this is not an error and I call it a replication infidelity, which goes back to an intrinsic property of the molecules, their structural flexibility. Another source of base substitution is a certain degree of chemical instability of nucleotides.

As a matter of fact, these sources of mutagenesis were among the first to be known, and many textbooks on genetics and evolution generalize and propagate the idea that spontaneous mutations are base substitutions going back to errors in replication.

LÉNA: I am struck by a number of words that you have been using. Maybe it is only semantics, but maybe it is more. You have the words 'error', 'help', 'purpose', 'goal', 'use'. I wonder if this is not somewhat anthropomorphic and what Paul Ricœur would have to say on the use of those words to describe molecular biology?

ARBER: Well, you may have seen that my talk was actually trying to make a bridge between science and culture, and I used some words that could be more easily understood. What I gave was not a fully scientific talk, although I tried to give you some evidence for my ideas.