



# Genetic engineering compared to natural genetic variations

**Werner Arber**

Biozentrum, University of Basel, Klingelbergstrasse 50-70, CH-4056 Basel, Switzerland

By comparing strategies of genetic alterations introduced in genetic engineering with spontaneously occurring genetic variation, we have come to conclude that both processes depend on several distinct and specific molecular mechanisms. These mechanisms can be attributed, with regard to their evolutionary impact, to three different strategies of genetic variation. These are local nucleotide sequence changes, intragenomic rearrangement of DNA segments and the acquisition of a foreign DNA segment by horizontal gene transfer. Both the strategies followed in genetic engineering and the amounts of DNA sequences thereby involved are identical to, or at least very comparable with, those involved in natural genetic variation. Therefore, conjectural risks of genetic engineering must be of the same order as those for natural biological evolution and for conventional breeding methods. These risks are known to be quite low. There is no scientific reason to assume special long-term risks for GM crops. For future agricultural developments, a road map is designed that can be expected to lead, by a combination of genetic engineering and conventional plant breeding, to crops that can insure food security and eliminate malnutrition and hunger for the entire human population on our planet. Public-private partnerships should be formed with the mission to reach the set goals in the coming decades.

## Contents

Introduction . . . . .	518
Principles of the Neo-Darwinian theory of evolution . . . . .	518
Towards molecular Darwinism. . . . .	518
The genetic script. . . . .	518
Definition of the term mutation . . . . .	518
Effects of mutations . . . . .	518
Molecular mechanisms of genetic variation. . . . .	518
Natural strategies of genetic variation . . . . .	519
Local sequence changes . . . . .	519
Rearrangement of intragenomic DNA segments. . . . .	519
DNA acquisition by horizontal gene transfer . . . . .	519
The tree of evolution . . . . .	519
A new evolutionary synthesis . . . . .	520
Natural reality actively takes care of biological evolution . . . . .	520
The duality of the genome . . . . .	520

*E-mail address:* w.arber@unibas.ch, Werner.Arber@unibas.ch.

From classical to modern biotechnologies . . . . .	520
The impact of reverse genetics on modern biotechnology . . . . .	520
Evaluation of conjectural long-term evolutionary risks of genetic engineering . . . . .	520
A road map for future agricultural biotechnologies . . . . .	521
References. . . . .	521

## Introduction

Genetic engineering was introduced around 1970 as a highly potent strategy for genetic research at the level of DNA molecules, the carriers of genetic information. This strategy consists principally of introducing nucleotide sequence alterations into DNA molecules, such as by site-directed mutagenesis and by splicing DNA segments from different locations in the genome or from different kinds of organisms (recombinant DNA molecules). Genetic engineering has rapidly become an efficient strategy for structural and functional studies in genomics.

Already at an early time, scientists raised the question of conjectural risks of their experimental approach. This led in February 1975 to an international conference held in Asilomar, California. There, conjectural risks were seen at two levels. On the one hand, short-term, rapidly manifested risks were proposed to be investigated, case-by-case, under laboratory conditions in analogy to the medically relevant diagnosis of pathogens and to investigations on the effects of toxic substances, avoiding any impact on the health of the investigators. On the other hand, long-term risks could be expected to become of evolutionary relevance after deliberate release of organisms carrying genetically modified (GM) DNA. For the assessment of such conjectural risks, monitoring was envisaged, as well as a comparison between the deliberate alteration of genetic information by genetic engineering and the naturally occurring spontaneous generation of genetic variants, which are the drivers of biological evolution. This comparison is the aim of the present article. It is a follow-up of earlier publications ([1,2]; see also [3,4]).

## Principles of the Neo-Darwinian theory of evolution

Any large population of living organisms contains individuals having suffered a genetic variation. Such variants can be identified by specifically altered phenotypic traits. These spontaneous mutants drive biological evolution. Together with their parental forms, their traits are the substrate for natural selection. The latter results from the environmental constraints that are exerted on living organisms by the physico-chemical composition of the environment and by the activities of other kinds of living beings in the natural ecosystems. Natural selection, together with the available genetic variants, guides the direction of biological evolution. Reproductive and geographic isolations represent the third pillar (besides genetic variation and natural selection) of biological evolution and they modulate the process of evolution.

## Towards molecular Darwinism

It is thanks to experimental work on microbial genetics [5] and in structural biology [6] that we have known for about 60 years that long filamentous molecules of DNA are the carriers of genetic information.

## The genetic script

DNA molecules are composed of linearly arranged sequences of four different nucleotides that form specific base pairs in the double-stranded form of DNA. Genetic information is contained in the linear sequences of these building blocks, comparable to the linear sequences of letters in our writing. Remaining with this metaphor, the genome (i.e. the entire genetic information) of a bacterium corresponds to one book, whilst the genomes of higher organisms correspond to many books, ranging up to encyclopedias of several hundreds to a thousand books. A classical gene, the determinant for a specific gene product, ranges between a few lines to about one page. As we will see below, this metaphoric comparison can help us in the comparison of genetic variations caused either spontaneously or by genetic engineering.

## Definition of the term mutation

Note that we use here the terms 'mutation' and 'genetic variation' synonymously. In classical genetics a mutation is identified by an altered phenotype that becomes transmitted to the progeny. By contrast, in molecular, reverse genetics a mutation is defined as an altered nucleotide sequence. Thus, it is advisable to be aware of this difference in the use of the term mutation.

## Effects of mutations

It is generally known that altered nucleotide sequences turn out to be only rarely favourable, useful for the organism that has suffered the mutation. Often, a mutation provides selective disadvantage by inhibiting to some degree the life processes. In extreme cases this can be lethal. Also quite often a new alteration in the nucleotide sequence has no immediate influence on the life processes. These are neutral, silent mutations. Consequently, we cannot identify evidence for a directedness of spontaneous mutations and the rates of spontaneous mutagenesis must be kept quite low under natural conditions not to eradicate life.

## Molecular mechanisms of genetic variation

Textbooks often state that spontaneous mutations represent errors or accidents which occur in the DNA, for example, upon DNA replication. In view of the now available, more profound knowledge on singled-out events of genetic variation, this concept of errors does not correspond to the reality. Particularly from experimental research with microorganisms, but increasingly also from DNA sequence comparisons involving evolutionally more or less closely related organisms, we know that many different specific molecular mechanisms contribute to overall genetic variation.

Some mutations are due to intrinsic infidelities of DNA replication. Short living isomeric forms of biological molecules represent a prominent source of replication infidelities. For example, a tautomeric imino form of the nucleotide adenine can no longer pair with thymine, but it can pair with cytosine. After returning

into its standard form, adenine's partnership with cytosine becomes a mispairing [7]. It is thanks to specific activities of repair enzymes that most such mispairings, sources for nucleotide substitutions, are rapidly eliminated after the passage of the DNA replication fork. Other disturbing effects on local nucleotide sequences, such as deletion or insertion of one or a few adjacent nucleotides and the scrambling up of a few neighbouring nucleotides, can also be attributed to intrinsic properties of the replication machinery.

Other genetic variations are attributed to intragenomic rearrangements of DNA segments. Such reshuffling of DNA segments is generally mediated by recombination enzymes (see 'Rearrangement of intragenomic DNA segments').

Still other genetic variations are due to the uptake of segments of foreign DNA. As a rule, this is also mediated to a large part by specific gene products (see 'DNA acquisition by horizontal gene transfer').

### Natural strategies of genetic variation

On the basis of our knowledge of specific molecular mechanisms contributing to spontaneous genetic variation, one can conceptually attribute each particular mechanism to natural strategies for generating genetic variants. As we will see, each of the three strategies here described contributes with a different quality to the occasional formation of genetic variants and thus to biological evolution.

#### *Local sequence changes*

Replication infidelities, such as those described in 'Molecular mechanisms of genetic variation', represent local sequence changes affecting usually only one or a few adjacent nucleotides. Chemical mutagens, either internal or environmental, often cause local sequence changes as well. Such changes can affect open reading frames, gene expression control signals or other sequences that are directly or indirectly involved in cellular functions. One can expect that only rather rarely will a local sequence change represent a favourable alteration and provide a selective advantage. But the rare, beneficial mutations represent, in general, a stepwise improvement of an available biological function.

#### *Rearrangement of intragenomic DNA segments*

Contributions to this kind of natural strategy of genetic variation are usually brought about by the action of recombination enzymes, that is specific gene products that we call here variation generators. Such enzyme systems with various specificities are found in all living organisms.

In the general recombination, more or less extended homologous stretches of nucleotides (often involving one line to about one page of the genomic library), become aligned, cut and repasted, so that recombinants are formed.

Mobile genetic elements, often involving a few lines to one page of the genomic library, are widespread in living organisms. These elements can occasionally transpose to another chromosomal location. Depending on the characteristics of the involved enzymes, this process may or may not involve further DNA sequence alterations. In the microbial world, one has already identified a large number of specific mobile genetic elements, each following its own specific mode of recombinant activities (e.g. see Ref. [8]).

Whilst site-specific recombination, in general, reproducibly splices DNA segments together at relatively short specific or consensus sequences, the underlying enzymes can very occasionally also use one of a large number of different secondary crossover sites. These latter, quite rare activities are a good source of evolutionally relevant fusions of different functional domains in the genetic information [8].

With regard to their contributions to the process of biological evolution, all these enzymatic variation generators can bring about an improvement or novel uses of available genetic capacities. For example, fusion between two previously separated functional domains (gene fusion) may lead to a novel ability, and the fusion of an open reading frame with a previously separated expression control signal can lead to a higher or a lower yield of the gene product concerned.

#### *DNA acquisition by horizontal gene transfer*

Microbial genetics took its fulgurant start some 70 years ago. It unravelled within one decade the basic principles by which prokaryotic microbial organisms can exchange genetic information. In transformation, free extracellular DNA can be taken up by so-called recipient bacteria [5]. In conjugation, a donor cell can pair with a recipient cell and thereby transfer parts of its genetic information into its partner cell [9]. In bacteriophage mediated transduction, a bacterial virus can serve as a gene vector after having incorporated donor DNA into infectious progeny viral particles [10]. Studies of these processes were facilitated by the availability of microbial mutants, so that recombinants could be identified between the involved donor and recipient bacterial strains. Whilst these processes proved to be efficient as long as donor and recipient strains belong to the same kind of bacteria, they also promote genetic exchange between more or less related microbes, although with much lower rates. As a matter of fact, several different natural barriers keep the rates of this so-called horizontal gene transfer at very low levels. Important barriers are, on the one hand, surface incompatibilities hindering the penetration of donor DNA into recipient bacteria, and on the other hand, DNA restriction-modification systems enabled to identify foreign DNA and to cut it into fragments. Only rarely can such a fragment find its way to integrate into the recipient genome before its rapid exonucleolytic digestion [11]. A last barrier acts at the level of expression of acquired genetic information: the functional harmony of the resulting hybrid must not be disturbed, otherwise natural selection will sooner or later eliminate hybrid forms from the concerned microbial population. Qualitatively, horizontal gene transfer can represent an extremely effective step in biological evolution, but for the abovementioned reasons, in reality it is allowed to occur only very rarely. Success of DNA acquisition is best if it occurs in small steps, involving some lines up to about one page of the book of bacterial genetic information.

#### *The tree of evolution*

With regard to the evolutionary contributions brought about by the DNA acquisition strategy, we draw the classical evolutionary tree with occasionally placed connectors between branches [12]. Hence, living organisms must have not only a common past, but also a common future, at least to some degree. As a matter of fact, there is increasing evidence that the strategy of DNA acquisition is

not limited to the world of microorganisms, but it also contributes to the biological evolution of higher organisms, sometimes spanning wide distances of evolutionary relatedness.

#### *A new evolutionary synthesis*

On the basis of specific knowledge on molecular mechanisms and natural strategies for the generation of genetic variants, one can envisage incorporating this knowledge into the Neo-Darwinian theory, in analogy to the modern evolutionary synthesis which around 1940 brought classical genetics together with the Darwinian theory of evolution and which resulted in the Neo-Darwinism [13]. The result of the new evolutionary synthesis can be called *molecular evolution* or *molecular Darwinism*.

#### **Natural reality actively takes care of biological evolution**

As we have seen, the overall genetic variation depends both on the availability of specific enzymes (acting as variation generators and as modulators of the rates of genetic variation) and on non-genetic elements including structural and functional flexibility of nucleotides, environmental mutagens and random encounter.

Enzymes are gene products. For the microbial world it has become clear that many of these gene products are inessential for the normal life of a cell from one generation to the next. Their biological function is clearly to foster biological evolution. We therefore call their genetic determinants evolution genes.

#### *The duality of the genome*

Unexpectedly we realise that not all of the genes carried in a genome serve for the fulfilment of the life of an individual during its lifetime. The products of evolution genes serve mainly for a constant, but slow evolutionary development at the population level. They serve for an expansion of life, for biodiversity. In other words, thanks to a well-balanced synergy between products of evolution genes on the one hand and non-genetic, intrinsic properties of matter and random encounter on the other, biological evolution steadily proceeds and nevertheless ensures to individuals a certain genetic stability, without which life would not be possible. We assume that in the long evolutionary history of life on our planet, evolution genes have been fine-tuned for their activities by second-order selection [14]. Organisms which had become genetically able to drive evolution by the three described, qualitatively different, natural strategies of genetic variation and to limit genetic variation to tolerably low rates, had an advantage over others, and this may have led to the functionally fine-tuned activities that we now observe in today's living organisms.

#### **From classical to modern biotechnologies**

Biotechnology takes advantage of biological functions and frequently uses the available knowledge to facilitate human life. Increasingly, care for sustainability of the development serves as guidance for biotechnological applications.

In classical biotechnological approaches, organisms were normally used as found in nature. Improvements of their envisaged activities could sometimes be reached by breeding techniques between related organisms. In more recent times, mutagens served to increase mutation rates and thus to procure a random improvement of the functions concerned and their availability.

#### *The impact of reverse genetics on modern biotechnology*

Reverse genetics makes use of components from genetic engineering. The sorting out of a particular segment of a genome and the carrying out of structural and functional studies with such a DNA segment, can lead to an understanding of its biological functions. This can be seen as fundamental research. In view of envisaged innovative applications, scientists may try in translational research to obtain improvements by site-directed mutagenesis, affecting the open reading frame of the gene in question. This can alter the gene product in a particular functional property. Alternatively, such mutagenesis exerted on the expression control signal may alter the yield of the envisaged product. In contrast to the possibilities of classical biotechnology, one can try in modern biotechnology to introduce the specific genetic information into another organism that might be more appropriate for the biotechnological production and further use of the envisaged products. These novel possibilities make modern biotechnological applications increasingly attractive.

#### *Evaluation of conjectural long-term evolutionary risks of genetic engineering*

Let us now compare the kinds of genetic variations carried out in genetic engineering with those acting in the natural, spontaneous generation of genetic variants. In both cases, the same three strategies of genetic variation are involved: small local sequence changes, intragenomic DNA reshuffling and acquisition of external, foreign DNA by horizontal gene transfer. Both in genetic engineering and in natural biological evolution, similar amounts of nucleotides are thereby generally involved, ranging from one letter to one or at most a few pages of the genomic encyclopaedia. In view of the implication of similar molecular mechanisms and similar amounts of DNA sequences involved in these genetic variations, one can expect that conjectural risks are also comparable for the natural biological evolution (including classical breeding techniques) and for genetic engineering. There is no scientific reason to claim that genetic engineering, as an efficient research strategy, would bear particular conjectural evolutionary risks. From our long-term experience, we know that neither natural evolution nor classical breeding activities have caused major, noted disasters in the living world. It is thus highly unlikely that such disasters could result from genetic engineering.

In this context, it is, nevertheless, advisable to maintain carefulness in human contributions to the process of biological evolution. This responsibility should equally concern contributions by genetic engineering and by classical breeding. Scientific knowledge is today available to test carefully in a case-by-case approach the kinds of alterations introduced into DNA sequences, and thus also into functional gene products, before their release into the environment for the benefit of humankind and of our natural environment. Available scientific knowledge and potent investigation methodology represent an efficient and effective basis for *a priori* responsibly carried out technology assessments before GM-organisms, either as produced by genetic engineering or as selected by classical breeding, become released into the environment. Any decision taken on such releases should be based on the specific biological functions involved, not on the ways by which the selected organisms were produced.

### A road map for future agricultural biotechnologies

In the long past history of agriculture, selection of food plants did largely follow the principle of trial and error. Random mutagenesis in the absence of knowing the physico-chemical basis of genetic information can nowadays be seen as blind genetics. As we have discussed in 'The impact of reverse genetics on modern biotechnology' and 'Evaluation of conjectural long-term evolutionary risks of genetic engineering', much more powerful research strategies are now available, both to stepwise alter genetic information and to assess the effects that such alterations can cause. In addition, rapid advances in genomics, proteomics and metabolomics provide us a wealth of knowledge on genetic functions and on nutritional requirements for our daily diets. This situation enables us to envisage programmes to specifically improve nutritional values of our common food plants. A convincing example is the so-called golden rice which provides us the required amounts of vitamin A [15]. In following this example, one can expect that it should be possible to enrich the nutritional values of our common food plants with various capacities to ensure nutritional requirements for the entire human population of our planet. At the same time, one should also envisage improving the health of the food plants themselves, both during their growth and during storage.

With this idealistic goal in mind, a road map has been described [16] that might serve as a guiding principle for the next few decades of agricultural development. The proposed road map respects environmental constraints such as the limited availability of fertile soils and of fresh water, and it also respects the preservation of a rich biodiversity and of the climate. In other words, the envisaged development is expected to be highly sustainable.

Under these conditions, priorities must be set for agricultural biotechnologies. A high priority should be given to the production of food for humankind. As we have already outlined, GM crops should be envisaged to have good health themselves and to ensure

high nutritional values, vitamins, minerals, essential amino acids, etc., which provide healthy, well-balanced food to the worldwide human population. If this goal can be attained, one can expect that eating habits may tend to shift towards largely vegetarian food. This will consequently render less pressing the production of animal food. The use of fertile soils for growing animal food can then be given a low priority.

High priorities could also be given to agriculture for biopharming, the growth of appropriately modified plants yielding products of medical relevance. Responsibly designed plants for bioremediation (amelioration of soil quality) should also be given a high priority. By contrast, and in view of ensuring food security without interfering with the goal for sustainability, low priority should be given for growing crops for obtaining commodities such as cotton and bioplastics. And last, but not least, low priority should be given to the production of biofuels.

It will be advisable for the political leadership, as speakers for the civil society, to form partnerships with the scientists and economists, to follow the road map drawn with the aim of guiding agriculture towards a sustainable future. This can ensure, on the one hand, durable food security for the human population and, on the other hand, the preservation of the environmental richness of the inanimate and the animate worlds. Scientific methodology and knowledge are rich enough to attain the set goal. Genetic engineering can contribute hand-in-hand with conventional breeding techniques to the envisaged development. A responsible, reliable assessment of envisaged introductions of GM crops can also be based on scientific methodology and knowledge. One can expect that the realisation of the envisaged development will have a good chance to be accomplished within a very few decades, provided the politicians drive the proposed action and favour the appropriate, scientifically based information of the general public.

### References

- Arber, W. (2002) Molecular evolution: comparison of natural and engineered genetic variations. The Pontifical Academy of Sciences, *Scripta Varia* 103 pp. 90–101
- Arber, W. (2006) The impact of microbial genetics on the development of genomics and biotechnology. The Pontifical Academy of Sciences, *Acta* 18 pp. 219–237
- Arber, W. (2003) Elements for a theory of molecular evolution. *Gene* 317, 3–11
- Arber, W. (2007) Genetic variation and molecular evolution. In *Genomics and Genetics*, (vol. 1) (Meyers, R.A., ed.), pp. 385–406, Wiley-VCH
- Avery, O.T. et al. (1944) Studies on the chemical nature of the substance inducing transformation in pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* 79, 137–158
- Watson, J.D. and Crick, F.H.C. (1953) Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. *Nature* 171, 737–738
- Watson, J.D. and Crick, F.H.C. (1953) The structure of DNA. *Cold Spring Harb. Symp. Quant. Biol.* 18, 123–131
- Arber, W. (2008) Stochastic genetic variations and their role in biological evolution. In *Predictability in Science: Accuracy and Limitations*, (Arber, W., Cabibbo, N., Sanches Sorondo, M., eds) pp. 126–140, The Pontifical Academy of Sciences, Acta 19
- Lederberg, J. (1947) Gene recombination and linked segregation in *E. coli*. *Genetics* 32, 505–525
- Zinder, N. and Lederberg, J. (1952) Genetic exchange in *Salmonella*. *J. Bacteriol.* 64, 679–699
- Dussoix, D. and Arber, W. (1962) Host specificity of DNA produced by *Escherichia coli*. II. Control over acceptance of DNA from infecting phage  $\lambda$ . *J. Mol. Biol.* 5, 37–49
- Arber, W. (1991) Elements in microbial evolution. *J. Mol. Evol.* 33, 4–12
- Mayr, E. (1982) *The Growth of Biological Thought: Diversity, Evolution and Inheritance*. Harvard University Press
- Weber, M. (1996) Evolutionary plasticity in prokaryotes: a panglossian view. *Biol. Philos.* 11, 67–88
- Ye, X. et al. (2000) Engineering provitamin A ( $\beta$ -carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* 287, 303–305
- Arber, W. (2009) The impact of science and technology on the civilization. *Biotechnol. Adv.* 27, 940–944