THE ROLE OF CHANCE IN EVOLUTION

GIORGIO BERNARDI

I would like to start this contribution on a personal note by mentioning that I come from one of the few, perhaps the only Institute in the world, the Stazione Zoologica of Naples, which was established in order to prove a theory, in our case Darwin's theory (1). After its foundation by Anton Dohrn in 1873, investigations at the Stazione concentrated on what was possible to investigate at that time, namely the morphology, the physiology and the embryology of marine organisms, their great biodiversity being the main reason for the choice of Naples as the seat of the Institute. For a century after the death of Anton Dohrn in 1909 practically no work on evolution was done. At the beginning of 1998 I took the direction of the Stazione Zoologica and started a Laboratory of Molecular Evolution which still is very active. I will report here on our work on genome evolution and its general implications.

THE ROLE OF CHANCE IN EVOLUTION

The first question one may raise about the role of chance in evolution is why this issue is so important. One may think about a number of explanations, but I prefer here to use a shortcut, by concentrating on the position presented in 1970 by Jacques Monod in his famous book *Le hasard et la nécessité* (2). There are three main reasons for this choice. The first one is the clarity of the ideas, the second the extreme stand and the third the discussion of its implications. These points make it easier to understand the problem under consideration here. Some key sentences clearly summarize the stand of the author: (i) *The origin of life on earth was due to a single chance event* and, since all living organisms descend from a common ancestor, (ii) 'the biosphere is completely separated from the inanimate environment', and 'Man knows to be alone in the indifferent immensity of the Universe, from which he emerged

by chance'. As far as the evolution of living organisms was concerned, Monod expressed the opinion that (iii) 'Mutations are accidents that happen at random. Since they represent the only source of changes in the genetic text, which is the only repository of inherited structures of organisms, it necessarily follows that chance is responsible for any novelty, for any creation in the biosphere', the conclusion being that 'Chance only is the source of every novelty, of every creation in the biosphere. Sheer chance, chance only, absolute but blind freedom at the very roots of evolution: this central notion in modern biology is not anymore a hypothesis among other possible or at least conceivable ones. This hypothesis is the only conceivable one, since it is the only one which is compatible with observation and experience. And nothing allows us to imagine (or to hope) that our ideas on this point will need, or will be subject to, revision'. Finally, Monod considered the implications of his conclusions and proposed an 'ethics of knowledge', which will be discussed at the end of this paper.

The best comment on Monod's book was made by Eigen (3) 'The only thing lacking in molecular biology was its integration into a general understanding of Nature. So far, such an attempt has been undertaken only once, by Jacques Monod. This was a fascinating and ambitious attempt, in which Monod did not shrink from drawing philosophical conclusions. It culminated in an apotheosis of chance'.

The classical evolutionary theories

The role of chance in evolution was not, however, a new problem. Let us look at which way mutations were visualized by the classical evolutionists. The most famous sentence in *The Origin of Species* (1) was the following: 'I have called Natural Selection, or the Survival of the Fittest, this preservation of favorable individual differences and variations and the destruction of those which are injurious variations'. This statement looks extremely simple, but Crick (4) remarked that 'Natural Selection is the basic mechanism that makes biology different from all other sciences. Of course anyone can grasp the mechanism itself, though remarkably few people actually do so'. Indeed, Darwin's sentence seemed to indicate a dichotomy, and was widely interpreted that way. The sentence was, however, immediately followed by another one, which is only rarely quoted: 'Variations neither useful nor injurious would not be affected by natural selection and would be left either a fluctuating element … or would ultimately become

fixed'. This still is the best definition of neutral changes. In other words, Darwin distinguished not two but three kinds of changes or mutations (which he called 'variations'): advantageous, deleterious and neutral.

Advantageous changes will tend to expand in the progeny, because the carriers and their progeny will reproduce more abundantly than average (this is the positive or Darwinian selection). In contrast, deleterious changes will tend to disappear from the population, because the carriers and their progeny will reproduce less abundantly (this is the negative or purifying selection). Finally, neutral changes may be fixed in the population (like advantageous changes) or disappear (like deleterious changes).

The idea of neutral changes was later obliterated by the neo-darwinians, the selectionists Fisher (5) and Haldane (6), only to be resurrected, later, by Kimura (7, 8) in his *mutation-random drift theory*. According to this neutral theory 'the main cause of evolutionary change at the molecular level – change in the genetic material itself – is random fixation of selectively neutral or nearly neutral mutants'; therefore, 'increases and decreases in the mutant frequencies are due mainly to chance'. As a logical consequence, this theory eventually replaced the *survival of the fittest* with the *survival of the luckiest* (9). Along the same line, King and Jukes (10) claimed in their *non-darwinian evolution* that 'most evolutionary changes in proteins may be due to neutral mutations and genetic drift' (the random changes in gene frequencies in a population). A significantly different position was taken by Ohta (11, 12) who proposed her *nearly neutral theory* according to which 'a substantial fraction of changes are caused by random fixation of nearly neutral changes, namely changes that are intermediates between neutral and advantageous, as well as between neutral and deleterious classes'. Fig. 1 (see p. 601) summarizes the points just mentioned.

It is now of interest to look at the experimental approaches used to develop the classical theories on evolution because of the tight links that exist between approaches, results and conclusions. Natural selection acts on the phenotype, namely the detectable characters (traits, features, properties) of living organisms. It is, therefore, understandable that the first approach to the study of evolution was based on morphological traits, a classical case being that of the beaks of the Galapagos finches, which show adaptations to different kinds of food, from hard seeds to soft vegetal tissues. After the rediscovery of Mendel's laws, the neo-darwinians relied on genetic characters. Only later a molecular approach was developed on the basis of the early protein and gene sequences, and this led to the neutral theory of Kimura. Indeed, the view that amino acids change linearly with

time in proteins (the molecular clock of Zuckerkandl and Pauling, 13), provided the very first hint in that direction.

The organization of the eukaryotic genome

A totally different approach moving from the molecular level of a few proteins and genes to the genome level was the one I started in 1959 by degrading DNA from mammals and birds with a DNase (14), and by fractionating DNA on hydroxyapatite columns (15). These experiments (probably the first ones in genomics) produced important results, such as the breakage of the genome into large fragments and the separation of double- from single-stranded DNA. Most of the following work was done, however, after our development (16) in 1968 of density gradient ultracentrifugation of DNA in the presence of sequence-specific DNA ligands (such as Ag+ ions), and our discovery in 1973 of the compositional heterogeneity of the bovine genome (17). Our *compositional approach* to the study of the genome, incidentally the only one that was possible at that time, was easily moved from the analysis of buoyant density profiles to nucleotide sequences as soon as these became available. The rationale of the compositional approach was that the base composition of the genome, the most elementary property of DNA, (i) is altered by mutations, insertions and deletions; (ii) influences DNA, RNA, protein and chromatin structure (see below); and (iii) can be precisely assessed on whole genomes and their domains. The conceptual simplicity of the approach is such that the results can be easily understood.

The compositional approach led to three major discoveries: (i) the vertebrate genomes (the only ones discussed here) are *mosaics of isochores* (18, 19), megabase regions (1 Mb is one million base pairs; the human genome is 3200 Mb in size) of fairly homogeneous GC level (Fig. 2, see p. 602); GC is the molar ratio (the percentage of the molecules) of guanine and cytosine in DNA); (ii) isochores belong in a few families, characterized by different levels of GC, dinucleotides and trinucleotides, and define a *genome phenotype* (20), namely the *compositional landscape* of the genome (see Fig. 3, p. 603); the GC-rich, gene-rich and the GC-poor, gene-poor isochores define two *gene spaces*, the *genome core* and the *genome desert,* that are correlated with all the basic structural and functional properties of the genome, the main ones being chromatin compaction, DNA methylation, gene distribution on the one hand, gene expression, recombination, replication timing on the other (see Fig. 4, p. 604); (iii) a *genomic code* (20; not to be confused with the

genetic code) correlates the compositions a) of coding sequences with those of contiguous non-coding sequences (*i.e.*, of 1% of the genome with the remaining 99%), b) of the three codon positions among themselves, and c) of coding sequences with the hydrophobicity and the secondary structure of the encoded proteins.

These discoveries (summarized in a book; 21) led to our conclusion that the genome is an *integrated ensemble*, with little or no room left for *junk* (22) or *selfish DNA* (23, 24). This is a completely new vision of the vertebrate (and more generally of the eukaryotic) genome, which has far-reaching implications. Indeed, (i) there is no way to create a compositionally compartmentalized genome, the mosaic of isochores, by random point mutations (namely, single base-pair changes); (ii) again no random process can lead to a genome phenotype or compositional landscape that is correlated with all basic structural and functional properties of the genome, and lastly, (iii) no random evolutionary process can lead to the compositional correlations mentioned above. In other words, the discoveries just presented rule out the *bean-bag view* of the genome (to paraphrase Mayr, 25), namely of a genome in which genes are randomly distributed in the bulk of non-coding sequences, a genome that is only endowed with additive and not with cooperative properties (21).

Genome evolution and the neo-selectionist theory

The ground was now ready to investigate genome evolution. The simple comparison of our early data (26) on vertebrate genomes (that we recently confirmed on the basis of full genome sequences, 27-30); led us to the discovery of two modes of evolution: the *conservative mode* and the *transitional mode* (31). The *conservative mode* is exemplified by a comparison of the isochore patterns of the genomes of Primates and Carnivores (Fig. 5, see p. 605). At least 50% base pairs changed during the time, 100 million years, comprised between their common ancestor and these two mammalian orders that independently diverged from it. The expectation from the randomness of neutral changes was a partial or total disappearance of the isochore families that were present in the common ancestor. Moreover, since nucleotide substitutions in vertebrates (and other organisms) favor GC→AT over AT→GC changes, this 'AT-bias' should also lead to lower GC levels. Instead, a remarkable conservation of isochore families was found in terms of GC levels and relative amounts.

This led us in a straightforward way to the *neo-selectionist theory* (32). As shown in Fig. 6 (see p. 606), this theory postulates a series of steps: (i) first of all, among AT-biased changes a number will accumulate to form local clusters; (ii) the 'last' AT-biased changes in the clusters, the *critical changes* transform clustered point mutations into regional changes that trespass a lower GC threshold; and (iii) cause changes in chromatin structure that expand over long distances. Fig. 1 (see p. 601) shows that the neo-selectionist theory incorporates the features of the nearly neutral theory of Ohta, adding, as a novelty, the critical changes, namely the superdeleterious changes that convert the clustered AT-biased point mutations into regional changes. It should be stressed that regional changes may also be caused by large insertions and deletions. The main point, however, is that chromatin changes are deleterious in that they affect some expression of genes located within the altered regions or in their neighborhood and may lead to negative selection of the carriers and of their progeny.

Since fish, amphibian and many reptilian genomes do not show the presence of the very GC-rich isochores that characterize the genomes of warm-blooded vertebrates (see Fig. 5, p. 605), a *transitional mode* of evolution in which isochore families underwent changes, must have taken place (see Fig. 7, p. 607). Back in 1986 we proposed (20) that: 'The formation and maintenance of the GC-rich isochores of warm-blooded vertebrates is due to natural selection, the selective advantages being the increased thermodynamic stability of DNA, RNA and proteins (GC-rich codons encoding aminoacids that stabilize proteins). In other words, the environment can mould the genome through natural selection'. The transitional mode involved both negative and positive selection, as discussed elsewhere (32).

An explanation as to why changes essentially affected the gene-rich isochores, is that these isochores are located (in the interphase nucleus) in an open chromatin structure, whereas the gene poor isochores are in a closed chromatin structure (33). Then, only the genome core needs to be stabilized by GC increases, the genome desert being stabilized by its own compact chromatin. While body temperature certainly is the *primum movens* of the compositional transitions that took place at the emergence of mammals and birds, other factors such as oxygen, salinity, pH, CO_2, may play a role in the compositional transitions which were found among fishes (see Fig. 5, p. 605).

To sum up, the *neo-selectionist theory* (i) provides a solution to the neutralist/selectionist debate, since it reconciles the nearly neutralist view of point mutations with selection at the regional level; (ii) is an epigenomic theory, in that the compositional changes in DNA affect chromatin struc-

ture and, as a consequence, gene expression, so leading to negative selection of the carriers and their progeny; and (iii) is an extension of Darwin's theory; in fact, the neo-selectionist theory may be visualized as an ultra-darwinian theory since even neutral and nearly neutral changes are eventually controlled by natural selection over evolutionary time. Needless to say, the neo-selectionist theory brings us back from Kimura's survival of the luckiest to Darwin's survival of the fittest (incidentally, a matter of satisfaction for somebody working at the Stazione Zoologica).

As any good theory, the neo-selectionist theory also made predictions: (i) that genome phenotype differences should be found in populations; and (ii) that some of them may affect the genomic fitness and cause genomic (not genetic) diseases (a typical one being cancer). The first prediction was confirmed by comparing two individual genomes: Venter's genome differs from the reference human genome because of a number of insertions and deletions that accumulate in GC-rich isochores (34). These may generate genomic diseases by affecting chromatin structure and, as a consequence, the expression of genes located within or next to altered regions, so reducing the genomic fitness of the carriers, without necessarily affecting the primary structure of coding and regulatory sequences.

Conclusions

We should now go back to our initial questions and see the answers that we can provide today. First of all, a currently accepted view is that in all likelihood the origin of life was not so much the *single chance event* visualized by Monod, as a necessity under the prevailing conditions (35). This establishes *a primordial link* between the inanimate world from which life arose and the living organisms. These are connected to each other by their common descent, and, far from being completely separated from the inanimate environment, are moulded by it through natural selection. In fact, we have shown that the genome itself is moulded by physical agents like temperature, oxygen, salinity, pH, etc. through natural selection.

Our findings lead to a largely deterministic vision of evolution, which is in contrast with the fully stochastic vision of Monod. Chance still plays a role in evolution through (i) *environmental chance events*, such as meteorite impacts, volcanic eruptions; (ii) *random drift,* the random changes in gene frequencies in populations*;* and (iii) *neutral and nearly neutral changes*; as in the case of random drift, these changes are evident when recent, or looked at on a limited time scale, but they vanish over longer time spans,

because they are eliminated by natural selection. Obviously, we are very far from the overwhelming role of chance postulated by Monod.

As a consequence, we are also very far from Monod's view on the ethical implications. Given his premises, Monod claimed that *true knowledge ignores values* and invoked an *ethics of knowledge,* whose only value is the objective knowledge itself. In contrast, knowledge contains values: knowledge of common descent of all living organisms links us with them and dictates our respect and love for them; knowledge of the moulding of living organisms by the environment, trough natural selection, links all of them to the inanimate world from which they derived in the first place. The *'old alliance'* with Nature, proposed by the *'animistic conception'*, far from being *'a projection of our brain on the inanimate world'* (as suggested by Monod), is the age-old intuition of links now established by Science.

I would like to finish as I started, on a personal note. I had the good luck of being acquainted with Jacques Monod over many years until his premature death in 1976. My admiration for him led me to change the name of the Institut de Biologie Moleculaire that I was directing in Paris to Institut Jacques Monod, as well as to organize several meetings in his memory (see, for instance, ref. 36). I would like to stress that the contrasting vision presented here was built on the scene set up by *Le hasard et la necéssité*, I could say on the shoulders of Jacques Monod. It is a great pity that we cannot have his viewpoint on our conclusions. I dare say, however, that he would have accepted them, based as they are on new facts, which were not available or conceivable at the time his book was published. I also venture to guess that he would have liked them, since one can feel that the pessimistic conclusions of the book were imposed by its internal logics but not necessarily liked by its author.

REFERENCES

1. Darwin C. (1859) *On the Origin of Species, a Facsimile of the First Edition* (Harvard Univ Press, Cambridge, MA); reprinted (1964).
2. Monod J. (1970) *Le hasard et la nécessité* (Editions du Seuil, Paris).
3. Eigen M. (1992) *Steps towards Life* (Oxford University Press, Oxford).
4. Crick F. (1988) *What Mad Pursuit* (Basic Books, New York).
5. Fisher R.A. (1930) *The Genetical Theory of Natural Selection* (Clarendon, Oxford).
6. Haldane J.B.S. (1932) *The Causes of Evolution* (Longmans, Green, and Co., London).

7. Kimura M. (1968) *Nature* 217:624-625.
8. Kimura M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge University Press, Cambridge).
9. Kimura M. (1986) *Philos Trans R Soc Lond* (Biol) 312:343-354.
10. King J.L., Jukes T.H. (1969) *Science* 164:788-798.
11. Ohta T. (1973) *Nature* 246:96-98.
12. Ohta T. (2002) *Proc Nat Acad Sci USA* 99:16134-16137.
13. Zuckerkandl E. and Pauling L. (1962) *Molecular disease, evolution, and genetic heterogeneity* in Kasha M. and Pullman B., eds. *Horizons in Biochemistry* 189-225 (Academic Press, New York).
14. Bernardi G., Champagne M., Sadron C. (1960) *Nature* 188:228-229.
15. Bernardi G. (1965) *Nature* 206:779-783.
16. Corneo G., Ginelli E., Soave C., Bernardi G. (1968) *Biochemistry* 7:4373-4379.
17. Filipski J., Thiery J.P., Bernardi G. (1973) *J Mol Biol* 80:177-197.
18. Macaya G., Thiery J.P., Bernardi G. (1976) *J Mol Biol* 108:237-254.
19. Bernardi G., Olofsson B., Filipski J., Zerial M., Salinas J., Cuny G., Meunier-Rotival M., Rodier F. (1985) *Science* 228:953-958.
20. Bernardi G., Bernardi G. (1986) *J Mol Evol* 24:1-11.
21. Bernardi G. (2004) *Structural and Evolutionary Genomics, Natural Selection in Genome Evolution* (Elsevier, Amsterdam); reprinted (2005).
22. Ohno S. (1972) *Brookhaven Symp. Biol.* 23:366-370.
23. Orgel L.E. and Crick F. (1980) *Nature* 284:604-607.
24. Doolittle W.F. and Sapienza C. (1980) *Nature* 284:601-603.
25. Mayr E. (1976) See letter.
26. Thiery J.P., Macaya G., Bernardi G. (1976) *J Mol Biol* 108:219-235.
27. Costantini M., Clay O., Auletta F., Bernardi G. (2006) *Genome Res.* 16:536-541.
28. Costantini M., Auletta F., Bernardi G. (2007) *Genomics* 90:364-371.
29. Costantini M., Di Filippo M., Auletta F., Bernardi G. (2007) *Gene* 400:9-15.
30. Costantini M., Cammarano R., Bernardi G. (submitted).
31. Bernardi G. (1995) *Annu Rev Genet* 29:445-476.
32. Bernardi G. (2007) *PNAS* 104:8385-8390.
33. Saccone S., Federico C., Bernardi G. (2002) *Gene* 300:169-178.
34. Costantini M., Bernardi G. (submitted).
35. de Duve C. (1995) *Vital Dust* (Basic Books, New York).
36. Quagliariello E., Bernardi G., Ullmann A., eds. (1987) *From enzyme adaptation to natural philosophy: heritage from Jacques Monod* (Elsevier, Amsterdam).

DISCUSSION ON PROF. BERNARDI'S PAPER

PROF. PHILLIPS: My question is about the conclusions, and it seems to me that conclusion n. 2 is the one that is most germane to the topic of your talk. So the conclusion seems quite remarkable. If I am understanding it correctly, you are saying that chance does not play much of a role in evolution. Is that your conclusion?

PROF. BERNARDI: No, I said that chance, of course, works, but is under control. Of course I do not pretend that point mutations are not random mutations initially. What I am saying is that, after a certain evolutionary time, when changes accumulate and cluster together changes also occur in chromatin structure and negative selection follows.

PROF. PHILLIPS: So basically your conclusion is, in a sense, reaffirming a kind of Darwinian…

PROF. BERNARDI: Absolutely. That is why I mentioned that I come from an institute that was set up to show that Darwin was right.

PROF. COLLINS: So, Giorgio, the focus on isochores clearly is an interesting way to look at DNA but I think also maybe potentially blurs out the details in the sense that evolution probably does not care too much about isochores, it cares about genes, it cares about selectable enterprises, so can you explain why it is useful, at this juncture, when we have complete sequences, to look at isochores as a specific element of the genome as opposed to drilling down to the more refined level.

PROF. BERNARDI: This is an excellent question. It would be totally unreasonable to deny that the changes in coding sequences as having an important effect. What I am saying, however, is that in the genomes of the kind we are considering now, where the amount of intergenic and intronic noncoding sequences represent 99% of the genome, there are effects which

have been neglected so far and which in fact have an importance as far as gene expression in concerned. I was saying that it is perfectly conceivable to have a coding sequence which is untouched, but the region next to it has been altered in terms of chromatin structure and this creates a problem in gene expression. Therefore, it is not obligatory to find a change in the coding sequence. Even if this is the most common case probably, there may be, situations in which things happen outside the coding sequence in the so-called non coding sequences which, as you know better than I do, there is an accumulation of sequences which, in fact, matter more than we thought even ten years ago.

BACTERIAL EVOLUTION: RANDOM OR SELECTIVE?

RAFAEL VICUÑA

INTRODUCTION

The publication of *The Origin of Species* by Charles Darwin in 1859 constitutes a fundamental milestone in the history of science. In this book, Darwin builds up his theory of evolution based on the objective statements that living organisms change, that changes are transmitted to the progeny and that reproduction of organisms frequently gives rise to progenies that are too numerous to permit the survival of all the individuals. Darwin then concludes that in general, those individuals that change in such a way that their fitness to the environment increases will have a better chance to survive and reproduce. Thus, variations that are beneficial will gradually accumulate by simple natural selection.

What struck most the world at large was not the realization that living organisms evolve; after all, a transformist theory had been advanced four decades earlier by the French naturalist Jean Baptiste Lamarck in his *Histoire naturelle des animaux sans vertèbres* but the substantial differences between the theories advanced by both scientists. According to Lamarck, during their lifetime organisms undergo changes that favor their adaptation to the environment. These changes, which are influenced by the environment, are then transmitted to the offspring. Lamarck also stated that the evolutionary paths of the different species are independent of each other and that evolution follows a natural path towards perfection. In contrast, Darwin proposed that there is no such tendency to perfection. Rather, variation of living organisms is gradual, passive, spontaneous, with no destination. Favorable traits would be transmitted through the progeny, whereas those that are detrimental would tend to disappear. Moreover, in sharp antagonism with Lamarck, Darwin proposed the theory of common descent.

That variation (or mutation, as we call it now) arises spontaneously with no influence from the environment and without regard for utility has been elegantly shown by Luria & Delbruck[1] and by Lederberg & Lederberg,[2] in studies that are considered classic contributions to the field of molecular genetics. What these authors described correspond to mutations that are said to be growth-dependent, because they exhibit a definable relationship to cell division and are considered to result from random errors of the DNA replication machinery.[3] Does this undeniable fact imply that there is no variation promoted by the environment, as Lamarck had put forward? For a long time, growth dependent mutations were considered to be the primary cause of Darwinian evolution and even today it is so portrayed in the non specialized literature. However, some decades ago, researchers began to observe mutations that arise in non-growing, nutritionally deprived bacterial cultures that were subjected to non lethal selective pressure. Unexpectedly, these mutations appeared to have arisen with certain specificity in order to allow a better adaptation to the stressful environment.

Studies at the molecular level later showed that the mechanisms implicated in adaptive genetic change offer a much higher versatility of variation than the sole growth-dependent mutations attributed to errors of the DNA replication machinery. Although any one would think that most mutations are expected to be detrimental, an increase in variation is needed to allow some members of the population to arrive at a phenotype suitable for survival and proliferation in the new environment.

A CHALLENGE TO RANDOM AND GRADUAL MUTABILITY

The first hint of mutations in non-growing cells was obtained by Ryan about fifty years ago.[4] He observed that cultures of *his*$^-$ *Escherichia coli* auxotrophs inoculated into medium lacking histidine continued to produce

[1] Luria, S., Delbruck, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 28, 491-511, 1943.

[2] Lederberg, J., Lederberg, E.M. Replica plating and indirect selection of bacterial mutants. *J Bact* 63, 399-406, 1952

[3] In this case, the term random is used in a loose way, since geneticists are well aware that an average genome possesses hot spots for spontaneous mutations.

[4] Ryan, F.J., Wainwright, L.K. Nuclear segregation and the growth of clones of spontaneous mutants of bacteria. *J. Gen. Microbiol*. 11, 364-379, 1954.

his$^+$ revertants during a period of ten days after inoculation. He did not investigate whether other mutations also occurred, although he confirmed that the revertants were not slowly growing mutants previously present in the inoculum. A couple of decades later, Hall & Clarke found that a deletion mutant in the *lacZ* gene encoding β-galactosidase, when incubated for several days in the presence of lactose, reverted to a phenotype that allowed metabolism of this sugar.[5] The *lacZ* gene encodes an enzyme called β-galactosidase, which breaks down lactose into its components glucose and galactose. The new phenotype was the result of two mutations in an operon called *ebg* (for evolved β-galactosidase), which specifies a second β-galactosidase of unknown function. A mutation in the gene *ebgA* activates the enzyme, whereas a second mutation in the gene *ebgR* inactivates the repressor of the operon. Considering that either of the single mutations did not represent any advantage to the cell, it is remarkable that the double mutants arose with a frequency much higher than expected. Later, in separate studies, Shapiro[6] and Cairns *et al.*[7] investigated reversion rates in *E. coli* cells with bacteriophage Mu inserted into a fusion between the regulatory segment of the arabinose operon and the *lacZ* gene.[8] In this system, excision of Mu prophage led to fusion of *araB* to *lacZ*, yielding a Lac$^+$ cell as long as arabinose was also present to act as an inducer. The evidence showed that incubation for several days in both sugars, but not in either of them alone, led to the appearance of colonies in which Mu had been excised, whereas cultures grown without starvation produced none.

Other examples followed. Benson incubated bacteria in medium containing maltodextrins as the only carbon source. Normally, these high molecular weight polymeric substances do not trespass the cell membrane. However, bacteria underwent mutations in the gene encoding an outer membrane porin that allowed their ready entry into the cell.[9] In

[5] Hall, B.G. and Clarke, N.D. Regulation of newly evolved enzymes. III. Evolution of the *ebg* repressor during selection for enhanced activity. *Genetics* 85, 193-201, 1977.

[6] Shapiro, J. Observations on the formation of clones containing araB-lacZ cistron fusions. *Mol Gen Genet* 194, 79-90, 1984.

[7] Cairns, J., Overbaugh, J., Miller, S. The origin of mutants. *Nature* 335, 142-145, 1988.

[8] Both the arabinose and lactose operons are missing in this strain, which therefore is ara- and lac-. However, upon deletion of the intervening Mu prophage, it can grow on lactose provided arabinose is present.

[9] Benson, S.A., Partridge, L., Miller, S. Is bacterial evolution random or selective? *Nature* 336, 21-22, 1988.

turn, Hall pursued his work analyzing other systems. One of them required double mutations for utilization of β-glycosides, namely, excision of an insertion sequence and a point mutation.[10] Incubation in solid medium, only when containing substrate, promoted both mutations allowing its metabolism. Hall also tested for the first time the production of mutations in anabolic genes.[11] Two *E. coli* strains, each one possessing single point missense mutations in genes encoding enzymes for the synthesis of tryptophan (the *trp* operon), exhibited elevated reversion frequencies during starvation of this amino acid. Reversions in the *trp* operon did not take place when cells were starved for cysteine and mutation rates in other loci did not increase during tryptophan starvation. Therefore, the increased reversion rate appeared to be specific to conditions where the mutations were advantageous. In a subsequent study,[12] the author showed that a strain carrying two missense mutations in the *trp* operon reverts 10^8 times more frequently than would be expected if the two mutations were the result of independent events.

At the same time of the latter studies by Hall, one of the most paradigmatic papers in the field was published by Cairns and Foster.[13] These authors measured the reversion of a frameshift rather than a point mutation in the Lac operon of *E. coli*, which in this case is carried in an F' conjugative plasmid. The strain, called FC40, is deleted for the Lac operon on its chromosome and at the same time is resistant to the RNA polymerase inhibitor rifampicin due to a mutation in the chromosomal *rpoB* gene. The mutants were found to vigorously revert to Lac^+ (about one revertant per 10^7 cells per day) when plated on lactose minimal medium, whereas no reversion to the wild type rifampicin resistance phenotype was observed. Conspicuously, Lac^+ mutants did not arise in the absence of selection, i.e., when lactose was not present in the medium.

One of the most striking features of these early studies was that the increased frequencies of the advantageous mutations were not accompa-

[10] Hall, B.G. Adaptive evolution that requires multiple spontaneous mutations. I. Mutations involving an insertion sequence. *Genetics* 120, 887-897, 1988.

[11] Hall, B.G. Spontaneous point mutations that occur more often when advantageous than when neutral. *Genetics* 126, 5-16, 1990.

[12] Hall, B.G. Adaptive evolution that requires multiple simultaneous mutations: mutations involving base substitutions. *Proc Natl Acad Sci USA* 88, 5882-5886, 1991.

[13] Cairns, J., Foster, P.L. Adaptive reversion of a frameshift mutation in *Escherichia coli*. *Genetics* 128, 695-701, 1991.

nied by mutations at other loci.[14] This apparent selectivity, in open contradiction with the prevalent doctrine of randomness, astounded researchers in the field. For example, Cairns *et al.*[7] dared to state that 'In this paper...we describe some experiments suggesting that cells may have mechanisms for choosing which mutations will occur'. Also: 'This experiment suggests that populations of bacteria in stationary phase have some way of producing (or selectively retaining) only the most appropriate mutations'. Cairns even proposed molecular processes that 'could, in effect, provide a mechanism for the inheritance of acquired characteristics'. One of them was completely ground-breaking, since it implied information transfer from protein to DNA. According to this model, a reverse transcriptase instructed by *some element* that monitors the protein products would retrotranscribe an mRNA variant encoding a useful protein. Cairns referred to these mutations as adaptive,[13] while they were called 'directed' mutations by the editors of *Nature*[15] and 'a unicorn in the garden' by Franklin W. Stahl.[16]

Undoubtedly, this idea challenged the traditional thinking about spontaneous mutation, although the possibility of non randomness in variation had never been completely abandoned. In fact, Delbruck himself had previously noted the distinction between selecting for phage resistance versus selecting for carbohydrate utilization, stating that 'in view of our ignorance of the causes and mechanisms of mutations, one should keep in mind the possible occurrence of specifically induced adaptive mutations'.[17] A. Weismann, the father of neo-Darwinism, stated late in his career that directed variation must be invoked to understand some phenomena, as random variation and selection alone are not sufficient explanation. In turn, the

[14] Typically, in these studies, a mutant bacterial strain that requires a nutrient is plated on solid medium that contains a very limiting supply of the nutrient. When the nutrient is exhausted, there is a sparse population of bacteria on the agar and further growth cannot occur unless a known mutation reverts. The first observable colonies are considered to be spontaneous mutants that were present in the population prior to plating. Further incubation of the plates for several days up to a month reveals the continuous appearance of new colonies in numbers that cannot be predicted by the Luria&Delbruck test. These late appearing colonies that arise in a non-growing population of bacteria that are subjected to a nutritional stress are said to result from adaptive mutation.

[15] Cited in Foster, P.L. Adaptive mutations: Has the unicorn landed? *Genetics* 148, 1453-1459, 1998.

[16] Stahl, F.W. A unicorn in the garden. *Nature* 335, 112-113, 1988.

[17] Delbruck, M. Heredity and variations in microorganisms. *Cold Spring Harbor Symp. Quant. Biol*. 11, 154 – , 1946.

eminent geneticist T. Dobzhansky expressed by mid 20th century that 'The most serious objection to the modern theory of evolution is that since mutations occur by chance and are undirected, it is difficult to see how mutation and selection can add up to the formation of such beautifully balanced organs as, for example, the human eye'.[18] Interestingly, in a speculative paper published earlier than Cairns' work, Fitch had stated that 'because mutations are advantageous during stressful times but genome wide mutagenesis would be deleterious, organisms probably have evolved a mechanism for selectively mutating only the genes of relevance'.[19]

As expected, the possibility that certain mutations in bacteria that were in stationary phase and subjected to non-lethal selective pressure might occur at higher rates when advantageous gave rise to a deep controversy.[20] This new type of mutation that came into sight more often when beneficial than when neutral appeared to vindicate the Lamarckian idea that the environment influences variation to improve adaptation. In this case, however, changes would obviously not occur as a result of use or disuse of a particular organ. Instead, they might perhaps arise from selection based on the presence of molecular variations within cells. On the other hand, one of the main arguments used by the supporters of adaptive mutations was that the classical experiment of Luria & Delbruck could not show the appearance of mutations during selection, since their protocol involved a lethal selection assay (resistance to bacteriophage T1). This assay gave no chance to detect additional mutations in cells that had not become resistant to viral infection.

Are adaptive mutations really directed?

The first hint that there was not reverse information flow that would instruct the cell how to mutate to attain successful survival was obtained by a reversion of an amber mutation in an episomal *lacZ* gene, both through

[18] Quotations by Weismann and Dobzhansky taken from: Wright, B.E. A biochemical mechanism for nonrandom mutations and evolution. *J Bact* 182, 2993-3001, 2000.

[19] Fitch, W.M. The challenges to Darwinism since the last centennial and the impact of molecular studies. *Evolution* 36, 1133-1143, 1982.

[20] See for example letters by several scientists and rebuttals in *Nature* 336, 21-22, 1988; *Nature* 336, 525-528, 1988 and *Science* 269, 285-289, 1995. Also: Lenski, R.E., Slatkin, M., Ayala, F.J. Mutation and selection in bacterial populations: alternatives to the hypothesis of directed mutation. *Proc. Natl. Acad. Sci. USA* 86, 2775-2778, 1989; Lenski, R.E., Mittler, J.E. The directed mutation controversy and neo-Darwinism. *Science* 259, 188-194, 1993.

intragenic mutations that eliminate the stop codon and by extragenic creation of a tRNA suppressor.[21] The latter necessarily had to be random, since there was no relationship between lactose metabolism and a chromosomal gene encoding a tRNA. In a subsequent study, Foster tested the mutability of a second gene (tet^S) also present in the plasmid harboring the *lacZ* gene mutant. She found that upon selection in lactose, tet^R mutants appeared at about the same rate as Lac^+ mutations.[22] These results showed clearly that selection was unnecessary for obtaining mutations in stationary phase, as originally thought. The concept of adaptive mutation was hence adjusted to mean those mutations that occur in non dividing cells during selection and are specific to the selective pressure. Mutants that arise in non dividing cells and that are either not adaptive, or have not yet been shown to be adaptive, were called stationary phase mutations.[23] Later, other Lac^+ revertants of the *E. coli* strain FC40 were found to carry mutations that were not related to selection.[24,25]

In turn, specificity of reversion of trp^- mutants was shown by the lack of reversion in cultures starved for other amino acids, as well as by the lack of appearance of other mutants during starvation for tryptophan. Out of 110 trp^+ revertants, Hall found only two carrying additional mutations.[26] However, he was somewhat cautious in the interpretation of these results: he stated that the explanation for the apparent influence of the environment in the selectivity of mutation did not necessarily have to be found in the two extreme choices that had been so far considered, namely randomness or directedness. He proposed to adopt the concept of 'Cairnsian' mutation to imply those sequence changes that occur with a higher probability when they are advantageous than when they are neutral. Later, citing a personal communication by J. Cairns, he speculated that the specificity could be

[21] Foster, P.L, Cairns, J. Mechanisms of directed mutation. *Genetics* 131, 783-789, 1992.

[22] Foster, P.L. Nonadaptive mutations occur on the F' episome during adaptive mutation conditions in *Escherichia coli*. *J Bact* 179, 1550-1554, 1997.

[23] Foster, P.L. Adaptive mutation: the uses of adversity. *Ann Rev Microbiol* 47, 467-504, 1993.

[24] Rosche, W.A., Foster, P.L. The role of transient hypermutators in adaptive mutation in *Escherichia coli*. *Proc Natl Acad Sci USA* 96, 6862-6867, 1999.

[25] Torkelson, J., Harris, R.S., Lombardo, M.J., Nagendran, J., Thulin, C., Rosenberg, S.M. Genome-wide hypermutation in a subpopulation of stationary cells underlies recombination-dependent adaptive mutation. *EMBO J* 16, 3303-3311, 1997.

[26] Hall, B.G. Spontaneous point mutations that occur more often when advantageous than when neutral. *Genetics* 126, 5-16, 1990.

explained by either selective capture or selective generation.[27] The former mechanism implies that mutations take place randomly and continuously during prolonged selection, but only those that are useful are captured by replication or recombination and immortalized by growth. Useless mutations have no way to express themselves. Selective generation, on the other hand, implies that sequence changes occur only in genes that are being actively transcribed. Indeed, one likely mechanism for directing mutations to specific genes requires their active transcription under nutritional deprivation (see below).

Systems involving mobile genetic elements represent a different situation. In the case of prophage Mu excision from the *araB-lacZ* fusion to allow growth on lactose when arabinose is also present,[6,7] the specificity of genetic variation is obvious. In the *egb* operon, it has been established that the gene *ebgR* encoding the repressor is a hot spot for the insertion of the mobile element IS30, whereas in the *bgl* operon the gene *bglF* reverts to wild type by excision of IS103. The latter event precedes mutations in the promoter (*bglR*), which will eventually allow growth in β-glycosides. In either of these situations, where movement of the mobile elements is stimulated by stress (see below), directedness could be explained by selective capture.

In spite of these clarifications, the controversy regarding the directedness of mutations followed for several years.[28] Even recently, Roth *et al.*[29] have been particularly critical in accepting that selection stimulates formation of new mutations. These authors prefer to think that what selection actually does is to allow faster growth of pre-existing mutants, with the parent strain remaining unable to grow due to the stringent conditions of the medium. However, the recent unraveling at the molecular level of several mechanisms involved in stress induced mutagenesis seems to leave no room for a controversy. It is now understood beyond doubt that stressful environments induce in bacteria genomic instability which results in mutants that are fitter than the parent strain to the adverse conditions.

[27] Hall, B.G. Adaptive mutagenesis: a process that generates almost exclusively beneficial mutations. *Genetica* 102/103, 109-125, 1998.

[28] See for example the series of papers by Rosemberg & Hastings, Ross & Andersson and Foster, with the corresponding rebuttals, in *J Bact* 186, 4838-4863, 2004.

[29] Roth, J.R., Kugelberg, E., Reams, A.B., Kofoid, E., Andersson, D.I. Origin of mutations under selection: The adaptive mutation controversy. *Annu Rev Microbiol* 60, 477-501, 2006.

A stressful environment induces adaptive mutations

Cells have different DNA repair pathways that are responsible for correcting sporadic mistakes arising as a result of DNA polymerase errors or through chemical modification of the bases. Therefore, mutations in the DNA are supposed to be transient, because they are normally corrected. However, under stressful conditions, these repair pathways are either down-regulated or become overwhelmed while taking care of abundant DNA damage.

There are several stress responses that intensify genetic variation in bacteria.[30,31] As mentioned previously, the molecular mechanisms leading to mutations in these pathways are different from those taking place in growing cells. All the previous findings of adaptation in non-growing cultures can now be interpreted under the light of one of these mutagenic pathways. In some cases, they may give rise to localized sequence changes, which have the advantage of avoiding non-adaptive mutations. The apparent selectivity observed in some of the laboratory studies may explain the original interpretation of directedness.

Perhaps the most thoroughly studied mutagenic pathway is the SOS response.[32] It is induced by extensive DNA damage, by cell saturation in rich medium, exposure to antibiotics and in aging colonies. About 30 genes encoding functions related to DNA metabolism are under the control of LexA repressor. Among them are those specifying DNA polymerases IV (*dinB*) and V (*umuC,D*), which are able to replicate damaged DNA although with low fidelity. Normally, the genes of the pathway are silent or are expressed at very low levels. The SOS response is triggered when the stressful environment induces RecA-dependent auto-proteolysis of LexA. If cells are proliferating, the two error prone polymerases increase the mutation rate by competing with the accurate DNA polymerase III, which replicates the chromosome under normal conditions. In non-growing cells, partial DNA synthesis by the mutagenic enzymes takes place during repair or recombination events. Some of the mutants arising will have a selective advantage for survival.

[30] Foster, P.L. Stress responses and genetic variation in bacteria. *Mutation Res* 569, 3-11, 2005.

[31] Foster, P.L. Stress-induced mutagenesis in bacteria. *Crit. Rev. Biochem. Mol. Biol.* 42, 373-397, 2007.

[32] Schlacher, K., Goodman, M.F. Lessons from 50 years of SOS DNA damage induced mutagenesis. *Nature Rev Mol Cell Bio* 8, 587-594, 2007.

Another important pathway is the general stress response.[33] In this case, the controller protein is not LexA but RpoS, a sigma factor (σ^S) that replaces the vegetative sigma factor σ^{70} of RNA polymerase. Sigma factors are critical for gene expression, since they are responsible for the selectivity of transcription by RNA polymerase. Nutrient limitation or stationary phase of growth results in the accumulation of polyphosphate (PolyP). This compound causes an elevation in the titers of σ^{70}, leading to higher levels of the error-prone DNA polymerase IV or to an inhibition of the expression of enzymes belonging to the mismatch repair (MMR) pathway. Both effects contribute to raise the adaptive mutation rate in bacteria.

Amino acid starvation also causes the buildup of (p)ppGpp, a phenomenon commonly known as the stringent response.[34] This rare nucleotide inhibits initiation of DNA replication and influences the selectivity of transcription by RNA polymerase. For example, it down regulates the synthesis of rRNAs and tRNAs while it also collaborates in raising the levels of RpoS. In addition, (p)ppGpp up regulates the operons for amino acid biosynthesis, which are normally subjected to end-product repression. It is well known that genes under transcription are more liable to mutate due to their partial single stranded character.[35] Thus, starvation for a specific amino acid makes its synthetic operon more susceptible to mutations. This may be the explanation for the 'directedness' observed by Hall in the reversion of the *trp* mutants.[11,12] DNA damage, starvation and high temperature (heat shock) also trigger a stress response dependent on a sigma factor called RpoH (σ^{32}). Among the genes controlled by σ^{32} is one that encodes GroE. This is a molecular chaperone that interacts with DNA polymerases IV and V (among many other proteins), protecting them from degradation by proteases and thus increasing mutagenesis.

There are three other mutagenic stress responses that are less well characterized. Two of them are specific for bacteria growing on solid media. One is called ROSE, an acronym for 'resting organisms in a structured environment'.[36] ROSE requires RecA and DNA polymerase I and it is independ-

[33] Hengge-Aronis, R. Signal transduction and regulatory mechanisms involved in control of the σ^S (RpoS) subunit of RNA polymerase. *Microbiol. Mol. Biol. Rev*. 66, 373-395, 2002.

[34] Braeken, K., Moris, M., Daniels, R., Vanderleyden, J., Muller-Hill, B., Michiels, J. New horizons for (p)ppGpp in bacterial and plant physiology. *Trends Microbiol*. 14, 45-54, 2006.

[35] Wright, BE. A biochemical mechanism for nonrandom mutations and evolution. *J. Bacteriol* 182, 2993-3001, 2000.

[36] Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P.H., Godelle, B. Role of mutator alleles in adaptive evolution. *Nature* 387, 700-702, 1997.

ent of DNA polymerase V and RpoS. Another one is called MAC ('mutagenesis in aging colonies') and it does not involve LexA, although it does require RpoS and DNA polymerase II.[37] A third response, the GASP phenotype[38] (growth advantage in stationary phase) relies on the SOS DNA polymerases II, IV and V and in an attenuated participation of RpoS. The GASP response allows survival of a small percentage of the bacterial population that consumes the debris of dying cells in long term batch cultures. Under these conditions, the birth and death rates are balanced. An increase in the mutation rate of cells in stationary phase is further supported by down regulation of the DNA repair pathways, some of which operate through intricate mechanisms that are highly energy consuming.[39]

The hypermutable state model

Hall proposed an additional argument to interpret the apparent directedness of adaptive mutations. It was what he called the hypermutable state model.[40] According to this model, although all non-growing bacterial cells in a selective medium are experiencing a stressful situation, only a minor subpopulation of them, perhaps between one in every 10^3 or 10^4 of cells enters a hypermutable state.[41] While in these circumstances, those bacteria that generate neutral or deleterious mutations die in a short time. However, if one of the mutations is a revertant that allows growth, the cell is relieved from the stress. It then proliferates exiting from the hypermutable state, building up just only growth-dependent mutations at a normal rate. Thus, the hypermutable state is transient. Eventually, the only cells that survive the stressful condition are those that never enter into the hypermutable state or those that do so and acquire a useful mutation. The fact that the frequency

[37] Bjedov, I., Tenaillon, O., Gerard, B., Souza, V., Denamur, E., Radman, M., Taddei, F., Matic, I. Stress-induced mutagenesis in bacteria. *Science* 300, 1404-1409, 2003.

[38] Finkel, S.E. Long term survival during stationary phase: evolution of the GASP phenotype. *Nature Rev. Microbiol*. 4, 113-120, 2006.

[39] Saint-Ruf, C., Pestut, J., Sopta, M., Matic, I. Causes and Consequences of DNA repair activity modulation during stationary phase in Escherichia coli. *Crit. Rev. Biochem. Molec. Biol.* 42, 259-270, 2007.

[40] Hall, B.G. Spontaneous point mutations that occur more often when they are advantageous than when they are neutral. *Genetics* 126, 5-16, 1990.

[41] Rosenberg, S.M. Evolving responsively: adaptive mutation. Nature Rev. *Genetics* 2, 504-515, 2001.

of mutations in selected revertants is notably higher than in the surviving cells that do not mutate the selected gene clearly satisfies the model.[42] It also adds evidence for selective capture rather than for selective generation.

The hypermutable state model has received support from Rosenberg's group.[43] According to these authors, the high mutation rate reaches its maximum with the coincident induction of the SOS and RpoS stress responses.

There are various mechanisms for adaptive mutations

Work in different laboratories has revealed that there are several ways by which bacteria can modify their genomes to relieve the selective pressure in a stressful environment. In other words, there are several types of adaptive mutations, each of them involving a molecular mechanism that sheds light into the seeming selectivity of mutation.

a) The episomal Lac system.[44]

As mentioned above, the *E. coli* FC40 strain carries a large conjugal plasmid which includes a fusion of the gene encoding the Lac repressor (*lacI*) with the *lacZ* gene encoding β-galactosidase. Therefore, it lacks the regulatory region of the operon and transcription starting from the promoter of *lacI* is constitutive. This construction is Lac^- because it carries a +1 frameshift in *lacI*, changing CCC to CCCC, although it is slightly leaky, conferring about 1% of wild type β-galactosidase level. The chromosome in the strain has a large deletion that encompasses the *lac* operon. When these cells are inoculated on solid minimal medium containing lactose as carbon source, colonies of Lac^+ mutants appear a few days later on the plate. In the absence of carbon source, Lac^+ mutations (as measured by subsequent plating on lactose) do not accumulate regardless the incubation time. Strain FC40 also reverts to Lac^+ during non-selected growth. In this case, mutations include duplication, deletions and large frameshifts,

[42] Drake, J.W. Too many mutants with multiple mutations. *Crit. Rev. Biochem. Mol. Biol*. 42, 247-258, 2007.

[43] Gallardo, R.S., Hastings, P.J., Rosenberg, S.M. Mutation as a stress response and the regulation of evolvability. *Crit. Rev. Biochem. Mol. Biol*. 42, 399-435, 2007.

[44] Foster, P.L. Stress-induced mutagenesis in bacteria. *Crit. Rev. Biochem. Mol. Biol*. 42, 373-397, 2007.

while mutations obtained during selection are almost exclusively -1 frameshifts. The latter are typically made by DNA polymerase IV (*dinB*), which is induced by the SOS and RpoS pathways. Adaptive mutations are severely reduced in GroE and polyphosphate kinase deficient cells, confirming the requirement for DNA polymerase IV. Under normal conditions, frameshift mutations are corrected by the mismatch repair system, which is insufficient or may be down regulated in stressed cells undergoing the transient hypermutation state. Mutants obtained under selection also differ from those arising during normal growth in that they require enzymes involved in the recombinational repair of double strand breaks, such as RecA, RecBCD and RevABC.

There are two models accounting for adaptive mutation in *E. coli* FC40 cells. One of them relies on the fact that the conjugal origin of the episome is subjected to continuous nicking. Occasional initiation of episomal replication at its vegetative origin is allowed by the energy provided by the leakiness of the Lac construction. Advancement of the replication fork towards the nick generates a double stranded break that is repaired by RecA, RecBCD and RuvABC recombination enzymes. Short patches of DNA synthesis required by this pathway are undertaken by the mutagenic DNA polymerase IV and by DNA polymerase II. This model accounts for the fact that the Lac construction needs to be in the episome in order to obtain adaptive revertants. A second mechanism leading to Lac^+ colonies of the FC40 strain consists in the 20-50 fold amplification of the *lac* locus.[45] These revertants appear somewhat later than the point mutants. Amplification does not require DNA polymerase IV or the other SOS-induced proteins, although it depends on RpoS, DNA polymerase I and the recombination proteins RecA, RecBCD and RuvABC. Interestingly, the amplified clones do not exhibit unrelated mutants as it is the case with the Lac^+ point mutants. Moreover, the Lac^+ phenotype of the amplified clones reverts to Lac^- upon re-plating in rich medium. Some investigators originally thought that amplification was an intermediate state in the formation of Lac^+ point mutants, but it was later shown that it consists on an alternative way to relieve the starvation stress by cells that never enter the hypermutation state.

[45] Hastings, P.J. Adaptive amplification. *Crit. Rev. Biochem. Mol. Biol*. 42, 285-311, 2007.

b) The transcription-dependent revertants of *trp* auxotrophs.

Amino acid starvation triggers the stringent response, which, as mentioned previously, up-regulates transcription of operons for amino acid biosynthesis. It has now been well established that transcription during prolonged starvation is mutagenic. The reason for this effect is that nucleotide bases are prone to undergo chemical modifications when present in single stranded DNA. For example, cytosine deaminates to uracil, which upon DNA replication, preferentially pairs with adenine instead of guanine. In turn, adenine spontaneously deaminates to hypoxanthine, which hydrogen bonds to cytosine rather than to thymine. In cells where the mismatch repair system is down regulated, these modifications remain in the DNA sequence.

Transcription generates localized single stranded structures in two ways.[36] One is the formation of a transcription bubble, where the DNA-RNA hybrid structure exposes the nontranscribed strand leaving it vulnerable to change. The other one, related to the negative supercoiling generated behind the transcription bubble, gives rise to stem-loop structures possessing susceptible unpaired bases. Since starvation for a particular amino acid specifically targets derepression of the corresponding operon, it is most likely that the adaptive missense mutations in the *trp* operon in Hall's studies are generated during transcription of this operon. This mechanism is coherent with the observed directedness of the revertant mutations.

c) Systems involving mobile genetic elements.

As mentioned previously, some adaptive mutations require either excision or insertion of DNA elements. Normally, molecular events of this kind are under tight control to avoid deleterious effects in the genome. However, stressful environments promote movements of such sequences,[46] providing the cells with an additional strategy for adaptation. For example, numerous studies have demonstrated that starvation elicits an increase in transposition frequency of mobile elements, which may be mediated by the RpoS or SOS responses. In the long term, this type of genome flexibility contributes to increase the genetic diversity of microbial populations.

[46] Shapiro, J.A. Genome organization, natural genetic engineering and adaptive mutation. *Trends Genet.* 13, 98-104, 1997.

Adaptive mutation and evolution

In proliferating bacterial populations, survival depends on efficient DNA replication, which requires high speed and fidelity. In contrast, a hostile environment where cells cannot multiply will favor the selection of mutants that are able to overcome the episode of crisis.

The basic difference between random and adaptive mutations is that the latter are beneficial by definition, since they increase fitness. Moreover, it has been observed that when adaptation requires more than one mutation, the appearance of the first one makes more expedite the production of those that follow. There are now three examples understood that illustrate this behavior: reversion of the *trp* double mutants, expression of the *ebg* operon and double reversion of *bgl* operon, all of them studied in Hall's laboratory. In each of these cases, reversion of the first mutation allows very slow growth. Then, selection operates to single out the second mutation which leads to rapid growth. This is undoubtedly a fine course of action for adaptation. No wonder evolutionary biologist Douglas Futuyma, excited by Hall's work on the evolution of the *ebg* operon to permit lactose metabolism, wrote: 'Thus, an entire system of lactose metabolism has evolved, consisting of changes in enzyme structure enabling hydrolysis of the substrate; alteration of a regulatory gene so that the enzyme can be synthesized in response to the substrate and the evolution of an enzyme reaction that induces the permease needed for the entry of the substrate. One could not wish for a better demonstration of the neo-Darwinian principle that mutation and natural selection in concert are the source of complex adaptation'.[47]

Common sense tells that the ability to accelerate variation in the genome offers a selective advantage for survival in a changing environment. Several studies, both theoretical and experimental, have confirmed this assertion. In this context, the hypermutation state could be particularly fitting because it increases the probability of obtaining an advantageous mutation when the majority of the cells undergoing a normal mutation rate do not produce it. A fine regulation of the hypermutation state lessens the likelihood of accumulating undesirable mutations.[48] First, it is transient,

[47] Futuyma, D.J. Evolution (Sunderland, M.A.: Sinauer Associates), pp. 477-478, 1986, cited by Miller, K.R. in *Finding Darwin's God*. Perennial, Harper Collins Publishers 2002.

[48] Foster, P.L. Adaptive mutation: implications for evolution. *BioEssays* 22, 1067-1074, 2000.

i.e., when adaptation to the medium is achieved, a return to low mutation rates is selected for. But also, it is restricted to space, as it is clearly exemplified by mutations induced by double stranded breaks, transcription of defined operons and movement of genetic elements. In spite of the clear advantages of confining mutation in space and time, there are occasions in which adapted mutants maintain a mutator phenotype. This outcome is thought to result from adaptive mutations originated in strains with a mutator allele, a property that would be transmitted by hitchhiking in conjunction with the favorable alleles they produce.[49]

It would be very difficult to establish the precise contributions of growth dependent mutations, adaptive mutations and horizontal gene transfer to bacterial evolution. This problem could perhaps be approached experimentally, although laboratory studies are generally short term, whereas microorganisms in their natural environments confront long periods of starvation. Having this limitation in mind, it is worthwhile to highlight recent results obtained by Yeiser *et al.*[50] with bacteria struggling to survive in stationary phase. These investigators confirmed that SOS-induced DNA polymerases II, IV and V enhance long-term survival and evolutionary fitness of bacteria under stress. When grown individually, wild-type and SOS DNA polymerase mutants exhibit similar cell yields and stationary phase survival patterns. However, when the wild type and the mutant strains are co-cultured and must therefore compete for nutrients, SOS polymerase mutants undergo a marked reduction in fitness and fail to express the 'growth advantage in stationary phase phenotype' (GASP). Since DNA polymerase V is the most mutagenic, it is remarkable that mutants of this enzyme are the most affected in the competition experiments. According to these authors, DNA polymerase V may provide the mutational raw material for natural selection in a manner superficially similar to the increase fitness accompanying the absence of the mismatch repair system.

[49] Kivisaar, M. Stationary phase mutagénesis: mechanisms that accelerate adaptation of microbial populations under environmental stress. *Environ. Microbiol*. 5, 814-827, 2003.

[50] Yeiser, B., Pepper, E.D., Goodman, M.F., Finkel, S.E. SOS-induced DNA polymerases enhance long-term survival and evolutionary fitness. P*roc. Natl. Acad. Sci. USA* 99, 8737-8741, 2002.

Concluding remarks

Unraveling the adaptive mutation phenomenon has allowed us to become aware that the complexity of living organisms is not the outcome of a sole random mutational process, as it is most commonly regarded. Instead, it has become clear that throughout evolution there have also been adaptive mutations stimulated by a variety of fine feedback mechanisms. These include activation of error prone DNA polymerases, down-regulation of DNA repair enzymes, gene amplification, movement of mobile genetic elements, development of a transient hypermutation state in some cells, localization of mutations in genomic space to minimize deleterious mutants, various types of recombination events, etc.

It is most likely that these induced mutations have had a key role in determining bacterial evolution, since natural habitats are often stressful due to a lack of nutrients or some other unfriendly condition. There is still a third kind of gene variation that is widespread in the microbial world and has played a decisive role in bacterial evolution, namely, horizontal gene transfer. In spite of its importance, however, the description of this phenomenon goes beyond the scope of this essay.

DISCUSSION ON PROF. VICUÑA'S PAPER

PROF. WITTEN: Do all mutations happen at random or does the organism have a library of things that it tries, which would be useful when it runs into different environments?

PROF. VICUÑA: A library? No, I don't think so. What microorganisms do under stress is to induce DNA polymerates that make mistakes. This implies that there will be many trial and error events. Most will be lethal, whereas a few will be useful. As soon as a useful mutation arises, the bacterium will start to proliferate. Therefore, it is not that they have a library and they can select which mutation or which gene they should change. Does that answer your question?

PROF. COLLINS: This is a fascinating story and it does seem to add some new understanding of Darwinian selection in a new way, in a fashion that might be referred to – and this is a word that is beginning to, I think, even be accepted more in mammalian evolution – the concept of evolvability that natural selection not only operates on specific changes, provided that there is some selective pressure, but that it is to the advantage of an organism to have a capability of evolving in unpredictable ways if some new pressure, some new niche arises, which is to say that it is to the advantage of organisms not to be squeaky clean in the way in which they handle their biology but to be prepared to make mistakes, to be prepared to have some stuff lying around that you are not really using that might come in handy when some new pressure arises as may be the case, for instance, with a lot of the transcription in mammalian cells for which it is not clear there really is a scientific function, but maybe it is just there in case you wanted to tinker with it. Would you agree that that is sort of the conclusion from this?

PROF. VICUÑA: I think so, I agree with it and you must be aware, of course, that,after the sequencing of the human genome, several genes were

found to encode these error prone polymerates. There are five or even more of them. They have been well characterised. We now know that all organisms produce DNA polymerates that do not copy DNA with fidelity, which are induced when the cell is in trouble.

PROF. W. SINGER: Actually, retrospectively it makes a lot of sense that evolution has worked on its own evolutionary mechanism, how could it have done otherwise.

PROF. M. SINGER: I think, in a way, this is complementary to what Francis has just said. As I recall, in a paper a couple of years ago, Susan Lindquist showed, in yeast, that certain mutations that did not have obvious effects under normal circumstances, did so when the yeast was stressed. Proteins that are produced in response to stress, can assist, in ways I will not go in to, those mutated other genes to fill in for the newly emerged needs. Do you remember that paper, Francis? I think it was in *Science*. It was presented both with data and with a certain amount of speculation about this system as a new tool in evolution. In a way, it responds to Ed Witten's question.

PROF. VICUÑA: Yes, if I may add something, this adaptive response has also been studied in yeast and other eukaryotes and it has also been proposed by Cairns as a model for the development of tumours in mammalian tissues.

PROF. CAVALLI-SFORZA: There is evidence that there are strains of bacteria that are more mutable than others with a frequency of even a hundred times higher. Now, if that is the case, which I believe is true, then if a stressful environment is encountered or produced somehow the bacteria mutants that arise may come more likely from the mutable part so natural selection automatically will select for more mutable strains, so that is a way in which automatically there is a greater adaptability of the organism as long as the mutant is unstable. There are other situations in nature where that is clearly so, because another way of increasing genetic diversity is to introduce recombination, whether by sexuality or by other methods. Now, there are organisms that ordinarily can reproduce both ways. If the environment is stable they tend to stay in the asexual reproduction. Whenever the environment changes, they shift to the sexual phase. One of them is *Daphnia*, which is an aquatic organism. So it is clear that when the environment is hard or stressful then increasing genetic variation is favoured.

FROM MICROBIAL GENETICS TO MOLECULAR DARWINISM AND BEYOND

WERNER ARBER

When Charles Darwin reflected on the process of biological evolution some 150 years ago, he could not know of the existence of genes. But he and some of his contemporary natural scientists had observed that individual organisms belonging to a given species showed obvious phenotypical variations. Darwin's theory of evolution postulates that the different variants and their parents are steadily submitted to natural selection. This means that variants which can better deal with their encountered living conditions are favoured, and can, in the long term, overgrow less favoured organisms in the natural ecosystems.

Independently of evolutionary biology, Gregor Mendel initiated classical genetics in 1866, based on the observation that some phenotypical traits became transferred into the progeny and that recombinants could show up upon cross-fertility between individuals with different traits (i.e. mutants). The genes that were postulated to represent the determinants for phenotypical traits remained for many decades an abstract concept. This was still the case when around 1940 classical genetics and evolutionary biology joined forces in the so-called modern evolutionary synthesis that resulted in the Neo-Darwinism (Mayr, 1982).

About at the same time, microbiologists reported that bacteria and bacterial viruses could also undergo mutation and were thus postulated to possess genes. This opened the possibility to experimentally explore the formation of recombinants in mixed cultures of different bacterial or viral mutants.

Already in classical genetics there was good evidence that genes were associated with chromosomes. These were known to contain nucleic acids as well as proteins (chromatin). A strong evidence that desoxyribonucleic acid (DNA), rather than proteins, is the carrier of the postulated genes came in 1944, when Avery *et al*. (1944) reported their experiments with pneumo-

coccal bacteria. These authors worked with two different bacterial types that were distinct by their traits. DNA was extracted from one of the strains and carefully purified from all attached proteins. When this DNA fraction was incubated together with intact bacterial cells of the second strain, some bacteria showed up that had acquired the trait characteristic for the first bacterial strain: these bacteria had been transformed. No transformation could be obtained upon incubation of the second strain with the purified proteins of the first strain. DNA was thus concluded to be the carrier of genetic information.

How the genetic information can be inscribed in DNA became obvious several years later, when Watson and Crick (1953) showed that DNA molecules are long filaments with double-helical structure. Two antiparallel strands composed of four different nucleotides are held together by hydrogen bonds ensuring a specific base-pairing between the neighbour nucleotides. On the basis of this discovery it became obvious that genetic information can be contained in the linear sequences of nucleotides. In addition, correct base pairing was suggested as the criterium for the transmission of the genetic information from the parental genome to the two daughter genomes upon DNA replication.

In the following decades molecular genetics was developed thanks to experimental investigations with microorganisms and soon also with eukaryotic organisms. This then led to genomics including DNA sequence analysis and investigations on gene functions. The thereby acquired knowledge can provide insights into impacts that spontaneous alterations of nucleotide sequences can have on specific phenotypical traits. This kind of research turned out to be quite informative for an understanding of molecular mechanisms that generate genetic variations. The products of genetic variants together with their parental forms represent the substrate for natural selection. It is thus appropriate to join now forces between Neo-Darwinism and molecular genetics to result in Molecular Darwinism.

Bacterial genetics was developed with only a few kinds of bacterial strains, particularly with *Escherichia coli*. Under laboratory conditions these bacteria propagate exponentially with a generation time of about 30 minutes between two cell divisions. Large populations can thus be obtained in one day. Since bacteria are haploid, having just one set of genetic information, spontaneously occurring mutants become phenotypically manifested quite fast. One can observe that, when a few hundred growing cells are reached, one new mutation shows up. With available research strategies the nature of newly isolated mutants can readily be analysed, both with

regard to the suffered DNA sequence alterations and with regard to their functional capacities under various growth conditions.

As more and more gene and genome sequences of various prokaryotic and eukaryotic organisms become available, sequence comparisons can give hints on the molecular mechanisms that might have been at the origin of the observed functional differences. Such comparisons are also welcome for bacterial strains that cannot be propagated under laboratory conditions. Still, it is of importance to base any conclusions from sequence comparisons on available knowledge on already identified molecular mechanisms that are at the source of newly generated mutants. Experimental data of such events can best be acquired in work with genetically well known bacteria and bacterial viruses.

It is good to mention here that in the relevant genetic literature two different definitions are used for the term 'mutation'. In classical genetics, a mutant displays an altered phenotype that becomes transmitted to the progeny. In molecular genetics, looking at DNA sequences, a mutation is usually defined by carrying an altered nucleotide sequence. In the meantime, we know that, as a rule, classically defined mutants have also an altered nucleotide sequence. But we also know that by far not all nucleotide sequence alterations lead to an altered phenotypical trait. Many novel sequence alterations are, often for known reasons, silent or neutral and have no immediate influence on life processes. On the other hand, there is a good consensus between researchers in the field that relatively few novel mutations are favourable and provide to the organism a selective advantage. Much more often new mutations are unfavourable. They can inhibit life processes to some degree or, in extreme cases, they can be lethal. These kinds of mutations provide a selective disadvantage. This situation gives us no evidence for a directive nature of spontaneous mutagenesis. In general, the spontaneous generation of genetic variants reflects some kind of randomness; it is not a directing response to an identified specific need. We will explain below that several specific molecular mechanisms contribute to the overall generation of genetic variations.

Figure 1 can guide us in the discussion on Molecular Darwinism. The top of the scheme shows the three pillars of biological evolution: Genetic variation as the driver of the evolutionary process, natural selection that directs evolution together with the available genetic variants, and geographic and reproductive isolations that modulate the process of evolution. The different living conditions and the effective size of the biosphere on our planet, as well as enzymatic repair processes, limit genetic diversity.

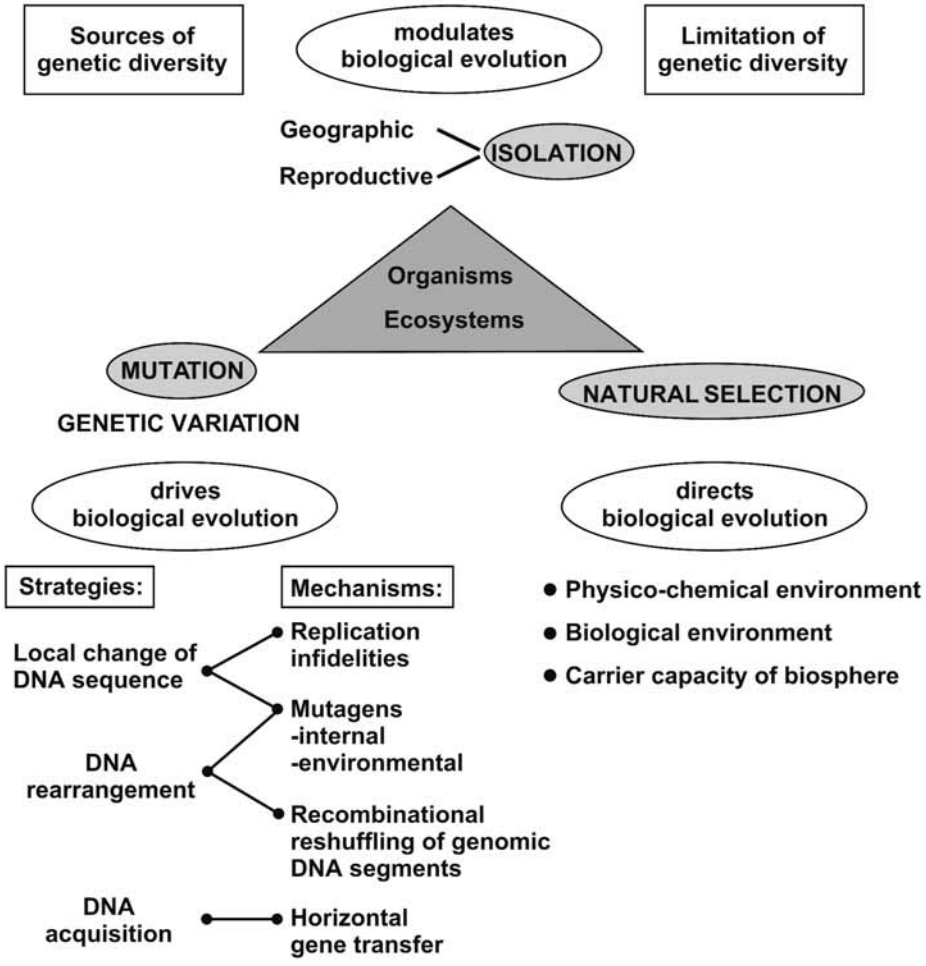

Figure 1. Schematic representation of elements involved in biological evolution and of the mechanisms and natural strategies of the generation of genetic variants (from Arber, 2008).

The lower part of Figure 1 shows on the left side in a condensed way the known sources of spontaneous mutagenesis. A detailed analysis to be commented below can lead to a classification of the different molecular mechanisms for the generation of genetic variation into three natural strategies of genetic variation: local sequence change, intragenomic DNA rearrangement, and DNA acquisition by horizontal gene transfer.

Before giving a more detailed explanation for these molecular mechanisms and strategies, it might be helpful for a better understanding to outline the main elements of the theory of molecular evolution (Arber, 2003, 2007) which is also called here Molecular Darwinism. We will see that a number of non-genetic elements contribute in reality to genetic variation. These elements are to a large extent intrinsic properties of matter, such as a certain degree of chemical instability of nucleotides. Another of these intrinsic properties relates to structural flexibilities of biological macromolecules, such as isomeric forms, in particular of nucleotides. Their tautomeric forms affect the proper base pairing in the double-stranded DNA molecules. Environmental chemical and physical (radiations) mutagens contribute of course also to spontaneous mutagenesis. Some of these mutagens are internal metabolic products (Smith, 1992). Still another factor is random encounter, e.g. of a mutagen with a cellular DNA molecule or of a gene vector with a target cell upon horizontal gene transfer.

Detailed studies of genetic variation processes have revealed that quite often, specific gene products are involved in genetic variation. These products of so-called evolution genes can act directly as variation generators and/or as modulators of the rates of genetic variation. Some examples will be discussed below. This outline shows that natural reality takes actively care of biological evolution. Genetic variation should not be attributed to errors and to accidents occurring to the DNA.

Local DNA sequence changes include the substitution of a nucleotide by another nucleotide, the deletion or the insertion of a nucleotide and also a scrambling of a few neighbouring nucleotides. There is good evidence that some of these sequence alterations occur upon DNA replication. It is known that enzymatic repair systems can rapidly spot the onset of these kinds of replication infidelities (Radman and Wagner, 1986). Upon the so-called repair, at least some repair enzymes can distinguish between the parental DNA strand and the newly synthesized strand. Consequently, they use the parental DNA strand as a master to put the affected nucleotide sequence in the newly synthesized DNA strand back into the correct parental order. Although these repair processes are quite efficient, they do not work with a 100% accuracy. This provides to cell populations a few rare local sequence changes in some individual cells, on the one hand, and to the individuals in the cell populations a relatively high genetic stability, on the other hand.

Let us now focus our attention on intragenomic DNA rearrangements. Various recombination enzymes are known to contribute to this kind of genetic variations. Generally speaking, these recombination processes can

affect DNA segments of various lengths, often containing one to several genes, and they can lead to the duplication, the deletion, the inversion or the translocation of a DNA segment, depending on the specific activities of the particular recombination enzymes at work.

By speaking on genetic recombination, one usually thinks at the so-called general or homologous recombination. In this reaction, the enzymes bring together DNA segments of a high degree of nucleotide sequence homology. DNA strands are then cleaved and spliced together across the two partners. In higher organisms these reactions are exerted in meiosis, when recombinants between paternal and maternal chromosomes are produced. In contrast, they do not work in mitosis during the normal DNA replication before each cell division. Bacteria also possess enzyme systems for homologous recombination. Again, in normal cell growth the enzymes become not readily expressed. But when breaks in the DNA molecules appear, e.g. after high energy irradiation, the so-called SOS repair becomes induced which produces a relatively high level of enzymes for homologous recombination. As a consequence, survival rates after irradiation rise, since intact genomes can be reconstructed by homologous recombination between sister DNA molecules that are present as already replicated genomes before cell division.

Mobile genetic elements are widespread in living organisms (Shapiro, 1983). These are DNA segments carrying normally one to several genes. The products of some of these genes are enzymes called transposases. Their activities can promote a translocation of the element to other chromosomal locations, sometimes in conjunction with a duplication of the element. These translocations are usually called transposition. Most bacteria carry in their genomes such mobile genetic elements, some of which are called IS (for inserted sequences) elements. Well studied *E. coli* bacterial strains carry in their genomes several specific kinds of IS elements, mostly in several copies of each kind. Interestingly, practically each kind of IS element (IS1, IS2, etc.) follows its own functional criteria, both for the selection of novel insertion sites on the DNA molecules and for the control of the availability of transposase activities at a low level, so that rates of transposition are actually very low. For example, IS30 (Caspers *et al.,* 1984) becomes inserted most readily into a specific, relatively short nucleotide sequence, although at much lower rates it can also insert elsewhere (Stalder and Arber, 1989). In contrast, IS2 prefers to insert in particular DNA regions of a length of a few thousand base pairs (Sengstag and Arber, 1983). But within these regions insertion can occur practically anywhere; the used inser-

tion sites show no distinct sequence homology (Sengstag and Arber, 1987). Transposons are mobile genetic elements that carry a segment with ordinary chromosomal genes between flanking elements that are responsible for their ability to transpose.

Transposition is not limited to intragenomic translocation, it can also occur to plasmids and to viral genomes during their residence in the bacterial cell. The impact of these possibilities on horizontal gene transfer will be discussed below. In this context, it is important to mention that some viral genomes can be counted to mobile genetic elements: They can insert into the host genome and at some later time excise again. Retroviruses that are widespread in higher organisms are a good example.

These considerations lead us to discuss processes of site-specific recombination. Indeed, some viral insertions into chromosomal DNA are to a high degree site-specific, and so is IS30, as we have seen. An interesting kind of site-specific recombination has been described as the basis for so-called flip-flop systems that are present in some bacterial and in some bacteriophage genomes (Glasgow *et al.*, 1989). We refer here to a flip-flop system that promotes the periodic inversion of a DNA segment. This segment is flanked on both sides by a 26 base-pairs long consensus (relatively high homology) DNA sequence. These flanking consensus sequences are carried in inverted order. The enzyme DNA invertase brings together these consensus sequences. The DNA strands become then cut and alternatively religated in the middle of the consensus sequences. This process results in the inversion of the DNA sequences carried in between the two consensus sequences. Inversion occurs back and forth every few generations in a growing microbial culture. This kind of flip-flop provides means to have two different genome organisations in a population of microorganisms. A strong evolutionary impact of this DNA inversion system, however, becomes obvious by the experimental observation that the enzyme-mediated DNA inversion can sometimes also occur between a consensus sequence and another, so-called secondary inversion sequence. Many different such sequences have been identified (Iida and Hiestand-Nauer, 1987; Arber, 1995). These secondary inversion sites do not show distinct similarities to the consensus sequence, and their spontaneous use shows at most some statistical reproducibility. This fact points to a certain specificity of the interactions. By using secondary inversion sequences, site-specific DNA inversion represents a source for novel gene fusions and for the assembly of an open reading frame for protein synthesis with an alternative expression promoter signal. We can conclude that these activities can be consid-

ered as active generators of genetic variants of evolutionary relevance. Note that in contrast to the regular flip-flop activity, the rates of using a secondary DNA site for DNA inversion are very low. These enzyme systems are a source for evolutionary novelty and they respect genetic stability of most of the individual cells in which they are carried.

With these presentations of a few well studied enzymatically mediated systems to promote occasional intragenomic DNA rearrangements we are impressed of the rich diversity of natural possibilities to provide novel genetic variations. In all these cases, the resulting genetic variants are, of course, substrates for natural selection. As we have already discussed, a majority of resulting genome orders may be lethal or of disadvantage. But it is the minority of winners, of variants providing selective advantage, that count for biological evolution.

Let us now look at horizontal gene transfer and its impact on the natural evolutionary strategy of DNA acquisition. As we have already discussed, microbial genetics has contributed much to today's knowledge in this field. Microbial genetics had a rapid start in the 1940s. In the already discussed transformation, free DNA molecules can be taken up by receptor bacteria, either actively or passively, depending on the particular microbial strain involved. In bacterial conjugation (Lederberg, 1947), two bacterial cells that can belong to different strains, meat physically. DNA from the donor cell can thereby become transferred to the receptor cell. A so-called fertility plasmid acts thereby as a gene vector (Hayes, 1964). Besides its own transfer through the conjugation bridge, the fertility plasmid can also provide the transfer of parts of the genome of the donor cell. A third possibility for horizontal gene transfer is its mediation by some bacterial viruses serving as gene vectors (Zinder and Lederberg, 1952). Again, we realize that nature was quite inventive with regard to the specific molecular mechanisms (Arber, 1994). In some of the processes, recombinant DNA molecules between viral and bacterial genomes are incorporated into viral particles (specialized transduction); in other instances, it is just a DNA segment taken from the donor genome that becomes incorporated into a viral particle (generalized transduction). Horizontally transferred DNA segments contain sometimes mobile genetic elements such as a transposon. This can facilitate an eventual incorporation of transferred DNA sequences into the genome of the receptor cell.

None of the horizontal gene transfer processes is specifically oriented to particular receptor bacterial cells. Transfer depends generally on a random encounter. However, between such an encounter and a stable integration of

the foreign genetic information into the receptor genome, several barriers seriously reduce the chance of acquisition of foreign genetic information. First of all, there is a requirement for surface compatibility of the receptor cell. In transformation the foreign DNA must find its way into the cytoplasm of the receptor cell, either by an active or a passive uptake, as already mentioned. In conjugation, the two mating partners must provide means for the building of a mating bridge. In transduction, the viral gene vector must find on the bacterial surface receptor sites that are required for successful infection. Secondly, many bacterial strains are equipped with one or even several restriction/modification systems (Arber, 1965). Restriction enzymes provide efficient means to identify if incoming DNA is foreign or if it had been produced in the same kind of bacteria. In the first case, the penetrating DNA molecules are cut into fragments by the restriction endonuclease. Within a few minutes the fragments are then further digested by exonucleases. However, at low rates foreign DNA fragments escape full digestion and succeed to incorporate at least some of their genetic information into the genome of the receptor cell. Generally speaking, successful DNA acquisition occurs only rarely and mostly in small steps, involving a part of a gene or one to a few genes at once. More or less random acquisition of foreign genetic information can often disturb the functional harmony of the cell in question. This is then the last barrier acting against successful acquisition. The hybrid resulting from acquisition will often have a selective disadvantage, less frequently a hybrid may have an advantage. This then represents a positive step in the process of evolution.

After having discussed examples for each of the three natural strategies of genetic variation, we can now compare the qualities of contributions made by each strategy to the evolutionary progress.

Local sequence changes offer the possibility for steps of improvement of available biological functions. Theoretically, local sequence changes could also represent a source for an occasional new biological function. But this can probably only become effective when the function in question starts to represent a substrate for natural selection.

DNA rearrangements can be seen as a tinkering with available capacities (Jacob, 1981). Novel combinations of functional domains from different genes may, for example, lead to a novel biological function. On the other hand, DNA rearrangements can also provide an alternative expression control signal to a functional gene. Such genetic variants may then express either higher or lower quantities of the gene product in question, as compared to the parental forms.

Favourable acquisition of foreign genetic information can be seen as a sharing in successful developments made by other organisms. In successful cases, the acquisition (in one step) of a biological function that the receptor organism did not possess, represents an extremely efficient contribution to the evolutionary progress. DNA acquisition as well as intragenomic DNA rearrangements might sometimes be a possible explanation for a sudden emergence of novel properties in evolving organisms.

The theory of molecular evolution postulates that evolutionary fitness may be reached when organisms are genetically equipped with the capacity to profit from all three natural strategies to generate genetic variants. For each of these strategies at least one, or better a few, specific mechanisms should be available.

In the light of this request one can postulate that in the course of the long past periods of evolution, the evolution genes, i.e. the sources for variation generators and for modulators of the rates of genetic variation, may have become fine-tuned for their functional activities by second-order selection (Weber, 1996). This means that those populations of organisms which had reached a certain degree of evolutionary fitness, were in advantage for adapting to changing living conditions. This can explain why organisms that live today are actually able to evolve and, nevertheless, to provide a relatively good genetic stability to the individuals of evolving populations.

In the description of molecular mechanisms and strategies for the generation of genetic variants, we have mostly referred to microbial experimental evidence. However, there are good reasons to assume that this acquired knowledge applies also to higher organisms. In recent times, more and more evidence for this expectation becomes available, particularly from DNA sequence comparisons. In this context, one can mention that some genetic variation generators nowadays also serve at the somatic level. A striking example is found in the somatic assembly of functional genes for specific antibodies of our immune system. In addition, some repair systems taking care of limiting rates of mutagenesis carry out their functions also in somatic cells. As far as horizontal gene transfer is concerned, one knows that some animal viruses can serve as natural gene vectors. In addition, symbiotic cohabitation of various microorganisms in animals and in plants is a very likely source for occasional gene transfer in one or the other direction.

A classical representation of long-term biological evolution is the tree of evolution. This tree usually shows the vertical flux of genes from the stem of the tree through the branches up to their ends, representing today's organisms with their enormous diversity. By taking care of the concept of

the evolutionary role of DNA acquisition, we have introduced more or less randomly placed connectors between branches as symbols for horizontal gene transfer (Figure 2). While in the vertical flux of genes the entire genomes are involved, upon the horizontal flux of genetic information only relatively short DNA segments become acquired, as we have already discussed. As we can expect from this modified representation of the tree of evolution, living organisms are not only interdependent by common roots in their past evolution, they are also interdependent in view of potential contributions to their future evolutionary progress by horizontal gene transfer. This new knowledge merits to become part of our understanding of biological evolution and it can enrich our world view.

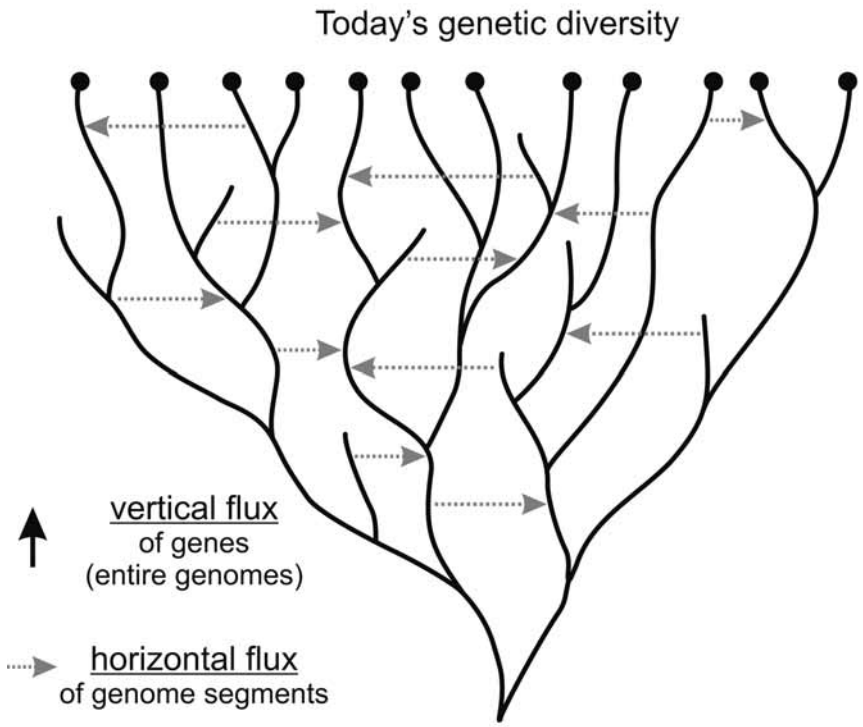

Figure 2. Actualized paradigmatic picture of the tree of biological evolution taking into account the evolutionary strategy of DNA acquisition by horizontal gene transfer (redrawn after Arber, 1991).

Another philosophical and world view aspect of molecular Darwinism is the notion of evolution genes. Although we know that some of their products serve also at the somatic level, other evolution gene products, such as at least some recombination enzymes and restriction/modification systems, clearly contribute only to the evolutionary progress of microorganisms. These genes are largely irrelevant for the bacterial life from one cell division to the next. We can thus conclude that genomes have a duality with regard to their content of genetic information. Many of their genes, such as housekeeping genes, accessory genes of use under particular life conditions, and in multicellular organisms the developmental genes serve for the fulfillment of the life of the organism. In contrast, the evolutionary genes ensure the capacity for biological evolution of the populations. Their products serve in cooperation with non-genetic elements for the expansion of life and for a slow, but steady, replenishment of biodiversity. Let me just mention that this philosophically interesting duality of the genome should not be taken as a strict classification of all the genes carried in the genome, since some gene products serve both for the needs of the individual life and for the capacity to evolve. However, the identified duality of the exerted functions can importantly contribute to a better understanding of the complexity of life and its evolution. From the scientific point of view, the living world of today reflects a long evolutionary path of permanent creation that may be based on a kind of self-organisation, and that must have its roots in the far past of the planetary evolution. The observed internal forces of the living world to undergo biological evolution gives us a guarantee that living organisms, at various stages of complexity, can continue to evolve and adapt to changing living conditions as long as such conditions will exist on our planet. Our actual knowledge on cosmic evolution predicts that appropriate conditions for organic life can still exist on our planet for about 5,000 million years.

REFERENCES

Arber, W. (1965) Host-controlled modification of bacteriophage, *Annu. Rev. Microbiol.*, 19, 365-378.

Arber, W. (1991), Elements in microbial evolution. *J. Mol. Evol.*, 33, 4-12.

Arber, W. (1994), Bacteriophage transduction. In: Webster, R.G. and Granoff, A. (eds.), *Encyclopedia of Virology*, Academic Press Ltd., London, pp. 107-113.

Arber, W. (1995), Genetic basis for evolutionary development of microorganisms. In: van der Zeijst, W.P.M. *et al.* (eds.), *Ecology of Pathogenic Bacteria; Molecular and Evolutionary Aspects*, Royal Netherlands Academy of Arts and Sciences, pp. 3-13.

Arber, W. (2003), Elements for a theory of molecular evolution, *Gene*, 317, 3-11.

Arber, W. (2007), Genetic variation and molecular evolution. In: Meyers, R.A. (ed.), *Genomics and Genetics*, Wiley-VCH, Weinheim, Vol. 1, 385-406.

Arber, W. (2008), Stochastic genetic variations and their role in biological evolution, In: Arber, W., Cabibbo, N. and Sánchez Sorondo, M., *Predictability in Science: Accuracy and Limitations*, Pontificia Academia Scientiarum, Vatican City, *Acta* 19, 126-140.

Avery, O.T., MacLeod, C.M. and McCarty, M. (1944), Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* 79, 137-158.

Caspers, P., Dalrymple, B., Iida, S. and Arber, W. (1984) IS30, a new insertion sequence of *Escherichia coli* K12. *Mol. Gen. Genet.*, 196, 68-73.

Glasgow, A.C., Hughes, K.T. and Simon, M.I. (1989), Bacterial DNA inversion systems. In: Berg, D.E. and Howe, M.M. (eds.), *Mobile DNA, American Society for Microbiology*, Washington, DC, pp. 637-659.

Hayes, W. (1964), *The Genetics of Bacteria and their Viruses*, Blackwell Scientific Publications, Oxford, UK.

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.

Jacob, F. (1981), *Le Jeu des Possibles*, Fayard, Paris.

Lederberg, J. (1947), Gene recombination and linked segregation in *E. coli*, *Genetics*, 32, 505-525.

Mayr, E. (1982), *The growth of biological thought: Diversity, evolution and inheritance*, Harvard University Press, Cambridge, MA.

Radman, M. and Wagner, R. (1986), Mismatch repair in *Escherichia coli*, *Annu. Rev. Genet.*, 20, 523-538.

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, *The EMBO J.*, 2, 67-71.

Sengstag, C. and Arber, W. (1987), A cloned DNA fragment from bacteriophage P1 enhances IS2 insertion, *Mol. Gen. Genet.*, 206, 344-351.

Shapiro, J.A. (ed.) (1983), *Mobile Genetic Elements*, Academic Press, Inc., New York.

Smith, K.C. (1992), Spontaneous mutagenesis: Experimental, genetic and other factors, *Mutation Research*, 277, 139-162.

Stalder, R. and Arber, W. (1989), Characterization of *in vitro* constructed IS30-flanked transposons, *Gene*, 76, 187-193.

Watson, J.D. and Crick, F.H.C. (1953), Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid, *Nature*, 171, 737-738.

Weber, M. (1996), Evolutionary plasticity in prokaryotes: a panglossian view. *Biol. Philos.*, 11, 67-88.

Zinder, N. and Lederberg, J. (1952), Genetic exchange in Salmonella, *J. Bacteriol.*, 64, 679-699.

DISCUSSION ON PROF. ARBER'S PAPER

PROF. SZCZEKLIK: Is there any evidence that horizontal transfer occurs in organisms higher than bacteria, multicellular?

PROF. ARBER: Absolutely. The retroviruses are known to be gene vectors, and evidence for an occasional acquisition of foreign genes comes from DNA sequence comparisons. Keep in mind that genetic variation can only occur at low rates. This renders experimental investigations on DNA acquisition quite difficult. Nevertheless, more and more biologists agree that horizontal transfer is a general phenomenon.

PROF. RAVEN: I do not exactly want to use the word 'species', but can you talk a little bit about how you view the diversity of bacteria philosophically? I mean, how many different ones should we name, how much do they differ from place to place?

PROF. ARBER: Well, probably less than 1% of bacterial types can be cultivated in the laboratory. Up to now experimentally accessible bacteria have been classified according to their abilities to use particular sugars, particular amino acids, and some other essential nutrients. This contrasts with the criteria to classify higher, sexual organisms on the basis of sexual fertility. More and more, genomics makes now entire genome sequences available, also from non-cultivatable bacteria. Future classifications will certainly largely be based on sequence comparisons. Already now, one can identify a kind of continuum of the genomes of different microbial isolates. This can be explained as a result of both the vertical and the horizontal transmission of genetic information. One can expect that scientific classification might end up with a large number of specific forms of bacteria.

PROF. PHILLIPS: There are millions of them.

PROF. ARBER: Yes, there are many, oh yes! I carry more bacterial cells with me, in and on my body, than the number of my own cells: roughly one kilogramme per adult. But this is important. The public opinion that bacteria are in general pathogens is completely wrong. From the scientific point of view bacteria are my friends in symbiosis, they help and facilitate my life. But how can I tell that to the people in the street who cannot observe the bacteria at their work?

PROF. PHILLIPS: I was confused by one point. You said that there was no evidence of directedness of spontaneous mutation and then you hastened to say, well, this seems to be somewhat in contradiction to what we have just heard from Rafael, but you said, 'I am not talking about stress situations'. Now, if I understood correctly, I thought that what you were saying was that everything was random but that stress, in a sense, increased the rate of mutation so that you could get out of whatever difficulty, is that right? So you would also agree that there is no directedness of stress mutations either, is that right?

PROF. VICUÑA: It is right, but in certain stressful conditions, mutations may appear to be directed.

PROF. ARBER: An observable response to a stress situation is rather an exception and, usually, if one studies the reasons intensively, one can often find a causal explanation why it is so.

PROF. PHILLIPS: Are you saying that you have seen evidence of that or that you can see how it might happen?

PROF. VICUÑA: There is evidence on how you can get directedness for a specific mutation. It depends on the system, but it would be an exception.

PROF. ARBER: But you should avoid generalising.

PROF. POTRYKUS: I have some problems with the general acceptance of horizontal gene transfer in multicellular organisms. It is certainly true that there is the possibility, but it can only have consequences if the transferred piece of DNA reaches the germline. Plants do not have a germline so the DNA ends up in somatic cells. It can have consequences if you can regenerate a plant from a somatic cell, but it is not the general natural mecha-

nism. So I am looking for alternative explanations for the widespread detection of homology. I am not convinced that this is the only explanation that there is horizontal gene transfer.

PROF. ARBER: I agree with you on some principle differences between higher organisms and bacteria with regard to the germ line. What you observe for higher organisms is possibly just another isolation phenomenon to keep the rates of horizontal gene transfer into germlines.

PROF. M. SINGER: So I was struck by the fact that you never mentioned *Archaea* in the talk. You talked about bacteria and I do not know whether you were including them or whether anything is known about horizontal transfer in *Archaea*.

PROF. ARBER: I must confess I am not sufficiently familiar with the literature, whether that has been clearly shown. Does someone else know? Yes, there are a few cases. For the time being the problem is that bacterial genetics is based on a handful of cultivatable strains. We then just extrapolate and often generalize the acquired knowledge. This, however, requires validation in future times.

Session III

INSIGHTS INTO HUMAN EVOLUTION

WHY IT IS USEFUL TO KNOW THE MODERN THEORY OF EVOLUTION

LUIGI LUCA AND FRANCESCO CAVALLI-SFORZA

1. DARWINISM

In the last two centuries there has been much discussion on the hypothesis, not unfamiliar to the ancients, that all living species originated from simple forms by a long process of evolution, that is, by transformation and differentiation into a great variety of species. Basic contributions came from Jean Baptiste Lamarck (1802) and Charles Darwin (1859). A major scientific step forward was accomplished through the understanding of the laws of inheritance in higher organisms, valid for the great majority of plants and animals, which we owe to the research carried out by the Czech monk Gregor Mendel. He was so far ahead of his time with regard to intelligent experimentation, that it took 34 years for Western scientists to appreciate, rediscover, and confirm with humbler experiments Mendel's original findings, which were communicated to the Natural History Society of Brünn, Moravia, in 1865, and published in its proceedings in 1866.

The introduction of *Drosophila melanogaster,* the fruit fly, as research organism by the group formed by T.H. Morgan at Columbia University, New York, made possible the rapid development of genetics after 1912, as the science of biological inheritance came to be called. In the twenties three geneticists: R.A. Fisher, J.B.S. Haldane and Sewall Wright set the mathematical foundations of the modern genetic theory of evolution, which were later enlarged by Motoo Kimura and many others. They thus applied Galileo's recommendation that, for scientific understanding, you must first learn the characters in which the world is written, and that the universe is written in mathematical language.

The demonstration in bacteria (1944) that DNA is responsible for biological inheritance; the discovery (1953) of the chemical structure of DNA;

and the research in the field of 'molecular genetics', that ensued, led to the examination of whole genomes, and provided new powerful means for studying their evolution.

Today evolution is no more a hypothesis and there are ample proofs that it is motored by natural selection. Knowledge of the sources of inherited variation and of the mechanisms that maintain it, at the same time favoring the transformation and differentiation of species, has greatly enriched a well-organized theory. The succession of evolutionary steps leading to the great variety of living organisms is being traced with astonishing precision thanks to the detailed analysis of whole genomes. That species change is no more a hypothesis or a debatable theory, and how and why they do so is becoming a matter of detailed proofs. Until a short while ago Genetics, once just the science of biological inheritance, was the Cinderella of Biology, but it has now become its central discipline, and has turned modern biology from a descriptive, morphological, 'qualitative' discipline with no theoretical background, into a highly sophisticated, quantitative science based on gigantic, exhaustive whole genome DNA data sets, and on very advanced techniques. Genetics also generated molecular biology, a slow but essential scientific machine that is systematically clarifying the very complicated network of metabolic pathways necessary to make a living organism develop, build, maintain and reproduce itself. To avoid confusion, it is worth adding that, while today we prefer to speak directly of DNA, earlier the structures responsible for inherited traits, then unknown, were called 'genes'. This word has now taken on a new meaning, limiting it to DNA segments making a protein with a specific function, but is still frequently used in its original looser meaning of inherited unit.

A recent development of genetic thinking that is referred to as 'epigenetics' is showing that the DNA of specific tissues can change during the development of an organism. Some of these changes, as in the formation of tumors (especially malignant ones) are definitely pathological. These and many other 'normal' processes that take place during development in 'somatic' DNA show how complex the process of development really is, ranging from somatic mutations and temporary partial prevention or modification of function of major parts of DNA, to the contributions of RNA to regulatory processes of gene action, and to occasional pathological deviations in the structure and function of proteins described under the name of 'prions'. In general, however, it is important that the DNA destined to pass information to future generations seems to be set aside fairly early in cells of the 'germinal' line, destined to produce 'gametes' (sperm and egg cells), and is basically excluded from these epigenetic developments.

The 'mystery of life' has now become very simple. A living organism is an organism capable of reproducing itself, generating other organisms that are almost identical to itself. The word 'almost' is added to indicate the fact that mutations are rare: they are chiefly very small errors in copying the hereditary patrimony, which is chemically a substance called DNA and is essentially a book of instructions on how to build a new organism, almost identical to the parent/s, a copy of which is transmitted by the parent/s to the children. The copies of DNA received by the children, that they use to direct their own development, are copied again for passing them on to their own children. Thus the copy errors made by parents in producing gametes accumulate into the master textbook the children will use, for their own development as well as for making new copies of DNA (with new errors) that will be passed on to their children and all descendants. Mutations are thus the main stuff evolution is made of, because they introduce all real novelties into the living world.

Most mutational changes taking place at every generation have little if any effect on the organism carrying them, or at least do not affect the adaptation of their carriers (i.e. are *selectively neutral*), but mutations that determine selection are those that affect the capacity to survive to reproducing ages, and/or the fecundity of the individual, because they alter automatically the composition of the next generation. In fact, changes of physical traits increasing survival probability or fertility (the number of children born), will generate relatively more descendants than the rest: they may therefore be spoken of as an evolutionary 'improvement' over the original types. Darwin, and independently another English naturalist, A.R. Wallace, understood around the middle of the XIX century that improved survival and/or fertility would thus *inevitably* cause evolutionary changes of living organisms over time and space, ensuring better adaptation to the environment/s. In fact those organisms that have more children than the original types must be, in some way, *fitter* than the ancestral type to the environment in which they live, and *if* the characteristics causing higher fitness are inherited by the progeny of the fitter types, their greater survival/fertility will increase their relative numbers in successive generations, causing a population change in time. Thus species will be transformed and will go on adapting ceaselessly to changes in the environment that demand different adaptations. Similarly, differentiation of a species in space will also arise in the course of time, wherever local environments differ.

In other words, *evolution due to natural selection is an automatic transformation of any species over time*, leading to differentiation in different

environments, due to the higher survival/fecundity of fitter types. Higher fitness is measured in the demographic terms of higher survival and/or fertility, but in terms of the structure and function of the fitter organism this must mean that the fitter type is somehow better 'adapted' than the original type to live in its particular environment. One can therefore observe an increase in the average adaptation of a population to its environment in the course of evolution, and R.A. Fisher showed that the rate of increase in adaptation can be predicted by the variation of what he called individual 'Darwinian fitness' in a population. He named this the 'fundamental theorem of natural selection', strictly valid under the condition that the selective advantages of the different types does not change. If this condition does not occur, more complicated theorems take over. The accumulation of changes over time because of natural selection can increase, or reduce the complexity of living organisms. More and more complex organisms have thus evolved, but parasites can lose many complex organs and functions, initially necessary for feeding and reproduction, because they can use those of their hosts. The general results are organisms that are very efficient, often with increased complexity that is useful for prospering, and we marvel at the apparent perfection of their structure and function.

Fisher noted that natural selection is a mechanism that causes 'improbability', by the accumulation of higher fitness over generations. Some observers think it is very unlikely that modern living organisms could ever arise by chance alone from much simpler organisms, forgetting that they had an extraordinarily long time available for building the organs that help them to live, and that they did so over many generations and in a great number of steps, most of which increased only modestly their survival and reproduction skills. This increase is constant nevertheless, even if it is mostly small and hard to notice, because of the very nature of living organisms, which can replicate themselves. Self-reproduction is constantly subject to natural selection, and consequently every generation contributes to some genetic improvement, in each species.

This modern synthesis of Darwinism and its translation in quantitative terms points thus to a process determined essentially by mutation and natural selection, that is, the spontaneous production of DNA changes and the automatic filtration of those that permit improved adaptation to the environment. This filtration takes place through the different survival and fecundity of carriers who are somehow better adapted and can pass that quality to their children. All mutation products that have fitness greater on average than that of the original type, will increase in relative numbers with

the passage of generations, at a speed in time an space that depends on how much greater the fitness of the mutant is over that of the original type (more generally, of the population average), and on migration. Those mutant types whose fitness is inferior to the population average must decrease in numbers and eventually disappear.

It is important, however, to note that fitness as a measure of adaptation is not, strictly speaking, a property of genes (DNA), and of genes only. More exactly, it is a property of the 'phenotype', i.e. the actual product of genes in the development of the organism, and it depends also on the environment in which growth, development and everyday life occur, including, especially in humans, behaviors culturally transmitted, i.e. learnt during development. An otherwise very good book by Richard Dawkins, *The Selfish Gene*, forgot to mention the caveat that natural selection directly affects phenotypes, not genes, an error which Dawkins later corrected.

Darwin was also impressed in drawing his conclusions by Malthus' observation that the number of children generated by any living organisms is practically always greater, often much greater than that allowed to live by the available resources. There is not enough room for all those who are born: some must die early or not reproduce. Natural selection can therefore be viewed also as a highly competitive struggle for existence, because not all children may manage to contribute to the next generation.

Some religious environments did not like this concept, because competition to survive seemed intuitively incompatible with a loving God. But what seemed most offensive to a large number of XIX century Anglican prelates (there was a famous exchange in 1860 between Julian Huxley and the bishop of Oxford, Samuel Wilberforce), was the inevitable conclusion that humans have common ancestors with animals, especially with our nearest Primates, like chimpanzees. Zoos were beginning to be built, and everybody could observe pictures or even living specimens of Primates. Recent research on Primates has actually shown that the gulf between the nearest Primate and us is not as profound as it seemed in the XIX century. The major difference objectively observed between us and other Primates is that they cannot develop an articulated and rich language like ours, and this may have proved a major limitation to the development of communication within different Primate families, and thus to cultural evolution.

The Bible gives humans a privileged position with respect to animals, by assuming our similarity to God. Jews were not allowed to use art to make representations of God, and this decreased the dangers of imagining men's similarity with God as physical, rather than spiritual and intellectual. In other cultures, when artists were given freedom of picturing the phys-

ical similarity of their Gods and showed them in human shape, there was increased potential for conflict between science and religion, as Gods inevitably became natural rather than supernatural beings.

The Bible also makes the acceptance of evolution impossible, if one takes literally the word 'days' in the statement on Genesis, and not as rough geological eras. Actually geology antedated biology by almost a century in making the literal interpretation of the beginning of Genesis scientifically obsolete.

The word 'Darwinism', as used today by its critics in philosophical or religious circles, is often plagued by a number of misunderstandings and abuses of the basic Darwinian concept of natural selection, that have little or nothing to do with Darwin's or the modern understanding of his theories. Discussion is useless with people who have not learnt that natural selection is a direct, inevitable, and automatic consequence of basic demographic processes. Darwin, by the way, knew nothing of Mendel's experiments, which were published five years after Darwin's first book, *The Origin of Species*, but remained practically buried until 1900. The word Darwinism is used correctly only if it refers to Darwin's idea of natural selection, remembering that he also did emphasize that the fundamental need for selection to be effective is limited to inherited traits. Darwin's ideas on inheritance mechanisms were inevitably vague but do not affect the validity of his understanding of natural selection.

2. THE MAJOR FACTORS OF EVOLUTION

Natural selection is not the only factor of evolution. Today we have considerable knowledge of the basic mechanisms of genetic change that give rise to the diversity of DNAs. For all we know today the errors of copy of DNA, which we call mutations, are spontaneous and *random,* in the sense that they are unpredictable, and are not necessarily directed, for instance, in an adaptive direction. Their rates of occurrence can be estimated, with some difficulty because mutations are also *rare,* and large numbers of individuals must be examined. There is a good reason for the rarity of mutations: living organisms are complex mechanisms and they need *all* their organs and functions to be reasonably efficient, to ensure their own survival. Hence errors of copy of DNA must be rare or mostly not dangerous, and in fact mutations are rare and most of them do not affect Darwinian fitness.

DNA is made of very long filaments (the chromosomes) formed by a chain of units whose chemical nature is that of a 'nucleotide'. There are four

types of nucleotides that can be aligned in any order, called A, C, G and T (the initial of their chemical names). It is common to suggest that DNA is a book of instructions for making a living organism, written in an alphabet formed by four letters. Human DNA is like a library made of 23 volumes (the 23 chromosomes). In sexual reproduction of 'diploid' individuals like us and the great majority of plants and animals, each individual receives one copy of each chromosome type from the father, and one from the mother, so that every cell in the human body, aside from the reproductive cells, numbers 23 *pairs* of chromosomes. All parts and units of DNA can mutate: the most common changes are called single nucleotide polymorphisms or *snps*, and are the replacement by mutation of a specific single nucleotide in a particular position on a particular chromosome by any other of the nucleotides in the set: A, C, G, T. *Polymorphism* means that both the ancestral type (allele) and the mutated allele are found in the population, usually in such frequencies that a study of 100 or even fewer individuals would find both alleles.

In our species *genetic* (= DNA) *diversity*, meaning the presence of polymorphisms, has been so far observed in about 0.5% of the billions of nucleotides forming the 23 chromosome pairs of the first man whose genome was fully investigated and published (Craig Venter). This variation of about 15 million nucleotide sites means that in 15 different million specific sites the contributions by Craig's father and mother were different, and are due to mutations that occurred many generations ago. Extending full sequencing of the genome to many more individuals will certainly increase this estimate of polymorphic sites. We call a site *heterozygous* when the paternal and maternal contributions differ and we use the percentage of sites that are heterozygous as a measure of the genetic diversity of an individual.

Mutation rates are a property of nucleotide sites: they can change under special conditions, and it is possible that they are adjusted by natural selection to optimal average values. It is interesting to note that mutation rates, if considered per unit of biological time, which is the generation time (the average age of reproduction) of the specific organism we study, tend to be of the same order of magnitude for many organisms, even though the difference in duration of a generation time between, say, bacteria and humans goes from thirty minutes for bacterial generation, to thirty years for humans: this means that the rate of reproduction is roughly half a million times greater in bacteria.

There have been efforts to show that mutation is not always random but tends to be adaptive, i.e. a mutation useful for the organism is more likely

to appear than other random mutations. There is no evidence today that this is true, in spite of many attempts (e.g. recent ones by Cairns). This is not to be confused with the fact that if a favorable mutation appears it will be picked up by natural selection and expand, until it becomes the norm in the species. One of the greatest geneticists of the last century, who unfortunately died about a year ago, Joshua Lederberg, worked on this problem in the last years of his life. The last time one of us had an opportunity to discuss it with him he said that it is clear that there are some 'funny things' in the mutation process judging from mutation rates, but nothing is clear. A reasonable guess is that a gene that is functionally active is more likely to mutate than one that is inactive, and one small attempt to test this hypothesis was made by Luca Cavalli-Sforza, but the effort remained inconclusive.

Mutation and natural selection, however, are not the only factors of evolution. The modern theory of evolution includes other factors: the major ones are drift, migration, and recombination.

Drift, more accurately called *random genetic drift*, is defined as the variation in the frequency of polymorphisms, through succeeding generations, which depends on the size of a population, intended as a social group whose members rarely marry outside the group, or accept foreign members for reasons other than marriage. This is or was in earlier times the tribe, basically a linguistic unit (tribe and language have the same name or names) that usually claims common ancestry. Population size, N, increased considerably during human evolution ca. 10,000 years ago, when a major change in food acquisition took place: the change from hunting-gathering and/or fishing to agro-pastoral economies, i.e. from food collection to food production. Until then, and throughout most of the evolution of the genus *Homo*, the size of N may have ranged from a few hundred to a few thousand per social group (the tribe), that is ca. 1000 as order of magnitude. The few surviving tribes of hunters-gatherers are of this size.

As we shall see, the evolutionary effect of drift is that of causing the reduction of genetic diversity, as estimated by the percentage of sites that are heterozygous in a sample of the individuals from the population. If prolonged indefinitely, drift would reduce genetic diversity of the population to zero, an ideal situation for a racist, who would probably consider attractive a greater genetic homogeneity of all individuals forming one's social group. But loss of heterozygosity is not at all desirable: the progeny of close relatives suffers from mortality and morbidity that are greater, the higher the degree of relationship of parents. By contrast, higher heterozygosity, found for instance in 'interracial' hybrids, is likely to show greater vitality under a

variety of respects – a phenomenon already well known to Darwin, and called hybrid vigor. Human social customs are usually geared to avoid too close relationship of husband and wife, and it has been estimated that tribes of size above 400 or 500 can escape damage caused by marriage among close relatives. Moreover, there is almost always some immigration, mostly by marriages with persons from other, usually nearby tribes.

In the last 10,000 years, the passage to agro-pastoral economy caused a considerable increase in population size, not far from a 1000 X factor. Tribes of hunter-gatherers have often maintained their original tribe name, which is usually also that of their language, but the new economy allowed considerable growth. In Nigeria, for instance, the four most important tribes (Hausa, Yoruba, Ibo, Fulani) have now more than ten million members each, but there are many much smaller ones.

Migration is another major demographic factor of evolution. When migration takes place among different tribes, it usually tends to reduce drift. Traditionally, much of it is due to marriage with a member of another tribe, or to work, which has recently been in constant increase. If the % of in-migration per generation of a tribe is m, the larger is m the more effective is the avoidance of drift effects. A larger population size, N, has the same effect, and their joint result of m and N in counteracting drift is measured by the product Nm. In Italy, Nm varies from 0.1 (in mountain isolated villages) to >2.9 (in towns of more than 100,000). (Observed data can be found in Cavalli-Sforza, Moroni and Zei, 2004).

Much migration occurs on an individual basis, especially when it is due to intertribal marriage, and is a very powerful factor that reduces drift effects. But there is a type of migration that acts in an opposite direction, generating new opportunities for drift: the migration of a group large enough to form a new colony. This takes place especially if the colony is far enough from the motherland, and contact with it is rare, for instance in the case when conflict was the reason for leaving the motherland. Puritans who escaped religious persecution founded some English colonies in North America, and the same was true of the French and Germans who joined the original Dutch founders of South Africa.

Long before any recent historical case, a special process of continuous migration accompanied several expansions of our species to the world. The oldest expansion of the genus *Homo* was from Africa to the Old World, Europe and Asia, about 1.7 million years ago. We know little about it genetically, because the earliest Eurasian human species, called *Homo erectus*, has probably left no direct descendants. The ancestors of our species, that

eventually became *Homo sapiens sapiens* (considered undistinguishable from anatomically modern humans), lived between 150,000 and 100,000 years ago in eastern Africa, and spread to all of Africa starting perhaps 100,000 years ago. But maybe just one tribe that must have been most advanced in language development started expanding about 60,000 years ago from East Africa and continued until it settled the whole world. While Australia and New Guinea were already settled by them 40,000 years ago, southern Chile, the most distant place from the East African origin, was reached 11,000 years ago, after crossing from Siberia to Alaska. The expansion covered a distance of about 25,000 km at an average speed of half a km per year: much of it probably took place along the coasts or rivers or oceans, and went faster as time passed. Major oceanic islands were reached later, mostly from S.E. Asia, beginning some 6000 years ago. Some very small and especially isolated island, like Pitcairn and Tristan da Cunha were settled only a couple of centuries ago by a dozen or so settlers, who afterwards increased in numbers at a regular rate.

The introduction of the agro-pastoral economy occurred at similar times in different areas of the world, and generated the major crops and domestic animals that still support us: wheat and barley, sheep, goats, cattle in the Middle East; rice, millet, chicken and many fruits in East Asia; maize, beans, squash, tomatoes and turkey in Mexico. They all probably developed from the same semi-conscious biological discovery: how living organisms are born. They mostly developed, probably independently, in near-tropical areas at mid altitude, where food was rich but population density outgrew the resources. The new economy spread slowly, about one km per year to Europe and to Central and south Asia, by a combination of demic diffusion (of people: the farmers themselves) and cultural diffusion (local hunter-gatherers learning about food production technology from immigrant farmers). In the Sahara, there were at an early time very sophisticated agro-pastoral developments, but the region dried up around 5-6000 years ago, and farmers had to go south. They were especially successful in West Africa, with limitations imposed by the poverty of the soil and the difficulties of raising crops and animals originating from the Middle East. Using local plants, agriculture reached the Nigeria-Cameroon boundary, where in the first millennium BC an ally joined it: iron use, coming from the Middle East via Egypt and Sudan. The Bantu expansion had its origin there and spread to central and southern Africa. But African agriculture remained poor, until manioc arrived in the XVIII century AD, probably brought by a missionary coming from South America. Manioc was domesticated in the central Andes, and

made possible the expansion to the South American plains via the major rivers, before conquering most of Africa in three centuries or less.

Agriculture changed the world. Hunter-gatherers were professionally nomadic, having to shift continuously to new hunting grounds. They traveled in small flexible groups, with no chiefs – a perfect democracy they still practice. People who travel all the time can own almost no personal property. But farmers had to settle near their fields, could build permanent houses, property became an advantage and a rule, and a variety of new jobs developed, requiring specialized skills. Societies acquired fixed caste structures, with chiefs, which reached the apogee in India, where the caste system has now disappeared, but only in towns.

The introduction of writing, the earliest in the Middle East and Egypt around 5000 years ago, began history. Metals soon followed, first copper then bronze and iron, all discovered above the Middle East, beyond the Caucasus. War, loot and piracy became a way of life, making defense necessary. Pastoral life separated largely from agriculture and went its own way, turning into a style of life in arid lands. This takes us to the history we learn at school, of which the Bible became a major record. According to some researchers, Genesis was written in two versions, later intermingled and partially contradictory, and it relates to the histories of two different tribes of farmers, one of which had partially reverted to hunting and gathering, or perhaps to a strictly pastoral life.

3. Natural selection and drift: the relative importance of adaptation and chance in evolution

3a. *Serial Founder Effect*

It seems likely that the so-called 'Out of Africa' expansion that settled the whole world and generated all presently living humans progressed by a series of repeated migrations of relatively small groups, which started out from the most peripheral colonies, settling not very far in uninhabited territory. This would allow a pioneering small colony to remain in contact with relatives and friends, in general with what was 'civilization' at the time. It is very unlikely or even impossible that there was admixture between modern humans and descendants of *H. erectus*, who must have had a very low population density throughout Eurasia at the time. There is so far no evidence of admixture of our species with Neanderthals, who lived in Europe at the

time it was first settled by *H. sapiens* and were certainly far more advanced than *H. erectus*. Neanderthal has now been shown to be sufficiently different from modern humans to be considered another species, although it separated from *H. sapiens* much later than any other branch.

There must have been a large number – at least hundreds, perhaps even thousands – of similar events of foundation of new colonies, one after the other, in the many directions in which expansion proceeded away from East Africa. Hunter-gatherers live in camps made of huts that are rapidly built, and move across a fairly large area that makes up their hunting ground. In the search for new ground, a group smaller than a tribe, and probably of small size, may have explored new territory at some distance from the mother tribe. If the new area was found suitable and the small group settled it, *a new opportunity for drift*, and therefore local loss of genetic diversity was created. In all cases when a group lives in an isolated island or region, or for social reasons (religious, political, etc.) breeds separately from the other local population, drift can create genetic as well as cultural differences, the magnitude of which depends on the size of the population.

This is shown by a large number of examples from medical genetics: quite a few instances of rare genetic diseases are found in genetically isolated populations in which a mutation arose a few centuries ago. If the group increased in numbers subsequently, it will be especially easy to find several cases of the same disease today. This is common in particular for recessive genes (that do not show in the heterozygous condition, but come to light in one out of four children, in marriages between two heterozygotes). Jewish people have traditionally good medicine and have discovered a number of new recessive diseases, some of which are found also in different populations, while others are present only among Jews, more often in individual Jewish groups that separated from each other in one of the several diasporas that spread Jews around the world in the last 2500 years.

Ashkenazi Jews, for instance, were subjected to one of the worst genocides in World War II, and their survivors are now mostly in the US and Great Britain. It is believed that they originated from a small group that migrated from Rome to central Europe, perhaps a thousand years ago. Genetic screening of members of the Ashkenazi community indicates that 50% of them are descendants of just four women. Several mutations that occurred probably during their expansion in N. Europe gave rise to a relatively large progeny carrying mutations rare elsewhere; some were not even found outside the Ashkenazi. These observations of cases of genetic diseases, found in a few populations that expanded recently, or more generally in ‘genetic isolates’, are referred to as ‘founder effects’.

3b. *Genetic and geographic distance*

It has lately been shown that the recent human 'Out of Africa' expansion has generalized the founder effect to the whole world. It is in fact reasonable to view the expansion of modern humans, from an original relatively small African tribe, as the sequential founding of small colonies, and therefore as a sequence of founder effects that ran across the whole world in the ca. 50,000 years period that it took to cover the 25,000 km between the place of origin of our species and the farthest places. The progression of the species by successive episodes of colonization, each of which gave rise to a founder effect because of drift, due to the usually small size of the early colonies, must have caused a linear fall of genetic diversity from Africa to S. America, first observed by Prugnolle and others (2003) by examining the HGDP (a collection of DNAs from 52 indigenous populations of the five continents, (L. Cavalli-Sforza 2004). The Stanford research team confirmed it by doubling the number of original observations obtained, with genes called microsatellites (a total of 783 of them, Ramachandran *et al.*, 2005) and later by examining 650,000 snps of the HGDP populations (Li *et al.*, 2008). The explanation offered, summarized by the name of 'serial founder effect', was tested by simulation. The average single founder step suggested by the simulation corresponds to an Nm of about 0.3, in reasonable agreement with anthropological information on surviving hunter-gatherer populations (Ramachandran *et al.*, 2005). There were most probably hundreds of these successive colonizations from beginning to end, on any of the different routes made by our African ancestors who settled the world, and the total of single founder effects must have been of many thousand.

The same papers (Ramachandran *et al.*, 2005, Li *et al.*, 2008) also showed that there is a very close correlation between genetic and geographic distances (measured as the crow flies, with entrance to the Americas by the Bering strait) of all the HGDP populations, when each is compared to each of the others. The correlation is 0.87 with microsatellites and 0.89 with 650,000 nucleotide sites. Such close correlations are most easily explained by simple drift, plus migration limited to geographically close tribes, and allow the suggestion that true natural selection effects during the great 'Out of Africa' expansion might amount at most to about 20% of the total genetic variation observed today among indigenous populations. This has been the first large-scale attempt to estimate the relative importance of selection versus drift in the origin of the genetic variation observed in a species. Our species is the one that lends itself best to such computation, because of the availability of the necessary demographic estimates of population sizes and migration, difficult to obtain in other species.

3c. *Drift in the Parma Valley*

One of the present authors (LLCS) was responsible for the very first attempt at measuring the relative selection/drift ratio in humans. The opportunity arose thanks to the information and support offered by one of his first students of Genetics at the University of Parma in 1951-52, the Catholic priest Don Antonio Moroni, who made him aware of the existence of demographic data, that had potential interest for genetic study, collected by parish priests over the centuries and available in the Catholic Church records. Demographic data from 74 parishes of the Parma Valley, covering the last 400 years, were used in the research: population sizes, migration, and frequencies of consanguineous marriages. Genetic distances among the parishes were calculated for 14 blood group genes then available, obtained from a total of 2875 individuals. The parishes varied in population size, from less than a hundred to several thousand individuals, with a strong stratification of village size and migration by altitude. It became clear that demographic data, based on 400 years of demography (population size and migration), could predict very well the genetic variation within and between villages (parishes) on the basis of drift alone.

In addition, computer simulations (Cavalli-Sforza and Zei) were made to test how long it would take, given the observed migrations and population sizes, to reach an equilibrium value. Both migrations and population size affect variation among populations (smaller village size increases variation among populations, here represented by parishes, but increased migration acts in the opposite direction, reducing it). The greatest variation among villages (parishes) is observed in the highest, mountainous part of the valley, where they are also smaller; in the intermediate altitude part (hills), the size of villages and the genetic variation among them are intermediate; while in the plains population density is highest and parishes are proportionately largest, and there is no measurable genetic variation among parishes over that expected by random sampling in a homogeneous population.

The computer simulation of the blood group data, starting from complete genetic homogeneity of the population, showed that the variation among mountain villages increased regularly over generations and came to a stop, as expected in conditions of equilibrium between drift and migration, after about 250 years (8 generations). The observed variation among villages agreed with that expected on the basis of the simulation. There was a mistake in the original study that gave a small difference, but it disappeared in the most recent analysis of the data (Cavalli-Sforza, Moroni and

Zei, 2004). This book contains all the data collected by our group in Italy in the last 50 years; they were gradually extended to much of the rest of Italy and to other sources of data, like surnames, dispensations for consanguineous marriages, etc.

With a population like the one studied for blood groups in the Parma Valley, and with the numbers of individuals tested, drift provided therefore a sufficient explanation of all the observed genetic variation for standard blood groups, leaving no evidence of natural selection. Some natural selection could be shown in early, classical observations on blood groups, but only by using special approaches, like mortality and morbidity of RH+ children born to RH- mothers. Such an effect would hardly show in the approach used in the Parma Valley. By contrast, in the analysis of 52 world populations, with all the genetic variation tested by 650,000 nucleotides, some natural selection effects did appear and are now being examined further in a paper being prepared for publication by J. Pritchard. Serial founder effect did not provide complete explanation of variation among the HGDP populations, but left a fraction of about 1/5 of the genetic variation potentially explained by natural selection, 4/5 being explained by drift.

For readers interested in the origin of this estimate, it is calculated from $1–r^2$ where r is the correlation coefficient between genetic and geographic distance, under the assumption that geographic distance can explain genetic distance entirely. But 1/5 is actually an overestimate for the contribution of natural selection, because a substantial part of it is explained by the red dots of Figure 1 of Ramachandran *et al.* deviating from the straight line, and they are due to the fact that the three oldest African populations have separated earlier and have been exposed to drift for a longer time than the rest, thus building a greater genetic distance from the other African populations.

3d. *A clear example of natural selection: lactose tolerance*

Direct study of individual genes known to be under selection shows it is possible to detect the place of origin of a mutation that is known to have increased in frequency because of higher fitness. By observing how it spread around, the selection coefficient (fitness value) can be calculated. Examples of natural selection clearly demonstrated so far are of individual genes that became known in other investigations, and the evidence comes from finding that mutants of the gene cause specific diseases. Among these, the most interesting one is for an snp that is a regulatory mutation of the gene making the enzyme lactase, which allows metabolizing the milk sug-

ar lactose. The enzyme-producing gene is located in the second chromosome, and a gene that regulates its production is located very close to it, within another neighboring gene (Peltonen *et al.*). The ancestral regulatory nucleotide site is responsible for suppressing the production of lactase after weaning, once milk is no longer available to the growing organism. The gene is found in all Mammals, as well as in the great majority of humans, because the consumption of milk after weaning is limited to a vast area centered around the Middle East, where sheep, goats, and cattle were first raised. In this area it is common to find a mutation of the regulatory gene, that does not stop the production of lactase after weaning, so that carriers of the mutant continue producing lactase and can therefore utilize the milk sugar for all of their life.

It has been shown that the mutation arose in an individual living somewhere in the Ural Mountains about 6000 years ago, probably a member of a reindeer shepherds' tribe that must have started consuming milk in adulthood. Adults of the ancestral type, that lose lactase production after weaning, suffer gastro-intestinal pains and other complications when they try to consume milk – at ages at which the lactase enzyme is no more produced – so they tend to abandon the custom. This condition is called *lactose intolerance*, while the capacity to consume milk as adults, without troubles and enjoying full benefit from the calories available upon digesting lactose, is called lactose *tolerance*. This capacity is especially advantageous in cold climates, which is where the mutation probably arose and therefore prospered particularly well.

The tolerance mutant is now very frequent in Scandinavia (90-95%), which is nearest to the place of origin, and in Great Britain, that saw the arrival of many Scandinavian Vikings. Its frequency decreases otherwise from the center of origin, being somewhat lower in other parts of northern Europe, close to 50% in northern Italy, and 20-25% in southern Italy, Sardinia, and other parts of S. Europe. The fitness increase determined by the mutation to tolerance has been calculated on the basis of the population size of the initial population to be between around 1.5 and 4% (Bodmer and LLCS, 1976. Other recent similar estimates have used other criteria). Similar recent estimates were obtained more recently, and this is one of the few advantageous mutations whose fitness has been estimated. It is interesting to remark that the selective advantage is realistic only in an environment where milk is available to adults for consumption. The environment is a special one, generated by human innovations, and there are probably many other examples of the same type.

3e. *Genetic variation between and within populations*

Further evidence that drift has a major effect is worth mentioning. It concerns the genetic variation *between*, and that *within* populations. The variation of gene frequencies among populations is estimated by a standard analysis of variance, and can be conceived as an average of the genetic distances between all possible pairs of populations examined. The genetic variation among populations has a close, formal relationship to the genetic diversity within a population. The variance between populations was estimated on HGD data with the 650,000 snps (Li *et al.*, 2008), separately for each of the 23 chromosomes. All the 22 autosomes (chromosomes other than the sex chromosomes XY) gave a variance between populations as a fraction very close to 11.7% of the total, with extremely little variation among chromosomes (standard error of the average ± 0.11%), with the only exception of the X chromosomes, which was 15.6% +- 0.53%, and will be discussed later.

In humans, the variation between populations is smaller than that observed in practically every other Mammal, for a good reason: differences among human populations have had very little time to build up, as the evolution of the species has been very short, and the separation among human populations is quite recent. The original observation that the fraction of variance between populations is very small was originally taken as the main reason to avoid using the concept of race for the human species (Lewontin, 1975). The first estimate used by Lewontin for the variance between populations in humans was 15%, and later results were also obtained on protein data for a long time, and were very similar to this value. Races are defined as relatively homogeneous subgroups of a species, clearly distinguishable from each other. They are sharply defined in domestic animals, where breeders have much interest in keeping their breeds homogeneous and easy to recognize. But the situation is very different in humans, where it seems impossible to establish useful races. Darwin had already noted that experts have trouble reaching an agreement when they try to classify humans into races, and mentioned that in his time the number of races varied from 2 to 63, according to different accounts. We cannot do any better with genes. Attempts at distinguishing races are also encouragements to racism, a serious social disease.

Our estimate of variation between populations based on DNA, 11.7%, is even less than the 15% estimated by Lewontin, working on proteins. Most of the older data are from protein polymorphisms: the genetic unit of transmission tends to be therefore the protein, which often has more than two alleles,

being long DNA segments usually made of hundreds or thousands of nucleotides (see for instance AB0 and many other blood groups, etc.); on the other hand, single nucleotide polymorphisms analyzed in DNA sequences usually have only two alleles, partly because mutations are so rare but also for technical reasons that are not relevant here. This consideration can probably explain the higher value of the protein data, compared with DNA.

3f. *Sex and recombination*

Recombination, the reshuffling of genes that accompanies exchanges of genetic material between individuals, is another powerful source of variation, to be kept different from mutation. Genetic differences arise through recombination because new combinations of variants appear, as different mutants at different nucleotide sites come together, and thus no true DNA novelties are involved, but simply exchanges between preexisting DNA segments. Yet, by bringing together different gene types, recombination allows to test an enormous variety of combinations, from which new genetic types with predictable and unpredictable advantages can arise. Every enumeration of the new combinations of genes made possible by recombination generates numbers that are more than astronomical.

In sexual reproduction, there are exchanges between the maternal and paternal chromosomes, but every progeny gets a complete set of DNA from each parent. In the absence of sexual reproduction, all descendants of a single individual are identical, and by tracing the genealogy of individuals of an asexually reproducing group or species it is possible to reconstruct when and possibly where the mutations occurred and created different genetic types (called 'haplotypes' when they are defined on the basis of more than one mutation for a specific chromosome). We have an equivalent situation in humans for the Y chromosome, a chromosome found in a single copy and in males only, which is transmitted from father to sons. In such a case one can go back from all Y chromosomes existing today to a single ancestor, from whose Y chromosome all Y chromosomes living today descend. It is not that there ever was a single male from whom we all descend, an Adam; but Y chromosomes descending from those of other men who were living at the same time as Adam have no descendants left today. As often enough some men have no sons, and more generally the number of sons varies from individual to individual, we can always find how far back we must go before we find a single common ancestor to all Y chromosomes existing today, and how long ago he may have lived.

The same can be done for mitochondria, cytoplasmic particles descended from an ancient bacterial symbiont, found in practically all Eukaryotes (animals, plants and fungi), which are transmitted by mothers only to all their children. Mitochondria can provide information on a 'mitochondrial Eve', but here again this should not be taken as evidence that at some time there lived only one woman, but simply that the mitochondria of all of us descend from that of just one woman. If one were tempted to infer that this is proof that the Bible section on Adam and Eve was right, one would be very disappointed to learn that Adam may have lived about 125,000 years ago, and Eve 175,000 years ago.

Y chromosome and mitochondria are very useful for understanding the evolution of modern humans. But they do not have the advantage of recombination, because they stand alone and cannot mix their genome with anybody's. We reproduce sexually, like most Eukaryotes, and this gives us the full advantage of recombination for all the other chromosomes. Each of us has two specimens of each chromosome, so that every cell in our body has practically 2 x 332=46 chromosomes, that is 23 pairs of chromosomes. Twenty-two of them are called *autosomes*, the 23rd is an asymmetric pair of chromosomes, made of two members of different size, shape and gene content: X and Y, which determine sex. This condition forces males and females to perform a special trick, called *reduction* or *meiosis*, when preparing *gametes*, or cells that will fuse to generate a new individual: sperm and egg cells. A gamete contains only one chromosome of each pair. Thus every gamete has 23 chromosomes, one for each pair.

Genes on different chromosomes behave independently from each other, as Mendel found in his experiments: we usually describe this as his third law, or the *law of independent assortment of different genes*. Morgan showed that this is true for genes located on different chromosomes, as well as for genes on the same chromosome, if they are located far enough from each other, but it happens less and less the closer they are to each other on the same chromosome. The fact is that assortment is possible for genes on the same chromosome only when a phenomenon called *crossing-over* occurs, in which the paternal and maternal members of the same chromosome pair exchange a sizeable chunk of DNA, so that genes that are close to each other are more likely to cross over in bulk, switching between corresponding chromosomes.

As remarked above, the number of possible combinations that can thus arise because of independent assortment of genes is incredibly high, and this is what made sex so popular, because it multiplies enormously the possibilities that natural selection can explore. William Hamilton has strongly sup-

ported an idea expressed by others before, that the real reason why sex has become so widespread is that our major enemy are parasites, and recombination enhances our possibilities to increase our resistance to them, by combining in the same individual different ways of fighting a specific parasite (e.g. biochemical, and/or many different immunological defences).

There is a simple way to convince us that this hypothesis is very reasonable, and probably correct. Consider the history of medicine in the last 150 years, after the discovery of microbial diseases, and the progress of surgery thanks to the introduction of hygienic measures and anesthesia. Prior to this the average life expectancy at birth was only slightly greater than that which was standard for a very long time, and is still true in the most primitive conditions: about 18-20 years. Today it is close to 80 years, four times more, in developed countries. The average number of children born per family had to be at least of 6 in order to keep the population from decreasing in numbers, remaining approximately stationary in size, because about 2 out of 3 of the children died before they could reproduce. We find the number of children to be in this range among modern hunter-gatherers, who do not reap the benefits of modern medicine (but still need to not reproduce at will, because the carrying capacity of their environment keeps getting narrower). On the contrary, with the very low mortality observed today in developed countries, the number of children born per family can be just a little bit higher than 2 per family, in order to keep population numbers stationary. This happens because mortality has decreased dramatically in developed countries since medical control of infectious diseases took hold. The impact of other sicknesses, such as heart diseases and cancer, has been decreasing to a far lesser extent, but these bear less on population growth, because they occur more frequently in post-reproductive ages.

The success of modern medicine in raising life expectancy points to the fact that parasites are the major risk that *any* species encounters, and therefore the one against which natural selection is mainly directed: all mutations that increase resistance to parasites will automatically be favored, proportionately to the number of lives they spare. But recombination is more powerful than mutation in producing novelties: by rearranging genes on chromosomes and assorting combinations of different mutations it gives a faster response to needs. Natural selection is there to favor automatically those gene types or combinations that increase the probability of survival. The big impact of risks due to the parasite load in the environment indicates that Hamilton's hypothesis may be correct in detecting the major culprit that made sex so popular, at least in Eukaryotes, where a marvelous

mechanism of gamete formation makes sex so efficient as a genetic mechanism, by making a precise recombination possible.

In organisms like Bacteria, that do not have such elegant mechanisms of gamete formation, more primitive yet efficient methods of DNA transfer or exchange have spread widely. One of them, the transfer of antibiotic resistance among bacteria, is extremely efficient and is the major danger to the efficacy of the most successful avenue of medical treatments that humans have invented. Recombination made possible by sex is good for humans, and for all victims of parasites in general, but is also good for parasites and their vectors.

3g. *Sex, drift and the 134 rule*

As mentioned above, tests on all the 23 chromosome pairs for 650,000 single nucleotides showed that the 22 non-sexual pairs (called autosomes and indicated in the following as A) showed very closely similar variation between populations, with very slight variation among autosomes, 11.7% of the total variance (Li *et al.*). The same variation was definitely higher for the X chromosome, close to 15.6% (Li *et al.*),

Why is the X chromosome more variable than autosomes among populations? The difference may seem trivial, about 15.6% instead of 11.7%, but these values have been estimated on tens of thousands of genes and are therefore very precise. Considerations like these can be extended to give the *a priori* expected value of the variance between populations for the various types of chromosomes, including the Y chromosome, which is transmitted in males as if they were a population four times smaller than that of the As, and 3 times smaller than that of the Xs. The variation among populations should be like that of the averages of samples of size 4, 3, 1 for A, X, Y, and therefore proportional to the reciprocals of these values, 1/4, 1/3, 1, which can also be written in the simpler form 1:3:4. This explains why the X chromosome has greater variation among populations than the average A, exactly like the ratio of the numbers 4 and 3. 4/3 equals 1.33, and should be equal to the ratio of the variations of X and A, which are 15.6% / 11.7% = 1.37.

Unfortunately we do not have adequate Y chromosome data for the 650,000 nucleotides, which should have a variance among populations equal to four times that of Y. But there are unpublished data collected by Chiaroni *et al.*, on the major haplotypes of Y chromosome in ca. 30,000 individuals belonging to 800 indigenous populations, which give a variance between populations of 38.9% +- 2.5%. This value has a fairly large stan-

dard error, and is only slightly smaller than expected by the 134 rule (4 x 11.7% = 46.8%). The difference is significant but the Y chromosome nucleotides on which it is based are not strictly comparable to those tested for autosomes and X; there are reasons that will be explained with greater detail elsewhere why the Y chromosome variance estimate could be smaller. Also this approach, therefore, confirms that drift plays a major part in determining human genetic variation among populations.

3h. *Kimura on molecular evolution*

In 1963-4 LLCS had Motoo Kimura as a guest in Pavia for eight months, and told him of the results of the observations carried out in the Parma Valley, showing that drift was responsible for probably all of the genetic variation observed for blood groups there. At the time a number of papers was being published reporting counts of amino-acid differences among proteins of different species, which were used for reconstructing evolutionary trees of a variety of species. Kimura had developed the idea, to be proved reasonable much later, that many mutations causing amino-acid replacements have very little if any selective effects, and a few years later he published a very elegant theorem (*Nature*, 1968) thanks to which he showed, based on this hypothesis, that the rate of molecular evolution is equal to the mutation rate. Of course it is not true that all or most mutations are selectively neutral, but it is true enough that his statement cannot be shelved, after some correction. When it was published, a symposium was convened at Berkeley, where practically every geneticist in the room reacted very loudly against this dethronement of natural selection. Today we have situations, like some of those here shown, in which it is very difficult to deny a role of chance greater than natural selection at least in some situations, without any attempt to really dethrone natural selection, which is the basis on which living organisms were built and prospered.

In 1970 a book by Jacques Monod appeared, named *Le hazard et la nécessité* (a title he borrowed from Heraclitus and applied to genetic evolution). As a molecular biologist, mutation was the only source of hazard he was familiar with; but it is a very powerful one. We now must add drift in its several manifestations: one might prove that it was active even in the situations that were so useful to Darwin for convincing himself and others of the power of natural selection. Here drift, considered more generally as a consequence of population size, can be shown to be very powerful in making the effects of natural selection particularly evident: it takes a much shorter time for a use-

ful mutation to replace the ancestral type in small, isolated populations as those of the Galapagos islands than in the larger ones inhabiting large expanses and whole continents, not to mention cosmopolite species like ours.

Our analysis of this big genome evidence, which is currently proceeding, is far from complete, but it tends to confirm that natural selection has not had great effect in causing genetic variation of modern humans. The expansion of modern humans has been accompanied by adaptations to local climate and diet, part of which are genetic, but more largely are the consequences of major cultural adaptations, for instance the use of fire, clothing, housing, and more recently government, urbanization, writing, war and transportation technology, which have all helped to decrease the need for purely physical adaptations, during the process of settling the whole Earth. It is difficult to state which part of biological evolution is today under control by cultural evolution, but it must be large. Our biological evolution may have been slowed down in some aspects, and greatly ignited and/or changed in others, by our unique cultural evolution.

3i. *Proofs that all mutations are spontaneous*

An experimental procedure introduced by Joshua and Esther Lederberg in bacterial genetics, called 'replica plating', has made it possible to show that mutations easily selected in bacterial populations and that are of considerable importance for us are those that determine resistance to antibiotics and in general to antibacterial agents, and are indeed produced by spontaneous mutations. The technique consists of using standard plates filled with a medium containing the usual nutrients for bacteria in addition to agar that makes the medium solid, and use them to grow bacteria on the surface of the agar as a patina, at most a millimeter thick. Areas in the plate where a mutation for resistance to, say, the antibiotic streptomycin has arisen can be easily discovered. One takes a sample of the patina grown on a normal nutrient agar plate, by applying to the surface of the patina a piece of tissue like velvet, or of filter paper, pulling it out .and transferring a sample portion of the patina to another fresh, sterile agar plate containing streptomycin (Sm), and making sure one identifies corresponding areas on the original, Sm-free agar plate and the one with Sm. On the latter, only Sm-resistant colonies will grow, wherever there was one or more resistant bacterial mutants. Although the mutation rate to Sm resistance is very low, the patina had a sufficiently large number of bacteria that many mutations to resistance occurred during the incubation of bacteria that produced the

patina, and may have generated locally descendants that are also resistant, if mutation to resistance is a spontaneous event (all descendant bacteria from the original mutant must be resistant, as expected for a genetic mutation). This technique will work with bacteria that tend to remain where they are born and do not move around. It becomes then possible to grow in the complete absence of streptomycin 'sibs' (co-descendants) of the resistant mutant who are also resistant but have never been in contact with streptomycin. In fact one can reasonably hope to find them, as they must be located in the area of the original plate corresponding to the position where the resistant colonies grew on the Sm-plate to which the original patina was replicated. In fact one does find them, and simple sequential repetitions of this replica plating procedure allow to enrich progressively the frequency of resistant mutants thus recognized, making it possible to select strains that are made entirely of resistant mutants of the original bacterial strain, and must have arisen spontaneously because they were never in contact with the antibiotic.

This experiment proves that bacterial resistance can arise spontaneously but does not prove that *all resistant mutants* are produced spontaneously. Transforming the experiment so that it is carried out in liquid medium rather than on agar plates, one can make the experiment quantitative (Cavalli-Sforza and Lederberg; 1954) and test if all mutants are produced spontaneously. The result was positive; initially it seemed that only a fraction of mutants were spontaneous, but it was later shown that, as one might have expected, this was due to the fact that resistant mutants, like the great majority of mutants, grow a little less fast than the original strain, and even a small difference of growth rate has profound effects on the results, given the very high growth rate of bacteria.

But in patients resistant cells can grow even if they are a little slower than the original type, as long as the presence of the antibiotic in the treated patient protects them, and later mutations make easily the resistant strain more competitive. It is worth stressing that we also know that multiple bacterial resistance to many antibiotics is now spread rapidly by non sexual or para-sexual mechanisms of 'lateral' transmission of DNA segments. Unfortunately this is becoming a major threat to the conquests of medicine in the last century, which made it possible to cause the most complete disappearance as causes of death due to infectious and parasitic agents.

The experiment was repeated successfully on chemotherapeutic-resistant tumor cells using cancer cells cultivated in vitro, and demonstrated that also this major cause of therapy failure is due to spontaneous mutations to resistance of cancer cells, similar to the phenomenon in bacteria.

4. Religions and evolution

A survey of belief in evolution inside a number of developed societies (Miller *et al.*, 2006) has given surprising results. Europeans show that the frequency of people who believe in evolution varies from roughly 60% to 90%, with an approximate average around 75%. Italy is near the European average. The most unexpected result is that the lowest percentage of believers has been observed in the United States (40%), lower than in the only Islamic country surveyed, Turkey (52%).

This result seems in stark contradiction with the level of development of science and technology in the United States, which is probably greater than in any other country, but its major cause is not difficult to locate: it is the influence of the southern Baptist religion and some other less important Christian sects. These religious groups do not accept any minor deviation from the strictly literal acceptance of the Bible. The Bible has not had that downgrading effect on the people who played the major role in generating it, Jews, who are far less affected by the first sentences of Genesis. The history of the settlement of the US, into which puritans of various origins took part, helps to understand why most States of the southern USA share a wide belief that the age of the Earth cannot be older than 6000 years, as estimated on the basis of Bible genealogies and of the initial statement in Genesis that the world was created in one week. A theme park in the southern US shows scenes of children of fewer than 6000 years ago playing with dinosaurs, a tale which is passed as 'science', and as such can only help to create idiocy. The 'intelligent design' theory is an important and influential part of this trend, and was probably catapulted to public attention by the interests of political lobbies.

Almost every religion did not accept Darwin's conclusions at the time they were produced, and there was widespread outrage, as Darwin of course had anticipated and feared. The Catholic Church was no exception at the time when Darwin's work was published and until the middle of last century, but in more recent times it has been going through a wide revision of its original stance. Recently its highest authorities have formally accepted that evolution is a fact, not a hypothesis, and the 2008 meeting of the Pontifical Academy of Sciences dedicated to evolution has contributed to reinforcing this statement, although there may continue to be subtler individual variations of opinion, as might be expected.

There remain however some basic differences of importance between religious and scientific views in the interpretation of the mechanisms of

evolution. The present paper tries to show that basic differences that are still common can be removed simply by more precise explanations.

Natural selection is the only evolutionary mechanism that generates automatic adaptation and is, in a sense, strongly deterministic in this direction. Practically all other evolutionary factors do not necessary help or oppose adaptation, and all contain elements that could be called 'chance'. In fact the findings of our research show that factors that can be described as chance are often quantitatively more important than natural selection in shaping our genome. This is still a cause of disagreement among geneticists, although the importance of chance is gaining support; and obviously of major disagreement with the very few scientists who are still fond of the Genesis 6000 years date. One reason to dislike the influence of chance, especially in some religious circles, seems to be the strength of admiration towards a hypothetical 'biological order'.

It should be more widely realized that often chance is introduced as the scientific way of treating situations in which the causal system is too complicated to be analyzed in detail, i.e. when it is complex enough to defy our descriptive skills. In this case use is made of statistical approaches that are known to be potentially of aid precisely when the causal system is too complicated to be tackled in detail, i.e. when there are too many causes that interact in producing the phenomena being studied. Probability calculus teaches, by well-known theorems, that in these situations continuous probability distributions, like e.g. the normal or Gaussian and the lognormal, may be useful. Statistical correlation methods can sometimes help in disentangling causes and effects, although experience shows they must be used with real caution, especially in human genetics, as exemplified in the classical case of the Intelligence Quotient (L.L. and F. Cavalli-Sforza, 1995).

Ignorance of causes is not an issue when chance is built into the specific phenomena under study by *random sampling*. Mendel knew that when he studied segregations of characters in crosses he had to look at large numbers of individuals, in order to beat irregularities generated by the random sampling process, and find the laws he eventually did find. He made a few mistakes that led him to overcorrect his data, as Fisher showed (1936), but they generated no mistakes in his major conclusions. There cannot be any question that when natural populations or experimental sample sizes are small we are going to find, on average, greater random oscillations in evolutionary processes due to genetic drift, perfectly predictable by probability calculus. We should not become unhappy or suspicious if in these cases chance takes its toll and may generate superficially strange results. Drift

may be defined simply as random samplings of gene frequencies accumulated over generations. It seems that even the fact that mutations are random (although there is a small chance, never really proved so far, that some – certainly very few – mutations may have a partly adaptive origin) should not trouble the minds of theologians.

Scientists are aware that ideologues do not accept scientifically ascertained facts when they are contrary to their favored beliefs. For this reason it is safe for scientists to refrain from political or religious ideology. It is necessary to keep science anchored to facts that can be observed with our senses (the world of nature), and to the search for rational explanations of them. The scientific way of proceeding democratically is the major guarantee of rationality. But scientists must stick to the reality of nature; they would betray science if they accepted supernatural explanations, which contain unverifiable hypotheses. Science cannot deal with supernatural facts, because they cannot be reproduced at will.

Ambitions, greed, prejudice, jealousy, dishonesty, dangerous ideologies (e.g. Lysenko's attempt at destroying Mendelism on the basis of Marxism principles) occasionally take the hand also in science. Still, there is a very good chance that sooner or later – maybe some time in the future, maybe after our death – truth will be recognized because of new, better experiments or simply because of a stricter use of logic by scientists.

One reason why some may consider chance a nuisance is that it seems to detract from, or even to destroy the idea that there is 'biological order', and other closely related assumptions which have a definite teleological flavor. It should be clear that it is better to avoid this kind of simplistic thinking that may easily invoke unnecessary supernatural explanations. Scientists can only try to interpret natural phenomena without recourse to supernatural causes, and nothing in biology has so far requested to resort to them, when enough time is dedicated to a problem. Louis Pasteur, to whom we owe so much in microbiology and medicine, and who was also a very devout believer, found himself unable to isolate chemically the enzymes active in the fermentations he had discovered because he found no ways of opening cells without destroying the enzymes they contain, and came to the conclusion that enzymes were created anew every time. This would positively have kept God's deputies very busy. But after Pasteur's death German chemists were able to develop subtler chemical methods of purifying enzymes and studying their structure.

What about the idea of 'biological order'? Is it really destroyed if we postulate that a lot of biological evolution takes place by chance? More than of

'order', when we marvel at the degree of perfection of certain organs and functions, e.g. of our eye, one should speak of 'biological efficiency'. Incidentally, our eye is a poor thing compared with the eye of most birds. And the organization of many organs and systems is far from perfect, in any species. The immune system, for instance, is a magnificent biological accomplishment that uses a new Darwinian structure, independent from our general development but operating inside us, for producing with special mechanisms of 'mutation' and natural selection new, specific antibodies against the parasites that attack us. But the system is not perfect and errors give rise to diseases (e.g. autoimmunity) that need medical help. Any biological mechanism has sufficient faults and imperfections that are hardly proof of divine intervention in generating them, as is perhaps in the intention of admirers of biological 'order'.

One of the best biologists of the XX century, François Jacob, together with another great scientist, Jacques Monod, discovered the mechanism whereby bacteria can produce a specific enzyme (e.g. lactase, that utilizes the sugar lactose) only when necessary, that is, only when the substance that the enzyme attacks, lactose, is present in the medium. After this breakthrough, other methods of regulating enzyme presence or action have been discovered (these enzymes are called 'inducible') while other enzymes are always present (and are called 'constitutive'). Inducible enzymes allow to spare bacterial energy and activity, and are useful especially if the enzyme substrate is seldom present, but require the ability to 'sense' the presence of the enzyme substrate in the medium – a primitive step towards rational organization of behavior. Jacob described the biochemical mechanisms he and Monod discovered as examples of 'bricolage' – do-it-yourself mechanisms that are assembled by using new tricks or old bits of machinery already available inside the organism, redirecting them to the new jobs. Usually this happens by exploiting new mutational changes that, if proved helpful, will be propagated by natural selection and can be improved further in many ways by new mutations. After a long series of improvements these mechanisms become rather efficient: the process by which efficiency is thus achieved is called simply 'trial and error', and we ourselves practice it many times when we busy ourselves with bricolage at home, to solve simple problems, usually of mechanical or electric nature.

Bricolage occurs all the time also in biological evolution, and not only in cultural evolution, where the name first arose, and where new ideas, small or big, have the same function as genetic mutations in biological

evolution. Again, in cultural evolution our innate and acquired tastes, which form our personality, affect the choice of new ideas, and we call the acceptance/rejection process of new ideas *cultural selection*, a clear analogue of natural selection. But the inventions and choices made by cultural selection are still subject to a higher check: and this is, of course, natural selection, which can destroy few individual lives when we accept excessive risks (e.g. drug overdoses, or houses falling down on the careless builders), or many lives, even the whole human species and many other living organisms (e.g. with the worst of all cultural choices: that of starting a major international nuclear war).

Confronted by the extraordinary examples of biological structure and function, many prefer to accept the idea of a direct intervention by God, for whom it must have been simple to create from scratch an apparently intelligent mechanism that works beautifully. But unless we try to understand the real mechanism, with all the complications that nature has put into it by its bricolage, we will not be able to repair its malfunctions: then we will give up medicine. This suggests to be critical of excessive admiration of biological order.

Probably the idea of biological order was a wrong impression generated by early taxonomists like Linnaeus, who first generated kingdoms and phyla, classes, orders, families, genera and species of living beings, all beautifully organized in a perfect hierarchy, reflecting original creation, of course, and therefore believed to be immutable. The reality is different: today. With a better knowledge of DNA, it has become impossible to build perfect hierarchies, and specialists disagree as strongly as ever, especially for the lower organisms But at least we understand why there are no perfect hierarchies: there has been a fair amount of 'lateral transfer', that is, acquisition of pieces of DNA, or whole sets of them, from other totally unrelated organisms. Thus some small organisms, which parasitized much larger ones at first, later probably became symbionts. Having become a forced and indispensable part of their hosts, they have lost their independence and even their identity, bu we cannot do without them. The two clearest examples are: mitochondria, that take care of a major part of energy production from simple sugars for all animals, plants and smaller Eukaryotes; and chloroplasts, that have the task of catching sun's energy to build substances that make plant and animal life possible. In spite of these difficulties generated by a complicated history, it is clear that analysis at the genome level is making the study of evolution an exact science.

4a. *What is chance, after all?*

While natural selection tends to always increase adaptation, mutations and other factors introduce strong random effects, which may also be called, with slightly different connotations, hazard or chance. We prefer the latter term: in Italian, the word for chance is 'caso', which has a similar origin. The English word 'chance' stems straight from the same French word, 'chance' (in old French this was 'cheoir', derived from Latin 'cadere', 'to fall', with the same origin as 'caso'. Hence the Italian word 'accadere', to happen, which is perhaps related to 'hazard').

Before we come to a full understanding of the relative importance of natural selection and chance in evolution, we should discuss the concept of chance further. Mathematically, the introduction of chance brings us directly to the probability calculus.

We have seen three major evolutionary factors which bring chance into evolution: mutation, recombination, and drift. To these, we may add cosmic events: for instance, we know that about every many million years a huge meteorite is likely to hit the Earth; this has apparently taken place a number of times in the history of life, with dramatic impacts on the course of evolution. Though the likelihood of these events can sometime be measured, there is no way to tell when the next one will take place or what developments will take place as a consequence.

We have two ways of dealing with the occurrence of chance in evolution. One way is that all phenomena that are determined by the interaction of a large number of causes, none of which is clearly identifiable, can still be brought to rational analysis (i.e. mathematically, by probability calculus). The second, more direct way is when we count numbers of individuals showing different characters. We then have 'sampling' problems, whereby results in terms of 'counts' of individuals will change unavoidably almost every time we repeat the same experiment. Here again: probability calculus gives clear, helpful predictions of sampling problems. In fact, random genetic drift is essentially a 'random sampling' problem, built into the way organisms produce the next generation. The sampling nature of reproducers who generate successive generations giving rise to drift is a classical statistical problem, complicated by the fact that the sampling effects accumulate over the successive generations: the difficulty is handled by mathematical methods dealing with 'stochastic processes', which were developed largely for dealing with genetic problems.

One can also describe the effects of chance by older statistical methods, for instance correlation between different variables: for instance, the strength

of inheritance by comparing the value of specific characters in parents and children. This can be done for qualitative traits like those chosen by Gregor Mendel, which were due to changes in individual genetic units, but were sharp enough to be defined by alternative adjectives, such as green vs. yellow for seeds, or tall vs. small for major differences in plant height. We are again struggling with sampling errors. Or we may struggle with variation in measured ('quantitative') traits like stature or any other anthropometrical trait, such as were chosen by Darwin's cousin, Francis Galton. In order to study the inheritance of quantitative traits, Galton and a statistician, Karl Pearson, developed methods that did not survive criticism by Fisher, who generated a large number of modern statistical methods, and also solved in his 1918 paper the problem of treating the inheritance of quantitative traits by incorporating Galton's approach into standard Mendelism.

4b. *It is probably a good thing that mutation is random*

It seems that not only scientists should take interest in the hypothesis that mutations are basically random. Any thinker dedicated to finding rational designs in the construction of the Universe should appreciate the idea that mutations are random events. By being random, mutation gives similar chances of being beneficial or not to all species. It is thus fair, giving equal chance of success to different species competing with each other, and to individuals of the same species. It is intuitively conceivable (but certainly difficult to prove), that this 'universal' democracy established by the randomness of mutations tends to prolong equally the probability of survival of all species and individuals. Thus, it also may give more stability to the system of all living organisms, that involves many millions of species. Species interact competitively but also need each other, and their numbers may vary greatly over time and space. And yet every species needs so many other species for its own survival that there is likely that there is a condition of general stability, permitting a slow, overall increase with time of biological mass as well as of general complexity.

In any case, there is evidence that mutation rates are under control by natural selection, and that at times when survival of a species is difficult mutation rates tend to increase. This is probably again an automatic reaction generated by natural selection: if mutants are favored by changes in environment and there is genetic variation of mutation rates in the population, increased selection of mutants may also automatically increase mutation rates because at least some of the mutants will have arisen in individ-

uals that are genetically predisposed to a higher mutation rate. This indicates another way in which natural selection may contribute to increasing adaptation. It is encouraging that the greatly improved possibilities of studying whole genomes will increase the chances of studying more accurately also mutation rates and their natural selection.

BIBLIOGRAPHY

Cavalli-Sforza, L.L. (2005) 'The Human Genome Diversity Project: Past, Present and Future'. *Nature Reviews/Genetics*. 6 (4): 333-40.

Cavalli-Sforza, L.L. and F. Cavalli-Sforza (1995), Postface of '*The Great Human Diasporas*'. Addison-Wesley Publishing Co. Reading, MS.

Cavalli-Sforza, L.L. and M.W. Feldman (2003) The application of molecular genetic approaches to the study of human evolution. *Nature Genetics* 33:266-75

Cavalli-Sforza, L.L. and G. Zei (1967) 'Experiments with an artificial population', Proceedings III International Congress Human Genetics, Johns Hopkins Press, Baltimore, Maryland, pp. 473-478.

Cavalli-Sforza, L.L., A. Moroni and G. Zei, (2004) *Consanguinity, Inbreeding and Genetic Drift in Italy*, Princeton University Press, Princeton, NJ.

Fisher, R.A. (1918) 'The Correlation between Relatives on the Supposition of Mendelian Inheritance'. *Trans. Roy. Soc. Edin*., 52:408-433.

Fisher, R.A. (1930) *The Genetical Theory of Natural Selection*. Oxford University Press.

Fisher, R.A. (1936) 'Has Mendel's work been rediscovered?'. *Annals of Science* 1: 115-137.

Jacob, F. (1981) *Le jeu des possibles*. Fayard, Paris.

Li, J.Z., Absher D.M., Tang Hua, Southwick, A.M., Casto, A.M., Ramachandran, S., Cann, H.M., Barsch, G.S., Feldman, M.W., Cavalli-Sforza, L.L., Myers, R.M. (2008) 'Worldwide Human Relationships Inferred from Genome-Wide Patterns of Variation'. *Science* 319: 1100-1104.

Miller, J., E.C. Scott and S. Okamoto (2006) A survey of belief in evolution. *Science* 313: 765.

Monod, Jacques (1970) Le Hasard et la Nécessité. Editions du Seuil, Paris.

Prugnolle, F., A. Manica and F. Balloux (2005) Geography predicts neutral genetic diversity of human populations. *Curr. Biol*. 15: 159-160.

Ramachandran, S., Deshpande, O., Roseman, C.C., Rosenberg, N.A., Feldman, M.W., Cavalli-Sforza, L.L. (2005) 'Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa'. *PNAS* 102:15942-7.

DISCUSSION ON PROF. CAVALLI-SFORZA'S PAPER

PROF. BERNARDI: Luca, I don't think there is any contradiction between your results and ours. We are looking at two very different scales. Looking at vertebrates means looking at five hundred million years of evolution. You are looking, of course, at human populations over a short span. You see the small details which we do not see at all on our much wider time scale. So I do not think there is any contradiction between the results.

PROF. CAVALLI-SFORZA: No, as a matter of fact I would also like to add that, when you say that all that is due to selection, in many cases there is really no proof that there is direct evidence of adaptation.

PROF. BERNARDI: Well, of course, I could not present everything but the evidence for evolution comes from the conservation which you can only assume to be the result of the elimination by negative selection of the changes, and I do not see any other way to get rid of those changes other than by negative selection.

PROF. CAVALLI-SFORZA: I'm not saying that those changes are not due to selection, I say that most frequently strong evidence that a particular genetic change is due to natural selection is difficult to acquire.[1]

[1] I have requested the Editors to add here, in the proofs, the following short note which is very relevant to the issue, and summarizes and extends somewhat the argument I tried to present in my main talk. Strong evidence of selection for a genetically determined trait requires proof that positively, or negatively selected traits really show increased, or decreased darwinian fitness of their carriers. This requires demographic estimates of fitness by analysis of survival and fecundity, a very difficult or impossible task in many situations, especially for mutants responsible for positive selection. Usually much evidence of selection is simply from correlations of phenotype and environment, as is the case, for instance, of body shape or size with climate. To put it simply, the nature of the trait suggests reasonable ideas for its adaptiveness to certain environments. But even if the intuitive answer derived from the nature of the trait seems unobjectionable, as for instance that wings are essential for flying, a closer analysis indicates that the development of wings came after a preliminary stage

PROF. GOJOBORI: Luca, could you give us a comment on *Neanderthal* genomes. I think Svante Paabo is now conducting sequencing of *Neanderthal* genomes, not only a single individual, maybe more than one, of course: then how can that information make an impact on your study? Obviously you are working on a gene frequency tree.

PROF. CAVALLI-SFORZA: All I know about *Neanderthal* is that there is no evidence that there has been any exchange but probably it may come if some individuals are found who may have been hybrid and so far there are none that have been found or tested. So it may well be that there was some exchange but, even if we find some hybrids, we still do not know whether we should consider them a separate species or not, because separate species, strictly speaking, requires that hybrids are not fertile, to be rigorous, but it is very difficult to reach that kind of rigour in many situations. But anyhow, I think, I don't know if *Neanderthal* developed a language and up to which point, but they probably did because I don't believe that language completely developed in the last hundred thousand years. There is evidence that the left part of the brain is developed in five out of six skulls that are 1.7 to 2 million years old, so it is likely that there was some early development of language and what happened more recently was only reaching the level of perfection that it has reached, like having syntax and so on. So I do not really think I can say more than that.

in which the beginnings were accidental or directed to other purposes, as was suggested in well known work by Lewontin. I would really like to know the global amount of selection that may have gone into evolution, and how it compares with other factors of evolution that were more or less entirely neglected, and this is more difficult to estimate. Our results on human evolution indicate that the input of natural selection in the recent peopling of the world by *H. sapiens sapiens* may have been much smaller than chance effects due to random genetic drift than one might expect, but they are limited to the intra-specific variation of one species, and this may be less than that found in inter-specific variation. Moreover our species is unique in several ways: In humans another very powerful general mechanism of adaptation exists beside natural selection, cultural evolution, which is not designed uniquely to improve the *species* survival, but depends to some extent on our whims. In human history there have been many catastrophes caused by hard follies of few individuals who gathered enormous power in their hands, and our society has not yet become able to protect itself from such risks. But leaving them aside, another consideration regarding natural selection in our species is that our species is in the nearly unique situation of solving many practical problems due to climate, food, or health etc. by cultural evolution, which is much faster than natural selection. This has probably lessened considerably the relative and absolute contribution of natural selection to the genetic variation we experienced, for instance, during the recent expansion of *Homo sapiens sapiens* to the whole world.

THE FUTURE OF LIFE

CHRISTIAN DE DUVE

In the two most recent plenary sessions of this Academy that I was able to attend, in 1996, when Pope John Paul II made his celebrated declaration: '*Evolution is more than a hypothesis*', and in 2002, when the present Pope, who was still Cardinal Ratzinger at the time, presented himself as a new member of the Academy, I have expressed some thoughts on the nature, origin, and evolution of life (de Duve, 1997, 2003). Today, in what is most likely my last participation in a meeting of the Academy, I will attempt to take a brief look into the future of life on Earth, as illuminated by our knowledge of its past and present. This topic is discussed in greater detail in a coming book (de Duve, 2009).

First, let me say a few words about the past. Life appeared on Earth at least 3.55 billion years ago, fairly soon after our newborn planet had become physically able to support it. Inaugurated by primitive cells of unknown origin, life remained unicellular for some 2.5 billion years, first in the form exclusively of prokaryotes (bacteria), to which, about 1.5 billion years later, were added the protists. These consist of much larger and more complex cells called eukaryotic and containing a nucleus, an elaborate membrane network, intricate cytoskeletal structures, and several cytoplasmic organelles, including lysosomes, peroxisomes, mitochondria, and, in photosynthetic organisms, chloroplasts. Many representatives of these microbes, both prokaryotic and eukaryotic, still abound in the world today.

Only about one billion years ago did eukaryotic protists first give rise to multicellular organisms. Plants led the way, soon followed by fungi and, 400 million years later, by the first animals. These started by blossoming into the rich world of invertebrates, of which one group eventually evolved into the first marine vertebrates, the fish, which, in turn, gave rise to the partly land-adapted amphibians, followed later by the fully land-adapted reptiles, from which arose birds, on one hand, and mammals, on the other.

Primates arose among the mammals some 70 million years ago, evolving to produce, in addition to a variety of apes and monkeys, a line, initiated some 6-7 million years ago, that led to the human species. Note the extreme lateness of this crucial event, which took place in the last 100th part of animal evolution, the last 600th part of the evolution of life.

The advent of humankind was signalled by several important acquisitions, including bipedalism, increased handiness, and, especially, a larger and more complex brain, which, in little more than two million years, almost quadrupled in size, from a volume of about 350 cm^3, the size of the brain of present-day chimpanzees, our closest relatives, to a volume of some 1,350 cm^3.

These acquisitions have allowed a fantastic evolutionary success, without equivalent in the entire history of life on Earth. Our early ancestors numbered about 3,000 when they separated from the Neanderthals, at a time estimated from recent DNA studies to lie between 800,000 and 500,000 years ago. There were about 10,000 of them 200,000 years ago, when 'mitochondrial Eve' and 'Y Adam' started *Homo sapiens sapiens* on its final evolutionary journey. They may have been on the order of 5-10 million, scattered over a good part of the world, when the first durable human settlements were created some 10,000 years ago. Since then, the human population has grown at an ever-increasing pace, reaching about half-a-billion in the time of Galileo, passing the one-billion at the start of the nineteenth century, and rising from less than 2 billion to more than 6.5 billion just in the last 100 years, coming to invade, occupy and exploit almost every habitable – or, even, uninhabitable – site on our planet. Ours is, by far, the most successful species – I leave out microbes – in the whole of biological evolution.

This success has a cost, briefly summarized in Table 1. We read or hear about it almost daily through the media. It is known to all of us and I need hardly elaborate. What I wish to do is extrapolate from the past and present to the future. If things continue in the same direction, there is little doubt that we are heading for disaster, soon to reach a point where we will be driven to extinction, together with a good part of the living world. If this happens, it will be nothing new in the history of life, including the recent history of humankind. These histories are landmarked by extinctions. But there will be a difference. Most likely, past extinctions were invariably associated with some kind of *failure* in the face of an external challenge (drought, glaciation, or other climate change, geological upheavals, meteorite impacts, epidemics, extermination by a more successful competitor, etc.). Our extinction, if it occurs, will be the consequence of inordinate evo-

TABLE 1.

The Cost of Success

1. Exhaustion of natural resources
2. Loss of biodiversity
3. Deforestation and desertification
4. Climate change
5. Energy crisis
6. Pollution
7. Overcrowded cities
8. Conflicts and wars

SUMMARY: IRRESPONSIBLE EXPANSION

lutionary *success*. We have developed to the point of endangering the ability of our planet to support us. If we go on following the same course, it can only lead to our doom.

Contemplating this ominous picture with the eyes of a biologist, I find a single culprit: *natural selection*. I use the word 'culprit' metaphorically – no guilt is involved – but, as will be seen later, the image is not entirely inappropriate. Natural selection is the process, now overwhelmingly established as a dominant evolutionary mechanism, whereby the forms of life that are most apt to survive and produce progeny under prevailing conditions obligatorily emerge from whatever set of organisms happen to compete for the same resources. All that is known about this process indicates that the variants on which it operates are accidentally generated, without intentionality or guidance, contrary to what is claimed by the defenders of intelligent design. Another key feature of natural selection, of special importance for our topic, is that it is governed entirely by *immediate* benefits. Natural selection has no foresight.

There is every reason to assume that humans are, biologically, products of natural selection, like all other forms of life. This implies that evolution has privileged in human genes traits that were immediately favorable to the survival and proliferation of our ancestors under the conditions that obtained there and then, regardless of later consequences. This is intrinsic to the process of natural selection. Note that I leave out traits that were

acquired by cultural evolution and transmitted by education. I shall turn to these later. Right now, I will deal only with genetically inscribed traits.

On an individual basis, human traits retained by natural selection included intelligence, inventiveness, dexterity, skillfulness, resourcefulness, and ability to communicate, all the qualities that have served to generate the fantastic scientific and technological achievements responsible for our evolutionary success. But the selected traits also included acquisitiveness, selfishness, greed, cunning, aggressivity, and any other property that ensured immediate personal gain, regardless of later cost to oneself or to others. The recent financial crisis has illustrated in a particularly dramatic fashion how such traits still flourish in the world today. On the other hand, genetic qualities whose benefits would become manifest only in the long run, such as far-sightedness, prudence, a sense of responsibility, and wisdom, were not singled out by natural selection. Their fruits would have appeared too late for that.

On a collective level, natural selection has favored traits, such as solidarity, helpfulness, cooperativity, tolerance, empathy, compassion, altruism, even personal sacrifice for the common good, that form the bases of human societies. But the selection of those traits has been mostly restricted to the members of given *groups*, united first by shared kinships and territories, and later by shared interests, a shared language and culture, shared beliefs, shared prejudices, even shared hatreds. The negative counterpart of those 'good' traits has been collective defensiveness, distrust, competitiveness, and hostility against members of other groups, the seeds of the conflicts and wars that have landmarked the whole of human history up to the present day.

In other words, the defects that endanger the future of our species and of much of the living world are *inborn*, written and sustained in our genes by natural selection. They were useful in the past, at a certain stage of our evolution but have become deleterious; they are a natural burden we assume at birth. I would like to suggest that awareness of these innate genetic defects inspired the notion of *original sin*. That is why calling natural selection the 'culprit', as I did earlier, is not entirely inappropriate, except, of course, that no culpability is involved. There is no Eve to blame, no serpent, only natural selection, which is mindless and without intention, devoid of foresight and responsibility.

Is there anything we can do? Fortunately, yes. Of all living beings on Earth, we humans are the only ones that are not slavishly subject to natural selection. Thanks to our superior brains, we have the ability to look into the future and to reason, decide, and act in the light of our predictions and

expectations, if need be against our immediate interest, for the benefit of a later good. We enjoy the unique faculty of being able to act *against natural selection*. The problem is that, in order to do this, we must actively oppose some of our key genetic traits, surmount our own nature.

It would be nice if we could correct our genetic defects by engineering, removing the bad genes and implanting good ones. We can do this to a limited extent with plants and animals; but we cannot possibly do the same with humans. We do not yet have a sufficiently reliable technology for human application. Even if we had it, we would not know what genes to modify in order to achieve a certain goal. Our knowledge of the genetic basis of psychological traits is still in its infancy. Even if we had this knowledge, there would be the problem of deciding who should benefit from the interventions. Finally, there are all the ethical objections such manipulations are likely to raise. So I won't waste time discussing this way out of our predicament. We are not ready, whether scientifically, technologically, or ethically, to create GMHs, genetically modified humans, for specified aims, as we do other organisms.

But there is another way out, provided by the fact that the structure of the human brain is genetically determined only in its general architecture. Its fine wiring takes place *epigenetically*, under the influence of the various stimuli to which the brain is subjected. Note that I use the word 'epigenetic' in its original meaning of 'added to the genetic', a meaning given to it by the developmental biologists who invented it and still used by neurobiologists; not in its new meaning of 'genetic, but not inscribed in DNA sequences', now accepted by many geneticists and molecular biologists.

It is known, from the work of Gerald Edelman, in the United States, and of Jean-Pierre Changeux, in France, that, in the developing brain, growing neurons continually send out extensions in various directions. Upon chance encounters between such extensions, the neurons form temporary connections, which are rapidly undone unless they happen to be repeatedly used, in which case they become stabilized as synapses. Thus, the stimuli to which the growing brain is subjected operate some kind of selection among the many interneuronal connections that are created by chance. The similarity with Darwinian selection has not escaped the authors. Edelman, for example, speaks of 'neural darwinism'.

Thus, the wiring of a human brain, which forms the underlying substrate of the thoughts, feelings and other mental processes the brain can experience, is largely determined by the impulses conveyed to it by the external stimuli to which the body is exposed. In a way, this has always

been known by all those who have had something to do with educating the young. Educators have always been aware of the importance of their work in the 'molding' of young brains. What is new is the realization that this process starts at birth, perhaps even before birth, and that parents, nannies, nurses, kindergarten personnel, elementary school teachers, baby sitters, that is, all those who deal with very young children, exert key influences on the wiring of the children's brains.

Thus, if we wish to create in young brains neuronal networks conducive to tolerance, sympathy, peacefulness, reasonableness, foresight, and wisdom, we must first do so with the parents and educators. Doing so in one shot is clearly impossible, but one can imagine initiating a self-enhancing movement that would progressively snowball into becoming worldwide. But for this to happen, the movement must be set in motion.

This brings me to my final message, of special significance within these walls. It concerns the role of *religions*. Historically, religions have always played a major role in the education of the young, even of adults. Even today, their influence in this domain remains tremendous. Religious leaders are, even more than the most powerful political leaders, uniquely placed to influence large crowds. When the Pope speaks, he reaches more than one billion individuals. Thus, he and the leaders of the other major religions are invested with an immense planetary responsibility. They are almost the only persons in the world who could play a decisive role in rescuing humanity from its suicidal course. They are particularly well placed to do this, in view of the millennia-old tradition of tolerance, love, and understanding that, originally, has been the main message propagated by the major religions.

Unfortunately, Churches have not escaped the genetic 'original sin' that plagues the whole of humanity. One cannot generalize, of course. There are important differences among the various religions. But each is, to a greater or lesser extent, tainted with authoritarianism, fundamentalism, doctrinal dogmatism, ethical rigidity, exclusiveness, extending, in some cases, to nationalism and strife, sometimes armed, even murderous.

The Catholic Church is not exempt from these defects. I hope that this statement, expressed within these venerable walls, will not be seen as disrespectful or unsuitable. This Academy was created to promote the free intercourse of ideas, within a framework of open-mindedness, intellectual honesty, and sincerity. With your permission and with apologies to those who disapprove, I will avail myself of this spirit, which corresponds to the true scientific attitude.

In my opinion, it is our duty, as members of this august body, to alert the higher Authorities to the extreme gravity of the menaces that weigh on the future of humanity and of planet Earth and to the urgent necessity of acting against those threats by all possible means. The facts (see Table 1) speak for themselves. They are evident and undeniable. We ignore them at our peril. The final outcome, if nothing is done to change the course of events, leaves little doubt. The Church, with its unique worldwide power and influence, bears an enormous responsibility in directing this course for better. If there is agreement on this point, this is a message our Academy, as advisor to the Holy See, could respectfully convey to the Magisterium.

PUBLICATIONS

C. de Duve (1997). *Life as a Cosmic Imperative*. Commentarii, Vol. IV, N° 3, pp. 311-320. Plenary Session on the Origin and Early Evolution of Life (22-26 October, 1996).

C. de Duve (2003). *The Facts of Life*. Scripta Varia, Vol. 105, pp. 71-91. Plenary Session on the Cultural Values of Science (8-11 November, 2002).

C. de Duve (2009). *The Future of Life on Earth* (Provisional title, to be published).

DISCUSSION ON PROF. DE DUVE'S PAPER

PROF. POTRYKUS: To my understanding of the real public power, you should address your call to the media, not to the Church.

PROF. DE DUVE: Well, one is not exclusive of the other. I think the Church has an enormous power and could be extremely helpful.

PROF. PHILLIPS: I hope you will understand that I mean this as a compliment, but you have given us more of a sermon than a talk, but I think it is a sermon that is well taken. It resonates with many of the thoughts that I have had and this is not the first time I have heard the idea that inappropriately evolved behaviour is the same as original sin. The thing that I am wondering about is that the message that you have preached to us seems like it is the same message that has been preached by religious leaders since the time of Moses, Buddha, Confucius and various other people. So how do you intend to convince people that there is something more urgent now than has been the case in all the times past when people made the same urgent pleas that you have made today?

PROF. DE DUVE: Well, let me make a few points. First of all, I am too old to start trying to convince people. I think people should just look at the evidence – it's there for everyone to see – and draw their own conclusions. The other point is, you say that I delivered a sermon, which is perhaps not a bad idea in these surroundings, but you have to remember that this, to me, has been my farewell speech. I am 91. This is, certainly, the last plenary session that I will attend. I felt that, as a farewell speech, I would, within these venerable walls, speak my own mind.

PROF. PHILLIPS: And I thank you for it.

THE EVOLUTIONARY ORIGIN AND PROCESS OF THE CENTRAL NERVOUS SYSTEM: COMPARATIVE GENOMICS APPROACH

TAKASHI GOJOBORI[1,2], KAZUHO IKEO[1,2] AND JUNG SHAN HWANG[1]

INTRODUCTION

Historically, it had been the most essential question to ask why we think we are ourselves. A famous phrase written by a French philosopher, René Descartes, in Latin '*Cogito, ergo sum*' (translated in English, 'I think, therefore I am'). This paradigmatic enigma given by Descartes about self-consciousness or self-recognition is still the central question for us today. Apparently there is no easy answer for this question from any perspective. However, we can question and resolve the biological problem of 'how our brains work and think' from the evolutionary standpoints, especially when we are able to study it at the genetic level.

Since the draft sequence of human genome project has been completed, biologists have focused on the post-genomics studies including proteomics, transcriptomics (gene expression profile), SNP (single nucleotide polymorphism), non-coding RNAs (eg. miRNA, siRNA), comparative genomics and etc. Among these studies, comparative genomics provide a powerful way to resolve the evolutionary questions (Koonin *et al.*, 2000). Sequence comparison across the species is a fundamental solution to understand the origin, the evolutionary differences between organisms and the complexity of biological systems. Particularly the currently advanced technology of second-

[1] Center for Information Biology and DNA Data Bank of Japan (DDBJ), National Institute of Genetics, Mishima 411-8540, Japan

[2] The Graduate University for Advanced Studies (SOKENDAI), Shonan Village, Hayama, Kanagawa 240-0193, Japan

generation sequencers such as 454, Solexa, SOLiD and Helicos allows one thousand human genomes to be sequenced just within two years. Therefore, we expect to have the number of genome sequences of various organisms increase significantly in the next few years. Taking advantage of comparative genomics, we attempt to understand the evolution of the central nervous system (CNS) or the brain of humans at the gene level. To do this, we address the following questions: (1) What is the origin of the nervous system (NS) as based on currently available sequence data and the advantage of homology search? (2) How old are the nervous system genes especially those that are also expressed in the human brain? (3) What kinds of genes are expressed in planarian, a primitive flatworm having the simplest brain?

Origin of nervous system – a nerve net

There are around 100 billion neurons in an adult human brain and they can be categorized according to the number of axonal processes extended from the perikaryon (the cell body). Depending on their localization in brain regions, these basic neuron types are further subdivided into specialized neuron types and also functionally diverse. Let us take pyramidal cells as an example, which are typically characterized by a spiny apical dendrite, a basal dendrite and a single axon. The morphologically identical populations of pyramidal cells are found largely in two distinct functional parts of the brain, neocortex and hippocampus. Further diversification of pyramidal cells can be found within the hippocampus in which a heterogeneous expression of genes is observed across the pyramidal cell layer (Lein *et al.*, 2007). The ancestor of the human brain is considered much simpler. It is believed to have a two-dimensional neural network that is somehow similar to the diffuse nerve net of basal phylum Cnidaria, having an average of less than 8000 neurons, no glial cells, no centralized nerve tissue and no anatomical compartmentalization (Holland, 2003, Telford, 2007) (Fig. 1, see p. 608). However, the neuroanatomical comparison has failed to show any homologous structures of the nervous system between human and Cnidaria, nor has the cell morphology given any clue due to the simplicity of cell types in cnidarians. Recently, gene expression data have proven the conserved body plan between vertebrates and cnidarians (Bode, 2001; Finnerty *et al.*, 2004; Kusserow *et al.*, 2005; Lengfeld *et al.*, 2009), suggesting that the origin of the body plan can be dated to the early Metazoa.

The phylum Cnidaria includes animals such as coral, sea anemone, sea pen, jellyfish and *Hydra* and they all share a sac-like body surrounded by two layers of epithelial cells (ectoderm and endoderm). They only have a single opening that functions as both mouth and anus. Cnidarians have the simplest nervous system named nerve net in which thousands of neurons make up a mesh-like network at both epithelial layers (Fig. 1 and 2A, see pp. 608-9). Jellyfish and certain species of *Hydra* (eg. *H. Oligactis*) also have neurons concentrate and form a ring around the mouth region (Fig. 1, see p. 608). This nerve ring is considered to be an intermediate structure between a nerve net and a ganglion. Since Cnidaria and Ctenophore are basal metazoans with the nervous system, many expect the origin of the nervous system to somehow resemble the cnidarian nerve net. From time to time hypotheses have been made to explain how a simple nerve net of cnidarians evolves into a bilaterian nervous system (Holland, 2003). Most ideas are based on the scenario proposed for the evolution of the bilateral body plan (Lacalli, 1995; Meinhardt, 2002; Holland, 2000; Martindale, 2005; Hejnol and Martindale, 2008). Holland (2003) has a good summary of all the hypotheses on the transformation of nerve net/nerve ring into brain and nerve cord. Among all, the majority of scenarios believe that Cnidaria has the ancestor-like nervous system and the bilaterian brain is originated from the nerve ring. Other minor scenarios either consider the whole cnidarian polyp as a brain or the nerve net is compressed to one side of the body axis and becomes the brain.

Does nervous system emerge from nowhere?

Yet, could the ancestor of the nervous system be simpler than a nerve net? The nerve cell (or neuron) is the fundamental unit of the nervous system. It is not found in the basal metazoans such as sponge and placozoan and thought to arise early in the Eumetazoa (a clade comprising all major animal groups except sponges and placozoans) (Fig. 3). The sponge has well-defined photosensory cells. At the posterior pole of demosponge larva, there appears a ring of monociliated, pigment-containing cells and these cells function as a photoreceptor and control the directional swimming of the larva (Leys and Degnan, 2001; Leys *et al.*, 2002). Unlike the sponge, placozoan contains four basic cell types and none of them is morphologically similar to the neuron or sensory cell. However, neural genes involved in neurosynaptic activity and biosynthesis are identified in the

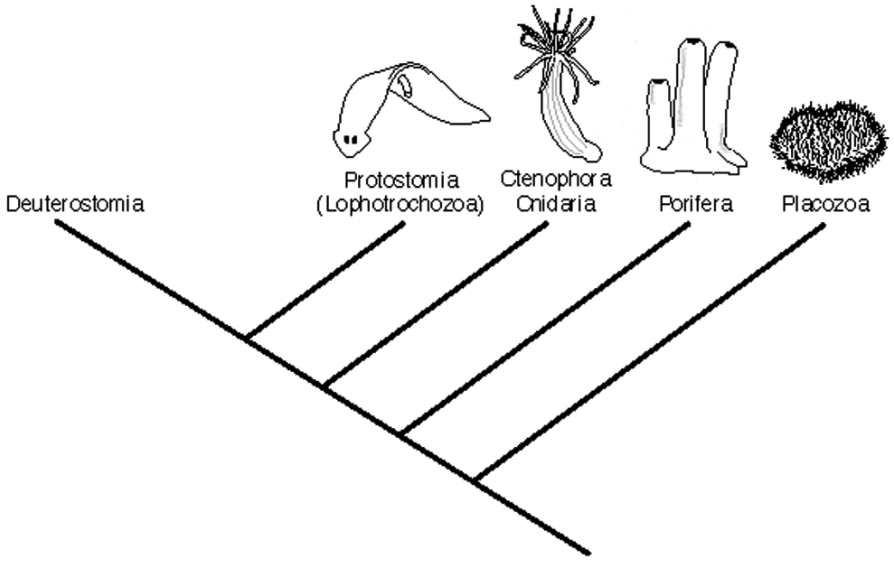

Figure 3. A classical taxonomy of basal metazoans.

placozoan genome (Srivastava *et al.*, 2008), and a *Proto-Pax* gene is expressed in a proliferating/differentiating region near the outer edge of placozoan cell body (Hadrys *et al.*, 2005). This molecular evidence suggests that neural genes predate the ancestor of nervous system. In fact, sponge and placozoan genomes encode a great deal of transcription factor genes that play a critical role in signaling pathway, embryogenesis and tissue specification of eumetazoan (Degnan *et al.*, 2005; Strivastava *et al.*, 2008). Therefore it is rather unlikely that the gene repertoire of the nervous system arises after the divergence of cnidarians but instead it emerges in the last common ancestor of Metazoa or even earlier. In other words, the repertoire of molecular factors that are essential for neuronal development and functions has already had a role in neuronal activities in the 'primitive cell' far before the emergence of the nerve cell in animals. The 'primitive cell' is referred to those having the potency but yet to develop into the neuron stem cell (Fig. 4). Later it evolves into two sister cells and one of them functionally diversifies into neuron stem cells. A similar view is found in a recent review, Arendt D. (2008) has proposed a scenario

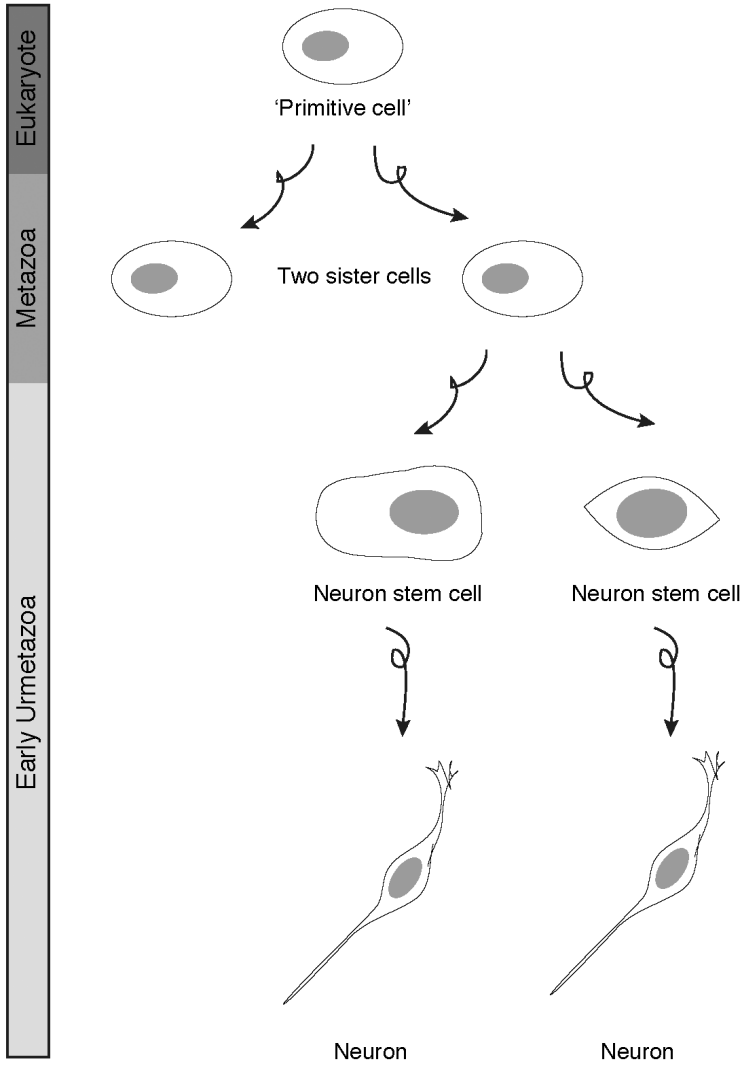

Figure 4. A scenario is proposed for the origin of the nervous system or neuron. A partial neural gene repertoire exists in the 'primitive cell' before the emergence of Metazoa. This 'primitive cell' evolves and diverges into two sister cells. One of the sister cells specializes its function via genetic modifications (eg. gene duplication, gene gain and loss, protein domain shuffling, horizontal gene transfer, etc) and further divides into two sister cells that contain similar potencies as neuron stem cells. Later in the evolution, one of two neuron stem cells might lose its ability to differentiate into a neuron, but the other one remains as a neuron stem cell (not illustrated in the figure).

of cell type evolution in which ancient metazoan cell types have multiple functions and later the cell type diversification within a species increases the number of functionally specialized cells. Arendt has further stated the importance of molecular signature (which refers to a set of differentiation and regulatory genes in sister cell types) in the evolutionary diversification of cell types. In fact, in our opinion, this molecular signature has to be imprinted in the ancient cell type before the emergence of sister cell types. Early sister cell types might retain the plasticity at the gene regulatory level. Not all cnidarians have neurons derived from one kind of cell types. The neurons of *Hydra* are differentiated from the interstitial cell (I-cell), a multipotent stem cell that lies between the ectodermal epithelial cells (David and Gierer, 1974). Nematostella, a marine cnidarian, lacks the I-cell and neurons are originated from the ectodermal and endodermal epithelial cells (stem cell lineages of cnidaria) (Extavour *et al.*, 2005; Marlow *et al.*, 2009). It seems plausible that the I-cell and epithelial cell are both sister cells and contain the same regulatory gene network that is capable of determining their cell fates and differentiating into neurons. Interestingly, this conclusion is supported by the observation that *Hydra* I-cell arose in the endoderm at the early embryonic stage, suggesting the endodermal origin of *Hydra* I-cell (Genikhovich, 2006).

In *Hydra*, I-cell and epithelial cell are having a relationship of sister cells as discussed above, while the differentiated products of I-cell, neuron and nematocyte (contain a stinging organelle that functions for prey capture and defense) can also be considered as secondary sister cells (Hwang *et al.*, 2007). This is not only because both are originated from the I-cell, but because they also both have specific expressions of *achaete–scute, prdl-b and COUP-TF* (Hayakawa *et al.*, 2004; Miljkovic-Licina *et al.*, 2004). The nematocyte has long been regarded as a sensory cell as it bears a mechano- and chemo-sensory receptor called cnidocil apparatus at the apical surface of cell.

In fact, the evolution of neural genes occurred far before the emergence of multicellular organisms, approximately 1,400 million years ago (Nei *et al.*, 2001). A genomic-scale analysis of nervous system (NS) specific genes shows that 35 out of 255 human NS specific genes (14%) appear prior to the split between metazoans and yeast (Fig. 5, see p. 610) (Noda *et al.*, 2006). Moreover in the same analysis, a sudden increase in the number of NS genes occurs before the emergence of vertebrates, and the majority of these NS genes are critical for protein binding or protein-protein interaction. Although the analysis is based on a small data set (255 human NS spe-

cific genes), the results support two conclusions: (1) A significant number of NS specific genes in yeast marks the ancestral complexity of the neural gene repertoire before the emergence of the nervous system, and (2) the evolution of the nervous system is mainly driven by the extensive gene gain.

The primitive brain of planaria

Although recent phylogenetic analyses have placed platyhelminth flatworms in the clade of Lophotrochozoa and not as basal to Bilateria (Fig. 3) (Baguñà and Riutort, 2004; Ruiz-Trillo *et al.*, 2004), it has a centralized nervous system that can be described as a 'primitive brain'. Planarian, a freshwater flatworm, contains a mass of cephalic ganglions in the head and a pair of ventral nerve cords (VNC) running parallel to the body axis (Fig. 1, p. 608). The cephalic ganglion has a bilobed structure with neuronal cell bodies that form the outer later (cortex) and nerve fibers that concentrate collectively in the inner core of the ganglion. Compared to the nervous system of cnidarians, the planarian central nervous system has evolved with several 'brain' features: (1) centralized neurons at cephalic region, (2) nerve cord, (3) neuron fibers surrounded by a layer of cortex, (4) lobes with commissural fibers, (5) glial cells, (6) motor neurons and (7) neurons with elaborated dendrites. Thus, the structure of the planarian nervous system has been well studied and characterized (Flexner, 1898; Oosaki and Ishii, 1965; Baguñà and Ballester, 1978). Not until recently, molecular tools including whole mount *in situ* hybridization, whole mount immunostaining, expression sequence tag (EST), microarray, and RNA interference are applied to study the detailed morphology and function of the planarian nervous system (Cebrià *et al.*, 2002a; Cebrià *et al.*, 2002b; Agata *et al.*, 1998; Mineta *et al.*, 2003; Nakazawa *et al.*, 2003).

What makes the nervous system (NS) complex?

In order to study the genes expressing in the planarian brain, we collect anterior tissue including the cephalic ganglia (above the neck) of planarians and conduct EST sequencing. Based on known NS genes, we have identified 116 genes out of 3101 that share significant homology to NS genes of other organisms (Mineta *et al.*, 2003). A further analysis of 116 NS-related genes has shown that more than 95% have their homologs

in humans, *Drosophila melanogaster* and *Caenorhabditis elegans*. These NS-related genes include those involved specifically in the brain morphogenesis and neural network formation, suggesting the possibility that the bilaterian central nervous systems are derived from a common origin.

Moreover, we also examined the gene expression in the anterior part (i.e. cephalic ganglia) of the planarian by using cDNA microarray containing 1,640 nonredundant genes (Nakazawa *et al.*, 2003). The use of planarian cDNA microarray has an advantage over the ESTs collection. Planarian cDNA microarray can be used to examine the novel genes expressed in the central nervous system. A total of 205 genes are differentially expressed in the anterior part and by using whole mount *in situ* hybridization, the top 30 genes show various regional expressions in the cephalic ganglia and the ventral nerve cords (Fig. 2B, **see p. xxx**). Many of the top 30 genes have an unknown function. The variety of expression patterns of the top 30 genes in the planarian brain demonstrates the highly organized nature and the complex neural activities of the planarian central nervous system.

In summary, the above data indicate that the planarian brain expresses genes related to those in the human central nervous system and it is also highly divided into distinct compartmentalized regions (i.e. functional domains) according to the gene expression patterns. One of the important features for a diffused nerve net (in *Hydra*) that evolved into a centralized nervous system (in planarian) requires the mechanism of axon guidance. Axon guidance allows the proper growth of the axon cone and the precise target reached by the axon. The axon guidance molecules in planarians such as *NCAM*, *slit, netrin and robo* play conserved and important roles in the maintenance of the nervous system architecture (Cebrià, 2007). RNAi interferences of these three genes in the planarian result in the failure to regenerate a normal brain. For example, no proper commissural connection is seen between cephalic ganglia and nerve cords in regenerating planarians after *Smed-roboA* RNAi (Cebrià and Newmark, 2007). Interestingly, sequences homologous to axon guidance genes such as *NCAM, robo, slit, netrin, Eph* receptor and NCAM are also identified in the *Hydra* genome (personal data). Thus, the emergence of axon guidance genes did not happen in the early ancestors of Bilateria but rather dates back to the Eumetazoa. It would be of great interest to know whether the axon guidance homologs of *Hydra* had conserved functions like those of planarians and other bilaterians. Perhaps the complexity of the central nervous system as compared to the nerve net is not due to the number of NS genes but to the dynamics regulation of the gene network. It should

be noted that the planarian also has lineage-specific NS genes. One of the examples is '*nou-darake*', a gene that belongs to the FGF (fibroblast growth factor) receptor family and is expressed in the cephalic ganglia and its surrounding tissues. The existence of '*nou-darake*' is important to restrict the brain tissue in the cephalic region (Cebrià, *et al.*, 2002a).

Concluding Remarks

For years, researchers have been struggling to isolate genes from an organism, to gain the genomic information of a gene and to compare the genes among many different species. Now with high-throughput sequencing, powerful analysis tools and large-scale data storage, we are able to collect a large amount of EST and genome data in a short time. These advanced approaches have found a solution to the current research on the evolution of the central nervous system. To understand the human brain, we believe that the study of lower organisms such as *Hydra* and planarian is essential and would provide useful knowledge of how the human brain evolves. In our studies, we predict that *Hydra* and planarian share at least half of their nervous system genes with humans. Clearly, many nervous system genes predate the emergence of Metazoa and the nervous system evolves as the gene network increases its complexity. For a future perspective, we believe that it is essential to construct a virtual 3D human brain. This 3D immersive environment would provide the gene expression map of each central nervous system gene against the anatomical, tissue- and single-cell levels of the human brain structure.

REFERENCES

Arendt, D. (2008) The evolution of cell types in animals: emerging principles from molecular studies. *Nat. Rev. Genet*. 9, 868-882.

Agata, K., Soejima, Y., Kato, K., Kobayashi, C., Umesono, Y. and Watanabe, K. (1998) Structure of the planarian central nervous system (CNS) revealed by neuronal cell markers. *Zool. Sci*. 15, 433-440.

Baguñà, J. and Ballester, R. (1978) The nervous system in planarians: Peripheral and gastrodermal plexuses, pharynx innervation, and the relationship between central nervous system structure and the acoelomate organization. *J. Morph.* 155, 237-252.

Baguñà, J. and Riutort, M. (2004) The dawn of bilaterian animals: the case of acoelomorph flatworms. *Bioassays* 26, 1046-1057.

Bode, H.R. (2001) The role of Hox genes in axial patterning in Hydra. *Am. Zool*. 41, 621-628.

Cebrià, F., Kobayashi C., Umesono, Y., Nakazawa, M., Mineta, K., Ikeo, K., Gojobori, T., Itoh, M., Taira, M., Sanchez Alvarado, A. and Agata, K. (2002a) FGFR-related gene *nou-darake* restricts brain tissues to the head region of planarians. *Nature* 419, 620-624.

Cebrià, F., Kudome, T., Nakazawa, M., Mineta, K., Ikeo, K., Gojobori, T. and Agata, K. (2002b) The expression of neural-specific genes reveals the structural and molecular complexity of the planarian central nervous system. *Mech. Dev*. 116, 199-204.

Cebrià, F. (2007) Regenerating the central nervous system: how easy for planarians! *Dev. Genes. Evol.* 217, 733-748.

Cebrià, F. and Newmark, P.A. (2007) Morphogenesis defects are associated with abnormal nervous system regeneration following roboA RNAi in planarians. *Development* 134, 833-837.

David, C.N. and Gierer, A. (1974) Cell cycle kinetics and development of *Hydra attenuata* III. Nerve and nematocyte differentiation. *J. Cell. Sci*. 16, 359-375.

Degnan, B.M., Leys, S.P. and Larroux, C. (2005) Sponge development and antiquity of animal pattern formation. *Integr. Comp. Biol.* 45, 335-341.

Extavour, C.G., Pang, K., Matus, D.Q. and Martindale, M.Q. (2005) Vasa and nanos expression patterns in a sea anemone and the evolution of bilaterian germ cell specification mechanisms. *Evol. Dev.* 7, 201-215.

Finnerty, J.R., Pang, K., Burton, P., Paulson, D. and Martindale, M.Q. (2004) Origins of bilateral symmetry: *Hox* and *Dpp* expression in a sea anemone. *Science* 304, 1335-1337.

Flexner, S. (1898) The regeneration of the nervous system of planaria torva and the anatomy of the nervous system of double-headed forms. *J. Morph*. 14, 337-346.

Genikhovich, G., Kürn, U., Hemmrich, G. and Bosch, T.C.G. (2006) Discovery of genes expressed in *Hydra* embryogenesis. *Dev. Biol*. 289, 466-481.

Hadrys, T., DeSalle, R., Sagasser, S., Fischer, N. and Schierwater, B. (2005) The trichoplax PaxB gene: A putative proto-PaxA/B/C gene predating the origin of nerve and sensory cells. *Mol. Biol. Evol*. 22, 1569-1578.

Hayakawa, E., Fujisawa, C. and Fujisawa, T. (2004) Involvement of *Hydra* achaete–scute gene CnASH in the differentiation pathway of sensory neurons in the tentacles. *Dev. Genes. Evol.* 214, 486-492.

Hejnol, A. and Martindale, M.Q. (2008) Acoel development indicates the independent evolution of the bilaterian mouth and anus. Nature 456, 382-386.

Holland, L.Z. (2000) Body-plan evolution in the Bilateria: early antero-posterior patterning and the deuterostome-protostome dichotomy. *Curr. Opin. Genet. Dev*. 10, 434-442.

Holland, N.D. (2003) Early central nervous system evolution: An era of skin brains? *Nat. Rev. Neurosci*. 4, 617-627.

Koonin, E.V., Aravind, L. and Kondrashov, A.S. (2000) The impact of comparative genomics on our understanding of evolution. *Cell* 101, 573-576.

Kusserow, A., Pang, K., Sturm, C., Hrouda, M., Lentfer, J., Schmidt, H.A., Technau, U., von Haeseler, A., Hobmayer, B., Martindale, M.Q. and Holsten, T.W. (2005) Unexpected complexity of the *Wnt* gene family in a sea anemone. *Nature* 433, 156-160.

Lacalli, T. (1995) Dorsoventral axis inversion: a phylogenetic perspective. *BioEssays* 18, 251-254.

Lein *et al.* (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445, 168-176.

Lengfeld, T., Watanabe, H., Simakov, O., Lindgens, D., Gee, L., Law, L., Schmidt, H.A., Ozbek, S., Bode, H. and Holstein, T.W. (2009) Multiple Wnts are involved in *Hydra* organizer formation and regeneration. *Dev. Biol*. *In press*.

Leys, S.P. and Degnan, B.M. (2001) Cytological basis of photoresponsive behavior in a sponge larva. *Biol. Bull*. 201, 323-338.

Leys, S.P., Cronin, T.W., Degnan, B.M. and Marshall, J.N. (2002) Spectral sensitivity in a sponge larva. *J. Comp. Physiol*. A 188, 199-202.

Marlow, H.Q., Strivastava, M., Matus, D.Q., Rokhsar D. and Martindale, M.Q. (2009) Anatomy and development of the nervous system of Nematostella vectensis, an anthozoan cnidarian. *Dev. Neurobiol.* 69, 235-254.

Martindale, M.Q. (2005) The evolution of metazoan axial properties. *Nat. Rev. Genet*. 6, 917-927.

Meinhardt, H. (2002) The radial-symmetric hydra and the evolution of the bilateral body plan: an old body became a young brain. *BioEssays* 24:185-191.

Miljkovic-Licina, M., Gauchat, D. and Galliot, B. (2004) Neuronal evolution: analysis of regulatory genes in a first-evolved nervous system. *Biosystems* 76, 75-87.

Mineta, K., Nakazawa, M., Cebrià, F., Ikeo, K., Agata, K. and Gojobori, T. (2003) Origin and evolutionary process of the CNS elucidated by comparative genomics analysis of planarian ESTs. *Proc. Natl. Acad. Sci. USA* 100, 7666-7671.

Nakazawa, M., Cebrià, F., Mineta, K., Ikeo, K., Agata, K. and Gojobori, T. (2003) Search for the evolutionary origin of a brain: planarian brain characterized by microarray. *Mol. Biol. Evol*. 20, 784-791.

Nei, M., Xu, P. and Glazko, G. (2001) Estimation of divergence times from multiprotein sequences for a few mammalian species and several distantly related organisms, *Proc. Natl. Acad. Sci. USA* 98, 2497-2502.

Noda, O., Ikeo, K. and Gojobori, T. (2006) Comparative genome analyses of nervous system-specific genes. *Gene* 365, 130-136.

Oosaki, T. and Ishii, S. (1965) Observations of the ultrastructure of nerve cells in the brain of the planarian, *Dugesia gonocephala*. *Z. Zellforsch*. 66, 782-793.

Ruiz-Trillo, I., Riutort, M., Fourcade, H.M., Baguñà, J. and Boore, J.L. (2004) Mitochondrial genome data support the basal position of Acoelomorpha and the polyphyly of the Platyhelminthes. *Mol. Phylogen. Evol*. 33, 321-332.

Srivastava, M. *et al.* (2008) The tricholax genome and the nature of placozoans. *Nature* 454, 955-960.

Telford, M.J. (2007) A single origin of the central nervous system? *Cell* 129, 237-239.

DISCUSSION ON PROF. GOJOBORI'S PAPER

PROF. ARBER: Thank you very much for this very interesting and prospective presentation. I think you provide here an important answer to a question which evolutionary biologists have raised since a long time: how do novel properties emerge. With that regard, thinking of your statement that a knock out mutant of planaria gave an increase in brain tissue in that animal, I just wonder, could one speculate that the relatively rapid increase in the size of the human brain is due to a particular deteriorating mutation of something which was inhibiting the proliferation of the brain before?

PROF. GOJOBORI: If you kindly allow me to make a wild speculation, I think our results show that all of the gene set existed. However, still, the connection between genes was not much. Therefore, I think there was a certain important gene which made the network easier. I would call it an epoch-making gene, which was able to form a gene network. We can now observe genes which are responsible for formation of the human brain. In particular, when the gene or genome duplication takes place, the number of genes increases. Therefore, the number of networks would instantly increase accordingly. This is speculation.

PROF. LE DOUARIN: It was very interesting. I would like to know how you obtained these mutant epithelial hydras and the second question is, are there genes which are expressed specifically in the nematocytes and not in nerve cells which are over the body of the hydras?

PROF. GOJOBORI: First, the epithelial hydra was obtained by chemical mutagenesis and, among many variants, successfully this particular mutant was obtained. Secondly, like the genes which are expressed in the neurocells and the nematocyte cells, certainly the nematocyte cells have a few sets of genes very much specific to hydra, so we speculate that those genes may be horizontally transplanted from other organisms. Although we do not know yet, according to what our phylogenetic tree shows, this is quite possible.

PROF. W. SINGER: I would like to make a comment. This was fascinating because it seems to show a parallel with brain development on the other end. You seem to indicate that it is not the invention of new building blocks but the way in which they interact in more and more complex networks that makes all the difference. We see exactly the same in the brain, there are no new building blocks but the complexity of interactions is increased and that makes all the difference. It is a beautiful example of emergence.

PROF. COLLINS: A very fascinating presentation in terms of the evolution of brain-specific genes. It is interesting to note this emergence seems to be particularly strong before the appearance of bony fish, but one wonders whether that is, in part, an artefact of the difficulty in identifying evolutionary matches as you get further and further apart and is it possible that, in fact, some of those genes that are only identifiable by the time you get to bony fish actually have homologues in similar species but they are too diverged for your computational methods to appreciate that match?

PROF. GOJOBORI: Thanks so much for this important question. We worried a lot about this possibility, therefore first we changed our homologic criteria, but still the same conclusion was obtained. The other evidence is as follows. It is possible that before primate emergence a huge number of genes might have emerged, because sequence homology is much easier to detect. But instead, somehow, before bony fish emergence it appeared. Therefore, of course, I think there are some biases, but still general features can be unchanged. The most important issue would be the number of genes, because we dealt with only four hundred genes. Therefore if we can successfully obtain more genes, it is a question whether the picture does not change much.

PROF. NIRENBERG: What kind of cell-membrane genes have evolved latest? What functions?

PROF. GOJOBORI: In the human brain?

PROF. NIRENBERG: Yes, in the nervous system.

PROF. GOJOBORI: I think it looks to me that ligand-receptor type genes have evolved latest. In particular, the genes that are related to the receptor system seem to have played an important function among membrane associate proteins.

PROF. VICUÑA: I have two questions. The first one is the following: when you study gene expression in human brains, how do you obtain the samples? The other question is whether gene expression in the brain changes according to the physiological conditions of the brain. For example, if a person has a mental disease, or is in a coma, or sleeping, etc. Have you conducted those studies?

PROF. GOJOBORI: I have to confess I did not conduct isolation of a brain from a human body by myself, because we used only the data. However, according to my understanding, mRNAs were obtained from the dead body of a healthy normal person with informed consent of closest relatives. Therefore, certainly, if we have a certain diseased brain it would be interesting to see. Of course, ethical problems should be addressed carefully. I think specifically expressed mechanisms should exist in a transcription regulation system. If we get to understand the regulation system then I think we might have some answers. But we do not know yet.

PROF. DEHAENE: It was a beautiful story of evolution that you told us. Surely along these steps there are some important inventions, and one of them, in the nervous system, is the invention of spiking or excitable cells, so that the cells can fire and send a very specific message to long distances, to specific targets. I was wondering whether your research was showing when the spiking of cells first appeared.

PROF. GOJOBORI: That is also an important question, thanks so much. We are now examining cellular structures by EM, particularly gap junction between cells. As you know, electric synapses and chemical synapses can be examined from an evolutionary point of view. We have some answers. Certainly we are now addressing this question.

PROF. M. SINGER: Thank you. I too enjoyed your talk very much. You touched on a subject which, I think, is probably fascinating, namely, why it is that brains wound up in the heads of animals. You touched on that when you talked about the mutation in the hydra that resulted in multiple eyes and also about the cells that, instead of being concentrated, are dispersed. I wonder, do you know anything about the nature of that gene?

PROF. GOJOBORI: That is a very interesting question. Again, we do not know the answer. I think certain gene sets seem to have been lost, in the

case of *C. elegans*. If a phylogenetic tree is correct and the planarian brain existed, then it must have had the formation of a brain, even before *C. elegans*. If that is the case, then the brain must have disappeared from *C. elegans*. Of course, researchers often say *C. elegans* has a brain, but in this case it means only a tiny number of cells. Therefore, if we can conduct transformation experiments in those organisms, then we may be able to answer the question of why the gene is located in the head. But we may be able to have the epitopic brain which is located in the tail, for example, by conducting transformation.

GENETIC AND EPIGENETIC SHAPING OF COGNITION – PREREQUISITES OF CULTURAL EVOLUTION

WOLF SINGER

Before entering the discussion of the evolution of our brains and the options for their epigenetic shaping I consider it appropriate to begin with an epistemic caveat. To the best of my knowledge there is consensus among neurobiologists that all mental phenomena including the highest cognitive functions are the product of neuronal processes. Likewise, social realities such as value systems and moral judgments are considered to be the products of interactions among human beings endowed with brains, the cognitive abilities of which allowed for the initiation of cultural evolution. If one accepts this position it follows that we can only perceive, imagine and comprehend what the cognitive abilities of our brains allow us to seize. Because brains – just as other organs – are the product of evolutionary adaptation, this implies that our cognitive abilities are with all likelihood constrained. Our brains are optimized to secure survival and reproduction in the narrow segment of the world in which life evolved. Coping with the highly specific challenges of an insecure and purely predictable world requires adoption of pragmatic heuristics that differ most likely from the cognitive strategies needed to assess a hypothetical 'objective' truth. Numerous experiments on perceptual illusions illustrate that such is indeed the case. Thus, the sobering conclusion seems to be inevitable that our cognitive abilities are likely to be highly constrained and idiosyncratically adapted to only a very small sector of the world. The world, as unravelled by scientific investigation, extends from infinitely small to infinitely large dimensions. Life, however, has evolved only within a narrow range that extends from micrometers to a few metres. Processes at this mesoscopic scale are dominated by the laws of classical physics and most of the dynamics that life has to cope with are linear. At this scale it makes sense to define states of matter as liquid, solid or gasous, to define space and time as separate categories, and to distinguish

between cause and effect. Our sensory systems extract in a highly selective way a few signals from our environment that we then experience as light, temperature, vibrations, sound, smells and tastes. Some of these sensory categories reflect an arbitrary subdivision of physical continua. Thus, we classify electromagnetic radiations with wave lengths between 400 and 700 nm as light and those with longer wave length as heat. To us these categories appear as natural properties of the world and even though we perceive only very narrow spectra of the available physical and chemical signals in our environment, we experience the world as coherent and continuous – a convincing example of the constructive nature of our perception.

Because our cognition has adapted to a narrow range of the mesoscopic world, it is difficult for us to develop intuitions for phenomena at other scales. Our intuition of objects is meaningless in the world of quantum physics just as our concept of causality and our intuition of space and time does not hold for the putative structure of the universe. We tend to believe that the rules and concepts that we infer from the mesoscopic world can be extrapolated to all the other dimensions but there is no guarantee that this is actually the case. It must even be considered that the way in which we reason and draw conclusions is a specific adaptation to the processes at the mesoscopic scale and perhaps not generalizable. Thus, it is very likely that there are natural boundaries to what we can perceive, imagine and understand. Where these limits are and what is concealed behind, will in principle remain unknown. There is, thus, ample space for metaphysics and belief, constrained only be what is actually known.

THE CONSTRUCTIVISTIC NATURE OF PERCEPTION AND THE SOURCES OF KNOWLEDGE

A large body of psychophysical evidence and neurobiological data indicate that perceiving is essentially a constructive process by which the brain attempts to interpret the sparse sensory signals conveyed by the various sensory organs on the basis of a huge amount of *a priori* knowledge (priors) that is stored in the functional architecture of the brain. What we perceive and how we perceive is by and large determined by context dependent expectancies and stored knowledge about the world. This raises the question of where the knowledge required for the construction of our percepts is derived from. Neurophysiological evidence indicates that knowledge and the rules for its application reside in the functional architecture of the brain. The term 'functional architecture' stands for the way in which nerve

cells in the brain are interconnected with each other. Unlike in computers that are often erroneously cited to explain the functioning of nervous systems, there are no structurally and functionally different subsystems in the brain that could be considered as central processors and the various storage devices such as memories for data and programs. In the brain there are only neurons, and connections and processing as well as storage functions are accomplished within the same networks. All computations are determined by the functional architecture of these networks. What matters is which neurons are interconnected, whether these connections are excitatory or inhibitory and whether they are strong or weak. The setting of these variables is also the basis of all the knowledge that is stored in the brain. Thus, the search for the sources of knowledge is reduced to the question of which factors specify the functional architectures of brains.

The most important of these factors is beyond any doubt evolution. Through evolutionary selection brain architectures have evolved which contain the knowledge and the application programs that the organism needs in order to cope effectively with the challenges of its environment. In this sense evolution can be considered as a cognitive process. Through adaptation of brain architectures to the requirements of survival in specific biotops knowledge about the world is acquired, stored in the genes and made available for the control of adapted behavior every time a new brain develops. The knowledge acquired through this process is of course implicit. We do not know that we have it because we were not around when it was acquired. Therefore, this knowledge serves as unconscious priors that determines all subsequent cognitive processes. An important consequence is that perceptions based on these implicit priors have the quality of being objective, unreducible and not relativatable. They are taken as representing undisputable truth.

Another important source of knowledge is developmental shaping of brain architectures addressed also as developmental imprinting. The human brain develops structurally until around age 20. This developmental process is characterized by a continuous making and breaking of connections whereby the selection of connections that are to be consolidated is guided by neuronal activity and hence by experience and interaction with the environment. This developmental process leads to a substantial modification and refinement of the genetically specified architecture of brains and thereby installs further knowledge in the brain – this time knowledge derived from interaction with the actual environment in which the organism evolves. Much of this knowledge is also implicit. Brain structures that

support episodic memories develop only years after birth which leads to the phenomenon of childhood amnesia. Children up to the age of about 4 years learn about the world but they keep no trace of the context in which they have learnt. They know but they do not know where their knowledge comes from. This is why early acquired knowledge – just as evolutionary knowledge – is implicit, serves as a source of unconscious priors for perception and thereby nourishes convictions that cannot be put to question.

This is not so for knowledge acquired through normal learning processes that begin once episodic memory functions become available and that persist throughout the entire life span. Knowledge acquired through this mechanism is explicit. Subjects are usually aware of having acquired the respective contents by experience and remember the context in which acquisition has taken place. Once brain development has come to an end, further learning is based on activity dependent modifications of the efficiency of existing connections and these changes are brought about by lasting modifications of the molecular machinery that mediates communication among nerve cells, i.e. synaptic transmission. These changes also go along with structural alterations but these are resolvable only at the ultra-structural level.

The layout of the functional architecture of brains is thus determined essentially by three factors, evolutionary adaptation, epigenetic shaping during postnatal brain development and normal learning processes. The resulting architecture in turn determines the various sensory categories according to which we classify sensory signals, the criteria for the definition of objects, the rules according to which brains detect contingencies in the outer world and form associations and finally, the way in which we reason, make inferences and assign values.

The following two figures illustrate the extent to which the *a priori* knowledge stored in the architecture of our brains determines the way in which we perceive.

The object in Figure 1 (see p. 611) is a mold used to produce candies. On the left side one sees the front aspect of the mold with the concavities and on the right the rear side with the corresponding convex protrusions. In reality, both pictures show the front aspect, but one picture is rotated by 180°. The reason for these very different perceptions is that the brain makes the *a priori* assumption that light comes from above. In this case contours that have the shadow above need to be interpreted as concave and those with the shadow below as convex. Thus, an implicit assumption determines what we perceive. Somehow this assumption is implemented in the processing architecture of the visual cortex but we are not aware of it.

Another even more striking example is shown in Figure 2 (see p. 611). It is hard to believe, but surfaces A and B have exactly the same luminance and this can be verified by covering all squares except A and B with white paper. The squares A and B appear as different because the brain sees the shadow that is caused by the cylinder on the right. Even though the amount of light reflected from surfaces A and B and impinging on the retina is exactly the same, the brain interprets the brightness of the two surfaces as different because it infers the following: given that there is a shadow, surface B must be brighter than surface A which has no shadow on it, in order to reflect the same amount of light. Thus, the brain 'computes' the inferred brightness of the surfaces but we are not aware of these computations. We just perceive the result and take it as real, i.e. we see B much brighter than A. One could spend hours with the demonstration of examples which indicate that the brain is generating inferences that we are not aware of, that it is permanently reconstructing the world according to *a priori* knowledge and that we, as perceiving subjects, have to take for granted what the system finally offers us as conscious experience. It is important to emphasize that this is not only the case with specially designed psycho-physical experiments but it is an essential feature of all our perceptual processes. We perceive the result of complex computational operations, and because we are unaware of both the priors and the rationale on which these interpretations are based, we tend to take for granted what we see. We do not realize that our percept is the result of complex computations that are based on assumptions and have difficulties to accept that what appears to be so evident and an invariable property of the perceived object is actually the result of a highly inferential and constructive operation.

The following Gedankenexperiment is meant to illustrate the adaptive value of such perceptual inferences. Imagine that red berries with a specific hue constitute a major food source and that red berries with a slightly different colour are poisonous. It is thus imperative to distinguish between the two sorts of berries and to be able to do this irrespective of daytime. The problem is that the spectral composition of sunlight is radically different in the morning, at noon and in the evening. Accordingly, the spectra reflected by the two kinds of berries differ at different times of the day and it may well be that the spectra of the poisonous berries produced by the morning light resemble the spectra produced by the good berries at noon. Thus, the only way to assure the distinction between the two at any time of the day is to interpret the reflected spectra as a function of the actual spectrum of the sunlight. The latter cannot be measured directly but it can be inferred from the comparison

of spectra reflected from familiar objects and *a priori* knowledge of their likely colour. Thus, by comparing the color of leaves, barks, rocks, the clouds etc. the system can estimate the spectral composition of the illuminating light source (the sun), take this into account when interpreting the spectra reflected from the two types of berries and only then compute the hue of the colour that is actually perceived. Through this complicated operation it can be assured that the good berries are perceived as having the same colour irrespective of illumination conditions. This is but one of a large number of examples which illustrate that what we perceive and interpret as invariant properties of objects is actually the result of a highly inferential and constructive process. Furthermore, these examples explain why it is advantageous for organisms to base their perception on *a priori* knowledge and pragmatic heuristics rather than perceiving the absolute, unprocessed values of the signals provided by our sensors that transform physical or chemical stimuli into amplitude modulated neuronal activity.

Causes and effects of cultural evolution

Because most of the priors that determine our perception of the world around us have been acquired during evolution we share them with the animal kingdom. Non-human primates for example but also members of other species such as cats, dogs and even insects make the same inferences and thus perceive the world in similar ways. There are, however, also important differences and these result from the fact that only human brains are exposed during their development to realities that were absent during biological evolution that has shaped our brains as well as those of animals – realities that are the product of cultural evolution. This raises two related questions: what are the cognitive abilities that allowed homo sapiens to initiate the process of cultural evolution and what are the consequences of the epigenetic shaping of human brains by their exposure to socio-cultural realities?

Over the last decades, a number of cognitive functions have been identified that are apparently not found in our nearest neighbors, the great apes, and thus with all likelihood are responsible for the initiation of cultural evolution. One of these functions is the ability to generate a theory of mind, to imagine what goes on in the mind of the respective other when she/he is exposed to a particular situation but does not signal through any perceivable signs what her/his thoughts, intentions or feelings are. Another important function is shared attention. If a human being directs his/her

gaze to a particular target or points towards it, a human observer is able to direct attention to the same target, understanding that both subjects are now sharing their attention. Dogs, probably because of domestication, are able to accomplish this very specific function but the great apes are not. Furthermore, human beings possess an unprecedented ability to generalize, to identify the common in the seemingly different and, therefore, are capable of forming abstract, symbolic representations. When monkeys learn to associate particular attributes with signals provided through one sensory modality, they usually have great difficulty recognizing the presence of the same attributes when signals are provided by a different modality. Humans accomplish such inter-modal transfer with great ease, probably because of the specific features of their cortical architecture that allow for easy exchange of information across the processing streams of the various sensory systems or because of the addition of association areas that allow for convergence of information from different modalities. The resulting ability of abstraction and symbolic coding is with all likelihood one of the prerequisites for the development of language. Other prerequisites seem to be the ability to represent complex sequences of nested relations which are at the origin of the comprehension and production of syntactic structures. Finally, human beings are capable of transmitting knowledge acquired during their lifetime through intentional instruction and education. Even the great apes learn essentially through imitation. Infant chimpanzees imitate nut cracking and even if they perform poorly, their mothers do not instruct their offspring but just continue to crack their own nuts.

This then raises the question of which changes in brain architecture might be responsible for the emergence of these novel cognitive abilities. When comparing the brains of the great apes with those of human beings, the only remarkable difference is the addition of new areas of the neocortex. Apart from that, there are no major structural changes and even the new cortical areas closely resemble with respect to their intrinsic organization those which exist both in humans and non-human primates. As outlined previously, the computational operations performed by a neuronal network are fully determined by its functional architecture and, therefore, it can be inferred that the new cortical regions operate according to the same principles as those that had already existed. Thus, the only options that these new areas offer are those that can be realized by implementing further nodes in the network. This could permit the generation of platforms for novel and more complex associations among the results obtained in parallel and previously unconnected processing streams or – if added on

top of processing hierarchies – the generation of meta-representations. There is evidence for both strategies and both are likely foundations for the enhanced sophistication of human cognition. This interpretation agrees with the evidence that the molecular composition of nerve cells, the mechanisms mediating signal transduction and the molecular machinery supporting modification of synapses by learning closely resemble those found not only in all vertebrates but also in molluscs and insects. With the realization of the canonical circuits that characterize cortical modules, evolution has apparently discovered a computational algorithm that is universally applicable both to the evaluation of sensory signals of different modalities and to the design and organization of executive acts. Moreover, and this seems to be particularly advantageous, this canonical circuit can support iterative, reentrant processing of the results generated by these very circuits and thereby allow for the virtually unlimited recombination of signals.

Epigenetic shaping, cultural diversity and tolerance

Together with anatomical modifications allowing bipedal gait that freed the front legs for duties other than locomotion, the development of the cognitive abilities listed above allowed *Homo sapiens* to initiate cultural evolution. Although at dramatically different time scales, the dynamics of biological and cultural evolution share certain similarities. In both cases, complexification and diversification of evolving structures were initially very slow but then experienced a dramatic acceleration. Once *Homo sapiens* appeared on stage, it took apparently tens of thousands of years to develop communication skills resembling syntactically based languages, social structures that allowed for labour sharing, tool making, sedentary lifestyles and the development of concepts that added a spiritual or metaphysical dimension to the material world. However, this period of slow differentiation underwent a phase transition about 30,000 years ago that led to an exponential acceleration of socio-cultural evolution with its countless ground breaking inventions. This acceleration suggests that evolutionary mechanisms that support autocatalytic processes became effective. One of them might have been the increase in population density. Increasing population density permitted the establishment of denser communication networks, the sharing of inventions, the development of cooperative strategies for a less time- and energy-consuming exploitation of resources and the reinvestment of the spared time and energy into exploratory activities that

rendered these early societies more and more independent of the hazards of nature. However, the most effective factors that catalyzed this unprecedented acceleration are with all likelihood the extremely protracted postnatal development of the human brain and the ability of human subjects to intentionally educate their offspring. In conjunction, these two mechanisms make it possible to translate knowledge acquired during lifetime into the functional architecture of the brains of the respective offspring. As outlined above, these modifications consist of changes in circuitry that determine the functional architecture of brains in very much the same way as genes. Thus, although the basic blueprint of our brains is not very different from that of our cave dwelling ancestors as the genetic outfit has not changed much over the last 30,000 years, our brains differ from theirs because of epigenetic modifications that our brains experienced while developing in a highly complex socio-cultural environment.

Right from birth our brains are exposed to a much more complex environment than the brains of our ancestors because of the countless artifacts that the various cultures have invented and added to nature. Moreover, our children are exposed to highly sophisticated languages that convey not only factual knowledge but also the experience with complex relational structures. And finally, there is intentional education that sets in right after birth and is intensified until it occupies nearly the whole wake time as children grow older. Thus, through the combination of epigenetic modifiability of brain architectures with intentional education, a mechanism is introduced in the evolution of *Homo sapiens* that permits reliable transmission of knowledge acquired during lifetime to the subsequent generation.

This is not the place to analyze in detail similarities and differences between genetic and epigenetic modes of information transmission. However, there is one important difference that I would like to highlight because it has far reaching consequences for our concept of tolerance. The knowledge about the world that has been acquired during biological evolution and that governs our perception of the world is similar for all human beings and we share this knowledge in various degrees with the animal kingdom. Although different species have evolved into different ecological niches, the constraints to which cognitive systems had to adapt were rather similar. This is why we usually agree with respect to the perception of phenomena characterizing the precultural world. We share the inborn priors with other human beings and, therefore, as reflected by the similarity of the genetically determined features of our brain architectures, rightly assume that other human beings perceive the world in very much the same way as we do. Still it may

occur in certain situations that subjects come to different conclusions concerning the perception of non-culture specific properties of objects. A color blind person for example bases her/his perception on different priors than a color competent subject. Both experience the same object in different ways and it would be hard for them to find out who is actually right. In this case, the dissent can be resolved by consulting 'objective measurement devices' and thereby including a third person perspective.

However, in case of the perception of realities that cultural evolution has generated, it is much less likely that all human beings agree. Priors installed by post-natal epigenetic shaping are much less likely to resemble each other than priors acquired during biological evolution. One of the hallmarks of cultural evolution is diversification. Accordingly, it is very likely that the priors acquired by early exposure to different cultures exhibit culture specific differences. As outlined above, the knowledge acquired during early development remains implicit because of childhood amnesia. Nevertheless, this implicit knowledge, just as the evolutionary acquired knowledge, will determine how subjects perceive the world around them. It follows from this that individuals raised in different cultures will base their perception on different epigenetically transmitted priors and, therefore, are likely to perceive realities, in particular those brought fourth by cultural evolution – the so-called social realities – in different ways. In situations where these perceptions are based on implicit priors, subjects will be absolutely convinced that the way in which they perceive a particular condition is the only way it can be perceived – just as we are convinced that there is only one way in which a particular object can be perceived. Subjects raised in different cultures with differing implicit priors about social realities will perceive the same social setting in perhaps very different ways, both experiencing their perceptions as evident and not questionable. However, in this case no 'objective measurement device' can be consulted. The categories of right and wrong become meaningless in this context. Both subjects have the same right to claim as correct what they perceive.

It is obvious that conflicts arising from diverging perceptions of the same social realities increase in frequency and severity as globalization forces different cultures to interact with each other. It is also obvious that the only recipe to cope with such conflicts is tolerance. However, the classical strategy to practice tolerance has been based on the implicit assumption that eventually a distinction between right and wrong is possible. If there is sufficient consensus about the perceived among members of a sufficiently large group of people, it is usually taken for granted that the

respective perception of conditions is correct. Deviating perceptions of others are then considered as false and it is believed to be a tolerant attitude if the dissenting minority is allowed to continue to maintain its 'false beliefs' as long as these do not really challenge the system of the majority. However, as history has shown over and over again, this non-reciprocal concept of tolerance does not solve but generates problems because of its humiliating effect on the tolerated minority. The worldwide surge of terrorism is but one of the many deplorable consequences.

The scientific evidence on the dependence of perception on priors and on the acquisition of priors by epigenetic shaping of brain architectures forces us to adopt new concepts of tolerance that are based on strict reciprocity. Perceptions that are based on implicit priors cannot be changed by argument, they remain evident to the subject and resist relativism. In addition, when it comes to the perception of social realities, distinctions between right and wrong, between correct and false perceptions are impossible. Therefore, members of all cultures have to be credited that what they perceive is correct, even if the respective perceptions diverge. Thus, mutual recognition and reciprocal tolerance are required. Tolerance needs to be granted on a mutual basis and may only be withheld when the respective other violates the rules of reciprocal tolerance. These rules, in contrast to the differing perceptions of realities, are objectivatable and can be codified. Rather than attempting to defend belief systems based on idiosyncratic perceptions of social and cultural realities mankind, if it were to cope with the tremendous problems of globalization, will have to invest massively into the definition and defence of rules securing reciprocal tolerance.

DISCUSSION ON PROF. W. SINGER'S PAPER

PROF. ARBER: Thank you for enriching our knowledge. I think that even at my age I will get some imprinting from what I hear in this room… this is quite nice. But I was a bit surprised to hear that brain development is completed at around the age of twenty.

PROF. BATTRO: Thank you, Wolf, for this thought-provoking presentation. Of course, this changed from biological to cultural evolution, it is key. I think it is what Prof. Zichichi said of Big Bang 3, perhaps, and the difference which is essential, as you said, in education – this is as a comment – it seems that the great change between our species and the other species is teaching. It is impossible to have a model of teaching in animals. What we say is that teaching is unique to the human species and this is why we need more work on the teaching brain, which is difficult, but we have a way out because, as you said perfectly well, children can teach because at 3 or 4 years old they already have a theory of mind. Without a theory of mind we really cannot teach. Therefore we can develop a protocol to do work on the teaching brain, we need more teaching brain research.

PROF. W. SINGER: I agree.

PROF. JAKI: I am puzzled by the fact that, although our brain is very specifically constrained in volume, in weight, in molecular composition, nevertheless that brain, which is the basic tool of our thinking, can perceive that it is constrained. My perplexity comes from a recollection of Goedel's incompleteness theorem. Nothing is more constrained than the laws of arithmetic, which is the basic form of all mathematics and Goedel discovered, in the late 1930s, that a mind, or the mind, which is bent on consistency, this is the basic rule of reasoning, can discover that it cannot achieve a full consistency within the laws of arithmetic. Do you know of any publication in brain research that considers the applicability of Goedel's theorem to the very problem you have discussed?

PROF. W. SINGER: I am not aware of a paper. I am only aware of epistemic circles.

PROF. WOLTERS: You differentiated or distinguished cognitions from reasoning and said that cognitions are a result of adaptation and have to be viewed in a functional perspective, which is different from truths. And then you said, well, reasonings may also be of that sort. My question is, what do you think of when you say 'reasonings', and this is related very much to your last point, where you pointed out that truth is cultural relative for quite a segment for things we say.

PROF. W. SINGER: Well, it is just by extrapolation, since I see that the same substrate that is responsible for perception is also responsible for reasoning. It is very likely that it is also adapted in an idiosyncratic way to a particular segment of the world in which certain contingencies are the case, from which would follow that the logic that we apply has evolved from experience with a narrow section of the world. Whether this is generalisable or not I don't know, but when I heard all the physicists talk on Friday and Saturday I sometimes had my doubts.

PROF. CAFFARELLI: I wonder if you gave any thought about devising some strategy for science education? For instance, there is this counterintuitive fact that if you teach science or mathematics specially, generalising from the particular seems to be natural way. Nevertheless it has been shown that I do not know at what level of age, if teenagers or children who are basically taught axiomatically, have a bigger power of generalising than the ones that have been…

PROF. W. SINGER: I have not given it any deep thought, but I thought it would perhaps be helpful to get our children used to complexity theory, so that they acquire an intuitive feeling for non linearity already from kindergarten, because the problems that we will have to deal with in the evolving world are of that kind and we are very poor in dealing with them.

PROF. MITTELSTRASS: A very short question. What distinguishes your principle of tolerance from the principle of relativism?

PROF. W. SINGER: Well, I think the principle of relativism admits that you can have different perceptions of the same thing and that there may be not

an absolute independent view that allows you to perceive things as they may be in reality. This notion of relativism forces one to defend concepts of tolerance That's how I would see it.

PROF. ABELSON: I thought you rather beautifully give the boundary conditions for the problems that Professor de Duve raised in his lecture, that is, in a world that is increasingly dangerous we have to evolve some shared values. The threat of global warming, for example, might lead to shared values.

PROF. W. SINGER: But I think we might be able to agree on certain architectural features of societies that need to coexist to make it possible for them to self organise towards stable states. I do not believe at all that it is possible to manage these complex systems, be it the economy or any social or political system, in a top-down fashion by some meta-intelligence who would know how to do it. Because you cannot really steer those systems, one needs to implement architectural features, like in the brain, that stabilise them through self-organisation. I do not know how they work but this is where we have to inquire, I would think.

THE LANGUAGE OF GOD

FRANCIS COLLINS

In choosing a topic for this landmark discussion, I took seriously the fact that we are here to talk not only about science, but how science interfaces with spiritual perspectives. I could have used my time to talk exclusively about genome science, because that field is undergoing enormous exponential growth right now. I will indeed talk about that, but I also would like to try to provide, from my own personal perspective, some comments about how these advances can be synthesized with belief in a Creator God. After all, the effort to explore such a synthesis is a major point of this meeting.

I often begin conversations about science and faith with a pair of images representing the two major worldviews that various peoples of the world are debating: one image is the rose window of a cathedral, with its beautiful radial pattern; the other is a view of DNA, a different one than you usually see, looking down the long axis of DNA and also showing quite a beautiful radial picture. There are many who argue at the present time that we have to make a choice between these two worldviews. Certainly, in my country, the USA, such shrill voices of opposition are heard much more commonly than those who argue for possible harmony.

Is it a mistake to try to discuss science and faith in the same room? I often reflect on the greatest commandment as spoken by Jesus, 'Love the Lord your God with all your heart, with all your soul, and with all your mind' (Matthew 22:37). Isn't doing science a way of loving God with all your mind? It certainly doesn't sound as if Jesus thought there was a conflict between faith and reason.

The Human Genome Project and the practice of medicine

The Human Genome Project, which I had the privilege of leading, had an audacious goal: to read out the entire DNA instruction book for *Homo*

sapiens, more than 3 billion base pairs. At the time of the beginning of this project the technology for doing this was clearly not in hand, so one could say this was a truly an ambitious objective. However, all of the goals of the Human Genome Project were achieved in April 2003. Throughout the course of the project, all of the DNA sequence from the human genome was made immediately available on the Internet every 24 hours, so that anyone who had ideas about how to use it for human benefit could begin work immediately.

The scientists who participated in the Human Genome Project hailed from six countries of the world. They, too, helped us identify where to go next. An iconic diagram featured in a *Nature* paper in April 2003 depicted a metaphorical building that we were now prepared to construct, resting upon the foundation of the Human Genome Project, but now applying that knowledge to biology, health and society.

Many of the 'Grand Challenges' outlined in that rather audacious publication have already been achieved, thanks to the rapid pace of genome research. Specifically, remarkable progress has been made in identifying variations in the human genome that are playing a role in risk of disease. Your genome and mine are about 99.6% the same. In that small percentage where we are dissimilar, most of those differences do not have medical consequences – but some of them do. For me, as a physician geneticist, a major goal was to try to identify what some of those genome glitches were that play a role in diabetes, heart disease, or cancer. While we had done a very good job of finding those glitches for diseases that were highly heritable, like cystic fibrosis and Huntington's Disease, until very recently we had not had much luck with the common diseases that fill up our hospitals and clinics. All that has changed in the last three years.

Building upon the success of the Genome Project, another project called HapMap provided a catalogue of human variation that made it possible in a comprehensive way – not based upon candidate genes, but looking at the entire genome – to scan and identify those variations associated with diseases that are non-Mendelian in their inheritance. The first success was age-related macular degeneration, mapped to chromosome 1 to a gene called 'complement factor H'. No one expected that gene to be involved in this disease, and yet a common variant in this gene is a major risk factor. Since that discovery, much has happened: in 2006 there were three more successes. With the full availability of the HapMap and the advent of very low-cost genotyping in 2007, discoveries really started to appear, and became a full-fledged deluge by 2008. As a result no less than 400 of these

well-validated genetic variations associated with common disease have emerged, mostly in the last two years, shedding dramatic new light on the causes of diabetes, heart disease, cancer, mental illness, autoimmune diseases, asthma, and many others.

These successes provide us with powerful new targets for therapeutics. They also present the opportunity to provide individuals with a refined estimate of their future risk of disease, depending on which of these variants they happen to carry. Already there are companies who offer you the chance to test your own genome for about a million different variants, for a cost as little as 400 US dollars. Whether that is premature or not is a matter of some debate; while the tests are scientifically based, most of the heritability of common diseases has not yet been uncovered, and there is limited evidence that knowing this information actually improves outcomes. But the era of personalized medicine is at hand.

As technology advances, we will soon be able to examine individual genomes in their entirety, identifying not only the common variants but the less common ones that play a critical role in disease risk. Professor Gojobori already presented information about the way in which DNA sequencing is advancing. This capability has made it possible to tackle problems in a comprehensive way that previously had not been feasible. An important area is cancer. Certainly we have known for a long time that cancer is quite literally a disease of the genome. It arises because of mutations in DNA. It takes an accumulation of several mutations over many generations of cell divisions to reach the point where that cell is truly malignant. If we really want to understand cancer, we need to develop a comprehensive catalogue of all the mutations in the cancer cell. Last year, the first paper describing the full sequencing of a cancer genome was published in *Nature*. It described the complete DNA sequence of a leukemia arising in a woman who had a very aggressive form of the disease. A number of new genes were found mutated in the cancer cells, and were not on anybody's previous list of oncogenes or tumor suppressors. From these findings it is clear that this comprehensive view is going to open up many new vistas in terms of the understanding of malignancy.

Another area that these sequencing advances now allow us to tackle is to look more closely at those non-human genomes that are on us or in us. There are hundreds of trillions of microbes on our skin, in our mouths, and in our gastrointestinal tracts. For the most part these organisms are synergistic with us and assist in maintaining our health. However, the balance between host and microbes can be deranged, and that can lead to illness.

The Human Microbiome Project is a new international program that aims to catalogue these microbial genomes, both in health and in disease. This has not really been possible in the past, as only a minority of these microbes are possible to culture in the laboratory. But they have DNA.

Technology promises even more disruptive advances for high-throughput, low-cost sequencing. An example mentioned by Professor Gojobori is a new approach from Pacific Biosciences that sequences single DNA molecules. I have recently seen a demonstration of this technology, which carries out DNA sequencing in real time using fluorescently labelled tags and massive parallelism. This promises to reduce the cost of sequencing another couple of orders of magnitude and bring it down to the point where a complete DNA sequence can be done for a thousand dollars or less, in a matter of a few hours.

So how will these advances play out in the practice of medicine? Discoveries about causes and treatment of each disease will move at a different pace, but I think we can expect things to happen pretty quickly. Already for some diseases, we are using the tools of genotyping and DNA sequencing to identify individuals at high risk. As just one example, those found to be at high risk for colon cancer can now be counseled to have annual colonoscopy beginning at age 30 (instead of the usual recommendation of age 50).

We also have the opportunity to use the tools of genetics to identify variations that will predict response to drug therapy. This is the field of pharmacogenomics, and promises to provide a better opportunity for a patient and physician to choose the right drug at the right dose.

I would predict, however, that the major, long term impact of the genomic revolution will be the discovery of new therapeutic opportunities, building on knowledge about biological pathways that are fundamental to disease pathogenesis. Some of these new treatments will be gene therapies, where the gene itself becomes the treatment. A recent exciting example of this is in the treatment of a particular type of blindness. But perhaps an even more widespread consequence of our new knowledge of the genome will be in the form of drug therapies, because of the new targets that are being discovered using the genomic approach.

It thus appears inescapable that medicine will undergo a major revolution in the course of the next ten years. Unfortunately, however, I do not think that the medical profession is currently well prepared to respond to this revolution, because of the disparity between the rapid nature of these discoveries and the relative slowness of the medical education system to incorporate them into training.

Evolution and the study of genomes

I would now like to turn to the evidence coming from these genome studies with regard to evolution, as that is a major topic of discussion at this meeting. If there have been legitimate doubts about whether Darwin's theory was correct, based upon so-called 'gaps' in the fossil record, those doubts have largely been swept away by the study of DNA. In fact, if Darwin had tried to imagine a compelling way to demonstrate the correctness of his theory, it is hard to see how anything outside of a time machine would have been better than comparative genomics.

Not only have we sequenced our own genome, but recent covers of *Nature* and *Science* magazines show successes for other genomes as well: the mouse, the chimpanzee, the dog, the honey bee, the sea urchin, the macaque, and the platypus. We have draft or complete genome sequences now for more than two dozen vertebrates. If you feed these genome sequences into a computer and ask it to create a relatedness tree between the organisms, it will produce a startlingly close match to evolutionary trees that have been generated from fossil data or from anatomical features.

But in my country, the USA, there are still many who reject the evidence that all of these organisms, including humans, are related by descent from a common ancestor. A recent poll shows that forty-five percent of Americans believe that the earth is less than 10,000 years old, and that humans were specially created by God. This view is in serious trouble, once one looks at the DNA evidence. Certainly, one could argue that God used the same motifs repeatedly to produce all of these organisms as acts of special creation, and that might explain the general relatedness at the DNA level. But when we look at the details, it is clear that this particular alternative view cannot be sustained. As an example, consider human chromosome 2. Chromosomes are the visible unit of heredity in a cell. We humans have 46 of them, made up in pairs. One can look under the microscope at a cell that is about to divide, and observe the chromosomes. It is noteworthy that human and chimpanzee chromosomes look a lot alike with regard to their size, their banding pattern and so on. The one exception, however, is that we have human chromosome 2 as our second largest chromosome, while chimps do not. They instead have two smaller ones. Gorilla chromosomes look similar to chimps; making us the outlier amongst primates.

There has been a prior supposition that perhaps in the lineage leading to humans there was a fusion of two smaller chromosomes giving rise to our chromosome 2. That finding has now been subjected to exquisitely

detailed analysis from the DNA sequence data. There are special sequences at the tips of all chromosomes. These are the telomeres; a particular sequence, TTAGGG, appears over and over again in order to prevent fraying as the cell divides. It is interesting to note that when you look at human chromosome 2, there are telomeric sequences in the middle, exactly in the position where you would predict such a DNA footprint would have been left by a fusion between two ancestral chromosomes.

Another revealing example of our common ancestry with other animals also explains why sailors contracted scurvy on those long sea journeys. If we look at the order of genes in multiple mammals around a particular gene called GULO, we will see the order of genes is the same in humans, cows and mice, as well as many other vertebrates. But this is an interesting example, because the gene GULO, which stands for gulonolactone oxidase, is a pseudogene in humans (and in other primates) – meaning that it has sustained a knockout blow, decapitating its front end completely so that it lacks the first part of the coding region. It is utterly nonfunctional. Well, the product of that gene normally catalyzes the final step in synthesizing ascorbic acid (vitamin C). Unable to make their own vitamin C because of the non-functional GULO gene, sailors developed scurvy when they did not have access to vitamin C. But the mice on the ship, possessed of a functional GULO gene, did just fine.

Looking at that data, it is extremely difficult to argue that we humans are created as a special separate lineage compared to other animals. One would have to infer that God intentionally inserted a non-functioning GULO gene in just the position to mislead us into thinking that descent from a common ancestor was correct. This model would put God in the position of being a DNA deceiver, which does not seem consistent with other basic tenets of religious belief.

Catholics are in general much more comfortable with the shared descent of humans and other animals, so I probably do not need to make this case so strongly to this particular audience. But for many protestant evangelical Christians in America, this is still not an easily accepted conclusion.

THE HARMONY OF SCIENCE AND FAITH

Let me turn now to another question. Simply stated, 'If evolution is true, does that leave any room for God?' Let me begin with a personal perspective. I was not raised in a religious tradition. Until my twenties, I con-

sidered myself an agnostic, and ultimately an atheist. It was actually my involvement in medicine that forced me to consider issues of life and death in more than hypothetical ways, and my involvement in science that convinced me that the purely materialistic approach can be unnecessarily limiting for the kinds of questions that we humans want to ask – such as why there is something instead of nothing. These intellectual explorations ultimately led me, to my great surprise, to Christianity.

It didn't take long for my colleagues to point out that they thought I was on a collision course between the scientific and spiritual worldviews. As a geneticist, evolution was fundamental to my understanding of biology. But didn't I know that evolution and faith were utterly incompatible? Certainly that case has been smoldering ever since 1859, and has been recently made rather loudly by some of my colleagues, such as Professor Dawkins.

In his book, *The God Delusion* (a rare book that does not require a subtitle), Dawkins uses evolution as one of his strongest arguments against the plausibility of God. He insists that once Darwin arrived at his theory of evolution the need to describe a Designer or Creator went out the window. But in my view and that of most thoughtful believers, Dawkins makes a category error by trying to use scientific arguments to weigh in on the existence of the supernatural.

Nearly two years ago, I engaged in a debate with Richard Dawkins for *Time* magazine. The exchange is still available on the Internet.[1] Ultimately at the end of it, Dawkins admitted this category error to a certain extent, recognizing that science cannot exclude the possibility of a supernatural God, even though he thought it highly unlikely. But he stated that if there was such a thing as a supernatural God, it would be much more grand than any of us could imagine. That's exactly the God believers are talking about, I said!

So we are back to the question, 'How can evolution and faith be reconciled?' If you will indulge me, I would like to provide a rather personal response. I understand the risk of doing so here, in front of esteemed scientific and theological colleagues. I am an amateur theologian and philosopher. But it seems to me that there is a readily-achieved synthesis that is entirely compatible both with what we know scientifically, and with what the basic Abrahamic principles say about God the Creator. Here it is: Almighty God, who is not limited in space or time (an Augustinian concept from 400 AD) created this universe with its parameters precisely tuned to

[1] See http://www.time.com/time/magazine/article/0,9171,1555132-1,00.html.

allow for the development of complexity over long periods of time. God thus endowed Creation with amazing potentialities. That plan included the mechanism of evolution to create the marvelous diversity of living things on our planet – and, most especially, human beings, with minds created in God's image. Evolution was sufficient to prepare the 'house' for all this, namely the human brain in all of its elegant complexity. But there was something missing until the additional spiritual component of humanity arrived. The story of the Garden of Eden is then a description of God's provision of additional gifts to humankind: free will, the soul, and – I know this will be controversial – the moral law. The moral law, the knowledge of right and wrong, is universal and unique to humanity, though its interpretation is strongly affected by culture. Biblically we learn in the story of Adam and Eve that we humans used our free will to break the moral law, leading to our estrangement from God. For me, as a Christian, it is Christ who provides the solution to that estrangement.

This synthesis of Biblical and scientific perspectives has traditionally been called 'theistic evolution'. But I don't think that is a great label. It turns a lot of people off because it sounds like evolution is the noun and theistic is the adjective, implying God is less important than Darwin. So, in my book *The Language of God*, I proposed an alternative term: *Bios*, meaning life, through the *Logos*, or the Word – or simply *BioLogos*, God speaking life into being.

As you may imagine, there are a variety of objections to this perspective. For instance, one often is asked: 'Didn't evolution take an awfully long time?' This question is a concern of many Evangelicals who cannot imagine why God would have taken so long to get to the point (humanity). They often ask, 'Why didn't God just snap his fingers and make it happen?' Well, again, if God is outside of time this is our problem, not God's problem. Another related objection is: 'Isn't evolution a purely random process?' This question seems to take God out of it. As one of several possible responses, I would posit that if God is outside of time, then randomness to us may not necessarily be randomness to God.

Intelligent Design proponents ask, 'Is evolution really sufficient?' In other words, aren't there biological structures, like the bacterial flagellum or the human eye, that are just too complicated for evolution alone to have produced? Each of these structures has many subunits, and when just one of them is knocked out, the whole thing stops working. So how could such complexity have arisen by natural selection alone? Well, those questions reveal a basic misunderstanding of the stepwise fashion by which such multiprotein complexes come into being. A recent paper from *Nature Reviews Microbiolo-*

gy points out how many of these intermediate steps are being discovered for the flagellum. Intelligent Design, in my view, is turning out to be a major misstep. It is both bad science, representing a God-of-the-gaps approach, and bad theology, portraying God as a rather inept Creator that had to keep intervening along the way to correct deficiencies in the original plan.

Proponents of evolutionary psychology have objected to my portrayal of the moral law as a signpost to God. Can't this be a consequence of evolution? Isn't altruism just a human behavior that has led to greater reproductive success of the species, and that's all? There are, to be sure, many aspects of altruistic behavior that are consistent with explanations provided by evolutionary psychology. They include: 'kin selection', which explains generosity to your relatives since you share your DNA with them, and if you help them be reproductively successful your own DNA is succeeding too; 'reciprocal altruism', which argues that our own altruism is often driven by a hope for some reciprocal benefit in the future from those we have shown kindness; and even 'group selection', which proposes that altruistic behavior of a group of individuals provides advantages to the whole group, even if it harms a few individuals' chances of reproductive success along the way. Martin Nowak at Harvard expounds on these models in his very interesting game theory studies. He concludes, however, that for group selection to work, one must be hostile to anyone who is not part of the group. But is that the kind of altruism we most admire in humans?

Imagine for a moment the person who, with great risk to themselves, reaches out to someone they do not know, someone who is part of another group. Evolution, ultimately, would predict hostility. But when we see this kind of radical altruism, we admire it. As an example from about a year ago, Wesley Autrey watched with horror as a young man standing on the subway platform in New York City went into an epileptic seizure and fell onto the tracks, with train No. 1 quickly approaching. Without hesitation, Wesley leaped onto the tracks. He covered the still seizing student with his own body and wedged them both between the tracks. The train rolled over them, and they both miraculously survived. Wesley was black. The student was white. They had never met. Stories like this one electrify us, and we are likely to point to such actions as representative of the best of human nobility. And yet, from an evolutionary perspective Wesley's action was a scandal, taking an enormous risk of sacrificing his own potential reproductive future to save someone he didn't even know.

A final objection to BioLogos, raised especially in my own Evangelical Christian circles, is the question about whether evolution conflicts with

Genesis 1 and 2. But as strongly as these concerns are raised, I see this as an unnecessary conflict. In this regard, I am greatly rewarded every time I open one of the four commentaries that St. Augustine wrote about Genesis. He was a theologian who thought deeply about this subject and who can hardly be accused of trying to retrofit his views into Darwin's theories – since St. Augustine wrote down his views on Genesis more than a thousand years before Darwin walked the earth. Augustine ultimately concludes that there is no way for any single interpretation of Genesis to be declared correct, and he provides a warning that ought to be heeded today by many churches, especially in my country. Augustine cautions, 'In matters that are so obscure and far beyond our vision, we find in Holy Scripture passages which can be interpreted in very different ways without prejudice to the faith we have received. In such cases we should not rush in headlong and so firmly take our stand on one side that if further progress in the search for truth justly undermines this position, we too fall with it'.

Finally, before concluding I would like to respond to Professor Zichichi's statements that took aim at the discipline of biology. Contrary to his view, I do believe that biology has arrived at a new phase of scientific rigor. The era of complete genomes, and the ability to understand life in a digital way, allows biology to take its rightful place as a truly quantitative science alongside physics and chemistry. Although this was not true a few decades ago, it is clearly true now. Evolution is at the core of these advances. I therefore associate myself with Theodosius Dobzhansky, one of the leading lights of evolutionary thought in the 20th century and a Russian Orthodox Christian, in his statement, 'Nothing in biology makes sense except in the light of evolution'. I do not know how we could do biological science at all without accepting the evolutionary paradigm. Nevertheless I agree that evolution does not have, and will never have, an answer to the 'why' question. That is a question that science cannot answer; it is a matter for faith to address.

Thank you, again, for the gracious invitation to join this distinguished group at the Pontifical Academy, and to spend time discussing these important worldview questions.

DISCUSSION ON PROF. COLLINS' PAPER

PROF. ARBER: Thank you very much for this interesting presentation and the outlook into the spiritual world.

PROF. M. SINGER: Tomorrow, when I talk about intelligent design, I will, try to show how it is that, in the U.S., many people who do not accept evolution disagree with Francis' conclusions.

PROF. COLLINS: Yes, they do!

PROF. PHILLIPS: So I would like to ask about your discourse about the development of morality, because you gave a picture that was very different from what Wolf Singer gave before.

PROF. COLLINS: Maybe, maybe not, that is a good question.

PROF. PHILLIPS: Well, anyway, the way I perceived it he was emphasising the diversity of moral understanding and encouraging us to take the point of view that there is no right or wrong. You, on the other hand, emphasised a certain commonality and, in fact, in describing how someone may object to the concept of God by explaining that commonality of moral understanding came about through some evolutionary process. So it seemed to me, at least, that you were emphasising commonality and that Wolf Singer was emphasising diversity and my own perception of moral understanding across human cultures is that I am much more impressed by the commonality than I am by the differences. In other words, the differences I see more or less as things that have to do with case law whereas the commonalities have to do with general principles. Well, so can you correct my impression? But I would like to have your perspective on that difference between what you said and Singer said.

PROF. COLLINS: it is a great question and I think, actually, the differences may not be very great. C.S. Lewis, in his book *The Abolition of Man*, has an

appendix that goes through the monotonous review of cultures down through history and across the world and their moral behaviour and the conclusion he arrives at, which, I think, is shared by those who have looked at that data carefully, is that the idea that there is such a thing as right and there is such a thing as wrong appears to be a universal human attribute. We do not find exceptions to the idea that there is such a thing as moral law. Where we find vast differences is how that is interpreted, which is what I took Professor Singer's discussion to mean, that that epigenetic modelling of the brain takes this fundamental law about right and wrong and it decides what goes in which category, and that is profoundly culturally affected by learning and by what your parents model for you. So you can see cultures that we see today as having done horrible things who at the time were convinced that they were behaving in a right fashion, based upon their own epigenetic modelling of the moral law in their brain or other forms of self delusion, perhaps, and I am sure our own culture has its own form of self delusion. But I do not think that that kind of variation in social mores or cultural mores can get us away from the fact that there is, apparently, in human commonality, this notion of a right and a wrong, again a notion which I find difficult to completely explain on totally materialistic grounds. And let me just say one other thing about that: if one wants to do that, and certainly the evolutionary biologists and my friend Dawkins will try to do so, you have to carry that all the way to its ultimate conclusion, which is a very uncomfortable one, which is that, in fact, good and evil are not real, that these are illusions, that we have been hoodwinked by evolution into imagining that there is such a thing and that we are driven by this evolutionary hoodwinking into certain kinds of behaviours that, perhaps, now that we are so smart, we should not have to be regulated by anymore. And yet, that is a very difficult thing even for the strongest atheists to say because they are quick in many instances to then point to religion and say, that is evil! Well, where do you get off, saying that anything is evil if you are a pure materialist evil goes out the window and so does good.

PROF. CIECHANOVER: Francis, about disease-related genes. I think several papers were published a few weeks ago showing that, in several aggressive cancers like glioblastoma multiforme and pancreatic cancer, there are about seventy or eighty mutations so how do we know which ones of them are causative and which ones are bystanders? Then, the next one is, of the causative, what is the minimal combinatorial or combinatory that is necessary in order to cause it? And then, if you can answer that, and about the predictability, not all of those mutations are inherited, some of them have

been accumulating for a long time, so it would be very difficult to predict susceptibility to disease by sampling at a certain time and you do not know what will be the detrimental time point that makes the predictability certain, so all this adds to a huge complexity of what I thought was a little bit simplistic picture that you described.

PROF. COLLINS: It is a fair point. I would say that the two papers you refer to from the Vogelstein Lab may suffer a bit from inadequate statistics to tell the difference between a so-called driver mutation, which is actually contributing to the malignant phenotype, and a passenger mutation, which is just something that has arisen during the course of cell division and does not actually have a consequence. I think a more careful analysis using larger numbers of tumours, coming on glioblastoma, for instance, from the cancer genome atlas, or looking at almost 200 such tumours, comes up with a shorter list. It is still a long enough list to be a bit daunting in terms of how to put this together. A similar study on 186 samples of adenocarcinoma of the lung published last week also comes up with a list of more than a dozen, but they actually fit rather neatly into four pathways and that is somewhat reassuring that there seem to be four fundamental pathways for adenocarcinoma of the lung where virtually all the tumours have a glitch in each of those pathways, it is just not always the same point where the pathway is disrupted. That would suggest that, if you can come up with therapeutics that target those pathways, you would not necessarily have to have a drug for each individual, which would be obviously untenable, you could come up with something more general. Similarly for the leukaemia patient that I mentioned. But a lot of sorting out has to be done. As far as predictability, almost all of the genetic glitches that are being discovered are somatic, and the ones that involve inheritable risks are much less numerous and for the most part, with some exceptions like BRCA1 or one of the HNPCC mutations, much of those other cancer predispositions are pretty small in terms of their odds ratios. They might imply a need to do better surveillance for a slightly higher risk, maybe we can personalise that approach, but they are not going to make a huge influence and again, the challenge I think we are facing, is how do we get to the point where every cancer is immediately sequenced to see exactly what is the array of mutations. Then how do we take that information and map it on to our set of therapeutics that target various pathways and try to get the right mix, and it is not going to be *a* drug for cancer, it is going to be combination therapy and there are all kinds of problems with that but that has to be the right answer.

PROF. WOLTERS: It occurred to me that you are subscribing to Stephen Jay Gould's NOMA principle.

PROF. COLLINS: No, no, no, I will resist that but go ahead!

PROF. WOLTERS: OK, but anyway, if you think that religion is the source of our knowledge of right and wrong I would contest this with respect at least to Christianity, because in the Bible we find quite different moral recommendations coming from God such as ethnocide, rape, killing the children of the enemies and so on, and we find all the very nice things we usually ascribe to the Bible, so it occurs to me that you need, in order to distinguish between right and wrong, a principle you do not find in the Holy Scriptures but you might find in philosophical discourses since the time of Aristotle going on to Kant and utilitarianism and other positions.

PROF. COLLINS: Good comment, and again I may have come across in some way as implying I think there is a firewall between the worldviews that are spiritual and the ones that are scientific. I do not see it that way, I do not know how I would live in that kind of circumstance where I had to be really clear about which part of my brain was active on any given Thursday afternoon. It seems to me that was an unnecessary and artificial division. More importantly, though, if you are approaching a particular question, you have to ask yourself, is that a scientific question or is it a philosophical or theological question and then use the right tools to try to address it. But I would very passionately argue for the need then to synthesise those worldviews into a harmony, not an artificial separation. I certainly agree that much of what I have said has been said vastly better by other philosophers going back to Aristotle and certainly to Kant, whose statement about his increasing awe at the starry heavens above and the moral law within very much describes what brought me ultimately to become a Christian.

PROF. MITTELSTRASS: Again, a very short question. What is the reason for the claimed necessity of synthesising different worldviews? Isn't man the animal who can live in different worlds, like the world of science and the world of faith, the world of science not claiming orientational knowledge and the world of faith not claiming scientific knowledge?

PROF. COLLINS: Well, perhaps we are getting into a bit of a semantic issue in terms of what is harmony and what is separateness. I interpret

Gould's NOMA proposal as proposing a very strong division between the two, that one really cannot allow any overlap at all in your thinking about the causation of important events or about the answers to important questions. I guess I am rebelling against that idea that the firewall has to be that severe. At the same time I take what you are saying as very much my own perspective that yes, I am delighted and blessed to live within both a scientific and a spiritual worldview, but I reject the idea that I have to decide at any given moment that I am doing one and I cannot think about the other one at that point.

PROF. BERNARDI: There was a famous episode, a long time ago, when Napoleon went to visit the French Academy of Sciences and you know very well that Laplace explained a number of things about movement of planets and so on and, at the end, Napoleon asked the question, and where is God in all you said? And Laplace replied that he did not need that hypothesis. He did not deny at all the existence of God, he simply said, there is no need for God to explain the movement of planets. I think that trying to use science to prove the existence of God is a fallacy and I do not think that the argument of evolution or the argument of the gene are good enough to prove something that is outside the scope. So, I think we should be very careful in not trying to see science or scientific points like evolution as providing any proof. There is no proof for or against it.

PROF. COLLINS: If I came across as suggesting that I had found one then I have misrepresented myself, because I certainly agree that proof for God's existence will not be found within the area of science. I do think it is fair to say, and maybe I am particularly reflecting about my own experience, that there are observations about nature, such as the Big Bang, such as the fine-tuning of the universe, such as the fact that mathematics actually works to describe the way the universe is put together, such as the fact that there is something instead of nothing, and maybe even the fact that there is a moral law, and that is probably the most contentious of my list, that are worth reflecting on. They are not proofs, they will never become proofs of God's existence, but they are somewhat of an antidote to the sweeping atheism which seems to be so prominent in the scientific community, so I think it is worth bringing them up from time to time.

PROF. SZCZEKLIK: I have a question dealing with the first part of your presentation concerning complex diseases, say, atherosclerosis or dia-

betes. Despite the enormous effort which has been done over the last years, genome-wide search etc, the value of genetic factors for clinical diagnosis and prevention is still very very limited. Could it be so because the phenotypes of the disease we clinicians diagnose, are very inhomogeneous, or the sequencing technologies are not sensitive enough to detect rare variants? Do you have some other explanation, or is it just too early to expect the solution?

PROF. COLLINS: It is a question that is on everybody's minds. We are both exhilarated by discovering some of the genetic risk factors for common disease which have been so elusive but also puzzled by why they account for such a small fraction of the heritability that other studies have indicated must be there. For diabetes, for instance, a disease my own lab has been working on for 15 years, we are exhilarated now by having 18 genetic risk factors that we are quite sure are right and yet, together, they only account for about 10% of the heritability that must be lurking somewhere in the genome. Some of this may be dilution by lumping together phenotypes that really do not belong together, and therefore reducing the signal but it cannot be all that. Some of it may be, in fact, rare variants of large effect, which our current methods are not detecting but sequencing will, or copy number variants which are not readily detected by our current technologies but that many people believe are going to be a lot more important than we had previously imagined. Or perhaps there are many of these very common variants with very small effects, a very long tail of the distribution. We are going to figure that out, that will become apparent. And then, at that point, the ability to make predictions about risk will start to get better. Right now it is small because we are so early on but yes, there is this issue and, reflecting on Vera Rubin's term, the people in genetics are talking about the 'dark matter' of the genome. Where is the rest of the heritability? It is in there somewhere, we just have not found it yet!

THE BUNCH OF PREHUMANS AND THE EMERGENCE OF THE GENUS *HOMO*[1]

YVES COPPENS

Firstly, I want to thank the Pontifical Academy of Sciences for inviting me to this very important, very interesting, quite exceptional meeting that I am enjoying very much. As you know, I am a paleontologist and, more precisely, a field paleontologist, so I will try, through the fossil record, to tell you what I think could have been the history of man, which means the history of the last 10 million years.

It is well known today, especially today, as you have seen, that Bonobos, *Pan paniscus* and chimpanzees, *troglodytes*, are the creatures closest to us in nature. In an evolutionary way of thinking, it means that they, and we, have common ancestors. Because all primates are tropical, and because Bonobos and chimpanzees are African, there is some probability that those ancestors would have been tropical and African. Moreover, the morphological, anatomical, physiological, genetic, molecular and even ethological distance between these cousins and ourselves allows to situate our last common ancestor somewhere in the upper Miocene, which means about 10 million years ago. So we have the place, Africa, and we have the time of existence, 10 million years ago, of our last common ancestor. We will travel through these ten million years in a chronological order in four parts: first part, 10 million years ago, the time of the last common ancestors; second part, the prehuman, before man on our side; third part, the emergence of man; and fourth part, the evolution and expansion of man.

First part, what do we know about the apes in Africa at these late Miocene times and who do we know could pretend to be the last common ancestors to man and chimpanzee? We have three main candidates to

[1] Transcript unrevised by author.

answer this question: *Chororapithecus abyssinicus* from Ethiopia, 0.7 to 0.1 million years old; *Nakalipithecus nakayamai* from Kenya, 8.99 to 8.80 million years old; and *Samburupithecus kiptalami* from Kenya, 9.6 million years old. This very modest fossil record does not allow, of course, to tell who, among these candidates, is the closest to the last common ancestor of chimpanzee and human. Let us say that we have just started to find some remains of African great apes of the right geological age, which are giving us an idea of what these famous grandparents would have looked like. As a matter of fact, we don't know either where they are really standing in the phylogeny, *Chororapithecus*, *Nakalipithecus* and *Samburupithecus* respectively. We don't really know even if they are preceding the divergence *Homo/Pan* or if they are already engaged in one of the two lines, or if they are engaged in another independent branch, having nothing to do with the *Pan* or with the *Homo* branches.

Second part, let us forget this common ancestor, let us forget, as well, the slice of 10 million years of pre-chimpanzees and chimpanzee to focus our attention on our side, the prehuman and the human side of the divergence. We will divide the ten million years of our affiliation into two major episodes: the prehuman one, from ten million years ago to one million years ago, and the human one, from three million years ago to the present, which immediately shows that the last prehumans coexisted with the first humans. The prehumans are magnificently documented by seven genera and fourteen species. I called this diversity, which is reflecting a diversity of ecological niches, a bunch, instead of a bush, because a bunch seems to me clearer than a bush and it is not a political statement.

The seven genera and fourteen species originated from Central, Eastern and Southern Africa: Chad, Ethiopia, Kenya, Tanzania, Malawi and South Africa. Chronologically they can be organized in three steps: the earliest from seven million to four million years ago, with *Sahelantropus toumaï*, *Orrorin* and *Ardipithecus*, a second step between four million to three million, lasting just one million years, and a third step between three million and one million years ago. The second step, between four million to three million years ago, is emerging after an opening of the landscape. The third step, between three million and one million years ago, is also emerging after drought and because of this drought it is more dramatic than the four-million-years-old opening. In order to try to appreciate the bunch of the fourteen prehumans as a group, we will examine the traits that they shared and the few ones that are dividing them. 1) All the prehumans were tropical without exception; 2) All the prehumans were African, without excep-

tion; 3) All of them, as far as we know, were upright, permanently upright; 4) All the earliest ones, as far as we know, were both bipedal and arboreal (it is the case of *Orrorin*, *Ardipithecus*, *Australopithecus afarensis* – Lucy); 5) All of them, but at different speeds, seem to become exclusively biped (*Australopithecus anamensis*, from Kenya and Ethiopia, 4 million years old, seems to have been the first not to climb anymore, so the first true exclusive biped); 6) All of them have a slowly increasing endocrinal capacity as well as a slowly more complex organization of the brain: more complex convolution, more complex irrigation; 7) All of them show a tendency of the face to reduce its prognathism, its projection, to reach a sort of orthognathism, flat face. *Kenyanthropus platyops* (flat face) and *Australopithecus bahrelghazali*, 3.5 million years old, seem to have been the first and the best in that reduction. But, as far as the teeth are concerned, we can follow two trends: a tendency to a reduction in the site of the cheek teeth, molars and premolars, and a tendency to an increased size, which means two clear adaptations to two different diets. In summary, the prehumans are tropical, African, upright. They seem to have been arboreal and bipedal before adopting, for ecological reasons, an exclusively bipedal locomotion. They all show a brain of increasing size and complexity. They all show a trend to reduce prognathism at different speeds according to their *phylum*. Some of them, at least and at last, show a trend to a reduced size of cheek teeth, while others show the reverse tendency.

Third part: as I just told you, in describing the third step of prehumans, an important drought happened between three million years and two million years ago. The global cooling of the earth appeared, for instance, in the study of oxygen isotopes ratio, oxygen 16, oxygen 18, in the test of microorganisms collected in deep-sea cores. This climatic change is also particularly well visualized in the sediments of the lower Omo River basin in Ethiopia, because the sediments of this lower Omo River are the only ones in tropical Africa to offer a clear, continuous deposit of these times. I worked there ten years, and during those ten years I collected fifty tons of bones. The Omo sedimentary sequence is a superb stratigraphical column more than 1 km thick, particularly well exposed because of tectonic reasons, particularly rich in fossils, including hominids, and particularly well calibrated by biostratigraphic, paleomagnetic and radioisotopic cross-checking dates. And all the fossils collected – I tell you, as far as I am concerned, fifty tons of them – are showing this cooling. I can give you the example, but I will not, of the evolution of elephants, rhinos, pigs, horses, bovines, primates and rodents, during that time, as well as the example of

the evolution of the frequencies of certain plants. All the animals are showing adaptation to a more open and drier environment, grass-eating adaptation, for instance. And all the plants are showing the same tendency towards a less and less humid climate, and a real drier and drier one. I will give you just an example, not to be too long. Having collected pollens in all these levels, I tried to do a ratio with the number of pollen of trees on the number of pollen of grasses and this ratio gave the result of 0.4 for three million years, and 0.01 for two million years, so I think this is clear enough. Well, the answers of the prehumans have been done in just the same direction, and they have been good enough to have given two answers. First answer, one prehuman, probably *Australopithecus afarensis*, but we are not sure, had chosen a larger size of the body and larger cheek teeth, the so-called *Robust Australopithecus*. Second answer, another prehuman, maybe *Kenyanthropus platyops*, maybe *Australopithecus bahrelghazali*, maybe *Australopithecus anamensis*, we don't know very well which one, maybe another one that we have not collected yet, had chosen a larger size of brain and teeth to eat a wider diet, meat included. It is man. The consequence of a larger brain is, of course, the emergence of more reflection and of something like consciousness. And consciousness is starting to build a new environment, the cultural environment, for the first time after four billion years of life, in its natural environment. Well, let me tell you that, for prehistorians, for me, culture is everything that is not nature. So, cultural environment for us means a technical, new environment, of course, but also an intellectual, spiritual, ethical, esthetic one and so on, so it is probably one of the Big Bangs of Prof. Antonino Zichichi, who told about the living matter with reason, and I called this event the '(H)Omo event', 'H' in brackets, because it was on the shore of the Omo river, which is amazing, that we collected this information and to remind of the pioneer role of the Omo area in this demonstration.

In summary, around 3 million years ago an important cooling appeared in the whole world, becoming an important drought in tropical areas, especially tropical Africa, and the whole ecosystem tried, of course, to adapt itself to this climatic change, to this new climatic environment. The prehumans were then part of this ecosystem, and one of these successful adaptations again is called 'man'. To answer Cardinal Christoph Schönborn, who unfortunately is not here, I would say *Homo* looks like a product of nature, a necessary adaptation to a climatic change but is a human being with, for the first time, this capacity of knowing that he knows, he is human since this beginning. So it is a sort of discontinuity in a continuum.

Fourth part: for biological and cultural reasons, the very first species of the genus *Homo* was more mobile than his ancestors because of his diet. He became a carnivore and had to run behind game. This very first species of *Homo* was more curious, also because of his better brain and the beginning of consciousness. He was more equipped because of his new manufactured tools. He was more numerous because of his successful adaptation to the climatic change we talked about. Therefore, he was more mobile, more curious, more equipped, and more numerous. I guess that it is the very first species, and not the second or the third, which moved out of its tropical birthplace and out of its ecological niche. And some environmental reasons can be added at that time to push *Homo* out of Africa, around 2.5 and 2 million years. For the moment we know stone tools in Israel, in one site, which could be, it is not sure but 2.2 to 2.3 million years old, and stone tools in China in three sites, which could be a little more than 2 million years old as well. We know stone tools in Algeria, around 1.8 million years old, stone tools and hominid remains, 1.8 million years old, in Georgia, in Java 1.8 million years old, in Italy 1.6 million years old. So, theoretically, I would think that it is this first *Homo* who had moved as soon as 2.5/2 million years ago, which means that I would not be surprised to meet this very first species of our genus anywhere and everywhere in the old world, in Africa, in Europe and in Asia, at dates between 2.5 and 2 million years ago.

The technical problem is that there are two first species of the genus *Homo*, *Homo habilis* and *Homo rudolfensis*, and I don't think that both moved. Because of a certain number of reasons I cannot develop here, I think finally it was *Homo rudolfensis* who moved. That means that, 2 million years ago, man was everywhere in Africa and almost everywhere in Eurasia, except maybe in the extreme north. And, as he was not demographically numerous enough to exchange genes everywhere and all the time, he first became *Homo erectus* where and everywhere he was, but this *Homo erectus* became *Homo neanderthalensis* in Europe, *Homo soloensis* in Java, *Homo floresiensis* in Flores and *Homo sapiens* in Africa and in Asia or in Africa only. And he probably became several other *Homo* in several other isolated places, isolated by sea or by ice, that we haven't yet discovered. So, at least, and it is a minimum, four humanities have been coexisting during several hundred thousand years, in some places maybe a million years without, of course, knowing that they were not alone: in Africa and in continental Asia, or in Africa only, as I told you before, in Europe for sure, in Java and in Flores. And, at last, *Homo sapiens* expanded his territory again

around 200 thousand years ago, starting from Africa, or around 50 to 60 thousand years ago if we took the date of the Middle East. He reached America on foot, and peopled the Americas and Greenland without any problem, without any competition; he reached Australia by boat and peopled Australia without any problem as well; and reached Europe, Java and Flores and met there the previous inhabitants. In the three places, after thousands of years, thousands of years of coexistence, without real fights, without any active competition, *Homo sapiens* won. *Homo neanderthalensis* disappeared around 30 thousand years ago, *Homo soloensis*, Java man, disappeared around the same time, *Homo floresiensis* disappeared a bit later, 15 to 20 thousand years ago, maybe because of his more important isolation. And since that time there is only one hominid genus, *Homo*, one hominid species, *Homo sapiens* and one hominid race, *Homo sapiens sapiens*, on the earth, so we can very well become racist, because being racist means being humanist in a way.

Thank you very much.

DISCUSSION ON PROF. COPPENS' PAPER

PROF. ARBER: Thank you very much also for your presentation and for keeping on time. You mentioned that you had collected large numbers of bones.

PROF. COPPENS: Yes, and teeth.

PROF. ARBER: On the other hand you concluded, for example, on an increased complexity of the brain organisation. Is that on the basis of just the size of the brain or is there any other information available? Is there, for example, DNA that can still be analysed nowadays in some of these bones?

PROF. COPPENS: Well, unfortunately not. For the moment, you know, paleogenetics is trying to develop but we are just reaching the Neanderthal step which is not bad, but which is only a one hundred thousand year old hominid. And, as you heard, I was excavating in levels which were between two and three million years and sometimes a bit more than that and, for the moment, we could not find, and there is no technical possibility to find and discover pieces of DNA. We would appreciate that because the phylum, the phyla that we are trying to build and draw with our bones and teeth is only based on anatomy and morphology and it is a bit light.

PROF. ARBER: Other questions?

PROF. COLLINS: Thanks for the very nice summary. Yes, in terms of what DNA may tell us the *Neanderthal* genome sequence is gradually being assembled, there is already enough data to be able to confirm the dating of the separation of the *Neanderthal* line from *Sapiens* as roughly five hundred thousand years ago and, at the present time, no evidence to suggest that there was crossbreeding between *Neanderthal* and *Sapiens*, although that has to be looked at every time we get more data, at the present time there

is no evidence that that is the case. There is also the ability to try to assess on a gene by gene basis, are there particular genes that have undergone more rapid evolution in recent times that might account for some of the unique features of *Homo sapiens*? So, for instance, the gene called FOXP2 which, if mutated in a human pedigree, results in the inability to use language in a very interesting way, with normal intelligence but inhibition of language production and understanding. That FOXP2 gene has undergone dramatically accelerated evolution since we separated from chimpanzees and people certainly point to that as a possible, molecular, very partial but interesting explanation for how language might have come about in recent times.

PROF. COPPENS: Thank you very much. It is *Homo erectus* which became either *Sapiens* or *Neanderthal*. There is no *Sapiens* at the bottom of *Neanderthal*. You know, for a long time we talked about *Homo sapiens neanderthalensis* and now we know enough to know that we could have talked about *Homo erectus neanderthalensis*, but surely not *Homo sapiens neanderthalensis*. *Homo* was only at the level of *erectus* when he divided itself because of isolation into *Homo neanderthalensis* in Europe, which was closed by ice, and *Homo sapiens* somewhere, and again Java Man and Forest Man.

PROF. VICUÑA: Thank you. I think my question was in the same line. The genetic data seems to show that *Neanderthal* and *Sapiens* had a common ancestor about six hundred thousand years ago. But the *Sapiens*' ancestors were in Africa and the *Neanderthal*'s ancestors were in Europe. So how can they have a common ancestor if by that time they were in different geographic regions?

PROF. COPPENS: Yes, that was the sense of my answer. You have *Homo erectus*, which is everywhere in the whole world, *Homo erectus* was in Europe, *Homo erectus* was in Asia and Africa but when Europe was closed by glaciers, by ice, there was a genetic drift in this small population and this was the genetic drift of *Homo erectus* that made *Homo neanderthalensis*, i.e. *Homo neanderthalensis* is really *the* European. We offered *Homo erectus* to Europe and Europe made *Homo neanderthalensis*, if I can say that.

PROF. PHILLIPS: You have told us the story of the human, of the hominid side of the divergence around ten million years ago, so what I am wonder-

ing is, what do we know about the other side, about the side that led to chimps and gorillas, and is there anything there that can give us insights into why the one side developed into people like us and the other side developed into chimps and apes and how different are those chimps and apes from the common ancestors?

PROF. COPPENS: I was hoping not to have this question! We have the impression to know quite a lot on the hominid side and we do not know much on the other side. The explanation can be that, on the other side, on the chimpanzee side, on that side there are forests at the moment and it is difficult to reach the sedimentary level through the forest and its humus, but that is just part of the explanation. There are a few bones which are now considered as possibly being part of an answer to the chimpanzee history, but they are very few for the moment and it is true that, in this beautiful divergence, we have a lot of people here, it is crowded here, and nothing on the other side. There are some bones that are considered to belong to hominids, which after a long debate, could probably come from the other side, but that is something else. As far as the difference is concerned, it seems to be really an ecological reason, that the common ancestor could have lived in a type of mosaic *paysage* with some forests and some open areas and these chimpanzees are clearly adapted to forests, wooded savannah areas, in their body, in their locomotion, for instance, it is very clear. On the other side, the hominid branch, what I am calling the pre-human, are adapted to much more open areas because at the same time they also climbed for a long time. This was a nice discovery, thanks to Lucy, Lucy was the first skeleton that allowed us to say that, well, the two locomotions are in the same body. It is the reason why they were at the same time upright and biped and still climbing for some million years and then, finally, exclusively biped.

PROF. DE DUVE: My question is, what is the present state of your East Side Story?

PROF. COPPENS: Yes, well, I cannot say that this question is good! You know, at a certain time – and some of you may know – at a certain time I thought that, having a common ancestor in central Africa at the time where the forests were developed from the Atlantic Ocean to the Indian Ocean, and having this big tectonic event which is called the Rift Valley – and the Rift Valley was not only a fault but it is especially mountains, mountains on the

Western shore of the rift, mountains going up to three thousand, four thousand, five thousand metres sometimes and this happening around 8 million years ago – there was a change in the rains and, of course, a change in the ecology and the ecology of the West and the East were not the same, the forests remained on the West and the savannah was clearly appearing – it is now very well documented – the savannah was appearing on the East and the East was becoming drier and drier every thousand of years. I thought that, having a common ancestor in these tropical areas at the time of the forest everywhere, the Rift could have been the reason, the separation, between one population and another and the change in ecological niche could have been the reason of the transformation of the ones on one side towards chimpanzee and the ones on the other side going to hominid. But, as you know, since 1995 there have been new discoveries in Chad and Chad is of course outside the Rift Valley, on the other side of the Rift Valley, so the explanation might be, as usual, more complicated than I thought. For the moment, Michel Brunet, who discovered the fossils in Chad, and myself are trying to understand what has been the role of the Rift Valley, because the Rift Valley has, for sure, played a role, but it was a sort of barrier sometimes and a sort of filter at another time. For instance, we are studying every group of mammal according to time and, as far as pigs are concerned, pigs are passing through the Rift, on the West and on the East they are the same but, for instance, anthracothere, which are a group close to hippos, anthracothere are on the West and could not pass the Rift, so the Rift Valley was really a barrier for them. And for hominids we do not know yet, because there are quite old hominids in Chad, *Sinanthropus* is seven million years old, and *Australopith gracile* is 3.5 million years old, so there are hominids between, let us say, seven and three million on the West and, in the East, the record is beautiful, as everybody knows. So, if the East Side Story has played a role, it is not the main role that I thought at a certain time.

F. CAVALLI-SFORZA: What do we know of the demography of these early humans, meaning the numbers that have been found in Africa rather than Asia and Europe, and do you think it is plausible, somehow, that the more successful development of the African *Sapiens* was due to the fact that the numbers were higher there, because it was obviously the place of original man, so more chance for mutation and successful adaptations?

PROF. COPPENS: Well, we do not know much, of course, about demography but, as in places like the Omo River Basin that I was telling you

about, we have enough material, enough stuff, thousands of bones, to have an idea of what could have been the composition of the ecosystem at that time. We know, in comparison with what the ecosystems are today, what are the proportions of herbivores, what are the proportions of carnivores and what could have been the proportion of omnivores and, according to that and again in these sites where the number of fossils is good enough and important enough, we were talking about the possibility of a few thousand inhabitants of *Homo habilis* on the whole Eastern Africa at the time of 2.5 million years ago. The movement could have been like the movement of the Inuit, well, maybe 50 years ago. I had a director at the Musée de l'Homme who worked in Greenland in the 1930s and he told me that a group of Eskimo, when the diet was good enough, was demographically increasing and then it grew too big for the area to get food so a small group went 30 or 40 km beyond and organised a new community and so on, so the movement could have been something like that. As far as *Homo sapiens* are concerned, I am not happy – I must say – about the origin of *Homo sapiens*, it is a bit confusing because having *Homo* in Africa expanding his territory beyond Africa 2 million years ago is OK, and it is quite clear, but having again *Homo sapiens* in Africa doing the same thing in the same direction seems a bit confusing and I just do not understand what this *Homo sapiens* could have done with the previous population of *Homo erectus* becoming a bit *Sapiens* also elsewhere. So 'Out of Africa' once is OK, 'Out of Africa' two is not very good. And the other question about *Homo sapiens* is that there is a confusion also in the literature between *Homo sapiens* and *Homo sapiens sapiens* and it is strange that we have *Homo erectus* becoming *Sapiens* in some places and *Homo sapiens* becoming *Sapiens sapiens* in Africa and leaving Africa again. So I do not know what we are talking about, if we are talking about *Sapiens sapiens* or only *Sapiens* and you know, for instance, in two places – I am working in the field in Mongolia at the moment – I have old *Homo sapiens* which have *Homo erectus* traits again, so it is a sort of transition between *Homo erectus* and a sort of *Homo sapiens*, a sort of *Sapientisation* in situ. And we have the same thing in China, which is the same big province actually, and we have the same thing in North Africa, in Morocco, where we have enough fossils there is really a transition between *Homo erectus*, who is completely *Homo erectus*, and *Homo erectus* who is a bit *Sapiens* and *Homo erectus* who is so *Sapiens* that he is a sort of *Sapiens* with some *Erectus* in it so it seems, really, that, by himself, *Homo erectus sapientisised* where he could and where he was, of course, and every-

where he was. So, again, I do not understand. I know the genetic arguments which are very strong saying that the population today is very homogenous but, again, I do not understand how these *Homo sapiens* or *Homo sapiens sapiens* leaving Africa dealt with, coped with the people who were already everywhere in Eurasia. So I do not know if the argument of saying that the number of people in Africa might be at the origin of the second movement is sound.