Scientific Insights into the Evolution of the Universe and of Life
Pontifical Academy of Sciences, Acta 20, 2009
www.pas.va/content/dam/accademia/pdf/acta20/acta20-cavallisforza.pdf

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.