

Library of Congress Cataloging in Publication Data

Jacquard, Albert.  
In praise of difference.

Translation of: Eloge de la différence.

Bibliography: p.

Includes index.

1. Human genetics. I. Title.

QH431.J27713 1984 573.2'1 84-7002

ISBN 0-231-05482-3 (alk. paper)

## CONTENTS

<i>Preface. The Improvement of the Human Race: An Ancient Objective</i>	vii
1. The Elementary Process: Parent Begets Child	1
2. The Collective Process: The Structure of Populations and the Succession of Generations	17
3. Future Prospects for Our Genetic Heritage: Real and Imaginary Dangers	33
4. Human Races: An Ill-Defined Concept	61
5. Evolution and Adaptation	85
6. Species Improvement: What Improvement?	107
7. Intelligence and Genetic Heritage	133
8. The Temptation to Act	159
<i>Appendix</i>	175
<i>References</i>	179
<i>Index</i>	183

Columbia University Press  
New York Guildford, Surrey  
© 1978 Editions du Seuil  
English translation © 1984 Columbia University Press  
All rights reserved

Printed in the United States of America

Clothbound editions of Columbia University Press Books are  
Smyth-sewn and printed on permanent and durable acid-free paper

## FOUR

### *Human Races: An Ill-Defined Concept*

THE MULTIPLICITY and complexity of humanity inspires the need to divide it into categories and groups, by assigning those individuals who resemble each other most to the same category. For this classification to have biological significance, those traits which allow one to see similarities must, of course, be hereditary and must also be to some extent stable from one generation to the next.

Initially, classifications could only be based on directly observable traits, the colors and shapes of individuals, for instance. Such classifications were often subtle and took account of complex parameters, but they were, of necessity, limited to the "realm of phenotypes." Taxonomists defined various races according to skin colors (black, white, or yellow), hair texture (curly or straight), the relationship between the length and breadth of the skull (dolichocephalic or brachycephalic), etc. Depending on the traits studied, the classes or "races" identified were different, and there were lively polemics between those who detected 4 principal races and 25 secondary ones and those who found 20, or 29, or 40 races.

The discoveries of genetics made it possible to outline the problem more precisely by giving a more objective basis to the concept of race: a race is a group of individuals who have a large part of their genetic heritage in common. This time, the classification is based on intrinsic characteristics of the various human groups, in-

dependent of the conditions under which they live; it concerns the "realm of genotypes." Therefore, conclusions sufficiently clear to win general agreement can reasonably be hoped for.

Unfortunately, scientists failed to discard the old concepts when interpreting the new observations, or, to use a biblical metaphor, they put the new wine into the old casks; in spite of the remarkable progress of knowledge, confusion on this point has merely worsened. Those biologists who have had the courage to go against generally accepted ideas (J. Hiernaux 1969; J. Ruffié 1976; or A. Langaney 1977, for instance, recently in France) have not received sufficient attention; public opinion is still marked by theories that are completely outdated but nonetheless have the authority of ancient myth.

### *Race and Racism*

To begin, it is useful to compare the two terms, race and racism:

- the first is the object of legitimate scientific inquiry, based on objective facts: the aim is to develop methods of classifying individuals with a view to defining human groups, "races," that are relatively homogeneous;
- the second evokes an attitude of mind, which, of necessity, is subjective: the different races are compared and evaluated with a view to establishing a hierarchy.

These are, of course, two distinct activities: it is possible to try to define races without being in the least "racist" in the sense that we have indicated. It is to be noted, however, that this is, in practice, only a theoretical possibility. The need to define races is rarely motivated by the taxonomist's single-minded concern with ordering all his data; it comes from the desire, so highly developed in our society, to differentiate the group to which we ourselves belong from other groups. It corresponds to the Platonic idea of "type." We can define the human species, but it is difficult to out-

line in any kind of detail the ideal human type; several types have to be considered: white, black, Indian, Eskimo, etc.

Without claiming that all responsible taxonomists fell to such excesses, we quote the following extracts from Crozat's *Géographie universelle*, published in 1827, just a century and a half ago, to illustrate to the point of caricature, where this notion of "type" can lead and what confusion it can create:

The Chinese have a wide forehead, a square face, a short nose, large ears and black hair . . . They are naturally gentle and patient but egoistic, proud . . . Negroes are generally well-built and robust, but lazy, deceitful, drunken, gluttonous and slovenly . . .  
The inhabitants of America are agile and swift; most of them are lazy and indolent, some are very cruel . . .

This is quite enough of such nonsense, which, it must be remembered, was written, not by a novelist expressing his personal opinions, but by a geographer concerned with producing scientific work. The point most clearly illustrated by these quotations is that a classification is usually based on a variety of criteria, some of which are objective, others subjective, and that it rarely avoids establishing a hierarchy: races are different, therefore some are "better" than others. The extremes to which some dictators have carried this notion are all too well-known.

They are merely translating the ideas of certain scientists into politics and action. Since Darwin, our perception of the progressive transformation of living things, plants, animals, or humans, is based on such concepts as the struggle for existence, the survival of the fittest, the elimination of the unfit, and the propagation, over many generations, of favorable traits. These concepts, which were initially developed with individuals in mind, have been extended almost unchanged, to groups of individuals, or races. Differences between races have come to be seen as the products of varying degrees of evolutionary development and to be thus perceived as inequalities. Even in casual conversation, the average person tends to make remarks which imply that racial inequalities are obvious: some races are superior (usually that of the speaker), others are inferior.

Of course, most French people claim sincerely that they are not racist; South Africans, North Americans, Germans, or Russians are terribly racist, but not us. It is just that we rightly feel superior to Arabs, blacks, gypsies, or Indians, not to mention various other groups who are less well-endowed than us by nature and "who are not like us," if you see what I mean. Let's face it, racism, that is the feeling of belonging to a biologically superior human group, is almost universal.

Examples of this unconscious racism abound. The one that surprised us most was undoubtedly this unexpected sentence taken from the chapter on the duties of a colonel in the *Règlement du service dans l'armée*: "The colonel . . . indicates the most appropriate means of developing patriotism: fortifying love of country and a sense of racial superiority . . ." This regulation was not directed at the German army in the Nazi period but at the French army, and the document was printed in 1957. That a sentence like this was approved by several government ministers and army chiefs-of-staff proves how natural it seems to most people to define a "French race" and to glorify it relative to other races.

A scientist who discovers that the data available to him do indeed confirm racial inequalities and therefore constitute a potential basis for a racial hierarchy should not hide this conclusion. Respect for truth is the ethic of science. In the opposite event, however, he must not hesitate to proclaim this truth and, when necessary, to combat entrenched ideas even when they are virtually unanimously accepted. It is important at this point to sum up: what does science, especially genetics, have to say about the concept of race?

### *What Does Classification Involve?*

Defining races involves classifying the teeming masses formed by the few billion people alive at present and their several billion ancestors.

Similarly, the definition of species implies that all living organisms must be divided into groups. However, in the latter case, we

have a rather specific criterion for deciding whether two organisms belong to the same species or not: their capacity (real or potential) for interbreeding. Some borderline cases are, of course, difficult to decide, but the notion of belonging to the human species is, on the whole, clear and comprehensible; all individuals belonging to this species, no matter how far apart in space and type, whether Australian aboriginals, Northern Eskimos, inhabitants of Tierra del Fuego, or Melanesians, are potentially interfertile.

There is, however, no such criterion for deciding whether two individuals belong to the same race or not. Nevertheless, we regularly make this kind of decision without hesitation; we need no special knowledge to decide almost unerringly whether a particular man passing by us on the street is Chinese or Arab or Indian. We are now going to examine the mental process that leads us to make this kind of classification.

We are aware of the vast number of different individuals belonging to our species; these individuals are much too numerous for our intellect to be able to compare them easily with each other; we therefore replace them with a set of categories, much smaller in number, in such a way that each individual belongs to one category and one only, and that all individuals within the same category are "similar."

Therefore, we are first of all obliged to specify what we mean by "similarity"; more precisely, we must indicate what criteria are being taken into consideration; if there is just one criterion, skin color on the inside of the arm, for instance, it is easy to measure the similarities. If, however, there are more than one, for example, this color and head size, we have to arbitrarily define a global measure that includes both these parameters simultaneously. A technique for defining this measure has been developed by mathematicians; it involves calculating a "distance": the smaller the distance between two individuals, the more similar they are. There are several formulae for making this calculation: depending on whether we are using "Euclidian distance" or "Manhattan distance" or "chi squared distance," a single set of data will yield several different sets of distances. The choice is practically infinite, so rich is the imagination of mathematicians.

Suppose that, having chosen certain criteria for a classification and a formula for calculating distances, we managed to determine all the distances  $d_{ij}$  between each individual  $i$  and each of the others  $j$  (for the 4 billion or so people alive at present, the number of distances will be in the order of 8 billion billions). The categories that we are trying to establish will be meaningful only if the distances between individuals within the same category are, at least on average, significantly smaller than between individuals in different categories. Here again a large number of methods for achieving this have been developed and each one leads to a different result.

The simplest method, which no doubt is also the closest to the intuitive, natural approach, involves the construction of a "tree": the two closest, or most similar, elements are first of all put together to constitute a class made up of just these two elements, then the closest classes are put together; in this way, the number of classes is gradually reduced until all that remains is one class that includes all the elements.

To illustrate this process, consider a very simple example: a species consisting of only 10 individuals  $a, b, \dots, j$ . Let us suppose that we have chosen the classification criteria; we have obtained the measure of each criterion for each individual; finally, we chose a "distance" formula that enabled us to calculate the 45 numbers  $d(a, b), d(a, c), \dots, d(i, j)$ , or distances, characterizing the degree of dissimilarity between individuals. We notice that the shortest distance is  $d(a, f)$ ; we therefore group  $a$  and  $f$  together in a "class" called  $K_1$ , which we now consider to be a fictitious "individual." Then, we calculate the 36 distances between  $K_1, b, c, d, e, g, h, i, j$ ; this time, we notice that  $d(K_1, c)$  is the shortest of these distances so we put  $c$  with the class  $K_1(a, f)$  to constitute class  $K_2$  and so on. To do this, we had to decide on a method for calculating the distance between a class, such as  $K_1$ , and an element, and for calculating the distance between two classes, which can be done in many different ways.

Eventually, having made all these arbitrary decisions, we get a tree similar to that illustrated in figure 5. How can it be used to define races? One further choice remains to be made, that of the

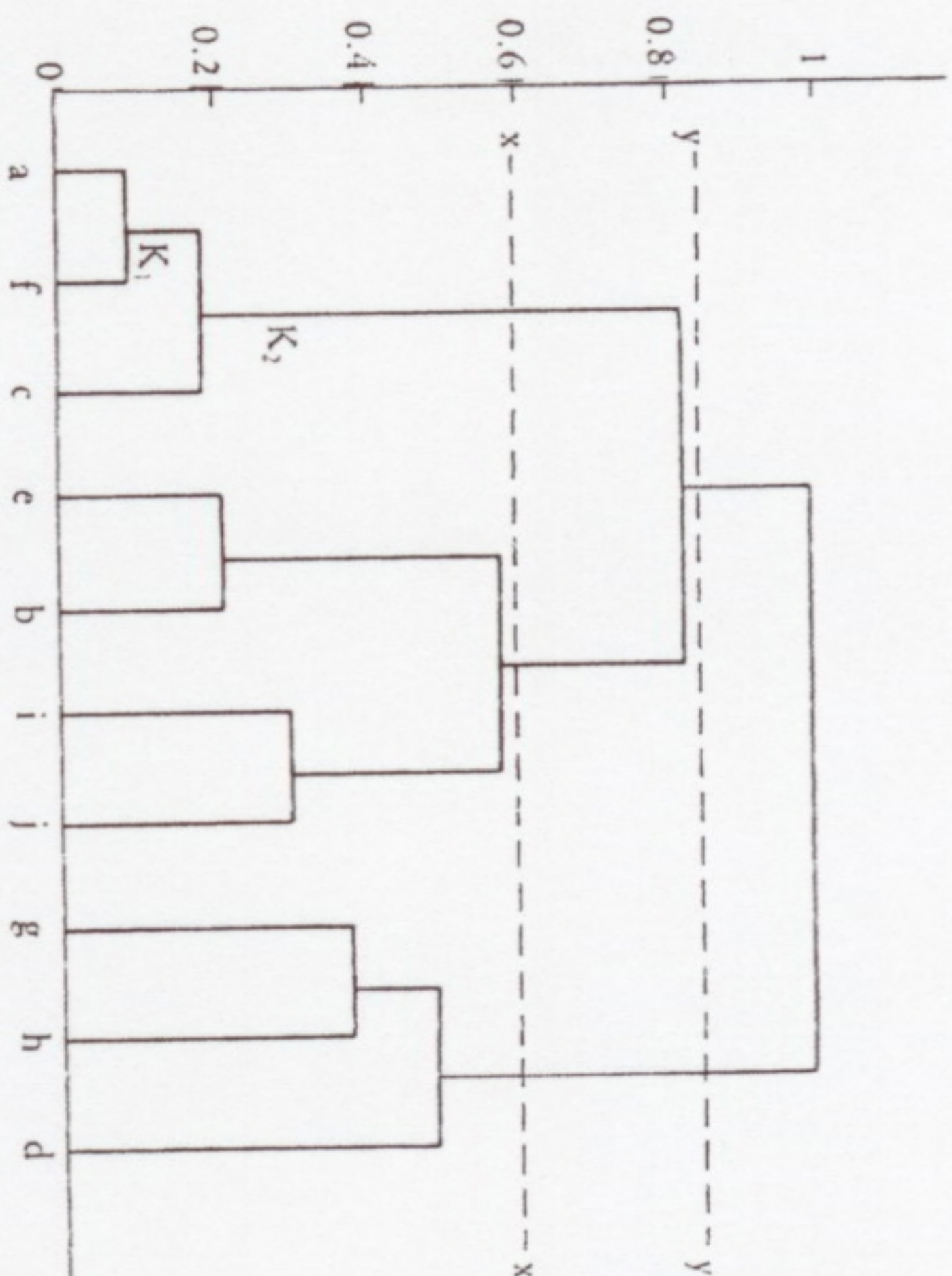


Figure 5

number of races to be defined, a number which, of necessity, is between 1 (race and species are thus one and the same thing) and 10 (as many races as individuals, which makes our efforts meaningless). If we wish to identify 3 races, we will cut the tree at the level of the  $xx'$  line, which gives us the "races" ( $a f c$ ), ( $e b i j$ ), and ( $g h d$ ); but if we prefer just two races, we must draw the line  $yy'$  which isolates the "race" ( $a f c e b i j$ ) from the "race" ( $g h d$ ); etc.

It is to be noted, however, that the level at which we drew the horizontal lines to group individuals or classes has a specific meaning: it represents the loss of information that occurs when the initial data concerning individuals are replaced by global data concerning classes. We see that, in our example, grouping in 3 races involves a loss of only 60 percent of this information (the line  $xx'$  which cuts the tree where it has three branches is indeed at the height 0.6), while grouping in two races involves a loss of more than 80 percent. In order to lose no information, one has to stay

at the zero level, which amounts to making no grouping; on the contrary, in order to group all ten individuals into a single category, all the information must be lost. Using the diagram, we can thus make an informed decision regarding the choice of the level at which to cut our tree, and therefore of the appropriate number of races.

This rapid explanation will seem very superficial to researchers accustomed to this kind of work. Over the past twenty years, techniques for analyzing data, for extracting a thread of meaning from a jumble of overabundant facts, have been developed and refined to a remarkable degree. However, our intention was not to give a comprehensive lecture but to show that the business of classifying, which may seem so simple and natural, is in reality very complex, and that the outcome depends on very arbitrary choices. It is not that the result of a classification is of no value, it is just that one needs to be aware of its relativity.

### "Phylogenetic Trees"

The mechanism involved in classifying, which we have just analyzed in the form of the construction of a tree, can be applied to any collection of objects whatsoever, whether it be the stock of a hardware store, the various languages spoken on our planet, the animals in a forest, or the individuals in our species. However, in the latter case, the aim is not only the grouping of similar individuals into relatively homogeneous classes, it is to find a set of historical facts: their genealogies, the complete pattern of their history, generation after generation.

Two individuals with common ancestors have received identical genes from those ancestors. This genotypic similarity will be reflected to some extent at the phenotypic level. When we make groupings based on comparisons between phenotypes, it is reasonable to hope that the closer two individuals are, the more common ancestors they have; by making a tree diagram, we will obtain a rough outline of their patterns of relationship and descent; we will draw what is called a "phylogenetic tree." Consider a population

that has undergone successive divisions during a process of fission similar to that outlined in figure 6. Each group, after a certain period of autonomy, divides into two populations which remain completely and permanently separate and which, in turn, later undergo similar division.

An individual belonging to group A in our diagram has more genes in common with an individual from group B than with an individual from C or G, because in the case of A and B, one finds common ancestors after going back a fewer number of generations. A study of resemblances between present-day populations can therefore be attempted with a view to reconstituting the history of their paths of descent. It is easy to see why this information is of such interest to historians and ethnologists who are eager to learn the origins of the people that they study.

This work has been accomplished with remarkable precision for the ensemble of species that constitute the living world, with each species being considered as a single homogeneous group. We now have trees displaying the whale as well as the fly, the human being as well as the trout, and also their distant common ancestors. Figure 7 shows a plausible tree for the appearance of various animal species, established by American geneticist D. Hartl (1977) based on structural differences between the various proteins common to all species. This tree is very similar to those that had been established by taxonomists based on anatomical comparisons. This re-



Figure 6

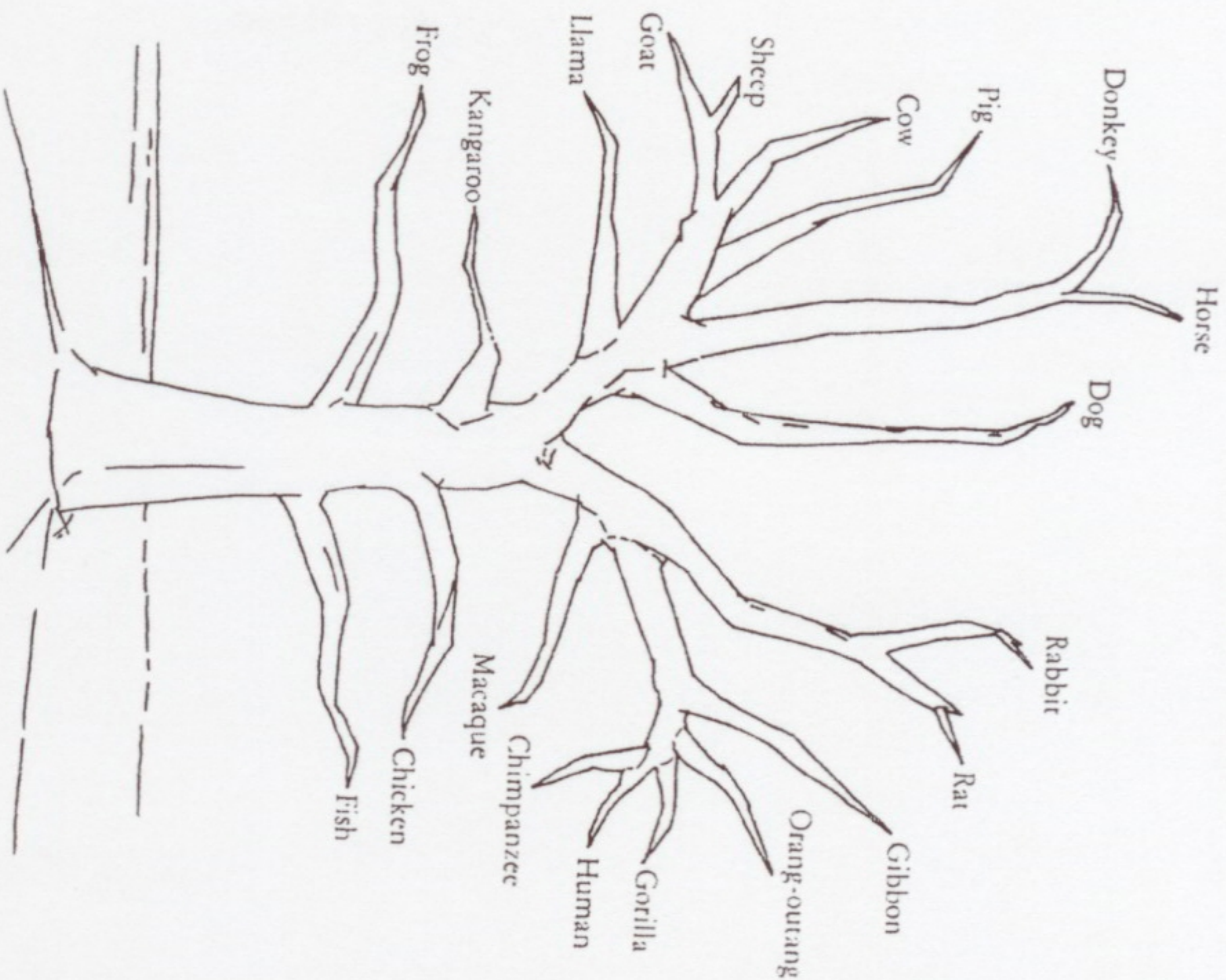


Figure 7

construction is facilitated by the fact that the various species satisfy the condition that we hypothesized for the drawing of a phylogenetic tree: populations are subject to divisions, but not to fusions; once separated, they remain so permanently. When, after chromosomal rearrangements or accumulated mutations, a new species appears, cross-fertilization with the original species is impossible (or, as in the case of the mule, the progeny are sterile, but

this sterility has the same consequences), and the genetic separation is complete.

On the contrary, when a population is split into two groups which remain interfertile and therefore still belong to the same species, though they evolve separately and gradually become two distinct "races," genetic exchanges remain possible between them, either due to migration or to the complete fusion of the two temporarily separate groups. This type of evolution cannot be represented by a tree of the kind shown in figure 6. It is, rather, a complex network, such as that represented in figure 8. This network cannot in any way be compared to a classification tree. Even the highly sophisticated mathematical techniques that made it possible to construct these trees are completely incapable of reconstructing networks that have been complicated by fusion between groups.

We will see that, in spite of this fundamental impossibility, many researchers have tried to use data gathered from populations alive today as a key to possible historical links between these populations. These attempts are not useless so long as their limitations are borne in mind. Any results derived from them must be considered merely as a confirmation of other data obtained from independent sources. In any case, the aim is usually not to study the phylogeny of humanity as a whole, but simply to analyze the relationships between the various groups living in a limited geographic area.

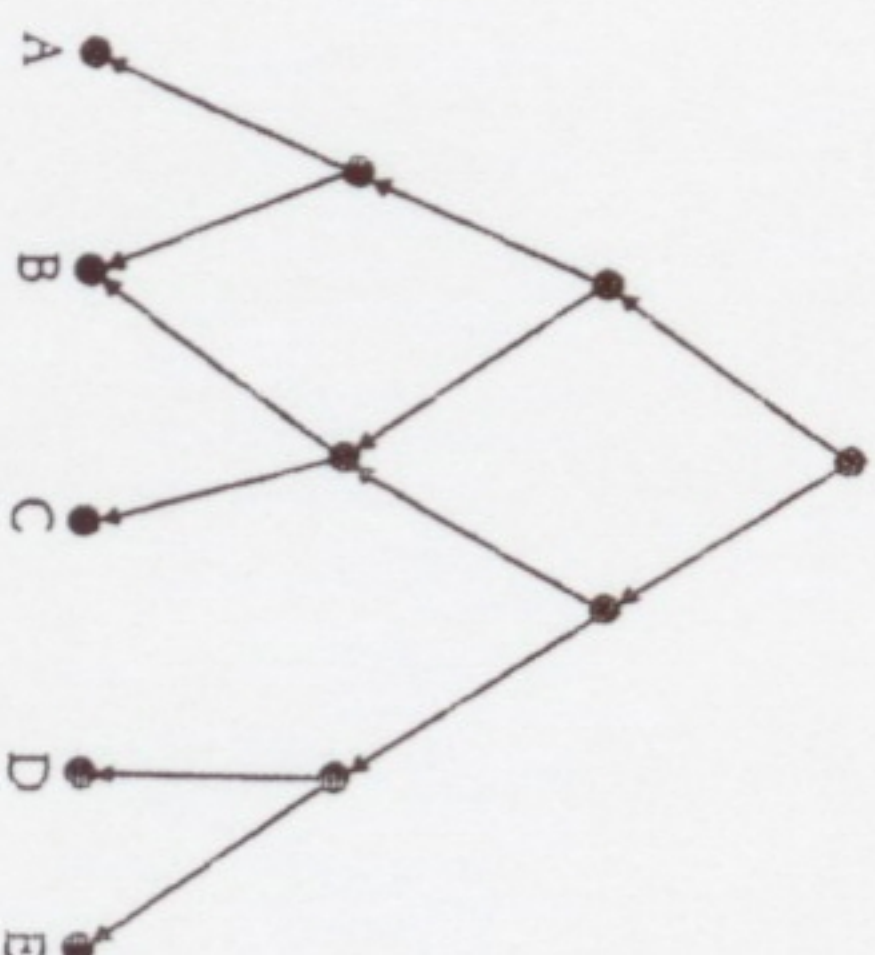


Figure 8

*The Question of Skin Color*

In defining races, the trait that is spontaneously brought to mind is that which is most easily observed: skin color. This trait is obviously hereditary, but the underlying genetic determinism is not fully understood.

The first point to be noted is that, contrary to widely-held opinion, the pigment melanin, which is responsible for skin color, is present in the epidermis of white-, yellow- and blackskinned people but in very different doses. The differences are therefore quantitative and not qualitative. There is considerable variation within each group and the difference between two people within a population can be much greater than the difference between the averages of two groups belonging to distinct "races." In a recent study entitled "La quadrature des races," André Langaney (1977) shows that it is possible to go without discontinuity from those people with the palest skin (Northern Europeans) to the darkest (the Saras of Tchad) with intermediate types being chosen from only two other populations (North Africans and Bochimian Bushmen of the Kalahari desert).

Studies on crosses between whites and blacks and between their descendants have shown that this trait follows a typically Mendelian pattern. It behaves as though it were governed by 4 pairs of genes with complementary effects; the actual mechanism is not doubt far more complex, but this simple model accounts very well for what is observed. Assume then that "whites" have eight *w* genes for white skin and "blacks" eight *b* genes for black skin. All the intermediate shades are possible, depending on the distribution of *w*s and *b*s within those eight genes.

Studies done on American blacks confirm this genetic model. This very heterogeneous group consists of all North Americans who have among their ancestors Africans deported as slaves between the beginning of the seventeenth century and the middle of the nineteenth century; they also actually have a considerable number of Europeans among their ancestors, since young female slaves frequently bore children that were fathered by their white masters. In fact, comparisons between the frequencies of certain genes in

African populations in the Bénin Gulf region (which were the chief source of slaves), in the Anglo-Saxon populations of Europe and in American blacks, have led to the conclusion that 25 percent of the genes in the latter group are "white."

For instance, in the case of the Rhesus blood system, the frequency of a certain gene called  $R_0$  is 63 percent for the Africans while it is only 3 percent for the Europeans. For American blacks, an intermediate frequency of 45 percent is found, which is consistent with the hypothesis that a quarter of their genes are European. This is, of course, only a general estimate; the actual proportions are no doubt very different according to regions and families. (The genetic input of whites is much greater among American blacks in the North and West than in the South.)

According to this hypothesis, each of the genes controlling skin color in an American black has 1 chance in 4 of being a *w* gene. The probability that all the 8 genes involved will be *w* is therefore equal to  $(1/4)^8$ , or about 1/65,000. In other words, of the 20 million or so American blacks, several hundred have only genes for white skin and are actually white. Similarly, the probability of having 8 *b* genes is  $(3/4)^8$ , or about 1/10: there are, therefore, only 2 million "American blacks" whose genes for skin color have come exclusively from the founder African population and who are as black as their ancestors. This calculation can be continued and the distribution of "blacks" according to the number of *b* genes (between 0 and 8) which they carry can be estimated. This distribution turns out to be very similar to the actual distribution of skin colors in this population, which proves that the "model with four pairs of genes" is a good representation of reality.

Populations with very dark skin are found mostly in Melanesia (i.e., the group of islands situated in the Southwest Pacific), in the Indian peninsula, and in Africa south of the Sahara, regions that are all close to the equator. This fact is used in support of the theory that black skin has a greater adaptive value in hot countries. Later, we will see that even this idea, which is so widely accepted, can be disputed. For the present, we merely point out that these three groups of populations do not, in any sense, constitute a "race"; apart from skin color, they are different in every respect: analysis

of their blood systems shows, for example, that they cannot be considered to be offshoots from one initial group; their "phylogenetic tree" cannot be represented as a single trunk with three branches. If this were so, traits other than black skin would have been transmitted from the ancestor population to all three groups, but this does not seem to be the case. This finding shows that no classification based on skin color alone can have biological significance; it is very inconvenient for those who imagine that a definition of races can be based exclusively on this criterion (which, of course, anthropologists no longer do); but there is no getting away from it.

Lastly, we find that even though skin color is the most obvious trait, and the easiest to compare, it involves only a tiny fraction of our genetic heritage (no doubt 8 or 10 genes out of several ten thousand); no other major biological trait seems to be linked to it; it cannot therefore be used as a basis for classifying populations in a meaningful way: how many individual and collective tragedies could have been avoided, and could still be avoided, if this simple fact had been, or were at last, accepted by everyone.

Are there not other physical traits, relatively easy to measure, which could be substituted for skin color as a basis for classification?

Height, head length and breadth, the relationship between the last two (the cephalic index being the basis for differentiating between "brachycephalics" and "dolichocephalics") and many other body measurements can be used to determine similarity and dissimilarity between individuals or between groups. However, the genetic determinism of those traits is far from clear, and, in the case of most of them, it is completely unknown. Given our present state of knowledge, it is impossible, and probably will be for a long time to come, to infer genotypes from information gathered on phenotypes.

Moreover, some of these traits, height for instance, are not at all stable, in spite of being genetically determined. In all industrialized countries since the beginning of this century, there has been an extraordinarily rapid increase in people's height. According to G. Olivier's (1977) recent study, the height of French conscripts at the age of 20 was:

HUMAN RACES: AN ILL-DEFINED CONCEPT				75
165.4 cm in 1880	165.8 cm in 1900	165.7 cm in 1920		
168.5 cm in 1940	170.0 cm in 1960	172.3 cm in 1974		

These data indicate that the increase is occurring more and more rapidly. It cannot be caused by genetic changes; only environmental influences (but which ones? one can do no more than guess) could have produced these changes in such a short time. The fact that height varies to this extent means that all hope of using it to compare different populations or to reconstruct their "phylogenetic tree" must be abandoned.

In confining itself to quantitative traits, the genetic significance of which, as we shall see in chapter 6, is difficult to interpret, anthropology risked becoming stuck in an impasse; advances in biochemistry opportunely supplied it with facts that made a new stage possible: these facts pertain to traits, essentially the blood systems, with a genetic determinism that is so clear-cut as to make the inference of genotype from phenotype much easier.

### *Blood and Its "Systems"*

The first blood-group "system" was discovered in 1900 (the same year that, by pure coincidence, the rediscovery of Mendel's laws heralded the development of modern genetics). The Austrian biologist Karl Landsteiner noticed that certain people's blood is apt to agglutinate that of certain others, which explains why accidents sometimes occur during blood transfusions. He was thus able to identify four "groups": A, B, AB, and O. Analysis of the transmission of this trait in families showed that, in each individual, it is governed by one pair of genes; for each of these genes, there are three possibilities: A, B or O; moreover, the O gene is recessive to the A or B genes. Therefore, the correspondence between the pair of genes carried (the genotype) and the trait manifested (the phenotype) is:

Genotype	Phenotype
AA and AO	A
BB and BO	B
AB	AB
OO	O

Twenty-seven years later, Landsteiner discovered a second system, called "MN," and, in 1940, a third, the well-known Rhesus system. Since the last war, discoveries have continued at an ever increasing rate. The study of the structure of hemoglobins, as well as of the properties of red blood cells, of white blood cells, and of serum, has led to the identification of more than seventy systems and the list is increasing every year.

The big advantage of these traits is that they give us information on genotypes. Even though this information concerns only a tiny number of genes and is therefore very limited, it allows us to compare populations, based on some objective measures that are independent of the effects of environment. A person who is born with an *A* and a *B* gene belongs to the group AB, regardless of whether he is young or old, starving or well-nourished, living in a tropical forest or in the Canadian North. To classify populations, all that is required is that sufficient data be accumulated, based on blood samples from the various human populations. This work has been undertaken by numerous teams who have left scarcely any "unknown territories" on the maps which chart their findings; Professor A. E. Mourant's (1976) atlas is proof of this. However, the samples from many regions are far from being representative and the results are imprecise. The task is, therefore, not completed and much work still remains to be done.

Nonetheless, some lessons can be learned from the data currently available, in spite of their incompleteness and imperfection. The first concerns the rareness of "marker" genes: a "marker" gene *g* is one which is found in a population *P* but in no other. It is, therefore, a specific trait which differentiates *P*: an individual with this *g* gene can belong only to population *P*. Note, however, that the converse is not true: *all* individuals in population *P* do not have the *g* gene and it can even be relatively rare in this population. Whether it was introduced from outside by a migrant or is the product of a mutation within, it is not necessarily very widespread. In spite of extensive research, relatively few "markers" have been found. The clearest example is that of a certain gene *a* for the system "Diégo" which was discovered in Venezuela in 1954. This gene, which reaches a frequency of 40 percent among certain

South American Indian tribes, is completely absent in Central Africa, as it is among Polynesians, Papuans, and Australian aborigines; in Europe, it is found only in rare cases; on the other hand, it is quite widely represented among most of the populations of Far East Asia.

Similarly, some genes associated with the *Gm* system, which we will be discussing in detail later, can be considered specific to Central Africa, for instance, *GmG* and *GmH*. Though it is not completely impossible to find one of these genes elsewhere, it is extremely rare.

In the case of all other genes, regardless of the system involved, no specificity has been found. What distinguishes two populations is not the fact of having or not having a particular gene, but the fact that the frequencies of this gene are different. The criterion is not one of "all or nothing" but, rather, "more or less."

To compare populations, we must therefore synthesize into a single criterion their greater or lesser similarity, which is determined by the degree of similarity between the frequencies of various genes within them. Consider an imaginary example, that of 4 populations in which the frequencies of 4 genes,  $a_1$ ,  $a_2$ ,  $a_3$ ,  $a_4$ , for a certain blood system are known. The frequencies, expressed as percentages, of each of these genes are indicated in table 1:

Table 1

Gene → Popula- tion	$a_1$	$a_2$	$a_3$	$a_4$
I	2	3	75	20
II	1	49	20	30
III	40	30	3	27
IV	27	30	40	3

The problem is how to decide which populations are most similar and which are most different. The difficulty becomes obvious when we try to compare the data: I resembles II for genes  $a_1$  and  $a_4$  but differs from it for  $a_2$  and  $a_3$ ; III resembles IV for  $a_1$  and  $a_2$  but differs for  $a_3$  and  $a_4$  and so on. To reach a decision, we need

to calculate a distance, that is, a number that increases as the populations become on average more dissimilar. We saw that there are many different formulae available and that they can produce widely different results; many population geneticists, especially British and Americans, use a distance called the "arc cosine" in such cases. Without going into technical details which are of little interest here, we can say that, for our example, the result is as follows:

$$\begin{aligned} d(I-II) &= 1 & d(I-III) &= 1.71 & d(I-IV) &= 0.86 \\ d(II-III) &= 1.05 & d(II-IV) &= 0.98 & d(III-IV) &= 0.76 \end{aligned}$$

the distance between I and II having been arbitrarily assigned a value of one. It is useful to graph these results in such a way that the populations are displayed as points separated by distances that are as similar as possible to the distances between populations. Here, we get 4 points like those on figure 9. The data available on many blood groups in a large number of populations thus allow us to calculate a set of distances and to draw genetic maps which sometimes show a surprising divergence between genetic distances and geographical distances.

Ph. Lefevre-Witier's work (1974) on the populations of North and West Africa is a good example of this kind of study: having isolated 26 "populations," he compared the frequencies of the various genes associated with 5 blood systems. The 325 pairwise dis-

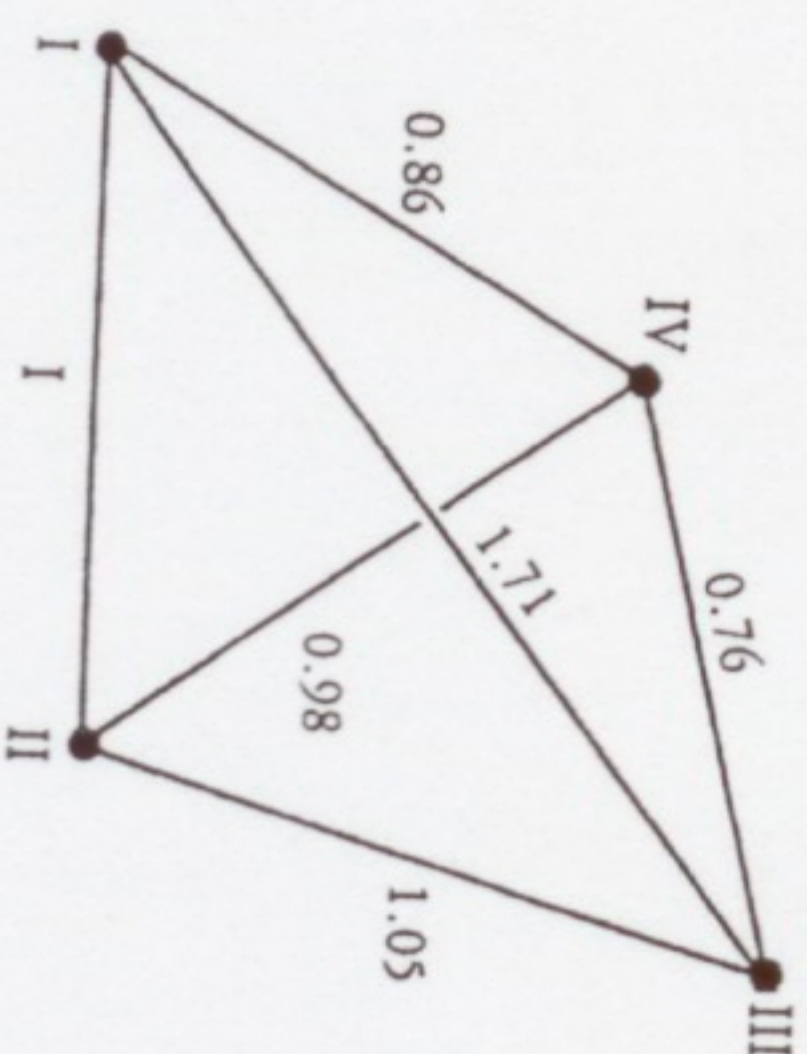


Figure 9

tances between these populations were used as a basis for the map featured in figure 10. From this map, we can see that the Kel Kummer Tuareg of Mali, the R'Gueibat of Algeria, and the control group of French from the eastern Pyrenees are very similar; at the other extreme of the distribution, there is a group consisting of the Gagou of Ivory Coast and the "Ikian," descendants of slaves originally from the Bénin Gulf and still living in the Tuareg tribes;

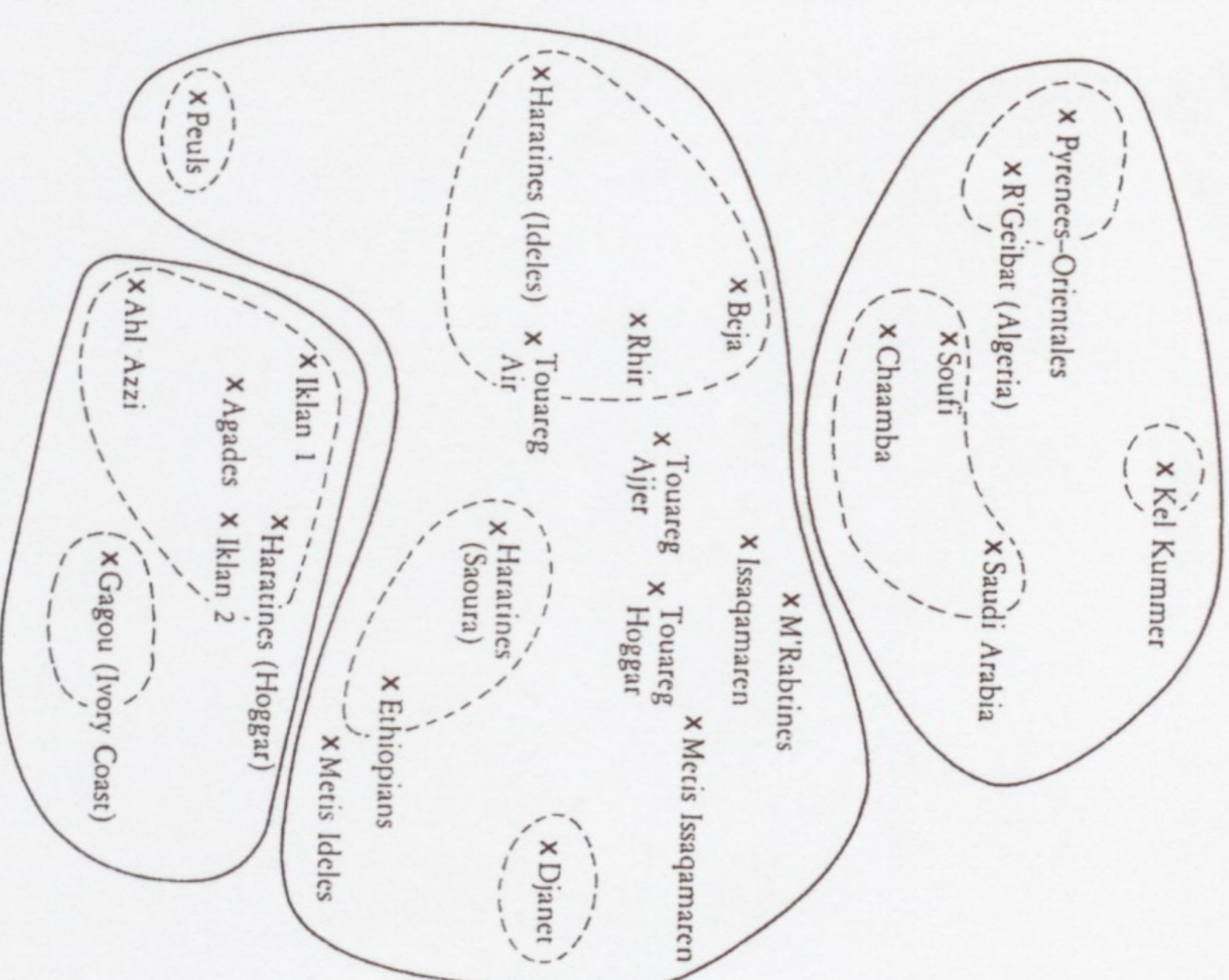


Figure 10

between these two extreme groups, there are various Saharan tribes, the Peuls, and an Ethiopian group.

This kind of map is enlightening and very useful for research, but it is easy to overestimate its significance. First of all, the samples used to represent the populations are often very small; in the case of small groups such as the Kel Kummer Tuareg or the Iklan, the sampling is sufficient to insure a good representation of the entire group, but certainly not in the case of populations that are as widespread and diverse as the Peuls or the Ethiopians. The points on the graph do not really represent these populations but rather particular samples of them; other samples might be situated on a completely different zone of our map.

Similarly, the graph might have been very different if data on other blood groups had been used. We objected that skin color is determined by only a few genes and is therefore not representative. We must avoid falling into the same trap with hematological data. The possible differences according to the blood groups considered are well illustrated by a comparison between the three best documented systems: Rhesus, Gm, and HL-A.

#### *Genetic Systems with a High Degree of Polymorphism: Rhesus, Gm, HL-A*

Collection of data on some systems, though extensive and worldwide, has led to the discovery of only a small number of different genes: only 3 genes have been found for the Duffy system which was discovered in 1950. For others, on the contrary, numerous genes were identified quite rapidly and new ones are constantly being discovered. These multi-gene systems are said to be "polymorphic."

The Rhesus system belongs to this category: the well-known "positive" and "negative" traits were quickly seen to be just one aspect of a whole so complex that the debate about the genetic mechanism involved is still not closed. Note that, up to the present, more than 20 different genes have been listed.

The Gm system, discovered in 1956, is under study in many

laboratories, for instance that of Claude Ropartz in Rouen (1971). This system is not a characteristic of the red blood cells but of the proteins in the serum, the immunoglobulins which recognize "foreign" substances and neutralize them. Some of these proteins, the IgGs, have variable structures: the study of their transmission within families led to the identification of 12 different genes (from A to L); this list is, of course, provisional.

The HL-A system has been studied very intensively because of its involvement in the rejection of tissue transplants. Since 1958, numerous teams, in particular that of Jean Dausset at Saint-Louis Hospital (1973), have successfully coordinated international studies which have led to the discovery of the underlying genetic mechanism: it is now generally accepted that 4 pairs of genes are involved. These genes are located very close together on chromosome number 6. For the first site, 20 genes have been identified, for the second 30, for the third 6, for the fourth 11; these numbers increase every year.

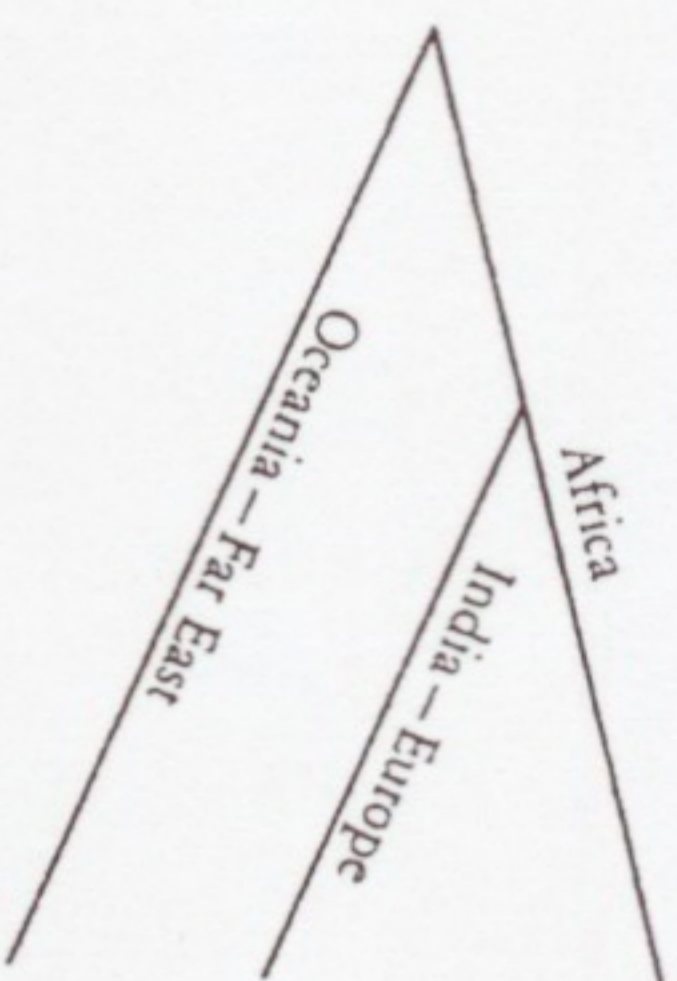
The remarkable richness of these systems makes them useful tools when it comes to comparing populations. However, this very richness makes their use more complicated: a very small sample suffices to tell us the frequency of the genes for the Duffy system in a particular population; but a huge sample is required if one wants to estimate the frequencies of the various Gm or HL-A genes, some of which may be extremely rare. In order to be able, without too great a risk of error, to assume that a particular gene is absent, and such assumptions are often crucial, samples from a large fraction of the population must be studied.

Without going into details, it should be noted, in agreement with A. Langaney (1977), that the information made available to us by these three systems on the colonization of the earth by human populations, and on their subsequent differentiation, is largely contradictory:

— the *r* gene for the Rhesus system is very rare in Oceania and the Far East, common in Africa, in India, in the Middle East, and above all in Europe (its frequency is at more than 50 percent among the Basques and the Bedouin of Sinai);

— the  $R_0$  gene, which seems to be due to a recombination that occurred somewhat later in human evolution, is at a high frequency in black Africa only.

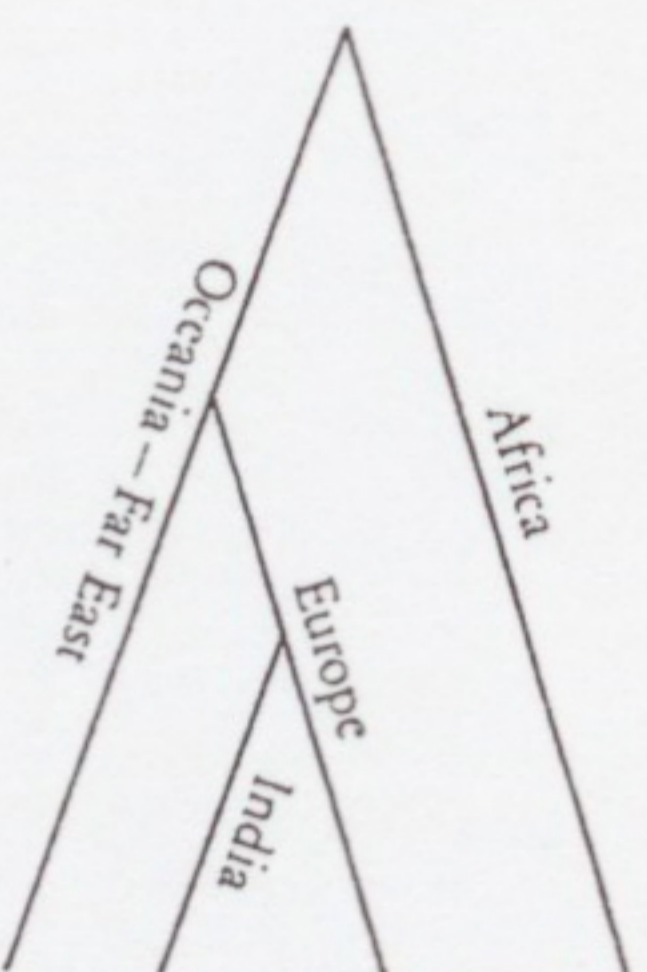
Based on this single system, the tree for the three big groups would take the shape shown in diagram 1.



If the Gm system is used as a basis for the classification the result is quite different:

- the  $A$  gene, the most frequent in Europe, is widespread in Asia and all of the Pacific; it is unknown in black Africa;
- the  $G$  and  $H$  genes which are common in black Africa are practically absent from the rest of the world;
- the frequencies of the various genes are very different in the Far East and in the India-Iran zone;

The tree based on the Gm system looks something like what is shown in diagram 2.



The HL-A system involves so many genes that the data that have been collected are difficult to interpret, since each population is highly polymorphic; there is no population with a simple genetic profile; the best that can be done is to note that the absence of certain genes makes Europeans and Africans seem relatively close, as against Far Eastern people; this gives a tree that is closer to that based on the Rhesus system than to that based on the Gm system.

### *Individual Diversity. Population Diversity*

At this point, the reader must surely have the impression that the accumulation of increasingly precise data and its treatment by increasingly complex processes have only made it more difficult to classify populations. The clear-cut division presented by the geography textbooks of our childhood into whites, yellows, and blacks is no longer acceptable. Has scientific research gone wrong?

The role of science is not to invariably give clear answers to all questions. Some questions are better not answered; in replying, even partially or imprecisely, to an absurd question, one becomes party to mystification and to an abuse of trust.

If the classification of humans into more or less homogeneous groups, of the kind referred to as "races," had real biological meaning, the role of biology would be to establish this classification in the best possible way; but this classification has no meaning. For it to do so, the history of humanity should have unfolded in the manner illustrated by the tree in figure 6: a series of successive splits. In reality, our present human populations were never separate for long enough to allow a significant genetic differentiation to take place. Individuals moved from one group to another and we saw that even a tiny trickle of migration can have major consequences.

Populations can be compared with respect to specific traits and the differences observed can then be analyzed; in some regions, the degree to which relatively separate populations have differentiated from each other can be studied; but this work does not lead

to an objective classification of the human species into "races." The best proof of the uselessness of attempts at defining races was no doubt provided by American researchers R. Lewontin (1974) and M. Nei (1975). They tried to analyze the total global diversity of all humanity into: (1) differences between the major commonly defined groups (whites, yellows, blacks), (2) differences between nations within one of these groups, and, finally, (3) differences between individuals within the same nation: the proportions are respectively 7, 8, and 85 percent. In other words, the total diversity found between humans is reduced by an average of only 15 percent when, instead of humanity as a whole, one considers only those people who belong to the same nation.

This result is worth thinking about: it is not between human groups, but between individuals that the level of diversity is at its highest. Of course, my friend Lampa, a bedik peasant from eastern Senegal, is very black and I am almost white, but some of his blood systems are perhaps closer to mine than are those of my next-door neighbor Mr. Dupont. Depending on which criterion I choose as a basis for comparison, the distance between Lampa and me will be greater or smaller than the distance between Mr. Dupont and me. The result obtained by Lewontin and Nei indicates that the biological distance which separates me from Mr. Dupont is, on average, only a fifth smaller than the distances which separate me from Lampa, or from one of my Japanese or Indian colleagues, or from an Australian desert hunter-gatherer. Does this small difference deserve all the attention we have been giving it for centuries?

## FIVE

# Evolution and Adaptation

THE UNITY of the living world and the interrelatedness of all species during the slow evolutionary process have come to be generally accepted as facts by all but a few fundamentalist groups who object, not on scientific, but on religious grounds. Where does this human species of ours come from? Before trying to reconstruct its past, consider the kind of universe in which it finds itself.

## *The Universe. Man*

According to astronomers, the universe which is accessible to our observation (and our tools of observation are now infinitely more powerful than our senses) occupies a vast space with a radius greater than ten billion light-years.<sup>1</sup> Within this space, matter is not distributed uniformly; it is aggregated into "lumps," the galaxies, of which there are many billions.

Our own galaxy is none other than the Milky Way; it is a major feature on our sky, not because it is larger than the others but because we see it from the inside, whereas the closest outside galaxy is 200,000 light-years away. The Milky Way, shaped like an im-

1. One light-year is equal to the distance covered by light in one year, that is, about ten thousand billion kilometers.