Cahiers de l'Urmis

20 | 2021
Race et biologie

Genomics and human diversity

Bertrand Jordan

Electronic version

URL: <https://journals.openedition.org/urmis/2397>

DOI: 10.4000/urmis.2397

ISSN: 1773-021X

Publisher

Urmis

Electronic reference

Bertrand Jordan, "Genomics and human diversity", *Cahiers de l'Urmis* [Online], 20 | June 2021, Online since 29 June 2021, connection on 01 July 2021. URL: <http://journals.openedition.org/urmis/2397> ; DOI: <https://doi.org/10.4000/urmis.2397>

This text was automatically generated on 1 July 2021.

Les contenus des *Cahiers de l'Urmis* sont disponibles selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

Genomics and human diversity

Bertrand Jordan

- ¹ Since the first complete reading of human DNA in 2003, after a decade-long effort involving hundreds of laboratories, DNA analysis techniques have progressed immensely and have since been applied on an increasingly large scale, for both basic research and its many clinical applications. Thus, genomics (the study of genomes) has examined the genetic heritage of many species, including our own. This has opened the way to reconsidering, more precisely and objectively, many anthropological issues, including those relating to the very existence and characteristics of human population groups. This article aims to provide an overview of the latest results of this research and their implications.

1. Human diversity at the DNA level

- ² As is widely acknowledged today, our genetic makeup is inscribed in our DNA, a suite of three billion ‘base pairs’ (combining the four bases identified as T, A, G and C) contained in our chromosomes and including some twenty thousand sequences (series of bases) that form our genes. This information is what determines the synthesis of the thousands of proteins underlying the structure and function of the two hundred cell types that make up a human organism: information transmitted to us from our parents and transferred by us to our offspring. We all carry the same genes, arranged in the same order on our chromosomes, but they are not strictly identical between individuals. Now that we can read (‘sequence’) an individual’s entire DNA at a reasonable cost (below one thousand euros), it is possible to compare different people’s DNA and ascertain that some three million differences exist between the DNAs of two randomly selected individuals.¹ These differences affect a single base, *i.e.* at a given point on a given chromosome, there is a given base (*e.g.*, G) in one individual, while in the other person there is a different base (*e.g.*, A). Three million out of three billion — one per thousand— represents the well-known 0.1% difference, which has led some thinkers to deduce that ‘there is no such thing as human races’ since the DNA of two

randomly selected people is 99.9% identical. As we will see, the truth is not so simple, and it depends very much on the meaning given to the word Race (Jordan 2008).

- 3 In fact, although there is *a priori* nothing negligible about three million differences, we can start by saying that the great majority of them are insignificant, with no impact on the phenotype (appearance, physiology, behaviour, etc.) of the person carrying this DNA. In fact, genes *per se*—sequences of bases coding for a protein or ‘coding sequences’— represent only about 1% of our DNA, with the great majority of the remaining 99% (sometimes called ‘junk’ DNA) being noncoding, with seemingly no specific role. Thus, differences located in these regions will have no functional consequence on the person carrying the corresponding DNA. Furthermore, even if the difference concerns a coding sequence, it is not necessarily significant: the ‘genetic code’ (which indicates which amino acid of the protein is specified by which DNA sequence) is ‘degenerate’ (in the mathematical sense of the term), meaning that several ‘codons’ (series of three bases in DNA) code for the same amino acid. For instance, histidine, one of the twenty amino acids present in the proteins, may be coded in DNA by CAT or CAG. As a result, if T is replaced by G in this codon, the same amino acid will be incorporated into the protein, which will not be different even though the DNA sequence is not the same: this is called a ‘silent’ difference (or mutation). In all, only some ten thousand of these three million differences are significant in that they change a protein formula. Modification of a protein (*e.g.*, a digestive enzyme) may make it more or less active, more or less sensitive to temperature, or even change nothing: not all of the several hundred amino acids constituting a protein are critical. If, however, the difference in the DNA generates a ‘stop codon’ (TGA, TAA or TAG) which indicates ‘end of message’, synthesis is interrupted and the corresponding protein is truncated and *a priori* non-functional: this is the case for about a hundred differences. This could appear to be catastrophic, but it must be remembered that each of our chromosomes exists in duplicate (one inherited from the father, the other from the mother): if a given gene is inactive, the homologous gene carried on the other chromosome can usually compensate. There can also be ‘missense’ mutations, which cause a change in amino acids and may sometimes (but not always) alter or suppress the protein’s function. In all, it is this set of genetic differences that leads to phenotypic diversity between individuals of our species: visible traits (*e.g.*, skin, eye or hair colour, adult height), as well as hidden differences (*e.g.*, blood type, lactose intolerance, tendency to obesity, vulnerability to certain forms of cancers, etc.). We must emphasise that although the DNA sequence determines many phenotypic elements, with a few exceptions, we are unable at present to deduce the phenotype from reading of the DNA sequence. It must also be recalled that genetic diversity is modulated by the environment, by living conditions and personal history—sometimes significantly as for height and obesity.
- 4 The study of human diversity at the genome level was able to develop only thanks to recent scientific and technological advances: sequencing, the first reading of a human genome, completed in 2003, then the development, in the following decade, of very sophisticated tools, ‘DNA chips’ or microarrays, microscopic DNA spots making it possible to rapidly establish the genetic profile of a given DNA sample. These systems can analyse several hundreds of thousands of particularly variable points in DNA (often called SNP, or Snip, for ‘Single Nucleotide Polymorphism’) and determine which allele (‘version’) of a SNP is present at each of these points. At present, these microarrays easily examine five hundred thousand points in human DNA, for a cost of a few dozen

euros per sample, which makes it possible to study thousands of people's DNA (Jordan 2008). More recently, thanks to amazing improvements in DNA sequencing techniques, the cost of complete sequencing of human genomes has dropped under one thousand euros —while it can be remembered that the first sequencing had mobilised hundreds of laboratories over a decade and cost several billion dollars. Complete DNA sequencing clearly provides a more precise image of a person's genetic profile (three billion points compared to five hundred thousand for a microarray), but it is not yet very widespread because of its cost and the complexity of its implementation for population studies bearing on thousands of people.

- 5 Determination of the genetic profile of several people —*e.g.*, using microarrays— opens the way to defining and calculating 'genetic distance' between two persons. Two genetically identical individuals (identical twins) will have the same allele for each of the five hundred thousand SNPs analysed, and zero genetic distance. Except in such cases, the genetic distance between two people depends on the number of different alleles identified by analysis, compared to the total number of SNPs examined. For two related individuals, the distance is smaller, depending on the closeness of kinship. At this stage, one can easily imagine that analysis of a population according to this approach makes it possible to examine the existence of 'clusters', *i.e.* groups of genetically similar people, as discussed in the next section.

2. Population studies and the appearance of clusters

- 6 Population studies performed on the basis of individual genetic profiles involve measuring genetic distances between people within that population, and presenting the results in an immediately intelligible way. To illustrate this approach and its results, I will consider a study published in 2012 by the *1000 Genomes Project*, an international consortium that analysed the DNA of 1,092 people of different origins —a detailed analysis, which involved complete sequencing of their DNA, thereby adding further precision to this study. The genetic profile of each of these 1,092 individuals was compared with that of each of the 1,091 others, calculating genetic distance between them in each case. The result was over a million genetic distance values: it is difficult to obtain a graphic representation of these results involving thousands of points to be positioned in a highly multidimensional space... This requires resorting to what has become a standard mathematical technique: decomposition into 'principal components'²: the method reveals the principal component, *i.e.*, the direction in which the dispersion of points is greatest, then the second component (the perpendicular direction with the greatest dispersion), and so on. In fact, the analysis often stops with the first two components, and the results can then be presented in a two dimensional space, *i.e.* as a graph on a sheet of paper, as shown in Figure 1:

Figure 1. Distribution of points indicating the position of DNA from 1,092 people analysed for the first two principal components (PC 1 and PC 2). Each point represents the position of one individual according to values for the two principal components (modified partial excerpt from Figure 4 of the supplement of *1000 Genomes Project* (2012), Creative Commons Licence).

- 7 Clearly, these points, each representing one DNA sample —one individual—, are not randomly distributed on the graph, nor are they all clustered together: they form a constellation with three main groups and clouds of lesser density between some of these groups. It should be emphasised that this result is based solely on the analysis of DNA, and that the origin of the individuals concerned was never taken into account. Thus, objectively, sub-groups of genetically similar people exist within the population studied.
- 8 The next step naturally entails examining the extent to which the groups thus revealed correspond to geographically and socially defined populations. This is shown using the same graph, where each population is identified by the shape and colour of the symbol.

Figure 2. Here we show again the data of Figure 1, but the origins of the different individuals are indicated, revealing three compact groups: European (GBR, CEU, FIN, in blue); African (YRI, LWK, in grey) and Asian (CHS, CHB, JPT in green). The points corresponding to other populations –African-American (ASW) and Mexican (MXL), in particular– are more scattered, which reflects their mixed origin. Abbreviations: GBR, British; FIN, Finnish; CHS, Han Chinese in Southern China; PUR, Puerto Rican; CLM, Colombian in Medellín; IBS, Spanish in Spain; CEU, residents of Northern European descent in Utah (USA); YRI, Yoruba in Ibadan, Nigeria; CHB, Han Chinese in Beijing; JPT, Japanese in Tokyo; LWK, Luhya in Kenya; ASW, African Americans in the South-Western USA; MXL, Mexicans in Los Angeles; TSI, Tuscans in Italy (partial excerpt from Figure 4 of the supplement of *1000 Genomes Project* (2012), Creative Commons licence).

- 9 Indeed, what immediately stands out is that the three zones of the graph with the highest density correspond to geographically defined populations: Europeans for blue points (GBR, FIN, CEU), Asians for green (CHS, CHB, JPT) and Africans for grey (YRI, LWK). Regarding the more widely scattered points, they concern African Americans in the United States (ASW), known to be often partly of European descent and thus positioned between Africans and Europeans. The same applies to Puerto Ricans (PUR), Mexicans and Colombians (MXL, CLM).³ In all, the analysis without *a priori* hypotheses of genetic distance applied to groups of people of diverse origins objectively reveals clusters, which correlate fairly well with ‘ethnic’ groups socially defined as European, Asian or African. In fact, these groups correspond with geographical origins (European, Asian, African) and are reflections of our species’ evolutionary history. After its emergence in Africa, two or three hundred thousand years ago, *Homo sapiens* (Modern Man) began spreading, mainly eighty or one hundred thousand years ago, gradually peopling the rest of the Earth. Natural selection acting in very different environments, genetic drift (the random loss of certain variants in small human groups) and perhaps sexual selection (choice of mates according to certain visible characteristics) led to the differentiation of human groups. Though limited, since a hundred thousand years is little on an evolutionary scale, this differentiation is nonetheless real.

- 10 One last point: the graph in Figure 2 gives the impression of separate populations, at least some of which show fairly clear boundaries with no overlap between them. Shall we agree with Arthur de Gobineau (1816-1882) who, in his *Essay on the Inequality of the Human Races* (1853), asserted that Humanity is 'divided' into different races? Of course not: it must be remembered that analysis into key components enables us to choose the representation that most differentiates the points corresponding to individuals. The very great majority of possible overviews of the same data do not differentiate clusters of points. And the assertion that genetic differences within a population are greater than the mean difference between two populations, which was taken to be the scientific proof that 'there is no such thing as human races,' remains valid, although it applies to characteristics considered in isolation. Almost all alleles (SNP variants) exist in every population, which, from this point of view, is almost as diverse as all humanity; but it is the frequencies of different alleles that vary between groups which make it possible to circumscribe them. To offer a simple example, individuals with light-coloured eyes can be found in a purely African population, but they are extremely rare, while they represent the majority among Northern Europeans.

3. From genotype to phenotype for individuals and groups

Individual diversity

- 11 The previously used example is improper, since eye colour is not determined by one gene, but by the expression of several—at least ten—genes. This is true for most visible characteristics (as well as many others), which leads to questioning how genotype (what is inscribed in the genes) determines or influences phenotype (an individual's morphological, functional, behavioural and other characteristics). So far, we have only considered genetic diversity, ascertained through DNA, but we must now examine how it is expressed and embodied in individual characteristics. We must begin by stating that, even today, elucidating this transfer remains very complicated for several reasons. First, as I have just said, a trait is usually determined (or rather influenced) by several genes: *e.g.*, despite its strong heritability, adult height is linked to over a thousand genes, eye colour to about ten and skin colour to over a hundred genes identified in different populations. And having identified a series of genes influencing one trait does not mean that DNA analysis makes it possible for us to predict this trait. That would require knowing the precise role played by each of these genes and how they may influence this particular trait, which is rarely the case. Furthermore, an individual's environment and personal history sometimes play a major role in the expression of a trait: the genetic potential for stature, for example, is achieved only if the child, then the teenager, has not suffered from any nutritional deficiency. Better diet in childhood is what explains that the Dutch grew from a mean height of 1.70 metres (in 1914) to 1.83 metres in 2014, rather than any significant change in their genetic heritage over that century, although the role of genotype remains significant. If we take the example of facial shape and features, strict genetic determination is obviously important, as attested by the resemblance between identical twins (with identical DNA at least at birth). Yet, at present, and despite certain statements that are more promotional than scientific, it is not possible to draw an accurate 'facial

composite' on the basis of DNA analysis (Jordan 2020). At best, it may be possible to estimate that someone's appearance may be more European, Asian or African, but this information is deduced from their ties with an ancestry group —whose approximate characteristics are known— rather than any precise knowledge of the genes determining appearance. The analysis of some forty genes can provide an idea of eye or hair colour and (more recently) skin colour (Jordan 2020). This is already quite considerable and may even raise ethical problems, but remains insufficient to characterise a face, despite the importance of DNA in determining its appearance.

- 12 Thus, individual genetic diversity combined with the multiplicity of lifestyles produces the great phenotypic variety we can observe in the human species. We are 'all related, all different'⁴ and these differences may provide an advantage or a disadvantage, depending on our living conditions. For Palaeolithic hunter-gatherers, being near-sighted must have been a major handicap, while having a mutation leading to severe neuromuscular degeneration after the age of forty⁵ had little incidence, given the short life expectancy at the time. For *Homo sapiens* today, the situation is completely reversed: although one individual is not intrinsically superior or inferior to another, he or she may be more or less adapted to a given environment.

Group diversity

- 13 How does such genotypic and phenotypic diversity manifest itself in human groups? This is where we find ourselves on shaky and even dangerous ground, which has given rise to a great many stereotypes which persist to this day: Asians supposed to be intelligent and cunning, Blacks lazy but musically gifted, Europeans industrious and authoritarian, etc. These sometimes insulting qualifiers are largely linked to a colonial past of oppression on the part of 'superior races' for the good of 'inferior races' as Jules Ferry asserted.⁶ In fact, to be able to analyse phenotypes in human groups, such groups must be defined by their genetic closeness, not as social categories. As shown in Figure 2, African Americans living in the South-Western United States (ASW) do not form a clearly defined genetic group (unlike the Luhya of Kenya, LWK), any more than Mexicans in Los Angeles (MEX) or Puerto Ricans (PUR). This is a fatal flaw in most studies conducted in the United States on topics such as differences in the incidence of diseases like diabetes or prostate cancer between different 'ethnic' groups. Since Race is a category featured in periodical population censuses (Figure 3), it is obviously convenient to take it as an approximation of identification with a group. But, as the above-mentioned examples illustrate, this approximation is misleading, and the delimitation evoked in this case by the term Race is actually nothing more than a social category of no specific value from a biological standpoint. In the case of African Americans, this social category is genetically heterogeneous because of historical factors: the sadly notorious 'one-drop rule', which remained in effect until World War II, and according to which someone with a single drop of 'Black blood' (one African ancestor) was listed as 'Black', despite possibly being genetically and phenotypically more 'White' than 'Black'.

9. What is Person 1's race?
 Mark ☒ one or more boxes **AND** print origins.

☐ White – Print, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc. ☐

☐ Black or African Am. – Print, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc. ☐

☐ American Indian or Alaska Native – Print name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, etc. ☐

☐ Chinese ☒ Vietnamese ☐ Native Hawaiian

☒ Filipino ☐ Korean ☐ Samoan

☐ Asian Indian ☐ Japanese ☐ Chamorro

☒ Other Asian – Print, for example, Pakistani, Cambodian, Hmong, etc. ☐

☐ Other Pacific Islander – Print, for example, Tongan, Fijian, Marshallese, etc. ☐

☐ Some other race – Print race or origin. ☐

Figure 3. The Race entry in the 2020 United States census form (excerpt from the form) is a self-declaration, and it is possible (though rare) to check more than one box. 'Person 1' is the signatory (Form available at: <https://www.census.gov/programs-surveys/decennial-census/technical-documentation/questionnaires/2020.html>).

- 14 What can be said of phenotypes linked to an ancestry group, and of the average difference between them? To the extent that they are seen as well-defined groups (with strong genetic proximity), there is actually some degree of equivalence between each group and visible phenotype: skin, eye and hair colour, facial type and features, etc. This 'mean' phenotype for each group results from natural selection (such as light skin for Europeans being favourable for vitamin D synthesis), 'genetic drift' (migrations resulting in loss of some alleles of certain genes) and possibly sexual selection (choice of partners according to certain aspect criteria). For a given phenotypic characteristic, there remains a large amplitude of variation within the group (some Africans from Africa are rather light-skinned, while others are very dark), but 'mean type' within one group is clearly distinguished from that of another. This can easily be observed by watching passers-by on the street, and can explain the scepticism triggered by categorical, blunt assertions such as 'there is no such thing as human races.'
- 15 So, what about deep phenotype, sensitivity to diseases, behaviour and aptitude? It could be assumed that such differences are less significant than those affecting appearance, in short that we differ in our 'bodywork' despite having the same engine and chassis. This could have resulted from a major role of sexual selection (based mainly on individuals' appearance), but genomic analysis has largely belied this hypothesis. We will look at the matter of disease later, but the study of physical and intellectual performance shows how difficult it is to differentiate the role of genotype and of environment and personal history. We can take an example that has generated considerable debate, that of intellectual performance. It is the subject of many, mainly English-language studies bearing on the heritability of IQ (intelligence quotient) and its

level in different populations. Two main conclusions have been drawn: first, that the heritability of IQ is significant, at least 0.5 (50%), then that its mean value differs between populations. We should note that 50% heritability does not mean that half of IQ is genetic and the other half environmental, but rather that in a population with a certain variability in IQ (90 to 110 for instance), half of this variation (*i.e.* ten IQ points) is linked to genetic heritage. As for mean IQ, depending on the studies, it seems 5 to 10 points lower in African Americans compared to Europeans or Asians. The fact that this is a highly heritable characteristic, coupled with the difference between two 'ethnic' groups led some people to assume that African Americans' intellectual inferiority was genetic. First, we must emphasise that IQ does not measure intelligence, but is a parameter that (more or less) correlates with it, and that an individual's IQ score is calculated on the basis of a series of tests (association, abstraction, logic, etc.) that are very probably affected by the person's culture (family, values, lifestyle, etc.). If we accept that this test is truly objective, and also overlook the genetic heterogeneity of African Americans, does the study prove that genes determine the differences between Europeans and Africans? Of course not: we must remember that, if the mean height of Dutch men grew by 13cm in one century, it was not because of any change in their genes between 1914 and 2014, but rather a change in their diet and living conditions. Yet adult height remains one of the most inheritable traits (80%) —but only with 'all things remaining equal.' Unfortunately, it is obvious that there is little 'equality' between being European and being African American in the United States.

- 16 This example illustrates the difficulty there is in defining what is due to heredity and what is due to environment when comparing human groups. Similarly, despite the predominance of Kenyans or Ethiopians among long-distance runners, it has not been possible to identify gene variants capable of explaining this success, so it is assumed that it results from a set of environmental and cultural factors interacting with a generally favourable genotype.
- 17 In the end, we can conclude that an individual's genotype largely determine his (her) phenotype but, with a few exceptions, we do not yet know how to use the former to deduce the latter. As for groups, they display many differences between mean values for specific phenotypic components, but it is generally very difficult to determine the extent to which these differences are due to genetic heritage, environment or both.

Genes, human groups and disease

- 18 An individual's genetic heritage has a role in their susceptibility to many health issues. This is particularly true for Mendelian or monogenic diseases, resulting from a change in a specific gene, *e.g.*, haemophilia, cystic fibrosis or Duchenne muscular dystrophy. Altered DNA from one parent, or resulting from a new mutation, can lead to the production of a defective protein that does not play its role in the organism and causes a clearly defined pathology. Such diseases are genetically determined in that all persons with that particular mutation will be stricken by that disease. They are also quite rare because those carrying them are severely handicapped, which means they are not favoured by natural selection. But there are many less debilitating and more widespread diseases, like diabetes, high blood pressure and Crohn's disease⁷ which have a significant genetic component: the risk of having diabetes, for example, depends greatly on a person's gene variants. In general, this genetic component involves tens or hundreds of genes, with each individual variant having a very small effect on the risk of

being affected, just as it is also highly dependent on environment (diet and obesity for diabetes). Finally, the risk of cancer is also influenced by genetics, as with breast cancer (where some mutations in the BRCA1 and BRCA2 genes strongly increase the risk) and many others (prostate cancer, in particular). Thus, we all carry an assortment of gene variants that, in a given person, may induce a risk of diabetes that is much lower than average, but also a high risk of prostate cancer.

- 19 We can also look at differences in the incidence of various diseases according to a given human group. Such differences are a reality for Mendelian diseases: for some of them, the frequency differs between groups, which reflects these populations' past history. Cystic fibrosis (inactivation of the gene coding for an 'ion channel') is most present among Northern Europeans and their descendants in the United States; sickle-cell anaemia (a mutation of the gene for haemoglobin) is found mostly in people of African descent. In the latter case, it can be explained by the fact that, when heterozygous,⁸ this mutation affords protection against *Plasmodium falciparum* infection, which causes malaria, a disease present essentially in warm moist regions, and is pathogenic only when homozygous.⁹ Similarly, it has been suggested that the mutation that causes cystic fibrosis when homozygous affords protection from infantile diarrhoea (once a major cause of infant mortality) when heterozygous. Other genetic diseases seem to result from relatively recent mutations (appearing in the past few thousand years) which remained confined to geographically or culturally isolated human groups, such as Tay-Sachs disease among Ashkenazi Jews. There are many other examples, which also concern non-pathogenic adaptive mutations, like the capacity for adults to digest milk (persistence of lactase activity) found mainly among Europeans and certain traditionally pastoral African groups.
- 20 For the most common diseases, influenced by many genes and living conditions, the situation is more complex. Numerous studies conducted in the United States (thanks to the existence of ethnic statistics) suggest that African Americans are likelier to suffer from hypertension and prostate or lung cancer than 'Caucasians',¹⁰ who, in turn, may be more affected by osteoporosis and skin cancer. But there are proportionally more obese people and smokers among African Americans, which increases the likelihood of hypertension and lung cancer. Since 'Whites' have a longer life expectancy, they are further at risk of being stricken by diseases associated with old age, like osteoporosis, etc. Because these two groups' environment diverges significantly, it is harder to assert the potential role of genes —except for skin cancer, linked to exposure to the Sun of light skin, whose colour is genetically determined. There have been attempts to take into account these socio-economic factors, but this is very difficult, since cultural identity also influences access to care, as well as the propensity to consult doctors and the choice of certain types of therapeutic approach, even for people of comparable economic level. Thus, it can be said that belonging to a social category (which is what the term 'African American' actually refers to) has a certain predictive value for different disorders —but nothing proves that these predictions are linked to this category's genetic makeup, which is anyway genetically heterogeneous in most cases.
- 21 In short, there are genuine differences in the frequency of diseases between socially defined human groups (often known as 'ethnic groups', to avoid calling them 'races'), but, aside from cases of Mendelian diseases in very specific populations, nothing proves that these differences are genetic in origin. In fact, this seems highly unlikely, given the genetic diversity of certain groups, such as African Americans (Figure 2).

4. Ambiguity of the concept of Race, reality of ancestry groups

A variable and ill-defined concept

- 22 Assigning a person to a socially defined Race relies largely on visible characteristics (skin, eye and hair colour, facial shape), which are genetically determined, although we cannot yet infer them all from DNA analysis. But such identification is also strongly influenced by the history of the society to which the individual belongs: it is by virtue of the ‘one-drop rule’¹¹ that Barack Obama is identified as the first Black president of the United States, although he is as White (through his mother) as he is Black (through his father). It also depends on the ‘narrative’ adopted by a social group to define itself as a racial group while, like the self-proclaimed Aryans of the Third Reich: such categories claiming to be biological in nature represent no more than an origin myth. And, of course, this concept has a tragic history: largely invoked to justify the practice of slavery and colonisation (the duty of superior races to civilise those that are inferior, according to French statesman Jules Ferry (1832-1893) —in agreement with most of his contemporaries), consubstantial later with Mussolini’s fascism and especially Nazism, underlying the Armenian and Rwandan genocides and, of course, the Shoah, an embodiment of true Evil. Thus, it strikes me as legitimate to ban this term, since it is inaccurate, likely to give rise to highly divergent interpretations and largely tainted by history. In this sense, it can be said that ‘there is no such thing as human races,’ if this term is given the meaning attributed to it by Gobineau and many others. It is not possible, however, to deny the existence of human groups defined by their DNA, reflecting the history of our species and displaying significant differences (Reich 2018).

‘Ancestry groups’, a solution?

- 23 How can we identify groups whose objective existence is proven (Figures 1 & 2) while avoiding the term ‘ethnic group’, which is often no more than a euphemism for Race and, if taken literally, refers more to a cultural than a biological entity? The term ‘ancestry group’ has not (yet) prevailed, but it has the great advantage of corresponding to a precise definition: a set of people who are genetically related through their common origin, often attached to a part of the world where this group has remained relatively isolated for several millennia. Over the centuries, such groups change (slightly) through the above-mentioned mechanisms (selection, genetic drift, sexual selection) and acquire certain common traits corresponding more or less, at least in their appearance, to ‘racial’ stereotypes. But, beyond such correspondence, ancestry groups are an objective reality, and their geographic origin can often be evaluated by comparing their genetic profiles with reference profiles defined by the study of ancestral populations. Ancestral populations are human groups whose history is sufficiently well known to assert that they remained static without significant intermixing over a few centuries or, increasingly, from DNA drawn from remains or from bone, for which detailed analysis has recently become possible (Reich 2018). The definition of ancestry groups is objective, unlike their relationship to past populations which depends strongly on the size and quality of the reference groups. Businesses like

23andMe, which are specialised in establishing such profiles for a moderate fee (US\$99), provide their customers with very detailed information (Figure 4) that should not be taken too seriously when, for example, it specifies 14.6% Native American descent, while omitting any margin of error, but which are generally valid: the origins of the person represented in figure 4 are indeed mainly (by order of importance) in Sub-Saharan Africa, Europe and Asia (where Native Americans originated).

Figure 4. An example of the result supplied by 23andMe after analysis of an individual's (Monica De Armond) DNA. The world map at the centre indicates to what the colours correspond; the breakdown into major continental descent groups is shown in the first concentric circle (white corresponding to unassigned components). The second concentric circle provides further details on origins, when it is possible to define them, as does the third circle. The apparent accuracy of the figures presented (e.g. 14.6% Native American origin) is illusory because the (unreported) margins of error are quite significant (taken from the company's site).

- 24 Unlike 'ethnic groups', 'ancestry groups' are not a politically correct substitute for the term 'human race'. Race, in the traditional sense of the word, refers to clearly distinct, homogeneous entities: to quote Gobineau in the mid-19th century (1853), humanity is 'divided' into races —which are not equal, the white race being obviously superior to all others. And the definition of Race, despite claims that it is grounded in biology, is very much influenced by behaviour, customs, dress, etc. —in short, Culture, as Claude Lévi-Strauss so well demonstrated (2001). Ancestry groups, however, are defined empirically, solely on the basis of DNA analysis, which can be linked to geographic origin only secondarily, with methods that remain subject to criticism and are highly dependent on the quality of reference groups. As we have seen, although visible phenotypes are quite largely tied to ancestry groups and certain health issues and metabolic traits may be virtually specific to certain groups (in which they appeared as a result of mutations), most significant elements are influenced by both genetic and environmental factors, thereby seriously complicating any causal analysis. With progress in genomic analysis techniques, it has become possible to assess the outcome of a large number of gene variants on a particular phenotypic trait, but it remains very difficult to distinguish what is influenced by genetics from what results from environmental factors in an individual's evolution, let alone that of a group, even when defined by objective analysis.

In short...

- 25 The information obtained from analysing human DNA sheds light on many as yet unresolved issues. It has revolutionised physical anthropology by revealing the complexity of human migrations and the mixing of populations over the recent millennia of our history through the study of ancient DNA (Reich 2018). As I hope to have demonstrated above, it deconstructs the notion of Race, which has largely prevailed in Western civilisation until the mid-20th century, while also disputing the somewhat simplistic assertion that ‘there is no such thing as human races.’ True, unbiased analysis of human DNA reveals the objective existence of groups whose members are genetically more closely related to each other than to other individuals; and indeed, the groups thus defined often correspond, although only roughly, to ‘racial’ —or ‘ethnic’ (to be more politically correct)— categories on the basis of the appearance and cultural identity of the individuals concerned. But the boundaries between these groups are hazy, with considerable internal diversity: they do not display the exclusive static hereditary nature of the traditional notion of race. This mirrors the history of our species which, though recent (only two or three hundred thousand years —which is quite short on the evolutionary scale), has a complex history of differential evolution depending on regions of the globe and their characteristics, but also influences by migrations and mixing between populations which have reshuffled the cards by recombining characteristics that initially appeared in distinct groups. This on-going, increasing ‘intermingling’ is not tending towards greater uniformity but, instead, is spreading the different alleles of so many genes, thereby boosting our species’ resilience, something we will need in the face of the challenges awaiting us...
-

BIBLIOGRAPHY

- 1000 Genomes Project Consortium. 2012. “An integrated map of genetic variation from 1,092 human genomes.” *Nature* 49: 56-65.
- De Gobineau, Joseph-Arthur. 1853. *Essai sur l'inégalité des races humaines*. Paris: Firmin Didot Frères.
- Jordan, Bertrand. 2008. *L'humanité au pluriel, la génétique et la question des races*. Paris: Ed du Seuil.
- Jordan, Bertrand. 2020. “ADN et portrait-robot: où en est-on?” *Med Sci (Paris)* 36: 813-816.
- Lévi-Strauss, Claude. 2001. *Race et Histoire* [1952]; *Race et Culture* [1971]. Paris: Albin Michel/Unesco.
- Reich, David. 2018. *Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past*. Oxford: Oxford University Press.

NOTES

1. For a good quantified discussion on the differences between individuals, see: <http://book.bionumbers.org/how-genetically-similar-are-two-random-people/>.
 2. https://fr.wikipedia.org/wiki/Analyse_in_composantes_principales.
 3. Native Americans (not featured in this figure) are genetically quite similar to Asians, who are at the origin of the peopling of the Americas some fifteen thousand years ago. The scatter of points corresponding to Mexicans and Colombians reflects their mixed Native American and European descent.
 4. From the title of an exhibition, *Tous parents, tous différents*, at Musée de l'Homme in Paris from 1990 to 2015.
 5. Like the mutation responsible for Huntington's chorea.
 6. Speech to the French National Assembly on 28 July 1885.
 7. Chronic inflammation of the intestinal wall.
 8. Mutated gene on one chromosome, normal on the other.
 9. Mutation on both chromosomes.
 10. European, White, etc.
 11. *A white person ... has no trace whatsoever of any blood other than Caucasian*, State of Virginia Racial Integrity Act of 1924.
-

ABSTRACTS

The sequencing of the human genome (2003) has been followed by a number of technical developments that allow detailed characterization (including complete sequencing) of the DNA of thousands of individuals. This has provided an estimate of human genetic diversity: approx. 3 million base substitutions within our genome that includes 3,000 million bases. Although the divergence between any two individuals is small, it is responsible for much of the phenotypic diversity observed within the human population, and current techniques make it possible to measure genetic distances between any two individuals whose genome has been analysed. When results from a large sample of persons are analysed by sophisticated multi-dimensional representation, fairly distinct clusters appear in the population, indicating groups of individuals that are more similar to each other than to persons from other groups. Although these groups are defined by genome analysis, without any *a priori* information on their origin, it turns out that they largely correspond to so-called “ethnic” categories that reflect genetic ancestry (e.g. Europeans, Africans and Asians). Every group is internally quite diverse, but analysis of thousands of genetic markers does indeed differentiate between them. “Race”, as commonly defined, is a very imperfect proxy for these ancestry groups, as it suggests high homogeneity within each group, implies assumptions about behaviour and “character” that have no scientific basis, and is heavily tainted by past uses in support of oppression and even genocide. Thus “ancestry” or “ancestry group” is a highly preferable term, especially now that more and more people are of mixed ancestry. These methods have been widely applied to current populations and to “ancient DNA” from historical samples, and have made possible great strides in understanding the history of our species.

Le séquençage du génome humain (2003) a été suivi de progrès techniques qui permettent aujourd'hui la caractérisation détaillée (y compris la séquence intégrale) de l'ADN de milliers d'individus. Cela a fourni une estimation de la diversité génétique humaine, soit environ 3 millions de différences dans notre génome qui contient 3 milliards de bases. Bien que la divergence entre deux individus soit faible (environ 0,1 %), elle est responsable de la majeure partie de la diversité phénotypique observée au sein de la population humaine. Les techniques actuelles permettent de mesurer la distance génétique entre deux individus dont l'ADN a été analysé. En appliquant des méthodes de représentation multidimensionnelle, les résultats obtenus pour un grand ensemble d'individus montrent l'existence de groupements (*clusters*) regroupant des personnes plus similaires entre elles qu'avec des personnes d'autres groupes. Bien que ces groupes aient été définis uniquement par l'analyse moléculaire de l'ADN, ils s'avèrent correspondre approximativement aux groupes dits « ethniques » (par exemple Européens, Africains et Asiatiques). La diversité interne de chaque groupe est importante, mais l'analyse de milliers de marqueurs génétiques permet de les distinguer. Le terme de « Race », tels qu'il est généralement compris, ne correspond pas à la réalité de ces groupes : il suggère une forte homogénéité à l'intérieur de chaque groupe, implique des notions de comportement et de « caractère » sans fondement scientifique, et est marqué par son utilisation ancienne en tant que justificatif d'oppression et même de génocide. « Ascendance » ou « Groupe d'ascendance » sont des termes bien préférables, d'autant plus qu'aujourd'hui beaucoup de personnes présentent un mélange d'ascendances. Les méthodes décrites ici ont été largement appliquées aux populations actuelles et à de l'« ADN ancien » provenant de restes humains ou même de fossiles, et ont permis de grands progrès dans la compréhension de l'histoire de notre espèce.