Disease Biology, Genetics, and Socioecology
https://www.sciltp.com/journals/dbgs

Review

Race/Ethnicity/Ancestry (REA) and Disease: A Complex Interplay of Social and Biological Factors

Michel Tibayrenc

MIVEGEC, IRD, CNRS, Univ. Montpellier, 34090 Montpellier, France; michel.tibayrenc@ird.fr

How To Cite: Tibayrenc, M. Race/Ethnicity/Ancestry (REA) and Disease: A Complex Interplay of Social and Biological Factors. *Disease Biology, Genetics, and Socioecology* **2026**, *2*(1), 1. https://doi.org/10.53941/dbgs.2026.100001

Received: 4 June 2025
Revised: 14 July 2025
Accepted: 15 August 2025
Published: 16 January 2026

Abstract: The connection between human genetic and geographical diversity on one hand, and disease distribution on the other hand, has been a subject of ongoing research. A key challenge lies in defining the relevant units of analysis (RUAs) within human populations to accurately assess the impact of disease. While "race" is often considered a social construct, biomedical studies frequently utilize major geographical populations—European, Asian, and African groups—As RUAs. Other groups like "Hispanics" or Native Americans are also commonly used. This review examines the findings of this prevalent approach and argues that it offers significant advantages over the historically dominant model that focused solely on European populations.

Keywords: anthropology; medicine; population genetics; sociology; subspecies; taxonomy

1. Introduction

The issue of human genetic diversity is a critical concern, particularly due to its connection with racism and discrimination. The 20th century witnessed the rise and devastating consequences of racist ideologies such as colonialism, the slave trade, and German Nazism. These historical events highlight the importance of understanding human genetic diversity in a responsible and nuanced way, avoiding the pitfalls of racial essentialism and fully recognizing the significant impact of social factors on health outcomes. While acknowledging the social construction of race, it is also important to recognize that genetic variations do exist across different geographical populations and can often play a role in disease susceptibility. Further research is needed to disentangle the complex interplay of genetic, environmental, and social factors that contribute to health disparities across different geographical groups.

2. A Population Genetics* and Phylogenetic* Framework for the Analysis of Disease Impact in Human Populations

Phylogenetic character mapping (PCM; [1]) involves a two-step process. First, the genetic and phylogenetic diversity of the species in question is characterized. Subsequently, genes and phenotypes of interest, particularly those associated with pathological processes, are mapped onto this established population genetics and phylogenetic framework.

Rather than analyzing genes* of disease susceptibility (or those assumed to be so) in an isolated manner, it is more relevant to first build a general population genetics framework of the species considered [2]. Neutral genes* and neutral polymorphisms* are the best choice for designing this overall picture because they are considered to be passive markers of the time elapsed ("historical markers"). In other words, differences in neutral genes between two populations or species are considered to be correlated to the time elapsed since these two populations or species shared common ancestors.

By definition, a neutral polymorphism is not impacted by natural selection*. Although it is extremely difficult to strictly verify this assertion, many genetic traits are classically considered as neutral. For example, in the genetic code, synonymous mutations* have no consequences on the amino acid encoded. Noncoding sequences (introns*,

spacer DNA* sequences) are considered neutral until further notice This is supported by the fact that their molecular clock* is generally faster than the molecular clock of coding sequences, which are prone to natural selection. The use of genes* that are highly selected for the purposes of population genetics could lead to very misleading results.

Another requirement for establishing a reliable population genetic framework of any species is to use a sufficient set of genes. Sufficient does not mean thousands, although the more, the better. It is interesting to note that the phylogenetic tree shown in Figure 1 was designed after the polymorphism of only 22 different genes. It sketches the same picture as the tree shown in Figure 2, based on a much broader range of markers. However, the use of only two or three genes could be grossly misleading. "Gene trees and population trees are not the same" [3].

Two parameters could lower the correlation between how long two populations have been separated and the genetic distance* by which they differ: (i) when the populations are not strictly separated and exchange migrants (the case for human populations), the genetic distances between them are underestimated, since migrants tend to homogenize the gene pools of the populations compared; and (ii) when founder effects* interfere, which increases genetic distances. Absolute datings ("these populations split apart 50,000 years ago") should therefore be cautiously considered, which is not always the case. Relative datings are more reliable. In Figure 1, it is can safely be inferred that English and Italians share more recent common ancestors than Africans and Chinese do.

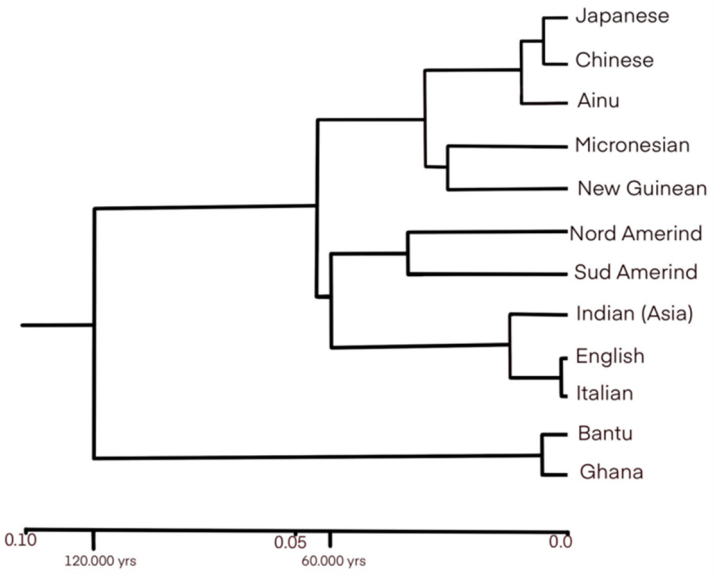

Figure 1. Phylogenetic tree depicting genetic distances for 11 isoenzyme loci and 11 blood groups among major geographical populations (Designed by Dr Jenny Telleria; after [4]).

By nature, neutral polymorphism has a limited or no predictive power on phenotype* polymorphism. On the other hand, it gives a general population genetic framework of the species under study, shows the degree of genetic similarity between populations, the rates of migration and genetic exchange, and enlightens many other relevant traits. As noted earlier, it is therefore convenient to have this general picture drawn before studying specific genes (for example in the human species, those genes that are specifically involved in the severity of diseases). A now outdated marker, isoenzymes*, has played a major role in clarifying the population structure of humans and a countless number of other organisms, including plants, animals, and pathogenic agents. It is worth noting that the results gathered by this old-fashioned genetic tool have been broadly confirmed, although with much more refinement, by the use of more sophisticated genetic and genomic markers.

2.1. A Now Classical Pattern

Isoenzymes made it possible to gather the main results that were confirmed and refined later by more advanced technologies such as microsatellites*, *Alu* insertion sequences*, or single nucleotide polymorphisms* (SNPs): (i) when historical markers are used, the main geographical populations (for example: Asians, European, Africans, Native Americans) are separated by genetic distances* that are comparable to the genetic distances observed between local populations of *Drosophila* fruitflies). Moreover, approximately 85% of human diversity is found within populations (for example, Europeans) and only 15% is due to differences between populations (for example, between Europeans and Africans) [4,5]. Genetic differences between populations can be measured by the *F*st component of the *F** statistics ([6]; Box 1 below).

Box 1. *F*-statistics.

A method proposed by [6] for describing the genetic population structure with 3 *F*-statistics, *F*is, *F*it and *F*st, whose relationships are quoted as follows:

$$(1 - Fit) = (1 - Fst)\ (1 - Fis)$$

*F*is can be defined as the correlation between homologous alleles between individuals in the local population, and *F*it as the allelic correlation in the total population. They are also called fixation indices (*F*i) and describe departures from Hardy-Weinberg expectations within local populations vs the total population:

$$FI = (h_o / h_e)$$

with h_o = observed heterozygosity and h_e = expected heterozygosity.

*F*i = 0 corresponds to Hardy-Weinberg* equilibrium; If *F*i > 0, there is a deficit of heterozygotes, due to either inbreeding (*F*is > 0) or inbreeding + population subdivision (*F*it > 0).

The *F*st component is a commonly-used measure of population subdivision (if *F*st = 0, *F*it = *F*is), and can be interpreted as the variance of allelic frequencies among populations.

In common language, *F*st is a measure of the genetic differences between two populations, for example, between people of European vs. African ancestry.

(ii) However, when phylogenetic trees are designed with isoenzymes and other, more modern markers, they recover the race/ethnicity/ancestry (REA) subdivisions classically used in physical anthropology [7] quite well (Figures 1 and 2). This results from the presence of certain alleles* that have different frequencies among geographical populations. This property (the presence of markers that are specific in a given group, or at least, that have a much higher frequency in this group) has been widely used with more modern techniques to elaborate sets of markers able to identify given geographical populations (ancestry informative markers* or AIMs) [8]. Various private companies market kits that can identify the ancestry of any individual and even their degree of admixture with a high rate of reliability. AIMs have been used in forensic medicine and are also applied to the compared genetic susceptibility of different geographical populations to given pathological traits (see below).

It is remarkable that analyses based on a wide range of modern molecular markers have fully confirmed the results obtained by the pioneering MLEE studies of the 1970s (see Figure 1), by (i) confirming that most of the genetic diversity recovered is within rather than between populations and (ii) by clearly identifying the ethnic groups* described by classical anthropology [7] (Figure 2).

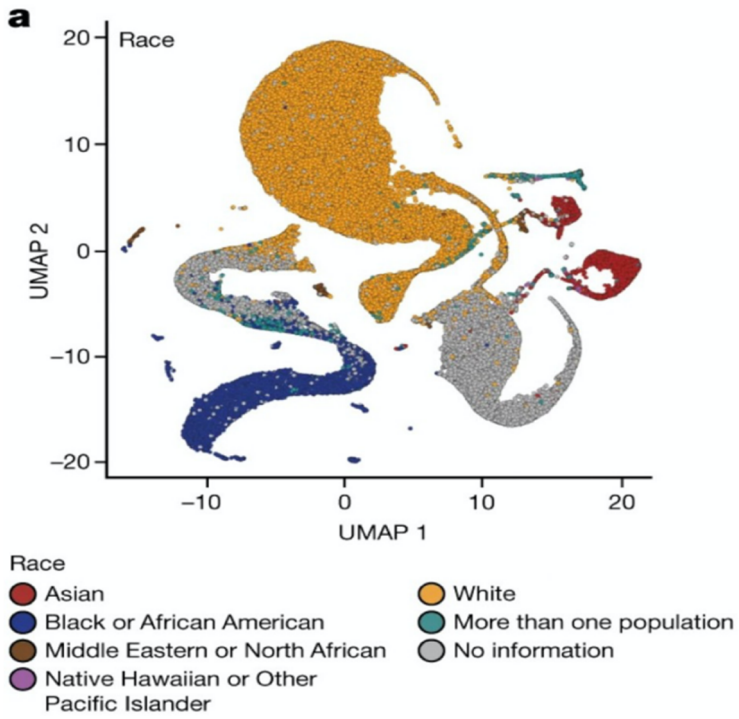

Figure 2. Characterization of major human geographic populations by state-of-the-art high-resolution genomic typing. As in the studies cited in Figures 4 and 5, the geographic groups distinguished by classical anthropology are clearly visible. African Americans appear grouped with Africans ("black") [9] (In free access).

2.2. The Contribution of Low-Frequency (LFV*) and Rare Variants* (RV*)*

A major bias of the *F*st statistic is that, particularly in [5] foundational study, it is based on so-called "common" variants, i.e., those found in all geographic populations, and whose overall frequency is greater than 2%. These variants are also characterized by a limited resolution power. In the pioneering era in which Lewontin wrote his famous article, geneticists had only this to work with. These variants, in addition to being "common" and of low resolution, have another particularity. They are generally, as already specified, "selectively neutral", that is, they do not lend themselves to natural selection. This is particularly true for isoenzymes, which were the "workhorse" of population geneticists from the 1970s onwards. These markers have no direct influence on the phenotype (all the essential properties of organisms, such as size or skin color). They are a kind of genetic "tick-tock" beating at a fairly regular rhythm. These markers are very effective in reconstructing the history of populations and their reciprocal filiations, in establishing genetic genealogies. Unlike isoenzymes, microsatellites, another category of "common", "historical" markers, can play a role in phenotypic expression, particularly in various pathologies [10]. Like isoenzymes, but with a higher resolution power than the latter, they are excellent population markers. They differentiate Europeans, East and South Asians, Amerindians and Africans into different groups or "clusters" [10,11]. They are also widely used in forensic medicine, for the identification of criminals [12].

The immense advances in genomics* have uncovered the existence of other categories of variants, which are of considerable evolutionary interest: these are the "low frequency" variants (frequency between 0.5 and 2%), and the "rare" variants (frequency less than 0.5%) [13,14]. These newcomers have two characteristics that are considerably informative for this article. Firstly, unlike "historical" variants, such as isoenzymes, they have a strong impact on the phenotype, including on "complex" diseases" (see further). Thus, these rare variants are thought to be involved in height [15], susceptibility to certain drugs [16], the genesis of diseases such as Alzheimer's, Parkinson's, psoriasis, autism, schizophrenia, as well as susceptibility to AIDS and other infectious diseases [16,17]. Secondly, these markers, defined by their frequencies at the general population level, tend to be specific to given geographic populations, in other words, to show strong differences in frequencies between populations [13,14,16,18–22]. Up to 86% of these variants are found in only one continental group [23].

It should be noted that these rare and low-frequency variants constitute the largest share of genetic variants known in the human species [24].

2.3. Conclusion: The Genetic Diversity of Human Populations Is Significant and Should Not Be Overlooked

The statement "human races do not exist" reflects a clear moral and political stance, and antiracism is undeniably a noble cause. However, this concise statement carries the risk of leading to the misconception that human genetic diversity is negligible. This is misleading. As above exposed, genetic diversity, particularly when considering low-frequency and rare markers, is abundant and unevenly distributed across continents and geographical locations. Regardless of the terminology used, even if the word "race" is to be avoided, major geographical populations (Africans, Europeans, Asians, etc.) statistically exhibit genetic differences. This genetic stratification manifests in varying allelic frequencies, including those of genes influencing key phenotypic traits. Human diseases serve as examples of this.

3. Human Genetic Diversity and Diseases

Diseases are phenotypes, representing inadequate responses to environmental pressures. Environments vary considerably across geographical locations and even more so between continents. Therefore, it is expected that the impact of diseases will generally differ among various human geographical populations.

Three main disease categories should be considered:

(1) Mitochondrial disorders: These diseases result from modifications in the DNA sequence of mitochondria*.
(2) Mendelian diseases: These are caused by mutations in a single nuclear gene or a small number of genes.
(3) Complex or common diseases: These arise from mutations in a large number of nuclear genes, with each individual mutation having a limited impact.

3.1. Mitochondrial Diseases

Mitochondrial diseases are a group of genetic disorders that affect the mitochondria. When mitochondria do not work properly, it can lead to a wide range of health problems. These diseases are generally rare.

Here are some of the main mitochondrial diseases:

3.1.1. Mitochondrial Encephalopathy, Lactic Acidosis, and Stroke-like Episodes (MELAS)

This condition affects the brain, muscles, and other parts of the body. Symptoms can include stroke-like episodes, seizures, muscle weakness, and cognitive problems.

3.1.2. Myoclonic Epilepsy with Ragged-Red Fibers (MERRF)

This disorder primarily affects the muscles and nervous system. Symptoms may include muscle twitching, seizures, weakness, and difficulty with coordination.

3.1.3. Leigh Syndrome

This is a severe neurological disorder that usually appears in infancy or early childhood. It causes progressive loss of mental and movement abilities.

3.1.4. Kearns-Sayre Syndrome (KSS)

This condition affects the eyes, muscles, and heart. Symptoms can include drooping eyelids, muscle weakness, and heart problems.

3.1.5. Leber Hereditary Optic Neuropathy (LHON)

This disorder leads to vision loss, usually in young adulthood. (Source: National Institutes of Health; NIH; https://www.ninds.nih.gov/search?search_api_fulltext=mitochondrial+disorders (accessed on 10 July 2025.))

Some disorders cannot be specifically called mitochondrial diseases. However, they are clearly linked to mtDNA polymorphism and maybe linked to geographical populations. A few examples follow.

A cohort of African Americans and "whites" identified through self-reported ancestry (SRA), that is to say: ancestry identified and declared by the individual themselves, was surveyed [25]. These authors found that blood levels of circulating cell-free mitochondrial DNA (ccf-mtDNA) were higher in the former group. Ccf-mtDNA is potentially associated with cardiovascular diseases, autoimmune disorders, cancer, pregnancy complications, and neurodegenerative diseases. The authors emphasize that race is a social construct, distinct from ancestry, as some African Americans may carry European mtDNA haplogroups*—an unsurprising finding in genetically mixed populations.

Hardelid, et al. surveyed mitochondrial OXPHOS pathway, gene, and haplogroup associations in the following populations: African Americans, Asian Americans, Latinos, Native Hawaiians, European Americans [26]. Stratified analyses were conducted by self-reported maternal race/ethnicity. Significant associations were found only in European Americans. Their findings suggest that collective mitochondrial genetic variation and particularly in the MT-CO2 and MT-ND2 genes may play a role in breast cancer risk among European Americans, but not among other geographical populations.

The study by [27] indicates that Mexican American women have an increased risk of mtDNA damage, likely linked to comorbidities such as cardiovascular diseases and diabetes. mtDNA damage may contribute to the heightened risk of Alzheimer's disease in this group. Previous research [28] has noted that Hispanics and African Americans have a risk of Alzheimer's disease that is twice as high as that of European Americans. In the study by [27], it remains unclear whether the greater burden of mtDNA damage is primarily due to comorbidities (partly influenced by environmental factors) or intrinsic properties of mtDNA.

The burden of ovarian cancer is higher in African American women than in European American women. The five-year survival rate is 38% for the former compared to 45% for the latter [29]. The authors note that comorbidities and living conditions likely play a significant role in these differences. However, they also suggest that mtDNA aberrations and mitochondrial proteins encoded by nuclear DNA are likely contributing factors.

Jaratlerdsiri et al. state that, due to maternal inheritance and ethnic-based diversity, the mitochondrial genome (mtDNA) contributes to inherited racial disparities [30]. They note that the burden of cancer is higher in African Americans than in other ethnic populations. Besides socioeconomic, educational, cultural and environmental factors, they hypothesize that mtDNA and nuclear DNA-coded mitochondrial proteins could also play a role in these differences.

3.2. Mendelian or Monogenic Diseases

This term has been coined after the name of the Czech monk Gregor Mendel (1822–1884), a pioneer of genetics through his masterpiece analysis of wrinkled and smooth peas. It refers to genetic diseases caused by a

single gene or a small number of genes. These are not the most common genetic diseases. In some cases, the disease is caused by a single mutation.

Mendelian diseases are characterized by particularly strong uneven geographic distributions [31]. SNPs specifically linked to Mendelian diseases tend to be low-frequency or rare variants [32]. This suggests that they must have highly uneven geographical distributions, since this is generally the case for low-frequency and rare variants, as we have already seen (see: "The Contribution of Low-Frequency and Rare Genetic Variants"). There is no satisfactory explanation for the uneven distribution of these Mendelian diseases among geographic groups.

3.2.1. Sickle Cell Disease*

"Sickle-cell anemia" (sickle cell disease) is highly illustrative for the topic of the present article. The persistence of this genetic variant in current natural populations can be explained by a case of balanced selection: homozygous carriers are affected by severe anemia, while heterozygotes are not ill and exhibit some resistance to malaria. This explains why this genetic trait persists in regions heavily infested by the parasite of malaria (4 species pertaining to the genus Plasmodium), most notably sub-Saharan Africa, where in some countries the transmission rate is very high. Sickle cell disease is also observed among North and South American individuals of African descent. However, in the United States, among African Americans, its frequency is somewhat lower than that observed in sub-Saharan Africa [33]. This may be due to the fact that African Americans have been protected from the selective pressure of malaria for several centuries. This shows that evolution continues to act even in modern times, and over short time scales. Sickle-cell anemia is a classic case of a Mendelian disease with uneven geographic and ethnic distribution. It is not limited to sub-Saharan Africa. It is also observed in certain regions of India, the Middle East, Saudi Arabia, and, more rarely, in Turkey, Greece, and North Africa [33]. But its frequency is especially high in Africa.

3.2.2. Hereditary Hemochromatosis* and Cystic Fibrosis*

These two Mendelian diseases are both much more common in Europe than in the rest of the world [18,21,31,34–38]. The first of these diseases is caused by iron accumulation in various organs, leading to tissue damage. Cystic fibrosis is caused by a mutation that leads to mucus buildup in the lung alveoli, resulting in respiratory failure. Manry and Quintana-Murci hypothesized that cystic fibrosis may have been maintained by a phenomenon of balanced selection, like sickle cell anemia, because this genetic trait may have provided some protection against cholera [39].

3.2.3. Phenylketonuria*

This disease is caused by a mutation that results in major metabolic disorders and mental retardation in affected children. It can be partly managed with a specific diet. According to [40], in the United Kingdom, the disease is more frequent in children of European descent, followed by those of Asian origin, and then those of African origin.

3.2.4. Huntington's Disease*

Huntington's disease is a Mendelian neurodegenerative disorder that is seen almost exclusively in Europeans [41].

3.2.5. Neurodegenerative Diseases Common among Ashkenazi Jews

Ashkenazi Jews (a historic Jewish population from Russia, Eastern Europe, Germany, and eastern France) represent a fascinating case regarding certain neurological diseases. This population is notable for having an average IQ 15 points higher than that of other Europeans, according to [42]. These authors speculated that Ashkenazi Jews have paid a high genetic price for these specific intellectual abilities. Through the mechanism of balanced selection, this high IQ would be associated with a heavy burden of Mendelian neurological diseases (Tay–Sachs, Gaucher, Niemann–Pick, mucolipidosis type IV). Ashkenazi Jews have specialized in financial jobs since the Middle Ages, as they were forbidden from owning land in those times. These financial occupations, in a world without computers, required high intellectual capacity. This fact, combined with endogamy* (the tendency to marry within one's group), driven by both persecution and tradition, and a founder effect* (the Ashkenazi community may have originated from a relatively small number of individuals), would have led to strong selection favoring genes believed to contribute to high IQ. According to [42], this selection process is the cause of the high percentage of neurological disorders in this community. Whatever the merits of this bold, although tentative, hypothesis, the frequency of these neurological disorders among Ashkenazi Jews is well-known in medicine, like

congenital hip dislocation in Bretons (inhabitants of the most Westernmost part of France). Ashkenazi Jews also show high rates of cystic fibrosis [43]. The Ashkenazi Jewish community, which is genetically quite homogeneous, has been at the forefront of personalized medicine, and has promoted a very efficient screening for Mendelian diseases and breast cancer [44].

The prevalence of certain Mendelian diseases among Ashkenazi Jews illustrates that broad geographic categories like Europeans, Africans, and Asians often lack the precision needed for meaningful disease studies [45]. Similarly, the Fulani people in Western Africa demonstrate a heightened resistance to malaria compared to other ethnic groups living in the same areas [46]. These examples highlight the importance of considering more specific population groups when investigating disease susceptibility and resistance.

3.3. "Complex" Diseases

Complex diseases, also called "common diseases", are more frequent than Mendelian disorders. They result from the action of numerous genes, each having a small individual effect. Moreover, the environment (lifestyle, diet, socioeconomic factors, education) often plays a considerable role in the onset of these diseases. Modern genomics makes it possible to distinguish between environmental and genetic factors underlying complex diseases. Type 2 diabetes and cardiovascular diseases, for which lifestyle is a key factor, are typical examples of such conditions. Autoimmune and psychiatric disorders, which appear to share certain genetic causes, are other examples.

Like Mendelian diseases, complex diseases have an uneven distribution across the world, with varying impacts across major geographic groups [47]. This is probably due to the fact that their heritability mainly is caused by low frequency and rare variants, while common variants play a more limited role [14]. As previously mentioned, these rare variants also tend to be specific to certain geographical regions. Here are a few examples illustrating the uneven repartition of complex diseases.

3.3.1. Type 2 Diabetes

Type 2 diabetes is different from type 1 diabetes (insulin-dependent), the latter being an autoimmune disease that mainly affects children. Type 2 diabetes, in contrast, has a polygenic background and is considerably impacted by lifestyle. Type 2 diabetes risk is highest among Africans, followed by Europeans, Middle Easterners, and finally Asians [48]. The origin of these differences is debated. The "thrifty gene hypothesis" [49] constitutes a classical explanation. According to this hypothesis, our hunter-gatherer ancestors, frequently facing hunger between random successful hunts, developed a high ability to store fat—a vital asset at the time that ensured survival during periods of scarcity, but problematic today in a world of fast food and food abundance.

Martini and Davis have criticized this hypothesis, arguing that hunter-gatherer ancestors did not experience famine [42]. It is worth noting, indeed, that contemporary hunter-gatherers are not undernourished. However, they do not face the harsh climatic conditions that ancient Europeans underwent during glacial periods. All these explanations remain quite speculative. Martini and Davis point out that type 2 diabetes rates are 2.5 times higher among Navajo Indians and four times higher among Australian Aborigines compared to Europeans [42]. This difference might be explained by an incomplete adaptation of these populations to an agricultural diet poorer in proteins and richer in carbohydrates than the usual hunter-gatherer food.

Duncan et al. noticed that among the Pima Indians, a higher proportion of European genetic ancestry, as evaluated through self-reported ancestry, is correlated with a greater protection against type 2 diabetes [18].

Whatever the precise genetic causes, the frequency of type 2 diabetes clearly is different among geographical populations.

3.2.2. Hypertension

As with type 2 diabetes, the repartition of hypertension varies among geographical populations. African Americans are more affected by this condition than Americans of European ancestry. It is obvious that lifestyle influences this pattern, as it differs considerably between these two communities. However, even when accounting for the impact of the environment, genetic factors remain important [50]. Large differences in the frequencies of blood pressure–related genes have been observed between Americans of European and African ancestries [51]. The "slavery hypothesis" has been proposed to account for the higher prevalence of hypertension among African Americans [52]. According to this hypothesis, individuals who were able to retain sodium more efficiently may have been positively selected during the long voyage from Africa, enabling them to better survive the adverse conditions of deportation by reducing vomiting and diarrhea. This higher sodium retention could now explain the higher prevalence of hypertension in African Americans.

3.3.3. Asthma

Respiratory diseases clearly illustrate the importance of considering patients' ancestry during medical examinations. Physicians use a device called a spirometer to evaluate lung function. Contrary to what [53] claim, lung capacity differs between geographical populations. Values considered normal for one geograpbical population may be considered as pathological for another [54].

Gonzaga-Jauregui et al. observed that among Americans of Mexican ancestry, a greater level of European genetic contribution provokes a greater severity of asthmatic conditions [21]. According to [55], genetic ancestry has a higher impact than environment on asthma prevalence. They suggested that populations originating from tropical areas living in northern countries are more affected by asthma, as is the case for African Americans.

3.4. Cancer

Cancer—or better said, cancers—represent a specific case of complex disease, where the environment has a considerable impact, to varying degrees depending on the type considered, and where geographic differences in distribution are also observed.

3.4.1. Melanoma

According to [56], the frequency of cutaneous melanoma (an extremely malignant tumor) among mixed-race "Latinos" is proportional to their percentage of European ancestry. Melanoma frequency is much lower among Africans than Europeans, suggesting climatic adaptation (to higher sunlight exposure) in Africans [33].

3.4.2. Breast Cancer

Mersha and Abebe observed that three mutations of the BRCA1 and BRCA2 genes are highly frequent among Ashkenazi Jewish women, conferring a predisposition to breast and ovarian cancers [44]. The presence of this genetic trait opens the door to early screening and possible preventive removal, an illustrative case of personalized medicine, discipline in which Israeli medicine has been a pioneer.

According to [57], African American women have a 40% higher incidence of breast cancer than American women of European ancestry, under equal socioeconomic conditions.

According to [55], the UGT2BA gene is at the origin of a higher risk of breast cancer among Nigerian women and, to a lesser extent, African American women.

Reid et al. observed a distribution difference in breast cancer predisposition genes between European and Asian women [58].

3.4.3. Colon Cancer

Maglo et al. noticed that colon cancer incidence is 20% higher, and mortality 40% higher, among African Americans compared to Americans of European ancestry [38]. These authors hypothesized an important genetic component for colon, endometrial, ovarian, prostate and pancreatic cancers.

3.4.4. Prostate Cancer

Risch et al. performed a genome-wide association study (GWAS) on a broad sample comprising 3630 Asians, 58,236 Africans, 211,342 Europeans, 58,236 Africans, and 23,546 "Hispanics" [59]. The polygenic risk scores (PRS) for prostate-specific antigen (PSA), used as a screening test for prostate cancer, was different among the four populations studied.

In a multiancestry survey on prostate cancer prevalence that included Asians, Europeans and Africans, [60] found that African ancestry–specific traits include a high tumor mutation burden, a higher percentage of genome alteration, and a greater number of potentially damaging mutations. In their study, prostate cancer was also more frequent among Africans compared to other geographic populations.

Shelton et al. observed a higher incidence of prostate cancer among Africans (identified by self-reported ancestry) [61].

3.5. Transmissible Diseases: A History of Co-Evolution with Two or Three Actors

3.5.1. General Information

Transmissible diseases are another specific case of complex diseases. They are caused by pathogens. These pathogens can be: (i) viruses (for example: dengue, influenza, yellow fever, AIDS, SARS-coronavirus-19); (ii) bacteria (bacterial meningitis, tuberculosis); (iii) parasitic protozoa (malaria, toxoplasmosis, human African trypanosomiasis or sleeping sickness; American trypanosomiasis or Chagas disease); (iv) parasitic worms or helminths (schistosomiasis, onchocerciasis, tapeworm); (v) microscopic yeasts and fungi (candidiasis, cryptococcosis, aspergillosis).

Transmissible diseases offer a model of choice for evolutionists, and in return, evolutionary sciences have considerably contributed to the knowledge and control of these diseases. This was the theme of the scientific journal for which I was editor-in-chief from 2000 to 2021, Infection, Genetics and Evolution (Elsevier; http://www.elsevier.com/locate/meegid (accessed on 10 July 2025)). Transmissible diseases are a case of co-evolution between the host and the pathogen. In the case of vector-borne diseases, that is to say, for which the pathogen is inoculated by a vector, such as malaria or human African trypanosomiasis, the phenomenon of co-evolution involves an additional actor. Vectors are most often insects (for example, mosquitoes of the genus Anopheles in the case of malaria), but they can also be other arthropods (ticks, crustaceans), and mollusks.

One speaks of co-evolution because the host (for example: humans), the pathogen, and the vector have had to adapt to each other over time. If the pathogen kills its host or makes it sick, it is not out of sadism: it is to survive. And in fact, its well-understood interest, to speak in finalistic terms, is that the host remains alive. As for the vector, it does not transmit the pathogen to do it a favor. A mosquito that bites you does not aim to contaminate you with yellow fever. It is simply a female taking a little blood to carry out its oogenesis. Finally, the host has developed a sophisticated array of immunological defenses to try to survive the infection.

According to the so-called “Red Queen” hypothesis [62], whose title is inspired by Lewis Carroll’s novel Through the Looking-Glass, our species is in a constant race against pathogens, and, of course, against vectors. “If you don’t move forward, you fall behind.”

3.5.2. The Considerable Role of the Environment

As is the case for other complex diseases, the development of transmissible diseases is strongly conditioned by the environment. This includes education, customs, nutrition, access to healthcare systems, socio-economic factors, political instability, and wars.

3.5.3. The “Diseases of Poverty”

Malaria, AIDS, and tuberculosis are labeled “diseases of poverty” because these diseases particularly affect low-income countries. Similarly, Chagas disease, which is observed from the southern United States to northern Argentina, is an illustrative example of the influence of socio-economic factors on the impact of a disease. I have been studying this disease for almost 50 years. The causal agent is Trypanosoma cruzi, a parasitic protozoan related to the agent of human African trypanosomiasis. This parasite is transmitted by blood-feeding bugs of the subfamily Triatominae, called vinchucas (a Quechua word) in the Andean countries. These unpleasant insects proliferate in disadvantaged habitats in Latin America. They favor traditional habitats built with “adobe”; a material made of clay simply dried in the sun. These adobe houses are the lot of poor populations in South America. This is why Chagas disease has been described as a “disease of underdevelopment.” It is no coincidence that transmission in houses in the United States is exceptional: North American houses have nothing to do with the Latin American adobe habitats.

In these “diseases of underdevelopment”, or “diseases of poverty”, the influence of the environment, particularly the socio-economic level, is considerable.

3.5.4. Genetic Inequality and Transmissible Diseases

The distribution and importance of pathogen and vector populations differ considerably between geographical areas. This has caused different selection pressures on the human populations concerned. We are genetically unequal in the face of infectious diseases, probably more so than for any other diseases. This is true at the individual level and at the population level. For the same infectious contact, all other environmental conditions being equal, one individual will contract tuberculosis, another will not. And populations of different geographical origins, in the same environment, will not react identically to the same disease.

3.5.5. How Do Humans Genetically Protect Themselves against Pathogens?

Three major diseases illustrate human genetic adaptations to its tiny enemies:

Malaria and sickle cell anemia (drepanocytosis) have already been discussed (see: "Mendelian diseases"). Let us recall that homozygous individuals for the sickle cell trait present a severe disease, while heterozygotes are healthy and partially protected from malaria. This is an illustrative case of balanced selection. Such protection is of course valuable in tropical Africa. Sickle cell anemia is still present in African Americans, even though this population is no longer exposed to malaria.

Sleeping sickness or human African trypanosomiasis is a calamity in intertropical Africa. Africans, unlike Asians and Europeans, have developed partially effective genetic tools to better resist this disease. However, African Americans, who are no longer threatened by this disease, suffer from a higher rate of kidney disease than Americans of European ancestry. This has been attributed to variants of the Apolipoprotein L1 (APOL1) gene, which would provide some protection against human African trypanosomiasis [63]

AIDS, caused by human immunodeficiency viruses 1 and 2 (HIV 1 and 2), is a rather counter-intuitive example of genetic protection against a disease, since Africans seem less resistant to this condition than Europeans. The explanation for this unusual fact lies in a gene deletion called CCR5 Δ32. This genetic trait is much more common in Europeans than in Africans [64]. Heterozygous individuals for this deletion have a slowed progression of the disease, and homozygous individuals are very resistant to HIV infection [65] According to [66], such a genomic variant is involved in the HIV viral load specific to African populations.

3.5.6. Other Transmissible Diseases Unequally Distributed According to Geographic Group

Cytokines are biochemical messengers involved in the immunological defense against infectious agents and in inflammatory processes. Tibayrenc highlighted numerous genetic differences related to cytokine metabolic pathways between a population of Tanzanians and a population from Western Europe [67]. This is undoubtedly due to the fact that Africans and Europeans have long faced radically different pathogen populations.

Tibayrenc found strong ethnic differences in Sudanese patients regarding the asymptomatic/symptomatic patient ratio for visceral leishmaniasis, a parasitic disease with a severe prognosis [68].

According to [69], there are differences in genetic susceptibility to influenza A virus and SARS-COV2 (Covid) between groups of ancestry.

Differences in innate susceptibility to COVID infection between geographical populations have been implicated by some authors [70,71].

According to [72], some of these differences related to COVID infection could be attributed to adaptive archaic Neanderthal introgression, with the genes in question being present in 16% of Europeans, 50% of Indians, and 63% of Bengalis.

It is also quite evident that environmental and socioeconomic conditions, as well as racial discrimination, also play an important role in SARS-COV2 infection [73,74]. This should not lead to an artificial opposition between genetics and environment, as [73–76] do. Analyzing the genetic causes of diseases does not imply neglecting environmental factors. Both should be jointly analyzed.

4. Conclusions

The exposition given in this article is intentionally simplified and abridged. However, it is based on the most recent research findings and relies on the considerable technological advances that have enriched this field of research. Genome sequencing is now routine, and its cost has been reduced 10 million times in 20 years [77]. All the most recent results point in the same direction: the genetic diversity of human populations is finely stratified, even down to very small geographical scales. Abundant migrations have extensively mixed the genes of humanity but have by no means homogenized them. Genetic markers can identify the geographical origin of any individual with extreme accuracy, and also their potential level of admixture. This very resilient picture has broad applications in forensic medicine. Furthermore, we have seen that human genetic diversity is of considerable importance in the prevention, epidemiological monitoring, and treatment of diseases.

The future of this medical field will need to rely on more integrative approaches, taking into account both genetic and environmental factors simultaneously. The rapid progress of modern technologies (proteomic analysis* of gene regulation and epigenetic phenomena, high-throughput sequencing, cloud computing, or mega-computing in general) paves the way for personalized medicine, finely targeted at specific individuals or populations. To achieve this, fields of knowledge that have been separate until now must be de-compartmentalized and opened to each other: evolutionary sciences and biomedical research, biology and human sciences. Our knowledge of human diversity, whether for fundamental science or for medicine, must not be restricted to biology.

This is all the truer given that the environment in the broadest sense (ecology, socio-economic conditions, education) has a fundamental impact on human diversity, alongside biology. As already emphasized, it is erroneous, even sophistical, to consider that taking into account the heritable component of diseases automatically leads [73,78] to minimizing the impact of the environment, what are called the social determinants of health [38]. Both must be taken into account in the genesis of diseases in an indispensable integrated approach. This is what classical medicine includes in the very fruitful concept of background ("terrain"), which encompasses not only genetics but also age, lifestyle, social context, and environment. I have coined the term "socioecology" to summarize all these complex non-genetic factors. This integrated approach is the specific theme of the new scientific journal of which I am the editor-in-chief: Disease biology, genetics and socioecology (https://www.sciltp.com/journals/dbgs/, access 10 July 2025). The present article aims to illustrate this need for an integrated approach in the study of human diseases.

Funding

This research received no external funding.

Data Availability Statement

No raw data. Only bibliographic references.

Conflicts of Interest

Given the role as Editor-in-Chief, Michel Tibayrenc had no involvement in the peer review of this paper and had no access to information regarding its peer-review process. Full responsibility for the editorial process of this paper was delegated to another editor of the journal. The author declares no conflict of interest.

Use of AI and AI-assisted Technologies

No AI tools were utilized for this paper.

Glossary

Allele	Different molecular forms of a given gene.
Alu insertion sequence	Dispersed repeated DNA sequences in the human genome, consisting of roughly 300 bp in approximately 3–5 × 10^5 copies, constituting roughly 5% of the human genome. These sequences are easily transposable. They are specifically cleaved by the restriction enzyme Alu I.
Ancestry informative marker (AIM)	AIMs are the subset of genetic markers that are very different in allele frequencies across the populations of the world. Most polymorphisms are shared among all populations, and for most loci, the most common allele is the same in each population. An ancestry informative marker is a unique combination of genetic markers that occurs mostly in particular founder population sets but may also be found at varying levels across all or some of the populations found in different parts of the world.
Cystic fibrosis	Genetic disease, affecting the glandular epithelia of many organs, particularly the lungs, leading to severe respiratory failure.
Endogamy	A situation in which reproduction occurs within genetically related groups (as opposed to exogamy).
Founder effect	When a small subsample of a larger population settles as a distinct population, its genetic diversity might be only a small part of the genetic diversity of the original population. For example, the present population of French Canadians was founded by only a few thousand French people, most originating from a limited number of French regions. Their original genetic diversity (and the diversity of their last names) was therefore lower than that of the entire French population. For this reason, some genetic traits that are rare in the general French population were by chance oversampled in the original founder population. This explains why some pathological traits (amyotrophic lateral sclerosis) are more frequent in Quebecois than in the present-day French population.
Gene	A DNA sequence coding for a given polypeptide. More broadly: any given DNA sequence.
Genetic distance	Various statistical measures inferred from genetic data, estimating the genetic dissimilarities among individuals or populations.
Genomics	While genetics concerns the study of individual genes, genomics concerns the study of the entire genome.

Haplogroup	Genetic population group that shares a common ancestor with a specific set of genetic markers. Haplogroups are typically identified through DNA variations in the Y-chromosome (passed from father to son) or mitochondrial DNA (mtDNA) (passed from mother to offspring). These groups help trace human ancestry, migration patterns, and evolutionary history. Haplogroups are often labeled with letters (e.g., Haplogroup R, Haplogroup L, Haplogroup U) and may have further subclades indicated by numbers and letters (e.g., R1b, U5a1).
Hereditary hemochromatosis	A genetic disease due to excessive intestinal absorption of iron, resulting in the deposition of this element in various organs such as the skin, heart and liver.
Heritability	Proportion of variation in a trait among individuals in a population that can be attributed to genetic effects.
Huntington's disease	A hereditary neurodegenerative disorder that results in severe impairment of physical and intellectual abilities.
Intron	A gene region that is not translated into a protein sequence.
Isoenzymes	Different electrophoretic variants of a given enzyme reflect genetic variability in the population being studied. These variants, or isoenzymes, migrate differently in an electric field within a medium such as starch gel or agarose. The differences in migration are due to variations in the enzymes' overall electric charge, which is determined by the individual charges of the amino acids that make up the enzyme. Consequently, electrophoretic differences indicate variations in amino acid sequences, which in turn reflect differences in the underlying gene sequences.
Microsatellite	A short DNA sequence, usually 1–4 bp long, that is repeated together in a row along the DNA molecule. In humans, as in many other species, there is great variation from one person to another (widely used in forensic applications for individual identification) and among different populations in the number of repeats. Numbers of repeats for a given locus* define microsatellite alleles*. There are hundreds of places in human DNA and in most other species that contain microsatellites. Microsatellites are fast-evolving markers, with a high resolution level, and are found in many different organisms.
Mitochondria	Membrane-bound organelles found in most eukaryotic cells. They are often referred to as the "powerhouses of the cell" because they generate adenosine triphosphate (ATP), the primary energy currency of the cell, through a process called oxidative phosphorylation. Mitochondria have their own DNA (mtDNA), which is inherited maternally and encodes some of the proteins essential for their function. In addition to energy production, mitochondria play key roles in cell signaling, apoptosis (programmed cell death), and metabolism.
Molecular clock	In its strict, original sense (more correctly called the DNA clock hypothesis), the concept that the rate of nucleotide substitutions in DNA remains constant or at least is a function of time. In a broader sense, simply the fastness of the evolution of the genomic part that codes for the variability of a given genetic marker. This fastness is driven by the rate of substitution/mutation and is influenced by natural selection. It can be regular or irregular.
Mutation	Any change in the structure of a gene, or of any segment of DNA.
Natural selection	Process first described by Charles Darwin that favors certain genotypes to the detriment of others over generations because they are better adapted to survive and have therefore more abundant offspring (they have a higher "fitness"). It is entirely driven by the interaction of an organism' genotype with the environment.
Neutral gene, neutral polymorphism	A gene/genetic polymorphism that does not undergo natural selection.
Phenotype	All observable characteristics of a given individual or a given population distinct from the genome. The phenotype is not limited to morphological characteristics and includes, for example, physiological parameters (blood pressure, muscular strength, etc.) or biochemical parameters (level of cholesterol, etc.). The phenotype is produced by the interaction between genotype* and the environment. The variable part of a given character that is driven by the action of the genotype constitutes the heritability* of the character under study.
Phenylketonuria	A rare genetic disease linked to a deficiency of the enzyme phenylalanine hydroxylase, leading to the accumulation of the amino acid phenylalanine in the brain and blood. It is a disease that can be cured by a diet low in phenylalanine, which allows for normal or at least improved intellectual development.
Phylogenetics	A branch of genetics that aims at reconstructing the evolutionary past and relationships of taxa or of separate evolutionary lines.
Population Genetics	The study of genetic variation across space and time within and among populations. This field focuses on individual genes and emphasizes the population or species as a whole rather than individual organisms. (See also population genomics.)
Population Genomics	The study of genomic variation across space and time within and among populations.

Proteomic analysis	Proteomics refers to the science that studies proteomes, that is, all the proteins in a cell.
Sickle cell disease, sickle cell anemia	A genetic condition common in people of African descent. It is also present in some regions of India. It affects hemoglobin, the protein that carries oxygen in red blood cells.
Single Nucleotide Polymorphism (SNP)	Polymorphisms resulting from single nucleotide variations in the DNA sequence, known as single nucleotide polymorphisms (SNPs), contribute to differences among individuals and populations. While most SNPs have no noticeable effect, others lead to subtle variations in features such as appearance, and some are linked to the risk of certain diseases. SNPs are commonly used as high-resolution population markers and are central to projects like the HapMap (described below). Various methods are employed to detect SNPs, including: (i) restriction fragment length polymorphism (RFLP), (ii) sequencing, (iii) denaturing high-performance liquid chromatography (dHPLC), (iv) mass spectrometry, and (v) array-based resequencing/microarrays. SNPs account for approximately 90% of all human genetic variation, and those with a minor allele frequency of ≥1% are typically found every 100–300 bases throughout the human genome. It is important to note that allelic frequencies vary across different human populations, meaning that a SNP that is common in one geographical or ethnic group may be much rarer in another (see ancestry-informative markers).
Spacer DNA	In eukaryotic and some viral genomes, untranscribed DNA segments that flank functional genetic regions or cistrons.
Synonymous mutation	Mutation that gives change in DNA but no change in protein due to the redundancy of the genetic code.

References

1. Abdellaoui, A.; Yengo, L.; Verweij, K.J.H.; et al. 15 years of GWAS discovery: Realizing the promise. *Am. J. Hum. Genet.* **2023**, *110*, 179–194.
2. Alexandre, L. *ChatGPT va Nous Rendre Immortels*; JCLattès: Paris, France, 2024.
3. Arama, C.; Maiga, B.; Dolo, A.; et al. Ethnic differences in susceptibility to malaria: What have we learned from immuno-epidemiological studies in West Africa? *Acta Trop.* **2015**, *146*, 152–156.
4. Armengol, L.; Villatoro, S.; González, J.R.; et al. Identification of Copy Number Variants Defining Genomic Differences among Major Human Groups. *PLoS ONE* **2009**, *4*, e7230. https://doi.org/10.1371/journal.pone.0007230.
5. Avise, J.C. *Molecular Markers, Natural History and Evolution*, 2nd ed.; Chapman & Hall: London, UK.
6. Barbujani, G.; Colonna, V. Human genome diversity: Frequently asked question. *Trends Genet.* **2010**, *26*, 285–295.
7. Bick, A.G.; The All of Us Research Program Genomics Investigators. Genomic data in the All of Us Research Program. *Nature* **2024**, *627*, 340–346. https://doi.org/10.1038/s41586-023-06957-x.
8. Boahen, C.K.; Temba, G.S.; Kullaya, V.I.; et al. A functional genomics approach in Tanzanian population identifies distinct genetic regulators of cytokine production compared to European population. *Am. J. Hum. Genet.* **2022**, *109*, 471–485.
9. Brothers, K.B.; Bennett, R.L.; Cho, M.K. Taking an antiracist posture in scientific publications in human genetics and genomics. *Genet. Med.* **2021**, *23*, 1004–1007. https://doi.org/10.1038/s41436-021-01109-w.
10. Bucheton, B.; Kheir, M.M.; El-Safi, S.H.; et al. The interplay between environmental and host factors during an outbreak of visceral leishmaniasis in eastern Sudan. *Microbes Infect.* **2002**, *4*, 1449–1457.
11. Budowle, B.; Sajantila, A. Short tandem repeats—How microsatellites became the currency of forensic genetics. *Nat. Rev. Genet*. **2024**, *25*, 450.
12. Byappanahalli, A.M.; Omoniyi, V.; Hooten, N.N.; et al. Extracellular vesicle mitochondrial DNA levels are associated with race and mitochondrial DNA haplogroup. *iScience* **2024**, *27*, 108724.
13. Chapman, S.J.; Hill, A.V.S. Human genetic susceptibility to infectious disease. *Nat. Rev. Genet.* **2012**, *13*, 175–188.
14. Choudhury, A.R.; Singh, K.K. Mitochondrial Determinants of Cancer Health Disparities. In *Seminars in Cancer Biology*; Academic Press: New York, NY, USA, 2017; Volume 47, pp. 125–146.
15. Cochran, G.; Harpending, H. The 10,000 Year Explosion—How Civilization Accelerated Human Evolution. In *Basic Books*; A Member of the Perseus Books Group: New York, NY, USA, 2009.
16. Corona, E.; Chen, R.; Sikora, M.; et al. Analysis of the Genetic Basis of Disease in the Context of Worldwide Human Relationships and Migration. *PLoS Genet.* **2013**, *9*, e1003447. https://doi.org10.1371/journal.pgen.1003447.
17. Dean, M.; Carrington, M.; Winkler, C.; et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. *Science* **1996**, *273*, 1856–1862.
18. Duncan, R.G.; Krishnamoorthy, R.; Harms, U.; et al. The sociopolitical in human genetics education. *Science* **2024**, *383*, 826–828.
19. Fu, W.; O'connor, T.D.; Jun, G.; et al. Analysis of 6515 exomes reveals the recent origin of most human protein-coding variants. *Nature* **2013**, *493*, 216–220.
20. Gentillini, M.; Duflo, B.; Danis, M.; et al. *Médecine Tropicale*, 4th éd.; Médecine-Sciences, Flammarion: Paris, France, 1986.

21. Gonzaga-Jauregui, C.; Lupski, J.R.; Gibbs, R.A. Human Genome Sequencing in Health and Disease. *Annu. Rev. Med.* **2012**, *63*, 35–61.
22. Burchard, E.G.; Ziv, E.; Coyle, N.; et al. The Importance of Race and Ethnic Background in Biomedical Research and Clinical Practice. *New Engl. J. Med.* **2003**, *348*, 1170–1175.
23. Graves, J.L.; Goodman, A.H. *Racism, not Race; Answers to Frequently Asked Questions*; Columbia University Press: New York, NY, USA, 2022.
24. Guha, S.; Rosenfeld, J.A.; Malhotra, A.K.; et al. Implications for health and disease in the genetic signature of the Ashkenazi Jewish population. *Genome Biol.* **2012**, *13*, R2. https://doi.org10.1186/gb-2012-13-1-r2.
25. Gurdasani, D.; Carstensen, T.; Tekola-Ayele, F.; et al. The African Genome Variation Project shapes medical genetics in Africa. *Nature* **2015**, *517*, 327–332.
26. Hardelid, P.; Cortina-Borja, M.; Munro, A.; et al. The Birth Prevalence of PKU in Populations of European, South Asian and Sub-Saharan African Ancestry Living in South East England. *Ann. Hum. Genet.* **2008**, *72*, 65–71.
27. Harris, K. Evidence for recent, population-specific evolution of the human mutation rate. *Proc. Nat. Acad. Sci. USA* **2015**, *112*, 3439–3444.
28. Hindorff, L.A.; Bonham, V.L.; Brody, L.C.; et al. Prioritizing diversity in human genomics research. *Nat. Rev. Genet.* **2018**, *19*, 175–185.
29. Hoffmann, T.J.; Graff, R.E.; Madduri, R.K.; et al. Genome-wide association study of prostate-specific antigen levels in 392,522 men identifies new loci and improves prediction across ancestry groups. *Nat. Genet.* **2025**, *57*, 334–344. https://doi.org/10.1038/s41588-024-02068-z.
30. Jaratlerdsiri, W.; Jiang, J.; Gong, T.; et al. African-specific molecular taxonomy of prostate cancer. *Nature* **2022**, *609*, 552–559.
31. Jia, G.; Ping, J.; Guo, X.; et al. Genome-wide association analyses of breast cancer in women of African ancestry identify new susceptibility loci and improve risk prediction. *Nat. Genet.* **2024**, *56*, 819–826.
32. Jordan, B. *L'humanité au Pluriel La Génétique et la Question des Races*; Éditions du Seuil: Paris, France, 2008.
33. Jorde, L.B.; Wooding, S.P. Genetic variation, classification and 'race'. *Nat. Genet.* **2004**, *11*, S28–S33.
34. Karlsson, E.K.; Kwiatkowski, D.P.; Sabeti, P.C. Natural selection and infectious disease in human populations. *Nat. Rev. Genet.* **2014**, *15*, 379–383.
35. Koch, L. Impact of genetic ancestry on viral infection response. *Nat. Rev. Genet.* **2022**, *23*, 72. https://doi.org/10.1038/s41576-021-00442-9.
36. Lewontin, R.C. The apportionment of human diversity. *Evol. Biol.* **1972**, *6*, 381–398.
37. Karlsson, E.K.; Kwiatkowski, D.P.; Sabeti, P.C. Association between mitochondrial genetic variation and breast cancer risk: The Multiethnic Cohort. *PLoS ONE* **2019**, *14*, e0222284. https://doi.org/10.1371/journal.pone.0222284.
38. Maglo, K.N.; Mersha, T.B.; Martin, L.J. Population Genomics and the Statistical Values of Race: An Interdisciplinary Perspective on the Biological Classification of Human Populations and Implications for Clinical Genetic Epidemiological Research. *Front. Genet.* **2016**, *7*, 22. https://doi.org/10.3389/fgene.2016.00022.
39. Manry, J.; Quintana-Murci, L. A Genome-Wide Perspective of Human Diversity and Its Implications in Infectious Disease. Cold Spring Harb. *Perspect. Med.* **2013**, *3*, a012450.
40. Martin, A.R.; Gignoux, C.R.; Walters, R.K.; et al. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* **2017**, *100*, 635–649.
41. Marouli, E.; Graff, M.; Medina-Gomez, C.; et al. Rare and low-frequency coding variants alter human adult height. *Nature* **2017**, *542*, 186–193.
42. Martini, R.; Davis, M.B. The DARC side of genetics in cancer: Breast cancer disparities. *Am. J. Hum. Genet.* **2024**, *111*, 1261–1264.
43. McLaren, P.J.; Porreca, I.; Iaconis, G.; et al. Africa-specific human genetic variation near CHD1L associates with HIV-1 load. *Nature* **2023**, *620*, 1025–1030.
44. Mersha, T.B.; Abebe, T. Self-reported race/ethnicity in the age of genomic research: Its potential impact on understanding health disparities. *Hum. Genom.* **2015**, *9*, 1. https://doi.org/10.1186/s40246-014-0023-x.
45. Moreno-Estrada, A.; Gignoux, C.R.; Fernández-López, J.C.; et al. The genetics of Mexico recapitulates Native American substructure and affects biomedical traits. *Science* **2014**, *344*, 1280–1285.
46. Mountain, J.L.; Risch, N. Assessing genetic contribution to phenotypic differences among "racial" and "ethnic" groups. *Nat. Genet.* **2004**, *36*, S48–S53.
47. Neel, J.V. Diabetes Mellitus: A 'Thrifty' Genotype Rendered Detrimental by 'Progress'? *Am. J. Hum. Genet.* **1962**, *14*, 353–362.
48. Nei, M. The theory of genetic distance and evolution of human races. *JPN. J. Hum. Genet.* **1978**, *23*, 341–369.
49. Nelson, M.R.; Wegmann, D.; Ehm, M.G.; et al. An Abundance of Rare Functional Variants in 202 Drug Target Genes Sequenced in 14,002 People. *Science* **2012**, *337*, 100–104.
50. Nichols, R. Gene trees and species trees are not the same. *Trends Ecol. Evol.* **2001**, *16*, 358–364.

51. Niedzwiedz, C.L.; O'Donnell, C.A.; Jani, B.D.; et al. Ethnic and socioeconomic differences in SARS-CoV-2 infection: Prospective cohort study using UK Biobank. *BMC Med.* **2020**, *18*, 160.
52. Omiye, J.A.; Lester, J.C.; Spichak, S.; et al. Large language models propagate race-based medicine. *NPJ Digit. Med.***2023**, *6*, 195. https://doi.org/10.1038/s41746-023-00939-z.
53. Ostrer, H.; Skorecki, K. The population genetics of the Jewish people. *Hum. Genet.* **2013**, *132*, 119–127.
54. Parens, E. The Inflated Promise of Genomic Medicine. Available online: https://www.scientificamerican.com/blog/observations/the-inflated-promise-of-genomic-medicine/ (access on 10 July 2025).
55. Pfaff, C.L.; Parra, E.J.; Bonilla, C.; et al. Population Structure in Admixed Populations: Effect of Admixture Dynamics on the Pattern of Linkage Disequilibrium. *Amer. J. Hum. Genet.* **2001**, *68*, 198–207.
56. Quintana-Murci, L. *Le Peuple des Humains*; Odile Jacob: Paris, France, 2021.
57. Reardon, S. Alzheimer's drug trials plagued by lack of racial diversity. *Nature* **2023**, *620*, 256–257.
58. Reid, D.M.; Barber, R.C.; Thorpe, R.J.; et al. Mitochondrial DNA oxidative mutations are elevated in Mexican American women potentially implicating Alzheimer's disease. *NPJ Aging* **2022**, *8*, 2. https://doi.org/10.1038/s41514-022-00082-1.
59. Risch, N.; Burchard, E.; Ziv, E.; et al. Categorization of humans in biomedical research: Genes, race and disease. *Genome Biol.* **2002**, *3*. https://doi.org/10.1186/gb-2002-3-7-comment2007.
60. Roberts, D.E. The art of medicine. Abolish race correction. *Lancet* **2021**, *397*, 17–18.
61. Shelton, J.F.; Shastri, A.J.; Ye, C.; et al. Trans-ancestry analysis reveals genetic and nongenetic associations with COVID-19 susceptibility and severity. *Nat. Genet.* **2021**, *53*, 801–808.
62. Shi, Y.; Niu, Y.; Zhang, P.; et al. Characterization of genome-wide STR variation in 6487 human genomes. Nat. *Commun.* **2023**, *14*, 2092. https://doi.org/10.1038/s41467-023-37690-8.
63. Shukla, P.; Singh, K.K. Uncovering Mitochondrial Determinants of Racial Disparities in Ovarian Cancer. *Trends Cancer* **2021**, *7*, 93–97.
64. Sudmant, P.H.; Rausch, T.; Gardner, E.J.; et al. An integrated map of structural variation in 2504 human genomes. *Nature* **2015**, *526*, 75–81.
65. Tennessen, J.A.; Bigham, A.W.; O'connor, T.D.; et al. Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes. *Science* **2012**, *337*, 64–69.
66. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **2015**, *526*, 68–74.
67. Tibayrenc, M. The Impact of Human Genetic Diversity in the Transmission and Severity of Infectious Diseases. In *Infectious Disease: Host-Pathogen Evolution*, Dronamraju, K., Ed.; Cambridge University Press: Cambridge, UK, 2004; pp. 315–324.
68. Tibayrenc, M. The Race/Ethnic Debate: An Outsider's View. In *On Human Nature: Biology, Psychology, Ethics, Politics, and Religion*; Tibayrenc, M., Ayala, F.J., Eds.; Elsevier/Academic Press: San Diego, CA, USA, 2017.
69. Tishkoff, S.A.; Kidd, K.K. Implications of biogeography of human populations for "race" and medicine. *Nat. Genet.* **2004**, *36*, S21–S27.
70. Vallois, H.V. *Les Races Humaines*; Presses Universitaires de France: Paris, France, 1976.
71. Van Valen, L. A New Evolutionary Law (1973). In *Foundations of Macroecology*; University of Chicago Press: Chicago, IL, USA, 2014; pp. 284–314.
72. Vasseur, E.; Quintana-Murci, L. The impact of natural selection on health and disease: Uses of the population genetics approach in humans. *Evol. Appl.* **2013**, *6*, 596–607.
73. Verma, A.; Huffman, J.E.; Rodriguez, A.; et al. Diversity and scale: Genetic architecture of 2068 traits in the VA Million Veteran Program. *Science* **2024**, *385*, eadj1182.
74. Wang, Y.; He, Y.; Shi, Y.; et al. Aspiring toward equitable benefits from genomic advances to individuals of ancestrally diverse backgrounds. *Am. J. Hum. Genet.* **2024**, *111*, 809–824.
75. Winther, R.G. A Beginner's Guide to the New Population Genomics of Homo sapiens: Origins, Race, and Medicine. *Harv. Rev. Philos*. **2018**, *XXV*, 1–18.
76. Wright, S. The genetical structure of populations. *Ann. Eugen.* **1951**, *15*, 323–354.
77. Xie, N. Building a catalogue of short tandem repeats in diverse populations. *Nat. Rev. Genet.* **2024**, *25*, 457.
78. Yudell, M.; Roberts, D.; DeSalle, R.; et al. NIH must confront the use of race in science. *Science* **2020**, *369*, 1313–1314.