Mapping Genomes through History[1]

Using Global Ancestry in a Contemporary Population Genetics Approach to Dengue

Ricardo Gomes Moreira

Abstract *In this chapter I examine the operational value of the concept of ancestry for a group of population geneticists from a Portuguese laboratory working closely with a Cuban team of virologists. Developed as an ethnomethodological inquiry about scientific concepts sustaining the work of these geneticists, it attempts to understand the rationale behind the use of broad classification categories for analyzing the genetic structure of populations and identifying potential genotype-phenotype associations of biomedical value. Arguing that there is an evolutionary dimension to the concept of ancestry that is consistently present in geneticists' naming practices, I propose that it is precisely this link of ancestry with evolutionary theory that enables geneticists to follow the historical trajectories of populations in order to gain an understanding of the selective processes that may underpin the role of specific genetic traits in disease and its clinical outcomes.*

Introduction

With the rise of biocomputational power and new automated (high-throughput) DNA sequencing technologies, the notion of ancestry has in recent decades become a central concept in human genetics. Particularly useful for population geneticists as a category for addressing biological diversity, "ancestry" enables laboratory practitioners to distinguish between genetic profiles of individuals and populations. This study describes and analyzes the practical and symbolic dimensions of this concept in a medical genetics context focused on dengue research, grounding the analysis in ethnographic laboratory fieldwork and qualitative analysis of biomedical and population genetics literature.

1 The research and writing of this chapter were funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, grant agreement No. AdG 695573, "The Colour of Labour," coordinated by Cristiana Bastos; and by National Funds through the Portuguese funding agency FCT, project "GeneAntropMag" (CIRCNA/CIS/0142/2019), coordinated by Luísa Pereira.

Aiming to clarify how geneticists understand "ancestry" as both a methodological and interpretative tool to account for certain dimensions of human diversity, it seeks a critical understanding of the logics and rationale underlying its use.

In recent decades, the idea of biogeographic ancestry has consolidated in population genetics as a core conceptual tool linking the historical and geographical origins of one's ancestors. It has become particularly useful for tracing remote forebears or mapping ancient migratory routes beyond memory. Several STS studies have examined how this concept is applied in geneticists' work. Some scholars warn of the risks of using "ancestry" in biomedical and genetics contexts in ways that may be interpreted as racial.[2] Others show that its conceptual flexibility and variable definition depend on methodological configurations and research designs, and that its referents and categorical distinctions hinge on related concepts such as "population" (Fujimura and Rajagopalan 2011; Panofsky and Bliss 2017; Lipphardt 2017; Byeon et al. 2021). Fujimura and Rajagopalan (2011) have persuasively demonstrated the analytical value of an ethnographic approach that considers this concept from the perspectives of scientists engaged in laboratory work. Examining how classificatory concepts such as ancestry operate in laboratory knowledge production reveals their contingent nature, shaped by technology, methods, research design, and—not least—theory.

Some STS scholars engaged in sustained critical examination of genetically applied notions of ancestry have recently produced scholarship focused on the North American context. They highlight the entwining of genetic ancestry classificatory concepts with US sociopolitical notions of race and their propensity to slip from laboratory into identitarian and politically charged contexts, where dynamics of power and relatedness may be intentionally or unintentionally reinforced by scientific practice (e.g., Nash 2015; Nelson 2016; Fullwiley 2024).

In this chapter, I propose to explore how the concept of biogeographic ancestry is used in a population genetics laboratory according to its practical applicability to the scientific problem of dengue, its objects of study, and various aspects of scientific work. My analysis draws on ethnographic methodology, consisting mainly of interviews with geneticists, direct observations, documentation of their views and commentaries on their own work, and analysis of publications by a population genetics team from Portugal. An ethnomethodological approach was adopted to yield insights into both the rationale for using a classificatory form of genetic ancestry and its possibilities across different stages of research: from planning, through collection building and statistical analysis, to comparison and writing of results.

The chapter examines conceptual and practical configurations of genetic ancestry, understood as procedurally formed by two mutually constitutive aspects—molecular and sociopolitical—combined to produce a scientific concept of analytical value and identitarian character.

The first section describes the configuration of a medical genetics project on dengue, in which this concept of ancestry was modeled and applied to name specific popula-

2 Cogent discussions on this topic include Fullwiley (2008, 2014), Gannett (2014), Wade et al. (2014), and Lewis et al. (2022). An earlier, important contribution to the debate was the edited volume by Koenig, Soo-Jin Lee, and Richardson (2008).

tion groups. The second explores how ancestry became useful as a research tool, associated with traits of immunological resistance and susceptibility, and gained significance through translation from epidemiological history and virology into population genetics.[3] The third identifies the importance of ancestry in linking laboratory research to evolutionary theory, enabling theoretical interpretation of results and connecting analyses to other projects that may lead to broader generalizations and hypothesis formulation.

Cuba Meets Europe: The Making of a Medical Genetics Approach to Dengue

Dengue virus (DENV) is recognized by international organizations as one of the most significant arthropod-borne diseases and a major global public health concern, affecting nearly all tropical regions. In the 2000s, about 2.5 billion people in over one hundred countries were considered at risk, while incidence patterns shifted as the disease became endemic in several regions and rates rose alarmingly (Guzmán et al. 2010, 57; see also Jaenisch et al. 2013).

In 2012, the European Commission funded a major dengue research project, organized as three independent consortia involving multiple scientific institutions. One consortium, DENFREE (Dengue Research Framework for Resisting Epidemics in Europe), aimed to provide "a new strategy for surveillance," "better control of DENV transmission," and determine epidemic factors and trajectories, as well as the underlying factors of "pathogenesis with respect to viral and host factors that can predict disease severity and prepare for further development of new vaccines [and] antiviral compounds" (Jaenisch et al. 2013, 2). In short, DENFREE sought to explore the genetic basis of resistance and susceptibility to dengue and to identify genetic configurations causal to the immunological determinants of infection outcome (Jaenisch et al. 2013, 2).

Led by a French team from the Institute Pasteur, the consortium included the Pedro Kourí Institute of Tropical Medicine (PKI) in Cuba and a team of population geneticists from IPATIMUP, Portugal, invited for their expertise in applying populational methods to medical genetics. Their task was to investigate possible genetic causes for different immunological responses observed in Cuban epidemiological records, which had been previously studied by Cuban biomedical scientists and virologists.

Cuba was selected as a case study partly because of its epidemiological history. Dengue has long been present in the Americas and Caribbean, with reports as early as 1699 in Panama (Guzmán 2012). The year 1827 marked the start of regular epidemic records. In Cuba, outbreaks were reported in 1906 (Sierra et al. 2017) and 1944–45, but the large-scale epidemic of 1977—with more than 400,000 cases reported—is regarded as the beginning of the modern presence of DENV on the island (Guzmán 2012).

Over the last five decades, Cuba has built substantial clinical, epidemiological, and virological knowledge on dengue, most of it developed at the PKI in Havana. Following the epidemics of 1981 and 1997, Cuban researchers documented distinctive aspects of dengue incidence, promoting the country as a relevant case study (e.g., Guzmán 2012).

3 I use the notion of translation in the sense proposed by Michel Callon and John Law, as part of the scientific work of connecting different spheres of interest (Callon 1984; Callon and Law 1982).

Internationally, they highlighted successful public health interventions that prevented endemic dengue, eliminated transmission chains, and disseminated research results through organizations such as the Pan American Health Organization, the Special Programme for Research and Training in Tropical Diseases, and the Dengue Vaccine Initiative (e.g., Guzmán et al. 1999; Guzmán 2012).

Some of these results were especially relevant to population geneticists, notably factors and patterns of risk regarding the Cuban population. In 1999, in the *WHO Dengue Bulletin*, Maria Guzmán and her team put forward what they called "some interesting observations" about the dengue hemorrhagic fever (DHF) epidemics of 1981 and 1997 (Guzmán et al. 1999), arguing that the "Cuban experience is probably unique." Among these were empirical demonstrations of population behavior under epidemic contexts, the crucial role of "secondary infections" in disease outcome, and the relevance of "race" as an indicator of potential severity. Based on epidemiological data, they hypothesized that "black people were at a reduced risk to manifestations of [dengue shock syndrome (DSS)] as compared to white people,"[4] an argument traceable to early twentieth-century outbreak reports.

Cuban virologists further explored these observations in a 2007 paper in the *Archives of Virology*, specifically focusing "race" as "a risk factor for dengue hemorrhagic fever." They state that:

> Cuban DHF/ DSS outbreaks have provided evidence of a reduced risk of people of Negroid race for DHF/ DSS compared to those of Caucasoid race. These observations from Cuban dengue outbreaks have significant epidemiological interest, as the differences in susceptibility to DHF/ DSS among racial groups in Cuba coincide with that reported in African and Black Caribbean populations. (Sierra et al. 2007)

As we will see, these conjectures became the touchstone of geneticists' approach to the DENFREE scientific problem. It is precisely the translation of these epidemiological observations into a population genetics framework that this chapter will explore. The fact that these observations, concerning population groups with different "racial" profiles, hold valuable hints for contemporary geneticists raises considerable questions: Why would differential skin pigmentation or phenotypic constitutions, considered in racial terms like "Caucasoid" or "Negroid," serve as an empirical basis for population genetics? What are geneticists seeking in associating "racial" profiles with differences in clinical outcomes and dengue severity?

Such questions, which require understanding how population genetics practices are applied in biomedical research, speak to the epidemiological and medical approach of the DENFREE consortium and become relevant within the European dengue research agenda. The Portuguese team's background shows important research lines in demographic history and medical genetics. Since the 2000s, this lab has extensively used mitochondrial DNA and Y-chromosome analyses to study genetic structure and variability in Portuguese and Iberian contexts, as well as in African Portuguese-speaking countries,

4 DSS is the most severe, life-threatening form of dengue, which follows as a series of complications of Dengue hemorrhagic fever.

producing narratives on demographic history, migratory routes, and past population encounters across continental regions. In the following decade, whole-genome technologies expanded the practical scope of genomic analysis. Genome-wide association studies (GWAS), for instance, were devised to identify DNA regions linked to phenotypic variations. The lab's publishing history also shows methodological strength in analyzing genetic structures in diverse contexts. Since the early 2010s, several projects targeted highly diverse populations, consolidating genetic ancestry as a central concept.

The integration of IPATIMUP's population genetics laboratory into DENFREE fit the scientific interests of both sides. The team's expertise in population structure met the needs of the consortium, while the research problem—rooted in epidemiological and virological observations at the population scale—aligned with the geneticists' preferred approaches. Another factor in the success of this partnership lies in the technological and methodological capacity of contemporary genomics, which can statistically link genotypical and phenotypical dimensions at population scales. With computational and statistical tools, geneticists can propose relations between specific DNA loci and biological traits selected during sample collection. In this sense, Cuban epidemiologists and virologists provided the starting conditions for a research design in which population geneticists could examine Cuban epidemiological and virological findings genetically.

In sum, the DENFREE project analyzed here was built on close collaboration between a Cuban team of virologists and a Portuguese team of population geneticists, jointly seeking answers to scientific problems posed by dengue.

Genetic Ancestry: Reframing Natural History in Dengue Research

As argued above, approaching dengue as a study of its "pathogenesis with respect to viral and host factors," in order to "predict disease severity," and aiming for the "development of new vaccines [and] antiviral compounds," meant directing research toward understanding genetic causes of different immunological responses, specifically the resistance and susceptibility patterns reported in Cuban epidemiological records and studied by PKI virologists.

This research did not begin with DENFREE, however. Over the preceding three decades, Cuban scientists had produced extensive work, summarized in a 2012 article listing major results since the first epidemic of 1981 (Guzmán 2012). These included populational and genetic configurations relevant to DENFREE's design, such as the observation that DHF risk factors included "white skin color," or that "higher association with DHF was found for persons of European descent, with stronger memory T-cell response and cross-reactivity in white compared with black dengue-immune persons." Other results identified biomarkers and immunologically relevant genes (ibid.).

As argued above, Cuba's selection as a case study was due less to public health impact (other serious candidates for that matter existed) than to the prior biomedical work of its physicians and virologists. The phenotypic diversity of the Cuban population, it can well be argued, was equally crucial. This diversity, long part of Cuban clinical and epidemiological records, was described in DENFREE as "admixed composition" (Oliveira et al. 2018). Geneticists considered such "admixture" advantageous, as they understood

it to be directly translatable into a genetic structure that could be empirically detected and then analyzed; this meant, in other words, that different populations with distinct genealogical histories and genetic profiles could be discerned within the contemporary Cuban population. Within DENFREE, an analytical framework was thus devised to relate population genetic structure (and ancestry categories) to observed clinical responses.

Thus, the DENFREE geneticists, using the classificatory concept of genetic ancestry, approached the problem as one of "ancestry-conferred DHF susceptibility/protection."[5] As they write: "admixed populations are a great advantage in cases of differential ancestry-conferred susceptibility/resistance to a disease through the use of admixture mapping" (Sierra et al. 2017, 3)—a methodological position that will be further considered here.

In molecular anthropology and human population genetics, genetic ancestry is often applied to statistically detect patterns of human biogenetic diversity, grounded in hereditary transmission and ascendancy in comparisons between population groups.[6] Despite being a concept developed and applied within the field of population genetics, the notion thus combines historical, sociopolitical, and molecular dimensions (Fujimura et al. 2010; Rajagopalan and Fujimura 2012).

I should note, however, that this dengue study illustrates the contingent character of genetic ancestry as it adapts to specific research goals. Unlike studies explicitly comparing groups defined along racial,[7] ethnic,[8] or national lines,[9] in which the idea of ancestry is thus superimposed on such sociopolitical labels in ways that "biologize" race or other identities (El-Haj 2007; Fullwiley 2007, 2008; Lipphardt 2017), DENFREE examined groups defined by clinical outcomes and probable hereditary variations explaining those outcomes. Here, researchers employed a notion of ancestry that linked "genes" to hypothetical past origins inferred from physicians' observations in order to identify and explain the relevance of these hereditary traits. Race was not the object of study but a shortcut toward the evolutionary theory of population genetics.

The link between ancestry categories and genetic material, traced through a populational approach, serves as a connection between specific ecogeographic environments and those molecular configurations that—because they constitute cases of variation—were likely positively selected over time under the influence of such environments. In seeking to understand the research process and its methodologies, I argue that, from the geneticists' perspective, identifying populations through ancestry categories is necessary to associate particular geographic and ecological origins with specific genetic features and to understand why such configurations have been transmitted and preserved across generations. Following the geneticists' theoretical position, this intermediate step—between identifying molecular (DNA) configurations related to clinical

5 See Sierra et al. (2017, 3). DHF and DSS are used as abbreviations for dengue hemorrhagic fever and dengue shock syndrome, respectively.

6 For an informative and detailed historical account about how these genetic variations have been studied by population geneticists and their sociotechnological aspects, see the study by Rajagopalan and Fujimura (2018).

7 See Gannett (2014) for a critique of the concept of global biogeographic ancestry.

8 See Lipphardt et al. (2021) for a critique of studies on vulnerable populations.

9 See, e.g., Novembre 2008.

outcomes and understanding the biological processes behind the genotype–phenotype nexus—makes the application of ancestry labels productive. These labels allow for the tracing of population histories and the formulation of evolutionary explanations for the causes, effects, and probable molecular pathways of certain genetic configurations and their phenotypic consequences. This rationale for ancestry categories, as a foundation for evolutionary explanation, is exemplified by the dengue research project discussed above.

In this case, Cuban samples were collected "during the 2006 dengue outbreak" and included "67 subjects from Havana … and 70 subjects from Guantanamo" with clinical symptoms, plus 32 asymptomatic individuals from Havana and 16 from Guantanamo, for a cohort of 185 individuals infected with one of the dengue virus strains (Sierra et al. 2017, 14). These samples, from hospital institutions, were classified according to the reported clinical outcomes of patients, which in this project is the fundamental information for the collection, with the additional stipulation that *all* samples came from patients infected with DENV. The primacy of clinical outcomes as the first classificatory criterion, and the most fundamental aspect in forming the object of study, is reflected in Figure 1, where genetic diversity of samples classified by national labels and analyzed with ancestry components is shown next to groups classified by dengue outcome severity.

In addition to DENV patient samples, the historical and scientific background of dengue epidemiological and medical research in Cuba provided DENFREE with essential empirical foundations. On that basis, geneticists analyzed variants and evaluated the relative weight of two ancestries—African and European—in the target Cuban population's genetic structure. Drawing on historical medical sources and Cuban research, they tested the hypothesis that genetic differences might exist between groups with different proportions of these ancestries, and that these differences might influence susceptibility and resistance to dengue. The hypothesis to be tested was made clear by quoting an early medical report:

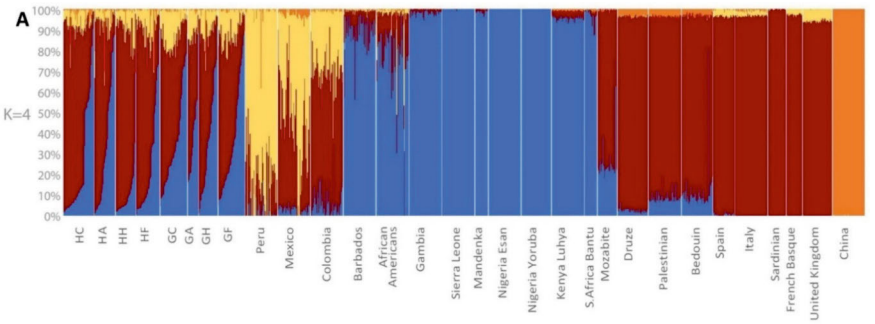
> As early as in 1906, it was reported that Cuban *dark-skinned* individuals showed a remarkable resistance against dengue disease compared with *light-skinned* individuals. This early observation was confirmed during the 1981 Cuban DHF/DSS epidemic of DENV-2 when *ethnicity* was recognized for the first time as a possible host risk factor and confirmed afterwards in several other dengue Cuban outbreaks. The low occurrence of dengue disease in *Haitians* and in *African populations* adds further support for this *ancestry* influence. (Sierra et al. 2017, 2–3, emphasis added)

This historical account and its implications for research design appear to have posed a terminological challenge to researchers, who in their published findings sometimes slip from the notion of ancestry toward ethnicity, skin color, and population designations. This is likely because historical records refer to what would be considered racial groups among Cuban populations in the early twentieth century. It is not clear why the notion of ethnicity emerged as an acceptable choice for geneticists, nor whether it was meant to identify broadly the physical characteristics or the genomically determined ancestry

of Cubans that earlier scientific works would have described in racial terms.[10] However, as another DENFREE publication shows, the idea of "genetic diversity" guiding the research design is entirely at odds with the notion of ethnicity. In published articles, when presenting and discussing results, the notion of ethnicity yields to an idea of diversity defined explicitly around genetic ancestry, in the form of *global ancestry*.

In essence, global ancestry is conceptualized as a partition of contemporary human populations into four major continental categories of genetic ancestry: African, European, East Asian, and Native American (Figure 1). The question that arises is whether these continental categories diverge fundamentally from the nominally similar racial classifications used by physical anthropologists in past centuries, or whether they represent a continuation of that history of racial classifications.

Figure 1: Representation of four different categories of biogeographic ancestry in different populations. Cuban samples are shown on the left side of the graph under clinical labels corresponding to different forms of dengue severity and sampling locations (Havana: HC, HA, HH; Guantánamo: GC, GA, GH, GF). Each vertical line represents an individual genomic analysis: African ancestry in blue, European in red, Native American in yellow, and East Asian in orange.



Source: Sierra et al. 2017.

I believe these are distinct types of classification. To answer this question, we must first understand the *nature of the "objects" to which these categories refer and the relevance of using such classifications, given the research questions and objects of study*. What do geneticists (as in Sierra et al.) mean by global ancestry, and why have they adopted a fourfold continental geographic division?

A crucial point in geneticists' work is that these categories are not meant to represent real populations but rather segments of DNA; and if they do not represent populations,

10 Terminological uncertainty sometimes arises and both terms are used interchangeably, in DENFREE's published works, as if they were equivalent, which can be confusing: "If an *ancestry* is associated with increased susceptibility, the disease locus will present a significant higher proportion of that ancestry in cases versus controls; when an *ethnicity* is protective, the disease locus will present a significantly lower proportion of that ancestry in cases versus controls" (Oliveira 2018, 3; emphasis added).

perhaps they should not be mistaken for racial labels. One reason to avoid such confusion is that, unlike the twentieth-century anthropological concept of "race," which defined broad population groups with fixed and unmixed hereditary biological traits (essentially a biological equivalent of "subspecies"), global ancestry categories are assigned to certain individual genomic features derived from presumed ancestral populations. In contrast to the physical and visual basis of racial classifications, geneticists present global ancestry probabilistically—as statistically inferred geographic origins for segments of DNA in present-day individuals, based on similarities with analogous segments found in populations from those regions.

While endorsing population-based methods and broad ancestry classifications, one interviewed researcher argues that, in studying genetic configurations—especially those with medical relevance—the use of labels may reflect population identity, but the use of ancestry is justified by its heuristic utility. That is, ancestry categories help differentiate historical trajectories of genes, which is essential for situating genetic configurations in time and geography. In this sense, populational approaches have adopted methodological designs that combine genomic sequencing technologies with historical and sociographic data to generate new biological knowledge with medical value. As one geneticist argued:

> Finally, we are able to say something about the selection effects. In early days we didn't have selection—a big concept that was introduced in biology and after became highly important in genetics; but we didn't have the necessary tools to study selection. Now, the chip gave us the power to study the entire genome in a random fashion. […] I don't think we are already at the level of personalized medicine. I think there is an intermediary step that implies studying genetics with biomedical implications at the level of population groups. (interview with DENFREE geneticist)

These researchers argue that although admixture has often been viewed as a "major confounding factor" (Sierra et al. 2017, 3; see also Price et al. 2006; Fujimura and Rajagopalan 2011), admixed populations may now offer a valuable research opportunity if admixture mapping is used as an auxiliary methodological tool. Rather than relying solely on principal component analysis (PCA) to "correct for stratification in the cohort" and adjust samples' "genotypes and phenotypes by amounts attributable to the ancestry" (Sierra et al. 2017, 2)—a procedure that makes cohort samples more comparable to controls—*ancestry mapping* is employed to exploit stratification itself (i.e., the internal variation of cohort samples), particularly when researchers suspect that phenotype and gene variability will show a strong link with ancestry.

When geneticists refer to the "great advantage" of admixed populations in this type of analysis, they mean that the technique of "admixture mapping" can yield promising results when applied to populations descended from two or more ancestral groups of distinct geographic origins. Given that different medical outcomes reflect phenotypic variations—such as susceptibility or resistance factors linked to genetic constitutions—admixture mapping is considered valuable for associating these phenotypic traits with specific ancestries. It is especially useful when the target populations exhibit a high degree of admixture. In short, admixture mapping is presented as one of the most efficient meth-

ods for identifying which ancestries confer resistance and which confer greater susceptibility to the pathogen. As Sierra at al. (2017, 3) note: "It has been shown that this test is statistically more powerful than traditional GWAS: around 250 samples can provide a 60% power to detect a two-fold risk due to ancestry, compared to the thousands of samples required in GWAS."

However, not long before this project began, admixed populations were often seen as obstacles to genetic analysis, particularly in genome-wide association studies (GWAS).[11] GWAS, used to identify genes of clinical or epidemiological relevance, typically required prior analysis and careful selection of samples so that cohorts and controls shared comparable genetic structures. Without such precautions, DNA regions or loci identified as statistically significant for disease outcomes could reflect structural genetic differences between those two sets of samples rather than meaningful trait variation.

Fujimura and Rajagopalan (2011) emphasized the importance of calibration tools in GWAS to address population substructure associated with different ancestries. This was the function of the software Eigenstrat, which geneticists applied to adjust data for ancestry. The adjustment allowed aggregation of samples sharing the same ancestry—and thus the same "population substructure"—ensuring that GWAS could identify genome regions genuinely associated with disease, rather than those reflecting different ancestry substructures of the sampled populations and controls. The central premise of Eigenstrat's statistical algorithm, as developed by Price et al. (2006) and discussed by Fujimura and Rajagopalan (2011, 12–15), is that distinct ancestry structures correspond to distinct genetic substructures (Price et al. 2006; Fujimura and Rajagopalan 2011, 12–15). Once principal component analysis (PCA) and Eigenstrat remove ancestry-related stratification within population samples, it becomes possible to disregard population labels, ancestry information, or identity categories in direct cohort-control comparisons (Fujimura and Rajagopalan 2011, 16).

But whether retaining or removing ancestry information, the relationship between population genetic structure and ancestry remains central to population genetics. With the growing volume and diversity of reference population data, it has become increasingly feasible to infer individual ancestry through genomic analysis. As shown in ancestry mapping, the transaction between ancestry and genomic information—what Fujimura and Rajagopalan call classificatory "slippage" (2011, 14–16)—moves in both directions: from molecular data to population ancestry, and from population ancestry back to DNA as molecular ancestry. This dynamic has significant epistemic implications for understanding how "genetic ancestry" functions as a classificatory tool.

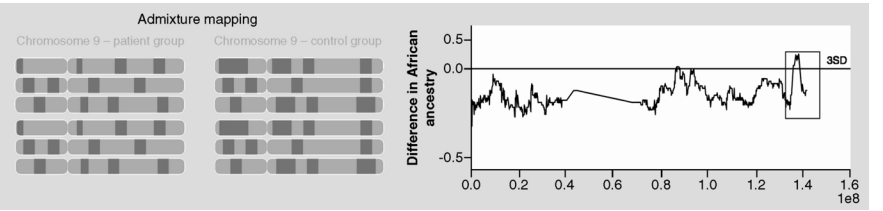
Mapping Ancestry: Tracing the Historical Temporalities and Geographies of Dengue Immunity

A brief general description of how ancestry mapping operates within the framework of populational research helps clarify how ancestries are attributed to specific portions of DNA, producing a chromosomal map. In laboratory practice, "admixture mapping"

11 GWAS is a common technique used in medical genetics to identify loci or genes with potential clinical significance.

is conducted after the standard sequencing of somatic DNA. Once genomic data are obtained for each individual sample, geneticists use computational analysis to identify chromosomal sections based on comparative differences, assigning each section an ancestry label from a predefined set relevant to the research question. These labels are applied using ancestry informative markers (AIMs)—previously identified genetic (molecular) locations that statistically vary across regional populations and are considered significantly more prevalent in certain populations.

Figure 2: Schematic representation of "admixture mapping" in chromosome 9, "comparing patient and control groups." Sections colored in orange (here: dark grey) indicate sections of the chromosome associated with African ancestry. The graphic on the right displays a line graph measuring statistical standard deviation in African ancestry (Y-axis) along chromosome 9 (X-axis), represented as an ordered sequence of hundreds of millions of base pairs. It shows significant deviation values—exceeding 3 standard deviations (3SD)—near the end of the chromosome, around position 1.4, corresponding to base pairs close to the 140,000,000th position.



Source: Oliveira, 2018.

As noted above, in studying the symptomatology of dengue epidemics, geneticists have associated ancestry categories with continental geographic regions. The global ancestry classification system gains relevance from the idea that phenotypic profiles of resistance or susceptibility to dengue reflect evolutionary dynamics—specifically, the selection of genetic traits shaped by the ecological contexts in which populations were, or were not, historically exposed to similar pathogens.

The role of evolutionary processes in selecting protective traits—or in the loss of inherited protection—is not only cited in published studies but was also explicitly addressed by geneticists during interviews. Within this project, such discussions frequently concern genes related to lipid metabolism. One interviewed researcher expressed a particular interest in evolutionary questions, especially in understanding how "African populations developed these variants, related to lipid metabolism, which protects them against infections." She noted that while two genes linked to lipid metabolism may confer dengue protection, many others are likely involved. She thus shifted her focus from studying migration and diversity alone to examining how genetic diversity impacts disease susceptibility.

By identifying chromosomal segments with AIMs corresponding to reference ancestral "origins," researchers can map the distinct ancestral genetic contributions inherited from previous generations, which can aid in comparing the sampled populations. As out-

lined in the DENFREE research design, mapping ancestries along the chromosomes of individuals with resistance or susceptibility to dengue allows geneticists to statistically correlate ancestry configurations in specific genome sections with variations in dengue severity as presented in individual and population records (Figure 2).

In this sense, admixture mapping is a method for extracting relevant information from genetic analysis, based on the idea that "ancestry blocks" can be identified across individual genomes, facilitating the tracking of relevant polymorphisms. These blocks—segments of DNA inherited from ancestral populations—form part of the chromosomal material of contemporary individuals and populations. When searching for a clinical phenotype potentially associated with a particular ancestry, some of these blocks may be statistically more prevalent in groups carrying the variants linked to that phenotype, reflecting higher proportions of that ancestry. If dengue protection were associated with African ancestry, as hypothesized by the Cuban researchers in the consortium, genomic analysis would reveal a higher proportion of African-ancestry DNA blocks in individuals with more favorable clinical outcomes. Protective genes would thus be expected within those same blocks. As Sierra et al. (2017, 3) note: "[A]ncestry blocks will be distributed at random across the genome, reflecting the admixture proportions of the parental ancestries, except in candidate gene locations where statistically significantly different proportions for the ancestry with higher disease levels will be observed in cases versus controls."

The central premise here is that, in order to determine how past populations have contributed differently to the biomolecular composition of present-day humans, geneticists must first conceptualize those ancestral populations in biomolecular terms. This involves constructing categories for ancestral groups and presuming knowledge of their "genetic composition," at least in terms of ancestry informative markers (AIMs) now found in contemporary individual genomes. Using admixture mapping, geneticists claim to identify the contributions of these ancestral populations within individual genomes. In theory, a population's genetic structure is the aggregate of its "composite" (admixed) individual genomes when considered together in a populational set.

A core principle of contemporary population genetics—criticized by STS scholars—is the linking of chromosomal segments to ancestry labels. This entails applying classification systems that carry social or identity categories (such as ethnic, racial, or national labels) into the domain of biological research. Critics argue that such research practices reproduce biological differences grounded in historical and ongoing processes of social differentiation (Duster 2006; El-Haj 2007; Fullwiley 2007, 2014, 2024; Nelson 2008; Reardon 2005; Schramm et al. 2012; TallBear 2013; Wade et al. 2014; Wailoo et al. 2012).

An examination of the methodological dimensions of genetics research in the DENFREE consortium—as described by geneticists in conversations, interviews, or published articles—reveals what may initially appear to be a simple reframing of racial categories (previously used by Cuban researchers) into global ancestries. This shift, however, constitutes a deeper methodological choice, not always made explicit in publications, but one that can be elucidated through the theoretical reasoning underpinning the evolutionary framework of population genetics. The first critical move by geneticists, allowing them to justify the link between genetic data and ancestry labels, is the appropriation of historical narratives that identify and establish the original popula-

tions serving as references for ancestry categories. These narratives—transcending the specifics of individual research projects and established across the discipline as foundational historical frameworks—function as "origin myths" for population genetics.[12] Broadly, they recount innumerable human movements and encounters over centuries or millennia, shaped by different temporal frameworks and addressing specific research questions concerning prehistoric periods, historical epochs, or the present. In these narratives, historical contexts are proposed as setting the stage for major past migration movements, generating encounters and admixtures among ancestral populations, which are then used to explain the formation of later composite generations. Geneticists argue they can trace these events molecularly in present-day populations, as DNA analyses reveal two or more *historically inferred* ancestral genetic contributions within individual samples.

In the context of the contemporary Americas or the New World (to which Cuba belongs), a key historical moment frequently invoked in the analysis of living populations is the onset of the modern age, marked by the early globalization that followed European expansion and the establishment of intercontinental oceanic routes. Seen as one of the historical conditions that led to the emergence of the "modern melting pot," this moment is recurrently memorialized in the oceanic voyages of Columbus or Magellan and finds its historical relevance in the development of oceanic traffic, which inaugurated a new phase in the circulation of people and things across continents and latitudes. This historical moment is frequently employed as a temporal boundary for defining ancestry references, grounded in a distribution of biological (or genetic) diversity attributed to the pre-Columbian world—a proposed 500-year-old distribution that enables the production of a meaningful, globalized analysis of genetic configurations across populations, accounting for evolutionary dynamics.[13]

Historians of science and STS scholars have frequently critiqued this framework. Their central concern is that this historical moment, while serving as a significant reference point for comparative analysis in population genetics, also marks the beginning of European colonialism and its associated violence, including the racialized exploitation of labor. Acknowledging the capacity of historical narratives to reproduce distinct worldviews and sociopolitical interpretations, critics have shown that global approaches

12 Some of these relatively well-known historical narratives include, for example, the "Mitochondrial Eve" theory, which refers to the 50,000-year-old movement of human populations originating in Africa, developed by geneticists in the 1980s (e.g., Cann et al. 1987); the multiple waves of migration from eastern into western Eurasia, known to have occurred throughout the Neolithic period and into the Middle Ages (e.g., Reich 2018, chap. 5); and, particularly relevant for the genetics of contemporary populations, the global population movements that followed the rise of capitalism and its expansion in the sixteenth century. When applied to specific research objectives, these narratives establish the key temporal and spatial contexts in which populations with distinct historical backgrounds encountered one another and produced "admixed" descendant generations.

13 See, e.g., Winkler et al. (2010). Other temporal frameworks in genetic analysis that focus on past populations have, however, enabled the tracking of older, meaningful historical moments of encounter and admixture, as in the case of ancient DNA analysis using biological material extracted from preserved ancient bones (Pääbo 2014).

to human biodiversity can, often inadvertently, perpetuate colonial or racial imaginaries (e.g., Reardon 2005; M'charek 2005).

As was evident to the DENFREE researchers from the beginning, the Cuban samples revealed a combination of two genetic ancestries inherited from African and European origins. Any other possible admixtures that may have occurred earlier in history—before African and European populations came together in the New World—were disregarded in the analysis. For the purposes of this dengue research, such prior encounters and admixtures were deemed either unknown or irrelevant. By taking the rise of the post-Columbian Americas as the "original" moment in the history of the present Cuban population, the European and African populations that mixed in Cuba over the past five centuries were hardly seen as the products of prior ancient admixtures. These populations were conceived as homogenous; earlier admixtures were effectively treated as part of European or African genetic heritage, since the phenotype analyzed for dengue resistance appeared from the beginning to vary distinctly between African and European ancestries, and not within each group.[14]

We should keep in mind that this understanding of global ancestry is explicitly molecular. As noted earlier, with the aim of identifying genes with phenotypic outcomes of medical or clinical relevance, DENFREE geneticists labeled certain DNA segments with global ancestry categories, indicating geographic origin and thus pointing to historical trajectories and circulations of genetic information. To understand global ancestry as used in biology and population genetics, it is crucial to recognize that these origins are attributed to "blocks" of DNA—not to individuals or populations. At this stage of genetic research, what is being classified is not people, but portions of the DNA molecule within individuals.

The process of tracing ancestries in genomes to determine the history of specific molecular configurations should not lead us to conflate molecular "identities" with the social or historical identities of people. Ancestry labels, as assigned through genomic analysis, refer solely to molecular configurations. Even when established through a populational approach, "genetic ancestry" concerns DNA and its variability. Therefore, in a deeper sense, we can ask: What is involved in the idea of ancestry? What do geneticists mean by genetic ancestry and its relation to DNA segments, and why is it so often misinterpreted—as much in scientific literature as in public reception—as a statement about social identity?

Because geneticists speak of an evolutionary path that led to the selection of particular DNA configurations over others, their concept of ancestry refers to an original population from which these DNA stretches disseminated—typically associated with specific geographic locations or sociocultural geographies. Although ancestry labels encode a geographic dimension—hence the term *biogeographic ancestry*—what emerges from the interviews, geneticists' publications, and other relevant works in population genetics is that evolutionary processes leading to differences in DNA molecular configurations are taken to justify the classificatory logic employed by the dengue researchers.

14 The *1000 Genomes* database was used as the reference for ancestry testing, with the European and African components represented by fifty samples each of Southern European (Italian) and Yoruba origin, used for genomic analysis and ancestry mapping (Sierra et al. 2017, 15).

These evolutionary processes, moreover, encompass dynamics such as mutation, selection, reproduction, and transmission, as well as environmental influences such as diet and pathogen exposure—all of which are thought to leave molecular traces in population-level genetic profiles.[15] The hypothesis I am advancing here is that ancestry is intended to identify population-geographic origins for particular DNA segments—not by ignoring evolutionary complexity, but by engaging it directly. While sociocultural classificatory labels such as ethnolinguistic categories may also be used, the global ancestry identities attributed to molecular sequences ultimately refer to hypothetical ancestral populations located in a specific geographic space at a defined historical time—populations understood as the "origins" of the DNA segments under study.

It is this complex technomethodological set of scientific norms and procedures—interrelating populations, molecules, historical narratives, and geographic and temporal locations—that must be disentangled in exploring what may lie within the concept of genetic ancestry.

Evolutionary Theory and Naming Populations in Genetics

The Theoretical Framework: Situating Change in Time and Geography

Clearly, the notion of global biogeographic ancestry was well-suited to the DENFREE project as a classificatory baseline for the molecular data analyzed by the project's geneticists. As demonstrated in the previous section, the DENFREE research proposal led geneticists to methodologically operationalize global ancestry categories deemed significant and to apply an ancestry mapping approach that, they argued, would enhance the statistical efficiency of GWAS. As shown in the team's published articles, GWAS was applied following ancestry mapping to identify genomic regions responsible for differences in dengue clinical outcomes (e.g., Sierra et al. 2017). Their conclusions pointed to the identification of loci believed to confer resistance to DENV infection, inherited as part of genome sections labeled as African ancestry blocks.

I interpret the use of four global ancestry categories by the research group as linked to the need to integrate a conventional medical genetics approach—centered on GWAS—with the evolutionary dimensions of genetic change and variation that, by disciplinary tradition, fall under population genetics. These dynamics of genetic change across time and space are particularly important to geneticists and are closely tied to selection events and past migrations. Such knowledge helps interpret molecular configurations and link them to specific phenotypic traits, thereby bridging the gap between genotype and phenotype. Oliviera et al (2018, 2–3) argue, for instance, that a genetically selected protection against a local pathogen in one population may, through migration and admixture, appear in a new region, where differences in ancestry between cases and controls reflect not sampling bias but a protective evolutionary mechanism.

15 See, e.g., Stoneking (2017) for a thorough, accurate and accessible treatment of molecular anthropology, the fundamentals of genetic variability, and evolutionary dynamics.

Beyond the strict medical genetics approach of using GWAS to identify genes that may play a role in specific pathologies, these population researchers are committed to understanding the role of different genetic configurations in immunological processes. To explore these links, and beginning with a connection between clinical phenotypes and epidemiological information based on visual traits, these practitioners produce sheets of genetic data that can be analyzed to identify loci in the genome related to this clinical information and phenotypes. Having identified these genomic locations, they must develop an understanding of the role of these loci as causal entities in producing differences in medical outcomes. With this problem in mind, the question researchers need to explore concerns dengue showing different forms and degrees of severity in populations with apparently different genealogical backgrounds. They attempt to clarify this connection by tracing the history of the genetic variants themselves and asking why such molecular differences exist, where they come from, and how following their trajectories might help explain the outcomes.

As noted above, GWAS tools typically identify genetic traits that may be relevant to certain diseases in population groups sharing similar genetic structures. But genetic structure similarity, which is crucial for GWAS to work correctly, is also tied to another feature of this method. GWAS configures a synchronic approach to genetics, in that it leaves time out of the analysis. It does not track genetic change over time—that is, it is not designed to incorporate evolutionary thinking into the interpretation of its results: mutation and selection processes are left at the door; what matters are the differences between the target populations (the cohort set of samples) under study and controls. This approach has limitations, as it disregards the theory of change that is the hallmark of population genetics. In fact, it is the trajectories of genetic transformations through time and the movement of populations through space—both of which allow researchers to relate geography, environment, and genetic change—that allow population geneticists to understand evolutionary processes, i.e., how certain genetic traits have been positively or negatively selected and how they have become present or absent in contemporary populations.[16] This knowledge can be of critical biomedical value and play a fundamental role in the search for clinical or pharmaceutical applications. If genetic and phenotypic data were related and made meaningful by considering the role of environmental pressures on past populations, differences could be read through evolutionary processes, and their molecular and metabolic consequences would be well positioned for study. Tracking genetic changes resulting from selection is made possible by ancestry information—in this case, by global ancestry. As DENFREE researchers proposed:

> Human populations are structured in three main groups, African, European and Asian. The independent selective pressures acting upon these groups can lead different genes to be selected in adaptation to the same pathogen, and those genes can interact in a common crucial pathway or in different pathways of additive importance to the disease process. (Sierra et al. 2017, 12)

16 On the role of evolutionary theory and thinking in the history of population genetics, see Reardon (2005).

While the genotype-phenotype relationship is likely the primary concern of geneticists focused on the medical dimension of genetic research, for population geneticists the problem is not merely whether certain genetic or molecular traits characterize a population or how frequently a profile appears. Another fundamental aspect—often absent in medical genetics—significantly shapes the work of population geneticists. Beyond genetic structure and profile frequency, these biologists are largely concerned with tracing processes of transformation across temporal frames. For population geneticists, evolution matters. They understand genetic traits and associated molecular configurations as products of evolutionary dynamics—that is, in Darwinian terms, as the outcome of transformation trajectories shaped by selective pressures and inheritance, and occurring over time at a particular rate of change. Genes—precisely because they are inherited and influenced by sexual-reproductive recombination—are seen as evolutionary products: the result of mutations and selection within populations characterized by a certain degree of variability. In sum, population scale and the temporality of mutations are the two fundamental aspects sustaining the evolutionary process, which is further guided by selection.[17]

I am arguing and attempting to demonstrate that it is this concern with evolution in Darwinian terms that renders the use of identity categories fundamental to the work of population geneticists.[18] For them, genes have a history—but that history can only be told in relation to the history of populations and their trajectories over generations, within which processes of biogenetic transformation, admixture, substitution, and adaptation can be observed. If variation and selection—the core premises of evolutionary theory—explain why certain molecular configurations are more common today in some populations and not others, and if genetic diversity arises through time-dependent processes, then the history of genes can only be pursued after the history of populations themselves.

How, then, can the history of DNA variation (or any molecule dependent on heredity) be explored, if not by identifying the populations that carry and transmit that variability through successive generations of sexual reproduction? For geneticists, there are clearly two productive ways to name such populations:

i. in relation to environment and geography, using categories of spatial location where specific molecular forms are found
ii. in relation to collective names for the populations themselves, often based on ethnolinguistic classifications

Geographic nomenclature typically refers to the places where populations live today rather than to hypothetical historical geographies, since the environmental origins that

17 In Darwinian biology selection is understood as "the only direction-giving factor in evolution" (Mayr and Provine 1980, 311).

18 Though tainted by the pernicious connotations introduced through early sociological appropriations, in contemporary biological sciences "evolution" simply means that living organisms change, that such change produces new species better adapted to shifting environments, and that this process is irreversible, meaning it does not revert to previous forms.

may have exerted selective pressures are rarely known with precision. Nonetheless, the geographic association with genetic configurations remains highly relevant, as heredity and evolutionary processes—such as drift, reproductive isolation, or selection on existing variation—usually occur in relation to ecogeographic conditions.[19]

It is not surprising, then, that the history of genes, as told by population geneticists, is partly expressed in a vocabulary that mirrors that used by historians to narrate population histories and distinguish between different peoples. These denominations are inevitably tied to names of geographical regions, cultural or linguistic classifications, and categories of political differentiation, all of which carry the symbolic weight of social identity.

Identifying Patterns and Tracing Connections

In relation to how geneticists name populations, two sets of scientific infrastructure are typically presented as fundamental to identifying genetic ancestry. The DENFREE genetics research program is no exception. The first set comprises chips with predefined SNP probes based on the HapMap,[20] used in genome-wide analyses to generate data on genetic variation. The second involves reference collections, which provide the comparative framework needed to estimate the proportion of different ancestries in analyzed samples.

I contend that to take advantage of these infrastructural and technological resources while engaging with evolutionary theory, the naming of present or past populations becomes a necessary step in producing knowledge from genetic data obtained in laboratory analysis. Time, along with geographical constraints (environmental factors, reproductive isolation), is central to the dynamics of genetic change in populations—commonly referred to as genetic drift. Thus, information that enables researchers to track the temporality of such processes is among the most valuable components of data. These naming practices—whether derived from fieldwork sampling or from primary and secondary historical sources—can provide the link connecting the present to historical events and contemporary populations to their historical predecessors. Published sources and interviews show that, in efforts to establish such connections and to relate contemporary populations to their genetic past and to history, ancestry has become one of the most useful and widely used concepts in recent decades.

In 2008, the American Society of Human Genetics (ASHG) issued a statement offering recommendations on ancestry testing. Following this, Fujimura and Rajagopalan

19 The correlation between genetic distance and geographic distance has long been considered one of the strongest influences on genetic variation, even in humans, who exhibit very low levels of intraspecific variation (Cavalli-Sforza et al. 1994, 121–125). Although some studies also show significant correlations with cultural and linguistic distances, geography—understood as spatial distance and natural barriers—has been regarded as the most relevant factor in explaining genetic distance (Tishkoff et al. 2009; Wang et al. 2012; Stoneking 2017, 130).

20 SNPs (pronounced "snips") are single nucleotide polymorphisms—variations at a single DNA position, either between individuals or between two chromosomes in one person (Stoneking 2017, 92). Chips (or SNP arrays) used in genomic analysis are "printed" with SNP probes that detect and determine the variants present in the analyzed samples.

argued that ancestry labels, though often necessary for calibrating population structure, are dispensable for epidemiological purposes (Fujimura and Rajagopalan 2011, 15–16).

However, it is reasonable to argue that these evolutionary explanations based on different background origins are indeed useful. I understand that ancestry labels matter to geneticists because they enable evolutionary interpretations of certain medical conditions. Geographic and historical information appears crucial for understanding evolutionary processes in relation to specific environmental conditions. But how, in practice, is this information important, and how does it relate to global ancestry categories? How is ancestry knowledge cognitively relevant?

The dengue research project offers a telling example. DENFREE geneticists observed that African ancestry conferred enhanced protection against dengue hemorrhagic fever compared to individuals of European background (Sierra et al. 2017, 12). In one of their published studies, DENFREE researchers discuss the possibility that another African-origin flavivirus—the yellow fever virus, which shows a 6.8 times higher mortality rate among Europeans—may have contributed to the selection of protective genetic variants, not only against itself but also against hepatitis C and dengue virus (DENV). Drawing on existing scientific evidence suggesting that this protective condition is linked to lipid metabolism, they use ancestry information to trace the link between DENV resistance and changes in lipid metabolism profiles:

> Signs of selection were already identified in West Africans for APOL1 and CD36 genes involved in lipid metabolism and probably driven by pathogen resistance. Our evidence adds two new genes to the differential lipid profile between Africans and Europeans, which play a role in infectious resistance. (Sierra et al. 2017, 12)

Without ancestry profile information on samples and populations, the cognitive articulation between lipid metabolism, infection outcomes, genetic profiles, and the specific genes involved would have been difficult to achieve. This information enables different sample sets analyzed in distinct research projects by separate teams to be related and compared, allowing solid statistical results to be collectively considered within broader epidemiological patterns. The classification of people has long been crucial for interpreting the epidemiological behavior of populations exposed to pathologies with unknown or unpredictable outcomes. One of the first steps in recognizing differential epidemiological behavior and identifying patterns of incidence, resistance, and susceptibility is developing methods for classifying and distinguishing groups within affected populations.

Without ancestry information, the findings on yellow fever and lipid metabolism profiles from other studies could not have been linked to this project. I argue that, without ancestry as a means to classify and relate genotypes, researchers would lack important clues about where to investigate differences in immune responses and their causes.

In the cited article, the DENFREE team explicitly stated that individuals with a certain lipid metabolic profile and a protective condition against some tropical viruses shared an evolutionary history with members of the Cuban cohort. They implicitly attributed this connection to information about the African ancestry of specific DNA segments and the loci identified within them (Sierra et al. 2017, 12–13). A link between

two distinct sample groups from separate projects was established through ancestry information, allowing the DENFREE team to relate lipid metabolism studied in one project to two newly identified genes and to virus protection.

Without ancestry mapping, this relationship could not have been explained; no evolutionary rationale would have led to the hypothesis connecting DENV resistance to analogous protection conferred by lipid metabolic profiles. These links—between pathogen resistance, genes, and lipid metabolism—would have lacked their theoretical evolutionary basis. As the DENFREE team assert, referencing "intense selective pressure" on "genes involved in lipid metabolism" that was "probably driven by pathogen resistance" (Sierra et al. 2017, 12), it was the identification of a common evolutionary background across samples from different research projects that led to the hypothesis connecting lipid metabolism to dengue protection and linking this protection to the "genes" identified through genomic analysis of the DENFREE samples.

Genetic differences are not merely the result of random variation at specific genome locations. Variations do not all carry the same significance, as some are shaped by selective pressures. European, African, or Asian ancestry are not simply labels distinguishing molecular or genetic configurations; geneticists also relate these molecular traits to specific environmental and ecological habitats that may account for such differences, thereby creating a context in which selective pressures and evolutionary processes can be identified and interpreted.

Conclusion

While the concept of genetic ancestry, as used by population geneticists, brings together molecular and sociopolitical dimensions, it is precisely this configuration that makes it scientifically useful as a methodological tool. In the context of a dengue research project aimed at identifying and understanding genetic configurations that might explain differences in disease outcomes along resistance and susceptibility lines, genetic ancestry became central to understanding population-level patterns of clinical outcome. Despite the racialized connotations of epidemiological observations, genetic ancestry categories were not used to distinguish racial groups but to differentiate DNA segments and trace their possible origins. Though linked to the genomes of dengue patients, these ancestry labels served to distinguish chromosome tracts that could become targets of further research into the molecular causes of dengue clinical outcomes.

This genetic difference was interpreted through the differential hereditary transmission of identified loci, enabling researchers to establish ecogeographic origins and to understand why those DNA locations conferred protection against DENV. While race as a sociopolitical category has informed epidemiological observations, it was the transformation of such observations into genetic (molecular) ancestry that enabled geneticists to understand genotype-phenotype relationships in ways that were relevant to underlying immunological processes.

The geographic information associated with ancestry and population history became genetically meaningful because it enabled evolutionary interpretations, supported generalizations, and generated hypotheses for further inquiry. Genetic (molecular) ancestry

nomenclature has followed paths shaped by historical, epidemiological, and geographic knowledge, while retaining links between labels and groups. Without these links, the relevance of such knowledge for interpreting genomic data through an evolutionary lens would be lost.

References

American Society for Human Genetics. 2008. "Ancestry Testing Statement." Accessed 24 June 2025. https://www.ashg.org/wp-content/uploads/2008/11/Statement-20081311-ASHGAncestryTesting.pdf.

Byeon, Yen Ji Julia, Rezarta Islamaj, Lana Yeganova et al. 2021. "Evolving Use of Ancestry, Ethnicity, and Race in Genetics Research—A Survey Spanning Seven Decades." *AJHG* 108 (12): 2215–2223. DOI: https://doi.org/10.1016/j.ajhg.2021.10.008.

Callon, Michel. 1984. "Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fisherman of St. Brieuc Bay." In *Power, Action and Belief, Sociological Review Monograph*, edited by John Law, 196–230. London: Routledge & Kegan Paul.

Callon, Michel, and John Law. 1982. "On Interests and their Transformation: Enrolment and Counter-enrolment." *Social Studies of Science* 12 (4): 615–625. DOI: https://doi.org/10.1177/030631282012004006.

Cann, Rebecca L., Mark Stoneking and Allan C. Wilson. 1987. "Mitochondrial DNA and human evolution." *Nature* 325: 31—36, doi:10.1038/325031a0

Cavalli-Sforza, Luigi Luca, Paolo Menozzi, and Alberto Piazza. 1994. *The History and Geography of Human Genes*. Princeton: Princeton University Press.

Duster, Troy. 2006. "The Molecular Reinscription of Race: Unanticipated Issues in Biotechnology and Forensic Science." *Patterns of Prejudice* 40 (4–5): 427–441. DOI: https://doi.org/10.1080/00313220601020148.

El-Haj, Nadia Abu. 2007. "The Genetic Reinscription of Race." *Annual Review of Anthropology*, 36: 283–300. DOI: https://doi.org/10.1146/annurev.anthro.34.081804.120522.

Fujimura, Joan H., and Ramya Rajagopalan. 2011. "Different Differences: The Use of 'Genetic Ancestry' versus Race in Biomedical Human Genetic Research." *Social Studies of Science*, 41 (1): 5–30, DOI: https://doi.org/10.1177/0306312710379170.

Fujimura, Joan, H., Ramya Rajagopalan, Pilar N. Ossorio, and Kjell A. Doksum. 2010. "Race and Ancestry: Operationalizing Populations in Human Genetic Variation Studies." In What's the Use of Race? *Modern Governance and the Biology of Difference*, edited by Ian Whitmarsh and David S. Jones, 169–183. Cambridge, MA: MIT Press.

Fullwiley, Duana. 2007. "The Molecularization of Race: Institutionalizing Human Difference in Pharmacogenetics Practice." *Science as Culture* 16 (1): 1–30. DOI: https://doi.org/10.1080/09505430601180847.

Fullwiley, Duana. 2008. "The Biologistical Construction of Race: "Admixture" Technology and the New Genetic Medicine." *Social Studies of Science*, 38 (5): 695–735, DOI: https://doi.org/10.1177/0306312708090796.

Fullwiley, Duana. 2014. "The 'Contemporary Synthesis': When Politically Inclusive Genomic Science Relies on Biological Notions of Race." *ISIS* 105 (4): 803–814. DOI: https://doi.org/10.1086/679427.

Fullwiley, Duana. 2024. *Tabula Raza: Mapping Race and Human Diversity in American Genome Science*. Oakland: University of California Press.

Gannett, Lisa. 2014. "Biogeographical Ancestry and Race." *Studies in History and Philosophy of Biological and Biomedical Sciences*. 47: 173–184. DOI: https://doi.org/10.1016/j.shpsc.2014.05.017.

Guzmán, Maria G. 2012. "Thirty Years after the Cuban Hemorrhagic Dengue Epidemic of 1981." *MEDICC Review* 14 (2): 46–51. DOI: https://doi.org/10.37757/MR2012V14.N2.11.

Guzmán, Maria G., G. Kourí, S. Vázquez, D. Rosario, J. Bravo, and Luis Valdés. 1999. "DHF Epidemics in Cuba, 1981 and 1997: Some Interesting Observations." *Dengue Bulletin* 23: 39–43, https://iris.who.int/handle/10665/148771

Guzmán, Maria G., Scott B. Halstead, Harvey Artsob et al. 2010. "Dengue: a continuing global threat." *Nature Reviews Microbiology* 8 (Suppl 12): S7—S16, doi:10.1038/nrmicro2460

Jaenisch, Thomas, IDAMS, Anavaj Sakuntabhai, DENFREE, Annelies Wilder-Smith, DengueTools. 2013. "Dengue Research Funded by the European Commission-Scientific Strategies of Three European Dengue Research Consortia." *PLOS Neglected Tropical Diseases* 7 (12): e2320. DOI: https://doi.org/10.1371/journal.pntd.0002320.

Koenig, Barbara, Sandra Soo-Jin Lee, and Sarah S. Richardson (eds.) 2008. *Revisiting Race in a Genomic Age*. New Brunswick and London: Rutgers University Press.

Lewis, Anna C. F., Santiago J. Molina, Paul S. Appelbaum et al. 2022. "Getting Genetic Ancestry Right for Science and Society." *Science* 376: 250–252. DOI: https://doi.org/10.1126/science.abm7530.

Lipphardt, Veronika. 2017. "The Body as a Substrate of Differentiation. Shifting the Focus from Race Science to Life Scientists' Research on Human Variation." *Varia Historia*, 33 (61): 109–133. DOI: https://doi.org/10.1590/0104-87752017000100006.

Lipphardt, Veronika, Gudrun A. Rappold, and Mihai Surdu. 2021. "Representing vulnerable populations in genetic studies: The case of the Roma." *Science in Context* 34: 69–100. DOI: https://doi.org/10.1017/S0269889722000002.

Mayr, Ernst, and William B. Provine. 1980. *The Evolutionary Synthesis: Perspectives on the Unification of Biology.* Cambridge, MA. and London: Harvard University Press.

M'charek, Amade. 2005. *The Human Genome Diversity Project: An Ethnography of Scientific Practice*. Cambridge: Cambridge University Press.

Nash, Catherine. 2015. *Genetic Geographies: The Trouble with Ancestry*. Minneapolis, London: University of Minnesota Press.

Nelson, Alondra. 2008. "Bio Science: Genetic Genealogy Testing and the Pursuit of African Ancestry." *Social Studies of Science* 38 (5): 759–783. DOI: https://doi.org/10.1177/0306312708091929.

Nelson, Alondra. 2016. *The Social Life of DNA: Race, Reparations, and Reconciliation after the Genome.* Boston: Beacon Press.

Oliveira, Marisa, Joana Ferreira, Verónica Fernandes, Anavaj Sakuntabhai, and Luísa Pereira. 2018. "Host ancestry and dengue fever: from mapping of candidate genes to prediction of worldwide genetic risk." *Future Virology* 13 (9): 647–655, DOI: https://doi.org/10.2217/fvl-2018-0073.

Panofsky, Aaron, and Catherine Bliss. 2017. "Ambiguity and Scientific Authority: Population Classification in Genomic Science." *American Sociological Review* 82 (1): 59–87. DOI: https://doi.org/10.1177/0003122416685812.

Pääbo, Svante. 2014. *Neanderthal Man: In Search of Lost Genomes*. New York: Basic Books.

Price, Alkes L., Nick J. Patterson, Robert M. Plenge, Michael E. Weinblatt, Nancy A. Shadick, and David Reich. 2006. "Principal Components Analysis Corrects for Stratification in Genome-Wide Association Studies." *Nature Genetics* 38 (8): 904–909. DOI: https://doi.org/10.1038/ng1847.

Rajagopalan, Ramya, and Joan H. Fujimura. 2012. "Making History via DNA, Making DNA from History: Deconstructing the Race-Disease Connection in Admixture Mapping. In *Genetics and the Unsettled Past: The Collision between DNA, Race and History*, edited by Keith Wailoo, Alondra Nelson and Catherine Lee, 143–163. New Brunswick, New Jersey and London: Rutgers University Press.

Rajagopalan, Ramya, and Joan H. Fujimura. 2018. "Variations on a Chip: Technologies of Difference in Human Genetics Research", *Journal of the History of Biology* 51: 841–873. DOI: https://doi.org/10.1007/s10739-018-9543-x.

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.

Reardon, Jenny. 2005. *Race to the Finish: Identity and Governance in an Age of Genomics*. Princeton and Oxford: Princeton University Press.

Schramm, Katharina, David Skinner, and Richard Rottenburg. 2012. *Identity Politics and the New Genetics: Re/Creating Categories of Difference and Belonging*. New York, Oxford: Berghahn.

Sierra, Beatriz, Gustavo Kourí, and Maria G. Guzmán. 2007. "Race: a risk factor for dengue hemorrhagic fever." *Archives of Virology* 152: 533–542. DOI: https://doi.org/10.1007/s00705-006-0869-x.

Sierra, Beatriz, Petr Triska, Pedro Soares et al. 2017. "OSBPL10, RXRA and Lipid Metabolism Confer African-Ancestry Protection against Dengue Haemorrhagic Fever in Admixed Cubans." *PLOS Pathogens* 13 (2): e1006220. DOI: https://doi.org/10.1371/journal.ppat.106220.

Stoneking, Mark. 2017. *An Introduction to Molecular Anthropology*. Hoboken: John Wiley & Sons, Inc.

TallBear, Kim. 2013. "Genomic Articulations of Indigeneity." *Social Studies of Science* 43 (4): 509–533. DOI: https://doi.org/10.1177/0306312713483893.

Tishkoff, Sarah, Floyd A Reed, Françoise R Friedlaender et al. 2009. "The Genetic Structure and History of Africans and African Americans." *Science* 324 (5930):1035–1044. DOI: https://doi.org/10.1126/science.1172257.

Wade, Peter, Carlos López Beltran, Eduardo Restrepo and Ricardo Ventura Santos (eds.) 2014. *Mestizo Genomics: Race Mixture, Nation, and Science in Latin America*. Durham and London: Duke University Press.

Wailoo, Keith, Alondra Nelson and Catherine Lee. 2012. *Genetics and the Unsettled Past: The Collision of DNA, Race, and History*. New Brunswick, New Jersey and London: Rutgers University Press.

Wang, Chaolong, Sebastian Zöllner, and Noah A. Rosenberg. 2012. "A Quantitative Comparison of the Similarity between Genes and Geography in Worldwide Human Pop-

ulations." *PLoS Genetics* 8 (8): e1002886. DOI: https://doi.org/10.1371/journal.pgen.1002886.

Winkler, Cheryl A., George W. Nelson and Michael W. Smith. 2010. "Admixture Mapping Comes of Age." *Annual Review of Genomics and Human Genetics* 11: 65–89. DOI: https://doi.org/10.1146/annurev-genom-082509-141523.