3

Eventualizing Human Diversity Dynamics

Admixture Modeling through Time and Space

Carlos Andrés Barragán, Sivan Yair, and James Griesemer

Since the introduction in the early 2000s of direct-to-consumer genomic ancestry (DTC) testing, human genomic knowledge has increasingly challenged popular imagination about what human diversity is—its past, present, and future.[1] Set in motion by academic research projects and by many small, medium, and large private companies, DTC has become a multibillion-dollar industry that prominently targets a sociocultural curiosity about DNA,[2] human origins, and a desire to have them narrowed down to the individual level. As producers and consumers of human genetic and genomic knowledge, we can be hopeful regarding what this type of knowledge can offer about our history and future as a species. Yet, optimism can be mistaken with a misguided sense that DNA data can exhaustively and unequivocally answer who we are by situating our individual histories within the history of *Homo sapiens*.[3] The blooming of DTC was possible through research work on human population genomic ancestry studies (HPGA). Both HPGA and DTC (illustrated by research enterprises such as the Genographic Project and companies such as Ancestry and 23andMe) have greatly complicated how actors and audiences (with multiple backgrounds and motivations) think and argue about complex and ambiguous concepts such as ancestry, ethnicity, history, identity (individual and/or collective), and/or race.[4]

Academic and public exchanges between life scientists, consumers, sample donors, bioethicists, journalists, lawyers, legislators, public servants, social scientists, and others around these concepts remain problematic due to the ways genomic knowledge is produced, disseminated, and consumed. By this we mean that *production* conditions—e.g., the conceptual, technical, and inference apparatuses used to sequence and interpret genomic data—are not necessarily made

explicit or understandable in research articles or genomic services because they are supposed to circulate mainly among specialists in HPGA, for whom such details constitute common background knowledge.[5] In the context of DTC products (tests, datasets, platforms, and narratives), the technological and statistical complexity behind linking individuals in time and space into accounts of ancestry is reduced to an oversimplified, and most of the time anachronistic, use of discrete ancestry-related categories, such as ethnic denominations, geographic locations, and/or nationalities. In the case of the *dissemination* of human genomic knowledge, journalists' insights can be limited by the characteristic brevity of their work or by sensationalist media approaches that contribute to the reification of the categories mentioned above. In the context of human genomic knowledge *consumption*, DTC producers and research participants/consumers are not necessarily interested in questioning or dissecting its "secret sauce," as human geneticist Spencer Wells described the theoretical and methodological scaffolding used to build and articulate narratives of "deep ancestry" in the context of the Genographic Project.[6] Doing so can compromise the overall perception of how robust these products and their findings are *vis-à-vis* the theoretical and statistical assumptions that HPGA studies require to render and interpret genomic datasets.

Perhaps one of the most widely shared assumptions regards the very concept of "ancestry." The fact that there is no effort to define ancestry in most contexts where it is being discussed turns paradoxical since, in the case of both academic scientific research (HPGA) and DTC, findings and products are supposed to teach multiple audiences more about it. In the last decade, several social scientists have pointed out how the concept is taken for granted and what the consequences of that are for multiple vulnerable communities as they consume and/or contest genomic knowledge.[7] Life scientists have pointed to similar caveats while being explicit about the limitations of the insight that genomic data and interpretative tools can offer.[8] There are also examples of collaboration between life and social scientists to deliver critical insights and criticisms.[9] The importance of ancestry is enhanced by the fact that its potential meanings are a base for individual and collective identity and the concept is crosscut by other polysemic concepts, such as ethnicity and race. Yet ancestry is not the only concept around which misunderstandings emerge.

There are other concepts that are perhaps less ubiquitous but that intersect and articulate with the concepts just mentioned in key ways. *Genomic admixture* is one of them. In a very general way, the concept captures a process through which human individuals from populations that have been separated for a long time breed together and produce offspring whose genomic lineages trace back to both populations. This concept led to the production of different technologies (methods) to track the frequency of disease-causing genetic variants by linking them to ancestral populations for contemporary "recently admixed" populations, such as African Americans, Mexican Americans, or Latinos. One of the names given

to such applications in human genomics has been *admixture mapping*. Since the 2000s, a shared goal in admixture mapping studies has been the construction of ancestry informative markers (AIM) for different ancestral populations (e.g., African, European, and Native American populations in the sixteenth century and their descendants in the Americas). The logic and the theoretical framework behind admixture are that the frequency of genomic variants thought to cause a disease might be higher in ancestral populations known for having a higher incidence of a given disease than in other ancestral populations not known for as high an incidence of such disease.[10] Although admixture studies have generally been framed and regarded as positive contributions to understanding human evolution,[11] the biological basis of diseases, and the socialization of DTC, some life scientists themselves have also critically emphasized the need to avoid deterministic interpretations of such genomic factors and datasets to think about ancestry and biomedical risk—something that general audiences in industrialized countries think about in equal terms and that helps to explain the popularity of DTC tests.[12]

The current body of literature on human admixture studies is enormous and includes applications that re-situated HPGA data and insights into areas such as biomedicine and forensics (respectively, HPGB and HPGF), whose analysis requires far more work than we can report and reflect on for this volume.[13]

In this chapter, we offer insights on how the concept of genomic *admixture* is currently being used in HPGA studies and how such uses open up spaces for misunderstandings as producers and consumers of genomic data re-situate findings to think and talk about ancestry and identity through time and space.[14] Within HPGA studies, we further narrow our scope by focusing on recent genomic research studying the peopling process of the Americas. Our aim is not to produce an exhaustive literature review of every single published article reporting on these migratory processes. The analysis provided here is a first step toward the tracking of research studies that aim to identify broad-scale (spatial and temporal) patterns of movement/migration/interbreeding. This will allow us to document and analyze key aspects of how the concept of admixture is understood and deployed to explore models of how humans populated the Americas—a process that is argued to have started 20,000–15,000 years before the present (BP). In this context, we ask how concepts and assumptions of admixture guide research design processes (e.g., modeling practices, assumptions behind choices for sampling strategies, and/or methods and tools deployed). Likewise, we track what assumptions about ancestry and identity are at play as other audiences re-situate findings that involve narratives about population admixture through time and space. These goals are part of a larger research agenda to understand how scientific knowledge is being re-situated between settings (e.g., laboratories, research institutes, companies) and/or between audiences with several degrees of expertise.[15] In this context, we understand scientific knowledge as a complex assemblage of objects that includes—but is not limited to—research questions, models, datasets, findings, visualizations, narratives, etc.

Our argument is that, as different actors re-situate *admixture* findings in and from HPGA, this *re-situation* complicates and disrupts the necessary contexts needed to evaluate the robustness of the genomic knowledge that is supposed to be generated in tandem with other scientific objects. This is the case whether admixture is set to travel as a genomic concept, or as an object that can be used to assess discrete or continuous states of genomic ancestry for individuals (and the aggregates they represent), or as a synonym of polysemic popular ideas of racial and ethnic mixture across the globe. In the form of questions: How well are genomic admixture data and findings traveling?[16] What sort of misunderstandings can happen if the concept is re-situated without also re-situating other objects (e.g., metadata about the populations being represented and sampled), or without necessary clarification about the assumptions it needs to be useful in a given workflow (e.g., in characterizing how homogeneous or heterogeneous a population can/must be as it changes through time and space)? We argue that, when admixture is set to travel and is re-situated, misrepresentations and misunderstandings can take place when producing HPGA and DTC work and critical work about them. Although there is no single model to predict how re-situations beyond the limits of specialists (i.e., life scientists) would take place, we contend that life scientists do have a vantage point to minimize potential misinterpretations of the concept as a euphemism for widespread colonial concepts of racial mixture, regionally illustrated by a plethora of other concepts, such as *mestizaje, métissage, mestiçagem*, and miscegenation.[17] This will require, for example, making explicit how the concept is understood in life scientists' grants and publications, how it is linked to the design of the research workflow (e.g., population representation, sampling, and metadata production), how it articulates the production of findings, and what the potential pitfalls are if the geographic and temporal scopes of analysis are challenged by other audiences.

What follows is a brief description of our analysis for the subsequent sections of this chapter. In the second section, we offer a genealogy of what admixture, as a genetic concept, is supposed to capture about human populations. We also make an argument about how its current uses afford different levels of abstraction during modeling processes and the enrollment of multiple assumptions to make them happen. In the third section, we describe and analyze published research outcomes that have used contemporary and ancient human DNA samples to study, complement, and contest research questions, models, data, and findings about the peopling of the Americas that in the past were mostly a specific domain of archaeologists, biological anthropologists, and paleontologists. In the fourth section we analyze how life scientists themselves and other key actors (such as journalists) re-situate genomic admixture findings and what consequences such processes can have in terms of robustness. Finally, in the fifth section, we discuss some strategies that can be used by life scientists to minimize potential misunderstandings

as their models, concepts, and findings are re-situated by other colleagues, actors, and large audiences.

THINKING WITH A CONCEPT: GENOMIC ADMIXTURE

> *Admixture: The formation of a hybrid population through the mixing of two ancestral populations.*[18]

The brevity of the definition of admixture offered by life scientist Mark Jobling and colleagues in their popular textbook *Human Evolutionary Genetics*—used for teaching undergraduate and graduate students in Anglo-Saxon contexts—stands as a stark contrast to the complexities of the genomic phenomenon it aims to capture.[19] The abstraction embedded in the concept starts with the circumscription of two genetically distinct populations (through time and space) and the emergence of a third (or more) population(s) through breeding. Semantically, the Latin root *admixtus* gives both the verb (*admix*) and noun (*admixture*) forms in the English language, which go back in time as far as the fifteenth century and mean the blending of two or more different things into a new one.[20] These are semantic dimensions that have older and wider historical and sociocultural contexts beyond the appropriation of the term in human population genetics since the 1960s.

Accumulated archaeological, paleontological, historical, and human population genetics research outcomes have shown that the peopling of the planet was possible due to complex processes of migration, settling, and further migration.[21] What the admixture concept adds to the recent genomic study of such large-scale processes is the ability to explore different models to track migration and population interactions that have contributed to shaping human genetic diversity—understood as the total amount of variation in genes or whole genomes of individuals within or among population(s)—and its structure. The application of such models becomes trickier as the elements of an admixture event (two isolated populations and a new one with multiple ancestries) require further characterization, usually offered in terms of genetic ancestry profiles that can have multiple population sources,[22] depending on how far in time and how wide in space such profiles or "diversity panels" are designed to go.

This is a good point at which to emphasize that such *genetic* or *genomic ancestry*—that is, the sources of genomic material within a genome (represented by a living tissue donor or by an ancient bioarchaeological specimen, for example)—is different from other characterizations of connections between individuals and the populations they could represent, such as *genealogical ancestry* or concepts such as *genetic similarity*. Geneticists Mathieson and Scally have emphasized the need to undo the conflation of these concepts, an outcome of the ubiquitous narratives set in motion by DTC, in order to avoid the oversimplification and misinterpretation

of genomic data turned into ancestry substantiations.[23] In the case of *genealogical ancestry*, the relationship created is between an individual and ancestors in their family tree and characterizations of interests such as nationality or surnames. Rather than referring only to an individual's pedigree, *genetic ancestry* is a subset of a family tree through which geneticists track genetic material that is inherited by an individual. On the other hand, the concept of *genetic similarity* between individuals (and the populations they could represent) is better understood as a "summary" of genetic variation built from multiple past or current individuals to represent specific populations (e.g., ancestral or admixed). This is a process that is susceptible to multiple contingencies that include explicit and implicit biases when deploying data about data—metadata—to design and execute research workflows (e.g., the naming, sampling, and representation of existing, deceased, or unknown populations). In the context of DTC, such summaries substantiate narratives of individual identity by conflating genetic and genealogical ancestry data in problematic ways. The most evident challenge is the assumption of population genomic continuity when making statements about "African," "European," or "Amerindian" ancestry that won't hold meaning when focusing on a finer scale (e.g., smaller than a continent) or for time periods that lack historical records.

These subtleties matter because, if a consumer of a genetic ancestry test uses its results to describe or corroborate their preconceived ancestry—individual, cultural—as *admixed*, such characterization may be the result of assumptions about genetic similarity to present-day individuals, rather than an account of ancient or past key admixture population events developed to track and understand the spread of *Homo sapiens* around the world. The latter is the kind of inference that matters most to some life scientists in HPGA interested in the peopling of the planet. Yet these conceptual distinctions in how genomic knowledge is being re-situated from HPGA to DTC contexts are not the only interesting challenges requiring some epistemological considerations when focusing on the concept of genetic admixture and admixture mapping as an application of ancestry identification.

Again, the most basic unit of a theoretical admixture event requires modeling with two geographically isolated populations that breed and produce a new admixed one that should reflect ancestors from multiple sources. However, retrospectively, it follows logic to assume that the two isolated populations were at some point (earlier in time) likely admixed from older populations. Likewise, prospectively, it is possible that the third population could become in the future an isolated one (geographically) and potentially an ancestral one in different admixture events. These aspects of the model prompt several questions. On one hand, how much time does it require for a population to become isolated enough to contribute to a new one through recombination? Is this something that is best estimated in terms of years, or in terms of generations? On the other hand, for how long does an interbreeding process between two isolated populations need to go on before it can be called an admixture *event?* Furthermore, do different admixture events

show lower or higher admixture levels, or are they simply different depending on the distribution of variants that amount to genomic diversity given how these are used to produce diversity panels? We are not interested in arguing that these types of considerations are unknown to life scientists when genomic admixture models are re-situated between labs and research programs.[24] However, its discussion is not so prominent in the training of new generations of researchers (graduate level) and is not necessarily made explicit in the dissemination of admixture mapping research findings in scientific research journals, which usually strive for brevity in content. Although command over these subtleties is achieved through practice (i.e., senior researchers who have accumulated theoretical, statistical, and computational modeling experience), our point is that this type of modeling subtlety and the larger set of scaffolded assumptions necessary to model must not only be made explicit but also contextualized in the process of data interpretation, for scientists and the public alike.

Modeling assumptions in HPGA in general or in admixture mapping in particular do not necessarily signify flaws in the scientific knowledge being produced. Assumptions can also be understood as key scientific objects that facilitate the production of models and should be appraised as the research workflow takes place. Unaddressed assumptions are the concerning instance, since they can lead to overinterpretation of datasets, which has been most evident so far in DTC. As a fairly recent area of research, practitioners of admixture mapping in HPGA are currently debating how its insights can be both descriptive and predictive and are thus setting research priorities. Our premise is that robustness can be built only by addressing the extent to which genomic data can be forced to speak about recent and ancient population admixture events.

The earliest challenge for the application of a basic model of admixture started when researchers pondered the identification of the best possible DNA donors to produce admixture studies. In abstract terms, such a scenario would require samples from the two isolated populations and from the new admixed one. However, that scenario is almost impossible to encounter when studying concrete human populations. When HPGA researchers started networking in the 1980s to produce large-scale research projects to answer questions about the peopling of the world (e.g., the Human Genome Diversity Project), one of the most basic consensuses reached was that the priority was to collect tissue from living populations that had managed to stay relatively isolated from global demographic processes set in motion by European kingdoms and their colonial enterprises.[25] Back then, the idea of sampling individuals representing genetically admixed populations, described as "melting pots," was anything but appropriate for the reconstruction of the history of human populations predating historical repositories. It took multiple technical and scientific developments, and no small number of debates among life scientists, to value concepts such as *admixture linkage disequilibrium* (ALD)[26] as potential strategies for characterizing ancestral populations.[27] ALD led to a method

known as *mapping by admixture linkage disequilibrium*, or MALD,[28] whose logic was substantiated by the association between an allele and a trait (marker) for the purpose of assigning gene(s) to a linkage group. A few years later, in 1998, another method was proposed that left aside the linkage disequilibrium between alleles and a trait and focused on the association between a local chromosomal ancestry and a trait.[29] This method was coined *admixture mapping*[30] and has been used to describe the burgeoning of a research program that applies ancestry identification to learn about the genetic basis of phenotypic variation (e.g., diseases) and to yield the consolidations of AIM panels for different "admixed" populations.[31]

The corpus of research produced during the 2000s and 2010s suggests that ancestry patterns found in so-called or historically self-identified admixed populations are useful for understanding larger evolutionary aspects of human evolution (e.g., timescales, mechanisms), whereas scientists had previously believed that only isolated populations (i.e., ethnic minorities across the globe) could serve this purpose.[32] Yet the interest in sampling ethnic minorities as a means to understand "deep" questions about humanity's journey didn't decrease in the past two decades. Its importance has actually been enhanced by the emergence of ancient DNA (aDNA) as a new "record" or "archive" and by the establishment of paleogenomics[33] as a research area in its own right for "rewriting" human evolutionary history.

In the case of the peopling of the world, biological anthropologists, archaeologists, paleontologists, and now human geneticists have been building models and extensive datasets over the past two decades to reconstruct and characterize the long migration journeys from what we today call Africa (100,000–60,000 BP) to the very end of South America (15,000–10,000 BP). Despite how new datasets complement and/or challenge previous findings, we can point out that a conceptual and methodological constant across these disciplines is the use of geographic areas (e.g., continents, regions, localities) as units of analysis to represent specific population(s) and thus to articulate specific admixture event(s). The point is that, in reality, such units of analysis could represent *multiple* populations and multiple admixture events (interpreted as ancestries). At stake is how, in order to narrate and visualize larger patterns of movements of people (whether sociocultural understandings consider them phenotypically different or similar), researchers silence details about smaller regional admixture processes. From an epistemological point of view, we can think about these choices as methodological trade-offs. As old and new models and datasets are evaluated, one aspect we want to emphasize is the current need to carefully address the spatial and temporal dimensions that articulate "admixture" as a genomic modeling enterprise.

In the next section, we focus on the peopling of the Americas to illustrate how admixture events are being modeled through time and space in a migration process that is estimated to have started 20,000 to 15,000 BP, depending on what set of archaeological, genomic, or paleontological findings are used.[34] We focus on the human population events that have come to stand as significant to narrate and

set in motion old and new models and research workflows about past migrations into North, Central, and South America (see table 3.1). We examine the mechanisms that make such events stable and the workflows that can challenge them. Certainly, our strategy overlaps, at a smaller scale, with Michel Foucault's critique of knowledge (e.g., historical) and sociocultural power. At the end of the 1970s, Foucault set in motion the concept of "eventalization" to capture a method to disrupt the self-evidence of historical constants (through the construction of events, universalities) by pointing out and visibilizing singularities.[35] In our analysis, we are tracking down specific re-situations of historical demographic events and the emergence of new ones as admixture mapping and paleogenomics practitioners explore new models and yield datasets and findings. What we call *eventualization*[36] in this chapter is the current state of innovation, validation, and contestation of models and evidentiary datasets about the movement and transformation of human populations in the not so new "New World." These are insights about events and findings that have not yet turned into "facts,"[37] yet they are well known among certain specialized and general audiences.

THE PEOPLING OF THE AMERICAS: MODELING ADMIXTURE(S)

In this section, we use a few published research outcomes that have challenged paleontological, archaeological, anthropological, and historical datasets by analyzing both contemporary and aDNA samples to characterize the peopling of the Americas from a genomic perspective.[38] This innovative approach is increasingly capturing the attention of funding agencies and of mass media to a point where the potential of genomics to provide answers is positioned higher than the ones held by the disciplines mentioned above.[39] Such hype not only allows the potential sensationalization[40] of genomic findings but also adds a great deal of complexity to the ways in which new genomic events (a sequenced specimen, a proposed migratory route, a new genomic ancestry) are supposed to be contextualized with larger or smaller existing events, regardless of their disciplinary origin (see table 3.1). This is the eventualization process that we also want to highlight, as HPGA modeling emphasizes or relativizes temporal and geographic scales as scientists infer from datasets.

Admixture through Contemporary DNA

In mid-July 2012, *Science* featured a news article by Anna Gibbons summarizing recent genomic findings about the peopling of the New World published online a few weeks before in a *Nature* article titled "Reconstructing Native American Population History."[41] Gibbons, a seasoned scientific journalist, emphasized that the 66 authors co-led by David Reich (at Harvard University) and Andrés Ruiz-Linares (then at University College London) brought to the spotlight a debate between

TABLE 3.1 Large-scale migration events for modeling the peopling of the Americas

Event	Temporal dimensions / dates	Geographic dimensions / locations	Population names / Demonyms	Presumed state of genomic diversity
f	After 1502	The New World (today's North, Central, and South America)	First peoples, Native Americans, Indigenous people, or Amerindians; Europeans; Africans; racial categories; national categories; ethnic categories; minority categories; etc.	Admixed**
e	1502	Forced migration from Africa to the New World	Africans	Isolated*
d	1492	Migration from the Old World to the New World	Europeans	Isolated*
c	After 15,000 BP	North America, south- and eastbound	First peoples, Native Americans, Indigenous people, or Amerindians	Isolated*
b	15,000 BP	Beringia	First peoples	Isolated*
a	Before 16,000 BP	Siberia / East Asia	Siberians / Asians	Isolated*

* This state could change to admixed depending on researchers' temporal and geographic scale (e.g., global, continental, regional, local).
** However, exceptions take place if researchers are sampling *assigned* and/or *self-identified* minorities (or isolated groups).

HPGA geneticists and archaeologists, biological anthropologists, and linguists over whether the first inhabitants of the Americas arrived in one wave of migration or in more.[42] Some HPGA researchers at the turn of the twenty-first century had leaned toward a model featuring one large-scale migratory wave—or a single founding population—since dataset comparisons were read then as showing that contemporary Native Americans across the continent were genetically similar.[43] This model countered, for example, a three-wave model proposed by interdisciplinary research teams twenty-six years earlier.[44] Reich, Ruiz-Linares, and their colleagues merged several datasets into a composite one with DNA samples from 493 Native Americans representing 52 populations, and from 245 individuals representing 17 populations in Siberia (for a total of 738 samples). The analysis of 364,470 single nucleotide polymorphisms in each of the 738 genomes allowed the researchers to argue that there were at least *three* migratory waves (or streams, as the authors call them) represented by three ancestral populations. The earliest one followed a coastal southward migratory pattern, and the latter two followed eastward directions after crossing a now extinct land bridge known as Beringia. The authors of the 2012 study framed their general findings as backing up the three migration waves suggested by the original model using interdisciplinary datasets,[45] rather than the one-wave model proposed using mitochondrial DNA and Y chromosome DNA.

We now briefly turn our attention to the modeling and curating adjustments that were required in order to make Reich and Ruiz-Linares's dataset work with an admixture model. The authors had to parse all samples to identify potential segments of recent European and/or African admixture based on historical timelines (see table 3.1) and mathematically "mask" data—exclude some alleles and segments—for all of the samples that were subjected to analysis.[46] In other words, in this particular research study and its main research question (Who were the first Americans?), the basic admixture model that starts with two isolated populations was turned into a model with three supposedly isolated populations (i.e., First Peoples, Europeans, and Africans at the turn of the sixteenth century) and potentially not one but several new generations of admixed populations based on chromosomal differences (segments) inherited from these three ancestral populations, or "deep lineages," as the authors sometimes called them. Indeed, the study carried out by such a large network of researchers at the time could have been considered "the most comprehensive survey of genetic diversity in Native Americans so far,"[47] but the findings also allowed the researchers to infer "back-migration" events for some of the populations that embodied each of the three proposed migratory streams. We want to emphasize that this is also an interesting contribution to the theoretical modeling character of admixture itself, one that highlights that its descriptive/evidentiary prowess diminishes when forced to map out a higher resolution, beyond general migratory events at continental scales and in search of regional details.[48] Yet this is far from being an end for the model.

On the contrary, it is a threshold for the further development and exploration of algorithmic tools (e.g., software), higher genomic mapping resolutions, the search for new specimens and samples, new data compilations, and the recalibration of old and new diversity or AIM panels.

In hindsight, this pattern of innovation in HPGA during the 2010s is what has promoted the increasing extraction and exploration of aDNA from biological-archaeological specimens collected in past and current research projects across the Americas.[49]

Admixture through Ancient DNA

By 2016, new technical breakthroughs and research outcomes using contemporary Native American and aDNA samples granted David Reich and Pontus Skoglund an opportunity to update the overview of the peopling of the Americas offered by Reich et al. four years earlier. Based on different aDNA specimens and further analysis of other contemporary samples from Native Americans across the region,[50] Skoglund and Reich argued:

> It is now clear that so many founder events and fluctuations in population size have occurred before, during, and after the peopling of the Americas that the evidence from one position in the genome—mitochondrial DNA, the Y chromosome, or any other location—is too subject to random changes in frequency (genetic drift) to provide a complete picture by itself. Only by taking the independent testimony of many locations in the genome simultaneously can we obtain a high-resolution picture of the deep past. The remainder of this article focuses on insights from whole genome studies of Native American population history. While these studies are still in their early days, they have already upended our understanding of key events. Application of ancient DNA technology promises further insights in years to come.[51]

Both researchers highlighted that the value of aDNA lies in offering a more complex picture of past and contemporary genetic structures by filling geographic gaps and complicating currently debated inferences about streams of ancestry/migration.[52]

Indeed, Skoglund and Reich argued that one of the most novel insights coming from the whole-genome approach listed in the 2016 update was the presence of a statistical signal linking contemporary Native American groups in the Amazon and Australo-Melanesians and Andaman islanders. To interpret that link, researchers suggested the theoretical existence of an ancient lineage, *Population Y*—short for the Tupi word *Ypyuéra*, the closest equivalent for the concept of "ancestor."[53] In this inferring scenario, Population Y also represents a wave of migration that contributed to early Australasians and First Americans but not necessarily through a north-to-south model of migration across the Americas. Since none of the aDNA samples available for the Americas at the time (2016) showed this type of ancestry, the authors hypothesized that a population with a more Australasian-related ancestry may have been present "in the broad geographic

area in order to contribute to the *founders* of Native American founders."[54] In other words, a three-layered source of ancestry. They also gestured at a possible alternative scenario: "pulses" of migration[55] through Beringia following a coastal migratory route and then through an ice-free corridor that allowed eastward exploration.

Conceptually, what all the new scattered genomic findings reported by Skoglund and Reich do for admixture modeling is signal that the premise of isolation for ancestral populations is very relative, particularly when aiming to understand populations and events that took place before the individuals represented by the aDNA inhabited the geographic area under analysis (e.g., Beringia first, then the Americas; or all the previous admixture population events prior to event *a* in table 3.1). Similarly, new smaller models need to be integrated within the larger admixture model with a view to start interpreting similarities and differences between patterns of genomic ancestry between individuals representing specific ancient or contemporary populations. Some of these models embed migratory patterns such as "pulses," "back-migrations," and "ghost populations,"[56] as represented in the case of Population Y.[57] Empirically, on the other hand, the emerging of newer aDNA datasets won't necessarily provide definite answers to larger and smaller genomic questions but most likely will challenge older, current, and newer inferences and the models behind them. Such a process may only be enhanced when archaeologists, biological anthropologists, and/or linguists aim to contextualize such datasets vis-à-vis their own, a painstaking process that itself justifies the production of multiple literature reviews and overviews around the main research question. This contrasts with what life scientists think new findings about the research question should generate.[58]

Both Skoglund and Reich generously stated for the audiences of the journal *Current Opinion in Genetics and Development* that "a true understanding of the population history of any group or region cannot be achieved through genomic studies alone, but requires a synthesis of insights from genomics with information from anthropology, linguistics, archaeology and sociology."[59] However, the conscious or unconscious use of the word *information* in the previous sentence rather than *data* or *insights* to describe research outcomes from areas other than genomics, or from multidisciplinary, interdisciplinary projects, points to larger disciplinary tensions that have been emerging as aDNA—a substance—is turned by life scientists into a research area in its own right (i.e., paleogenomics) and extending beyond the scope of *Homo sapiens*. The core of the tension centers around the idea that, for some life scientists in HPGA, genomic datasets are more informative and valuable than paleontological, archaeological, and historical datasets.[60] This alleged authority to "rewrite history" influences not only how concepts like genomic admixture or genomic ancestry are instrumentalized in research workflows that include or exclude assumptions, models, and findings from other disciplines (e.g., anthropology, archaeology, linguistics, paleontology), but also how findings produced with those concepts are set in motion to travel among several

audiences. Most of the time, such impacts take the form of oversimplifications, essentializations, and misunderstandings that can point out key epistemological challenges that intersect the workflows of archaeologists, biological anthropologists, geneticists, and paleontologists alike.

A concrete example of such common challenges for HPGA researchers involves the production and use of metadata—again, broadly understood for the purpose of our analysis as data about data—to characterize and represent the individuals behind contemporary and aDNA samples and the larger groups they are supposed to represent. In table 3.1 we have listed a few key large-scale migration events that several generations of researchers interested in the peopling of the Americas have produced and used as a temporal and geographical grid for the exploration of new models,[61] the contextualization of new findings, and the production of inferences. Each event, from the oldest (*a*) to the most recent (*f*), is associated with modeling (presumed) state(s) of genomic diversity that in turn enabled the identification of larger patterns of genomic transformation and usually delivered results as ancestry percentages. Each association requires assumptions from other disciplines that allow HPGA researchers to circumscribe DNA samples as representative of specific groups of people and no others. Likewise, it requires assumptions to circumscribe segments within such DNA sequences as representative of past ancestries to other groups of people through time and space. The following considerations are only a few structural conceptual challenges HPGA researchers should engage when carrying out research workflows and interrogating newer findings.

If we consider migration event (*b*), for example, we can ask when it makes sense to stop calling migrants exploring Beringia "Siberians" and start calling them "First Peoples" or "Native Americans." Archaeologists and paleontologists have offered evidence of how to trace and differentiate these kinds of transformations through time, but such datasets do not necessarily overlap with the patterns of difference offered by genomic analyses, either for aDNA specimens or for contemporary genomes.

If we consider migration events *(d)* and (*e*), researchers focus on complex genomic admixture processes that started taking place in the sixteenth century and that in the present can be reduced to the characterization that some contemporary populations in the Americas are admixed (e.g., "Latin Americans," as viewed from North America). What we want to emphasize here is that modeling at such large temporal and geographic scales pushes to the background the admixture stories behind standard categories such as "Native Americans," "Europeans," and "Africans." In other words, it is well supported through archaeological, bioanthropological, and historical records that such groups were not genomically homogeneous at the moment of their encounter. On the contrary, each of them represented complex states of genomic diversity or multiple genomic ancestries.

If we consider the migration events contained in (*f*) as a point of departure to model and reconstruct ancestral populations, the admixed characterization of

certain groups of people across the Americas can quickly become less tangible. For this, HPGA researchers only need to increase the geno-graphic resolution and pay attention to the historical regional migrations and breeding processes that substantiate the emergence of countries' populations represented through individual samples in regional studies and composite datasets.[62] The genomic diversity reported with larger datasets for each country contest the assumption that "Latin Americans" show a homogeneous admixture genomic pattern.[63]

Similarly, this general pattern gets challenged when researchers ponder the fact that there are populations in each of these countries that consider themselves *direct* descendants, without admixture, of one of the three ancestral/continental populations. Should these self-reported distinct populations be included as separate populations or as "versions" of the continental ancestries looked for in "admixed" individuals analyzed by HPGA researchers working in the region? Would such a maneuver make sense from a modeling/biostatistical point of view? As we have seen, for Reich and some of his colleagues, this approach could work if the right algorithmic tools are deployed to exclude or mask certain genomic ancestries for recent admixture events, as they did for Native American samples.[64]

TRACKING THE RE-SITUATION OF ADMIXTURE(S)

In 2018 Reich published a book written for general and specialized audiences that aimed to answer both questions he used as a title: *Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past*.[65] The subtitle already elevated aDNA from a biological material with evidentiary power into *the science* that could answer the previous questions. Reich's justification for such a claim was the "ancient DNA revolution" that he and his lab members had helped to forge in the previous years by making aDNA "industrial."[66] By that he meant the streamlining of extraction, sorting, and sequencing fragments of aDNA that allowed the comparison of individuals at a pace not experienced before and that quickly surpassed previous article production and publication. Although Reich encouraged his readers to avoid processing the overview of the revolution and the answers to the research questions as definite (based on some findings and inferences), the tone, as he moves in chapters through continental areas, is dismissive, for example, of what archaeologists and biological anthropologists have contributed and about what they might contribute in the future. This we read as counterproductive at times, considering that it is in an interdisciplinary counterpointing of models, workflows, datasets, and findings where the discussion of concepts such as genomic admixture can yield more robust outcomes for all researchers invested in HPGA. A concrete example of this is the way in which archaeological datasets can relativize genomic ones (i.e., inferred ancient ancestries from a few or single specimens) by showing that these do not necessarily represent well-circumscribed groups—in sociocultural terms—through space and time.

The overall arguments in the book were, by design, meant to be controversial at many levels. One expected controversy exploded on March 23, 2018, when the *New York Times* Sunday Review published excerpts of the book, mostly from chapter 11, "The Genomics of Race and Identity."[67] The Sunday Review piece was titled "How Genetics Is Changing Our Understanding of Race." With some adjustments to make it work as a stand-alone piece, Reich boldly argued that, as a geneticist working on contemporary and ancient DNA, he knew "that it is simply no longer possible to ignore average genetic differences among races."[68] He did state that race was *also* a social construction, but argued that this characterization had been turned by well-meaning researchers into an orthodoxy rooted in "outdated" genetic studies by figures like Richard C. Lewontin (1929–2021). Lewontin had found in 1972 that most genetic variation was linked to differences between individuals and only a very small fraction was linked to differences among racial (continental) groups represented in his study.[69] Reich argued further that, because of the fast-paced turnout of data/evidence he and his colleagues were producing, human civilization should be prepared for more *findings* showing that differences among populations do exist.[70] Needless to say, the speed with which the book and digital versions of the abbreviated piece traveled was matched only by equally fast reactions, debates, and counterarguments,[71] and promoted more attention to the book and further interrogation of its structure and evidentiary foundation. Producing a detailed map of the impact of Reich's arguments on race could yield significant insights on multiple topics but would derail our main purpose in this chapter. However, we must say that Reich's positioning on the concept of race and racism in the book and in the Sunday Review stood in high contrast with the previous scientific reports on admixture and the peopling of the Americas, in which both were absent.[72]

In the previous scientific reports that we used to build and articulate some of our reflections, Reich and colleagues kept a somehow straightforward and overall consistent use of the concept of genomic admixture and a cautious take on the yields of admixture mapping and loci ancestry assignment to answer the questions of who the humans who made it to the Americas were and how they did so. Yet Reich's individual voice in the book overemphasized the current and future descriptive power of aDNA and the predictive power of genomics for medicine and/or highly contentious research areas linking biology to cognition and/or behavior and socioeconomic status:[73] "With the help of [DNA sequencing technology], we are learning that while race may be a social construct, differences in genetic ancestry that happen to correlate to many of today's racial constructs are real."[74] Furthermore, as he weaves an argument for genomic differences between contemporary populations, we read implicit conceptual maneuvers that render opaque the concept of *admixture*, for both ancient and recent samples.

This is an argumentative process in which patterns of ancient genomic ancestry are read as waves or streams of migrations not only to *interpret* genomic structure

between specimens but also to substantiate theoretical populations into clearly circumscribed ones—culturally first and now racially. We say theoretical populations because admixture patterns can go beyond the genome of the specimen under consideration and be used to make inferences about genomic ancestry that are not yet substantiated with material culture (e.g., Population Y, or "ghost populations" in the context of the peopling of the Americas). From another angle, Reich's arguments about biological differences among current populations can be read as elevating *findings* to *facts* in order for genomics to have larger stakes in terms of predictive power. Arguing that aDNA studies show an "exquisite accuracy"[75] undoes the cautious approach followed by Reich and his coauthors in the research papers mentioned above, where they explored, inferred, and contextualized genomic ancestry data with other types of datasets. When concepts such as ancient "Native American ancestry" are re-situated and used to substantiate the idea of a contemporary homogeneous population—in economic, sociocultural, and now racial terms across the Americas—we witness how the overall finding of *continual* ancient admixtures is rendered, at best, opaque. (See Lisa Ikemoto's chapter for an analysis of case studies in which the idea of "pure" races is revealed as a fiction.) In this context, the impossibility of tracking other scientific objects and assumptions, dropped in the process of re-situation, opened doors for multiple audiences to doubt and question the robustness of the overall workflow of HPGA and HPGB as represented in Reich's individual work.[76]

Reich's own re-situation of some of the research findings he produced in collaboration with colleagues—in multiple institutions and countries—to back up his individual perspectives about the biological differences between populations shows us several important things. In terms of the concept of genomic admixture, the content and fate of the Sunday Review piece show us that, although the core goal of the model is to reconstruct sources of genomic ancestry regardless of what a past ancestral and admixed population looked like—the phenotypes of its members[77]—genomic admixture as a concept (and as a process) can be easily oversimplified and misinterpreted as synonymous with racial mixing. What this facilitates is substantiating HPGA findings to corroborate lay stereotypes about "racial populations" and correlations with disease prevalence and biomedical risk, such as those discussed in Tina Rulli's chapter and in the conclusion to this volume. The oversimplification of genomic admixture and the misinterpretations that ensue don't take place only when life scientists themselves re-situate components of their workflows. HPGA findings are re-situated by multiple audiences. A relevant example of this type of re-situation was set in motion with the publication of British journalist Nicholas Wade's book *A Troublesome Inheritance: Genes, Race and Human History* by Penguin Press.[78] In this book, the long-term contributor to the science section of the *New York Times* re-situated multiple HPGA authors, findings, and debates and elevated some of them into reiterations of racial stereotypes about behavior.[79] We point out Wade's case here because, interestingly enough,

Reich uses Wade's re-situation to illustrate "irresponsible" and "racist" stereotypes that have no genetic evidence backing them up.[80] Reich was even one of the 138 life scientists who drafted and signed an open letter published in the *New York Times* Book Review, some of whom had been cited by Wade.[81] At stake here is the need to compare in the future how both of these types of re-situation (by practitioners and by journalists) end up being re-situated in turn by other audiences and what is lost and gained, silenced, and enhanced in the process as they travel through society.[82]

CONCLUSION: ON HOW NOT TO OVERSIMPLIFY GENOMIC ADMIXTURE

Reich's own re-situation of some of the workflows he co-produced with multiple researchers allows us to gesture at the epistemological frailty of re-situating concepts, models, and findings in HPGA. In order to make an argument about the biological basis of some of the global racial assumptions behind populations, and to frame genomic descriptive power into predictive power for biomedical purposes, Reich set findings to travel without other companion objects (e.g., models and assumptions) to a point where they lost robustness. HPGA knowledge requires that the objects in HPGA workflows travel in bundles in order to maintain clarity about the exploratory character of the research area. When that doesn't happen, objects such as findings sit far from the confirmatory position they are expected to have in larger debates (from the peopling of the Americas to the biological and genetic basis of race). Our aim is to blame not the re-situations of genomic knowledge in general but, as we have seen, the *conditions* of it. The settings in which knowledge is set to travel could be improved by taking humble steps that could help audiences gain clarity about how HPGA workflows have been designed and therefore insights into what the extent of inferences could be for objects like models, datasets, and findings. In the specific cases of HPGA admixture mapping studies, one such step could be for researchers to make explicit in their scientific articles and publications how the concept of genomic admixture is being deployed conceptually and practically. This should also include the process of *ancestry assignment*[83] for both ancient and contemporary samples and the larger production of inferences about the populations such samples are supposed to represent.

Our proposal is inspired by more general calls inviting life scientists to clearly state in their research outcomes how they are understanding and using concepts such as race and ethnicity in HPGB and HPGF.[84] These calls have been motivated overall to address biases in how the two concepts—turned in the twentieth century into governmental categories at a global scale—have been granted at the same time too much descriptive and attributive power to assess health risk. Descriptive power in the sense that both categories have been used to identify economic, sociocultural, and political conditions that qualify health outcomes and health infrastructure. Attributive power, on the other hand, in the sense that race and

ethnicity have been used to point out causal mechanisms in disease prevalence. Furthermore, both instances have a looping effect: descriptive uses can easily reinforce attributive explanations that inappropriately inform clinical applications. The social spaces that this transformation aimed to impact were multiple, including schools where new generations of health practitioners are being trained; public and private organizations overseeing provision of health services and/or grant funding; and publishers in control of scientific journals. Although the outcomes of all these efforts in the last two decades varies greatly from country to country due to local challenges that will require scientists to adopt similar models to enforce research and clinical guidelines, it is fair to point out that editors and editorial boards of scientific and biomedical journals have shown willingness and have taken concrete actions toward requiring authors to clarify and justify their use of racial and ethnic categories.[85] In retrospect, perhaps the most positive outcome of these efforts has been to keep the discussion alive so that editorial boards continue revising and updating their guidelines on how to report both categories.[86]

Going back to HPGA and the concept of admixture, we would like to see specialized journals require authors to elaborate on how the concept is conceived and incorporated into their workflows so as to ensure transparency in a process that very easily can be rendered opaque, as authors aim to elevate the exploration of migratory and genomic ancestry models into findings. We anticipate that clarifications about the concept of genomic admixture and the populational processes it seeks to track may not by itself suffice to ensure transparency. In most cases, clarification about admixture would also need clarification and contextualization about other theoretical concepts that, more often than not, pass unaddressed in both specialized and public opinion. This is the case for *population*. Population(s), understood at a theoretical level required for modeling genetic admixture to signify a group of individuals in which random mating is possible, at times also signifies a group of individuals that share sociocultural or biological characteristics. Such nuances facilitate the oversimplification of concepts and genomic processes, particularly when weaving datasets and findings from other disciplines' workflows into one's own work. As we have previously pointed out, oversimplifications open doors for misunderstandings and reifications (e.g., about *ancestry* in general, about genomic ancestry and sociocultural identity at both collective and individual levels). However, scientific and biomedical journals have predesignated sections for all the different types of articles they publish (e.g., methods, discussion, and online supplemental materials), where HPGA researchers can make explicit how they understand and use concepts like genomic admixture and population and the way they produce findings. Unfortunately, these sections (other than "discussion") are not likely to be read by readers and *re-situators* in audiences outside the small circle of specialists who conduct this kind of research.

We are not naïve about expecting that the inclusion of detailed reflections about scientists' concepts will prevent misunderstandings in how multiple actors could

re-situate findings or any other isolated aspect of the workflow. Or, for that matter, how researchers decide to re-situate aspects of their own research outcomes (e.g., Reich). However, having clarifications about how genomic admixture was used and about what cannot be inferred from it affords the possibility of revisiting aspects of the workflow that could prevent other scientists, critics, and consumers of HPGA knowledge from misinterpreting findings and narratives.

Likewise, if misunderstandings and oversimplifications have taken place, such a record in future publications will allow third parties to track down how and where they happened. In other words, there will be an opportunity to partially reconstruct the process of re-situation and find the conceptual and analytical seams that were turned into oversimplifications and potential misunderstandings. Several life scientists have considered and embraced this as an approach that enhances both the robustness and the ethics of their research activities, particularly when working among vulnerable communities.[87]

Bringing clarity to concepts like genomic admixture (or population, for that matter) keeps expanding discussion spaces already opened by the questioning of race and ethnicity as meaningful categories in research areas such as HPGA, HPGB, and HPGF. This is a small but much-needed step, since the boundaries between genomic ancestry and biomedical data are being blurred by some life scientists and by private DTC companies to a point where public opinion reads ancestry and genomic ancestry mostly in terms of disease risk. For the particular research question about the peopling of the Americas, these efforts will prevent misinterpreting the current theoretical scaffold around ancient migration events (see table 3.1, events *a* to *e*) in order to substantiate individual and collective identity through the lenses of race (table 3.1, event *f*) or even ethnicity when used in theory as a politically correct synonym for race.

NOTES

Carlos Andrés Barragán's and James Griesemer's work on this topic is supported by the US National Science Foundation (NSF) under Grant No. SES-1849307–0. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

1. Dorothy Nelkin and M. Susan Lindee, *The DNA Mystique: The Gene as a Cultural Icon* (Ann Arbor: University of Michigan Press, 1995).

2. Kathy Hudson, Gail Javitt, Wylie Burke, and Peter Byers, "ASHG Social Issues Committee: ASHG Statement on Direct-to-Consumer Genetic Testing in the United States," *American Journal of Human Genetics* 81, no. 3 (2007): 635–37; Henry T. Greeley, "The Future of DTC Genomics and the Law," *Journal of Law, Medicine and Ethics* 48, no. 1 (2020): 151–60.

3. See Mark Fedyk's chapter in this volume for a complementary perspective on this topic.

4. See Keith Wailoo, Alondra Nelson, and Catherine Lee, eds., *Genetics and the Unsettled Past: The Collision of DNA, Race, and History* (New Brunswick, NJ: Rutgers University Press, 2012).

5. James R. Griesemer, "A Data Journey through Dataset-Centric Population Genomics," in *Data Journeys in the Sciences*, ed. Sabina Leonelli and Niccolò Tempini (Cham, Switzerland: Springer, 2020),

146–67; James R. Griesemer and Carlos Andrés Barragán, "Re-situations of Scientific Knowledge: A Case Study of a Skirmish over Clusters vs. Clines in Human Population Genomics," *History and Philosophy of the Life Sciences* 44 (2022), article 16.

6. Spencer Wells, *The Journey of Man: A Genetic Odyssey* (London: Allen Lane, 2002); *Journey of Man*, directed by Clive Maltby (Alexandria, VA: PBS Home Video and National Geographic Society, 2003); Spencer Wells, *Deep Ancestry: Inside the Genographic Project* (Washington, DC: National Geographic Society, 2006).

7. E.g., Jennifer Reardon, *Race to the Finish: Identity and Governance in an Age of Genomics* (Princeton, NJ: Princeton University Press, 2005); Catherine Bliss, "Mapping Race through Admixture," *International Journal of Technology, Knowledge and Society* 4, no. 4 (2008): 79–83; Bliss, *Race Decoded: The Genomic Fight for Justice* (Stanford, CA: Stanford University Press, 2012); Duana Fullwiley, "The Biologistical Construction of Race: 'Admixture' Technology and the New Genetic Medicine," *Social Studies of Science* 38, no. 5 (2008): 695–735; Sandra Soo-Jin Lee, Deborah A. Bolnick, Troy Duster, Pilar Ossorio, and Kim TallBear, "The Illusive Gold Standard in Genetic Ancestry Testing," *Science* 325, no. 5936 (2009): 38–39; Sahra Gibbon, Ricardo Ventura Santos, and Mónica Sans, eds., *Racial Identities, Genetic Ancestry and Health in South America* (New York: Palgrave Macmillan, 2012); Catherine Nash, *Genetic Geographies: The Trouble with Ancestry* (Minneapolis: University of Minnesota Press, 2015); Kim TallBear, *Native American DNA: Tribal Belonging and the False Promise of Genetic Science* (Minneapolis: University of Minnesota Press, 2013).

8. E.g., Deborah A. Bolnick, "Individual Ancestry Inference and the Reification of Race as a Biological Phenomenon," in *Revisiting Race in a Genomic Age*, ed. Barbara A. Koenig, Sandra Soo-Jin Lee, and Sarah S. Richardson (New Brunswick: Rutgers University Press, 2008), 70–85; Graham Coop, Michael B. Eisen, Rasmus Nielsen, Molly Przeworski, Noah Rosenberg, et al., letter to the editor, *New York Times*, Book Review, August 18, 2014, https://nytimes.com/2014/08/10/books/review/letters-a-troublesome-inheritance.html; Marcus W. Feldman and Richard C. Lewontin, "Race, Ancestry, and Medicine," in *Revisiting Race in a Genomic Age*, ed. Barbara A. Koenig, Sandra Soo-Jin Lee, and Sarah S. Richardson (New Brunswick, NJ: Rutgers University Press, 2008), 89–101; Ian Mathieson and Aylwyn Scally, "What Is Ancestry?," *PLoS Genetics* 16, no. 3 (2020): e1008624.

9. E.g., Deborah A. Bolnick, Duana Fullwiley, Troy Duster, Richard S. Cooper, Joan H. Fujimura, Jonathan Kahn, Jay S. Kaufman, Jonathan Marks, Ann Morning, Alondra Nelson, Pilar Ossorio, Jenny Reardon, Susan M. Reverby, and Kimberly TallBear, "The Science and Business of Genetic Ancestry Testing," *Science* 318, no. 5849 (2007): 399–400; Lee et al., "The Illusive Gold Standard"; Gibbon, Santos, and Sans, *Racial Identities*.

10. E.g., J. Claiborne Stephens, David Briscoe, and Stephen J. O'Brien, "Mapping by Admixture Linkage Disequilibrium in Human Populations: Limits and Guidelines," *American Journal of Human Genetics* 55, no. 4 (1994): 809–24; Paul M. McKeigue, "Prospects for Admixture Mapping of Complex Traits," *American Journal of Human Genetics* 76, no. 1 (2005): 1–7; Michael W. Smith and Stephen J. O'Brien, "Mapping by Admixture Linkage Disequilibrium: Advances, Limitations and Guidelines," *Nature Reviews Genetics* 6, no. 8 (2005): 623–32; Cheryl A. Winkler, George W. Nelson, and Michael W. Smith, "Admixture Mapping Comes of Age," *Annual Review of Genomics and Human Genetics* 11 (2010): 65–89; Line Skotte, Thorfinn Sand Korneliussen, and Anders Albrechtsen, "Estimating Individual Admixture Proportions from Next Generation Sequencing Data," *Genetics* 195, no. 3 (2013): 693–702; Katherine L. Korunes and Amy Goldberg, "Human Genetic Admixture," *PLOS Genetics* 17, no. 3 (2021): e1009374.

11. The topic has been formalized for a general audience—for example, in the *New York Times* with a popular column titled "Origins" under the direction of Carl Zimmer, a journalist and an author who specializes in evolution and heredity. See the scope of the column at https://www.nytimes.com/column/origins. He is also the author of *She Has Her Mother's Laugh: The Powers, Perversions, and Potential of Heredity* (New York: Dutton, 2018).

12. E.g., Bolnick et al., "Science and Business of Genetic Ancestry Testing"; Bolnick, "Individual Ancestry Inference"; Lee et al., "The Illusive Gold Standard"; Jenny Reardon and Kim TallBear, "'Your

DNA Is Our History': Genomics, Anthropology, and the Construction of Whiteness as Property," *Current Anthropology* 53, Supplement 5 (2012): S233–45.

13. E.g., Mónica Sans, "Admixture Studies in Latin America: From the 20th to the 21st Century," *Human Biology* 72, no. 1 (2000): 155–77; Indrani Halder and Mark D. Shriver, "Measuring and Using Admixture to Study the Genetics of Complex Diseases," *Human Genomics* 1, no. 1 (2003): 52–62; Francisco M. Salzano, "Interethnic Variability and Admixture in Latin America—Social Implications," *Revista de Biología Tropical* 52, no. 3 (2004): 405–15; Guilherme Suarez-Kurtz, ed., *Pharmacogenomics in Admixed Populations* (Austin: Landes Bioscience, 2007); Alkes Price, Nick Patterson, Fuli Yu, David R. Cox, Alicja Waliszewska, Gavin J. McDonald, Arti Tandon, Christine Schirmer, Julie Neubauer, Gabriel Bedoya, Constanza Duque, Alberto Villegas, Maria Cátira Bortolini, Francisco M. Salzano, Carla Gallo, Guido Mazzotti, Marcela Tello-Ruíz, Laura Riba, Carlos A. Aguilar-Salinas, Samuel Canizales-Quinteros, Marta Menjivar, William Klitz, Brian Henderson, Christopher A. Haiman, Cheryl Winkler, Teresa Tusie-Luna, Andrés Ruiz-Linares, and David Reich, "A Genome Wide Admixture Map for Latino Populations," *American Journal of Human Genetics* 80, no. 6 (2007): 1024–26; Winkler, Nelson, and Smith, "Admixture Mapping Comes of Age."

14. For social scientific and STS analyses of the impacts of DTC tests, see, for example, Barbara A. Koenig, Sandra Soo-Jin Lee, and Sarah S. Richardson, eds., *Revisiting Race in a Genomic Age* (New Brunswick, NJ: Rutgers University Press, 2008); Wailoo, Nelson, and Lee, *Genetics and the Unsettled Past*; Reardon and TallBear, "'Your DNA Is Our History.'"

15. We understand re-situation as a process of accommodating the direct or indirect transfer of objects of knowledge from one site/situation to (one or many) other sites/situations. See James R. Griesemer and Carlos Andrés Barragán, Standard Grant: A Case Study of How Re-Situation of Scientific Knowledge from Human Population Genomics Works, NSF grant SES-1849307, 2019–2024; Griesemer and Barragán, "Re-situations of Scientific Knowledge."

16. Following Peter Howlett and Mary S. Morgan, eds., *How Well Do Facts Travel? The Dissemination of Reliable Knowledge* (Cambridge: Cambridge University Press, 2011).

17. See Jayne O. Ifekwunigwe, ed., *'Mixed Race' Studies: A Reader* (New York: Routledge, 2004); Carlos Andrés Barragán, "Untangling Population Mixture? Genomic Admixture and the Idea of *Mestizos* in Latin America," *Gene Watch* 28, no. 2 (2015): 11–13, 20–21.

18. Mark Jobling, Edward Hollox, Toomas Kivisild, and Chris Tyler-Smith, *Human Evolutionary Genetics*, 2nd ed. (New York: Garland Science, [2004] 2014), 609.

19. Jobling et al., *Human Evolutionary Genetics*.

20. *Oxford English Dictionary*, accessed February 8, 2024, https://oed.com.

21. E.g., Richard G. Klein, *The Human Career: Human Biological and Cultural Origins* (Chicago: University of Chicago Press, [1989] 2009); Luigi Luca Cavalli-Sforza, Menozzi Piazza, and Alberto Piazza, *The History and Geography of Human Genes* (Princeton, NJ: Princeton University Press, 1994).

22. Some life scientists refer to *ancestral populations* also as *parental populations*. In this chapter, we decided to use the former term since it avoids overemphasizing kinship metaphors that can obscure the uses of genomic admixture as a model at different timescales.

23. Mathieson and Scally, "What Is Ancestry?"

24. An example of this is when HPGA models, methods, and findings have been re-situated by other life scientists to other areas such as HPGB and HPGF.

25. E.g., Luigi Luca Cavalli-Sforza, Allan C. Wilson, Charles R. Cantor, Robert M. Cook-Deegan, and Mary-Claire King, "Call for a Worldwide Survey of Human Genetic Diversity: A Vanishing Opportunity for the Human Genome Project," *Genomics* 11, no. 2 (1991): 490–91; see also Luigi Luca Cavalli-Sforza, "Human Genome Diversity: Where Is the Project Now?," in *Human DNA: Law and Policy*, ed. Bartha Maria Knoppers, Claude M. Laberge, and Marie Hirtle (Boston: Kluwer Law International, 1997), 219–27.

26. In a very general sense, *linkage analysis* is the mapping of genes (using pedigrees, for example) or traits on the basis of their tendency to be co-inherited with polymorphic loci. On the other hand, the

concept of *linkage disequilibrium* (LD) describes the nonrandom association between alleles in a given population because of the tendency to be co-inherited.

27. E.g., David C. Rife, "Populations of Hybrid Origin as Source Material for the Detection of Linkage," *American Journal of Human Genetics* 6, no. 1 (1954): 26–33; Ranajit Chakraborty and Kenneth M. Weiss, "Admixture as a Tool for Finding Linked Genes and Detecting That Difference from Allelic Association between Loci," *Proceedings of the National Academy of Sciences* 85, no. 23 (1988): 9119–23; Winkler, Nelson, and Smith, "Admixture Mapping Comes of Age."

28. Stephens et al., "Mapping by Admixture Linkage Disequilibrium"; Winkler, Nelson, and Smith, "Admixture Mapping Comes of Age."

29. E.g., Paul M. McKeigue, "Mapping Genes That Underlie Ethnic Differences in Disease Risk: Methods for Detecting Linkage in Admixed Populations, by Conditioning on Parental Admixture," *American Journal of Human Genetics* 63, no. 1 (1998): 241–51; Winkler, Nelson, and Smith, "Admixture Mapping Comes of Age."

30. Here it is important to distinguish between *ancestry assignment* and *admixture mapping*. The former, also referred to as *admixture proportion*, describes the process of estimating in a sample the proportion of the genome that is assigned to a particular ancestry. It requires the use of AIM and diversity panels, their doctoring, or the establishment of new ones.

31. E.g., Heather E. Collins-Schramm, Bill Chima, Takanobu Morii, Kimberly Wah, Yolanda Figueroa, Lindsey A. Criswell, Robert L. Hanson, William C. Knowler, Gabriel Silva, John W. Belmont, and Michael F. Seldin, "Mexican American Ancestry-Informative Markers: Examination of Population Structure and Marker Characteristics in European Americans, Mexican Americans, Amerindians and Asians," *Human Genetics* 114, no. 3 (2004): 263–71.

32. Carlos Andrés Barragán, "Molecular Vignettes of the Colombian Nation: The Place(s) of Race and Ethnicity in Networks of Biocapital," in *Racial Identities, Genetic Ancestry and Health in South America*, ed. Sahra Gibbon, Ricardo Ventura Santos, and Mónica Sans (New York: Palgrave Macmillan, 2012), 41–68.

33. See Susanne Hummel, *Ancient DNA Typing: Methods, Strategies and Applications* (Berlin: Springer-Verlag, 2003); Beth Shapiro, Axel Barlow, Peter D. Heintzman, Michael Hofreiter, Johanna L. A. Paijmans, and André E. R. Soares, eds., *Ancient DNA: Methods and Protocols*, 2nd ed. (New York: Human Press, [2012] 2019); Charlotte Lindqvist and Om P. Rajora, *Paleogenomics: Genome-Scale Analysis of Ancient DNA* (Cham, Switzerland: Springer, 2019).

34. This is a very active and therefore contested temporal model. New paleontological, archaeological, bioanthropological, and genomic findings are frequently emerging and framed as providing new evidence or resolving debates. For example, at the time of preparation of this manuscript, *Science* published an article reporting that the ancient human footprints found in what today is White Sands National Park in New Mexico were left as early as 23,000 years ago. Using radiocarbon analysis on conifer pollen seeds associated with the footprints, researchers confirmed that the footprints are evidence of earlier human entrance(s) to North America, before the well-documented migrations through the Beringia land bridge that occurred at the end of the Ice Age. Jeffrey S. Pigati, Kathleen B. Springer, Jeffrey S. Honke, David Wahl, Marie R. Champagne, Susan R. H. Zimmerman, Harrison J. Gray, Vincent L. Santucci, Daniel Odess, David Bustos, and Matthew R. Bennett, "Independent Age Estimates Resolve the Controversy of Ancient Human Footprints at White Sands," *Science* 382, no. 6666 (2023): 73–75. The workflow in this project was a response to skeptical reactions by critics to previous research findings that proposed a similar timeline but used radiocarbon analysis of aquatic plant seeds. Matthew R. Bennett, David Bustos, Jeffrey S. Pigati, Kathleen B. Springe, Thomas M. Urban, Vance T. Holliday, Sally C. Reynolds, Marcin Budka, Jeffrey S. Honke, Adam M. Hudson, Brendan Fenerty, Clare Connelly, Patrick J. Martinez, Vincent L. Santucci, and Daniel Odess, "Evidence of Humans in North America during the Last Glacial Maximum," *Science* 373, no. 6562 (2021): 1528–31.

35. Michel Foucault, "Questions of Method," in *Power: Essential Works of Foucault, 1954–1984*, vol. 3, ed. James D. Faubion (New York: The New Press, [1978] 2000), 223–38.

36. The difference in spelling is on purpose to signal the reader that our take on the concept is different from Foucault's.

37. See Howlett and Morgan, *How Well Do Facts Travel?*; Sabina Leonelli and Niccolò Tempini, eds., *Data Journeys in the Sciences* (Cham, Switzerland: Springer, 2020).

38. In general, contemporary DNA samples in the studies we dialogue with come from individuals representing Siberians, Native Americans and/or Indigenous peoples, and "admixed" populations across the Americas. Ancient DNA samples, on the other hand, come from several archaeological sites and specimen repositories located across these regions. For a theoretical and methodological view of aDNA research, see Hummel, *Ancient DNA Typing*; Shapiro et al., *Ancient DNA*; Lindqvist and Rajora, *Paleogenomics*.

39. See Michael Balter, "New Mystery for Native American Origins," *Science* 349, no. 6246 (2015): 354–55; Ewen Callaway, "Migration to Americas Traced," *Nature* 563, no. 7731 (2018): 303–4.

40. For a philosophy of biology take on sensationalization and aDNA, see Joyce C. Havstad, "Sensational Science, Archaic Hominin Genetics, and Amplified Inductive Risk," *Canadian Journal of Philosophy* 5, no. 3 (2022): 295–320.

41. Ann Gibbons, "Genes Suggest Three Groups Peopled the New World," *Science* 337, no. 6091 (2012): 144; David Reich, Nick Patterson, Desmond Campbell, Arti Tandon, Stéphane Mazieres, Nicolas Ray, Maria V. Parra, Winston Rojas, Constanza Duque, Natalia Mesa, et al., "Reconstructing Native American Population History," *Nature* 488, no. 7411 (2012): 370–75.

42. The findings of Reich et al. were also highlighted by other journalists in other scientific and news outlets; see, for example, Linda Geddes and Michael Marshall, "Once, Twice, Thrice into the Americas," *New Scientist* 215, no. 2873 (2012): 12; Robert Lee Holtz, "Early Americans Arrived in Three Waves," *Wall Street Journal*, July 12, 2012; Nicholas Wade, "Earliest Americans Arrived in Waves, DNA Study Finds," *New York Times*, July 12, 2012.

43. E.g., Erika Tamm, Toomas Kivisild, Maere Reidla, Mait Metspalu, David Glenn Smith, Connie J. Mulligan, Claudio M. Bravi, Olga Rickards, Cristina Martinez-Labarga, Elsa K. Khusnutdinova, et al., "Beringian Standstill and Spread of Native American Founders," *PLOS One* 2, no. 9 (2007): e829; Nelson J. R. Fagundes, Ricardo Kanitz, Roberta Eckert, Ana C. S. Valls, Mauricio R. Bogo, Francisco M. Salzano, David Glenn Smith, Wilson A. Silva Jr., Marco A. Zago, Andrea K. Ribeiro-Dos-Santos, Sidney E. B. Santos, Maria Luiza Petzl-Erler, and Sandro L. Bonatto, "Mitochondrial Population Genomics Supports a Single Pre-Clovis Origin with a Coastal Route for the Peopling of the Americas," *American Journal of Human Genetics* 82, no. 3 (2008): 583–92.

44. Joseph H. Greenberg, Christy G. Turner II, and Stephen L. Zegura, "The Settlement of the Americas: A Comparison of the Linguistic, Dental, and Genetic Evidence," *Current Anthropology* 27, no. 5 (1986): 477–97; see also Ann Gibbons, "Geneticists Trace the DNA Trail of the First Americans," *Science* 259, no. 5093 (1993): 312–13.

45. E.g., Greenberg, Turner, and Zegura, "Settlement of the Americas."

46. Reich et al., "Reconstructing Native American Population History," 374.

47. Reich et al., "Reconstructing Native American Population History," 373.

48. This type of development is hardly captured in news articles due to its abstract character. Yet tracking it turns out to be very useful for understanding how modeling assumptions evolve while following a workflow to reconstruct genomic admixture.

49. See David J. Meltzer, *First Peoples in a New World: Colonizing Ice Age America* (Berkeley: University of California Press, 2009); David Reich, *Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past* (New York: Pantheon Books, 2018); Jennifer Raff, *Origin: A Genetic Story of the Americas* (New York: Twelve Books, 2022).

50. Pontus Skoglund, Swapan Mallick, Maria Cátira Bortolini, Niru Chennagiri, Tábita Hünemeier, Maria Luiza Petzl-Erler, Francisco Mauro Salzano, Nick Patterson, and David Reich, "Genetic Evidence for Two Founding Populations of the Americas," *Nature* 525, no. 7567 (2015): 104–8.

51. The authors were also highlighting the value of pursuing whole genome sequencing since it produces multilocus rather than single-locus data. See Pontus Skoglund and David Reich, "A Genomic View of the Peopling of the Americas," *Current Opinion in Genetics and Development* 41 (2016): 27–35, 28.

52. This is a point that other life scientists also share, from a theoretical point of view, and have publicly declared: "Jennifer Raff, an anthropological geneticist at the University of Kansas in Lawrence, says that the emerging picture of the Americas is less a revision of the earlier models and more an elaboration. 'It's not that everything we know is getting overturned. We're just filling in details,' she says" (Raff, in Callaway, "Migration to Americas Traced," 304). See also Raff, *Origin*.

53. Skoglund et al., "Genetic Evidence for Two Founding Populations of the Americas," 106; Skoglund and Reich, "Genomic View of the Peopling of the Americas," 31–32.

54. Our emphasis. Skoglund and Reich, "Genomic View of the Peopling of the Americas," 31.

55. We interpret "pulses" in Skoglund and Reich's update article (2016) to be smaller and dispersed—in time and space—waves of migration.

56. Reich, *Who We Are*, 81–83, 96–97.

57. Skoglund et al., "Genetic Evidence for Two Founding Populations of the Americas"; Skoglund and Reich, "Genomic View of the Peopling of the Americas."

58. E.g., Raff, in Callaway, "Migration to Americas Traced," 304; Raff, *Origin*.

59. Skoglund and Reich, "Genomic View of the Peopling of the Americas," 33.

60. Reich argues in connection with the concept of "race" that "To understand why it is no longer an option for geneticists to lock arms with anthropologists and imply that any differences among human populations are so modest that they can be ignored, go no further than the 'genome bloggers.' [Their] political beliefs are fueled partly by the view that when it comes to discussion about biological differences across populations, the academics are not honoring the spirit of scientific truth-seeking. The genome bloggers take pleasure in pointing out contradictions between the politically correct messages academics often give about the indistinguishability of traits across populations and their papers showing that this is not the way the science is heading." Reich, *Who We Are*, 254–55.

61. This is far from being an exhaustive effort to group all the migratory events that can matter for HPGA researchers working in the region, and therefore it shouldn't be read as representing a settled analytical scaffold.

62. E.g., Reich et al., "Reconstructing Native American Population History"; see also Barragán, "Untangling Population Mixture?"

63. See, for example, the findings reported by two studies led by Andrés Ruiz-Linares back then at University College London: Sijia Wang, Nicolas Ray, Winston Rojas, Maria V. Parra, Gabriel Bedoya, Carla Gallo, Giovanni Poletti, et al., "Geographic Patterns of Genome Admixture in Latin American Mestizos," *PLOS Genetics* 4, no. 3 (2008): 1–9; Andrés Ruiz-Linares, Kaustubh Adhikari, Victor Acuña-Alonzo, Mirsha Quinto-Sanchez, Claudia Jaramillo, William Arias, Macarena Fuentes, et al., "Admixture in Latin America: Geographic Structure, Phenotypic Diversity and Self-Perception of Ancestry Based on 7,342 Individuals," *PLOS Genetics* 10, no. 9 (2014): e1004572. For an analysis of these two studies, see Carlos Andrés Barragán, "Lineages within Genomes: Situating Human Genetics Research and Contentious Bio-Identities in Northern South America" (PhD diss., University of California, Davis, 2016).

64. Reich et al., "Reconstructing Native American Population History."

65. Reich, *Who We Are*.

66. Reich, *Who We Are*, xix.

67. Reich, *Who We Are*, 247–73.

68. David Reich, "Race in the Age of Modern Genetics," *New York Times*, Sunday Review, March 23, 2018, https://nytimes.com/2018/03/23/opinion/sunday/genetics-race.html. Title of the online version: "How Genetics Is Changing Our Understanding of 'Race.'"

69. Reich referred to Lewontin's analysis of protein types in blood from individuals representing so-called racial groups—Africans, Asians (East), Asians (South), Australians, Eurasians (West),

Native Americans, and Oceanians—and found that 85 percent of the variation in protein types could be accounted for by variation within racial groups and only 15 percent among the groups. Richard C. Lewontin, "The Apportionment of Human Diversity," in *Evolutionary Biology*, ed. Theodosius Dobzhansky, Max K. Hecht, and William C. Steere (New York: Springer, 1972), vol. 6, 381–98. Also see Richard C. Lewontin, *Biology as Ideology: The Doctrine of DNA* (New York: Harper Perennial, [1991] 1992).

70. He elaborated more on this argument a week later in a short follow-up comment in the *New York Times* in which he addressed a few of the comments posted to the online version of his first article. See David Reich, "How to Talk about 'Race' and Genetics," *New York Times*, March 30, 2018, https://nytimes.com/2018/03/30/opinion/race-genetics.html.

71. For an example of how sociocultural anthropologists received Reich's arguments, see Jonathan Kahn et al., "How Not to Talk about Race and Genetics," *BuzzFeed News*, March 30, 2018, https://buzzfeednews.com/article/bfopinion/race-genetics-david-reich. Other public figures' reactions were compiled by the editor of the *New York Times* as "Race, Genetics and a Controversy," *New York Times*, April 2, 2018, https://nytimes.com/2018/04/02/opinion/genes-race.html.

72. See Reich et al., "Reconstructing Native American Population History"; Skoglund et al., "Genetic Evidence for Two Founding Populations of the Americas"; Skoglund and Reich, "Genomic View of the Peopling of the Americas."

73. See Emily Merchant's chapter in this volume for a thorough discussion of these topics.

74. Reich, "Race in the Age of Modern Genetics."

75. Reich, "Race in the Age of Modern Genetics."

76. See responses in Kahn et al., "How Not to Talk about Race and Genetics"; *New York Times*, "Race, Genetics and a Controversy."

77. It is important to clarify that in the context of "racial" takes to characterize current human populations, the phenotype is translated into an overgeneralization of some physical characteristics of individuals that are supposed to represent a given "race." In current biological theory, on the other hand, "phenotype" is understood as a set of observable characteristics in an individual resulting from interactions between the individual's genotype and the environment.

78. Nicholas Wade, *A Troublesome Inheritance: Genes, Race and Human History* (New York: Penguin, 2014).

79. Wade, *Troublesome Inheritance*.

80. Reich, "Race in the Age of Modern Genetics."

81. Coop et al., letter to the editor; see also Griesemer and Barragán, "Re-situations of Scientific Knowledge."

82. This is an endeavor we're working on, but expanding on it will overflow the scope and space for our analysis in this edited volume. However, we want to offer a glimpse of how the process of re-situation can turn into a rabbit hole. In response to Reich's Sunday Review article, Wade wrote a letter to the editor of the *New York Times* stating the following: "At last! A Harvard geneticist, David Reich, admits that there are genetic differences between human races, even though he puts the word race in quotation marks. Obvious as this may seem, American academics for decades have insisted that race is a social construct, and have vilified as a racist anyone who says otherwise. After covering the human genome project for this newspaper for many years, I wrote a book, *A Troublesome Inheritance*, which explained that there is indeed a biological basis to race, a fact that Mr. Reich has now echoed, though without acknowledgment. He even tries to portray my book as racist, which it is not. Acknowledging that race has a biological basis is a salutary advance. Opposition to racism should rest not on the lie that races don't exist but on principle, allowing science to proceed without hindrance. It is those who believe that free scientific inquiry will turn up something terrible who should check their consciences. The human genome's forceful message is one of unity: All races are but variations on a single theme." Wade, in the *New York Times*, "Race, Genetics and a Controversy." Wade seized the opportunity to use

Reich's re-situation of findings to legitimize arguments synthesized in his book (Wade, *Troublesome Inheritance*) and to dispute Reich's framing of Wade's book as racist. At stake in this brief skirmish is how each author values findings in HPGA and when it is objective to use such findings to substantiate differences between so-called "racial" populations.

83. See n. 30 above.

84. E.g., Simon M. Outram and George T. H. Ellison, "Anthropological Insights into the Use of Race/Ethnicity to Explore Genetic Determinants of Disparities in Health," *Journal of Biosocial Science* 38, no. 1 (2005): 83–102; Outram and Ellison, "Improving the Use of Race/Ethnicity in Genetic Research: A Survey of Instructions to Authors in Genetics Journals," *Science Education* 29, no. 3 (2006): 78–81.

85. See Outram and Ellison, "Improving the Use of Race/Ethnicity in Genetic Research"; Andrew Smart, Richard Tutton, Paul Martin, George T. H. Ellison, and Richard Ashcroft, "The Standardisation of Race and Ethnicity in Biomedical Science Editorials and UK Biobanks," *Social Studies of Science* 38, no. 3 (2007): 407–23; "Instructions for Authors: Reporting Demographic Information for Study Participants," *Journal of the American Medical Association*, accessed February 8, 2024, https://jamanetwork.com/journals/jama/pages/instructions-for-authors#SecReportingRace/Ethnicity; editorial, "Why *Nature* Is Updating Its Advice to Authors on Reporting Race or Ethnicity," *Nature* 616, no. 7956 (2023): 219.

86. E.g., editorial, "Why *Nature* Is Updating Its Advice to Authors."

87. E.g., Andrew Smart, Deborah A. Bolnick, and Richard Tutton, "Health and Genetic Ancestry Testing: Time to Bridge the Gap," *BMC Medical Genomics* 10 (2017), article 3; Mathieson and Scally, "What Is Ancestry?"