1

Introduction

The researcher in the field

On December 15 1996, I went to Munich Airport to meet a professor in population genetics. She had travelled from Tel Aviv to visit the laboratory where I was conducting my research. After we had found each other in the crowd, we took the train back into the city. Professor B.-T. turned out to be a very pleasant person and quite soon we found ourselves in animated conversation. She told me about the rare DNA samples that she had brought along and where she had collected them. The members of the laboratory were looking forward to the samples, specifically because they were running short of male samples from these populations. She had heard that I too was going to use the samples for my research project. I told her about my study and what I had uncovered so far. At the same time, I started to feel a little uncomfortable. I felt the urge to "reveal" my "identity" to her. Because I was not just a member of the laboratory: I was also studying it. However, before I could do so, Professor B.-T. was eager to learn where I came from. I told her that I lived in Amsterdam but that I was originally from Tunisia. A little shy but curious, she asked me whether I was also from "one of those interesting populations." I had to disappoint her there, but I told her about the genealogical history of my family, which dates back over a couple of hundred years and goes back into Lebanon.

Two years later I was visiting Professor B.-T. in Tel Aviv. She invited me to her laboratory and introduced me to her group. I learned that her laboratory housed one of the consortia of the Human Genome Diversity Project (referred to in this book as the Diversity Project) where they were growing cell lines of various population samples. In addition, when she introduced me to her colleagues, I was surprised that I was introduced not as a social scientist from Amsterdam but as a member of the laboratory in Munich.

The stakes and the argument

In 1991, a group of population geneticists embarked on an international project designed to map human genetic diversity.[1] The initiators of the Diversity Project were interested not only in mapping contemporary genetic diversity as such but also in studying how the current diversity had evolved, and how people and genes had spread over the world. Knowledge of the origins of populations, as one of the initiators of the project has stated, would have "enormous potential for illuminating our understanding of human history and identity."[2] By tracing similarities and differences in the DNA of various groups of people, geneticists hoped to reconstruct where humans came from, along which routes they migrated and when, and how different groups of people relate to one another. To do so, a special emphasis is placed on the study of "indigenous peoples" and "isolated populations." They are deemed the "treasure keepers" of original information, which, in the course of history, had gradually been obscured in other large groups because of migration and admixture. Isolated populations are held to be conservative in this respect by geneticists.[3] As distinct populations, their DNA is considered to be representative of human genetic diversity at large and, therefore, convenient for attaining the goals of the Diversity Project.

The Diversity Project was launched with rhetoric of preservation, time pressure and alarm. In June 1991 the journal *Science* published an article entitled *A Genetic Survey of Vanishing Peoples*, which opened: "Racing the clock, two leaders in genetics and evolution are calling for an urgent effort to collect DNA from rapidly disappearing populations."[4] One of them, the Stanford population geneticist Luca Cavalli-Sforza, argued that "if sampling is too long delayed, some human groups may disappear as discrete populations. . . . At a time when we are increasingly concerned with preserving information about diversity of the many species with which we share the Earth, surely we cannot ignore the diversity of our own species."[5]

However, the Diversity Project soon ran into trouble. It was faced with a variety of criticisms, especially from indigenous and environmental organizations. It was soon dubbed "The Vampire Project," referring to the collecting of blood samples.[6] Furthermore, this naming seemed to suggest that the people sampled were ill-informed and misled by geneticists and that the samples were collected for interests other than those of the sampled groups. In the television documentary *The Gene Hunters*, the Professor of Medical Ethics at Massachusetts Institute of Technology George Annas said: "We're taking from them their DNA, which we now consider like gold. It's even worse than standard colonialism and exploitation, because we are taking the one thing that *we* value,

and after we take that, we have no real interest in whether they live or die." In that same documentary, the spokesperson for the Arhuaco People, Leonora Zalabata, stated: "Our land, our culture, our subsoil, our ideology, and our traditions have all been exploited. This [the Diversity Project] could be another form of exploitation. Only this time, they are using *us* as raw material."[7] The criticism led to heated debates about the social and ethical aspects of the Diversity Project. In 1993, the Rural Advancement Foundation International (RAFI) as well as other political agents urged geneticists to incorporate indigenous organizations in every step of the Diversity Project and to reassess its scientific and ethical implications. By the mid-1990s, many other organizations, including the Bioethics Committee of UNESCO and the US National Research Council (NRC), were calling for strict regulations on how to sample and handle the information obtained.[8] The project had also become part of a debate about commercial revenues in science, such as the patenting of human genes and the development of drugs for specific diseases.[9] Geneticists, however, emphasized that their initiative had no commercial interests, nor would they accept funding from commercial agents.[10] They argued that the knowledge resulting from the Diversity Project might contribute to the understanding of genetically inherited diseases but its major goal was an investigation of genetic diversity and the history of human migration. This "pure science" approach was looked at with suspicion, for example by Ray Apodaca, a spokesman of the National Congress of American Indians. Countering the "pure science" claims he stated: "We know where we came from, and we know who we are, and we think we know where we are going. Why do we need to know anything else? I mean, is this for their benefit? It certainly isn't for ours."[11]

In the face of this criticism, the Diversity Project met initial problems finding financial or other support within the scientific community and institutions.[12] In Europe, the Human Genome Organization (HUGO) proved at an early stage to be willing to finance a series of workshops in order to assess the project's scientific values. Some US organizations, such as the US Science Foundation and the US Department of Energy, followed this by funding three planning meetings.[13] However, whereas the European Union had supplied 1.2 million dollars to set up laboratories where European genetic diversity could be studied, the project was put on hold for several years in the USA. By the end of 1997, however, a committee of the US NRC had evaluated the project and found that it should receive financial support within American national borders, provided that it met ethical and legal restrictions placed on genetic research funded by federal agents.[14] While few research projects received financial support, a number of diversity consortia for the storage of samples and the growing of cell lines were established, such as the one I encountered

in Tel Aviv. Thus, although haltingly and dispersed, the Diversity Project started.

This book is about the Diversity Project. More specifically, it deals with genetic diversity in scientific practice. Prompted by the issue of "conserved genes" and the mapping of similarities and differences between populations, it focuses on what genetic diversity is made to be in scientific practice. This brief review of the controversy shows some of the political stakes in the Diversity Project. Rather than a study of that controversy and of the different politics involved in the debate outlined above – however important and interesting in its own right – this book aims at tracing the politics of genetic diversity in laboratory routines. It investigates the daily practice in which humans, samples and technology are aligned to produce the stuff of which the power and prestige of science is made. The argument pursued throughout this book is that genetic diversity is not an object that lies waiting for the scientist to discover, nor can it be treated as a construct of scientists. Genetic diversity is enacted in a complex scientific practice. It is not only dependent on the scientist and the DNA but also on the various technologies applied to produce it.[15]

Let me briefly illustrate the relevance of technologies for the Diversity Project. For instance, the haste with which geneticists attempted to "conserve" human diversity before "isolated populations" ceased to exist as such cannot be explained exclusively in socio-cultural terms, or as a sudden interest in (bio)diversity. What is at stake is not the fact that the lives of these groups of people are endangered or that their integrity is threatened because they nowadays tend to migrate and mix more frequently with other groups than in previous times – a so-called "death by reproduction" as Corine Hayden[16] aptly termed it – nor is it that these groups only came to the attention of geneticists in the late 1980s. Many of the geneticists participating in the Diversity Project had already been studying and comparing these populations previously and some had even stopped doing so in the 1970s because they "ran out of data."[17] With the technology available at that time, these scientists could acquire no more information from the samples they had. What *did* change by the end of the 1980s was the availability of new technologies. The introduction of revolutionary technologies to the field of genetics had made it possible to produce new "data" based on the samples already collected and also brought within reach a study of diversity on a much larger scale. What these technologies are and how they affect genetic diversity is, therefore, at the center of this book. Consequently, rather than *whether or not in our genes*,[18] the question addressed is *how* in *whose* genes? Before going into the details and the organization of this book, let us first go back to the Diversity Project to have a second look at how it is (intended to be) organized.

The Human Genome Diversity Project

The Diversity Project did not emerge in isolation. Many more genome projects were launched in the 1990s and before.[19] Most powerful is the *Human Genome Project* (HGP), aimed at *the* human genome, which had been presented to the world in June 2000 by science, commerce and politics.[20] Since the Diversity Project was announced by its initiators as a response to the HGP, I will elaborate on this. The goal of the HGP was a map of the complete human genome.[21] The sequence map would function as a reference genome against which all human individuals could be located and compared. As *the* reference, it would provide the genetic terms in which all individuals would be expressed.[22] One of the initiators, the geneticist Walter Gilbert, presented the HGP as the ultimate means to know oneself. He insisted most strongly that molecular biologists would have the final answer to what it is that makes us human, namely the DNA. One of his frequently quoted statements was that: "one will be able to pull a CD out of one's pocket and say, 'Here is a human being; it's me!'"[23]

The compact disc (CD) metaphor is obviously a pregnant one, not only because it allowed Gilbert to make his argument tangible during his presentations by actually pulling a CD out of his pocket but also because it underlined the technical aspects of genomes and genetics. However, riding on that metaphor, the political stakes are not only in knowing the information it contains but also in how and where the CD is produced. What kinds of polymerized substance, stencil-plate and printing technologies contribute to the CD? How can it be played and what kind of equipment is necessary? How can it be read and who will be able to read it? Who will have a CD? What about the possibility of copying it? Will the result be a copy or a clone? Also, what kind of place will the CD-of-life take in the collections of those who have many different CDs? Will it be able to compete with a CD containing a family photo album, with one bearing a game called *Doom* or with that of a singer called *Fairouz*? What kind of practices make the one CD more important than the other? These questions encompass more and more people and things and make them and the relations between them part and parcel of the CD-of-life. In addition, since the goal of the HGP is to produce *one* CD, a question raised within and outside the confines of genetics is, whose CD is it going to be?

> The first complete human sequence was expected to be that of a composite person: it would have both an X and a Y sex chromosome, which will formally make it a male, but this "he" would comprise autosomes [non-sex chromosomes] taken from men and women of several nations – the United States, the European countries, and Japan. He would be a multinational and multiracial melange, a kind of Adam II, his encoded essence revealed for the twenty-first century and beyond.[24]

This was written by Daniel Kevles, half ironically, in *The Code of Codes*, a now classical edited volume about the HGP. However some geneticists outside the realm of the HGP claimed that "[t]he Human Genome Project aims to sequence 'the' human genome with DNA taken mainly from individuals likely to be of European ancestry in North America and Europe. But, like all brothers and sisters, all humans have slightly different genomes."[25] They, therefore, suggested another genome project, the Diversity Project, which "wants to explore the full range of genome diversity within the human family."[26]

Studies of human genetic diversity are not new and go back to the beginning of the twentieth century, when they were based on blood groups. In addition, DNA-based research flourished from the mid-1970s onwards.[27] Hence the initiation of the Diversity Project followed from ongoing research. Yet every project has a myth of origin:[28] there is a date of birth and there are great men involved; there is a vision and there are allies inside and outside the field; there is a world to be gained and ghosts to be exorcized. What follows is the origin myth of the Diversity Project.

The Diversity Project was initiated in 1991 by a group of American geneticists among whom were the late Allan Wilson (Professor of Biochemistry at Berkeley University) and Luigi Luca Cavalli-Sforza (Professor of Population Genetics at Stanford University). Together they found more colleagues welcoming their plan to map genetic diversity of human populations on a worldwide basis.[29] The values of this initiative (referred to in the quote as the HGD Project) were summarized as follows.

> The main value of the HGD Project lies in its enormous potential for illuminating our understanding of human history and identity
>
> The resource created by the HGD Project will also provide valuable information on the role played by genetic factors in predisposition or resistance to disease
>
> The HGD Project will bring together people from many countries and disciplines. The work of geneticists will be linked in an unprecedented way with that of anthropologists, archaeologists, biologists, linguists and historians, creating a unique bridge between science and the humanities
>
> By leading to a greater understanding of the nature of differences between individuals and between human populations, the HGD Project will help to combat the widespread popular fear and ignorance of human genetics and will make a significant contribution to the elimination of racism.[30]

A central question of population genetics is how did humans migrate out of Africa to 'colonize' other regions in the world, and when did these events take place.[31] The idea is that human genetic makeup is indicative of historical events and vice versa: that the contingency of human history is reflected in the DNA. By tracing similarities and differences in the DNA fragments of various

populations, geneticists hoped to provide another (a better?) account of human history. Culture and nature are thus married-up in the Diversity Project.

> There is a cultural imperative for us to respond to that opportunity and use the extraordinary scientific power that has been created through the development of DNA technology to generate – for the benefit of all people – information about the history and evolution of our own species.[32]

To reach this goal the initiators aspired to create an internationally organized project: a project based on technologies and knowledge developed within the realm of the HGP and capable of redirecting the work conducted in the field of population genetics. As early as 1991, the Diversity Project was "adopted" by HUGO, which had been established in 1989 within the HGP. To assess the potentials of the project in Europe, HUGO set up an ad hoc committee in the autumn of 1991. This committee was charged with organizing a series of workshops where various aspects of the project were to be discussed and evaluated, such as the methods of sampling and the storage of the samples, the technologies to be applied and the processing of the information, and the social and ethical aspects of the project. The committee was also asked to conduct a pilot study, using already existing samples, to establish the relevance and added value of the project and to adjust the protocols for the forthcoming research.[33] In the first five years, the project as a whole was estimated to cost 25–30 million American dollars. HUGO provided 1.2 million dollars to organize the workshops and to conduct a pilot study. Additionally, HUGO helped to create a friendlier political climate for the project to get started. Parallel to this, a number of agencies in the USA had supplied some funding for the organization of three workshops to deal with the sampling strategy, the selection of the populations to be studied and the technologies and ethics of the project. The Diversity Project is now organized in a number of regional committees responsible for their own initiatives.[34] Whereas the European regional committee was receiving European Union support as early as 1992, the North American regional committee had to wait until 1997 for federal support.[35]

Making a genetic map of the world

A major goal of the Diversity Project is to make a map of the world that shows genetic relief and contours. Such a map would reconstruct human migration out of Africa and the spread of humans and their genes around the world; the intention was to be able to assign different *populations* to different loci on that map. Yet its two initiators, Cavalli-Sforza and Wilson, already had conflicting

ideas about the sampling strategy; that is, about what a *population* is. Whereas Cavalli-Sforza had strong ideas about how to define a population, namely on the basis of linguistic criteria, Wilson argued against any presupposition about what population is. In an interview with *Science*, Wilson stated: "We should abandon previous concepts of what populations are and go by geography. We need to be explorers, finding out what is there, rather than presuming we know what a population is." Hence his idea was that what population *is* should be the outcome of genetic research and not the start. He, therefore, suggested a grid sampling based on geographical distances (100 miles).[36] The grid approach, however, was considered too costly in terms of time and money, and categorization according to linguistic criteria was regarded as the most appropriate.[37]

In addition, in the NRC evaluation of the Diversity Project, the term population was bracketed.

> The term "population" has many meanings; it is most often used to designate a body of persons (or other organisms) that have a common quality or *characteristic*, to designate a group of interbreeding organisms, or to designate a group of persons (or other organisms) that occupy a specific geographical locale.[38]

Taking linguistic criteria as characteristics, geneticists were faced with 5000 different populations.[39] Yet, as in the case of a geographical grid, sampling, storing and studying the cell material of all these groups did not seem feasible. The initiators, therefore, decided to focus on just 500 populations. The selected populations should do the following.

> . . . answer specific questions about the processes that have had a major influence on the composition of current ethnic groups, language groups, and cultures . . . [This suggests a study of] populations that are anthropologically unique; populations that constitute linguistic isolates; populations that might be especially informative in identifying the genetic etiology of important diseases; and populations that are in danger of losing their identity as recognizably separate cultural, linguistic, or geographic groups of individuals.[40]

These qualities not only give clues about what it means to be genetically representative or to enable the tackling of "interesting" questions. They also suggest that the linguistic criterion is highly invested with various notions about the social, the cultural and the biological. In an article published in the *Scientific American*, Cavalli-Sforza reported on the correspondence between the distribution of genes and that of languages among populations. Elaborating on the transmission of genes, language and culture from one generation to the other, he distinguished between a vertical and a horizontal transmission, the first being a transmission between parents and offspring, and the latter a transmission

between unrelated individuals. Whereas genes can only be transmitted vertically, culture and language may be passed on by either path. While identifying the difference between "isolated populations" and populations that have undergone admixture, he stated:

> In the modern world horizontal transmission is becoming increasingly important. But traditional societies are so called precisely because they retain their cultures – and usually their languages – from one generation to the next. Their predominantly vertical transmission of culture most probably makes them more conservative.[41]

Hence language is not just an arbitrary means of distinguishing between groups of people: it is deemed to correlate with the genes. More specifically, this correlation is held to be even more elegant when applied to the Diversity Project's main object of study, namely "isolated populations." By analyzing and comparing the similarities and differences found in various of these populations, geneticists hope to gain insight into "genetically complex" populations: populations that are less isolated, less unique and less easy to categorize and to study. It seems that those who are not considered to be connected to the global traffic of humans and things, especially those in far-off places, are considered best sources for understanding how genetic "melting pots" must have come about.[42] Based on the idea that genetic diversity (just like language and culture) is better preserved in "isolated populations" and the idea that all humans belong to one "genealogical family" originating from Africa, these populations are assigned the role of origin and resource.[43] They are thus considered to be more homogeneous and their genetic makeup to be more conserved. However, how can they represent an overall human diversity, such as aimed at by the Diversity Project? In addition to their homogeneity and conserved genes, the genetic makeup of different "isolates" in different parts of the world is held to represent specific moments in the history of human migration. These migration events may also be represented in intermixed groups but their effect on the clustering of genes tends to be blurred through population admixture. This indicates that representing human genetic diversity at large can only be done if different "isolated populations" from different parts of the world are taken into account.

The emphasis placed on "isolated populations" is relevant for studies of diversity not only in the context of human history but also in that of genetic diseases. In a document issued by the Diversity Project, this relevance is phrased as follows: "Every time we ask whether a particular genetic marker is associated with a disease, we need to know about the normal control population. The need for this comparison increases with the diversity of the population."[44] Therefore, in order to understand the mechanisms of inherited diseases in genetically

diverse populations, "isolated populations" may function as normal control populations. With the help of such information, geneticists hope to trace where and when specific mutations occurred and whether they lead to the same effects: that is, they also cause diseases in the control population. However, where the specific genes related to a disorder are not known, the role of an "isolated population" might be different. For example, if such a population is susceptible to a specific disease, studying that particular population and not one where genetic diversity is larger may be seen as an application of the reductionist method of the natural sciences.[45] Applied to an object of research, this method consists of reducing complexity to a small number of controllable variables that *can* be studied in a laboratory context. In line with this, "isolated populations" rather than other control groups would function as resource material.[46] As a geneticist once explained to me: "It would be crude to place a wall around Friesland [a province in the Netherlands] and observe what happens to its inhabitants, made isolated. These populations live isolated by nature and can give us insight into the development of various diseases." Although geneticists would consider these populations interesting for studies in their own right,[47] within the context of the Diversity Project they occupy the position of resource and can be seen as a "natural" laboratory for the rest. Whether the aim is to reconstruct the migration history of humans, to preserve (knowledge about) human genetic diversity or to study human genetic diseases, the Diversity Project makes some populations into a more appropriate resource than others.

Studying genetic diversity within the context of a project does not just affect what may be considered a population, what a population is and how it is deemed to contribute to its research, it also affects genetics as a field. Within the Diversity Project, geneticists had to decide upon how to sample, how to store the samples and what kinds of technology will be used to study the samples. To create a project, they simply had to work together and standardization is an important condition for achieving that.

The Diversity Project intended to collect 10 000–100 000 samples from the 500 populations under study. The sampling was delegated to the regional committees, who were asked, where possible, to work together with "local" scientists and anthropologists in the field.[48] When the samples left these regions, they were not to travel alone: they should be accompanied by information about the region and about the sampled individual. The samples were to be accompanied to central storage areas by information regarding "sex, age (or approximate year of birth), current residence, place of birth, linguistic affiliation [of these individuals and] current residence, place of birth, cultural affiliation, linguistic affiliation [of the individual's] biological parents."[49] Thus, the study of the diversity