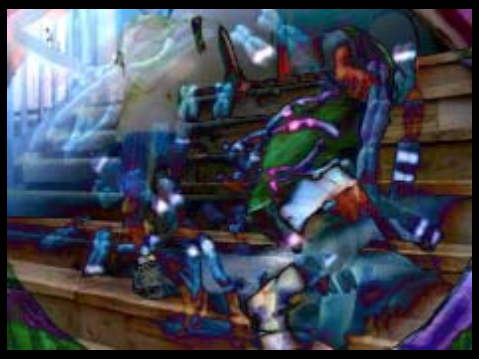
Sequencing the Trellis: The Production of Race in the New Human Genomics



Brady Dunklee
In partial completion of the requirements for honors.
Brown University. December, 2003.
Chris Amirault, advisor.

Sequencing the Trellis: The Production of Race in the New Human Genomics

Brady Dunklee

In partial completion of the requirements for honors.
For the independent concentration "Productions of Biological Knowledge."
Chris Amirault, Ph.D., Departments of Education and Modern Culture and Media, Director Brown Fox Point Early Childhood Education Center.

Brown University
December, 2003

Cover Illustration: Composite images from

-United States. National Institutes of Health (NIH). National Human Genome Research Institute. Internet video: "Exploring Our Molecular Selves." February, 2001. Component of NHGRI's educational resources. Available at http://www.genome.gov/10000002. 2/14/03.

- Risch, N., Burchard, E., Ziv, E., Tang, H. "Categorization of humans in biomedical research: genes, race and disease." *Genome Biology* 2002. **3**:7. http://genomebiology.com/2002/3/7/comment/2007.1

Note on the Title: "Trellis" refers to an analogy that NHGRI director Francis Collins uses to describe race and human evolution, emphasizing mixture between "races," in opposition to evolutionary trees which emphasize divergence. "Sequencing" refers to the main activity of recent genomic research, and is meant to suggest both this activity and the differentiation of groups of people, which is the subject of this thesis.

For Caitlin Dunklee, who is dangerous and knows it.

Acknowledgements

This thesis owes its life to Chris Amirault, friend of clarity, enemy of obfuscation and textual defibrillator extraordinaire. I owe my education at Brown to his teaching, generosity and friendship, and don't know how to thank him. Thanks to Jennifer Reardon for her readership, her scholarship, and for being who I want to be when I grow up. Her work is a rising threat to simplistic thinking and irresponsible science everywhere. Biologists, science studies scholars, and members of any number of sexes and genders owe Anne Fausto-Sterling a debt of gratitude for her work. I take her readership as a great honor, and thank her for it. I thank Joan Richards, Peter Heywood, Lundy Braun, Rachel Morello-Frosch and Barbara Herrnstein Smith, for their professorship, which has made my education at Brown wonderfully difficult. Thanks to Paula Kavathas, Shing Kwok, Latha Narayanan, Peter Glazer, and Barbara Beitch for old favors that have gotten me far. Thanks to Mike "aborted coda" Jackson, revolutionary and shining example of how to write a thesis. Thanks to Alison Barnstable, who has been my friend for three or four lifetimes and has survived all of them with me somehow. Much respect to Jackie Mahendra, Kara Wentworth, Adeline Goss, Sam Solomon and Jess Stites, friends and organizers of science studies at Brown. Thanks to my sister, to whom this thesis is dedicated. Thanks to my mother and father, who should know by now how grateful I am and how much I love them.

Table of Contents

Table of Figures

♦

—We are also bound to seek perspective from those points of view which can never be known in advance, which promise something quite extraordinary, that is, knowledge potent for constructing worlds less organized by axes of domination. — DONNA HARAWAY

—With whose blood were my eyes crafted? —DONNA HARAWAY

—Foucault says knowledge wasn't made for understanding, but for cutting. For the Wendy's worker, that's especially true. The Wendy's worker *knows […]* and his knowledge is alive with results. It is only for *us* to understand. In understanding, we construct, justify, and secure ourselves *above work.* This is how we conceal the knives and restrict their use to the production of delicious results.
—JOE WENDEROTH, from LETTERS TO WENDY'S

—He has his mother's mitochondria, but his father's centrioles.
—PETER HEYWOOD, considering a birth announcement for his son

♦

Thesis Statement

Human genomic science has emerged in the past decade as a powerful new biological field, combining molecular and population genetics with advanced information technologies, allowing DNA sequencing and analysis in a rapid, high throughput fashion. In addition to producing a vast quantity of scientific data, the Human Genome Project and other efforts in human genomics have produced claims about the social implications of their work. The result has been a complex expert discourse on the nature of the human.

A particularly rich subset of this discourse has addressed the meanings, use and reality of race and ethnicity in light of new genomic knowledge. A great variety of positions on racial and ethnic difference have been put forth, best known of which is the contention that race is biologically meaningless.

This thesis shows that this claim is not the whole story. Genomic discourse has, since its beginnings, deployed and produced race in a constant, if variegated manner. A "technology of difference" has been produced, a set of terms, meanings, and ways in which knowledge is structured and authorized, whose collective action is to differentiate people racially and ethnically.

This thesis examines this technology of difference, showing that genomics is in fact *making race*, and demonstrating some of the ways in which it does so.

My approach is an analysis of discourse, which addresses terminology, formal configurations and epistemology in the literatures produced by genomic scientists. The dominant characteristic in this discourse is instability. Meanings, forms, and claims shift and change on a variety of levels.

This thesis shows that surprising patterns can be seen in this instability, and that instability is itself a constitutive factor giving strength and cohesion to the genomic production of human racial and ethnic difference.

I suggest, further, that now is a crucial time for interventions to be made in the genomics of human difference. Those who want an end to race, or who want positive, livable transformations of race, can find both opportunity and danger in these new differentiations.

Introduction

I. Unifications

In late June, 2000, the publicly financed Human Genome Project (HGP) and its rival, the private corporation Celera Genomics announced simultaneously that they had completed working drafts of the human genome. Although it would take HGP until April 2003 to produce a full reference edition, the 85% draft was widely hailed as revolutionary. Comparisons to the moon landing and the Manhattan Project were frequently made.

The Human Genome Project was planned in 1988 through the National Institutes of Health (NIH), funded by NIH and the Department of Energy (DOE) in 1990, and given full institutional status as the National Human Genome Research Institute (NHGRI) by the Department of Health and Human Services (DHHS) in 1997. By 2000 the public Project had "read" eighty-five percent of "The Book of Life," far ahead of schedule.

In most accounts, the speed with which a working draft was achieved is attributed to the pressure exerted on the public consortium by competition with Celera. The company was founded by Applera Corporation and J. Craig Venter in 1998, for the purpose of sequencing the entire human genome in three years, beating out the public project. The relationship between the two was strained. The fact that Celera could make use of all of HGP's data, which was electronically available, while HGP had access to none of Celera's, was a particular source of tension. To use *Nature's* description, "it was portrayed as the scientific rivalry to end all scientific rivalries."(Macilwain, 2000)

By June of 2000, however, unity had become the order of the day. There was rejoicing in Bethesda, Rockville and Washington, and adulation in *Science*, *Nature* and the New York Times.

The leaders of the parallel projects agreed to a simultaneous announcement of draft completion, putting an end to the competitive "race" between the two. They endeavored to downplay their rivalry, and emphasize the significance of their mutual achievement in the massive media coverage that ensued. Craig Venter and NHGRI's director Francis Collins appeared at a White House ceremony, shook hands and reflected on their accomplishment (see Figure I.1). There was a sense that the differences reflected in their rivalry were belittled by the universality of the genome.

> "At the White House event, Collins [a born-again Christian] struck a spiritual note: 'It is humbling for me and awe-inspiring to realize that we have caught the first glimpse of our own instruction book, previously known only to God.' Venter was philosophical: 'The complexities and wonder of how the inanimate chemicals that are our genetic code give rise to the imponderables of the human spirit should keep poets and philosophers inspired for the millenniums.' (Macilwain, 2000)

Everywhere the language was grandiose, often Biblical, and almost always textual.[1] The New York Times said that "the manual with the specifications for the human species" had been read, the "set of instructions embodied in the ancient chemical deoxyribonucleic acid," comprising "the book of life" were in hand. (Wade, 2000) It was felt among the biological community, at least, that something had been achieved which was important, lasting, and universal among humans.

[1] Lily Kay, a science studies theorist, addresses this use of language in genomics. She writes: "The human genome is now generally viewed as an information system, as a book of life written in DNA language, or DNA code, to be read and edited. A 1989 *Nova* PBS television program, "Decoding the Book of Life," promotes the human genome project as a scriptural mission."(Kay, 1998, in Biagioli, 1999) Informational metaphors, in Kay's conception, have extended beyond metaphor in producing knowledge about humans, lending themselves to "visions of informational man." "Though remarkably compelling and productive as analogies, 'information,' 'language,' 'code,' 'message,' and 'text' have been taken as ontologies."

The sense of importance and universality were by no means limited to the sphere of technical and medical research that the genome's sequence promised to revolutionize. Scientific literature and the popular press were brimming with political implications, social ramifications, and the spirituality and philosophy expressed by Collins and Venter.

Neither was it a momentary sort of exuberance. Almost a year later, in February of 2001, the journal *Nature* published a special human genome issue. Its cover was a DNA double helix formed by a mosaic of photographs of human faces (See figure I.2). Collins, in a Short Course introductory lecture on genomics and the Human Genome Project, explained the way in which the image was assembled, demonstrating some of the same sense of human universality evoked in the original fanfare around the genome:

> This cover of Nature, by the way, was chosen by us rather specifically, to convey a message. The message is, of course this is DNA - you recognize the double helix - but if you look closely, this is actually a mosaic, where the tiles of the mosaic are made up of the faces of people from all over the world. And they are people from every ethnicity, and culture, and form of dress, and age, and gender that you can think of, and that really was what we wanted to say. This is the people's genome. This is all of our genomes. This is our shared inheritance. Sure, it's about DNA, but it's really about human beings. And I think that message continually is one that seems very important to this International Public Sequencing Consortium. (Collins, 2001)

The genome here is an olive branch of sorts—a commonality that crosses "every ethnicity, and culture, and form of dress, and age, and gender that you can think of." Divisions, boundaries, conflicts—difference—are rendered small by comparison. The "people's genome" is not only a chemical analysis, a massive set of data, the results of a decade's research, but a "shared inheritance." DNA is its formal medium, "but it's really about human beings."

The theme of friendship, commonality and happiness is echoed in another educational project out of NHGRI. Also in February of 2001, the Institute released a teaching toolkit aimed

at "high school students and the interested public." Several multimedia presentations were featured, including the short video "Exploring Our Molecular Selves." This video is intended to give an introduction to the basics of genetic biology, and situate it within a human context. It is included here for the same reasons—to provide an introduction to what genomics is and does, and begin to more closely examine the ways in which genomics has articulated what it means to be human.

The opening and closing shots of the video feature two boys sitting on the stairs of a porch, smiling broadly (See figure 1.3 for stills from this video). They are identifiably members of different "races:" the boy on our right is "African American," the boy on our left is "Caucasian."[2] The camera closes in on the right-side boy, moving closer and closer to his eye. As this happens, a female narrator's voice is heard:

> Ahhh, life. A sunny day, a simple friend, and a complex biological story. The human genome project is a way of exploring our molecular selves. Almost all of our cells—the muscle cells that let us smile; the brain cells that perceive the humor in things; the cells of our eyes that take it all in—contain a complete set of all our genes: the genome.

Biology in this conception, intercalates our experience in complex ways. Friendship and humor, the best of human life, the external connections among us, are all mediated by our internal commonalities, "our molecular selves." That which is best about humans finds its core in that which we all share. The Human Genome Project provides a way to explore that which is at the core of our common humanity.

---

[2] One of the main points of this thesis is that "race," and racial categories like "African American" and "Caucasian" are historically specific ideas with shifting definitions. My use of this racialization is not intended to imply that these are stable, inherent components of the boys. I believe, however, that these two boys were chosen largely *because* of their "race," in order to represent "diversity," another racial term that I will discuss extensively. While the race of the boys is never stated in the video, I intend "African American" and "Caucasian" to be attributed to its makers. If this way of making race also marks me—indicates the ways in which I see racially— all the better to show the power of racial discourse, and to make my perspective explicit.

By the time the word “genome” is pronounced, the camera has zoomed in to an extreme close-up of the right-side boy’s eye. When “genome” is said, computer graphics commence, and at high speed, with light effects, we enter the internal, microscopic space of a cell, and then a nucleus. Three dimensional computer graphics illustrate the geography of sub-cellular biology.

The biology lesson that we enter into is not a description of this individual’s cellular anatomy. Nor is it a description of the humanity of this individual in molecular terms. The space is portrayed as *everyone’s* cellular anatomy, and *all of our* molecular selves. The phrasing is always plural, the claims universal:

> If we could journey inside ourselves—into a cell—we would see twenty-three pairs of chromosomes packed into a nucleus. Each chromosome contains a long coil of DNA. If all the chromosomes were unwound, the DNA in just one of our cells would stretch six feet long. The DNA double helix contains four kinds of building blocks—an A always pairs with a T, a C with a G. DNA contains information to make every part of our bodies with its four letter language. Each of our thousands of genes codes for a specific part. RNA polymerase copies the information in a gene into a messenger molecule: messenger RNA. The building blocks of messenger RNA and DNA are called bases. The bases on one strand of DNA specify the order of bases on the new strand of messenger RNA. The DNA always stays inside the nucleus, but messenger RNA travels out into the cytoplasm. There a protein-making machine called a ribosome can read messenger RNA to make a particular protein. Every three bases of the messenger RNA molecule codes for an amino acid. Proteins are made of amino acids. tRNA molecules help translate the language of DNA and RNA into the language of proteins. tRNA molecules bring the right amino acids that the ribosome links together to make a protein. Proteins are the laborers. Some form structures like tendons and hair—others perceive

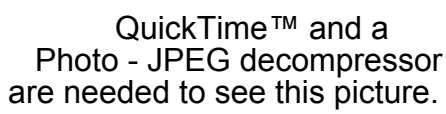
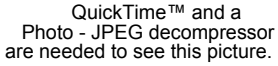


Figure I.1 Craig Venter of Celera Genomics, left, shakes hands with Francis Collins of NHGRI, right, at a ceremony at the White House, June 2000. Image Source: *Nature*: Mcilwain, 2000.

Figure I.2
Cover of Nature, February 15, 2001. The mosaic includes the faces of Mendel, Watson and the Beatles. Image source: *Nature:* Kemp, 2003.



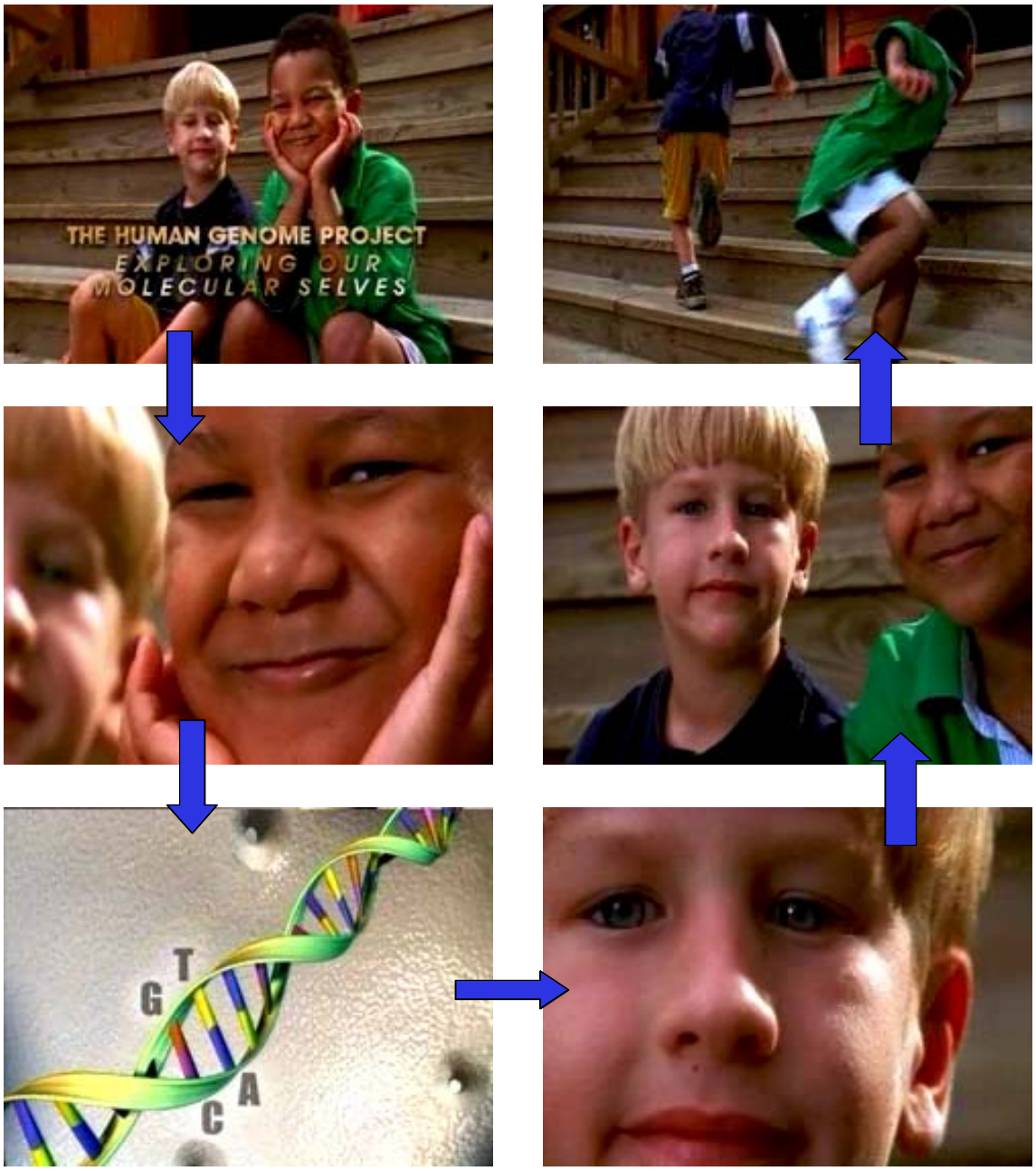


Figure I.3. Stills from "Exploring Our Molecular Selves," a film produced by NHGRI as part of a free educational toolkit for high school students. Arrows denote order of appearance. Source: NHGRI website, http://www.genome.gov/Pages/EducationKit/video/qt/3D.mov.

light, scents and flavors, control chemical reactions, and carry messages between cells.

This is the story of DNA in the cell. DNA makes RNA, RNA makes proteins. The process is illustrated in vivid detail. Throughout, DNA lies at the core, directing the labors of the proteins. These are autonomous structures and processes, which lie at the heart of human beings.

Having covered the mechanics, the video moves on to the Project. There is good news. The coded human commonality has been deciphered:

> To understand our molecular selves, scientists have read the three billion letters making up the DNA in the human genome. Different sets of genes, interacting with complex environmental factors, influence things like our looks, personalities, and risks for diseases like cancer and heart disease. [Here the camera starts to draw back from the cell, beginning the light-show again in reverse] A growing understanding of our genes and all they do will help us understand [the camera emerges, this time through the left-side boy's eye] the complexity and the wonder of life. A simple friendship—a sunny day—ahhh, life. [boys get up and run up steps.]

Everything has come full circle. Not only have we journeyed into one person, and seen the structure at the heart of our molecular selves represented there, by that individual. By making this internal journey, we have emerged again *through another person*. We have established a link not only to the molecules mediating our friendships, our humor, our sunny day. We have established a link between us and our friends—a link that transcends bodies, individuals. A link that transcends even race.

After the journey, the boys run up the stairs together, putting a finishing touch on these images of progress, friendship and humanity. There is the implication of growing up with DNA as a guide, a structuring component of development. Along with it there is the suggestion of human and scientific progress, moving upwards towards better things.

"Exploring Our Molecular Selves" is full of hope that is more than an awkwardness or naivete characteristic of high school instructional videos. The medical hopes are spelled out: that

> light, scents and flavors, control chemical reactions, and carry messages between cells.

knowledge of the genome will lead to better treatments for diseases like cancer and heart disease. The humanist hopes, I think, are more to the point. The uniting of people, the wonderment at seeing what we really are, true answers, real progress—these are the promises of the genome. And not-so-subtly, the video promises a new hope in the genome for an end to racism.

The dynamics of race in this video are not mere tokenism. The cooperation of HGP and Celera, the celebrated finding that 99.9% of the genome is universal, the uncovering of truths that link and produce our molecular selves—these were not the only unifications emphasized at the time of draft completion in 2000.

It was contended in many circles that the Human Genome Projects had made race itself obsolete. Craig Venter, in particular, has been an advocate of this position. In an interview with the newspaper The Australian, for example, he said "In the five Celera genomes, there is no way to tell one ethnicity from another," evidence that “the concept of race has no genetic or scientific basis.”(Wires, 2000)

It was a claim echoed on the basis of slightly different evidence from the Projects by President Clinton in his speech at the White House ceremony with Collins and Venter.

> I believe one of the great truths to emerge from this triumphant expedition inside the human genome is that in genetic terms, all human beings, regardless of race, are more than 99.9 percent the same. What that means is that modern science has confirmed what we first learned from ancient faiths. The most important fact of life on this Earth is our common humanity. My greatest wish on this day for the ages is that this incandescent truth will always guide our actions as we continue to march forth in this, the greatest age of discovery ever known. (Office of the Press Secretary, 2000)

The journey inside in “Exploring Our Molecular Selves,” and the “triumphant expedition inside the human genome,” situated in this “greatest age of discovery ever known,” are trips that bring

us back to ourselves.[3] Race, by the travels of this new age could be finally overcome. Unification, in these conceptions, was on the horizon.

II. Divisions

The story of race in human genomics, and the older story of race in biology, are not tales of racial innocence lost. Genomics has employed racial thinking before, during and since 2000, the year when anti-racist discourse in genomics reached its zenith. This is not to say that genomics has been *racist*, but that the use and reproduction of a system that divides people into biological units called "races" has been steady. At the same time, the values attached to the use of these categories, the meanings associated with them, and the ways in which they are produced have occupied a variety of positions.[4]

Even in the hopeful gestures of 2000, in which race is pronounced dead to science, the use of a racial system can be identified. I read"Exploring Our Molecular Selves" as making a strong use of race in the casting of the two boys. There is a molecular connection between the two, a common humanity mediating their friendship that is extended to all of us. The genome is positioned as something larger, more fundamentally important than race. But race remains as a structuring principle, even where humanity, as located in *our* genome, supercedes it.

Race is also used in the practice of genomics research, not just in the representation of this research and its implications. Both Celera and NHGRI used race in choosing the humans for their Human Genome Projects. The projects required DNA samples from donors in order to

[3] T.S. Eliot's Four Quarters is widely quoted in genomics literature. Francis Collins' Short Course lecture, excerpted above, is one example (2001): "We shall never cease from exploration, and at the end of all exploring, we shall arrive at where we started but know that place for the first time."

[4] In this introduction I make no distinction between "race" and "ethnicity" as categories. The use of these terms in genomics is compared and contrasted in the first two chapters.

carry out their mapping and sequencing endeavors. The groups collected their DNA in different ways.

Celera assembled its genome sequence using the DNA of five people. According to the New York Times, "Celera [...] sequenced the DNA of five individuals from different ethnic groups -- two Caucasian men and three women, one black, one Hispanic and one of Chinese descent -- though the principal sequence has been done on a single anonymous male. Call him Celer.[5]" (Wade, 2000) The consensus sequence, meaning the the final draft combining data into a unified reference version, for Celera was produced by a comparison and averaging of all of the participants, after the Celer sample was completed.

NHGRI was particularly concerned with preserving the anonymity of its donors. For this reason, existing libraries of DNA were excluded from the later stages of the project, because the donors' identities were widely known, and because prior informed consent could not be extended to a project with such far-reaching anticipated effects. The strategy they settled on involved taking out an ad in a newspaper in Buffalo New York for donors to contribute to "the worlds biggest science project—deciphering the genetic blueprint for human life." (Clark, 2000)

Many applicants volunteered, ten women and ten men were sampled anonymously, and one anonymous man's DNA was selected for initial seqencing. The Buffalo volunteer's DNA was used to make the draft version in 2000, but he was not intended as the only donor:

> Another project launched last year [also coordinated through NHGRI], a $48 million collaboration of pharmaceutical companies and the Wellcome Trust of London, will weave many more layers of genetic code from another group of 24 ethnically diverse volunteers into the final sequence, all to be accessed for free through GenBank to any researcher in the world. In the end, the threads of DNA from the Buffalo volunteer, whoever he is, will be enmeshed in the colorful tapestry of the human genome, mixed in and indistinguishable from the DNA from all the other donors. And it really won't matter,

[5] Rumors have circulated about the identity of Mr. Celer. Although unconfirmed, it is generally believed that Celer is J. Craig Venter himself. (e.g., pers. comm. Reardon,2003)

scientists say, because we aren't really all that different, after all. (Clark, 2000)

The subjects for both projects are identified in two ways in these two articles: gender and ethnicity. The Celera 5 are "Caucasian men," and "black," "Hispanic" and "of Chinese descent" women. Celer himself, is not identified "ethnically," although if he is Craig Venter, or if he is one of the five (the article is not clear on this point), he is "Caucasian." The group whose DNA will interweave with the Buffalo volunteer are "ethnically diverse." But the Buffalo volunteer—chosen to be a male because males have both X and Y chromosomes— is not identified in this way. We do not know explicitly if Celer and Buffalo-man were Caucasian, although in context this is implied by the "ethnic" identification of the other donors.[6] Ethnic "diversity," here indistinguishable from "racial" diversity, is an additive. It is incorporated pre-hoc in the Celera process, and post-hoc in HGP's.

Race occupied an important position, then, in the practices of the first genome projects, and in the discourse produced by them. While anonymity is maintained completely for the public project, and almost completely for the private, the identifying labels of sex and race remain, and constitute criteria for selection. Since 2000, there has not been a new resurgence of race into an existing vacuum. Rather, we can identify shifts in the emphasis and particular uses of race since that time. Race has come to occupy a more obvious and central position in successor projects to HGP, and in research enabled by its accomplishments. As this process has occurred, discourse on race has also shifted.

As a producer of widely circulated public discourse, Celera's prominence in the public sphere has been short-lived. After its success as a competitor with HGP, the company and its

[6] This implication depends in part on an idea of the "unmarked category," a feature of discourse in which white men are seen to occupy a default position, a mode of existence which is unmodified by race (or gender). This default status is a central component of power relations of white privilege in the U.S. (i.e., McIntosh, 1988; Haraway, 1988)

spokesman have faded from the headlines, and withdrawn somewhat from the sort of philosophical pronouncements represented in Venter's quotes above. The focus for Celera has been commercial, developing it's online genomics "Discovery System," and focusing, with its parent company Applera, on medical and biotechnical applications for its genomics database.

NHGRI, on the other hand, has continued to produce political and humanist statements, and has continued to fund Ethical Legal and Social Implications (ELSI) research. Discourse on humanity still characterizes public genomics' representation of its research, with race playing a leading role in many ways.

With the reference genome largely finished, a major area of investigation that NHGRI began to pursue was human genetic variation. Turning from the 99.9% of DNA that all humans shared in common to the 0.1% that made everyone different was a logical next step, particularly when this 0.1% contained the genetic factors that modified disease risk. The Project had been funded, after all, largely on the basis of its potential to revolutionize medicine. In order to begin producing medical results, difference, not overriding human unity, had to be broached. It has been in the context of variation research that race has been employed most commonly by genomics, and in writing about this research that genomics has made its most recent claims about the nature of race.

As with the original Human Genome Project, the aims of successor projects have been global in scale. Examples include the Haplotype[7] Map, and the SNP[8] Consortium's project,

[7] "Haplotype" refers to internal structures within genomes. It turns out that complex arrays of variants tend to travel in "blocks" on the genome, such that identifying a sample of the single base variants can allow researchers to infer the presence of the whole block. This simplified variation research immensely, and the HapMap project was set up to characterize these blocks. The results, it was claimed, would greatly facilitate disease association studies.

[8] SNP, pronounced "snip" stands for Single Nucleotide Polymorphism, and means a single location on the genome in which one base varies between individuals. The SNP Consortium's project was to characterize a large number of SNPs.

which is referenced in the quote above as the "$48 million collaboration of pharmaceutical companies and the Wellcome Trust of London" that introduced "diversity" to the original genome.(Clark, 2000) Research into variation has taken as its goal a total picture of human variation worldwide. Since it is impossible to sequence the genome of every human being, a process of representation has been required to meet the goal of globalism. There needed to be a system in which variation could be sampled in a way that everyone was included.

The solution was population-based sampling. "Populations," generally referred to as "geographically defined," were identified, sampled and sequenced. The subsequent data from the HapMap and SNP Consortium were categorized in terms of these "populations." For the SNP Consortium, and the first phase of the HapMap, the categories fell into three sets: African, Asian and European. In later phases, the HapMap has identified a larger number of groups, in smaller geographic focus.

Race enters into this process in a number of ways. Lundy Braun describes a cycle in which population-specific genetic variants, of the type produced by these studies, depend on racial thinking to make sense of their data:

> Although efforts to categorize human genetic diversity have historically relied on skin color, even skin color is not a marker for race. Increasingly, researchers are attempting to identify population-specific genetic variants. But to make sense of these variants, it is necessary to correlate gene frequencies with knowledge of human history, a history that is superficial and informed by Western assumptions about non-Western peoples. Identifying and interpreting frequency distributions of genes is both a scientific and a social process, guided by scientific and social assumptions. […] The meanings attached to race and ethnicity shape in a fundamental way experimental design, interpretation of scientific findings, training of health practitioners and biomedical researchers, and development of public health policy. (Braun, 2002)

Organizing knowledge about human variation among populations requires genetic markers to define these populations. But in order to find genetic markers, researchers must organize populations first using social categories. To study human genetic diversity—that is the "natural," biological distribution of genetic difference— scientists need historical, socially defined categories of difference that are most often rooted in histories Braun identifies as superficial and often imperial. "Race," a system that asserted a hierarchy among peoples, and helped to justify genocide and colonialism, participates in the production of knowledge about genetics. "Populations" in the *human species* depend on "races" in the *human.*

There are, of course, other ways of constructing populations. Occupation, location of residence, preferred hobbies and affinity for spicy food are potential examples. But it is not believed that sampling based on these sorts of categories would produce any sort of valuable *genomic* information about variation on a worldwide scale. These sorts of division are not seen as representative. Studying people who come from a variety of places on the globe[9] is seen as the way to produce a global analysis of genetic difference. This, in turn, requires the production of population categories, which depend on race in the ways that Braun outlines.

Two major processes of division are required for genomic studies of human variation to take place. Because there is too much DNA for it to be feasible or desirable to analyze every part of it, studies of many people must divide the billions of nucleotides in any given genome into that which is of interest, and that which can be ignored or assumed identical for everyone. Because every individual is genetically different, in order to produce a survey of difference, people must be divided into groups that can be

[9] "Come from," of course, is a term in need of definition, which does not apply to everyone equally, and depends on notions of race. Ancestry, origin, descent, geography are concepts with shifting definitions and usages in genomic discourse, and ideas of race more broadly. These issues will be discussed extensively in the coming chapters.

proxies for this difference. So divisions must take place both biochemically and socially.

The field of genetics has developed many types of systems for dividing and organizing DNA biochemically, providing a rich set of options for scientists doing this work. These systems include techniques, chemicals, machines, resources, discourse, informatics and institutions, all of which interrelate in complex ways. To give an extended example, polymerase chain reaction (PCR) refers to a process for selecting and "amplifying"—that is making a great number of copies of—a small piece of interesting DNA. This requires an expensive piece of equipment called a thermal cycler, which is usually purchased by institutions rather than individuals, often using money from grants that come from other institutions. Crucial chemicals used in the PCR process are custom designed to amplify only the DNA segment of interest. To order these chemicals, called primers, one needs to translate the DNA of interest into sequence, rendering the chemical structure of a particular piece of a deoxyribonucleic acid chain in the letters A, C, T and G. This can either be done using a database of already known sequence, such as Genbank, where one can access the HGP's data, or by sequencing a sample. Sequencing is a technique as complicated as PCR, that *involves* PCR, and can be done by the lab in question or a contracted third party. Once a sequence is produced, it can be mobilized electronically and interpreted in such a way that a company that manufactures oligonucleotide primers can use a complex set of machines, chemicals, techniques, resources, discourse and informatics to custom manufacture the materials that are required to amplify the interesting DNA. The original lab can then carry out PCR, study the DNA of interest, put their data in databases, attach meaning to this DNA, use it as evidence, and produce knowledge with it.

PCR is only one technique for genomic division available to genomics, and it is one that participates in hundreds of other processes. PCR is a complex organizational process, whose

result is differentiating the genome into a small area of interest and a vast area that can be ignored successfully. Everything in the genomics of variation, from a single PCR run to the institutional structure of the HapMap can be viewed in this way, as technologies for differentiating—and thereby organizing— the genome.

The emergent ways in which genomics differentiates and categorizes people do not involve this sort of technical complexity. These are social processes, whose complexity is more centered in words, assumptions and interactions.[10] The purpose of this thesis is to examine the ways in which the field of human genomics differentiates people along racial and ethnic lines. I refer to the processes by which it does so as a "technology of difference."

III. Contexts

In order to understand the relation of genomics to race, in order to characterize the genomic technology of difference, it is necessary to place them in intertwined historical, political and scientific contexts. I present a brief sketch of the history of race in the biological sciences, followed by a look at contemporary developments that relate genomics to broader social situations. This is a whirlwind tour, whose historiography has simplifications and omissions. The themes I wish to emphasize are the historical instability and multiplicity of biological "race" concepts, the relations between biology and state power, and the existence of a variety of racisms

[10] PCR also involves words, assumptions and interactions, and is a social process in the ways suggested in the example above. And categorization of people in genomics makes use of technologies that include PCR. I use the distinction between social and technical to indicate that it will be more productive to analyze the social and discursive components of racial categorization if we wish to understand the production of race. In the same way, an understanding of genomic organization begins most effectively, but does not end, with a technical approach.

and anti-racisms in the production of race. These themes will importantly inform an investigation of current racialized genomics.

"Race," as a Western system of thought that categorizes and organizes people, most often hierarchically, has been produced and employed by the life sciences since they were in their infancy. It has not been a matter of race exerting a biasing influence on otherwise rational scientists and natural philosophers. Race has not come in from "society" to "science," but has been produced and enforced by scientists *in* society.

> Racial thinking and scientific thinking, after all, grew together. Recent scholarship in the history of the idea of race has claimed that "race" is a recent development in Western thought. The notion that the human species can be divided into biologically discrete units and that these can be hierarchically arranged is a product of ideas that only came to fruition in the sixteenth century and gained power in the centuries since. Historians of science will note that the idea of race grew concomitantly with the growth of modern science in the Western tradition. The drive to categorize the new discoveries in the age of exploration and cultural notions of "savage" and "civilization" were enmeshed from the start. (Jackson, 2002)

This introduction to a compilation of essays from *Isis* and *Osiris* points to the early and fundamental engagement of racial thinking and the human sciences. Race and the human sciences grew as systems after Columbus in ways that were dependent on each other. Audrey Smedley has read this relationship as one of predication: "All of the human sciences—biology, psychology, anthropology, and sociology—in their early stages were predicated in some fashion on what is identified here as the racial worldview" (Smedley, 2002 in Stockwell, 2002). This worldview allowed and structured a host of knowledge-making practices in these sciences, whose results in turn, served to justify a host of racist practices (i.e. Stockwell, 2002).

In making this observation, we should avoid the implication that "*the* concept of race" was developed, and speak rather of multiple, shifting concepts of race (Reardon, 2004). Carolus Linneaus, considered the founder of biological taxonomy, published one of the earliest racial

categorical systems in the middle eighteenth century, placing Man into one species of four races. *Homo sapiens americanus* were "red," and "ill-tempered," *Homo sapiens europeaeus* were "white, serious and strong," the *asiaticus* race were "yellow, melancholy, and greedy," the *afer* "black, impassive and lazy" (Linneaus in Stockwell, 2002). A five category system was introduced by J.F. Blumenbach in 1781, consisting of "Caucasian," "Mongolian," "Ethiopian," "American" and "Malay" varieties. (Blumenbach in Stockwell, 2002) A polygenetic system, in which the categories were held to represent independently originated species was advanced in 1799 by Charles White. A three category polygenetic scheme, consisting of "Negroes," "Caucasians" and "Mongoloids" was posited by early genetics, anthropology and evolutionary theorists. Race, throughout its history, has been unstable, encompassing many concepts (Stockwell, 2002; Stepan, 1982). These schemata were never value-neutral, as is made obvious in Linnaeus' racial characterizations. Most often they made an explicit racial hierarchy, with European races, variously defined on top, African races, variously delineated, on the bottom (i.e. Stepan, 1982)

Darwin's approach to race was paternalistic and hierarchical, although he opposed genocide and slavery (Darwin, 1896). This passage from *The Descent of Man* demonstrates some of the complexity of Darwin's racial views, attributing criminal traits at times to the "civilized," and kindness to "savages," while maintaining the dichotomy between them:

> Slavery, although in some ways beneficial during ancient times, is a great crime; yet it was not so regarded until quite recently, even by the most civilized nations. And this was especially the case, because the slaves belonged in general to a race different from that of their masters. As barbarians do not regard the opinion of their women, wives are commonly treated like slaves. Most savages are utterly indifferent to the sufferings of strangers, or even delight in witnessing them. It is well known that the women and children of the North-American indians aided in torturing their enemies. [...] Nevertheless, besides the family affections, kindness is common, especially during sickness, between members of the same tribe, and is sometimes extended beyond these limits. Mungo Park's touching account of the kindness of the negro women of the interior to him is well known. (ibid.)

Not only did the specific categories of racial systems change throughout scientific history, but the meanings attributed to them changed as well.

Furthermore, the meanings attributed to Darwin's ideas changed, as they were mobilized for a great variety of political ends. An adequate history of the relations among Darwinian thought, social Darwinism, the eugenics movement, and Nazi *rassenhygiene* is out of the scope of this thesis. What is important to emphasize is that the idea of Darwinian evolution, has been unstable and contested to this day, as have its uses in the production of race.

Eugenics occupies an important position in any history of race and biology in the twentieth century.[11] The movement, based in part on Darwinist ideas, advocated the genetic "improvement" of the human species. Positive eugenics mandated that people of desirable traits breed with each other, negative eugenics mandated that "defective" people, bearing traits like "feeble-mindedness," "imbecility" and blindness either not reproduce or be sterilized (i.e. http://www.eugenicsarchive.org/eugenics/).

Eugenics' primary contribution to ideas of race was to produce a scientific platform against "race-mixing" in support of anti-marriage and anti-miscegenation laws in the U.S. Madison Grant wrote *The Passing of the Great Race* in 1916 as a eugenic polemic against miscegenation, warning that the "white race" was headed for destruction.

> The cross between a white man and an Indian is an Indian, the cross between a white man and a negro is a negro... When it becomes thoroughly understood that the children of mixed marriages between contrasted races belong to the lower type, the importance of transmitting in unimpaired purity the blood inheritance of ages will be appreciated at its full value. (Grant, 1916 in ibid.)

[11] In an interesting connection, a website dedicated to memorializing the abuses of eugenics, "Image Archive on the American Eugenics Movement," run by Cold Spring Harbor Laboratory (which had been the center of American eugenics) is funded by an ELSI grant from NHGRI http://www.eugenicsarchive.org/eugenics/)

President Calvin Coolidge, influenced by Grant, agreed, citing "biological laws" as prohibiting miscegenation. The authors of Cold Harbor's eugenics history website also cite Virginia's "Racial Integrity Act of 1924" as produced by eugenic activism. These relations between biology, race and the state should alert us to the depth of influence and power, and the complex relations that exist among these three.

German race science during the Third Reich illustrates this point further, and it drew direct influence from the American eugenics movement (i.e Allen, in Jackson, 2002). *Rassenhygiene* is eugencis that, besides improvement of the genetic stock of a race, aims at multiplying the numbers of the race (Weiss in Jackson, 2002). German eugenecists were not only Nazi, but included socialists and others on the left (ibid.). The roles that doctrines of *Rassenhygiene* played in the Holocaust are well-documented, and are roles to which I cannot do justice. Biological race participated directly in the deaths of many millions.

The Holocaust represented for the West a defining crisis of race. Human biology and anthropology felt themselves to be particularly in crisis (Reardon, 2001; pers. comm, 2003). The ways in which this crisis was negotiated drew on a recent tradition of anti-racist thinking about race in the human sciences. Meetings at the United Nations Educational, Scientific and Cultural Organization (UNESCO) in 1949 and 1951 brought together panels of experts in physical anthropology and population genetics. The statements that were published by UNESCO in 1950 and 1952 have been construed by many historians of science to constitute a crystallization of anti-racist thought in the sciences, and the turning point at which race was withdrawn from biology in favor of a population-based approach (UNESCO, 1951, 1952; Haraway, 1997; Stepan, 1982)

Jennifer Reardon has shown this narrative of the withdrawal of race to be inaccurate (Reardon, 2004). The "population-based approach" constituted the movement of race *into*

populations, not the *replacement* of race with populations. While liberal humanism, and anti-racism were dominant in expert discourses of race, the existence and usage of concepts of race remained prevalent.

One major turning point that the UNESCO Statements did represent, however, is a new sort of boundary work, in which experts were held to have the only legitimate access to the truth about race. Admissible criteria for race were increasinly placed inside bodies and inside populations, territories over which experts exercised a certain exclusivity. Racism, it was held, resulted from abuse of scientific findings by the public. It became the responsibility of scientists to educate the public about the proper uses of race in order to prevent misinterpretation. As this thesis will show, this inside-outside structuring of knowledge about race has a strong resonance with current racial discourses in genomics.

But even this subsumption of race into population, and the non-racist or anti-racist tone favored by the UNESCO Statements were not maintained stably during this period. Other concepts and uses of race were produced during the same period. In the American context, a simultaneous and overtly racist biological and medical project was undertaken, which has served to exemplify a certain kind of racial abuse (Shavers, 2001).

The Tuskegee experiments, run by the federal government's Centers for Disease Control (CDC) from 1932-1972 were designed to test the progression of untreated syphilis. In order to do so they told 399 syphilitic men, all raced as "Negro," that they had "bad blood" (their disease was never revealed to them), and that they were receiving curative injections, which did in fact exist at the time. They received only placebos. The researchers' interest was primarily post-mortem, and their subjects were allowed to die, usually slowly. This has stood as the exemplar of racist biomedical practice in the United States, and it remains the most prominent such example from the twentieth century. Its legacy is often cited in genomics ELSI literature

outlining ethical issues in racially structured testing.

Several more recent developments are also important in locating genomic racial thought in its broad social context. One important political and legal context is the enactment of the NIH Revitalization Act of 1993. This legislation required that "women and members of racial and ethnic minority groups be 'included as subjects' in each clinical study funded by the agency from 1995 onward."(Epstein 2002) It also required that each clinical trial be "designed and carried out in a manner sufficient to provide for a valid analysis of whether the variables being studied in the trial affect women or members of minority groups, as the case may be, differently than other subjects in the trial." (NIH 1993 in Epstein 2002) This legislation mandates racial thinking in medical research for explicitly anti-racist ends. It is particularly relevant for NHGRI as an institute in the NIH umbrella. As this thesis will show, research into the medical implications of genomics is an important driving factor the genomic production of race.

An important, and seemingly contradictory biological context is the claim that the majority—eighty-five percent—of genetic variation in humans occurs within, rather than among racial population groups. This conclusion, first reached by Lewontin in 1972 in his work with blood types, was confirmed by Barbujani in 1997, using 109 polymorphisms in 16 populations (Braun, 2002). The variation described by Lewontin and Barbujani occurs in only 0.1 percent of the genome; every individual is 99.9 percent genetically identical. This finding has led to the positions that race is biologically meaningless as Lewontin asserts, and that unity should be emphasized above racial difference, as Clinton suggests in his speech in 2000.

We can identify an important tension between the NIH Revitalization Act and Lewontin and Barbujani's finding. The latter is used to indicate that race should be de-emphasized or phased out. The former provides a set of legal and medical imperatives to study race in every clinical application.

Further complicating matters, a large body of research focused on the health effects of racism has shown that people of color in the United States—that is everyone who is raced as non-white—have worse health outcomes than people raced as white (i.e. Bullard, 1987). This situation is most commonly referred to as "racial disparities," or just "disparities" in health. The Department of Health and Human Services' Healthy People Initiatives were designed to eliminate these disparities, and NHGRI and NIH are participants. Racial health disparities and the Healthy People Initiatives aimed at addressing them contribute in complex ways to the genomic production of race (i.e. NHGRI, 2001).

A final, and vital development that frames the genomic production of race is the legacy of the Human Genome Diversity Project (HGDP). The Diversity Project was conceived by Luca Cavalli-Sforza, a world expert in population genetics, in 1989. It received support from the leaders of the early Human Genome Project, after Cavallli-Sforza responded to their concerns that it would stoke a new racism by arguing the converse, that HGDP would undermine racism. The Project's goal is summarized by its organizers:

> The HGD Project is an effort by anthropologists, geneticists, doctors, linguists, and other scholars from around the world to document the genetic variation of the human species worldwide. This scientific endeavor is designed to collect information on human genome variation to help us understand the genetic makeup of all of humanity and not just some of its parts. The information will also be used to learn about human biological history, the biological relationships among different human groups, and may be useful in understanding the causes of and determining the treatment of particular human diseases (HGDP North American Committee, 1993-4).

In order to pursue this goal, the organizers planned to collect hair-root, white blood cell, and cheek tissue samples from more than 700 groups of "indigenous peoples" worldwide

(Haraway, 1997). Genomic data from these samples could be analyzed and used to build knowledge about human evolution and human difference.

The Project lost funding after massive pressure from indigenous political activists as well as criticism from within the scientific community. A number of arguments were mobilized in opposition. First, HGDP positioned its subjects as "vanishing populations," meaning that they were nearing extinction, and the genetic information they harbored could be lost forever. This idea that their cultures were "vanishing" was not well taken by those studied. "The Vampire Project," as it was dubbed in reference to its blood collection, did nothing to benefit the groups studied medically or otherwise. Indigenous involvement, beyond the level of sample donation was almost absent, and concerns about commercialization, biopiracy and even ethno-specific biowarfare were high. Scientific criticism focused in large part on the definition of "populations" by their "isolation" which was viewed by many as inaccurate and unscientific. As a result, HGDP failed while HGP thrived (e.g. Haraway, 1997; Reardon, 2001; HGDP NAC 1993-1994)

Because HGDP failed, and because my focus is more recent, this thesis does not take the discourse produced by the Diversity Project into account. However, the legacy of the Project has ramifications, I think, throughout racialized genomics. This is especially evidenced in discussions of group consent in ELSI and agency literature from NHGRI. I believe it also informs the justifications that current genomicists use in pursuing racially and ethnically based sampling approaches, as with the HapMap Project. As we will see, the current justificatory frameworks strongly emphasize medical benefit to the people sampled. And while "diversity" is a highly valued item, in sample-based research (as opposed to internal affirmative action-style inclusion programs), it is seldom posited as an end-in-itself.

Here, then, is a sketch of the historical and contemporary contexts in which racialized genomics is operating. Instability, an engagement with the state, a variety of evaluations of race, and a complex set of legacies characterize this history. How, then, does a technology of racial difference exist and operate within the context of these legacies and tensions?

The response has been heterogeneous and internally contradictory. In many instances, the utility, reality and accessibility of race categories for genetic research have been contested in recent years. Most of this criticism operates on a level that shares terminology and particular discursive patterns with the work it criticizes. Although disagreements on emphasis and content are common, proponents and opponents of racial categorization in genomics share elements of a technology of difference. This thesis theorizes a contemporary emergent mainstream, while acknowledging and emphasizing the role of its internal contradictions and controversies.

IV: Materials and Methods

Syllogistic logic finds it central aphorism in the statements: A) All men are mortal, B) Socrates is a man, C) Socrates is mortal. This is the transitive property, such that if X = Y, and Y = Z, than X = Z. In analyzing the production of race in genomics, and the epistemology mediating that production, syllogistic logic will get us nowhere.[12] Race in genomics has a logic all its own.

My approach is a formal analysis of discourse, whose patterns point to elements of epistemology. Discourse, defined as the system of meanings through which words and knowledge signify and are communicated, is the most readily accessible arena in which the production of race can be examined. By observing the forms of discourse, it is possible to theorize the epistemological structuring principles involved in making knowledge. Discourse is

[12] Credit to Chris Amirault for suggesting this introduction (Amirault, pers. comm. 2003).

not itself epistemology, but helps to structures scientific knowledge in complex, fundamental ways.

Genomic discourse is literally quite accessible. Literature of a wide variety is freely available electronically—indeed free distribution is an important aspect of the culture of public sector genomics, and most of the sources of this thesis were readily found online.

Besides accessibility, I chose discourse as a subject because it comprises a great bulk of knowledge-making tools for science, informing the techniques, equipment and social interactions that comprise scientific life. An extended example, distinct from race, can demonstrate some of the ways in which discourse works as "tools" in this sense. To return to PCR, if I say, "*Thermus aquaticus* polymerase does not denature in a thermal cycler, allowing for PCR replication of DNA," I am using language designed to engage with epistemology. *"Thermus aquaticus" (Taq)* is a species of bacteria found in hot springs at Yellowstone. Using its name, I give information about its taxonomy, a system of categorization that relates to evolutionary theory and morphology, two systems of epistemological systems. "Polymerase" is an enzyme that synthesizes double-stranded DNA from single-stranded DNA. The "-ase" suffix tells me that it is an enzyme, knowledge about DNA tells me what it does, both operating in a specific system of biochemistry that organizes what I know. A "thermal cycler" is a machine designed to accomplish the "polymerase chain reaction" (PCR). The machine creates heat environments at timed intervals to facilitate the chemical reaction that "amplifies" DNA such that the machine plus the process cause a particular strand of nucleotides to be copied exponentially. *Taq*'s polymerase was chosen for this process because it doesn't break down in heat, because it lives in hot springs.[13]

[13] This "because," of course bears a fundamental relation to evolutionary theory, illustrating another way in which language habits and knowledge structures engage with each other.

Discourse in this example structures an organism (*Taq)*, a chemical (polymerase), a machine (the thermal cycler) and a technique (PCR). It also structures me, and my biologist friends, by providing us with a vocabulary in common designed to indicate taxonomies and epistemologies. The "-ase" suffix is a particularly good example—my friends and I can tell that polymerase, hydrogenase, RNAse and topoisomerase are enzymes, and that dynein, bradykinin, luciferin and melanin are proteins.

This example focuses on the context, definitions and epistemological valence of particular terms. This is the approach of the first chapter. By examining the words and categories deployed by genomicists when they are making racial and ethnic determinations, we can begin an inquiry into the conditions of those differentiations.

Terminology is interesting too, because little controversy has emerged around its particulars—the debate has focused mainly on race itself as a biological category or non-category, and less on the forms which scientists have used for race.

This thesis illlustrates genomics using categories of race and ethnicity which have distinct, and multiple contexts and meanings. These categories are inseparable from historical concepts of race and ethnicity, but operate in particular ways within genomics. These are the building blocks which genomics is using to make race. "The meanings attached to race and ethnicity shape in a fundamental way experimental design, interpretation of scientific findings, training of health practicioners and biomedical researchers, and development of public health policy."(Braun, 2002) This is an effort to theorize and locate these meanings, acknowledging that internal contradictions and disagreements are the rule and not the exception in this emerging literature. Even within this instability, we can identify trends forming, fading and clashing. The point is to investigate a set of categories and keywords that will be seen in operation in the second chapter as the formal configurations "nested proxies" and "perspectival phasing." These

are particular patterns of instability which this thesis uses both as evidence that genomics is making race, and as characterizations of the ways in which this is done.

The fourth chapter is a characterization of the genomic epistemology of race. I argue that both novel and historic means of structuring racial knowledge are used and produced by genomics.

The discourse in question comes from the writings of genomicists in peer-reviewed articles both print and electronic, committee reports, websites, multimedia presentations and films, testimony, requests for applications and the popular press. These were arrived at by a variety of mainly internet-based search techniques, including Pubmed and Lexis-Nexis database searches, and extensive exploration in genomic websites. NHGRI's www.genome.gov has been especially useful. A variety of secondary sources within scientific publications, the popular press, literary theory, history of science, and science studies theory have also informed this inquiry.

There is no strong focus on technology and technique, political economy, laboratory culture, ethnography of genomicists, in-depth historical analysis or comparisons to other sciences. The absence of these types of inquiry is problematic, but is meant in no way to devalue them. Although discourse has been an especially productive site for this research, there is no intended assertion that discourse is the most important, most indicative level on which scientific epistemology operates, and there is certainly no implication intended that science is *purely* discursive. Science studies is often accused of taking this last position, and although I have encountered very little work that actually does so, I want to clarify that it is far from my position in this paper.

One difficulty in pursuing research that aims to theorize an aspect of the epistemology of an entire field is the tension between generalization and embodiment. Describing sweeping

theoretical objects like “technologies of difference” risks glossing over complex social and political locations of individual scientists, who differ in complex ways from each other. In turn, performing a detailed ethnography of particular genomicists risks de-emphasizing theoretical tools which may be useful in working towards anti-racist science. In this paper I use quotes from individual, located authors to typify aspects I observe of a new, variegated genomic attitude towards race. I attempt to fully contextualize passages, and avoid the implication that genomics is somehow monolithic, or performed in an abstract stratosphere devoid of subjects. Still, I attempt to draw preliminary conclusions that generalize. I intend this with respect and concern for the authors I quote.

I should say that unless otherwise noted, no author’s work is included here for the sake of a broad dismissal as *racist*. Racism is neither a precise enough, nor an expansive enough concept to use in examining what appears to be a new area in the production of *race*.

Chapter 1: Categories and Keywords in the Genomics of Race

I. Transferals

"Race" is not a preferred term in genomic literatures. While it is used—often, in some areas of study and by some authors—it is used generally with a sense of caution, trepidation or at least complication. It is seldom expressed as a given and unproblematic category, particularly not in American public sector literature.[14] Francis Collins, director of the National Human Genome Research Institute (NHGRI), conveys some of this tone in the transcript to his Genomics Short Course presented in 2001:

> So again, this is a bit of a difficult message, I suspect, for you to convey to students; it's hard for me to get this exactly right - both, to say two things are true. First of all, that it is improper to draw precise boundaries around any particular ethnic or racial group and say they're biologically different. Science won't support that. But at the same time, it is correct to say there may be differences in frequencies, of a quantitative sort, of particular disease-susceptibility genes in one group compared to the other, that may play a role in why there are health disparities. Both those statements are true, and yet it is I think fairly challenging to get that message across in a fashion that is convincing to people, and it doesn't seem paradoxical.
>
> Well, let me go on to: Where are we going anyway, with genomics? (Collins, 2001)

[14] Some medical literature that uses information from genomics, and is often cited by genomics literature, does not reflect this sort of caution.(i.e. Yancy, 2001; Exner, 2001). I will discuss some of these examples further in later chapters. However, within genomics, even the strongest construction of the reality of race that I have seen, in a paper by Neil Risch, et al (2002), devotes itself to an explanation of the validity of race. That racial validity is seen as requiring an explanation indicates the extent to which race is treated as problematic in genomic discourse, and begins to indicate that genomics is producing race anew, rather than simply deploying it.

In fairness to Dr. Collins, this is the transcript from a spoken presentation, rather than a written passage. Still, in sections of the talk which bear no relation to race, Collins' language is more direct, more certain and less hesitant. The caution represented here is visible in the professional and agency literature as well, to varying degrees. Saying "race" is not quite taboo, but not quite casual or common.

Much of the terminology used in the genomic discourse around race can be viewed as a means to negotiate this discomfort with the term "race." One effect of this attitude—or at least one indication of it—is that other terms are more commonly deployed in "race's" place. For example, in this and the surrounding paragraphs, "race," is used twice, and "group" is used nine times to replace it. "Group" stands in for "race" as a kind of euphemism. The most commonly deployed euphemism for "race" in genomic literature is "population." Other euphemisms include "community," "descent," "ancestry," and "geographic origin."

As this chapter will show, euphemisms here are not quite synonyms, but encompass a more productive, complex set of discursive activities that allow shifting emphases, expansion, contraction and substitution. Population, community and group, for example, tend to expand the connotations of "race" when they act as euphemisms. They allow for categorizations that include a broader and more nebulous set of attributes and boundaries than "race." "Geographic origin" or "ancestry" can also euphemize race, but in a way that narrows its connotations, specifying the criteria of categorization to include relationships of parentage, location, history and diaspora.

This chapter will also show that none of this replacement occurs in a stable sort of fashion—even within terms like "ancestry" that tend to narrow meanings, meanings are apt to shift. Terms have multiple meanings in the genomic discourse of race. Very often

they refer to both political and biological knowledge at the same time. "Diversity," for example, can refer both to a political value, drawing from liberal humanist discourse, or to the abundance of genetic variation. This instability of meaning allows a great deal of work to be done towards the production of racial difference.

Of particular concern in this chapter is the use of terms as transfer points between biological and social knowledge. It is not only that they have implicit connotations that are both biological and political, not that they are just coincidental double entendres, but that they serve as transfer points that link genomic claims about political and social situations with claims about molecular and population genetics. Replacements of the term "race" are not only euphemistic concessions to a kind of squeamishness, or "political correctness," they are sites at which tensions among different sorts of knowledges are resolved.

Collins, for example, identifies a tension between the existence of racial health disparities in the United States—and NHGRI's resolve to alleviate them—and his claim that to "draw precise boundaries around any particular ethnic or racial group" is unscientific. At first glance, the issue of racial health disparities could be described as social or political, related to social stratification and racism. The issue of unscientific racial boundaries could be described similarly as biological, related to the genetics of populations and to biological knowledge about such genetics.

But this distinction between biological and social sorts of knowledge in this passage cannot be effectively made. Complex negotiations between claims about types of knowledge traditionally understood as "biological" and "social" are made throughout the passage. These negotiations have two points at their core:

> First of all [...] it is improper to draw precise boundaries around any particular ethnic or racial group and say they're biologically different.

Science won't support that.

This first point is a claim about what is appropriate to biological knowledge practice. But the claim points to itself as a social issue in several ways. It identifies a discursive activity that biologists should not participate in: "draw[ing] precise boundaries." Doing so implies a social and a discursive engagement with biology, a particular mode of classifying that is warned against, because "science won't support that." Collins could have said "boundaries don't exist," or "boundaries are not biological," but in negating a particular social and discursive action, he instead voiced "science" as something involved potentially in the support these sorts of actions. "Ethnic or racial group[s]" should not be precisely bounded, but their existence is at least social and political—they exist as something that *could* be bounded, biologically. These cannot be categorized as claims that are political, social, discursive on one hand, and biological, scientific, natural on the other. They occupy dual spaces, and the terms "ethnic or racial group," and "precise boundaries" are transfer points between these sorts of claims. The transfers continue in the following sentence, with different results:

> But at the same time, it is correct to say there may be differences in frequencies, of a quantitative sort, of particular disease-susceptibility genes in one group compared to the other, that may play a role in why there are health disparities.

"Group" here euphemizes "racial and ethnic group." Biological differences may exist, within the criteria Collins sets up as "frequencies," between these groups, whose medical situation is itself political, and is referenced immediately as "disparities." "Disparities" as a term transfers meanings between biomedical claims about health and sociopolitical claims about health stratification.

Together these sentences do a variety of work towards the production of race. Through euphemism and implicit transferal of meanings between the "social" and the

"biological," terms with many valences help negotiate complex tensions. The result is the production of a certain way of knowing race. Race cannot be bounded strictly, but gene frequency can differ among races. Races don't have proper edges, races don't have genes all to themselves, but the amount of some genes in races can differ, playing a role in medical disparities.

As this chapter will show, terms like "population," "diversity, "minority" and others can be equally powerful in contributing to the production of race in some of the same ways cited above: by transfering between different sorts of knowledge claims, by having multiple valences and definitions within discourse, and by euphemizing other categories including "race" itself. These types of terminological instability are central elements in the technology of difference.

II. "Race" and "Ethnicity"

"Race" as a stand-alone term, an unmodified noun, appears very rarely in the literature. "Racial groups" is more common than "races," and employing specific geographically-based names for races like "African," "Asian" and "European," is more common still. This pattern of usage contributes to the tone of trepidation. "Race," when it is used by genomicists as a term is mainly used in three senses: as an independent variable with limits, as a controversial "concept" whose meanings differ between science and society, and as a misconception.

When genomicists use "race" in its sense as an independent variable, they use it to structure experimental research on the distribution of genetic variation. Such research is usually medical in focus, or has medical benefit as its long-term goal, as with the HapMap's aim of facilitating future disease association studies. A medically oriented

genomicist might use "race" as a term in establishing his variables and study groups with respect to a particular disease. HapMap collected sequence data on single nucleotide polymorphisms (SNPs) by sampling three racial groups: African, European and Asian. This racial approach has been controversial, and has usually been described as being useful, but limited (Couzin, 2002; Foster, 2001).

Morris Foster, professor of anthropology at the University of Oklahoma and editor of American Indian Quarterly, has written ELSI papers that identify these limits, and describe race as a "heuristic starting point" for variation research. The term "race" is positioned as a classificatory tool for researchers with limited, but significant usefulness:

> Categories such as race and ethnicity are useful as heuristic starting-points for the investigation of biological relatedness. In attempting to approximate the range of human genetic variation, social categories such as these offer a practical way of approaching the study of that diversity and of defining inclusive criteria for participant recruitment. It is this practical consideration and the imperfect (sometimes very imperfect) fit between social and genetic definitions of a population that give rise to a series of scientific and ethical issues in contemplating the development of a human haplotype map with identified populations. (NHGRI: Foster, 2001.)

The "heuristic starting point" sense of "race" described here defines race with respect to science, and puts the term "race" in a position that negotiates social and biological knowledge. Race is identified as a social category that is a practical interpretive tool from which to launch investigations of biological difference. Race is not itself biological difference, but biological difference may be accessed through the social and discursive lens of racial categories. "Race," in this sense of the term, is a tool at variation researchers' disposal.

While the use of racialized categories as tools and heuristics has been frequent in the genome projects, the term "race" is seldom employed in their description. It is much

more commonly found in discussions of the implications and legitimacy of such projects, with the projects themselves most often calling their racialized variable groups by geographic or ancestral names. So the practices represented in the "heuristic starting point" sense of race are common, but the "heuristic starting point" sense of the specific term "race" is unusual.

Use of the term "race" as a "concept" inviting pronouncement and controversy has been less restrained. When genomic authors describe race as a "concept" or "notion," most often they mean traditional concepts that are located in society rather than science. Often they cite the potential of genomics to weaken these popular concepts or "constructions." "Race" here is a static conceptual tradition outside of science that can be cleared up by genomic data with proper education:

> Several speakers suggested that the information about different racial groups might in fact be helpful in showing how much of the genetic make-up is in fact held in common. "If we educate the public properly it will undermine the traditional concepts of race, and race will begin to fall by the wayside in the traditional sense," said Dr. Phyllis Epps, a health lawyer at the University of Houston. Dr. Lander said, "I think these data have tremendous potential to deconstruct simplistic notions of race and ethnicity." (New York Times: Wade, 2001)

"Traditional concepts" and "simplistic notions" are located with the public. The public possesses a way of thinking about race that is simplistic and traditional, and takes the form of discrete concepts. There is a sense that these ideas of race are "myths," old, stable stories that have been handed down and can be debunked.[15]

[15] Francis Collins uses "race" most often in this sense as a concept that exists with the public and engages with genomic knowledge in ways that require careful mediation by scientists. In his article in *Nature* on the final completion of HGP, he writes "Race is a largely non-biological concept confounded by misunderstanding and a long history of prejudice. The relationship of genomics to the concepts of race and ethnicity has to be considered within complex historical and social contexts."(Collins, 2003) Genomics here

That these concepts are "traditional" and "simplistic" implies that new complex notions are possible, using these data. The old problematic or illegitimate *beliefs* can be swept away by new *knowledge*, given the proper educational treatment by genomics. This new knowledge about race can draw from factual, natural, real entities of human difference and similarity, sweeping away the myths.

The ways in which genomicists are negotiating "race" as "concept," located with the public, with scientific claims about racial difference are at the crux of the production of race in genomics. The dynamic between that which is cast as social construction and that which is cast as biological reality takes complex forms in which these scientists make truth claims that are not separated by this distinction. Claims are made about the social as much as the biological in questions of "race," and the term "race" in its sense as "concept" emerges as one transfer point among these spheres, in which any such line is transgressed and reconfigured.

The third common usage—less common now than the two others— of "race" identifies it as a misconception. It describes "race" as entirely social myth with no existence in biological reality. It is not just the separation of simplistic notions and enlightened notions, but a truth claim based in biological knowledge about human difference that race is *only* mythical. Race here is constituted as an "illusion," a "social construction," without basis in natural reality (i.e. PBS, Race., 2002; Syracuse, 2002).

The truth claims employed in making this argument include the prevalence of admixture (a term indicating interbreeding between racialized groups), the predominance of individual diversity, and the greater proportion of intra-group rather than inter-group diversity. Admixture is deployed to argue that race boundaries are too fluid to be

---

is something that relates to outside notions of race as concept, but bears no culpability in their production.

biologically legitimate (i.e., Disotell, 2000).[16] Individual diversity can be positioned as so vast and complex as to overwhelm the imposition of biological races (i.e., Syracuse, 2002). Finally, genetic diversity within groups is held to far outweigh genetic diversity among groups so that the definition of such groups is damaging and unscientific (i.e., Lewontin, 1972; Barbujani 1997).

At first glance, the appearance of these types of anti-race critiques appears to frustrate an attempt to theorize a mainstream of genomic ideas about race and ethnicity—they simply appear contradictory. It is my contention that they are contradictory on significant levels, but that they share a terminology, a set of discursive patterns, and a certain epistemology that allow them to resolve such contradictions, and unite them in making race.

Even when the term race is used as a "misconception," race is configured in new ways with respect to genomic knowledge. Race is *produced*, as an entity that is purely mythical and controverted by this expert discourse. Race is made by genomicists into something new which is not genomic.

The term "race" in genomic literature has a distinct set of meanings: an almost taboo utterance; a tool of significant, if limited use; a concept differently positioned for genomics and "the public;" and a misconception. "Ethnicity" has a parallel place as a term. "Ethnicity" occurs most commonly paired in the phrases "race and ethnicity" or "racial and ethnic." Here the two categories are not quite synonymous, but very few authors state the difference between them explicitly when they write these terms. An example comes from the May 2001 edition of the *New England Journal of Medicine*. Two articles in that issue examined differences in cardiac drug response among black

[16] Other authors, including Neil Risch (2002) and Francis Collins (2001, 2003) identify admixture, but position it as complicating, rather than undermining the existence of race.

Americans and white Americans (Yancy, 2001; Exner, 2001). NEJM also carried editorials both for (Wood, 2001) and against (Schwartz, 2001) this way of applying race in medical research. Wood's article "Race Differences in the Response to Drugs—Pointers to Genetic Difference" does so in genomic terms, citing differences in allelic frequency. As with Collins' "differences in frequency of a quantitative sort," the ways in which "racial and ethnic groups" in Wood differ is by the frequency with which a given genetic trait occurs, rather than by having unique genes. This criterion of "racial and ethnic" differentiation is key in a wide range of new genomic research into human genetic variation, most prominently the Haplotype Map.

In presenting his views on the subject, Wood exemplifies the paired usage of "racial and ethnic," highlighted here in italics:

> Genetic differences among *racial and ethnic groups* usually reflect differences in the distribution of polymorphic traits, which occur at different frequencies in different populations, rather than a trait unique to a particular *racial or ethnic group*. The underlying genetic determinants of the response to a drug and, specifically, *racial and ethnic differences* in that response are beginning to be unraveled and do indeed appear to be based on the varying distributions of polymorphisms in drug receptors or drug-metabolizing enzymes among different *racial and ethnic groups*. Persons whose genes encode drug-metabolizing enzymes that lack normal activity have impaired metabolism of the drug substrates of the affected enzyme. Polymorphisms of the enzymes responsible for drug metabolism are distributed differently among different *racial and ethnic groups*, so the proportion of people with impaired metabolism differs among these groups. (Wood, 2001, emphases mine)

Normalcy and impairment of drug response is located here in polymorphic enzymes. Genes become "quasi-pathogens" altering normal reactions to pharmaceuticals (Yoxen, 1984). The frequency of quasi-pathogens—the frequency of impairment vs. normalcy—

maps to "racial and ethnic groups." Frequency, in this way, could make one "racial and ethnic group" more normal than another.

In Wood's piece, *racial and ethnic* is a slurred reference. Ethnicity and race are linked concepts with an equalized accountability for impairment and an equalized definition. Nowhere in the article does the author distinguish between "racial" and "ethnic," although in several places "racial" appears alone, serving an apparently synonymous function as the two categories used jointly ("physicians lacked guidance as to the optimal therapy for patients of different racial backgrounds;" "Any racial difference in the ability to tolerate carvedilol;" "racial differences in the response to drugs," etc).

This slurring allows his points to be extended to both terms, implying an equal and homologous sort of difference. Differentiations can occur through the mediation of "frequency," and the stakes can be as high as normalcy and impairment. The slurring of the categories leaves an open-ended ground for making these differentiations. The combination of "race and ethnicity" with "frequency" appropriates a great deal of space for the production and description of racial difference.

In this context, Wood's editorial is a particularly strong example of "race and ethnicity" used and advocated as an independent variable to structure variation research in a medical context with genomic influence. The piece is unusual in representing none of the hesitance in using "race" that is characteristic of most literature closer to primary genomics.[17] Its use of "frequency" differences, and its elision of "race and ethnicity," however, is strongly representative.

[17] That is, literature that communicates new knowledge from genomic science, rather than making use of this knowledge for other purposes, as Wood does.

"Ethnicity," then, runs a parallel role with "race" in structuring variables in research. "Ethnicity," though, does not occur in the "misconception" context described for race above. "Race" is a misconception for some, but ethnicity exists. Schwartz (2001), who writes that "race is a social construct, not a scientific classification," and that "a racial designation in the context of medical management not only defies everything we have learned from biology, genetics, and history but also opens the door to inequities in medical care," writes in his response to the race articles in NEJM:

> "Some geographically or culturally isolated populations can properly be studied for genetic influences on physiological phenomena or diseases. The Pima Indians, who have unusual susceptibility to non-insulin-dependent diabetes mellitus, and the people of Gambia, in whom polymorphisms in the *NRAMP1* gene influence susceptibility to tuberculosis, are examples. But even these cases are complex, since non-genetic factors influence the outcome. Among these many factors is culture: for instance, the germline *BRCA1* mutations that render Ashkenazi women susceptible to breast cancer owe their prevalence in that population the to fact that for countless generations, Jews have married Jews. (Schwartz, 2001)

"Ethnicity," as defined by geographic and cultural isolation, is within the realm of proper medical genetic inquiry. "Prevalence," here, like frequency, is used as a criterion of difference, although the *NRAMP1* gene polymorphism is voiced as almost unique to Gambia. But "ethnicity," as defined by isolation that alters prevalence, is advocated as a legitimate object of study.

Interestingly, the criteria that legitimize ethnicity for Schwartz, also exclude "race" from medical legitimacy. Black people in North America, he continues, are not genetically defined in this sense "as a unique population or race," because they have none of these elements of isolation. The carvedilol and enalapril studies in NEJM are therefore without validity.

The *proper* operation of ethnicity as an independent variable and biological object of study, is used to cast "race" as a misconception. Witzig echoes this position in his paper for the American College of Physicians "The Medicalization of Race: Scientific Legitimization of a Flawed Social Construct" (1996), advocating for ethnicity in his critique of race in medical usage:

> In contrast [to race], ethnicity is a concept that incorporates social, religious, linguistic, dietary, and other variables to identify individual persons and populations. Ethnicity may be able to impart clinical clues to diagnosis if the clinician taking the history is well informed and open minded. Ethnic boundaries are dynamic and imprecise, and a strict methodological approach to ethnicity that is equal to the approach required for the study of other variables is necessary if the concept of ethnicity is to be clinically useful. (Witzig, 1996)

This might be called the "heuristic starting point" definition of "ethnicity," in which it is a "concept" that serves as a useful way to approach biological difference if the researcher is "well-informed and open minded." "Ethnicity" here is certainly not the misconception that "race" is, but must be understood to have limits.

"Ethnicity" in Witzig's sense, is not commonly used in recent genomic literature. It may be that the failure of the Human Genome Diversity Project has made "isolated populations" an unpopular target, and has stimulated a turn away from ethnicity and towards race (Reardon, pers. comm. 2003). Whether or not this is the cause, the slurred reference "race and ethnicity," is by far the most common current usage of "ethnicity."

Witzig, Wood and Schwartz all have medicine as their primary focus, but the last two are particularly influenced by genomics. Wood uses differential allelic frequency and evidence from molecular genetics to help his argument for the validity of "race and ethnicity." Schwartz uses genomic evidence to support "ethnicity" and reject "race:"

> The publication of the first draft of the human genome should force an end

> to medical research that is arbitrarily based on race. The Human Genome Project now gives us the power to uncover the true origins of genetic variations; linking them to race has become passè. And instead of using polymorphisms to seek racial distinctions, we can spark real progress in clinical resarch by using genetic variations to track down clinically relevant alleles and pathogenic mutations. (Schwartz, 2001)

Schwartz's passage here illustrates a way in which genomics is making race, and being deployed in determinations about race. The newly available ideas and technologies of genomics can "give us the power to uncover the true origins of genetic variations," separable finally from "race," which genomics has made "passé." Genomics is positioned as a new and superior arbiter of true origins in knowledge about human difference—a position of great power in a medical context. The Human Genome Project has made racial links passé, and medicine can make better progress without it. Race is a misconception because of genomics.

Besides such positionings of the Human Genome Project as racial arbiter, there is an interaction of terminology and knowledge among genomics, medicine, federal science agencies and "the public." This interaction is forming and reinforcing the basis of a mainstream whose discourse delimits and structures ideas about human difference. "Race," "ethnicity" and "race-and-ethnicity" are terms with particular meanings in this discourse. These terms are used as tools, as "concepts," for "race" as a misconception, and for "ethnicity" to support the assertion that "race" is a misconception. Genomic knowledge and terminology can be used for purposes that seem at odds with one another. Small elisions like "race and ethnicity" can do a large amount of work. Terms can be shared, swapped and changed.

There is a range of discussion in the literature that uses these terms, but the terms themselves are defined within particular, if multiple parameters. By a variety of

transferals, including interactions among various spheres of knowledge production inside and outside science, dual and multiple definitions of terminology, and between the biological and the social, a genomic technology of racial and ethnic differentiation emerges.

III. Populations, Groups and Communities

"Population" is a key term in genomic writing. It is used when making technical claims about genomics, and in discussions and claims about the social and political implications of the field. In this way, "population" has both biological and social meanings.[18]

"Population," as this section will show, is also a term through which biological and social meanings of "race" are transferred and combined. Not only is "population" used in discourse that does different things, from agency literature on ELSI issues, to peer-reviewed articles on genomic sequence, but "population" links and transfers among different sorts of meanings within this discourse. "Population" can be a euphemism for race, but more often functions as a transfer point between social and biological knowledge about race.

[18] This section mainly addresses the use and strengths of the *term* "population," rather than the process of constructing and bounding populations in science. Lisa Gannett, a philosopher of science, has written extensively on this subject, saying "Populations are made pragmatically and variably constituted in different sorts of investigations of species genome diversity. Genes become bounded in space and time in ways that fulfill aims, interests, and values associated with particular explanatory contexts." (Gannett, 2000) The bounding of genes in space and time is fundamental to the process of race production in genomics, and has a remarkable flexibility in application. However, "population," as this section shows, defines and bounds a variety of things in addition to genes, giving the term a far-reaching and unstable sort of strength.

"Communities" and "groups" carry out similar discursive functions, and overlap in meaning with "population." They have emphases that can differ importantly. As terms, "community," "group" and "population" are central participants in the discursive production of race in genomics. One major source of their strength is their ability to substitute for "race" in complex ways.

Some of the ways in which "population" can substitute for race and ethnicity are illustrated in Jennifer Couzin's article for Science's May 24, 2002 issue, "New Mapping Project Splits the Community." Couzin reports on a conflict unfolding over the newly initiated HapMap project. The idea for this NHGRI-coordinated successor project to HGP came from the finding that genomic variation has a structure in which SNPs—individual, single-base differences in genetic code—tend to group together in large sections with only several versions called haplotypes. For a given section of chromosome 21, for example, there might be only three our four haplotypes in humans, each containing thousands of SNPs. The idea of the HapMap Project is to characterize all human haplotypes in such a way as to allow genomicists to genotype people using only a few representative SNPs from which they could infer the variation structure of a whole genome with some degree of statistical certainty. The data from HapMap will be used in a wide variety of disease association studies.[19] Couzin's article focuses on the controversy over the project's effectiveness and cost.

[19] The HapMap Project can be viewed in some ways as a medicalized successor to the Human Genome Diversity Project. The initial phase haplotypes individuals from three races, and the second proceeds to a finer level of focus, in which "scientists will test whether the haplotypes they find in those very large populations also appear in about 10 others."(Couzin, 2002) This study design is justified in medical terms—the basic argument is 1. haplotyping will greatly facilitate disease association studies, allowing genetic predispositions to be easily located at an exponentially higher speed, 2. while all "populations" have all haplotypes, the *frequency* of haplotypes may vary, and 3. in order to ensure that no "population" benefits while others do not, samples must be structured

Several passages in the article illustrate the use of "population" as a racial or ethnic euphemism in genomic discourse. In structuring and representing an early survey of haplotypes, "population" here replaces what is usually described as "ethnicity:"

> His Whitehead colleagues David Altshuler, Lander, and others wasted no time investigating. In a paper published last spring in *Nature*, they described common haplotypes in a group of northern Europeans and a Nigerian population called Yorubans, arguing that haplotypes varied somewhat between the two--a vestige of evolutionary history. (Couzin, 2002)

"Population" here references a place called Nigeria, and a group of people called Yoruban. The categories "Nigerian" and "Yoruban" have a variety of historical and discursive complications that the passage does not indicate.

Nigeria is the most populous nation in Africa, with around 130 million people. Its history as a nation is brief. According to a Library of Congress reference website,

> Like so many other modern African states, Nigeria is the creation of European imperialism. Its very name--after the great Niger River, the country's dominating physical feature--was suggested in the 1890s by British journalist Flora Shaw, who later became the wife of colonial governor Frederick Lugard. The modern history of Nigeria--as a political state encompassing 250 to 400 ethnic groups of widely varied cultures and modes of political organization--dates from the completion of the British conquest in 1903 and the amalgamation of northern and southern Nigeria into the Colony and Protectorate of Nigeria in 1914. (Library of Congress, 1991)

Nigeria achieved independence from Britain in 1960, and post-colonial political history has been complex. Political power in the country has been contested along "ethnic" lines.

---

along the lines of populations. (e.g., NHGRI, 2002) By approaching some of the same research procedures as HGDP using a medical, rather than an anthropological justifications, HapMap accomplishes some of HGDP's aims, while highlighting not only the medical benefit to "populations" if they are sampled, but the risk of accruing medical disadvantage if they are not. This discourse of medical risk/benefit is a major framework in which the genomic production of race is located.

The category "Yoruba" is generally agreed to be one of the three major "hegemonic" ethnic affiliations within Nigeria, the others being Hausa-Fulani and Igbo. "Hegemonic" refers to their being the three main power groups in a tripolar arrangement, around which other Nigerian peoples align themselves. Within this context, "Yorubans" are generally described as an ethnic group.

But the terms of "ethnic" identity in Nigeria are contested, both by Western anthropologists and within Nigerian society. Even the number of languages is debated, and lies within a range from 250 to 450.

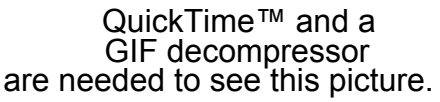
Classification of "Nigerian" "groups" has followed linguistic, religious and political lines. Prior to 19th Century British colonization, an anthropological perspective holds that "Yoruba" was a Hausa reference to their neighbors that enveloped a broad heterogeneous group of city-states centered at Ile-Ife, with similar religions and languages. The Yoruba became more consolidated in this perspective, only under British political structures. Yoruban nationalists dispute this history, tracing a centralized set of religious, political, geographic and linguistic connections that extend five hundred years before the European slave trade. Census data in Nigeria is usually contested, and there is a broad range of claims about the number of Yorubans, ranging from 15 million to 40 million. Using "Yoruban" as a "population" in the HapMap, then, rolls a number of historical complications into one category. Not only does "population" substitute for "ethnicity," but also the unspoken "ethnicity" is unstable, complex and contested.

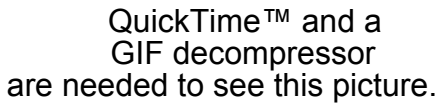
But "population" performs other substitutions in Couzin as well. The Yoruban "population" category is presented as equivalent with two others:

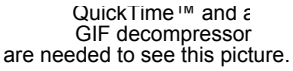
> In March, NIH began soliciting grants for the first of the map's two stages. The first will be to create haplotype maps of the genomes of three

> populations: those of northern and western European ancestry, Japanese and Chinese, and Yorubans. In the second stage, scientists will test whether the haplotypes they find in those very large populations also appear in about 10 others. (Couzin, 2002)

"Northern and western European ancestry" describes people with a historical kinship—an "ancestry"— to a multinational geographic region with a variety of cultures and languages, no matter where they live at the time of study. "Japanese and Chinese" refers to specific nations, but does not specify ancestry. This "population" also includes a wide range of languages, cultures and land mass. "Japanese and Chinese" includes roughly 1.4 billion people (UN, 2003). Northern and western Europe as regions include roughly 210 million people, if their diaspora is not included (Ibid.). Census data for Nigeria is contradictory, but it is likely that Yorubans number in the tens of millions. The numbers of the three "populations" here differ each by an order of magnitude. Their constructions each involve different uses of nation, region, culture, language and ancestry. What is the common link? It is well-illustrated in the third example from Couzin:







This illustration accompanies the article, whose focus is not on the construction of "these populations" as a source of conflict, but on the effectiveness and value of the technique involved. Why, then, choose this illustration? Why do we not understand "these populations" as school-children, sugar cane farmers and folksingers? What system of meaning allows us to move from photographs of three individuals of different ages and sexes to a strong sense of what "these" means without further explanation?"

The link, of course, is race. Yorubans are black, those of "northern and western European ancestry" are white, and "Japanese and Chinese" people are Asian. Negroid, Caucasoid and Mongoloid in a classical schema. In a broad sense, "population" here is acting as a proxy for race. Narrowed specifically to the Yoruba example, "population"

proxies for ethnicity. Incorporating the Japanese and Chinese nationalities, "population" might best be understood as euphemizing the common genomic slur "race-and-ethnicity."

This article, although it is secondary reporting, provides broad, easily visible and non-technical examples of common ways in which "population" is used in primary technical, ELSI and agency genomics literature to euphemize race.

The use of "populations" does not always encompass all of the categories identified above—race, ethnicity, geography, ancestry, culture, language, nationality, biological relatedness, etc. "Population" is often a more specific proxy for one or two of these types. For example, the NHGRI in partnership with Howard University is conducting the Africa America Diabetes Mellitus Study (AADM), in which

> DNA from 400 pairs of affected siblings has been collected from West Africans in Ghana and Nigeria. These West African populations were selected because they are thought by many anthropologists to be the founding population of modern African-Americans and have fewer dietary and nutritional confounding variables than the current African-American population. (NHGRI, 2003)

Three types of population are presented in this passage: "West African populations," represented by siblings in Ghana and Nigeria, "founding populations," and "the current African-American population."

The category "West African" refers to a region of a continent, and is made specific by the nations Ghana and Nigeria, presumably selected because they are English-speaking, easing complex logistics. West Africans are asserted to be relatively homogenous in diet and nutrition, facilitating genetic diabetes studies. "Population," here, is set up by geography, attributes of diet, and with reference to nationality.

This particular "population" is chosen not only because of its diet, but more importantly because it is asserted by anthropologists *to be* another kind of category, a

"founding population." "Founding population" here is an intermediate sort of category between the "West African population" and "the current African-American population." In this setup, two separate populations are bridged by another sort of population, establishing three linked groups,[20] each with different attributes.

The "current African-American population" has "dietary and nutritional confounding variables," which the "West African population" has fewer of. West Africans *are* founders at the same time as they are West Africans: "they are thought by many anthropologists to be the founding population," rather than to be descendants of people who *were* the founding population. I believe this description of West Africans being, rather than descending from "populations" whose "founding" by the slave trade ended more than a century ago, subtly echoes an understanding of Africa that holds African countries to be "primitive," or unchanging. In any case, this phrasing points to the fact that "population" can be used in both synchronic and diachronic senses, even in linked, adjacent categories.

"Population" here is not a euphemism for "race" or "ethnicity," but engages with racialization in several ways. "African-American," for example, is defined as a "race" in official practice in the United States (i.e., OMB, 1995; Risch, 2003). But all three populations in the passage are also routinely categorized as belonging to the same race (i.e., Risch, 2003). This is a rather large instability, with both a part, and that part combined with other parts asserted to equal the same whole.

"Population" negotiates these complications with racial classification, in such a way that biomedical and genomic investigation can proceed as if these "populations"

[20] Chapter 4 addresses the relationship between the one-drop rule, which the United States has historically used to determine blackness, and this notion of "founding populations"

were stable and comparable. "Population," then, is a homogenizing force that allows different categories to be treated in the same way.

The term also mediates among biological and social sorts of knowledge, such that both anthropological and genomic knowledges take it here as an object of study, and make truth claims about it. It transfers meanings between different sorts of knowledges, generally without an acknowledgement that this has been done.

This transferal is well evidenced in ELSI discourse. It is not only that "population" is used both in this sort of writing as well as writing concerned with technical questions, as in the AADM introduction. It is that "population" is used with both biological and social sorts of meaning in either form of discussion.

The ELSI literature produced in parallel with the Human Genome Project's sequence data and technical papers uses "population" in a variety of applications, including those related to race and ethnicity. ELSI literature has been produced internally by NIH and NHGRI to set goals and make recommendations for ethical action, independently by various interested parties, and on commission from NHGRI. A 2001 report from NHGRI submitted to The National Center for Minority Health and Health Disparities, uses "diverse populations" to signify non-white people as a group to be recruited for jobs in research, a constituency in need of outreach, a group which can be impacted by human genetic variation research, and a group in need of education:

> In 1998, when NHGRI started an initiative to study human genetic variation, an NHGRI staff group was convened to make recommendations to the director about how the institute should proceed to assure that diverse populations were included at all levels of its research programs and other activities. Several goals were identified: to establish meaningful interactions between NHGRI and members of diverse populations; to support research on the ethical, legal, social and health policy issues involved in the use of information about human genetic variation; and to

> increase the knowledge about genetic variation and its implications among diverse groups, researchers and policy makers. (NIH, 2001)

"Population" here is used both as a scientific concept and as a political one. "Diverse" here, and in similar agency literature, is used as a synonym for racialized people who are not racialized as white. "Diversity" in genomic literature also refers to a range of genetic difference between individuals—that is, "variation" and "diversity" are linked technical concepts. In this passage "diverse populations" are two terms, each with strong technical and social significance in genomic literatures, linked together to specify scientific, ELSI, and internal political goals.

"Diversity" as a double-entendre involves an exchange of credibilities. "Diversity" is valued within American liberal humanist discourse, and its invocation gives an air of political legitimacy. It also avoids the use of the word "race," helping alleviate the apparent discomfort with the term in genomics. "Diversity" in its genomic sense gives an air of scientificity to political discourse of this type, bringing the authority of the technical side of genomics to bear on discussions of value. "Diversity," in this way, kills two authority birds with one stone, and extends the already flexible term "population" to substitute in another way for "race."

"Isolated" and "special" often modify "population" so as to perform discursive work that is similar to what is done by "diverse." "Isolated populations" generally correspond to particular types of "ethnicity," as used by Witzig and Schwartz, not "race," or "race and ethnicity" as used by Wood. Isolation is defined both geographically, as with people living on remote islands like Tristan da Cunha, and culturally as with a group which tends to have children only with people who define themselves similarly, like the

Ashkenazim. Isolation is cast as strong evidence for genetic homogeneity.[21] Discussions of "isolated populations" in ELSI and agency literature most often focus on the need to achieve consent from those studied in a culturally sensitive manner.

The category "special populations" is an umbrella term for isolated populations and racial-and-ethnic minorities. In this sense, it is another type of euphemism for "race," although it produces a flexibility that extends its connotations beyond race to other types of difference. The term "special" also mobilizes a certain ethical imperative to carry out policies and programs that accommodate its status.

The term is in common usage at NHGRI, and it has a program dedicated to the pursuit of this imperative. "The mission of the Special Populations Research Program is to address issues related to the genetic study of minority and/or isolated populations." (http://www.genome.gov/10000829) The main projects of the program are to coordinate collaborations with Howard University, and Antioquoia University in Colombia, studying diabetes, hereditary colon cancer and ADHD. "Special populations," similar to "diverse populations," is a term with political and technical implications, that engages with and modifies race both in discourse and in practice.

This sort of engagement of "populations" with race has been the site of considerable controversy and critique. Critiques have targeted the fit of social categories with biological categories at the transfer point of "population:"

[21] The construction of "isolation," has been controversial, most prominently in the Human Genome Diversity Project, which set out to sample several hundred such "isolated populations," which were described as "vanishing." HGDP failed both in response to opposition from the "vanishing" and "isolated people" it intended to study, and from opposition within the scientific community. This failure may be seen as a reason that current genomic projects are more commonly structured with respect to racial classification, than they are to "ethnic" or "isolated" classification.

> However, social mechanisms exist that can make socially defined populations imperfect as proxies for biological relatedness when developing a haplotype map or conducting other genetic studies. These include: migration; intermarriage; cultural preferences for claiming only one's father's identity or one's mother's; fictive and adoptive kinship; instrumental and situational choices in asserting alternative identities; colonialist, racist, and nationalist ideologies that impose new identities on subjugated populations or population subgroups; and economic, religious, and other barriers to interaction (and reproduction) within larger social categories. Thus, even after a human haplotype map is constructed, socially defined populations still will differ from genetically defined populations – a fact that is well understood by geneticists but not necessarily by the general public. A genetic or biological view of population history will not be the same as a social view of population history, although the two versions will involve some of the same members. (NHGRI: Foster, 2001)

This critique locates the social definition of populations with the "general public," and the genetic definition of populations with geneticists. Foster presents a neat dualism that masks complex interactions between general and technical knowledge, but acknowledges the problems with the use of "population" as a transfer point between the two. Most genomic critiques of "population" definition around issues of race employ the dualism of lay and scientific knowledge, and focus on the problem of "misinterpretation" by the public. Population, then, is a site of controversy, but within the limitations structured by this dualism.

The construction of "communities" has received less such critique, than "population." As terms, "community" and "population" are at times used interchangeably, but "community" is most often used to indicate a particular power structure. Communities make decisions, and in conducting research on communities, it is considered ethical to gain approval via a culturally sensitive process. In light of the failure of HGDP to achieve approval from its study populations, a significant body of

literature is dedicated from NHGRI is dedicated to the dynamics of this process. A definition of community is offered:

> Communities are composed of individuals with shared interests and interactional patterns that constitute an internal social dynamic that permits (among other things) collective decision-making processes. Communities and populations may overlap in a number of different ways. For example, populations may be composed of more than one community (e.g. a research population described generally as "African Americans" that is composed of multiple smaller communities) or may fail to correspond with any existing community. (NHGRI: Foster, 2001)

Community here is not a synonym for race—"African Americans" are not a singular community, although they are composed of particular African American *communities*. Rather, communities may or may not be structured within racial categories, or more broadly within "populations" which are constituted as racial-and-ethnic. "Community" in genomic discourse is a corollary of race, but not reducible to race.

"Group," is the catch-all term in genomic literatures of race. "Group" can be used interchangeably with "population" and its various connotations, "community," "race" and "ethnicity." It can also be modified in a variety of directions—"racial groups" and "ethnic groups" among others. "Group" as a non-technical term can be employed for humanist pronouncements, political debates, committee reports, peer-reviewed genomics papers, and others. I have found no examination of problematics in the term "group" within genomics, although it stands in for "population" and "race" in critiques of those categories.

In summary, "population" is used in genomic literature with a wide range of flexibility. It is a transfer point between the "social" and the "biological," at times a euphemism for race and ethnicity. It is used in a technical sense and in a political sense.

Critiques of "population" within genomic discourse tend to occur within a defined set of parameters structured along a dualism between science and the public. A distinct vocabulary of "population," including the terms "special" and "isolated" populations has currency at NHGRI. The categories "diverse," "community," "group," and "minority" (discussed in the next section) are corollaries to "population." Taken together, these terms in context add linguistic and logical mechanisms to a genomic technology of difference, allowing a mobilization of determinations about race for a variety of ends.

III. "Minorities" and "Inclusion"

Two major topics are discussed in the literature of genomics: scientific or technical concerns related to DNA sequence, and political or humanist concerns related to the meaning, extraction and social operation of that sequence with respect to the human. Technical claims are devoted generally to the detailing, computation, application and organization of the genome "itself," to the medical implications of the genome, and to variation in the genome among people. Humanist claims by genomicists—in their professional capacity as genomicists—speak to ideas about what it means to be human in light of genomic work, and how to do genomic work in light of humanist ideas. For example, this humanist statement focuses on the value of emphasizing similarity in society, informed by genomic knowledge: "When we look around at the people who surround us in our multiethnic society, we tend to focus on differences rather than similarities. […] It is a much more profound revelation to realize how similar we are at the most fundamental molecular level."(Mansoura and Collins, 2001) The "we"

employed means everybody, and the statement is a liberal humanist claim addressed generally to society and informed by molecular biology. Claims of this variety are also addressed to the genomics community internally, and are used in formulating organizational policy within NHGRI and NIH. Claims related to sequence and its biological ramifications can be usefully differentiated from humanist claims with relation to sequence and its social ramifications.

Although a distinction is useful, I do not mean to imply a neat dichotomy—the two subjects are often discussed within a single paper, and are used to inform and structure each other in a variety of ways.

A dichotomy is not the case, but discourse in genomic literatures can be structured differently depending on whether a claim is humanist or technical. A term can be used exclusively for one category and not another, or carry a different significance between categories of claims. "Race," "ethnicity," "community" and "population," introduced above, do not generally behave this way and show up in a variety of contexts. "Minority" is an example of a term that has several meanings within humanist and political claims, but is seldom used in technical ones.

The term "minority" is most commonly used within "humanist" areas of discussion that deal with racial and ethnic inequality, and policy and ethics for genomic and medical genomic research practices in light of such inequality.

"Minority" in context tends to travel with three terms common in NIH and NHGRI literature: "inclusion," "health disparities" and "representation" and its corollaries "underrepresentation" and "disproportion." These terms are used in negotiating ideas of racial inequality in the literature of genomics.

"Inclusion" is the term used by NIH to describe an imperative in a variety of applications that are similar to Affirmative Action. "Inclusion" is used both alone, and with "diverse populations," "racial and ethnic minorities," "various groups," and other categories to mean the involvement of people who are not raced as white. That is, the term "inclusion" can signify this kind of involvement with or without modifying clauses. Inclusion is advocated by NIH and NHGRI for activities including conducting research, finding participants for research, and educating the public about the results of research. Inclusion almost always involves recruitment—in order to produce inclusion, members of "diverse populations," etc. must be actively sought out and involved with the scientific and medical communities. This activity is mandated and informed by the 1993 NIH Revitalization Act mentioned in the first section of this chapter.

There is, then, a separation implied between federal agency biomedicine and "minorities," an inequality which promoting "inclusion" can remedy. By pointing this out I do not mean to deny the inequality or oppose inclusion, but rather to describe a particular way of framing the issue which identifies two groups: those to be included and those to do the including. "Inclusion" is one term setting the context in this way for the use of "minority" in genomics.

"Health disparities" are also involved in defining the terms of this discussion. Health disparities are racial and ethnic differences in disease outcome and medical service and access in the United States. Disparities are well-documented and often severe. Healthy People 2010 set the goal of eliminating health disparities by that year. In context, disparities are spoken of as a strong reason to study racial and ethnic difference in applied genomics. Disparities mandate inclusion in clinical research participation.

"Representation" modifies the idea of inclusion to factor in demographic proportion. Minorities are "under-represented" if their percentage of the United States population is not mirrored by their numbers in a given study or set of researchers. Disproportion is more often used in the context of disease, as a corollary of *health disparities*. A "disproportionate" number of minorities may have a particular health outcome, with respect to whites.

In context, three features of the use of "representation" and "disproportion" are common: 1) The level of focus in their arithmetic is national—population percentages refer to the United States, rather than any other geographic level of focus. 2) The measure is with respect to whites. 3) "Representation" implies appropriateness—that is, over- and under-representation are problems of inequality, and "representation" as defined with reference to the U.S population and whites, implies fair treatment. Representation and disproportion, then, provide norms in negotiating inequality.

"Inclusion," "disparities" and "representation" as terms structure discourse of racial and ethnic inequality and justice in literature from NIH and NHGRI. "Inclusion" is a process for ameliorating inequality. Disparities are a medical type of inequality, which mandate action such as "inclusion." Representation provides a norm for the process of inclusion. Together, these terms provide a context for the typical usage of the term "minority."

One particularly rich example of the use of "minority" from the literature is the National Advisory Council for Human Genome Research's Action Plan. Approved in May 2001, the plan "outlines clear goals for the inclusion of underrepresented minority groups in research training, research collaborations, and education and outreach activities supported by all components of the institute." (http://www.genome.gov/10001192) The

goal is "to increase the number of underrepresented minorities that are trained to pursue research in the fields of genomics and/or ELSI research." (//www.genome.gov/10001707)

In this example, "minority" refers to groups as well as individuals. The aim is inclusion, produced through recruitment, and the level of focus is on both individuals and groups referred to as "minorities." The distinction between these two senses of the word minority is almost never made. This is another example of terminological instability producing an effect that widens the possibilities of discussion on race.

A number of activities were approved in the Action Plan to achieve the goal of inclusion. These include educational materials specifically targeted to minority students in K-12 schools, a webpage dedicated to genomics minority issues, outreach to minority organizations, and the production of articles for minority publications. The following section, entitled "Minority Conferences" gives a strong sense of who is a "minority" in NHGRI's conception.

> NHGRI will establish a presence at conferences targeted toward minorities (e.g. Society for Advancement of Chicanos and Native Americans in Science National Conference, Annual Biomedical Research Conference for Minority Students) by hosting a visible and active exhibit booth; organizing genomics symposia; compiling an attractive brochure that highlights opportunities for minority students in genomics; and hosting roundtables or hospitality suites so that students have a chance to talk with staff from NHGRI. NHGRI will actively seek out opportunities to give presentations to groups such as Zeta Phi Beta, the National Association of Hispanic Nurses, the Association of American Indian Physicians, the Intercultural Cancer Coalition, and other organizations serving the communities appropriate for ELSI research. NHGRI will participate in at least three conferences per year aimed at minority or underrepresented communities. (http://www.genome.gov/10001707)

"Minority" here includes concepts of race, but also includes ethnicity, nationality and culture. The term encompasses all non-white groups and individuals in the context of the United States population. "Minority students," and "minority or underrepresented

communities" are simply those students and communities that are not white. Together, they represent an identifiable constituency for whom recruitment and outreach leading to inclusion is warranted.

The term "minority" here participates in one way of negotiating inequality. It frames the issue as an inequality between whites and non-whites, allowing a variety of differences, individuals and peoples (such as "Native American physicians" and "Hispanic nurses" above) to be subsumed into a larger, dichotomous way of writing.[22]

In summary, the term "minority" is used most commonly in genomic discourse for humanist claims, and particularly those related to the proper courses of action for genomics organizations. "Minority" is most commonly used in negotiating issues of inequality, and participates in a particular framing of the issue as dichotomous. The terms "inclusion," "health disparities," and "representation" are used to denote actions, mandates and norms in writing about issues of inequality. Together, they contribute to a vocabulary that structures genomic discussions of racial and ethnic difference in an American context.

IV. Geographic Ancestry

Several authors have identified the potential for genomics and biotechnology to reframe and redefine questions of race. José Morales, in a report on the social implications of

[22] The names of contemporary racial and ethnic classifications in the United States are used to varying degrees in genomic literature. The terms black, white, African-American, Hispanic, Latino, Native American and Asian are used in both technical and humanist claims. At times they are used with the sense of caution described for "race," at times they are used more liberally. I have seen no particular pattern to their usage, and I don't have the sense that there is a particular genomic definition or connotation in their deployment, except in cases where they are deliberately redefined, as with Risch, et al (2002). Although they come into play very strongly in the next chapter, proper names of races do not appear as categories or keywords specific to genomics, so I do not discuss them further here.

genomics to the *Second National People of Color Leadership Summit*, described this reframing as largely geographic:

> One of the many impacts of biotechnology concerns the issue of identity. The existential question "Who am I?" will be answered by asking the question "where do I come from?" This will have noteworthy consequences for the social reality of people of color in the United States. (Morales, 2002)

In all of the discussions and debate around racial and ethnic classification in genomics, geographic ancestry has sparked the least controversy, and seems to be viewed as most based in biological reality. Whereas the biological legitimacy of race is strongly debated, and genomic claims about its nature are contradictory, geography—where people are from, where people's ancestors are from— is written as a solid, legitimate entity available for study.

At the broadest level, geographic ancestry has three centers: Africa, Asia and Europe. It is not that these are strictly bounded locations which produce strictly bounded groups. Rather, these locations are the centers of three nebulous groups that intermix at their boundaries.

The technical term for interbreeding between geographic groups is *admixture.* Admixed populations have multiple geographic ancestries with reference to these three groups. Non-admixed populations are more geographically pure. A number of studies have produced genomic admixture tests that claim to identify the proportions of geographic groups represented in individuals—that is, how much a person is African, European or Asian. These tests have forensic, anthropological, medical and geneaological applications. The example below describes a disease-association application. It also

illustrates the involvement of ideas of race and ethnicity with geographic ancestry and the idea of "founding populations":

> Population linkage disequilibrium occurs as a consequence of mutation, selection, genetic drift, and population substructure produced by admixture of genetically distinct ethnic populations. African American and Hispanic ethnic groups have a history of significant gene flow among parent groups, which can be of value in affecting genome scans for disease-gene discovery in the case-control and transmission/disequilibrium test designs. Disease-gene discovery using mapping by admixture linkage disequilibrium (MALD) requires a map of polymorphic markers that differentiate between the founding populations, along with differences in disease-gene allele frequencies. We describe markers appropriate for MALD mapping by assessing allele frequencies of 744 short tandem repeats (STRs) in African Americans, Hispanics, European Americans, and Asians[…] The map will be useful for studies of diseases, including prostate and breast cancer, diabetes, hypertension, and end-stage renal disease, that have large differences in incidence between the founding populations of either Hispanics or African Americans. (Smith, et al, 2001)

Admixture here, rather than arguing against "genetically distinct ethnic populations" as it does in some critiques of race (see section II), is a form of combination of such distinct entities. "African Americans represent an admixed population with significant genetic contributions from both African and European ancestors"(Smith, et al, 2001) African and European ancestry, geographically based, are objective quantities, and Smith offers genomic markers for studying their relative distributions.

Admixture here actually works to reinforce the reality of distinct geographically based races, because current mixture requires distinct, pre-existing groups to do the admixing. In this way, "admixture" helps to produce race.

This type of geographic classification in the genomic conception is uncontroversial. The significance of admixture among geographically defined categories is open to interpretation, but the centrality of geographic ancestry appears commonly accepted.

V. Chapter summary

I have described in this chapter specific genomic usages for the terms race, ethnicity, "race and ethnicity," diversity, population, special populations, isolated populations, community, group, inclusion, health disparities. representation, minority, geographic ancestry and admixture. Together, these terms offer a sketch of the vocabulary used by genomicists in making claims about human racial and ethnic difference. In application, these terms structure concepts of the biology of race, the social and biological elements of classification, the ethical conduct of research, racial inequality in the United States and the geographic elements in biological relatedness. These are the building blocks in a genomic technology of racial and ethnic difference. The terms are involved in setting parameters for discussion, suggesting questions for research and actions for researchers, framing concepts of difference in particular ways, and suggesting criteria for what claims are legitimate. These are the ways in which terminology makes race.

Chapter 2: Formal Configurations: Nested Proxies and Perspectival Phasing

I. Theoretical Framework

Here I move from a focus on the terminology that forms some of the building blocks of genomic discourse, to a broader view of the patterns within that discourse. While the terminological approach allowed an examination of forms and features, a study of configurations allows a focus on the actions and operations that the technology of difference takes. By analyzing the ways in which discourse is configured, we can see the technology of difference in operation—differentiating people, and establishing the parameters and criteria by which it does so. The configurations establish particular patterns of instability in meaning. This instability is crucial to the production of race.

One way in which the assembly of discourse can be examined is at the level of the *statement*, with a focus on specific claims that genomicists make about race, and the forms these claims take. Another level of analysis is the *document*, examining statements, frameworks, terms and contexts assembled by authors in individual papers. The similarities and differences in these modes of assembly among different documents could be used to analyze discursive forms.

My approach emphasizes the *passage*, a piece of writing excerpted from a document for a particular purpose. I select passages from documents and focus on the patterns of their configuration. The passages I examine are not always delimited by the unity of their content, such that the writing I select refers only to one set of conclusions. Particular aspects of form, voice and terminology also provide my criteria for selection, so that the content of a passage is sometimes fragmentary. Basing excerpted quotations largely on formal aspect allows the patterns in which discourse is assembled to be brought into specific relief. Using this method makes it easier to draw conclusions about the conditions of assembly that are at work here. Understanding the formal conditions of presentation lets us understand how race is produced, and brings us closer to understanding epistemology.

There are two predominant patterns in form in the genomic discourse of race at the level of the passage, which I call *nested proxies* and *perspectival phasing*. The two are formal configurations—heterogeneous, but distinctive manners of form in writing. These formal configurations are useful in examining and theorizing the genomic discourse of race, and they are characteristic, although not exhaustive, of the emergent mainstream. Nested proxies link successive, non-equivalent categories together, producing a large degree of flexibility and the appearance of consistency in passages concerned with race. Perspectival phasing is a rapid change in authorial voice, so that the presentation of an author's perspective shifts in the space of a passage. It is an alternation between the erasure and the foregrounding of perspective.

My idea of perspectival phasing has been influenced by Donna Haraway's essay "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective"(Haraway: 1988). In this essay, Haraway, a prominent science studies

scholar and feminist theorist, elaborates the idea of "the god-trick" as a characteristic discursive habit and self-image in contemporary technoscience. The god-trick is the escape from perspective, a claim to the ability to see a universal truth, the habit in which authors transcend their bodies and speak from everywhere and nowhere at the same time. With the visual technologies of current science, "vision […] becomes unregulated gluttony; all perspective gives way to infinitely mobile vision, which no longer seems just mythically about the god-trick of seeing everything from nowhere, but to have put the myth into ordinary practice (ibid.)." The god-trick is not only a way of writing in which the author's location is erased, but a set of associated visual practices.

Haraway's argument is for the return to partial vision and embodied location as the condition of credibility. "only partial perspective promises objective vision. This is an objective vision that initiates, rather than closes off, the problem of responsibility for the generativity of all visual practices." If the practice of vision in science *produces* things—if vision is "generative"—then the only way in which just accountability, and objectivity can be achieved is through privileging partial perspective.

At first glance, genomic science seems to be the diametric opposite of partial perspective. Its efforts are explicitly "global." *The* Human Genome Project renders universal truths about *the* human. The visual technologies and the disembodied narrator[23] in "Exploring Our Molecular Selves" are the god-trick in practice. We travel through an eye, into a cell, see through ribosomes, see the underlying four "letter" structure in DNA,

[23] The narrator in the video is a woman, an unusual choice for the Voice Of Science given Haraway's characterization. Women and people of color, she argues, are most often denied the privilege of disembodiment in technoscientific regimes. Only white men get to leave their bodies and see from everywhere and nowhere. That the narrator is a woman should not indicate to us that Haraway is wrong, and certainly not that genomics is a new feminist paradise, but that our narratives about science need continual adjustment, and our old stories may prove inadequate to an understanding of emergent genomic science.

and exit again through the eye of another person, eliciting light-shows the whole time. "We" can be everywhere, and "Our" vision is truly god-like in this sort of visual practice.

But in the genomics of race, another sort of trick is visible in the discourse. In certain parts of certain passages, the authorial voice is presented as omniscient— absent yet everywhere. However when the god-trick appears, it tends to be part of a larger pattern, in which authorial voice oscillates, often in extreme degrees, from disembodied and naturalized, to located and socialized. Perspectival phasing is in this way a "Cheshire Cat trick," named for the creature that appears and disappears, in whole or in part, and in a different place each time.

My idea of perspectival phasing rests on the observation that an author can present herself in different ways in discourse generally, including at the level of particular statements, and in the space of a document. Writing can make an author's presence as the source of perspective obvious or erased. For example, in a view discussed at length below, Elijah Saunders, says "genetics may set the stage for [racial disparities in] developing hypertension, [but] I favor the environmental-socioeconomic explanation as a cause for high prevalence of the disease (Brower, 2002)." This way of configuring discourse calls attention to its perspective by its pronouns, its use of "favor" and "explanation," and the uncertainty implied in "may." Saunders as an author is foregrounded. In contrast, the sentence "African-Americans respond better to a new drug, eplerenone," from the same passage, has no indication of perspective, and is simply declarative.

The extent to which writing indicates that it is a discursive activity—the "discursivity" of writing—can shift rapidly in a way that is parallel and linked to changes in authorial voice. For example, the phrase "We categorize Africans as" emphasizes

categorization, a discursive practice, whereas the phrase "Asians are" gives no indication of discursivity.

Changes in authorial voice and the indication of discursivity seem to travel together in the genomic discourse of race. When one shifts, the other usually does as well. Taken together, authorial voice and discursivity produce a presentation of perspective.

Single, consistent perspectival presentations are *not* characteristic of the genomic discourse of race, either as a whole, as individual documents, or even within the small spaces of passages. The genomic discourse of race is characterized by shifting, inconsistent presentations of perspective. This process of shifting is what is meant by "phasing." I chose this term because it implies a certain rhythm that is suggested in much of the literature, as well as connoting appearance and disappearance, exposure and erasure.

Perspectival phasing both contributes to, and results from the production of race. As a site of instability, this pattern works, often in connection with nested proxies, to strengthen race in complex ways that this chapter examines.

The formal configuration *nested proxies* refers to the habit of replacing heterogeneous terms with each other. "Proxies" are the relationship between terms, when one term stands in for another within the space of a document or passage. "Nesting" happens when these substitutions occur in succession, one category taking the place of the previous, while maintaining a connection to the first. In writing nested proxies, authors use categorical terms in a linked succession within a document, producing both technical and humanist claims, as well as more generalized features of documents.

A proxy is not quite a euphemism—"race" and "population," for example, are seldom used as synonyms to mean precisely the same thing. Rather, a discursive arrangement is employed that allows a category like race to be linked to a category like population progressively, and by implication rather than explication. Explanations for the use of proxies are not generally offered; categorical replacements are simply deployed.

My concept of nested proxies has benefited from several critical papers within genomic literature, which discuss "proxies" in reference to race. Morris Foster, professor of anthropology at the University of Oklahoma and editor of American Indian Quarterly, has written extensively on the ethics of community consent in genomic research, and issues in socially identifiable populations, and is the author of the "heuristic starting point" conception of race, in which racial structures are considered arbitrary, though practical means to design genomic research. When Foster uses the term "proxy," he means the use of social categories as modes of representation for biological realities. He describes this sort of proxy as problematic but often useful:

> Clearly, deciphering the relationships that may exist between social classifications and biological categories is not a simple matter. The biological significance that a social distinction may have for one purpose can dissolve when those same social categories are used to answer other biological questions. Thus, it may be appropriate to use social categories as a proxy for biological relatedness (or unrelatedness) in some circumstances but not in others. (Foster, 2002)

Foster's conception of "proxy," then, is the process of making social classifications such as race and ethnicity surrogates for biological similarity or difference. These two categories, biological and social are identifiably separate things, but making one a proxy for the other may be legitimate in certain research circumstances.

Foster's definition is useful for this thesis, in that categories linked to each other in nested proxies often bridge the "biological" and the "social." Things that are generally agreed to be biological, like geographic origin or ancestry often stand in as proxies for things that are generally thought of as social, such as nationality or self-reported racial identity. Nested proxies often embody this kind of transferal among categories that are traditionally understood as occupying either end of the nature-culture split.

But by describing nested proxies I do not mean to indicate simply a troubled correspondence between socially constructed classificatory systems and the real biology of difference. Instead, proxies in my schema point to the linkage of heterogeneous discursive categories *within* discourse, not the inconsistencies between words and the world. It is more in the dynamics of these discursive interconnections, than in their engagement with material reality that race is being produced in genomics.

Sankar and Cho in their piece "Towards a new Vocabulary of Human Genetic Variation" for Science's policy forum section, identify and critique such a discursive process in their concept of nonequivalence. Their position is that a certain kind of instability damages scientific accuracy, and forms the bad biology of race:

> Accessing a particular set of conditions with a variable requires choosing the right variable and using it consistently. While this may seem obvious, race-related genetic research does not always observe this rule. For example, the initial reference to race in an article is often to the racial identity of individual subjects, sometimes described as "self-assigned" by subjects. A subsequent reference to race might appear in the classification of genotypes associated with groups of the self-identified subjects. The final one might appear in the discussion section that generalizes the findings to different racial groups, i.e., massive world populations, such as Euro-Americans or Asians.
>
> Nonequivalent use of labels is illustrated by the common juxtaposition of terms such as "white" with "African-American," where skin color and geographic location are treated as equivalent. Another example is the juxtaposition of "Asian-American" with "Mexican-

> American," which implies that people of Asian ancestry now living in the United States represent a level of genetic diversity that is equivalent to that of people of Mexican ancestry now living in the United States. Such examples indicate a need for more consistent attention to definition of groups and to the need to explain the rationale for their equivalence. (Sankar and Cho, 2002)

The description of moves from the initial reference to racialized individual subjects, to genotyping, to a final statement about massive racial world groups, is one type of nested proxy. Their analysis of implied equivalence is especially useful as an analytical tool, allowing us to see hidden implications in the juxtaposition of group names. By opposing two or more categories, the argument is made that one "represent[s] a level of genetic diversity that is equivalent to" that of the others. Implication of equivalent difference is a powerful factor in the production of race, and is an example of the type of consistencies produced by the nested proxy configuration.

Sankar and Cho view "nonequivalent uses of race within one research report" as one of three major problems caused by the complexity of race as it overlaps between scientific and popular use. The others are "inverting the relationship between genetics and race, or studying race as an end in itself; and […] an overemphasis on race." In their own critique, however, the authors imply equivalence between the discursive process of nonequivalent use, and the problems that are both discursive processes, and the *results* of various kinds of processes: inversion, the study of race as an end in itself, and the overemphasis on race.

Most importantly, they characterize the influence of the trends they identify as primarily destructive or confounding—or at the very least as scientifically illegitimate. I argue in this chapter that these things—nonequivalent usage or nested proxies—are in fact productive, that they serve to structure genomic knowledge about race in particular

ways, allow certain conclusions, and resolve certain contradictions within this knowledge. That rather than stymieing knowledge production, these formal configurations actually empower it.

The following sections illustrate examples of nested proxies and perspectival phasing in action, and elaborate on their structure and dispersal. These formal configurations are the processes through which race is made in genomics. A surprising range of strengths and abilities in the production of race—a large variety of race-making moves—is enabled by, and based in these instabilities.

II. Making Difference Within Race

This section describes some of the ways in which nested proxies and perspectival phasing work together to attribute characteristics to, expand on, and bound concepts of race. The production of these types of difference, internal to racial categories, is one major capability of these formal configurations.

Examples of this sort of capability can be found in all of the major genomic positions on the reality of race, from social construction to fundamental biological reality. Nested proxies and perspectival phasing can even support and structure multiple positions within individual articles.

The reality of race is construed in a variety of ways, for example, in Vicki Brower's (2002) discussion of genomic and medical race debates in "Is health only skin-deep?" a news article for *European Molecular Biology Organization* (EMBO) Reports. She offers a passage summarizing the views of Elijah Saunders "himself an African American" (and the only racially identified person in the article) regarding hypertension

and race.[24] Saunders is quoted as saying "genetics may set the stage for [racial disparities in] developing hypertension, [but] I favor the environmental-socioeconomic explanation as a cause for high prevalence of the disease." He goes on to say that race is not a "genetic situation," but a potential "marker" for a genetic situation. This view locates race as a "heuristic," to borrow Foster's term, such that it can be useful as a surrogate for biological reality without itself being a biological reality. Race, in this sense, is a transfer point between natural and social categories.

But this is not the only viewpoint presented in Brower's review of Saunders' points. Brower goes on to cast racial groupings in a very *real*, medically distinct light. Nested proxies are at the center of this movement. I italicize the categorical terms here to emphasize the linked succession among them:

> "Furthermore, as *many blacks* are more salt-sensitive and more likely to have low plasma renin levels, *they* are less sensitive to ACE (angiotensin-converting-enzyme) inhibitors and ARBs (angiotensin II receptor blockers), which makes the disease more difficult to control in *this population.* […] *African-Americans* respond better to a new drug, eplerenone. (Brower, 2002, emphases mine.)

At first glance, a summary of this passage might read "Because *African-Americans* tend to be salt-sensitive and have low renin levels, they do not do well with ACE inhibitors or ARBs, so hypertension is hard to treat in *African-Americans*—they respond better to

[24] This identification points to the distribution and conditions of credibility in genomic discourses of race. By being "African American himself," Saunders acquires a particular authority in his positions on race. The implied history of racial injustice is mobilized to incline us to give more credence to a defense of race that comes from a racialized person. The fact that none of the other authors are identified means that, because it is the discursive default, they are "white." So while Saunders words are given the privilege of raced perspective, the others have the privilege speaking from non-raced perspectives. This points to a complex interplay of power relations that exist in racial discourse which cannot be reduced to "racism," but engage with systems of racial privilege and embodiment.

eplerenone." This summary is a claim about one stable category, "African-Americans," and their medical attributes.

But the passage is not stable, and not about one category. Instead, multiple categories are deployed successively, and their meanings shift. "Many blacks" are more[25] salt-sensitive. "They" respond badly to two different anti-hypertensives. Therefore, hypertension is a difficult problem to address in "this population." "African-Americans" do better with the anti-hypertensive eplerenone.

"Many blacks," "They," "this population" and "African-Americans" are not euphemisms or synonyms. Rather, they are proxies. The terms stand in for each other without equaling each other. Their succession moves the argument from a claim about metabolic chemistry in a proportion of black people, through a claim about a drug response in a nebulous "they," to a claim about a more general "population" (not the proportion of a group expressed in the "many blacks"), and finally to a claim that "African-Americans" as a group respond better to a particular drug. The movement of claims is not explicitly inductive or statistical. It does not say "a greater proportion of African-Americans exhibit XYZ characteristic, and so for the purposes of this research we describe African-Americans as XYZ."

The argument moves by simply changing its categories in a particular way, concluding with a claim that provides the sense that the whole passage has been produced using the same terms. These movements are nested proxies, and they link heterogeneous terms so as to create the effect of consistency. The immediate result in this passage is a new attribute of a racialized group of people. Beginning with "many blacks," the

[25] "More" here implies a baseline—something to be *more than*. I think white normativity supplies it here, the concept that whiteness provides a norm of health from which others diverge. I will discuss this in detail in the final chapter.

movements of proxy formation get us to "African Americans" as a whole, who now have a new medical trait. The medical trait, in turn, is used to reinforce the appearance of a stable category. If we can ascribe scientific traits to a category, then the category itself feels scientific.

Equivalent usage would not allow this attribution to be reached, and would do nothing to reinforce the existence of "race." Nonequivalence is actually productive, rather than confounding. The links among "many blacks," "they," "this population" and "African Americans" impair nothing in the argument. In fact, they *enable* the point, they *allow* the conclusion. Equivalent usage would read:

> "Furthermore, as *many blacks* are more salt-sensitive and more likely to have low plasma renin levels, *many blacks* are less sensitive to ACE (angiotensin-converting-enzyme) inhibitors and ARBs (angiotensin II receptor blockers), which makes the disease more difficult to control in *many blacks*. *Many blacks* respond better to a new drug, eplerenone. (Brower, 2002, emphases mine.)

The strongest implication of this argument is about eplerenone, not race. We can see that the drug works well for salt-sensitive or low-renin people, but the repetition of "many blacks" reads more like a distraction. Because the statements refer only to "many," there is no new medical attribute that marks a whole "race," and no new articulation of racial differentiation. This equivalent usage does nothing to reinforce the appearance of a stable category. It may operate within this sort of assumption—that "blacks" exist as a discrete biological group—but its claims cannot be mobilized back to buttress the categorical boundaries.

It takes the addition of the chain of categorical proxies, for a new criterion of racial drug response differentiation to be produced. Proxies move us out of a proportion of people within a race, to the whole group itself.

Along with this movement comes a medical imperative to make use of it. If we *know* these things about medical traits of a race, we are obligated to make use of them. We'd be racists not to. This sort of thinking structures many of the justifications for race-based genomics research.[26]

The medical imperative, and the new racial attribution, both enabled by nested proxies, feed back and give the impression that stable race exists. Instability in this instance is deeply stabilizing.

The result is that a new particular of racial differentiation—a new facet of race itself—is created. Genomics produces race. Nested proxies make it possible.

Perspectival phasing also occurs in this passage, representing another kind of instability, and allowing the production of race in related ways. The rest of Brower's paragraph helps to show the oscillation in the presentation of perspective:

> At the meeting of the International Society for Hypertension in Blacks in Miami this year, the Association of Black Cardiologists presented a similar heart failure study specifically targeted to *African Americans* in which the drug BiDil was found to have a significant effect only in *blacks*,[27] who some believe to have low nitric oxide activity. But even given these results Saunders cautioned against categorical assumptions that *race* is a determining factor. 'We're not talking about race as a genetic situation. We're talking about it as a *marker'* he said." (Brower, 2002. emphases mine)

[26] The Haplotype Map's population sampling approach was justified largely in these terms, such that *not* testing by geography could lead to unfair distributions of medical benefits.

[27] The interchangeability of "blacks" with "African-Americans" appears to be a consistent aspect of the discourse of genomics.

In addition to the transitions among categories I describe as proxy formation (which extend here through the italicized terms), we see a rapid passage through different ways of treating perspective and truth claims. The introductory sentences to the first half of the passage portray an initial belief located in an individual—"I favor the environmental-socioeconomic explanation"— producing a foregrounded perspective. But the text turns immediately to a core of objectively detailed facts, erasing perspective: "many blacks are more salt-sensitive." Next the argument moves to facts attached to their professional and organizational sources: "Saunders recently presented a study showing…" the Association of Black Cardiologists presented a […] study […] in which the drug BiDil was found to…" In the second part of the passage, this tone is echoed, where the organizations "International Society for Hypertension in Blacks" and "The Association of Black Cardiologists" are behind truth claims. Although authorship is cited, their results are voiced as truth, not claims. Following the core of erased perspective, these sentences present a sort of mixed perspectivity, placing their authors in view, but without conceptuality or uncertainty. Closer to the end of the passage, authors, uncertainty and self-conscious discursivity come back into the forefront, and facts are replaced with claims. "Some believe" that "blacks" have low NO activity. "Saunders cautioned against categorical assumptions." "We're not talking about *race* as a genetic situation. We're talking about it as a *marker*." Taken as a whole, these quotes represent very different, phasing presentations of perspective in the course of the passage.

The effect of this phasing is to naturalize some claims and socialize others. In this passage, all of the points that call race into question, or express limits on its reality are voiced as social, and point to their own discursivity. "I favor the socioeconomic explanation," "Saunders cautioned against categorical assumptions," and "we're talking

about [race] as a marker" are the phrases that weaken claims about biological race. The source of the perspective is emphasized in each case, with "I," "Saunders," and "we." The discursivity of claims by is emphasized as well, with "favor," "explanation," "categorical assumptions," "talking" and "marker" calling attention to their status as processes in discourse.

Conversely, all but one of the points strengthening the reality of race have perspectives that are erased wholly or in part, and make no reference to their discursivity. The entire first half of the passage emphasizes the reality of race, and is voiced without perspective, and without discursivity. The effect is that the claims appear as facts without owners, and the scales are tipped for the reality of race.

In this context, perspectival phasing strengthens the production of race, in concert with nested proxies. Not every use of this formal configuration serves this function, and not every use of nested proxies operates to move an argument from a proportion to a whole. Their uses in this passage show some of the ways in which instability can enable the production of race.

The passage from Brower focuses on one particular "race." It specifies it, fixes an attribute to it, and narrows its criteria of differentiation to a medical trait. Nested proxies can also broaden race, making it ambiguous and flexible.

A passage from a request for grant applications (RFA) from NHGRI (1999) exemplifies this sort of expansion. The RFA solicited applications for "Studies of the Ethical, Legal and Social Implications of the Research Into Human Genetic Variation," and asked a series of questions that applicants' studies should address in two subsequent sections labeled "Interpretation and Use of Information on Sequence Variation" and "Potential Uses of Information on Sequence Variation in Clinical Settings."

> -Will individuals from *diverse populations* be more or less vulnerable to various forms of stigmatization and discrimination as a result of the availability of information on sequence variation?
>
> -How will *diverse communities* be affected by the potential commercialization and patenting of DNA sequence variation information?
>
> -Are there ways to minimize harm to individuals and *groups*, while maximizing possible benefits?
> […]
>
> -How will individuals with *diverse cultural and socioeconomic backgrounds* perceive the development of genetic tests and treatments that are designed for use in specific *socially or geographically defined populations*, in which a particular DNA sequence variation is prevalent?
> […]
>
> -Are there strategies to increase the likelihood that *diverse communities* will benefit from health information and services derived from sequence variation research? (NHGRI RFA, 1999, emphases mine)

The word "diverse" has a particularly broad range of applications in this passage. "Diverse" modifies "populations," "communities," and "cultural and socioeconomic backgrounds." At first glance it seems to have a stable meaning, repeated throughout

On closer inspection, we can see that "diverse" is not stable. It is not merely a synonym or euphemism for "racial" or "racial and ethnic." Nor is it precisely a proxy for race, in the sense that it substitutes, non-euphemistically, for an antecedent. "Diverse," instead, suggests racial and ethnic difference, while leaving the possibility of other forms of difference open.

"Diverse communities," to take one example, may "be affected by the potential commercialization and patenting of DNA sequence variation information." The term "communities," refers most often to political constituencies, particularly in regards to consent. NHGRI is offering a grant to study the impacts of commodified information on genetic difference in "diverse communities." The emphasis is on racial and ethnic

groups—"minorities" would probably be the term deployed, if "racial and ethnic" were the exclusive types of "diversity" referenced here—but the ambiguity of "diverse" allows the question to be extended beyond minorities, to other modes of difference. This ambiguity does work, allowing the strong suggestion of race and ethnicity as the main ideas in question, while maintaining the flexibility to extend to other types of difference.

Diversity links this ELSI issue with a variety of others, making the implication that they are similar sorts of issues, affecting similar, stable sorts of categories of people. This consistency is implied, then, but exactly who this consistent group is made up of is never specified. Race is implied, but race exists in an ambiguous, undifferentiated field.

Proxies nested in this passage produce this sort of ambiguity. It is the links among the categories in italics, not their links back to specific raced groups, or their use as social surrogates for biological difference that produces race in the passage. "Diverse populations," "diverse communities," "groups," "diverse cultural and socioeconomic backgrounds," "socially or geographically defined populations," and "diverse communities;" work together to set a particular type of expansive reference a particular *field* of difference, instead of making a specific *claim* about difference, such as we saw in the quote from Brower.

In context these nested proxies frame a field of inquiry without strictly bounding it—they suggest that race is their proper referent, without defining it, using its name, or excluding other fields of difference. The effect is that the meanings of this passage are expanded, ambiguous and polite. The production of race here happens through the location of difference in this sort of expanded terrain. The repeated use of "diverse" gives the sense of consistency in a highly flexible, interlinked presentation of difference.

We should not confuse instability, non-equivalence or inconsistency with weakness. The discursive heterogeneity in this passage does work, structures knowledge, frames questions. Materially it helps to direct the flow of funds from NHGRI to ELSI researchers, allowing a certain breadth of focus, and implying certain aspects of tone and terminology for applicants to perform.

If nested proxies in the Brower example allowed a specific, narrowed racial attribution, and produced a general, expanded racial implication in NHGRI's RFA, they can also be seen structuring a specific racial schema that encompasses all human groups. Nested proxies can help to characterize, diffuse and distribute, and bound race.

Neil Risch's opinion piece for the online journal *Genome Biology* "Categorization of humans in biomedical research: genes, race and disease,"(2002) makes use of nested proxies in producing racial boundary work. He is one of a small number of genomic authors whose ideas fall outside of the heuristic-concept-misconception schema of "race" I identified in the previous chapter. His position is that "the greatest genetic structure that exists in the human population occurs at the racial level."[28] In making these claims, Risch

[28] Risch's paper has received some strong critique from outside the field of genomics, including from social theorist Jacquelyn Stevens, who I quote in a discussion of the one-drop rule in the next chapter. Within genomics, Risch has been a subject of controversy, but not dismissal. In a piece for the New York Times, "Race Is Seen as Real Guide to Track Roots of Disease," Stephen O'Brien, a geneticist at the National Cancer Institute (NCI) "said that the conclusion that race was not a valid concept 'comes from honest and brilliant people who are not population geneticists.'"

"That doesn't mean they are insincere," Dr. O'Brien said. "It's just that they haven't really looked at it. What is happening here is that Neil and his colleagues have decided the pendulum of political correctness has taken the field in a direction that will hurt epidemiological assessment of disease in the very minorities the defenders of political correctness wish to protect." [¶] Others play down the medical usefulness of racial differences. "We can't wish away these boundaries," said Dr. Aravinda Chakravarti, a population geneticist at Johns Hopkins. "But I'm not convinced that knowing these boundaries is necessarily useful for genetic research."[¶] Dr. Risch concludes his review by noting that every race and ethnic group within a race has its own set of diseases and clinical priorities, which a new arsenal of genetic tools is poised to address. "We need to value our diversity rather than fear it," he writes. "Ignoring our differences, even with the best of intentions, will ultimately lead to the disservice of those who are in the minority." (New York Times: Wade, 2002) Diversity here conflates social and genomic discourses, and is mobilized against the "politically correct" assertion that race is not biological. This is another example of the strengthening race through the unstable meanings of terminology.

makes use of a variety of nested proxies. The linked substitutions among terms in this passage are not racial categories, but shifting types of relationships that are used to produce racial categories. The shifts in the description of "ancestry" are particularly interesting:

> For our purposes here, on the basis of numerous population genetic surveys, we categorize Africans as those with primary ancestry in sub-Saharan Africa; this group includes African Americans and Afro-Caribbeans. Caucasians would include those with ancestry in Europe and West Asia, including the Indian subcontinent and Middle East; North Africans typically also are included in this group as their ancestry derives largely from the Middle East rather than Sub-Saharan Africa. 'Asians' are those from eastern Asia including China, Indochina, Japan, the Philippines and Siberia. By contrast, Pacific Islanders are those with indigenous ancestry from Australia, Papua New Guinea, Melanesia and Micronesia, as well as other Pacific Island groups further east. Native Americans are those that have indigenous ancestry in North and South America (Risch, 2002)

These are the basics of Risch's schema of race. His conclusions derive from his analysis of a number of population genetics studies not his own, which "have recapitulated the classical definition of races based on continental ancestry." He introduces a number of novel modifications of this schema around the problem of "migrations blurr[ing] strict continental boundaries," but maintains that "the existence of […] intermediate groups should not […] overshadow the fact that the greatest genetic structure that exists in the human population occurs at the racial level," which is continental, classical, and accurately captured by self-categorizations.

On the surface, this passage gives the strong impression of consistency. There is none of the ambiguity of NHGRI's "diverse" groupings. The entire passage consists of what so many other genomic documents lack—clear definitions of the racial categories

employed. But instability is in fact at the core of Risch's argument, as shown in Figure 2.1, which dissects the clauses of each sentence.

The link to race, in the second column of the diagram rests on concepts of "ancestry." "Ancestry," according to Risch, "refers to the race / ethnicity of an individual's ancestors, whatever the individual's current affiliation."[29] But there are

[29] "Race / ethnicity" in this sentence is a further source of instability in Risch. Immediately prior to this sentence, he draws a distinction between "race" and "ethnicity," and then uses them interchangeably.

Figure 3.1. Diagram of racial schema in Risch, et al. 2002

#	Name of Category	Link to Race	Geo. Criteria of Race
1.	We categorize Africans as	those with **primary ancestry** in	sub-Saharan Africa
2.	Caucasians would include	those with **ancestry** in	Europe and West Asia
3.	Asians are	those **from**	eastern Asia
4.	Pacific Islanders are	those with **indigenous ancestry** from	Australia, Papua New Guinea, Melanesia, and Micronesia.
5.	Native Americans are	those that have **indigenous ancestry** in	North and South America.

Continued..

#	Intro to extension	Extension
1.	this group includes	African Americans and Afro-Caribbeans.
2.	including the	Indian Subcontinent and Middle East;*
3.	including	China, Indochina, Japan, the Philippines and Siberia.
4.	as well as	other Pacific Island groups further east.
5.		

*Super-extension: "North Africans / typically also are included in / this group / as / their ancestry / derives largely from / the Middle East."

several types of ancestry that are presented here as equivalent. Africans have "*primary* ancestry in," Caucasians have only "ancestry in," Asians are "those *from*," Pacific Islanders are "those with *indigenous* ancestry from," Native Americans "*have* indigenous ancestry *in*."

In addition, North Africans' ancestry is something that "derives largely" from a place. There are two means of referencing "indigenous ancestry:" one can have it in a place, or be with it from a place.

These attributions of ancestry subtly invoke received notions about race. "Pacific Islanders" and "Native Americans" are usually understood to be "indigenous people," a term that refers to a colonial history, and generally refers to a "tribal" social structure. Western ideas about economic and cultural development also influence our use of the word "indigenous," such that indigenous societies are usually "less developed" or "primitive." Using "indigenous" to modify ancestry for Pacific Islanders and Native Americans mobilizes these underlying connotations, thereby strengthening the production of race.

The slight modification "indigenous," within a nested proxy configuration, allows Risch's schema to conform to, and draw from a whole history of the construction of "indigeneity" in the New World. The backing of this history gives it strength and specificity that it would not have without the use of nested proxies. If all of the types of ancestry were "indigenous," this history would either not be involved, or it would weaken the production of race.

Caucasian *indigenous* ancestry, for example, contradicts ordinary ideas of who is "indigenous," and its use would reverse the process of invoking received racial ideas. Its use would call these ideas into question, and weaken the connection between existing

racial discourse and the Risch's new racial architecture. Nested proxies allow complex new discourses of difference to be produced, at the same time as old and far-reaching discourses of racial difference are invoked to support the new. All of this is done in a way that produces the sense of consistency throughout.

"Primary ancestry" seems to be geared specifically to a North American context, so as to include African Americans. Unlike Asians, Africans cannot be "from" Africa, if African Americans are to count as members. African Americans are not "indigenous," because this usually refers to a colonial relationship in which the colonized were already present in a place before colonizers arrived. "Primary" avoids the implications of "from" and "indigenous," and makes a concession to "admixture."

In fact, Risch discusses admixture of African Americans at length, emphasizing the "Caucasian" contribution to this "population." While this admixture has been frequent, he argues that this should not distract us from the underlying racial structure, that we should be mindful that African Americans *really are* Africans. "Primary," admits admixture, but mobilizes this sense of essential, or at least majority ancestry. "Primary" makes adjustments to a North American conception of race, so that African Americans can be included as Africans. By invoking our ideas about who is African, and subtly modifying the terms of his definitions, Risch strengthens his claims, still within the context of one nested proxy configuration.

The voicing of "Asians" as "from" eastern Asia also mobilizes our ideas about who Asians are, partaking of a certain history of nationalism and recent immigration. While Caucasians have ancestry in Europe and Western Asia—and have become North Americans— Asians are *from* Asia and remain foreign in this production of race.

It might be argued that I am reading too much into these subtle modifications. Whether or not these different sorts of ancestral relationships involve the discursive formations I have speculated on, they are not the same types of ancestry, and are treated as if they were. The links between the terms are nested proxies, and the conditions of these linkages are not made explicit, but left implicit. There is an equivalence implied. There is the appearance of consistency, in the presence of heterogeneity. There is a succession of substitutions that moves the passage through a variety of types of relation in a subtle way.

Nested proxies in Risch grant a certain flexibility to the ways in which ancestry is viewed. They provide a certain shelter to a construction of ancestry by linking several approaches together, homogenizing their difference through close linkage, and producing a final, flexible view that masks its own complications.

Another element of complication, another source of instability helping in the production of race, is perspectival phasing in the passage. It begins "for our purposes here, on the basis of numerous population genetic surveys, we categorize Africans as[…]," strongly indicating its own discursivity by "categorize," and the presence of its author explicit by the pronoun "we." As shown in the first column of figure 2.1, "Name of category," this foregrounded perspectival presentation dissolves as the passage progresses. The "we" implied in "for our purposes," and stated in "we categorize Africans," the use of words and concepts implied in "categorize," and the contingency suggested by "on the basis," give way first to "Caucasians would include"—with some authorship still indicated—and finally to "Asians *are*," "Pacific Islanders *are*" and "Native Americans *are*." The effect is a movement from a strong presentation of perspective, to an erasure at the end.

This configuration neatly describes the arc of a perspectival phase. By modulations in voice and terminology, the text moves from a perspective that is foregrounded to a perspective that is erased. The move corresponds with a unitary set of nested proxies and a coherent set of claims, and works with them to produce race by transferring between natural and social constructions, and by conferring credibility..

The use of perspectival phasing allows a transferal between naturalized and socialized constructions of race. "Categorizing" is undergone, and it has "purposes," and a "basis," at the same time as races "are" independent, natural groups in the world. The change in voicing at the end of the passage is not simple abbreviation, but alters the space in which race is made, combining socialized categorization and naturalized "continental" groups. The natural and the social have a transfer point between each other in these simple changes in the presentation of perspective, and the result is a combined, hybrid racial knowledge. The passage leaves us at a *naturalized* construction of race, but in getting to that point incorporates *socialized* constructions as well.

Perspectival phasing also gives the passage credibility by giving a humble, located sort of perspective first, and then using the credibility gained in this modesty to support naturalized truth claims. By making this an implicit process—no warning being given for the sudden change in voice—the change can go unnoticed, and the benefits of both perspectival presentation and erasure accrue to the claims. That is, Risch gets the best of both worlds: the credibility of accounting for his own methods, in which "we categorize" is an admission of responsibility, and the disembodiment of the god-trick, in which "generative" racial vision can extend to everywhere, and erase its source, lending the sense of universality. Not a bad deal for such small discursive shifts.

The extent to which these two kinds of instability—nested proxies and perspectival phasing— are able to mask themselves, as well as other sorts of inconsistencies is quite striking. The passage above gives a set of criteria that define racial groups. Risch describes them as "continental." But as the third column in Figure 2.1 shows, the central geographic components of Risch's are not, in fact, *continents*. Sub-Saharan Africa is a part of a continent delimited with reference to a desert. Europe and West Asia are a continent and a half (plus a subcontinent)—assuming that West Asia and eastern Asia are equally sized, and not defined only via direction. Assuming this, Asians are defined by half a continent. Australia, Papua New Guinea, Melanesia and Micronesia are, respectively, a continent, a very large island, and two federated nations that encompass many small islands each. North and South America are, roughly, two continents. In fact, if continent = X, then Risch's continental schema is: [some X] : [1.5 X + sub-X] : [0.5 X] : [1 X + island + 2 groups of islands] : [2 X]. These are the criteria for a new production of race.

An comparison of Column 5, "Extensions," also reveals instability and non-equivalence. The categories "African-Americans and Afro-Caribbeans," "Indian Subcontinent and Middle East," "China, Indochina, the Philippines and Siberia," and "other Pacific Island groups further east" extend the definitions of race beyond the larger geographic criteria discussed above. But the categories here are not all geographic. "African-Americans and Afro-Caribbeans" can refer to cultures, political constituencies, diasporic histories, and simply race itself—that is, "African American" can itself mean a *racial* group, and in this passage a racial group is extended to include it. The Indian Subcontinent is a large landmass that juts out from an even larger landmass, and contains several nations, including India, which has the second largest population in the world.

The Middle East is defined with reference to a direction from Europe, lying in between the "Near East" and the "Far East," categories that arose during European imperialism. The Middle East also references a set of cultures, usually implying that they are Muslim. The Middle East encompasses many nations. China is the nation with the most people in the world, Indochina is a region defined by a peninsula, the Philippines are a nation of thousands of islands, Siberia is a region of the nation Russia. "Other Pacific Island groups further east" refers to archipelagoes in the world's largest ocean, which are not part of Micronesia, Melanesia, Australia or New Guinea, but are east of them. If the main criteria in making race are not "continental," but varied sorts of geographical entities, the criteria in their extensions are not even consistently geographic, and include culture, nation, direction, region and race itself.

All of this heterogeneity and instability is presented using configurations that give it the appearance of consistency, and allow it to do work. The result is a production of a full racial architecture, a strong claim about the reality of race, and a set of social and natural ways of structuring knowledge that facilitate reproduction of this kind of race. Credibility is achieved in the shifts of discursive tone, and the presentation of particular kinds of evidence. Risch's article has been received with respect in genomics, including from genomicists high up in NHGRI's echelons, such as Aravinda Chakravarti (New York Times, Wade, 2002). Instability, heterogeneity, nested proxies, perspectival phasing, configurations of nonequivalence are the *dominant* characteristics of this piece. Should we not conclude then, that instability is itself powerful?

These three passages demonstrate the breadth of activity that nested proxies take up in the production of race. They are three examples of mainstream discursive configurations, from three writing styles—journalistic, ELSI and technical-- concentrated

in contexts that make them particularly easy to see. In their containment, and in the density of their categorical replacements, they are somewhat exceptional. Nested proxies are more commonly dispersed over larger textual spaces, on a document-wide level, rather than within a paragraph. This does not mean that the examples given are unique and peerless, but that this type of proxy use is a special case of a broader embedded type of discursive configuration, whose activity is vital to the genomic technology of racial difference.

III. Making Difference Around Race

Besides helping to produce difference within race, nested proxies and perspectival phasing can help to structure and determine the constructions of race in a broad social context. Difference, in this case, is made *around* the racial categories, rather than within them.

. When the Human Genome Project was completed in April, 2003, Congress declared April 25th National DNA Day (http://www.genome.gov/11007057), and Francis Collins and three other genomicists at NHGRI published in *Nature* "A vision for the future of genomics research: a blueprint for the genomic era."(Collins, 2003) This article outlined "III Grand Challenges" for the newly inaugurated era, each with several subheadings. Race was addressed in "Grand Challenge III-2: Understand the relationships between genomics, race and ethnicity, and the consequences of uncovering these relationships." This passage presents race in terms of its position in society, and the criteria by which it may known. It also differentiates among the types of knowledge, and the social positions occupied by the people who can know about race. Its production of race occurs *around* racial categories, rather than *within* them.

> Race is a largely non-biological concept confounded by misunderstanding and a long history of prejudice. The relationship of genomics to the concepts of race and ethnicity has to be considered within complex historical and social contexts.
>
> Most variation in the genome is shared between all populations, but certain alleles are more frequent in some populations than in others, largely as a result of history and geography. Use of genetic data to define racial groups, or of racial categories to classify biological traits, is prone to misinterpretation. To minimize such misinterpretation, the biological and sociocultural factors that interrelate genetics with constructs of race and ethnicity need to be better understood and communicated within the next few years.
>
> This will require research on how different individuals and cultures conceive of race, ethnicity, group identity and self-identity, and what role they believe genes or other biological factors have. It will also require a critical examination of how the scientific community understands and uses these concepts in designing research and presenting findings, and of how the media report these. Also necessary is widespread education about the biological meaning and limitations of research findings in this area (Box 6), and the formulation and adoption of public-policy options that protect against genomics-based discrimination or maltreatment (see Grand Challenge III-1).(Collins, 2003)

The reality of race here is "largely non-biological." The biological component implied in the "largely" is declared when "populations" stand as a proxy for "race:" "certain alleles are more frequent in some populations than in others." The passage produces criteria of racial difference in this way; populations are not defined because of their allelic frequency differences, but already defined populations characteristically have different frequencies. The proxy between race and population allows this move.

Changes in the presentation of perspective also occur here, and facilitate the differentiations happening around race. Examining the first four sentences individually will give us a better sense of how this change in perspective proceeds:

> Race is a largely non-biological concept confounded by misunderstanding and a long history of prejudice.

At the level of authorship, the perspective here is absent, scientific—the authors of this claim are not brought to light. But the words specifically emphasize discursivity and social contingency, giving perspective, if not to the statement itself, then to the people and history in which the

"largely non-biological concept" is situated. The "largely" and "misunderstanding" imply a perspectival position occupied by the subjects of this statement. The effect is that perspective is partially foregrounded. Although the author is not brought to the forefront, his discourse is.

This tone in perspectival presentation is maintained in the second sentence:

> The relationship of genomics to the concepts of race and ethnicity has to be considered within complex historical and social contexts.

The term "genomics" has at least two valences here: genomics as a *field*, and as the *biological facts* uncovered by this field. Genomics in both senses may have a "relationship" with the largely non-biological "concepts of race and ethnicity[30]," negotiating the terrain of "complex historical and social contexts," while remaining distinct from A) the non-biological, B) concepts of race and ethnicity, and C) historical and social contexts. In this sentence, genomics may not be strictly abstracted from A, B and C, but it does constitute a different kind of thing, by virtue of having a "relationship" with the others. The elements of a foregrounded perspectival presentation are present in both genomics and its others, but we sense that the perspective which genomics has (by virtue of being able to do things, by operating within a context, by having relations) is of a more privileged type than these others (by virtue of their conceptuality, largely non-biological nature, and by-definition-sociality). Although its relationships are troubled, genomic knowledge has a certain privilege in making knowledge about race.

As if to perform this privilege, all of a sudden, the terms are changed, and perspective disappears:

[30] "Race and ethnicity" here replaces "race" in the previous sentence, and is subsequently replaced by "populations." "Race, ethnicity, group identity and self-identity" occurs later in the passage. "Concepts" and "constructs" are used interchangeably. These linked substitutions are further examples of nested proxies.

> Most variation in the genome is shared between all populations, but certain alleles are more frequent in some populations than in others, largely as a result of history and geography

While race is conceptual, "populations" are autonomous or self-evident categories. This is a declarative sentence, without authorship or discursivity indicated. In this particular perspectival abstention, that which makes race "largely non-biological" (variation sharing among "populations") is displaced with that which makes race biological and genomic. "Certain alleles *are* more frequent in some populations." The foregrounding of perspective in the first paragraph is, in one sentence, reversed, perspectival phasing happens, and the work of justifying the investigation of racial difference in genomics is done.

And again, in an instant, the perspective comes back:

> Use of genetic data to define racial groups, or of racial categories to classify biological traits, is prone to misinterpretation.

Genomicists here *define* and *classify*. These are explicitly discursive activities. "Racial groups" are discursive objects that may be formed by genetic data; "racial categories" are discursive tools that may organize biological traits. Although authorship of the statement itself may be absent, its topic is the ways in which genomicists can use race in their status as authors, and discursivity is present in this action.

The effect of this phasing is that genomics comes to occupy a broad range of flexible positions in respect to race. Some of its knowledge is socialized and discursive, while its knowledge about "populations" is naturalized. The effect is a variety of transferals across the social and the biological, and this back-and-forth action gives rise to new, complex differentiations *around* race. One level on which these differentiations occur is within genomic knowledge. Another is a set of distinctions between public and genomic ways of knowing race.

The public is differentiated from the genomic community when "individuals and groups" "conceive" and "believe," but "the scientific community" "understands," "uses these concepts," educates about the "biological meaning and limitations." The sorts of shifts in perspective involved in establishing these differences are common in genomic discussions of "misinterpretation."

What does it mean to say racial data are "prone to misinterpretation?" What is the corresponding correct interpretation implied by this "mis-"? Who is doing the misinterpreting? "The public" is never specifically named, and this omission allows a certain flexibility in the attribution of misinterpretation. But the strong suggestion in the course of the passage is that the public is the guilty party. The ways in which perspective is voiced both differ around and help to form the basis of this suggestion of demarcation between the public and genomics. The results of this particular perspectival phasing process are a dichotomy of knowledge processes, represented in Figure 2.2.

Figure 2.2. Perspectival Differentiation in Collins (2003).

The Public	**Genomics**
The site of "**misinterpretation**"	**-minimize misinterpretation** **-understand, communicate**
-individuals and cultures: -**conceive** of race, ethnicity group identity and self-identity -**believe** role of genes or other biological factors	-scientific community -**understands** concepts -**uses** concepts
-**needs** widespread education	-**needs to offer** widespread education

While the public misinterprets, genomics minimizes misinterpretation. While individuals and cultures conceive and believe, the scientific community understands and uses. The public needs education, and genomics needs to offer it. The second half of the passage does little shifting in

terms of the voice of its own perspective, but provides a strong differentiation in the perspective it ascribes to others. The "use of genetic data to define racial groups, or of racial categories to classify biological traits" by genomicists is voiced as problematic not in its own regard, but in its use by subjects whose perspectives Collins' foregrounds more than the perspectives of genomics. This foregrounding is accomplished through the pluralism suggested by "individuals and cultures" versus the unity implied by a "scientific community" in need of a "critical examination." To "believe" and to "conceive" imply at least partiality and bias, if not simple wrongness, while to "understand" and to "use" are more neutral. While the perspective of the scientific community is not *erased*, it is backgrounded with respect to the perspective of the public. The upshot is another production of difference around race, in which knowledge practices are changed depending on who is knowing.

This type of differentiation between the perspectives of the public and genomics is more strongly represented in a report from an NHGRI conference called "Developing a Haplotype Map of the Human Genome for Finding Genes Related to Health and Disease." (NHGRI, 2001) The section "Describing the contributing populations" gives a high measure of perspective to the public, and attributes a strong measure of discursive work to scientists:

> Regardless of whether the individual samples would be identified by their population of origin, any populations that contribute to a haplotype map must be described in a way that does not reinforce the mistaken perception that populations are genetically distinct, well-defined groups. Because people take in information most readily when it confirms their stereotypes, terms related to race and ethnicity must be used with precision, sensitivity, and care. Populations should be described as specifically as possible; for example, if a group of Chinese-Americans in Hawaii were studied, the population should not be labeled simply "Chinese." This specificity of description is crucial to minimize the risk of essentialist definitions of race, which assume that all individuals of a race are genetically similar.

As with the passage from Collins, "the public" here is not named specifically, but indicated in context. There is a division between the "people [who] take in information most readily when it

confirms their stereotypes," and the genomicists who must use "terms related to race and ethnicity" "with precision, sensitivity and care." This distinction is somewhat complicated by the final sentence, in which the trouble with "essentialist definitions of race" is voiced passively. It is clear the genomics must undertake a "specificity of description," but who produces the "essentialist definitions of race," and for whom do they hold currency? Although the work of definition here is explicitly discursive, there are no authors attributed to this work, even by the public vs. science implications evidenced in the rest of the passage. The effect is a transition from a foregrounded discursivity—identification, description and definition are its marks—and a differentiated, implied but clear attribution of perspective—marked by "mistaken perception" and "stereotypes" on the one hand, and use, "precision, sensitivity and care" on the other—a transition from these presentations of perspective, to a mixed and contradictory statement in which discursivity is foregrounded and perspective erased.

Several paragraphs later, in a section labeled "Samples from real populations," discursivity too is erased, in a statement that seems to directly contradict the form and content of the previous passage:

> To obtain the most representative samples, it is important not to use samples of convenience, but to choose samples from real populations. The populations that contribute samples should be chosen based on the goals of the haplotype map, and the samples should be collected with appropriate population consultation and informed consent. (NHGRI, 2001)

What, precisely, is a "real" population? How did we get from populations that "must be described in a way that does not reinforce the mistaken perception that populations are genetically distinct, well-defined groups" to the notion of the "real?"

I believe that perspectival phasing facilitates this sort of move. The perspectival presentation in the first passage ranges from foregrounded to mixed—in the second passage,

perspective is phased out. The two passages perform very different types of work—the first emphasizing the limits of race in the context of an inside / outside distinction, the second suggesting a normative approach to a legitimate study of (racial) populations. Perspectival phasing has the ability to link contradictory statements, and provide shelter for certain types of controversial assessments.

Many examples exist of this sort of move from the limits to the legitimacy of race. One of the more dramatic shifts in statements about race and ethnicity occurs in the Food and Drug Administration's (FDA) draft guidelines on the collection of racial and ethnic data in clinical tests. (FDA, 2003) This is an example in which instability not produces work within discourse, knowledge and scientific practice, but helps to explain and justify a pending piece of agency regulation.

The opening paragraph emphasizes its own discursivity, and describes the racial and ethnic categories it recommends adopting as nonscientific:

> The OMB stated that its race and ethnicity categories were nonanthropologic (in other words, not scientifically based) designations but, instead, were categories that described the sociocultural construct of our society. The Department of Health and Human Services (HHS) chose to adopt these standardized categories for its agencies that report statistics, as the categories are relevant to assessing various health related data, including public health surveillance and research. FDA believes that the use of the OMB categories will facilitate comparisons across clinical studies analyzed by the FDA and with data collected by other agencies. Collection of data using standard categories may enhance patient safety by helping the Agency evaluate potential differences in drug response among subpopulations and may help facilitate analyses seeking differences in response.

Race and ethnicity here are "categories," "designations," and descriptions of the "sociocultural construct of our society," that are not only "nonanthropologic," but not "based" in science. The authors of the knowledge that is employed are made clear, their reasons for doing things explicit, the discursivity of their activities well-known. Race is a social thing that has potential medical implications, and needs to be studied as such.

But in the very next paragraph, race and ethnicity categories are posited as *completely* scientific, and racial differences relevant to drug testing are enumerated. These are differences located *within* racialized groups that are mobilized to support actions and knowledge *around* them. The statements are evidence justifying a recommendation that racial and ethnic data be collected in clinical trials.

> Some differences in response to medical products have already been observed in *racially and ethnically distinct subgroups* of the U.S. population. These differences may be attributable to intrinsic factors (e.g., genetic, clearance 8),extrinsic factors e.g., diet, environmental exposure, sociocultural issues),or interactions between these factors. For example, in the United States, *Whites* are more likely than persons of *Asian and African heritage* to have abnormally low levels of an important enzyme (CYP2D6) that metabolizes drugs belonging to a variety of therapeutic areas, such as antidepressants, antipsychotics, and beta blockers (Xie 2001). Additionally, after using some drugs in the psychotherapeutic class, slower enzyme metabolism (CYP2C19) has been observed in *persons in the United States of Asian descent* as compared to *Whites* and *Blacks* (Xie 2001).Other studies have shown that *Blacks* respond poorly to several classes of antihypertensive agents (beta blockers and angiotensin converting enzyme (ACE) inhibitors)(Exner 2001 and Yancy 2001). *Racial* differences in skin structure and physiology have been noted that can affect response to dermatologic and topically applied products (Taylor 2002). Clinical trials have demonstrated lower responses to interferon-alpha used in the treatment of hepatitis C among *Blacks* when compared to other *racial subgroups* (McHutchison 2000 and Reddy 1999).

Nested proxies are at the core of this presentation of evidence. Links and substitutions are made among the terms and categories "racially and ethnically distinct subgroups," "Whites," "persons of Asian and African heritage," "persons in the United States of Asian descent," "Whites and Blacks," "Blacks," and "other racial subgroups." All of these categories are presented in a way that implies their equivalence. The various studies cited are positioned as comparable, identifying differences in the same sorts of things, even while a variety of relationships and categories define these things. No distinction is made, for example, between "heritage" and "descent." The category

"Blacks" is used interchangeably with the "persons of […] African heritage." "Subgroups," are "racial" and "racially and ethnically distinct."

The effect of these proxies is to allow the racial differentiation to encompass many sorts of difference, giving flexibility and breadth to the evidence. This unstable configuration allows the studies to be compared—if consistency were practiced, this could not be done, since different terms and constructions of race are used in their reports. Nested proxies link a variety of evidence together, allowing it to be homogenized and mobilized for a regulatory justification.

Perspectival phasing also participates in this process. Despite the presence of cited authorship (Xie, Exner, Yancy, Taylor and McHutchison), authorial voice and indication of discursivity in knowledge practices, with authors labeling, classifying, defining, etc.—all of these are absent, and so perspective is erased. As with Collins, Brower, and the passage from NHGRI, the section that describes the limits of race presents a foregrounded perspective, and the section that advocates for the legitimacy of racial differentiation erases perspective. The effect is that racial differentiations are naturalized when the reality of race is emphasized, and socialized when their limits are stressed.

Through changes in perspectival presentation and nested proxies, race is produced both in terms of internal differentiations, and external positioning. Through the mediation of this instability, an FDA regulation produces and uses racial knowledge, rooted in genomics, to justify and initiate practices with respect to race.

Perspectival phasing and nested proxies are not simply a stylistic anomalies, nor mere rhetorical devices. These formal configurations constitute a discursive site of power in the genomic technology of racial and ethnic differentiation. Transitions along

axes of terminological instability and perspectivity structure and enable certain ways of differentiating, both within and around race. The result is a production of race from all angles. Racial attributes are specified, racial difference is invoked and expanded upon, racial architectures are proposed and justified, criteria for the reality of race are formed, inconsistencies are masked, a differentiated social position for knowledge about race is made, policies involving race are recommended, and transfers are made between traditionally biological and social categories of knowledge. That instability can do all of these things, that particular patterns of discursive configurations can define and shelter these constructions of race, gives us an indication of how strong they are, and in turn, how far the genomic technology of difference can reach.

Chapter 3: Instability and Discourse

I. Reading and Writing

This thesis has shown some of the ways in which terminology and formal configurations in genomic discourse allow the production of race. The analysis so far has focused on terminology and configuration separately. Considered jointly, they constitute a discursive system whose defining characteristic is an instability that facilitates a variety of racial articulations.

In synthesizing the conclusions so far, there is an apparent disconnect between the analyses of the first two chapters. Chapter 1 is concerned with defining terms. Chapter 2 identifies a formal configuration, nested proxies, in which the meanings of these terms are elided. How can terms both have discrete definitions, and be linked together in such a way as to produce meanings that substitute for their definitions?

In part, these conclusions correspond because discrete definitions and elided meanings are not always coincident. Nested proxies do not occupy every space in every document of genomic discourse. Genomic terms with discrete, if multiple, definitions can occupy textual spaces in which no nested proxies are employed. Meanings can change depending on whether terms are placed inside or outside of proxy relationships.

Another part of the answer to the question of correspondence is that the theory of nested proxies and the critical vocabulary of genomics offered by this thesis are theoretical tools, intended to give analytical teeth to investigations of racialized genomics. As analytical tools, there are limits to their use, and problems with their correspondence.

But it is not necessary to say that these are *only* theoretical tools to explain the problem of correspondence between discrete and elided meanings. The crucial distinction between the analyses produced by the first and second chapters is a difference in reading practice.

Reading practices, for the sake of this thesis, are the systematic interpretive processes by which meaning is read from discourse.[31] These are processes that enact complex relations between individuals, culture and words in the production of meaning.

These processes can be viewed at the level of the individual and at the level of groups of people. Biologists, for example, have reading practices different from semiotic theorists. Biologists reading biology employ reading practices different from biologists reading semiotic theory. And, most importantly for the subject of this thesis, biologists can employ different reading practices in reading biology, even within the space of one document, and even simultaneously.

Different reading practices can produce different meanings. A biologist reading a journal article critically, for example, might engage with that discourse in a way that allows an analysis of underlying assumptions. Other reading practices employed with the

[31] Introducing the process of reading moves the scope of this thesis out of strictly discursive analysis into an interactive and sociological realm. This opens a whole new set of issues, which I am not in a position to pursue fully. While the introduction of reader response theory is useful to make this point, my central focus remains within discourse.

same article could allow meanings to be produced about the author's intentions, a generalized summary of the main points, an analysis of style, etc. These sorts of practice are seldom pure and isolated, and combine with each other in meaning-making responses to discourse.

This thesis has employed two major reading practices in its analysis of genomic discourse. A critical analysis of terminology resulted in multiple discrete meanings. A critical formal analysis produced the theory of nested proxies, in which meanings are substituted and fluid.

Looking at building blocks, terms, categories, keywords, allows a theory of "what gets said," and "what it means." Reading in this way produces definitions, shades of meaning, different senses in which words can mean things. Generalization is possible from this analysis towards a critical vocabulary of terms and contexts in racialized genomics.

Looking at discourse in action, looking for patterns and formal configurations, allows a theory of how—technically—these things are said. Reading for configurations not only allows a new sort of perspective to arise, but new kinds and valences of meaning to appear. Relations, claims, substitutions, erasures, links, elisions, all of which do and mean things become available to this reading practice.

These meanings are not mutually exclusive with the first. Because reading is done in different ways simultaneously, because understanding is possible both for terms in themselves, and for relations among terms, we can say that both kinds of meaning can exist at the same time.

Another reading of discourse, which asks not only what is meant, or how this meaning is achieved, but what the conditions of possibility are for such meaning,

produces a different analysis again. Not a new type of meaning, as with the first change in reading practice, but a look at the structuring principles of knowledge that allow such meaning to be produced. This is an analysis of epistemology, which is the subject of the final chapter.

The example from Brower is a good illustration of this process of changing reading practices (2002). The quote from Elijah Saunders is a direct definition of race: "we are not talking about race as a genetic situation. We are talking about it as a marker." Using a reading practice that focuses on "what is said" and "what it means," produces an analysis of the term "race." In this context race is a "marker," and race is not a "genetic situation." Similar definitions and usages can be found in the rest of the literature, if a similar reading practice is employed. In this way, a generalized definition for "race," can be produced. From this starting point, it is possible to identify and articulate other definitions and usages, and establish a schema that labels this the "heuristic" idea of race. This reading practice defines, compares, distinguishes, schematizes and generalizes. It is not necessary to claim that these practices are absent or neutral in order to claim that these are accurate, useful ways to analyze meaning. Neither are they mutually exclusive with other ways of reading.

Turning to patterns, looking at the whole passage, armed with definitions for terms, allows the configuration of nested proxies to come into resolution. Terms are stretched out, proxied for, elided, erased, made flexible in the service of a larger meaning about drug response and risk for heart disease. It's not that the individual terms don't mean what they used to mean. It's not that the first analysis was wrong. It's that meaning is not independent of reading. Reading for formal configuration, the terms in themselves no longer bear the main weight of signification. When we read in this way,

the relations among terms are that which *means,* and different types of meaning emerge. It is true that race is a kind of "heuristic." It is also true that race is the endpoint of a number of elisions that destabilize this meaning and say different things about it. So the meanings of the terms do change. They change depending on reading practice.

These distinctions explain the apparent disconnect between the previous two chapters. Several reading practices participate in the production of meaning in racial genomics. Consequently, it is necessary to employ several reading practices in order to understand the discursive production of race.

But reading practices are not the only processes through which genomics relates to discourse. The analysis up until this point has moved among assertions about genomic discourse in general, and assertions about discursive examples attributed to particular authors, including statements about the authors' institutional and social positions. The link between genomic authors and genomic discourse has been left unstated, or at best implied that examples from certain authors exemplify certain trends of the general discourse.

Regarding the question of exemplification, particular examples have been chosen as illustrations of wide discursive patterns. For this reason, the passages quoted have been selected as fragments in which these patterns are especially clear and prominent. In this sense, the passages are by no means average, but are exemplary of patterns that are often more muted or dispersed. This method of characterization deliberately highlights certain elements of discourse, in the same way that divergent reading practices produce emphases on different levels of meaning.

At first glance, the relationship between authors and discourse does not seem similarly complicated. It seems that genomic authors write genomic discourse, in a one-

directional process. But in the same way that reading practices involve an exchange of meanings among discourse and the reader, the relation between discourse and authors is one in which power relationships are mutually exchanged.

Discourse exerts power within authors by mediating the parameters within which claims are made, empowering authors to operate within a set of meanings, and providing structures for the production of knowledge. Authors exert power within discourse by reproducing elements of discourse, lending professional credibility to statements and ways of making statements, and producing new terms and forms.

Scientists have an especially broad latitude in their power to produce new terms and forms. Science as a field produces knowledge that frequently depends on a novel vocabulary. The sorts of phenomena that scientists articulate often do not have pre-existing terminologies that are adequate to their purposes. New terms and taxonomies are commonly made in the service of articulation.

Also, science as a profession involves expert status, which, within a complex set of criteria, can give weight to new discourse which scientists produce. A non-scientist can come up with new ways of talking about fungal diseases, cellular structures or Martian geology all she wants, but there is little chance that his new discourse will be reproduced or ramify broadly, without the social accreditation of scientific authority.

This should not be taken to imply a monolithic "science," which is unitary, general and transhistorical. Historical differences in the production of knowledge, and differences among contemporary fields of scientific knowledge are complex and vital. There is no one "scientific method," and no one "science" which have worked together to produce discourse in a regular fashion.

It is possible, however, to make the generalization that scientists produce discourse in unique ways, at a higher rate than in many other professional spheres, and with a higher level of privilege attached to the results.

We can also say that discourse and authors exert power on each other, although not on equal, circular terms, but within a complex field of relation. With respect to genomics, we can say that discourses of race exert power by empowering authors with a set of meanings, setting parameters within which meanings can be made, providing frameworks for research practices, and informing the production and reproduction of knowledge. We can say that authors exert power by advocating, modifying, deploying, invoking and lending credence to discourses of race in a variety of forms. We can also say that in special circumstances, genomicists can produce new discourses of race.

The contention of this thesis is that the technology of difference that is a combined way of looking at authors, discourse, epistemology and knowledge is doing something more extraordinary than *using* race. Genomics, I argue, is *making* race.

The remainder of this chapter reviews and synthesizes the attributes of discourse I have theorized, whose dominant characteristic is instability. This instability is crucial in both the reproduction and the production of race.

II. Articulate Instability

Instability is the dominant formal and terminological theme of genomic discourses of race, and the theme that has dominated this inquiry. Instability means that the meaning produced within discourse shifts, changes, fluctuates and comes in a heterogeneous variety of packages. Instability means that there is no single referent—whether "in

discourse" or "in nature"[32]—around which significance revolves. It is not simply that the people to which racial categories refer changes, although this is the case. It is not simply that the categories themselves change, and different labels are used for roughly the same things, although this is also the case. It is not only that claims are contradictory and various, although there are many divergent claims in the literature. Rather, instability is all of these things and others besides. It is instability that *resolves* instability, flux that *removes* its own contradictions, and shifts in meaning that *empower* themselves and other meanings. Instability, in short, is articulate.

That instability is characteristic of any particular racial discourse will come as no surprise to race theorists. Instability is the very life of race. What may be surprising or novel are the forms which instability takes here, the politics in which it is produced, its particular relation with a new science, and its capacity to *make* race.

That race is strongly involved with scientific practice, discourse and epistemology will similarly be of no surprise to scholars working in the various interdisciplinary fields of science studies. What may be novel in this analysis are discursive habits quite distinct to race genomics, and the somewhat unusual combinations of epistemology elaborated in the next chapter.

In order to move towards some more credible claims for novelty, and to synthesize the conclusions of the previous chapters, it is useful to review and summarize the various kinds of instability so far discussed. Instability happens in at least five forms in the genomic discourse of race: multiple definitions and contexts of use for terms,

[32] Quotation marks here denote that discourse and nature, although probably impossible to distill into separate spheres, are useful ways to characterize the ways in which contexts in which claims are made, as having to do with words or having to do with material things. This is the same type of move as the technical/political, and biological/social distinctions made in previous chapters among types of discourse.

nonequivalent categories of comparison, nested proxies, perspectival phasing, and contradictory claims.

The multiplicity of definition in terminology provides both flexibility and strength to genomic discourse of race. The availability of terms that euphemize or suggest race without being the term "race," allows authors to avoid a certain discomfort with the term. Saying "diverse populations," for example, is less touchy than saying "different races."

Besides allowing a certain polite avoidance, "diverse populations" can include other systems of difference, while maintaining an emphasis on race in context. There is, then, an instability among terms—"race," "population," "community," the ever-flexible "group," etc.— as well as instability within terms, such as the different senses of "population." These modes of instability have functions that include avoidance, implication, and a certain broadening or making flexible of the points under discussion.

Terms also allow the topics of discussion to extend well beyond the boundaries traditional to scientific fields. Rather than limiting writing to the topics of individual genomes, allelic distributions, functional variation and medical implication with relation to race, genomic authors also write about ELSI concerns, institutional integration, racial inequalities and what it means to be human. In some situations the terms overlap, as with "variation" or "population." In others, their meanings are limited to one context or another, as with "admixture" or "minority." It seems that more terms in racial genomics cross over the technical/humanist, biological/social boundaries than remain within one or the other. It also seems that the overlap facilitates an exchange of credibilities—the authority of scientificity is transferred to political statements, and the humanitarian tone of liberal humanist politics marks technical ones. Instability in this context of exchange, then, broadens the topics able to be addressed by genomic discourse, and mediates a kind

of cross-fertilization of knowledge and authority. Again, instability is making things *possible*, not hindering them.

At times the multiple sense of terms are explicitly differentiated. The delineation of "race" versus "ethnicity" in Schwartz (2002) and Witzig (1996), and in the first part of Risch (2002) are examples, in which "ethnicity" is marked off as a distinct from race, if often coincident. Ethnicity for the first two is a more legitimate subject of inquiry than race, and for Risch of secondary importance. It could be argued for these examples, then, that "ethnicity" is not an example of instability, but of simple articulation.

But the elided "race and ethnicity" is at least as, if not more common than an articulated scheme of distinction between the two. Fascinatingly, they even coexist. For example, Risch proceeds immediately from distinguishing the two to conflating them, all within the confines of one paragraph (2002). And so another axis of instability emerges, in which distinctions among terms can fade in and out.

This instability of terminological distinctions seems to happen in some cases as it becomes rhetorically convenient. In other cases, these shifts happen and for no apparent reason at all. This would seem to indicate an *inherent* instability in this racial discourse, one that is not only functional, but also simply characteristic.

Whether functional or characteristic, unstable terms also participate in the attainment of divergent conclusions that can be expressed with the same sorts of terminology. Risch and Morris Foster, for example, have both produced commentaries about the significance of race. Risch argues that it is the most fundamental genetic structure in the species, Foster that it is a potentially useful heuristic starting point. Although there are tonal differences—for example, Foster does not invoke his own "objective perspective" multiple times as Risch does— the major forms of their

vocabulary are the same. The divergence in their claims is disagreement, not instability *per se*, but instability is a fundamental component in its production. Foster and Risch, after all, are both talking about *race.*

The use of nonequivalent categories of comparison in these sorts of discussion is another sort of instability related to the multiplicity of terms. "Nonequivalence" is somewhat provisional, intended to indicate the comparison of heterogeneous categories in assertions about racial differences. It is not intended to mean the successive substitutions of categories that form nested proxies, but an immediate juxtaposition of categories that identifies them as comparable entities.

Good examples of nonequivalent comparisons include the construction of populations for the Haplotype Map, as "Yoruban," "Japanese and Chinese" and "northern and western European." This method of comparison involves heterogeneity within these categories, as well as among them. "Yoruban" is a contested term that involves historical, cultural, religious, ethnic and population-size inconsistencies. The three categories differ in number of members each by an order of magnitude, and each refer to different sorts of difference.

The way in which these nonequivalent categories are presented for comparison implies consistency and comparability. More than that, the linkage of incommensurable terms in particular ways *produces* consistency and comparability.

This production is part writing practice, part reading practice. Writing, scientists name units of difference, frame them in particular ways, and bring the units into comparison and attribution. Readers participate in the process by naturalizing the comparisons. That is, for heterogeneous comparisons to work, they must be accepted as either natural orders, or useful frameworks for explaining natural orders. The production

of knowledge about difference, then, lies both with the author and the reader. I think that in our role as readers we most often naturalize heterogeneity—instability—of this type without evaluating the terms of comparison. This response draws from a long history of naturalized racial discourse, which has made such comparisons ordinary. Within discourse, and in the reader response to discourse, nonequivalence masks itself, resolving its own contradictions. By simply offering comparisons without explaining their criteria, authors position nonequivalent categories as having no criteria in need of explanation. From this starting point, authors can use their units of difference for a number of purposes, including using them as independent variables in research, and in making particular claims about difference.

Nonequivalent categories of comparison tend to equalize types of difference, making them comparable. Metaphorically, they make apples and oranges into red apples and green apples, which differ in ways that can be compared.

To extend the metaphor, nested proxies tend to make apples and oranges into juice. The form of instability produced within nested proxies turns divergent individual categories into fluid, mobile relationships. Their use extends beyond the production of categories for comparison, allowing claims to be made, tones to be modified, and implications to be expanded in the production of race.

The relationships established by nested proxies, like those enacted by nonequivalence, combine reading and writing practices. The writing component of nested proxies involves the assembly of incommensurable terms in succession, such that they can substitute for each other without making the process of substitution explicit. But in order for these substitutions to work, the reader must read them *as* substitutes, and help to produce a meaning in this context. Terms can be arranged in close proximity until our

hands are exhausted and our computers obsolete— without a reader linking them in causal, logical, tonal, attributive or other relationships, substituting one for another, they are not *proxies*, they are just disparate statements.[33]

The instability embodied in the nested proxies formal configuration has its contradictions resolved in dialogue with reading practices, and allows a number of relationships to be established with respect to racial difference.

Perspectival phasing is another level of instability that participates with nested proxies and the others in the formation of racial difference. In some cases, this configuration functions to naturalize some such differentiations, and socialize others. Several examples of this process show perspective being erased at strategic points, which leads to the naturalization of race. For example in the passage from Collins (2003), the following sentence, which presents the crucial justification for racial genomics research, is flanked on either side by statements that highlight their perspective. This sentence, however, erases its perspective, naturalizing racial difference: "Most variation in the genome is shared between all populations, but certain alleles are more frequent in some populations than in others, largely as a result of history and geography." In this sense, phasing may be aligned such that certain controversial or sensitive things are made to seem natural and objective, and others— often statements that mesh well with genomic humanism— seem more social and discursive.

But perspectival phasing cannot always be expressed in terms of what it allows, resolves, facilitates and makes. It produces effects, but not always in a systematic or regular fashion. It seems that phasing is a form of instability provoked by the instability

[33] It is important to consider as well, that an author is also a reader of her own work. She employs reading practices to know whether what she writes coheres. Again, the relationships among authors, readers and discourse involve complex exchanges of power.

of the racialized contexts in which it appears, in addition to contributing to it. This formal configuration, in other words, seems to be most often a *result* of the surrounding instability inherent in genomic race, such that authors occupy different positions in their own texts in response to the other shifts within those texts. And once established, the unstable presentation of perspective can then feed back and contribute to that which provokes it.

In the contexts of all these sorts of instability, and by means of them, claims about race in genomic literature are themselves unstable. Seemingly opposed constructions of the meaning of race, the structure of race and the validity of race in genomics occur throughout the literature. This is particularly true when one compares statements leading up to the 2000 HGP working draft completion, with the subsequent shift in focus to variation, and the initiation of HapMap in particular. But such contradiction and controversy also appears without reference to this shift, across the past decade. Even within a single issue of a journal (e.g., NEJM, May 3, 2001), and even within a single article (e.g., Risch, 2002; Collins, 2003), claims seem to contradict each other. And yet no one's grants are revoked, no one's research is stymied, and no one's head explodes from the sheer impossibility of paradox.

The central reason that such discursive contradictions are possible is that instability is articulate. Instability produces possibilities for statements, makes pathways for producing structures of difference, and both constitutes and resolves contradictions. The system works. It is a functional technology of difference with a broad range of capabilities.

In order to produce this discursive instability that functions in the articulation of race, particular epistemological conditions must be met. Epistemology encompasses the

structuring principles of knowledge, that guide its production and work with discourse to shape its expression. The genomic epistemology of race is the subject of the final chapter.

Chapter 4: Epistemology

I. Definitions and Methods

Epistemology is that which guides the production of knowledge and provides the conditions of its possibility. Epistemology provides the "rules" in knowledge production, influencing ideas of what knowledge is valuable, how it is produced, what counts as evidence in this production, and how knowledge becomes authorized.

Although epistemology acts like "rules" in the sense that it gives structure to knowledge, genomic epistemology of race does not resolve into discrete rules and statements. There is no one formula for what counts as evidence for the reality of biological race, for example. The ways in which evidence is produced and presented for race are contested and complex, so that it is impossible to say that in genomic knowledge, any one evidentiary criterion is privileged.

Even if it does not resolve into simple structures, epistemology guides the production of knowledge in ways that can be characterized on the basis of a discursive analysis. It is possible to perform this analysis because knowledge is structured and represented within discourse. Epistemology and discourse exist in a complex network with each other.[34] Nested proxies, for example, are configurations within discourse, where categories are elided and meanings linked and altered. But nested proxies can themselves constitute a kind of evidence. Individually the claims about hypertension and medical response in the example from Brower would not work to support the linked claim on the reality of race, or the attribution of a medical difference to a racialized group (2002). With the nested proxies configuration in place, the claims coalesce to become discursively grounded evidence both for a construction of "race" as "marker," and for hypertension as a feature of a racial group. Epistemology and discourse work together to allow this sort of evidence to be produced within the configuration. Discourse and epistemology exert influence on each other.

Because of these integral power relationships, an analysis of discourse indicates some of the epistemology by which knowledge is structured. The patterns of discursive

[34] Although they have not been the subject of this inquiry, factors like economy, material objects, practices, institutional structures, culture and politics also engage with epistemology in ways that are as complex and vital as discursive relations.

instability that this thesis has theorized arise within certain conditions of possibility. Epistemology allows the shifts in meaning that characterize genomic discourse. Examining discursive patterns for their functions, dispersals and conditions of possibility leads to a characterization of the epistemology that guides the production of knowledge about race.

Like discourse, the dominant theme in the genomic epistemology of race is iself instability. To illustrate this point, this chapter presents characterizations of epistemology using historical modes of racial thinking, including the one-drop rule, white normativity and racial essentialism. The function of these modes of thinking is to provide contexts, to mediate justifications, and to organize knowledge. They are productive sites, in that they are systems whose dominant power is to allow, rather than to forbid results. These historical characterizations of epistemology, while strongly indicated in some ways, are inadequate in explaining the epistemology that structures and allows genomic knowledge. The inadequacy of these explanations is evidence that race produced in these ways is new knowledge that uses elements of old racial thinking without simply recapitulating them. This is a novel epistemology of race that allows discursive instability. This novelty provides further support to the main idea of this thesis, that genomics makes race.

The chapter concludes with an analysis of the presentation of evidence in racialized genomics. The presentation of evidence is described as occupying three evidentiary spaces. Evidentiary spaces are the locations in which truth claims are made and adjudicated. In the case of genomics, they are the spaces of the genome, the social world, and bioinformatics. These are not purely discursive spaces—they combine

discourse, material, economy, people, practice, etc. They are the terrains on which the production of racial knowledge is played out.

The characterizations that lead to these claims about epistemology are meant to be read as speculations. They are potential beginnings for further lines of research—"heuristic starting points" in the study of genomic knowledge. The central intention of these speculations, and this thesis, is to show genomics *making race*. The contentions presented here about epistemology are meant as evidence that such a production is undertaken, and as conjectures about the anatomy of that process.

II. One Drop

The one-drop rule is a system of thought that has furnished the United States' primary answer to the question "who is black." (Davis, 1991; Omi and Winant, 1986; Rockquemore and Brunsma, 2002) The one-drop rule is specific to North America, as well as specific to the categories "black, "Negro," "colored,"[35] "Afro-American" and "African American." People in other countries are raced differently, and people of races that do not include these categories associated with "black," are not raced in this way in the United States. Also referred to as hypo-descent,

> This descent rule requires Americans to believe that anyone who is known to have had a Negro ancestor is a Negro. We admit nothing in between… "Hypo-descent" means affiliation with the subordinate rather than the superordinate group in order to avoid the ambiguity of the intermediate identity…The rule of hypo-descent is, therefore, an invention, which we in the United States have made in order to keep biological facts[36] from

[35] "Colored" being a category different from "people of color," which now refers to those not raced as white.

[36] "[B]iological facts" here are particularly relevant to the topic of this thesis. Biological facts are positioned as true evidence for a notion of race as fiction. Biology is anathema to race, particularly with respect to the dictates of hypo-descent. Rockquemore and Brunsma provide a more recent echo of this argument: "Race, and therefore racial

> intruding into our collective racist fantasies. (Harris, 1964 *in* Omi and Winant, 1986)

Two complementary causal explanations are generally given by historians and race theorists for the rise of the hypo-descent system during colonial and antebellum slavery. The first is economic: it was financially useful for slave owners to class people with "mixed ancestry"[37] as Negro,[38] because it expanded the number of people able to be enslaved within contemporary justificatory frameworks. The second, linked explanation

---

categorization as it is commonly understood, is simply not a biological reality. Because race is a social delusion, it requires an elaborate set of rules and regulations to maintain [in context, this refers especially to hypo-descent]. These rules, of course, would be unnecessary if it were biologically real."(2002)

This point is complicated by the presence of two easily conflated definitions of "biological." The term combines "pertaining to the field of biology" and "having to do with the true reality of life-systems" Arguing that race and its rules do not relate to the true reality of life-systems may be a powerful strategy. Arguing that race does not pertain to the field of biology has been supported by a dominant narrative in the history of science, which holds that following World War II, a new liberal humanist settlement, crystallized by the UNESCO Statements (1952), resulted in the removal of race from mainstream biology, and its replacement with population.(e.g., Haraway, 1997; Stepan, 1982) Race, in this conception, has returned only lately to biomedicine, mirroring "the fierce resurgence of explicit racist, sexist, and class-biased discourse of many kinds all over the world, and exuberantly in the United States."(Haraway, 1997) Jennifer Reardon has argued convincingly against this narrative, asserting that race was never "replaced" with "population," but rather *absorbed* into it (2001). Also, despite a loosening of the connections between race and culture (cultural hierarchy in particular), race maintained currency within biology, and population genetics especially. Another point that undermines the assertion that race does not pertain to biology as a field, is simply that biologists and doctors *use* the one-drop rule, as discussed below. The evidence presented throughout this thesis shows that biological knowledge and race in its many formulations, are not mutually opposed.

[37] Scare quotes here, because even ancestry has no stable referent. Ancestry can mean "race," "geneaology," "ethnicity" and "nationality." The nested proxies in the Risch (2002) example should demonstrate this, in which ancestry of several types and origin are conflated. Also, "mixed ancestry" as a term carries none of the connotations of sexual violence by slave owners, in whose context the system was formed. "The slave-owning mentality included a belief that white male slave owners had the right to sexually 'use' their black female slaves at will. As a result, the vast majority of interracial sex consisted of exploitative unions between white male slave owners and their black female slaves."(Rockquemore and Brunsma, 2002)

[38] Following Jennifer Reardon's (2001) lead, I preserve period terminology both to avoid anachronism and to avoid implying a stable referent.

is ideological: the maintenance of doctrines of white supremacy depended on notions of purity, racial and sexual hygiene, and categorical stability generally.(again, Davis, 1991; Omi and Winant, 1986; Rockquemore and Brunsma, 2002, as well as Haraway, 1997) In this context, notions of racial "miscegenation" arose as preventative taboos. But the brunt of the taboo applied only to Negro male relations with white females.[39] Hypo-descent provided a means to resolve the effects of this one-sidedness, while maintaining racial purity as well as white supremacy. Classing all "mixed" children as Negro resulted in two, clearly defined races, and one clearly defined power relationship, all untroubled by intermediates.

The application of hypo-descent has continued since emancipation, performed both by white people and by people whom it has made black. Excluding small, local "mulatto" categorizations that held brief tenure in some parts of the antebellum South, the one-drop rule has applied totally in the post-Civil War U.S. until very recently, when it has begun to apply *almost* totally. The biracial movement that is the subject of Rockequemore and Brunsma's book "Beyond Black," has established some spaces within the system, but was unable to achieve a multiracial category in OMB's 2000 census, largely due to opposition from African American activists, who argued that it would undermine gains in political organization in militating against racism. This should draw our attention to the contemporary multivalence of the rule. A system rooted in white

[39] The term "white" gained currency around 1680 in colonial America, supplanting "Christian," and then "English" and "free." Its greatest period of instability was the late 19th century in response to immigrations. "[P]olitical and ideological struggles emerged over the classification of Southern Europeans, the Irish and Jews, among other 'non-white' categories. Nativism was only effectively curbed by the institutionalization of a racial order that drew the color line *around*, rather then *within*, Europe."(Omi and Winant, 1986) The terms "white" and "Caucasian" may be entering a new period of instability within genomic discourse, as evidenced in classifications such as Risch's inclusion of the Indian subcontinent.

supremacism has been reappropriated as a unifying factor in combating white supremacism.

Hypo-descent appears to be unique in the world. Two brief examples illustrate differing regimes of racialization. Members of the same immediate family in Brazil can be classed as being different races, according to a "phenotypic" analysis based largely on skin color (*cor* most commonly, or *raça*).(Omi and Winant, 1986; Parra, et al. 2003) In this system, race has nothing whatsoever to do with ancestry, but with a schematized way of interpreting appearances.

In Australia the racial regime has functioned to reduce, rather than expand the number of people classed as black. The Australian government undertook a seventy-year program (roughly 1900-1970) designed to re-socialize "mixed" Aboriginal people (blacks), into subservient positions within white society. This was done by police abductions of children and infants, and relocation into government or mission schools. Once acculturated, they could have children with whites, and in several generations their offspring would *become* white. Rather than maintaining or expanding a racial group, the policies whose results are known as "the Stolen Generation" were designed to gradually eliminate Australian blacks, a move consistent with a variety of approaches to genocide in that country. Conversely, many people who identify as racially black (although pan-cultural terms like Murri and Koori are more commonly used) are read as racially white by white Australians. (e.g., Graham, pers. comm. 2002)

In summary, the one-drop rule is a nationally and historically particular epistemology of blackness—a technology of difference that produces knowledge about who is black. Its roots lie unambiguously in white supremacism and slavery, but its current embodiments are politically complex. The one-drop rule plays a role in structuring genomic knowledge.

One consistent usage of the Rule in the literature occurs in the notion of "founding populations" for African Americans. "Founding populations," refer—unstably—to the people who first colonize a piece of land. They reproduce, history happens, and we are left with current populations, who can trace their ancestry back to these founders. "West African" populations—especially those from Nigeria and Ghana— are commonly asserted to represent the founding populations of African Americans for the purposes of medical genomics. To see the use of hypo-descent it is useful to trace this assertion in context.

Two examples of the use of founding populations from NHGRI's "Five-Year Strategic Plan for Reducing Health Disparities," a report submitted to The National center for Minority Health and Health Disparities (established by the 1993 NIH Revitalization Act) illustrate this point (NHGRI, 2001). The sampling approach for NHGRI's collaboration with Howard University in the "Africa America Diabetes Mellitus Study" (AADM), is discussed:

> "Because of the high frequency of environmental risk factors for Type II diabetes in the African American population, it is more productive to study genetic risk factors in West Africans, since they are thought by many anthropologists to be the founding population of modern African Americans and have fewer dietary and nutritional confounding variables."

A process of analogy occurs here—a process, actually, of proxy formation. West African populations can serve as proxies for African Americans not only discursively, but in practice. West Africans are particularly valuable to this research goal because they are asserted to be more homogenous in terms of diet and nutrition.

Although the assertion employs a certain idea of history, the analogy is placed out of time: the group called West African is not voiced as having ancestors who were the

founding population of African Americans, but rather as *being* the founding population. So there is a second proxy at work, that of modern day West Africans for West Africans during the slave trade. What knowledge practices allow this proxy relationship to be established?

Two bounded populations are required. The bounding may be strict, or rest more on a strong tendency. In the second instance, there must be some sort of core at the center, a stable axis about which a population can be defined. Homogeneity is linked to this—there must be some unity, some sameness implied in bounded populations, and it must persist across history. Ancestry, in turn is required for homogeneity. There must be stable, direct lines of ancestry that link African Americans to their founding population, and West Africans to previous West Africans. They must be parallel, with a minimum of divergence. Epistemologically, there must be a system for doing all of these things in order to arrive successfully at a notion of founding populations.

At first glance, the historical operation of the one-drop rule should confound or at least trouble all of these assertions—bounding, homogeneity and ancestry. For "African Americas," boundaries have been transgressed and repositioned, "homogeneity"—if such a genetic situation actually existed before the tenure of the North American slave trade—is even more doubtful now, and to assert direct unbroken ancestry is to be ahistorical. The history of Ghana and Nigeria complicates the situation even more, in that neither existed as nations prior to independence from Britain in 1960 (Library of Congress website, 2003). The current nation of Nigeria is estimated to encompass 250 to 400 "ethnic groups" (ibid). A margin of error of 150 whole ethnic groups should does nothing to build confidence in bounding and homogeneity. Furthermore, the whole of human genetic variation is asserted, *by genomics,* to have arisen and to persist in Africa,

and 85% of human genetic variation is asserted, *by genomics*, to exist within, rather than between races (e.g., Jorde, et al, 2001; Disotell, 2001; Barbujani, 1997).

In spite of all this difference, not only are West Africans and African Americans made to exist as stable categories, but a proxy can be made between them. What allows these purifications in the face of such complex histories? My contention is that the one drop rule is at play. What other force in American cultural history *asserts*—in terms of genetic kinship—bounding, homogeneity and ancestry for African Americans?[40] Figure 4.1 summarizes these arguments, showing the questions that must be answered in order to produce such relationships, and the ways in which the one drop rule furnishes them.

The idea of founding populations represents an assertion about what and who African Americans *really are*, an essence that requires a notion of hypo-descent to cohere, to maintain historic continuity. In this context, the one-drop rule's influence can allow other proxies as well. In the following passage, also from the Five-Year Strategic Plan, people from Barbados serve as proxies for African Americans, and both are proxies of West African people:

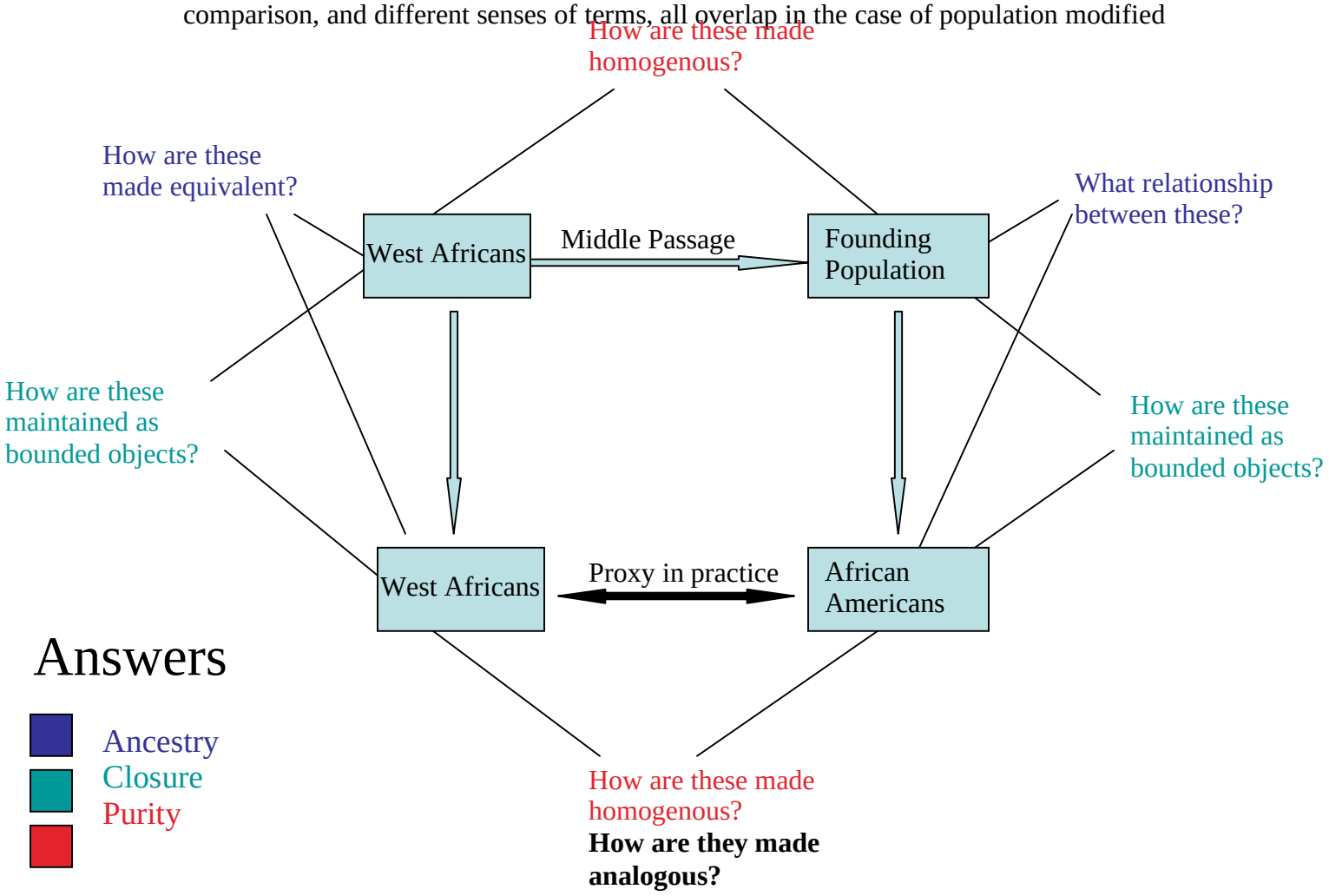
> "The island of Barbados, West Indies, has a relatively homogenous, predominantly black population originating from similar West African regions as the US African American population. The people of Barbados represent an ideal population for genetic-based studies of diseases prevalent in populations of West African ancestry. Prostate and breast cancer are the most common cancers in Barbados. Prostate cancer is substantially more common in African American, and especially Afro-Caribbean men, than in Caucasian or Asian men. Breast cancer, while not

[40] Let me emphasize the *genetic* and the *descent* here. This argument does not refer to political, cultural or community level unity, coherence, similarity, etc. Nor does it negate African *identities*. One can assert that one is, and live as an African on a variety of levels. The ways in which genomics takes up and modifies these issues are of more central concern to this thesis. I believe genomics' racial determinations have the potential to effect the ways in which raced lives are lived and that analyzing them is critical at this juncture.

more common in blacks than whites, leads to higher mortality in black populations."(NHGRI, 2001)

Here are discourses of homogeneity, race, population, history, geography, origination, idealization, disease and relation. What allows these heterogeneous types of category to be linked? Instability again is articulated. Nested proxies, nonequivalent categories of comparison, and different senses of terms, all overlap in the case of population modified

Figure 4.1: One Drop Rule and Founding Populations in Genomics



The epistemological system of the one drop rule provides answers to problems posed by the genomic construction of founding populations. As a system it offers assertions of ancestry, closure and purity, concealing the complex ancestral relationships that it allows under the category "African American." This assumes, of course, that "ancestral relationships" are somehow less complex within categories such as "West African."

by "black," "US African American," and regions of origination.[41] Proxies in practice are also present, with the people of Barbados serving as surrogates for African Americans. Hypo-descent plays a strong role in allowing these moves.

The picture of the genomic use of hypo-descent becomes more complicated in other areas. Jacquelyn Stevens (2003), a social theorist, accuses Risch, et al of employing the "scientifically indefensible notion of a 'one drop' rule indicating blackness," on the basis of this passage:

> Gene flow from non-Caucasians into the US Caucasian population has been modest. On the other hand, gene flow from Caucasians into African Americans has been greater; several studies have estimated the proportion of Caucasian admixture in African Americans to be approximately 17%, ranging regionally from about 12% to 23%. Thus, despite the admixture, African Americans remain a largely African group, reflecting primarily their African origins from a genetic perspective (Risch, 2002).

African Americans and Caucasians are identified here as two bounded populations, with directional gene flow between them. Stevens' contention is that this construction of gene flow is a direct example of hypo-descent, because

> If you have parental genes from Group A and parental genes from Group B contributing to a child, it is ridiculous to state that genes from one group have gone into genes of another group, since of course the child has genes

[41] This report provides an interesting exception to the use of the term "minority" to reference racial inequality. Topping the list of performance measures for the project is the criterion "recruitment of *minority* families to participate in the Barbados Prostate and Breast Cancer Study." This use of "minority" occurs immediately below the section quoted, in which Barbados provides an "ideal population" because its people are "*relatively homogenous, predominantly black*"—ideal because black people are *not a minority at all.* And so we can say that "minority" here serves as a proxy for "black."

> from both (and this is assuming that these are originally "pure" groups in the first place, which is also incorrect) (2003).

This analysis is more a polemic against the one drop rule, than it is an understanding of its operation in Risch, or its historical operation. Historically, the one drop rule has made it possible for genes in one "group" to go into genes from another "group," and not the other way around, because that is the way that these "groups" were defined. Enslaved "Negro" women were often impregnated by "white" slaveowners, and their children became part of the "Negro" "group," despite equal genetic contributions from each parent. What Stevens' argument usefully implies, however, is that this way of building groups is at least scientifically unusual, and can only be explained by the involvement of the one drop rule.

That involvement is not as simple as the naked deployment of the "old canard" Stevens identifies. On the basis of a study by Parra, et al (1998), Risch claims that "despite the admixture, African Americans remain a largely African group, reflecting primarily their African origins from a genetic perspective." This is a statement of emphasis, in which genetic evidence is mobilized to support a prioritization of what the group *largely* and *primarily* is. This emphasis is Risch's addition—it is not the point of Parra's paper. Parra tested both gene-frequency indicators and population-specific alleles for African and European descent, in order to gauge the level of Caucasian admixture in African Americans, among other goals. The central point of the paper was an emphasis on complexity in regional admixture proportions among African Americans—an ascription of difference, not the unity Risch ascribes to a range as broad as 12% to 23% Caucasian admixture.

But even in the presence of this complexity, Parra's paper itself does the work of homogenization in several ways. First, it takes "African Americans" as a self-evident

category. There is no definition for this category, and there is no description of the way in which sampled individuals' race was determined. Second, these data are averages from North American population samples by city, ranging from 47 to 175 individuals/population. Data for internal difference within these populations are not presented. This group focus masks a historical diversity of "descent" in who is categorized as African American. For example in one famous case, a woman with 1/32 African American ancestry was denied permission by a court to have her legal race categorization changed to Caucasian (Davis, 1991). There is no basis in the presentation of Parra's data on which to know what the range among individuals raced as African American might be for the genetic admixture markers the study is based on. Averaging here is a means of homogenization, even when Parra's emphasis is complexity across these geographic foci. The one drop rule—or at least an analogous epistemological process— is manifested here in the sense that it has defined who is African American, it has homogenized difference within the category, it has maintained boundaries despite "mixed" "descent," and it has given the sense of categorical stability. Similar processes are at play in Risch's piece, although employed for different purposes.

The one drop rule is not an adequate picture of this epistemology. It is not only that homogenized categories and directional gene flow are presented, but that admixture, even as the central focus of a study, can coexist with such entities. Admixture, as discussed in the second chapter, is in essence a genetic term for miscegenation. Intermediate categories that are the product of admixture are present in almost the same breath as discourses of purification that represent, or at are at least influenced by hypo-descent. Even within Risch piece, with its strong construction of racial architecture intermediates are present and even emphasized:

> Populations that exist at the boundaries of these continental divisions are sometimes the most difficult to categorize simply. For example, east African groups, such as Ethiopians and Somalis, have great genetic resemblance to Caucasians and are clearly intermediate between sub-Saharan Africans and Caucasians. The existence of such intermediate groups should not, however, overshadow the fact that the greatest genetic structure that exists in the human population occurs at the racial level." (Risch, 2002)

The genetic resemblance of Somalis and Ethiopians to Caucasians may occupy a subordinate position in this racialist account, but the one-drop rule's purification is not at play. In fact, although this passage directly addresses race and Africa, it has little if anything to do with phenotypic blackness, given the skin color of the majority of people living in the nations Somalia and Ethiopia. There is another construction of blackness at work in Risch. While explicitly dissociating himself from a "merely cosmetic" account of race, he argues for the legitimacy of self-reporting practices for "white" and "black" people in relation to genomic research.

> Racial categorizations have never been based on skin pigment, but continent of origin. For example, none of the population genetic studies cited above, including the study of Wilson *et al*, used skin pigment of the study subjects, or genetic loci related to skin pigment, as predictive variables. Yet the various racial groups were easily distinguishable on the basis of even a modest number of random genetic markers.(*Ibid)*

This passage relates specifically to the categories "white" and "black," and Risch is arguing for their legitimacy, and against their being construed as having *anything* to do with skin color. He emphasizes that black people have a large range of skin colors as evidence. Doing so, he acknowledges and rationalizes an effect of the one-drop rule that he employs in one part and denies in another. For Risch, and for genomics, hypo-descent is only part of the story.

Clearly, admixture does not necessarily lead to the delegitimization of race in genomic discourse. There must be two or more racial or ethnic populations to do the admixing. There has even been a recreational admixture test based on SNP profiling. Marketed by DNAPrint Genomics, Inc. for $290, the test gives people the proportions of their ancestry in terms of "five geographical areas [that] correspond to the major human population groups or races, those of 'Native American, East Asian, South Asian, European, sub-Saharan African, etc.' according to the company's Web site."(New York Times: Wade, 2002)

Admixture in genomics does not carry with it the sense of crisis or scandal associated with miscegenation in racist discourse. Parra goes to great lengths to emphasize his anti-racism, including extensive histories of slavery in his article for the American Journal of Human Genetics, an unusual platform for such histories (1998). Risch justifies his racializations in terms of reducing health disparities. Collins, in his short course lecture uses his image of the "trellis" to portray admixture with respect to human evolutionary history, using "population" as a proxy for race.

> When people draw the history of human populations as a tree - you've probably seen such tree structures - that may be a useful shorthand, but it really isn't right. Because gene-flow doesn't occur down one branch and then never back again, and does some other branch - the way the gene-flow occurs is backwards and forwards. Our history as a human species is really more of a trellis, than it is of a tree. (Collins, 2001)

This passage represents an almost celebratory attitude towards boundary fluidity, not explicable in terms of an epistemology of hypo-descent. Although the Rule involves itself in complex ways in the production of blackness in genomics, it coexists with ideas of admixture that would seem to otherwise oppose it, producing an instability that allows other instabilities. That epistemologies of blackness, ancestry and miscegenation are not

resolving into historical systems of thought indicates that something new is being produced. That difference and unity are being constructed in ways that are both old and new, strengthens the case that race is being made.

III. White Normativity

White normativity is the casting of bodies and populations that are raced as white, and especially those gendered male, as normal. Theories of white privilege often center this way of thinking as a method by which privilege maintains its hegemony. That white men are positioned as the unmarked, default category confers certain effects of social power.

White normativity in medicine casts white male bodies as the most normal in terms of health—a standard from which other bodies diverge. This does not necessarily mean that white men are positioned as the healthiest group. Deviations may occur in either direction.[42]

Some aspects of medical white normativity surface in genomic discourse, most commonly in the use of implied baselines in racial comparisons. A baseline in epidemiology is the average health status of a group of people, usually in the absence of a particular putative risk factor. Establishing a baseline allows one to use a group of people as a control in health research. Variable groups can then be compared with the baseline. For example, a group of non-smokers could be surveyed for lung cancer incidence. The frequency of incidence could then be used as a baseline for comparison with smokers.

[42] Although support for this claim will require further research, it strikes me that bodies and populations raced as Asian are often positioned as abnormally *healthy*. This would seem to be consistent with traditions of Orientalist discourse.

Confounding variables, which typically include geography, occupation, sex, age, diet, amount of exercise, etc., could be eliminated either statistically or by sampling adjustments. The deviation from the baseline of cancer incidence could then be used as evidence that smoking causes cancer. Elevated incidence from the baseline is evidence that the risk factor is associated with disease etiology.

That baselines are "implied" in the genomic discourse of race, means that racially and ethnically defined groups are often asserted to have disease outcomes that are "elevated" or "more prevalent," without specifying a baseline from which they diverge. I think that baselines are furnished in these cases, in a combination of reading and writing practice, by a system of white normativity.

Ethnically defined groups, in particular the Ashkenazim and Pima Indians are commonly identified as having elevated disease incidence. The Ashkenazim have a higher incidence of breast cancer and Tay-Sachs Disease. The Pimas have elevated diabetes rates. Racially defined groups are also identified in this way, such as African Americans with hypertension and sickle cell anemia. In cases where no baseline is made explicit, medical white normativity plays a role. If, as in the Brower passage "many blacks are more salt-sensitive and more likely to have low plasma renin levels, [and] they are less sensitive to ACE (angiotensin-converting-enzyme) inhibitors and ARBs (angiotensin II receptor blockers)," a normal baseline must enter in. The normal category from which comparisons are drawn is the unmarked category of whites.

But white normativity is not adequate as a description of genomic epistemology. Two papers, both describing racial differences in response to cardiac therapy for blacks,

in the May 2001 edition of the New England Journal of Medicine establish baselines explicitly[43]:

> Heart failure is a substantial public health problem among black Americans. It is more common in the black population than in other populations in the United States, affecting approximately 3 percent of all black adults in this country. Symptoms of heart failure develop in blacks at an earlier age than they do in nonblacks, possibly because blacks are more likely to have hypertension and diabetes and to have ventricular hypertrophy and vascular injury than nonblacks." (Yancy, et al, 2001)
>
> Black patients with heart failure have a poorer prognosis than white patients, a difference that has not been adequately explained. Whether racial differences in the response to drug treatment contribute to differences in outcome is unclear. (Exner, et al., 2001)

The baseline for Yancy, et al. is "nonblacks," for Exner et al. it is "whites." In making the terms of comparison explicit, even while no definitions for these terms are offered, moves this piece away from the unmarked category implicit in silent baselines, and away from white normativity.[44] Whites and nonblacks may be more healthy (and therefore more normal) in this assertion, but the argument does not rest on a concept of a priori normativity.

Adding to the complications, there is at least one disease, hemochromatosis, undergoing genomic analysis that is largely specific to whites. In a consensus statement in JAMA, coauthored by Francis Collins, this is described:

---

[43] These two papers are not in fact genomics literature, but racially structured drug response research in medicine. They are, however, frequently cited in race-based genomics as evidence of genetic difference among races. (i.e., Risch, 2002; Wood, 2001) Furthermore, pharmacogenomics, which is focused on differences in drug response, quite often structures its research racially. These papers can be viewed in this light as precursors to pharmacogenomics.

[44] In response to its publication of these two articles, NEJM began requiring definitions of, and justifications for using race in research that the journal publishes. (?????)

> Data from Europe and North American suggest that hemochromatosis is more common in whites than in members of other races, and may be most common in people of Celtic origin. A recent study suggests that the prevalence of hemochromatosis in people of Hispanic descent may be similar to that in whites. Hemochromatosis is reported to be less common in African Americans than in European Americans, although a distinct non-HLA-linked primary iron overload syndrome has been described in African Americans and in sub-Saharan Africans. The usefulness of TS [transferrin saturation] as a screening tool and appropriate threshold levels are likely to vary for people of different racial or ethnic groups. (Burke, et al, 1998)

In addition the use of proxies in this paragraph—for example "Europe and North America," "whites," "races," and "Celtic origins" are linked in an interesting fashion—it is noteworthy that this undermines a notion of the consistent use of white normativity. Nothing could be further from engaging in such an activity than identifying a disease as primarily white. The category is marked and attributed with medical characteristics. The picture, then, is a complex one in which white normativity plays a role, but is inadequate as a characterization. This should furthers the case for a novel production of race.

IV. Racial Essentialism

Essentialism is defined by the Oxford English Dictionary as 'the belief in real essences of things, esp. the view that the task of science and philosophy is to discover these and express them in definitions." Essences have a number of definitions. The most relevant to the question of "racial essentialism" include " specific being, manner of existing, 'what a thing is'; nature, character," "'substance' in the metaphysical sense; the reality underlying phenomena; absolute being," and "the most important indispensable quality or constituent element of anything; the specific difference." Racial essentialism, then, is the

belief that race is an underlying and fundamental reality, a difference that is constitutive and specific. Howard Winant uses the idea of racial essentialism to construct a theory of contemporary racism:

> Today, a racial project can be defined as racist if it creates or reproduces hierarchical social structures based on essentialized racial categories. This approach recognizes the importance of locating racism within a fluid and contested history of racially based social structures and discourses. It allows us to recognize that there can be no timeless and absolute standard for what constitutes racism, because social structures undergo reform (and reaction) and discourses are always subject to rearticulation. This definition therefore does not invest the concept of racism with any permanent content, but instead sees racism as a property of certain political projects that link the representation and organization of race—that engage in the "work" of racial formation. Such an approach focuses on the "work" essentialism does for domination, and the "need" domination displays to essentialize the subordinated. […] [P]rojects may be considered racist […] only if they meet the criteria I have just outlined: in other words, essentialization and subordination (which are always linked) must be present. (Winant, 1998, emphasis in original)

Racial essentialism, then, is a crucial component of racism. A racial project must essentialize a group, and couple this essentialization to the work of subordination in order to be considered racist within this framework.

Using Winant's theory, three linked questions about the genomic epistemology of race can be asked. 1) Are these racial projects *essentialist*?, 2) Do these projects participate in processes of *subordination*?, and 3) Are these racial projects *racist*? In answering the first question, the difficulty that we face, is negotiating the difference between "essentialist" and "biological." Are they synonymous? Is a claim about racialized hereditary drug response differences or allele frequency distributions inherently essentialist? Are *genes* essences?

As this thesis has shown, genomic discourse accommodates a spectrum of claims about the legitimacy of race, from fundamental genetic architecture, to useful concept, to heuristic, to misconception. Perhaps it is useful to say that one side of the spectrum leans towards essentialism, the other towards a notion of race as discursive or mere illusion.[45] But to do so would ignore the instability of meaning within individual papers, and at different points by the same authors. Nested proxies and perspectival phasing allow essentialism and constructivism to coexist. "African-Americans respond better to a new drug, eplerenone" is made to agree with "'We're not talking about race as a genetic situation. We're talking about it as a *marker."*(Brower, 2002) Recall the passage from Francis Collins in which

> [I]t is improper to draw precise boundaries around any particular ethnic or racial group and say they're biologically different. Science won't support that. But at the same time, it is correct to say there may be differences in frequencies, of a quantitative sort, of particular disease-susceptibility genes in one group compared to the other, that may play a role in why there are health disparities. Both those statements are true, and yet it is I think fairly challenging to get that message across in a fashion that is convincing to people, and it doesn't seem paradoxical. (Collins, 2001)

The statement about frequencies could be construed as essentialist. The true specific differences among races, which influence American health disparities, are asserted to lie in a biological, essential notion of *frequency*. But, then again, distribution differences be construed as "essences?" The disease susceptibility genes are not asserted to belong

[45] Winant critiques both ends of an analogous spectrum among social thinkers: "But obviously problematizing race is not enough. We must steer between the Scylla of thinking that race is a mere illusion, mere ideology (in the sense of false consciousness) on the one hand, and the Charybdis of thinking that race is something objective and fixed. Both of these positions have their temptations, and by no means only for those who would deny the significance of race. The former position ("race as an illusion") is upheld today not only by neoconservatives but also by radical theorists of race such as Anthony Appiah and Barbara Fields. In the work of these scholars, whatever its other merits, there is little recognition of the autonomy and depth of racialization in the U.S. The latter view ("race as objective") is accepted not only by biological determinists and scientific racists, but also by many social scientists (some of them quite progressive)…"

exclusively to certain groups, representing the *foundation* of their difference. Rather genes that every racial group has are asserted to be more common in some than in others. This leaves the notion that there are racial groups, supported again by the evidence of frequency.[46] The result for genomic epistemology is an unstable, circular sort of reasoning, a kind of feedback loop between essentialism and its opposite. The result is a type of racialization that depends, not on essentialism (or its problematic "mere illusion" opposite), but on a process of shifts between these poles, a dynamic, unstable sort of phasing. There are elements of essentialism to be found, but if they are viewed in static isolation, the more important epistemological process of forming racial difference in flux among a large number of changing positions is missed.

What can be said, then, about projects of subordination and domination, the other component of Winant's schema for racism? I have seen no evidence of efforts at subordination projects in genomics, whether conscious or otherwise. In fact, remedying racial health disparities in the U.S. is one of the central arguments justifying racial genomics research.

We should ask, however, to what extent genomic knowledge of race—whether "essentialist" or otherwise—can feed into the regimes of subordination that produce

[46] In his article on NHGRI's website "Ethical Issues in Developing a Haplotype Map with Socially Defined Populations," Morris Foster worries explicitly about producing a genetic essentialism for racial categories. As with so much of this material, the risk lies with the public, not with genomics: "Using broad, socially defined populations to structure participant recruitment for a haplotype map project, and retaining racial and ethnic identifiers, may lead non-scientists who become aware of the project to reify those social categories as biological constructs, fostering an unintended genetic essentialism in the way the public understands such categories as race and ethnicity. That essentialism could obscure the important fact that the "boundaries" between groups are highly fluid and that most genetic variation exists within all groups - not between them. Even the smallest socially defined population will have multiple haplotypes, and haplotypes will be shared among different, socially defined groups."(Foster, 2001)

racial health disparities. This is not intended to recapitulate the distinction between "the public" and "science." Regimes of subordination do not respect this type of distinction, although it may be possible to identify different ways in which such regimes are produced and participated in by scientists and non-scientists. While racial production is a process that involves "essentialist," "constructivist," and other sorts of positions, the products of this knowledge may contribute in unexpected ways to such regimes, however anti-racist their conception.

Is genomics racist? By Winant's criteria, no. There are no *stable* processes of essentialization, linked to no *current* regimes of subordination. But we can identify pieces of essentialism engaged in shifting processes, and we can identify potentials for uptake into regimes of subordination. Perhaps this epistemological position should be described as racism waiting to happen.[47]

So, in the same way that the one drop rule and white normativity are inadequate characterizations of genomic epistemology, racial essentialism falls short of the mark. This should be interpreted to support the central point, that genomics is not only *using* race (in its manifestation as essential), but producing it.

V. Three spaces

The presentation of evidence for race and racial attributions draws in part from an old evidentiary development in population genetics. Reardon describes a process taking place among population geneticists beginning in the 1950s in which an inner space was

[47] There is a significant body of ELSI literature concerned with this potential. Almost invariably, it locates racism as a product of *misinterpretation* by social and political forces outside of science. While it is heartening that the study of E., L., and S. Implications is supported by genomics, and that their topic includes race, this approach is inadequate.

carved out for the adjudication of race, in the form of genes and gene frequency (2004). Race became legible for mainstream biology, and human population geneticists in particular, as a matter of alleles, with reference to their distributions in populations. A new evidential space was developed in this way.

Lay perception, phenotype, culture and appearance could all be shaken loose from this space. So too could lay people, politics and governments. In light of the crisis in conceptions of biological race induced by the Holocaust, these disaggregations served as useful boundary projects. Scientific race could be walled off from the misuse of nationalist racism, might could be separated from right. But race remained, not in *a* concept, but in a set of complex formations and knowledge technologies, centered, for science, in a new genetic interior.

"The human genome" can be read as a recent version of this inner evidential space. In Risch et al's article, for example, race takes place not in the "merely cosmetic" surfaces of phenotype, of skin pigment, the realm of appearance (2002). Race takes place in genomic structures determined by "continent," such that races are "easily distinguishable on the basis of even a modest number of random genetic markers; furthermore, categorization is extremely resistant to variation according to the type of markers used (for example, RFLPs, microsatellites or SNPs)" (Ibid.). In the rarified, coded space of the genome, categorization ultimately is positioned as resistant, pure and accessible to the initiated.

In this sort of presentation of evidence, genetic science has been called reductionist. "Reductionism" is an accusative term. It means that the complications inherent in a topic are reduced to something simplistic, something *mere*. "Reductionism"

is a dangerous slope to slide down, a process that blinds the reducer to what the accuser sees as important. No one wants to be reductive.

In the case of genetics and genomics, reductionism has a relationship with determinism, although the two are not the same. Genetic determinism, also an accusative, relates mainly to ideas of behavior and culture. It is the belief that culture and behavior are determined by genes, that our fate is "written," and may be read in our DNA. Reductionism, in part, is the process by which we arrive at such a simplistic, determinist conception. A determinist reduces the complexities of culture and behavior to genes.

Reductionism as an accusation can extend beyond a fated cultural determinism. With respect to disease etiology, a genetic reductionist would reduce the causes of disease to genes, obscuring a variety of important epidemiological factors (Tesh, 1990). The contributions of environment, occupation, class, racism, diet, lifestyle, psychology, among others, would be delegitimized in favor genetic explanations. Reductionism might erase these entirely, or acknowledge them but confer a special privilege on the genetic.

Leveling the accusation of "reductionism"—especially in its determinist sense—at genomic discourses of race gets us no closer to an understanding of how race is produced epistemologically and discursively. Unsurprisingly, there is a level of privilege accorded to DNA-based explanations; these are *genomicists* after all. It is useful to ask on what sort of assumptions a genetic presentation of evidence depends, what it can mask, and what social and medical stakes are involved (Tesh, 1990). It is useful as well to examine the ways in which this privilege is produced, and to be alert to what is at stake medically and socially.

I have seen no instances of genomic discourse attributing cultural and behavioral characteristics, via genetic explanation, to racialized groups. This is not the sociobiology seen in *The Bell Curve*, in which genetically based intelligence differentials are asserted to characterize different races (Herrnstein & Murray, 1994). Claims about racialized intelligence, culture,[48] and behavior are foreign to mainstream genomic discourse. Genetic determinism is not the norm for these types of attributes, and genetic reductionism in this sense does not apply.

Genetics are privileged—although not exclusive—in discourses of disease etiology, including racialized attributions that center on differing genetic frequency in "populations." Genomic evidence for race is presented most often in the context of disease, medical disparities and differential risk.

But the adjudication of race takes place for genomics not only in medical terms, or in the inner terrain of the genome, but in a space that points to its own discursivity and sociality. Social constructions and heuristic starting points occupy the same papers as single nucleotide polymorphisms and haplotypes. And Risch, if not an extremist, does at least occupy an extreme end of the mainstream in betraying very little discursivity. Social and discursive evidence has been demonstrated in both formal and informal situations. Claims about words and interactions can occupy everything from technical papers to congressional testimony. The instability embodied in movements in and out of sociality and discursivity represented by perspectival phasing, is complemented by a variety of positionings within this evidential space.

[48] "Culture" here means genetically based cultural attributes and qualities. There are many claims about *culture*, but in terms of race they are separated from genetic causes. For example, ELSI discourse around "communities" frequently refers to the need for cultural appropriateness when achieving consent from constituencies to be sampled.

In addition, this thesis has said very little about the silent partner of bioinformatic technologies. But "cyberspace," to use the cliché, permeates genomics. Haraway describes this situation:

> If universal humanity was plastic under the sign of the population at midcentury, then human nature is best described as virtual in present, end-of-the-millennium regimes of biological knowledge and power. Specifically, human nature is embodied, literally, in an odd thing called a genetic database, held in a few international locations such as the three large public databases for genetic map and sequence data: the U.S. Genbank©, the European Molecular Biological Laboratory, and the DNA Data Bank of Japan. […] GenInfo, developed by the U.S. National Center for Biotechnology Information of the National Library of Medicine, is a kind of metadatabase containing both protein and nucleic acid sequence data "to which other databases can add, refer, annotate, interpret, and extrapolate."(Haraway, 1997)

Informatics are more complex than repositories or research tools. The idea of a "metadatabase," to which other databases can "add, refer, annotate, interpret, and extrapolate" should point us towards the idea that informatics is an actor in genomic epistemological processes. It is not only that the genome exists electronically. Informatic services are available which in produce and mediate evidence. For example, as of this writing, I can perform searches through the NIH's National Center for Biotechnology Information's (NCBI) website, which can compare and contrast nucleotide and amino acid sequence in all reading frames and a variety of combinations. I can do this at the moment for 246 bacterial, 18 archaeal, and 44 eukaryotic genomes. (http://www.ncbi.nih.gov/entrez/query.fcgi) The information I produce in this process can be mobilized, used as evidence, and structure further research.

The genome, the social and discursive and the informatic are the terrains on which genomic knowledge about race is being produced and contested. I think perhaps the spaces of "the genome" and "the social" can account to some extent for the phenomenon

of perspectival phasing. Biochemists, after all, are new at this whole "social construction" business. Phasing can be read as a negotiation of tension between two relatively new spaces, "the genome" and "the social"[49] from the perspective of genomicists, which are in the process of sorting themselves out. Race itself can be viewed as a transfer point between those spaces, an object of those negotiations. In this sense, race occurs as an accretion of evidential shifts, a sediment produced at the intersection of different epistemologies.

[49] Quotation marks denote that these are spaces marked by genomicists—I do not refer to a negotiation between what is "actually genomic" and what is "actually social."

Conclusion

Despite the initial promise, and widely held belief that the Human Genome Project draft completion of 2000 would bring an end to the legitimacy of biological race, this thesis has shown that genomics is a powerful site for the production of racial and ethnic differentiations. Genomics is making, rather than debunking race. The stakes of this production are high, and its implications far-reaching. Genomics embodies a powerful social and scientific authority, relates in complex ways to the U.S. government and corporate interests, consumes massive amounts of funding, and resonates with humanist and genetic discourses. All of these structural and discursive positions of strength, along with the sheer breadth and volume of genomic knowledge, indicate that these productions of race will ramify.

Most immediately, the ramifications of genomic race will display themselves in biology as a field. By all accounts, a "genomics era" has been inaugurated for biomedicine. Both the public HGP and Celera explicitly set out to influence and inform further research. The genome was positioned as a tool of extraordinary utility for biological knowledge production. It seems likely that, in addition to informatic accessibility and volume, the authority given to genomic knowledge within biology will extend past simple sequence to encompass determinations about race as well. The

reaction to this authority may take the form of contestation, but the terms of discussion and the evidential battlefield, will in all likelihood be genomic.[50] Experimental practices, discursive terms and configurations, epistemological formations, reading practices—the forms and content of genomic knowledge—all have the immediate potential to be taken up and employed within the field of biomedical research.

Such cross-fertilization is already in effect. Genomic microarray technology is beginning to be used for a wide range of applications, for example. Microarrays allow measurements of the expression—that is, whether genes are producing messenger RNA— of many thousands of genes rapidly and simultaneously. It is not only that the technology itself has been employed, but that the explanatory structures it facilitates and embodies, in which knowledge about molecular biology is located, valued and produced within expression analyses, travel with it. Genomic discourse, wrapped up with informatic and biomolecular technologies, has circulated into an expansive diaspora of biomedical practice.

Toxicogenomics and pharamcogenomics are hybrid fields that enact such circulation. Toxicogenomics, whose central experiments involve microarrays, is a hybrid between human genomics and environmental toxicology. Its focus is the genomic response to chemical exposure, and the ways in which this response mediates "individual susceptibility" to disease. The field has its own nationally funded project, the Environmental Genome Project, orchestrated by the National Institute for Environmental Health Science (NIEHS). Pharmacogenomics, also a field that employs microarrays, combines human genomics with pharmaceutical research, aims to produce individually tailored drugs, and commonly uses race as an independent variable. Genomically based

[50] "Genomic" here refers to the field of knowledge, and all of the ways in which its discourse presents evidence for race, not only the genome.

and informed clinical medicine is emergent. These hybrids between genomics and toxicology, environmental studies, pharmacology and clinical medicine illustrate the ramifications of genomic knowledge and discourse within biomedicine.

The stakes of racial production within such hybridized projects in the genomic era revolve not only around new forms of race in biomedicine, but new racializations of risk and health. Medicines, environmental hazards, disease susceptibilities, constructions of biological normalcy and impairment and many other productions of biomedicine are already being affected by genomic knowledge. There is no reason to think that new productions of race will not travel with them. BiDil, the "ethnic drug" by Nitromed is an example. Bidil is a cardiac medicine for black people. Its existence rests on a racially differentiated construction of risk, disease and biology. Within biological fields, new race can be linked to a variety of new knowledge. There is no reason to think that such a process of linkage will naturally happen on equal terms, with equal benefits to those raced. Conscious racism is not required for differential accruals of benefit—a variety of institutional, economic and accidental factors can converge to make new racializations manifest themselves as inequalities among the racialized. If BiDil is for black people, where are the medicines for Asians? That is, if a particular racialization is accepted, and put into medical practice, new ethnospecific treatments are as much a question of withholding as administering. If a drug is only good for one new "race," it must not be given to the other "races." If one "race" is particularly genomically susceptible to disease in response to a particular environmental contaminant, and can organize against its introduction into their "community," another race may be disempowered in the same action. Biomedical race, even when constructed by means of statistical allelic frequencies, has no way of guaranteeing justice. And if the genetic attributes of new

racial categories are naturalized, organization against unjust outcomes will have no enemy but nature to organize against. Although racialized people may organize for more biomedical research to be directed to their concerns, the roots of racial problems will lie with no oppressors. Paul Rabinow describes this condition— independent of race— as "biosociality," in which the "new genetics will cease to be a biological metaphor for modern society and will become instead a circulation network of identity terms and restriction loci" (Rabinow, 1998). In such a context,

> [I]t is not hard to imagine groups formed around the chromosome 17, locus 16,256, site 654,376 allele variant with a guanine substitution. Such groups will have medical specialists, laboratories, narratives, traditions, and a heavy panoply of pastoral keepers to help them experience, share, intervene, and "understand" their fate. Fate it will be. It will carry with it no depth (Ibid.)

The possibility of such a biosociality carries risks enough, without considering the potential for it to become biomedically racialized. The circulations of power within such a framework involve life, death, health and the production of meaning around all of these. Race, of course, has always involved life, death, health, and lived meaning as well. It is not necessary to invoke narratives of genetic dystopia like *Brave New World* and *GATTACA*, to say that the potential combination of race and biosociality may enable sweeping, dangerous transformations.

Such transformations are not limited to biological science. Although genomic technologies like sequencing machines and microarrays may not extend much beyond highly funded laboratories, genomic discourse and knowledge can travel beyond such small spaces in surprising ways. That they should do so has been both an explicit and implicit hope of genomic authors. Unlike so much specialized knowledge in recent science, genomics speaks often to the universal, global human, albeit through a

specialized lens. In *Power/Knowledge*, Foucault spoke of a new "specific intellectual," replacing the "universal intellectual" as the contemporary social position of the scientist (Foucault, 1977). The specific intellectual was no longer the "rhapsodist of the eternal," but the "absolute savant" and "strategist of life and death," whose social power was located in concentrated, expert knowledges rather than a position as "consciousness/ conscience of us all." It strikes me that the prominent genomicists, and Francis Collins in particular, do not fit well into this schema. Collins' frequent references to God, the Book of Life, the universal human, and liberal humanist politics make him a "rhapsodist" in a way unbecoming to an "absolute savant," while the deep specificity of genomic practice, and the sequence "code" that is the field's central product are hardly the "great writing" characteristic of pre-war universal intellectuals. Whether the social positions of genomic scientists in general will develop into a new sort of public intellectual remains to be seen, but the position of Collins as scientist, governmental official and humanist should indicate that new spaces for social authority may become available to genomics.

Because of the broadened scope of genomics, both in terms of the social position of its practicioners, and the wide political valence of their discourse, the stakes of the genomic production of race extend beyond the biomedical, and beyond Rabinow's biosociality. The Ethical, Legal and Social Implications program is an explicit acknowledgement of this power. As this thesis shows, the ELSI question of ramifications in racial thinking usually rest on a dualism between the public and science, in which the public misinterprets genomic knowledge. In this context, "reification" of racial categories by the public is a common way to frame the issue. Reification means that (largely) socially constructed categories will come to be viewed as biological or "real" by the public in response to genomic investigations that employ these categories. Although

this view is strongly limited by its dualism, it points to the potential for genomically authorized racializations to ramify in spheres that are not explicitly "biological" or "scientific." ELSI discourse often emphasizes the potential for misinterpreted or reified categorizations to stoke racism, feeding into supremacist and essentialist ideologies. This sort of racist uptake is a possibility and a concern.

But the larger danger in the non-biomedical ramifications of race is subtler than new biological buttresses for racism. Race itself is a locus of violence. "Race is a fracturing trauma in the body politic of the nation—and in the mortal bodies of its people. Race kills, liberally and unequally; and race privileges, unspeakably and abundantly" (Haraway, 1997). This is not the whole story— race has been appropriated effectively for political struggle in many instances, by organizations like the Black Panther Party. Race engages in complex ways with lived realities, producing identities and subjectivities that have had positive effects. Studies of the biological and medical effects of racism, such as those carried out by the environmental justice movement, have used ideas of race in order to advocate for more just social outcomes. But history shows that in its greatest bulk, the various Western racial systems have operated to justify violence, exploitation, heightened privilege and colonization for the dominating race. This legacy should make us wary of any new racial productions. Genomically produced racial architectures and attributions may affect social relations in unexpected ways.

If naturalized constructions of race gain greater currency in genomic science, the claim that race is nonbiological will no longer be available to race theorists. Such a removal of scientific authority would pose difficult issues to negotiate for anti-racist theory and practice. The argument that race is postulated on a scientific fiction has been powerfully employed in anti-racist work. This rhetorical investment in scientific

authority may backfire if scientific authority shifts towards a biologicization of race. Even genomic knowledge that presents race as social or discursive may prove harmful or difficult for race theorists in its concepts of "misinterpretation," and its analysis of "disparity." The issue may be framed in ways that are ineffective or detrimental to the priorities of race studies and anti-racist activism. The production of genomic race is a powerful and unpredictable force being introduced into an already complex political situation.

In light of these high stakes and unpredictable consequences within and around biomedicine, genomics cannot be left to its own devices, particularly not when there are only 4,000 genomicists nationally. Some scholars in biomedicine, science studies and social theory have begun to respond, and this thesis is intended to contribute to this growing field. But it is not enough to say that genomic science produces race. It is not enough to deconstruct the discursive processes by which it does so, to take stabs at its epistemology, to say, "here is instability, revealed at the heart of your so-called 'science.'" It is not enough to engage in boundary work, to divide science from "science," to purge the pseudos- and the charlatans, to debunk the bunk. The price in simplification is too high, the lost allegiances too valuable.

Biology is not the enemy. It is not a question of delegitimizing the field, or advocating for its political neutralization. Neither of these approaches is effective, and if pursued, both would fail. Oppositional politics are not enough. The task is to advocate for a full realization of and accountability for the social and political generativity and authority of genomics. The activity of the field in creating race and negotiating its social position must be studied in such a way that genomics can take responsibility for its productions. In a field with such potential to effect social and scientific transformations,

we must have theories of how these transformations operate, and we must have means of conscientious intervention that extend beyond institutionally financed ELSI programs. We, as biologists, race theorists and "science students," must understand race in science and society in more complex ways than "misinterpretation," "reification," and the need to educate "the public."

In his essay on biosociality, Paul Rabinow describes knowledge about the genome as geared from its origin towards intervention:

> The object to be known—the human genome—will be known in such a way that it can be changed. This dimension is thoroughly modern, one could even say that it instantiates the definition of modern rationality. Representing and intervening, knowledge and power, understanding and reform, are built in, from the start, as simultaneous goals and means. (Rabinow, 1998)

To appropriate this characterization in a way that it was not intended, knowledge about *genomics*, produced both by its practicioners and its critics, must be known in such a way that it can be changed. At the heart of the theories we produce together about the operation of scientific knowledge must be a passionate commitment to transformation. Such a hope does not need scientistic utopianism or chauvinism, it does not need an anti-technoscientific attitude often referred to as contemporary Luddism,[51] it does not need a rejection of genomics based in discourses of "nature" and the "natural," and it cannot be accomplished through reformism[52] or critique alone.

---

[51] This common usage misunderstands the purposes of the Luddite movement, which sabotaged industrial machinery not out of a reactionary technophobia, but as a means of social protest against capitalist domination during the early British Industrial Revolution.

[52] Jacquelyn Stevens has suggested a policy change at NIH, for both its agencies and grantees, in which the Institutes withhold funding from those that publish "claims about genetics associated with variables of race, ethnicity, nationality, or any other category of population that is observed or imagined as heritable unless statistically significant disparities between groups exist and description of these will yield clear benefits for

One approach to building knowledge about the processes and ramifications of genomic race production— and building it in such a way that it can be changed—is the organization of educational and activist communities around these issues. That there is a national need to "build communities" has become a platitude in contemporary liberal discourse. It is intended here to mean a more radical and vital restructuring process than the cliché would indicate. It means not only a temporary assembly of people for a purpose, but a means of social organization that establishes itself inside scientific structures, and empowers itself in knowledge and praxis. The goal is not invasion and coup, but the establishment of meaningful, equalized dialogue among people with every sort of stake in genomic race—not a tokenized supplement to genomics, and not a refutation, but the epistemological integration of many knowledges and perspectives, whose insights can be operationalized.

To be sure, this is easier said than done. The institutional and disciplinary establishments in which genomicists are situated come well equipped with walls. The integration of state and corporate power with science, lengthy professional initiation processes, complex bureaucracies, esoteric vocabularies, constructions of expertise and strongly felt disciplinary loyalties make genomics a difficult ground to build on. As evidenced by the Science Wars, impassioned rhetorical policing can be easily provoked in these quarters.

But there are many techniques available from the traditions of grassroots and labor organizing and radical pedagogy that can be applied to such an endeavor. And

---

public health." (Stevens, 2003) While her arguments for the benefits of a "heuristic practice allowing for the disaggregation of race from heredity" through this reform are well-taken, such a reform does not address the myriad ways in which race is produced that involve no independent variables. Furthermore, this reform involves no dialogue with genomicists and medical researchers, and provides restrictions, rather than new, better possibilities for the production of knowledge.

despite the Science Wars, there will be no Pinkertons to shoot visiting activists on the NIH campus in Bethesda. Our object, after all, requires no wrenches in the sequencing machines.

To organize activist educational communities around and within genomic knowledge first requires a willingness to transgress disciplinary boundaries. There is a vital and growing body of academic research that does so under the rubric of science and technology studies. And there are the beginnings of a precedent for such border-crossings in the ELSI program, which funds researchers from many scientific and non-scientific fields. The breadth of topics covered by genomicists themselves indicates a fluidity at the bounds of the discipline that could be capitalized on.

Disciplinary boundaries are not the only lines that must be crossed, however. Many people—possibly *most* people in the United States—have stakes in the genomic production of race. Any attempt to organize activist genomic communities must avoid the potential to replace one small, powerful elite with a slightly expanded powerful elite. "Inclusion" and "diversity" must be produced, but not by their definitions in genomic discourse. People working towards an activist understanding of genomic race must form communities that include those from the academy and outside it, and people who have been and are becoming raced in every sort of fashion, on equal terms. This is not "diversity" as a comfortable euphemism for "people of color," or as a value that takes races as stable categories whose members should be "represented." Nor is it the sort of "inclusion" that is framed within the science/public dichotomy, in which white science does the including and people of color get the including done to them. The process must be one of convergence, initiated and executed by members of all "races," and many occupations, in which diversity of perspectives is integral rather than introduced.

Again, this is easier said than done. Even to begin the process will require a broad and concerted educational effort. Subjects from molecular biology to dismantling white privilege must be tackled. Learning cannot take place in a unidirectional sense, in which members of one field are instructed by another, or members of "the public" are taught by the specialists. Educational communities must *produce* knowledge in a dialogic way, not administer and receive it. Paulo Freire, educational theorist, described such a process in *Pedagogy of the Oppressed* (1970). While the situation at hand is not one of oppressed versus oppressor, the educational praxis he advocates is precisely the learning process in which communities should produce knowledge about genomics:

> A revolutionary leadership must accordingly practice *co-intentional* education. Teachers and students (leadership and people), co-intent on reality, are both Subjects, not only in the task of unveiling that reality, and thereby coming to know it critically, but in the task of re-creating that knowledge. As they attain this knowledge of reality through common reflection and action, they discover themselves as its permanent re-creators. In this way, the presence of the oppressed in the struggle for their liberation will be what it should be: not pseudo-participation, but committed involvement. (Freire, 1970)

"Re-creating" knowledge and "committed involvement" are precisely the goals I mean. While there is no strategic benefit in describing these community organizing projects as having a "revolutionary leadership," and there is currently no one "oppressed" by genomics in need of liberation, the knowledge to be produced "in such a way that it can be changed" must be transformative and may be "revolutionary." The aim, as Haraway says, is building "knowledge potent for constructing worlds less organized by axes of domination" (1988).

But the process need be nothing so grandiose. At its core, organization in response to the genomic production of race can take forms as simple as reading groups

and discussion circles. They have to be reading groups and discussion circles that negotiate difficult, at times intractable, differences, and involve high-ranking intellectuals and officials, but at base they are only groups of people making knowledge together, like any other community.

How could such communities be established and maintained? The question requires another thesis, but several possibilities suggest themselves. Non-governmental organizations, and particularly those associated with the environmental justice movement, are in a good position to pursue such a project. The ELSI program itself is another good starting point, albeit one which must be navigated carefully. The academy can provide a useful launching ground as well. With whatever financial and institutional backing available, the difficulty will be in finding practical ways to organize and involve people across an array of differences.

There is a place for strong critique, for opposition, and for the lonely scholarly analysis of discourse that this thesis represents, inside and outside such social organizing. I hope by this work to contribute to an understanding of the discursive and epistemological means by which race is produced in genomics, so that it may be employed for such critique, opposition, and further analysis. And much further analysis is required. The goal that it must pursue is not only deconstructive. We must undertake the passionate, cooperative, at times enraged pursuit of transformative, livable ways of knowing difference and producing humanity. To begin this process, we need friends, allies and communities. We need new, better unifications.

♦

Bibliography

1. Allen, Garland E. "The Eugenics Record Office at Cold Spring harbor, 1910-1940: An Essay in Institutional History." Science, Race, and Ethnicity: Readings from Isis and Osiris. Ed. Jackson, John P. Chicago: The University of Press, 2002. 301-341.

2. American Association of Physical Anthropologists. "AAPA Statement on Biological Aspects of Race." *American Journal of Physical Anthropology*, 1996, 101: 569-570.

3. Anderson, Matt; Fulchon, Chinita; Moscou, Susan; Neuspiel, Daniel. "Race in Clinical Data." Position statement: http://www.racesci.org. 2/23/03.

4. Anderson, Matthew; Moscou, Susan; Fulchon, Celestine; Neuspiel, Daniel. "The Role of Race in the Clinical Presentation." *Family Medicine*. 2001, 33 (6): 430-434.

5. Angier, Natalie. "Not Just Genes: Moving Beyond Nature vs. Nurture." New York Times. 25 February, 2003. F 1.

6. Barbujani, Guido; Magagni, Arianna; Minch, Eric; Cavalli-Sforza, Luca. "An apportionment of human DNA diversity." *Proc. Natl. Acad. Sci. USA.* 1997, 94: 4516-4519.

7. Branca, Malorye. "Race and Genomics: Marketing Ploy or A Route to Individualized Therapy?" BIOIT World. 30 April, 2003. http://www.pbs.org/race/000_General/000_00-Home.htm. 9/4/03.

8. Brandt, Allan. "Racism and Research: The Case of the Tuskegee Syphilis Experiment." In Susan Reverby (ed.) *Tuskegee's Truths: Rethinking the Tuskegee Syphilis Experiment*. Chapel Hill and London: University of North Carolina Press. 15-38.

9. Braun, Lundy. "Race, Ethnicity and Health: can genetics explain disparities?" *Perspectives in Biology and Medicine.* 2002, 45(2): 159-174.

10. Brower, Vicki. "Is Health only skin-deep? Do advances in genomics mandate racial profiling in medicine?" *European Molecular Biology Organization (EMBO) reports*. 2002, 3(8): 712-714.

11. Burke, Wylie; Thomson, Elizabeth; Khoury, Muin; McDonnel, Sharon; Press, Nancy; Adams, Paul; Barton, James; Beutler, Ernest; Brittenham, Gary; Buchanan, Allen; Clayton, Ellen Wright; Cogswell, Mary; Meslin, Eric; Motulsky, Arno; Powell, Lawrie; Sigal, Elliott; Wilfond, Benjamin; Collins, Francis. "Hereditary Hemochromatosis: Gene Discovery and Its Implications for Population-Based Screening." Consensus Statement in *JAMA*, 1998 280 (2): 172-178.

12. Cha, Ariana Eunjung. "Race Plays Role in New Drug Trials: Treatment by Genetic Origin, Ethnicity Divides Medical Profession." Washington Post. 28 July, 2003. A 1. Available at http://www.racesci.org/in_media/race_drug_trials.htm.

13. Clark, Cheryl. "Single thread, tapestry of life, Anonymous man from Buffalo gives wings to Human Genome Project." The San Diego Union-Tribune. 2 February, 2000: E1.
14. Collins, Francis and Mansoura, Monique. "The Human Genome Project: Revealing the Shared Inheritance of All Humankind.." From the 7th Biennial Symposium on Minorities, the Medically Underserved and Cancer. Washington, DC, February 9-13, 2000. *Supplement to* ***Cancer*** 2001, 91(1) 221-225.

15. Collins, Francis. "Ahead of schedule and under budget: The Genome Project passes its fifth birthday." Review in *Proc. Natl. Acad. Sci. USA*, 1995. 92: 10821-10823.

16. Collins, Francis. "Contemplating the End of the Beginning." *Genome Research*. 2001, 11:641-643.

17. Collins, Francis. "Contemplating the End of the Beginning." Commentary in *Genome Research*, 2001. 11: 641-643.

18. Collins, Francis. "The Human Genome Project and Beyond" transcript from 2001 Genomics Short Course, delivered 8-7-2001, location not specified. Transcribed for U.S. NIH. http://www.genome.gov, 2/23/03

19. Collins, Francis; Green, Eric; Guttmacher, Alan; Guyer, Mark on behalf of U.S. National Human Genome Research Institute. "A vision for the future of genomics research." *Nature*. Advance online publication, 14 April 2003, doi:10.1038/nature01626 and 422: 835-837.

20. Collins, Francis; Guyer, Mark; Chakravarti, Aravinda. "Variations on a Theme: Cataloging Human DNA Sequence Variation." *Science* 1997, 278 (5343)*:* 1580-1581.

21. Collins, Francis; Morgan, Michael; Patrinos, Aristides. "The Human Genome Project: Lessons from Large-Scale Biology." Viewpoint in *Science* 2003, 300: 286-290.

22. Collins, Francis; Patrinos, Ari; Jordan, Elke; Chakravarti, Aravinda; Gesteland, Raymond; Walter, Leroy; DOE and NIH planning groups. "New Goals for the U.S. Human Genome Project: 1998-2003. Review in *Science* 1998, 282: 682-689.

23. Collins-Schramm, Heather; Chima, Bill; Operario, Darwin; Criswell, Lindsey; Seldin, Michael. "Markers informative for ancestry demonstrate consistent

megabase-length linkage disequilibrium in the African American population." *Hum. Genet.* 2003, 113: 211-219.

24. Collins-Schramm, Heather; Phillips, Carolyn; Operario, Darwin; Lee, Jane; Weber, James; Hanson, Robert; Knowler, William; Cooper, Richard; Li, Hongzhe; Seldin, Michael. "Ethnic-Difference Markers for Use in Mapping by Admixture Linkage Disequilibrium." *Am. J. Hum. Genet.* 2002, 70: 737-750.

25. Cooper, Richard; Kaufman, Jay; Ward, Ryk. "Race and Genomics." Sounding Board: *NEJM*. 2003, 348 (12): 1166-1170.

26. Couzin, Jennifer. "New Mapping Project Splits the Community." News Focus: *Science* 2002, 296: 1391-1393.

27. Darwin, Charles. The Descent of Man. New York: D. Appleton and Company, 1896.

28. Davis, James F. Who is Black? One Nation's Definition. University Park: Pennsylvania State University Press, 1991.

29. Disotell, Todd. "Molecular Anthropology and Race." *Annals of the New York Academy of Sciences*. 2000, 925 (1): 9-24. http://www.annalsnyas.org/cgi/content/abstract/925/1/9

30. Dolan DNA Learning Center: Cold Spring Harbor Laboratory. "Image Archive on the American Eugenics Movement." http://www.eugenicsarchive.org/eugenics/. 1//19/03

31. Duster, Troy. "Medicine and People of Color: Unlikely mix---Race, biology and drugs." San Francisco Chronicle. 17 March, 2003. http://sfgate.com/cgi-bin/article.cgi?file=/chronicle/archive/2003/03/17/ED263680.DTL 9/4/03.

32. Duster, Troy. "Some Social Implications of the Molecular Biological Revolution." http://www.aacu-edu.org/SENCER/pdfs/Backgrounders/MolecularBiologicalRevolution.pdf 3/13/03.

33. Epps, Phyllis. "White Pill, Yellow Pill, Brown Pill: Pharmacogenomics and the Changing Face of Medicine." Presented at the conference The Challenges and Impact of Human Genome Research for Minority Communities, presented by Zeta Phi Beta Sorority, Inc, National Education Foundation, July 7-8, 2000, Philadelphia. Available at http://www.ornl.gov/hgmis/publicat/zetaphibeta/epps.html. 2/23/03.

34. Epstein, Steven. "Bodily Differences and Collective Identities: the Politics of Gender and Race in the United States." Publication pending in *Body & Society*. 2001

35. Ewbank, Douglas. "Demography in the Age of Genomics: A First Look at the Prospects." Prepared for discussion at the Meeting on bioindicators, Committee on Population, National Research Council, February 10-11, 2000. Available at http://www.nap.edu/books/0309071992/html/64.html. 10/15/03.

36. Exner, Derek; Dries, Daniel; Domanski, Michael; Cohn, Jay. "Lesser Response to Angiotensin-Converting-Enzyme Inhibitor Therapy in Black as Compared with White Patients with Left Ventricular Dysfunction." *NEJM*, 2001 344 (18): 1351-1357.

37. Foster, Morris and Freeman, William. "Naming Names in Human Genetic Variation Research." Insight/Outlook in *Genome Research*. 1998, 8: 755-757.

38. Foster, Morris and Sharp, Richard. "Race, Ethnicity, and Genomics: Social Classifications as Proxies of Biological Heterogeneity." *Genome Research*, 2002, 12:844-850.

39. Foucault, Michel. "Truth and Power." Power/Knowledge. New York: Pantheon Books, 1977. 109-133.

40. Foucault, Michel. History of Sexuality. New York: Vintage Books. 1978. 136-151.

41. Foucault, Michel. The Birth of the Clinic: An Archaeology of Medical Perception. New York: Random House: Vintage Books, 1994.

42. Freire, Paulo. Pedagogy of the Oppressed. New York: Continuum, 1970, 2000.

43. Friend, Tina. "Genome Mapmakers raced to middle ground." USA Today. 27 June, 2000. D 10.

44. Gannet, Lisa. "Making Populations: Bounding Genes in Space and Time." Available at http://philsci-archive.pitt.edu/archive/00001069/. 5/03.

45. Gannet, Lisa. "The normal genome in twentieth-century evolutionary thought." *Stud. Hist. Phil. Biol. & Biomed. Sci.* 2003, 34: 143-185.

46. Griffith, Victoria. "Genes reveal ethnic origins." The Financial Times (London). 20 December, 2002: International Economy, 7.

47. Griffith, Victoria. "Wires cross over genes as information on ethnic groups pours in." The Financial Times (London). 2 November, 2002: A1.

48. Grover, Jan Zita. "AIDS: Keywords." AIDS: Cultural Analysis, Cultural Activism. Ed. Crimp, Douglas. Cambridge: MIT Press. 1989.

49. Graham, Mary. Murri elder, scholar and activist. Murri Murra Cultural Center. Brisbane, QLD, Australia. Personal communication. February 2002.

50. Guttmacher, Alan and Collins, Francis. "Genomic Medicine — A Primer." Review Article in *NEJM*, 2002, 347 (19): 1512-1520.

51. Guyer, Mark and Collins, Francis. "How is the Human Genome Project doing, and what have we learned so far?" Review in *Proc. Natl. Acad. Sci. USA*, 1995, 92: 10841-10848.

52. Haraway, Donna J. "Race: Universal Donors in a Vampire Culture." Modest_Witness@ Second_Millenium.FemaleMan© Meets_OncoMouse™. New York: Routledge, 1997. 213-266.

53. Haraway, Donna. "Situated Knowledges: the Science Question in feminism and the Privilege of Partial Perspective." 1988. The Science Studies Reader. Ed. Biagioli. New York: Routledge, 1999. 172-188.

54. Hedgecoe, Adam and Tutton, Richard. "Genetics in Society/Society in Genetics." Guest editorial: *Science as Culture*, 2002, 11(4): 421-428.

55. Helmuth, Laura. "Map of the Human Genome 3.0" *Science*, 2001, 293 (5530): 583-585.

56. Herrnstadt, Corinna; Elson, Joanna; Fahy, Eoin; Preston, Gwen; Turnbull, Douglass; Anderson, Christen; Ghosh, Soumitra; Olefsky, Jerrold; Beal, M. Flint, Davis, Robert; Howell, Neil. "Reduced-Median-Network Analysis of Complete Mitochondrial DNA Coding-Region Sequences for the Major African, Asian, and European Haplogroups." *Am. J. Hum. Genet.* 2002, 70:1152-1171.

57. Holden, Constance. "Race and Medicine." Editorial in *Science*. 2003, 302: 594-596.

58. International Human Genome Sequencing Consortium. "Initial sequencing and analysis of the human genome." *Nature*. 2001, 409: 860-921.

59. International SNP Map Working Group. "A Map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms." *Nature* 2001, 409: 928-933.

60. International Union of Anthropological and Ethnological Sciences. "Proposed Replacement Statement for the Unesco Documents on Biological Aspects of Race." http://www.leidenuniv.nl/fsw/iuaes/08-race.htm. 4/12/03.

61. Jackson, Deborah. "Human Genome Breakthrough." *Border Voices of the BEC.* Border Epidemiology and Environmental Health Center. 2001 http://www.nmsu.edu/~bec/genome.pdf

62. Jackson, John P. "Editor's Forward." Science, Race, and Ethnicity: Readings from Isis and Osiris. Ed. Jackson, John P. Chicago: The University of Press, 2002. 1-4.
63. Jacqueline Stevens. "Racial Meanings and Scientific Methods: Changing Policies for NIH-sponsored Publications Reporting Human Variation." *Journal of Health Politics, Policy and Law*. 2003, 28(6): 1033-1088.

64. Jorde, L.B.; Watkins, W.S.; Bamshad, M.J. "Population genomics: a bridge from evolutionary history to genetic medicine." *Human Molecular Genetics*, 2001, 10 (20) 2199-2207.

65. Juengst, Eric. "Group Identity and Human Diversity: Keeping Biology Straight from Culture." *Am. J. Hum. Genet.* 1998, 63: 673-677.

66. Kagawa-Singer, Marjorie. "From Genes to Social Science: Impact of the Simplistic Interpretation of Race, Ethnicity, and Culture on Cancer Outcome." From the 7th Bieennial Symposium on Minorities, the Medically Underserved and Cancer. Washington, DC, February 9-13, 2000. *Supplement to* ***Cancer*** 2001, 91(1): 226-232.

67. Kaplan, Erin Aubry. "Black Like I Thought I Was." LA Weekly. 7 October 2003. available http://www.alternet.org/story.html?StoryID=16917 10/10/03.

68. Kay, Lily. "In the Beginning Was the Word? The Genetic Code and the Book of Life." 1998. The Science Studies Reader. Ed. Biagioli. New York: Routledge, 1999. 224-233.

69. Kemp, Martin. "The *Mona Lisa* of modern science." Feature in *Nature*. 2003, 421: 416-420.

70. Kolata, Gina. "Using Genetic Tests, Ashkenazi Jews Vanquish a Disease" New York Times, 18 February, 2003. F 1.

71. Kristof, Nicholas. "Is Race real?" New York Times 11 July 2003. Available at http://www.racesci.org/in_media/is_race_real.htm. 9/29/03,

72. Leffall, Le Salle. "Foreword." From the 7th Bieennial Symposium on Minorities, the Medically Underserved and Cancer. Washington, DC, February 9-13, 2000. *Supplement to* ***Cancer*** 2001, 91(1) 189-192.

73. Lehrman, Sally. "The Reality of Race." *Scientific American* 13 January, 2003. http://www.sciam.com/article.cfm?articleID=0002A353-C027-1E1C-8B3B809EC588EEDF. 3/13/03.

74. Lewontin, Richard. "The apportionment of human diversity." *Evolutionary Biology*. 1972, 6: 381-398.

75. Lonjou, C., Collins, A., Morton, N.E. "Allelic association between marker loci." *Proc. Natl. Acad. Sci. USA.* 1999, 96: 1621-1626.

76. M'charke, Amade. "Technologies of Population: Forensic DNA Testing Practices and the Making of Differences and Similarities." *Configurations*, 2000, 8:121-158.

77. Ma, Xin; Wright, John; Dou, Shunjun; Olsen, Paul. Tecter, Larry; Adams, Gerald; Graviss, Edward. "Ethnic divergence and linkage disequilibrium of novel SNPs in the human *NLI-IF* gene: evidence of human origin and lack of association with tuberculosis susceptibility." *J. Hum. Genet.* 2002, 47:140-145.

78. Macilwain, Colin. "World leaders heap praise on human genome landmark." News in *Nature*. 2000, 405: 983-984.

79. Marks, Jonathan. "Folk Heredity." In Race and Intelligence: Separating Science from Myth. Ed. Fish, J. New York: Lawrence Erlbaum, 2002. 95-116.

80. Mcintosh, Peggy. "White Privilege Unpacking the Invisible Knapsack." Excerpted from "White Privilege and Male Privilege: A Personal Account of Coming To See Correspondence through Work in Women's Studies." 1988. Available at http://www.utoronto.ca/acc/events/peggy1.htm. 11/18/03.

81. McLeod, Howard and Evans, William. "Pharmacogenomics: Unlocking the Human Genome for Better Drug Therapy." *Annu. Rev. Pharmacol. Toxicol.* 2001. 41: 101-121.

82. Morales, Josè. "Genomic Justice: Environmental Justice Policy on Human Biotechnology." Presented to the Second National People of Color Environmental Leadership Summit – Summit II. Available at http://www.ejrc.cau.edu/summit2/Summit%20Exec%20Sum%202002.pdf. 2/03.

83. Morrison Institute for population and resource studies. HGDP North American Committee. "Human Genome Diversity Project: Frequently Asked Questions." 1993-1994. Available at http://www.stanford.edu/group/morrinst/hgdp/faq.html.

84. Nature Publishing Group (no author listed) "Census, race and science." Editorial *Nature Genetics*. 2000, 24(2): 97-98.

85. NitroMed, Inc. "NitroMed Initiates Confirmatory BiDil© Trial in African American Heart Failure Patients: First prospective heart failure trial conducted exclusively for black patients." 17 March, 2001. http://nitromed.com/newsindex.html, 9/4/03.

86. Olson, Steve. "The Genetic Archaeology of Race." The Atlantic Monthly. April 2001: http://www.theatlantic.com/issues/2001/04/olson-p1.htm 2/23/03.

87. Omi, Michael, and Winant, Howard. Racial Formation in the United States: from the 1960s to the 1980s. New York: Routledge and Keagan Paul, 1986.

88. Owens, Kelly and King, Mary-Claire. "Genomic Views of Human History." Viewpoint in *Science*. 1999. 286: 451-453.

89. Pääbo, Svante. "The Human Genome and Our View of Ourselves." Opinion in *Science* 2001, 291(5507): 1219-1220.

89. Parra, Esteban; Marcini, Amy; Akey, Joshua; Martinson, Jeremy; Batzer, Mark; Cooper, Richard; Forrester, Terrence; Allison, David; Deka, Ranjan; Ferrell, Robert; Shriver, Mark. "Estimating African American Admixture Proportions by Use of Population-Specific Alleles." *Am. J. Hu. Genet.* 1998, 63: 1839-1851.

90. Parra, Flavia; Amado, Roberto; Lambertucci, Josè; Rocha, Jorge; Antunes, Carlos; Pena, Sèrgio. "Color and genomic ancestry in Brazilians." *PNAS.* 2003, 100(1): 177-182.

91. *Peacework Magazine*, no author indicated. "Constructs of Race Difference." February 2002. http://www.afsc.org/peacework/ 2/14/03.

92. Philipkoski, Kristen. "Gene Map Presents Race Concerns." Wired. http://www.wired.com/news/technology/0,1282,41619,00.htm. 5/5/03.

93. Pigliucci, Massimo and Kaplan, Jonathan. "On the Concept of Biological race and its applicability to humans." Available at http://philsci-archive.pitt.edu/archive/00001078/. 5/03.

94. Pollack, Adrew and Altman, Lawrence. "Large Trial Finds AIDS Vaccine Fails to Stop Infection." New York Times. 24 February, 2003. http://www.nytimes.com. 2/24/03.

95. Rabinow, Paul. "Artificiality to Enlightenment: from Sociobiology to Biosociality." 1992, 1998. The Science Studies Reader. Ed. Biagioli. New York: Routledge, 1999. 407-416.

96. Race-The Power of An Illusion: Episode One: The Difference Between Us. Public Broadcasting Station (PBS). Transcript available at http://www.newsreel.org/transcripts/race1.htm. 11/19/03.

97. Reardon, Jennifer. Personal communication, Brown University. March, November, 2003.

98. Reardon, Jennifer. Race to the Finish: Identity and Governance in and Age of Genetics. Princeton University Press, forthcoming, 2004.

99. Risch, N., Burchard, E., Ziv, E., Tang, H. "Categorization of humans in biomedical research: genes, race and disease." *Genome Biology* 2002. **3**:7. http://genomebiology.com/2002/3/7/comment/2007.1

100. Risch, Neil. "Searching for genetic determinants in the new millennium." *Nature*, 2000, 405: 847-856.

101. Rockquemore, Kerry Ann, and Brunsma, David L. Beyond Black: Biracial Identity in America. London: Sage Publications: 2002.

102. Root, Michael. "The Use of Race in Medicine as a Proxy for Genetic Difference." http://www.philsci-archive.pitt.edu/archive/00001094. 5/03.

103. Rosenberg, Noah; Pritchard, Jonathan; Weber, James; Cann, Howard; Kidd, Kenneth; Zhivotovsky, Lev; Feldman, Marcus. "Genetic Structure of Human Populations." *Science* 2002, 298: 2381-2385. Also, "Supplementary information for "Genetic structure of human populations," *ibid*.

104. Rothstein, M.A and Epps, P.G. "Pharmacogenomics and the (ir)relevance of race." *The Pharmacogenomics Journal*. 2001, 1: 104-108.

105. Rotman, David. "Genes, Medicine, and the New Race Debate." MIT: Technology Review. 24 July, 2003. Available at http://www.techreview.com/articles/rotman0603.asp. 10/15/03.

106. Sankar, Pamela and Cho, Mildred. "Toward a New Vocabulary of Human Genetic Variation." Policy Forum article in *Science*, 2002, 298: 1337-2002.

107. Schwartz, Robert. "Racial Profiling in Medical Research." Editorial in *NEJM*. 2001, 344(18): 1392-1393.

108. Shavers, Vicki; Lynch, Charles; Burmeister, Leon. "Factors That Influence African-Americans' Willingness to Participate in Medical Research Studies." From the 7th Biennial Symposium on Minorities, the Medically Underserved and Cancer. Washington, DC, February 9-13, 2000. *Supplement to* ***Cancer*** 2001, 91(1) 221-225.

109. Shelton, Brett Lee and Marks, Jonathan. "Genetic Markers not a Valid Test of Native Identity." Indigenous Peoples Council on Biocolonialism, date of original publication not indicated. http://www.ipcb.org/publications/briefing_papers/files/identity.html. 2/14/03.

110. Smith, Michael; Lautenberger, James; Shin, Hyoung Doo; Chretien, Jean-Paul; Shrestha, Sadeep; Gilbert, Dennis; O'Brien, Stephen. "Markers for Mapping by Admixture Linkage Disequilibrium in African American and Hispanic Populations." *Am. J. Hum. Genet.* 2001, 69: 1080-1094.

111. SNP Consortium, LTD. Various sites at http://www.snp.cshl.org, including "TSC Allele Frequency Project," November 2001. http://snp.cshl.org/allele_frequency_project/afp_summary_nov2001.shtml. 2/22/03.

112. Soo-Jin Lee, Sandra; Mountain, Joanna; Koenig, Barbara. "The Meanings of 'Race' in the New Genomics: Implications for Health Disparities Research." *Yale Journal of Health Policy, Law and Ethics*. 2001, 1: 33-75.

113. Stepan, Nancy Leys. "Race and Gender: The Role of Analogy in Science." Science, Race, and Ethnicity: Readings from Isis and Osiris. Ed. Jackson, John P. Chicago: The University of Press, 2002. 5-21.

114. Stepan, Nancy. The Idea of Race in Science: Great Britain, 1800-1960. London: Macmillan: Archon Books, 1982.

115. Stevens, Jacqueline. "Symbolic Matter: DNA and other linguistic stuff." *Social Text*. 2002, 20 (1): 105-146.

116. Stocking, Clinton. "Concepts of Race in Culture and Biology: From Carolus Linneaus to Claude Levi-Strauss." University of Chicago, 2002. Available at htttp://www.cmet.com/conceptsrace.htm. 11/19/03.

117. Syracuse University. "All of us are related, each of us is unique." 2002. Available at http://www.allrelated.syr.edu/fulltext.html. 2/03.

118. Thorisson, Gudmundur; Stein, Lincoln. "The SNP Consortium website: past, present and future." *Nucleic Acids Research*. 2003, 31(1): 124-127.

119. United Nations. United Nations Educational, Scientific and Cultural Organization (UNESCO). The Race Question in Modern Science: The Race Concept: Results of an Inquiry. Paris: UNESCO, Imprimerie des Arts et Manufactures, 1952.

120. United Nations. United Nations Educational, Scientific and Cultural Organization (UNESCO) Statement on Race. Proposed by participants at the Scientific Workshop of the International UNESCO-Conference Against Racism, Violence and Discrimination, June 8 and 9, 1995. http://www.uni-oldenburg.de/biodidaktik/race.html. 4/12/03

121. United Nations. United Nations Scientific, Educational and Cultural Organization (UNESCO). Address by Sane, Pierre to World Conference Against Racism Durban, South Africa, 21 August-7 September, 2001. http://www.un.org/WCAR/statements/unescoE.htm. 4/12/03.

122. United Nations. World Health Organization. Genomics and World Health: Report of the Advisory Committee on Health Research. Geneva, 2002.

123. United States. "Executive Order 13145: To Prohibit Discrimination in Federal Employment Based on Genetic Information." National Human Genome Research Institute, 2000. http://www.nhgri.nih.gov/NEWS/Executive_order/index.html. 2/14/03.

122. United States. Centers for Disease Control (CDC). "Use of Race and Ethnicity in Public Health Surveillance Summary of the CDC/ATSDR Workshop." 25 June, 1993. available at wysiwyg://87/http://www.cdc.gov/mmwr/preview/mmwrhtml/00021729.htm. 2/03.

124. United States. Department of Health and Human Services. National Institutes of Health. Request for Application: HG-98-001. "Methods for Discovering and Scoring Single Nucleotide Polymorphisms." 9 January 1998. Available at http://grants.nih.gov/grants/guide/rfa-files/RFA-HG-98-001.html 3/03

125. United States. Department of Health and Human Services. National Institutes of Health. National Human Genome Research Institute. National Human Genome Research Institute. "ELSI Research Program: ELSI Grant Abstracts." http://www.genome.gov/10001796, http://www.genome.gov/10002224. 2/12/02

126. United States. Department of Health and Human Services. National Institutes of Health. National Human Genome Research Institute. National Human Genome Research Institute. Press Release: "Bringing the Human Genome Project into the Classroom: Human Genome Project Unveils Multimedia Educational Kit for High School Students and Public." February, 2001. http://www.genome.gov/10002193. 2/14/03.

127. United States. Department of Health and Human Services. National Institutes of Health. National Human Genome Research Institute. National Human Genome Research Institute: "Dissertation Research Grants for Underrepresented Minorities in the Ethical, Legal, and Social Implications (ELSI) of Genetics Research." 2002, http://grants1.nih.gov/grants/guide/pa-files/PA-02-048.html. 2/12/03.

128. United States. Department of Health and Human Services. National Institutes of Health. National Human Genome Research Institute. National Institutes of Health and National Human Genome Research Institute: Teacher's Guide: "Human Genetic Variation." http://science-education.nih.gov/Customers.nsf/HSGenetic?OpenForm&CS_11=false&CS_12=false&CS_13=false&CS_22=false& 2/14/03.

129. United States. Department of Health and Human Services. National Institutes of Health. "Social and Demographic Studies of Race and Ethnicity in the United States." 2003. http://grants1.nih.gov/grants/guide/pa-files/PA-03-057.html 2/23/03.

130. United States. Department of Health and Human Services. National Institutes of Health. Powerpoint Presentation. "Sex/Gender and Minority Inclusion in NIH Clinical Research: What Investigators Need to Know!" http://grants1.nih.gov/grants/funding/women_min/training/sld001.htm 2/7/03.

131. United States. Department of Health and Human Services. National Institutes of Health. Online hub for NIH regulations "Sex/Gender and Minority Inclusion in NIH Clinical Research" "http://grants1.nih.gov/grants/funding/women_min/training/ 11/21/03.

132. United States. Department of Health and Human Services. National Institutes of Health (NIH). "FY 2002 Hearing on Health Disparities." Testimony from Francis Collins to the House Subcommittee on Labor-HHS-Education Appropriations. 4 April, 2001. Available at http://www.hhs.gov/budget/testify/b20010404i.html. 2/14/03.

133. United States. Department of Health and Human Services. National Institutes of Health. National Human Genome Research Institute. "NHGRI: Request for Applications (RFA HG-99-002): *Studies of the Ethical, Legal and Social Implications of Research Into Human Genetic Variation*." 29 April, 1999. Available at http://aeolus.nhgri.nih.gov/page.cfm?pageID=10001017. 2/03.

134. United States. Department of Health and Human Services. National Institutes of Health. National Institute for Environmental Health Sciences. National Center for Toxicogenomics and the Environmental Genome Project. http://www.niehs.nih.gov/nct/concept.htm. 4/03.

135. United States. Department of Health and Human Services. National Institutes of Health. National Human Genome Research Institute. "International Consortium Launches Genetic Variation Mapping Project: HapMap Will Help Identify Genetic Contributions to Common Diseases." NIH New Advisory, October 2002. Available at http://www.nih.gov/news/pr/oct2002/nhgri-29.htm 3/03.

136. United States. Department of Health and Human Services. National Institutes of Health. National Human Genome Research Institute. "Understanding Human Genetic Variation." Teacher's Guide. http://www.science.education.nih.gov/supplements/nih1/genetic/guid/genetic_Variation3.htm 2/3/03.

137. United States. Department of Health and Human Services. National Institutes of Health. National Institute of General Medicine. "Report of the First Community Consultation on the Responsible Collection and Use of Samples for Genetic Research." 25-26 September, 2000. Available at http://www.nigms.nih.gov/news/reports/community_consultation.html. 3/03.

138. United States. Department of Health and Human Services. "FY 2004 Budget in Brief." http://www.hhs.gov/budget/04budget/fy2004bib.pdf 2/12/03.

139. United States. Department of Health and Human Services: Food and Drug Administration: "International Conference on Harmonisation; Guidance on Ethnic Factors in the Acceptability of Foreign Clinical Data; Availability." *Federal Register* 1998, 63 (111): 31790-31796.

140. United States. National Institutes of Health (NIH). National Cancer Institute (NCI). "Statements from the President's Cancer Panel Meeting: The Meaning of Race in Science—Considerations for Cancer Research." http://deainfo.nci.nih.gov/ADVISORY/pcp/09apr97.htm 4/12/03.

141. United States. National Institutes of Health (NIH). National Human Genome Research Institute. Internet video: "Exploring Our Molecular Selves." February, 2001. Component of NHGRI's educational resources. Available at http://www.genome.gov/10000002. 2/14/03.

142. United States. National Institutes of Health (NIH). National Human Genome Research Institute. Internet video: "Variation." February, 2001. Component of NHGRI's educational resources. Available at http://www.genome.gov/10000002. 2/14/03.

143. United States. National Institutes of Health (NIH). National Human Genome Research Institute. Internet video: "History." February, 2001. Component of NHGRI's educational resources. Available at http://www.genome.gov/10000002. 2/14/03.

144. United States. National Institutes of Health (NIH). National Human Genome Research Institute (NHGRI). Foster, Morris. "Ethical Issues in Developing a Haplotype Map with Socially Defined Populations." Available at http://www.genome.gov/10001683. 11/14/03.

145. United States. National Institutes of Health. McManus, Rich. "'First Order' Wisdom Meets 'Second Order' World: Once Declared Meaningless, Race Still Influences, Says Duster." Interview with Troy Duster. *NIH Record*. 12 June 2001. http://www.nih.gov/news/NIH-Record/06_12_2001/story03.htm. 4/12/03.

146. United States. Office of Management and Budget (OMB). "Guidance on Aggregation and Allocation of Data on Race for Use in Civil Rights Monitoring and Enforcement." OMB Bulletin 00-02. http://www.whitehouse.gov/omb/bulletins/b00-02.html. 3/13/03.

147. United States. Office of Management and Budget (OMB). "Recommendations from the Interagency Committee for the Review of the Racial and Ethnic Standards to the Office of Management and Budget Concerning Changes to the Standards for the Classification of Federal Data on Race and

Ethnicity." *Federal Register* 7/9/97, part 2: 36873-36946. http://www.whitehouse.gov/omb/fedreg/directive_15.html 3/13/03.

148. United States. Office of Management and Budget (OMB). "Standards for the Classification of Federal Data on Race and Ethnicity." August, 1995: http://www.whitehouse.gov/omb/fedreg/race-ethnicity.html 3/13/03.

149. United States. Office of Management and Budget (OMB). "Standards for the Classification of Federal Data on Race and Ethnicity." *Federal Register*. 6/9/1994. http://clinton3.nara.gov/OMB/fedreg/notice_15.html. 3/13/03.

150. United States. Office of the Press Secretary. "Remarks by the President, Prime Minister Tony Blair of England (via satellite), Dr. Francis Collins, Director of the National Human Genome Research Institute, And Dr. Craig Venter, President and Chief Scientific Officer, Celera Genomics Corporation, On the Completion of the first Survey of the Entire Human Genome Project." Delivered at the White House East Room. 26 June, 2000. Available at http://www.ornl.gov/sci/techresources/Human_Genome/project/clinton2.shtml. 2/23/03.

151. United States. Office of the Press Secretary. Clinton, William. "Remarks by President Clinton on Genetic Discrimination." Delivered Ferbruary 8, 2000 at a meeting of the American Association for the Advancement of Science. Available at http://www.geneticalliance.org/geneticissues/clintonremarks.html. 1/26/03.

152. United States. Department of Health and Human Services. "HHS Policy for Improving Race and Ethnicity Data." Office of Information Resources Management, 1997. http://www.hhs.gov/oirm/infocollect/nclusion.html. 2/7/03.

153. United States. Department of Health and Human Services. National Institutes of Health. National Human Genome Research Institute. "National Human Genome Research Institute Five-Year Strategic Plan for Reducing Health Disparities." Submitted to The National Center Minority Health and Health Disparities, for The NIH Comprehensive Strategic Plan To Reduce and Ultimately Eliminate Health Disparities. 15 November, 2001. Available at http://www.genome.gov/10001492. 3/03.

154. United States. Department of Health and Human Services. National Institutes of Health. National Human Genome Research Institute. "NHGRI Special Populations Research Program Collaborations with Howard University." Available at http://www.genome.gov/10000831. 9/03.

155. United States. Department of Health and Human Services. National Institutes of Health. National Human Genome Research Institute. "Workshop on Human DNA Sequence Variation." March 31- April 1, 1997. Available at http://www.genome.gov/10001478. 3/03.

156. United States. Department of Health and Human Services. National Institutes of Health. National Human Genome Research Institute. "Developing a Haplotype Map of the Human Genome for Finding Genes Related to Health and Disease." 18-19 July, 2001. http://www.genome.gov/10001665. 3/03.

157. United States. Department of Health and Human Services: Food and Drug Administration: "Guidance for Industry: Collection of Race and Ethnicity Data in Clinical Trials: Draft Guidance." *Federal Register* 2003. http://www.fda.gov/cder/guidance/5054dft.pdf 1/23/03.

158. Wade, Nicholas. "For Genome Mappers, the Tricky Terrain of Race Requires Some Careful Mapping." New York Times. 20 July, 2001: A 17.

159. Wade, Nicholas. "For Sale: A DNA Test To Measure Racial Mix." New York Times. 1 October, 2002. F 4.

160. Wade, Nicholas. "Gene Study Identifies 5 Main Human Populations, Linking Them to Geography." New York Times. 20 December, 2002: A 1.

161. Wade, Nicholas. "Race Is Seen as Real Guide to Track Roots of Disease." New York Times. 30 July, 2002. wysiwyg://34/http://www.nytimes.com/2002/ ...30RACE.html?pagewanted=print&position=top. 9/29/03.

162. Wade, Nicholas. "Reading the Book of Life: Now, the Hard Part: Putting the Genome to Work." New York Times. 27 June, 2000: F 1.

163. Wade, Nicholas. "Reading the Book of Life: tools Already in Use; Whose DNA Is It? In a Way, Nobody's." New York Times. 27 June, 2000: F 2.

164. Weiss, Sheila Faith. "The Race Hygiene Movement in Germany." Science, Race, and Ethnicity: Readings from Isis and Osiris. Ed. Jackson, John P. Chicago: The University of Press, 2002. 225-268.

165. Wertz, Dorothy. "Genetic testing in the workplace: The Lawrence Berkeley Labs case." *Gene Letter* April 2000, 1(3).

166. Wires, Stephen Brook. "Race has no place in genome – decoding humanity." The Australian. 28 June, 2000: Features 14.

167. Witzig, Ritchie. "The Medicalization of Race: Scientific Legitimization of a Flawed Social Construct." *Annals of Internal Medicine* 1996. 125:675-679.

168. Wood, Alastair. "Racial Differences in the Response to Drugs—Pointers to Genetic Differences." *NEJM*. 2001, 344(8) 1393-1395.

169. Wyatt, Stephen; James, Lovell; McGinnis, LaMar; Weinberg, Armin. "Introduction." From the 7th Biennial Symposium on Minorities, the Medically

Underserved and Cancer. Washington, DC, February 9-13, 2000. *Supplement to **Cancer*** 2001, 91(1) 193-194.

170. Xie, Hong-Guang; Kim, Richard; Wood, Alastair; Stein, Michael. "Molecular Basis of Ethnic differences in Drug Disposition and Response." *Annu. Rev. Pharmacol. Toxicol.* 2001, 41: 815-50.

171. Yancy, Clyde; Fowler, Michael; Colucci, Wilson; Gilbert, Edward; Bristow, Michael; Cohn, Jay; Lukas, Mary Ann; Young, Sarah; Packer, Milton; for the U.S. Carvedilol Heart failure Study Group. "Race and the Response to Adrenergic Blockade with Carvedilol in Patients with Chronic Heart Failure." *NEJM*: 2001, 344 (18): 1358-1365.

172. Yoxen, Edward. 1984. "Constructing Genetic Disease." In Troy Duster and Karen Garett, eds., Cultural Perspectives on Biological Knowledge. Norwood, NJ: Ablex. 41-62.

173. Library of congress, 1991, history of Nigeria http://lcweb2.loc.gov/cgi-bin/query/r?frd/cstdy:@field(DOCID+ng0012)