Murray, Charles. Human Diversity: The Biology of Gender, Race, and Class. Twelve, 2020.

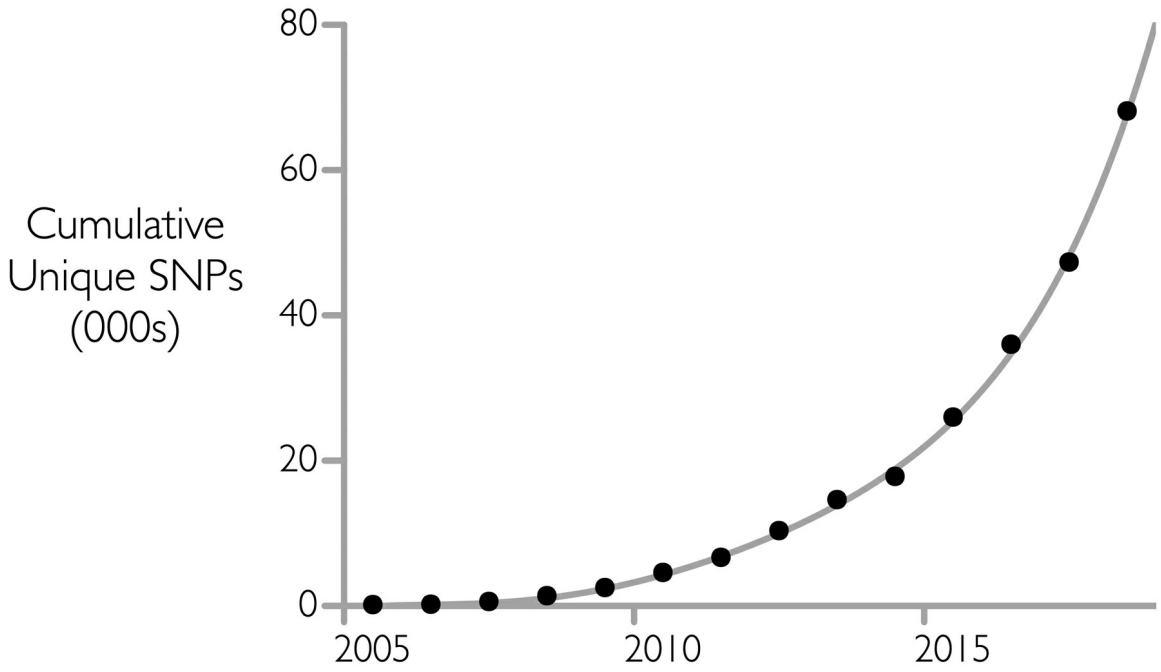
9

The Landscape of Ancestral Population Differences

Proposition #7: Continental population differences in variants associated with personality, abilities, and social behavior are common.

This chapter is about raw ancestral population differences in SNPs that are statistically related to cognitive repertoires—"raw" meaning that a great deal of work remains to be done before the significance of such differences is understood.

Until a few years ago, this topic was still *terra incognita.* Only a handful of statistical relationships between specific SNPs and cognitive traits had been identified. But the growth in that number has been phenomenal, paralleling the growth in the number of SNPs associated with diseases and physiological traits. To illustrate what's been happening, consider the GWAS Catalog. It was begun in 2008 by the National Human Genome Research Institute, part of the U.S. National Institutes of Health.[1] The first year with a published GWAS was 2005, when two studies reported a grand total of two SNPs.[2] In 2018 alone, the GWAS Catalog added 17,182 previously unidentified SNPs. Here's what the history looks like:



Source: Author's analysis, GWAS Catalog.

As of the end of May 2019, the catalog included 3,469 studies reporting 136,286 variants. The total number of unique variants was 89,544. And that's just a fraction of the total number of variants that have been associated with phenotypic traits at lesser levels of statistical significance. That total is over a million, residing in databases maintained by university and private sector research centers scattered around the world.

Hardly any analyses of this burgeoning knowledge base have compared results for different continental ancestral populations (which for readability I will subsequently abbreviate to "continental populations"). Researchers have been wary of such comparisons because the results can't be trusted. Paradoxically, the reason they can't be trusted has indirectly become the reason that continental population differences will soon be studied intensively.

Why Continental Population Differences Will Be Studied

For the last quarter century, medical researchers have been grappling with evidence that what's true about a disease for one ancestral population isn't necessarily true for another. It's called *population stratification.*

The problem was first suspected in the mid-1990s. Medical researchers looking for candidate genes routinely compared the genetics of a group with the disease being studied with a comparison sample of people without the disease. In choosing those samples, ancestry was initially not a consideration. But as researchers got access to more genetic information, they began to worry that their results were being contaminated because the ancestral populations in the samples had different genetic profiles.[3] A back-and-forth debate ensued in the technical literature. By 2004, the weight of the evidence had become clear. As one of the early DNA-based studies concluded, "Even small amounts of population admixture can undermine an association study and lead to false positive results. These adverse effects increase markedly with sample size. For the size of study required for many complex diseases, relatively modest levels of structure within a population can have serious consequences."[4]

POLYGENIC SCORES

Polygenic scores will be discussed in detail in chapter 14. For now, think of them as analogous to test scores, but based on combined allele frequencies instead of combined answers to test questions. A polygenic score for schizophrenia (for example) measures the genetic risk of schizophrenia.

A decade later, the first studies using polygenic scores verified an explanation for population stratification that generalizes far beyond the study of diseases: Polygenic scores for one continental population don't work as well for other continental populations no matter what the trait may be. In technical terms, the predictive validity of a polygenic score deteriorates as the genetic distance between the test population and the comparison population increases, consistent with population genetics theory.[5] For example, a polygenic score based on a test population of English and Italians usually generalizes accurately for French and Germans, not so accurately for Chinese and Indians, and least accurately for the genetically most distant populations from sub-Saharan Africa.[6]

Population differences in predictive validity could reflect natural selection, genetic drift, or gene × environment interactions. Population geneticists have had strong scientific motivation to learn more about those differences but have been frustrated because the artifacts produced by population stratification are so common.[7] Statistical analysis can correct for some of population stratification's effects, but the only full solution is to have large samples from all the ancestral populations that are being compared. The problem is that genomic data have typically been collected from people who lived in the nations where geneticists worked, dominated by Europe and the United States, which in turn meant that large genomic databases were overwhelmingly people of European ancestry.

The collection of large samples from non-European populations was on the back burner through the first half of the 2010s. It's understandable—samples in the hundreds of thousands are logistically demanding, and the foundations and government agencies with deep enough pockets to fund such samples have not (until recently) put them on their agendas. There also has been a lack of urgency: Geneticists have been kept busy with an ample supply of GWA research projects that can be done with European samples.

Then murmurings about underrepresentation of non-Europeans in genomic databases began appearing. They reached a broad audience in 2018 when British geneticist David Curtis charged that by using European samples, "UK medical science stands at risk of being institutionally racist."[8] In 2019, an article by a team of American geneticists in *Cell*, "The Missing Diversity in Human Genetic Studies," widely picked up by the media, detailed the many ways in which the bias toward European samples "effectively translates into poorer disease prediction and treatment for individuals of under-represented ancestries."[9]

It now appears likely that large samples from underrepresented populations—notably Africans and South Asians—will be available soon (China and Japan have been building such databases on their own). When they come online, ancestral population differences related to disease are going to be studied minutely.

Those same databases will potentially allow researchers to study genetic differences in personality traits, abilities, and social behavior across continental populations. That potential is likely to generate cross-cutting pressures. For highly charged topics such as IQ, many people will continue to urge that studying population differences does more harm than good. But

what happens if findings from European samples about cognitive-related traits such as depression, autism, or schizophrenia lead to more effective treatments for Europeans but not for other populations? It will be ethically imperative to study the genetics of mental disorders in other populations as well, which means studying the ways in which they differ from Europeans. The idea that geneticists could ignore ancestral population differences indefinitely was always implausible. It is now out of the question.

Differences in Allele Frequencies Within and Across Continental Population

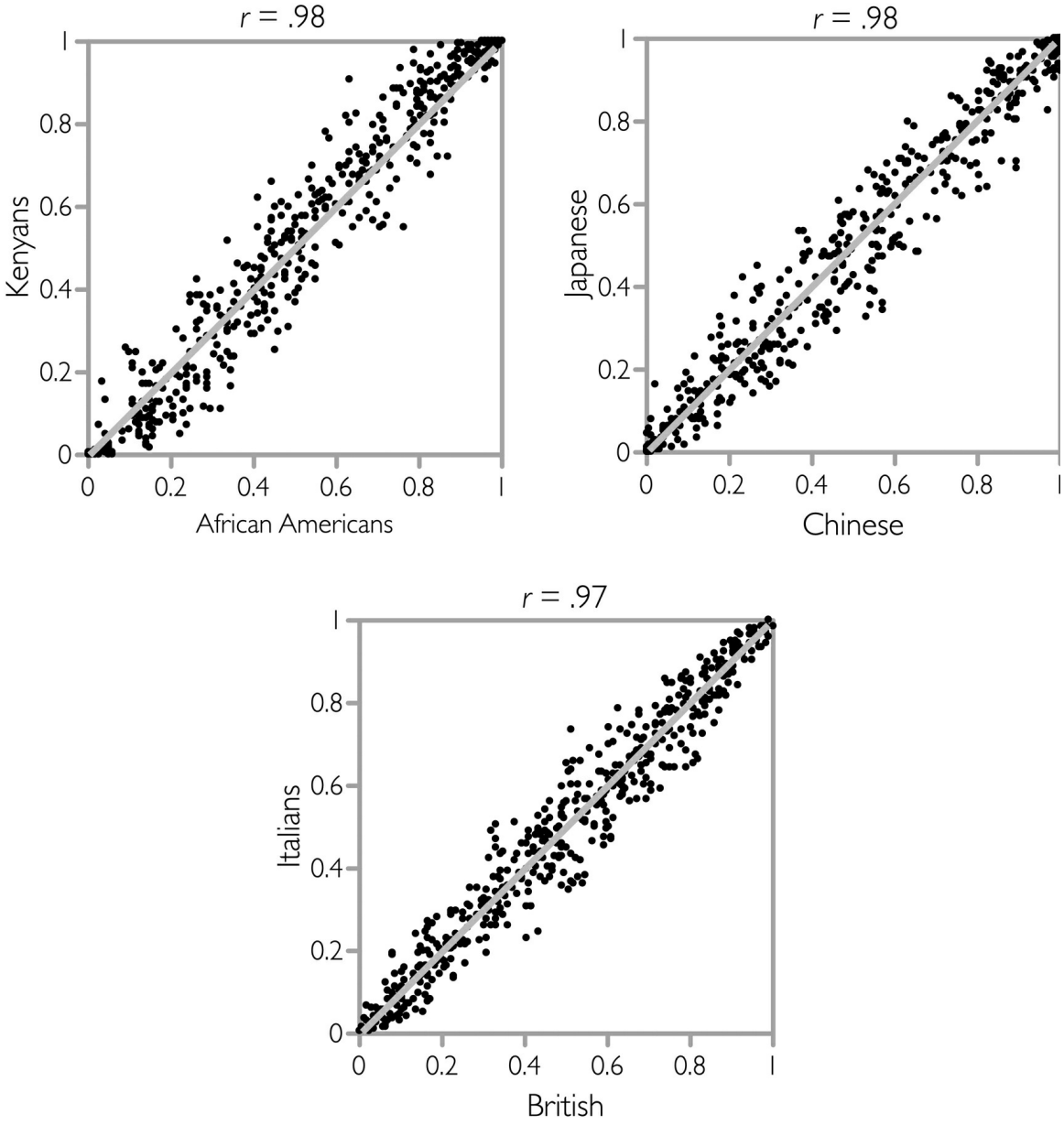
When SNPs cause differences in phenotypic traits, evidence for that role surfaces first in differences in target allele frequencies. The *target allele* is usually defined as one that is associated with an increase in the magnitude or intensity of a trait. If the topic is diabetes, the target alleles are the ones associated with an increase in the risk or severity of diabetes. If the topic is IQ, the target alleles are those associated with increases in IQ scores. Other labels used in the literature include *risk allele*, *effect allele*, and *increasing allele*. As in chapter 8, I express target allele frequencies exclusively as proportions ranging from 0 to 1 rather than as percentages of chromosomes.

My purpose in this discussion is limited to the wording of Proposition #7 as it applies to common SNPs: Continental population differences in target allele frequencies associated with personality, abilities, and social behavior are common. *I am not presenting proof that those differences cause phenotypic differences*, but showing you how different the situation actually facing geneticists is from the impression you may have when you hear that "race is a social construct." Virtually all traits, whether physiological, related to disease, or related to cognitive repertoires, exhibit many large differences in target allele frequencies across continental populations.

Comparing Subpopulations from the Same Continent

I'll use a specific example, schizophrenia, as an entry point to the topic. First,

consider the landscape for subpopulations within the same continental population. The following graphs show what happens when the target allele frequencies for two populations are plotted against each other for three within-continent pairs: Kenyans and African Americans, British and Italians, and Chinese and Japanese.



Source: Author's analysis, GWAS Catalog, and Phase 1 of the 1000 Genomes Project. A total of 962 SNPs in the GWAS Catalog are associated with schizophrenia. For this and the subsequent graphs, I chose 500 to plot (962 in a small graph would produce too many overlays,

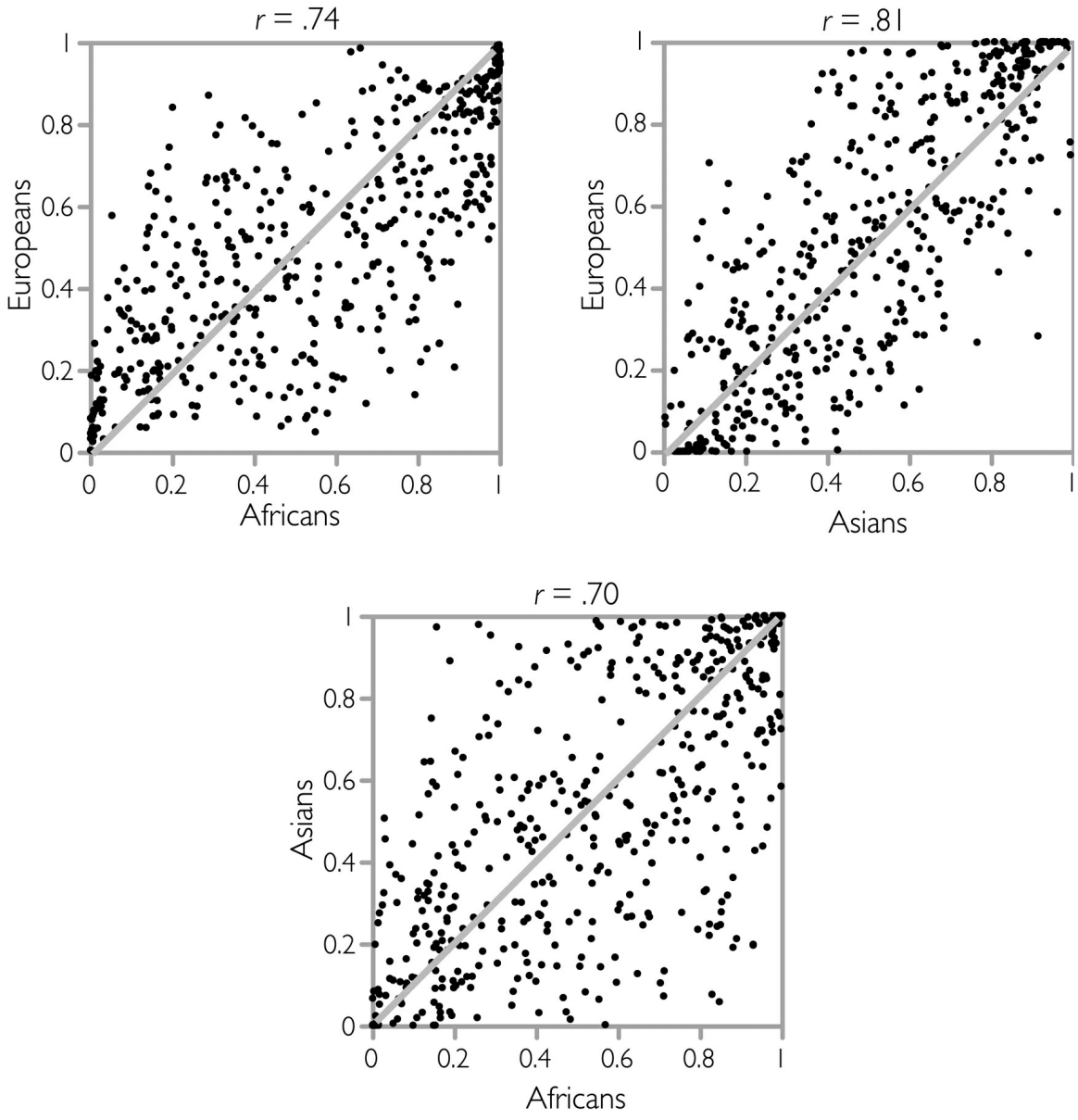
obscuring the pattern).[10]

The diagonal line identifies SNPs for which the target allele difference is zero. As you can see, the actual differences are closely bunched to either side of the diagonal on all three graphs.

Scatter plots like these imply extremely high correlations between the two sets of target allele frequencies, and indeed they are high: +.98 for the African and Asian pairs; +.97 for the European pair. These results are typical. Taking all of the unique SNPs for all traits that are part of both the GWAS Catalog and 1000 Genomes—a sample of 43,543 SNPs—the average correlations were +.98 for the three African pairs, +.98 for the six European pairs, and +.99 for the three Asian pairs.[11]

Comparing Continental Populations

Now look what happens when we repeat the exercise, but comparing Africans with Asians, Asians with Europeans, and Europeans with Africans.



Source: Author's analysis, GWAS Catalog and Phase 1 of the 1000 Genomes Project.

The landscape is completely different. The cross-continent correlations are all high by the standards of social science, but even correlations of +.70, +.81, and +.74 (which are the ones represented in the figure) are associated with large differences between target allele frequencies. Why did I choose schizophrenia for the example? Because the three correlations for schizophrenia are nearly the same as the correlations for all SNPs related to cognitive repertoires in the GWAS Catalog: +.71 for Africans and Asians,

+.76 for Africans and Europeans, and +.81 for Asians and Europeans. The schizophrenia example is typical, not extreme.

And that's the nut of what I am trying to convey with Proposition #7. We don't know what these differences mean yet (with a few exceptions to be taken up later), but the image fostered by "race is a social construct" does not apply. The raw material for investigating genetic sources of population differences in phenotypic traits consists of differences in target allele frequencies. For subpopulations within continents, the raw material is meager. For continental populations, the raw material is abundant.

An Operational Definition of "Large"

To demonstrate that abundance, I need a summary statistic for conveying how many SNPs fall far from the diagonal in the scatter plots. I settled on an operational definition of "large" that defines "large" relative to differences within continental subpopulations: *A difference in target allele frequencies is called "large" if it is bigger than 99 percent of the target allele frequency differences found within continental subpopulations.* To calculate that number, I used all 43,543 unique SNPs in the GWAS Catalog that are also found in Phase 1 of the 1000 Genomes Project. Combining all of the 12 pairs of within-continent subpopulations produced a sample of 522,516 pairs of target allele frequencies. Twenty percent of the absolute differences in target allele frequencies were less than .01, 63 percent were less than .05, and 88 percent were less than .10.[12] Ninety-nine percent were less than .19—to be more precise, less than .186. Thus my operational definition says that the smallest between-continent difference that is "large" is anything greater than .186. For convenience, I will round up and use .20 as the criterion. It's easier to remember.

In other words, if Asians have a target allele frequency of .45 for a certain SNP and Europeans have a target allele frequency of .65 or higher on the same SNP, that difference qualifies as "large." If Europeans have a target allele frequency of .25 or less, that difference also qualifies as "large." What's important is the absolute difference between two populations.

How many SNPs show that large a difference? The following table shows the results for 112 phenotypic traits grouped into three types of

noncognitive traits and three types of cognitive traits. The noncognitive traits are major diseases such as breast cancer and Parkinson's disease, physiological biomarkers such as height and weight, and blood parameters such as red cell count and metabolite levels. The cognitive traits are cognitive disorders such as depression, cognitive ability (both IQ and neurocognitive functioning), and personality features such as risk-taking tolerance and life satisfaction. The note gives details.[13]

TARGET ALLELE DIFFERENCES QUALIFYING AS "LARGE" (.20+)

Physiological Traits

No. of Unique SNPs: **13,431**
Total: **33%**
African-Asian: **37%**
European-African: **33%**
Asian-European: **30%**

Diseases

No. of Unique SNPs: 3,718
Total: 33%
African-Asian: 38%
European-African: 33%
Asian-European: 30%

Biomarkers

No. of Unique SNPs: 5,298
Total: 35%
African-Asian: 39%
European-African: 35%
Asian-European: 31%

Blood parameters

No. of Unique SNPs: 4,415
Total: 31%
African-Asian: 35%
European-African: 31%

Asian-European: 28%

Cognitive Traits

No. of Unique SNPs: **9,628**
Total: **36%**
African-Asian: **39%**
European-African: **37%**
Asian-European: **32%**

Cognitive disorders

No. of Unique SNPs: 2,594
Total: 35%
African-Asian: 38%
European-African: 37%
Asian-European: 31%

Mental abilities

No. of Unique SNPs: 5,715
Total: 36%
African-Asian: 39%
European-African: 36%
Asian-European: 32%

Personality features

No. of Unique SNPs: 1,319
Total: 38%
African-Asian: 42%
European-African: 38%
Asian-European: 35%

Source: Author's analysis, GWAS Catalog and Phase 1 of the 1000 Genomes Project.

When comparing the three continental populations, about a third of all target allele differences are at least .20.[14] Note that .20 is the smallest difference that qualifies. The mean difference among those that qualify is .33 for both the physiological and cognitive traits.

The results for this subset of traits generalizes to all 2,147 traits in the GWAS Catalog as of May 2019 that also had SNPs represented in Phase 1 of

the 1000 Genome Project. For the combined noncognitive traits, 32 percent of target allele differences across continental populations qualified as large. For the combined cognitive traits, 34 percent qualified as large.

One other feature of the results generalizes: The three continental pairs are consistently ordered. Africans and Asians have the highest proportion of large differences, Asians and Europeans have the smallest proportion, and Africans and Europeans are in between. This is consistent with the theoretically expected relationship between geographic and genetic differences between populations discussed in chapter 7.

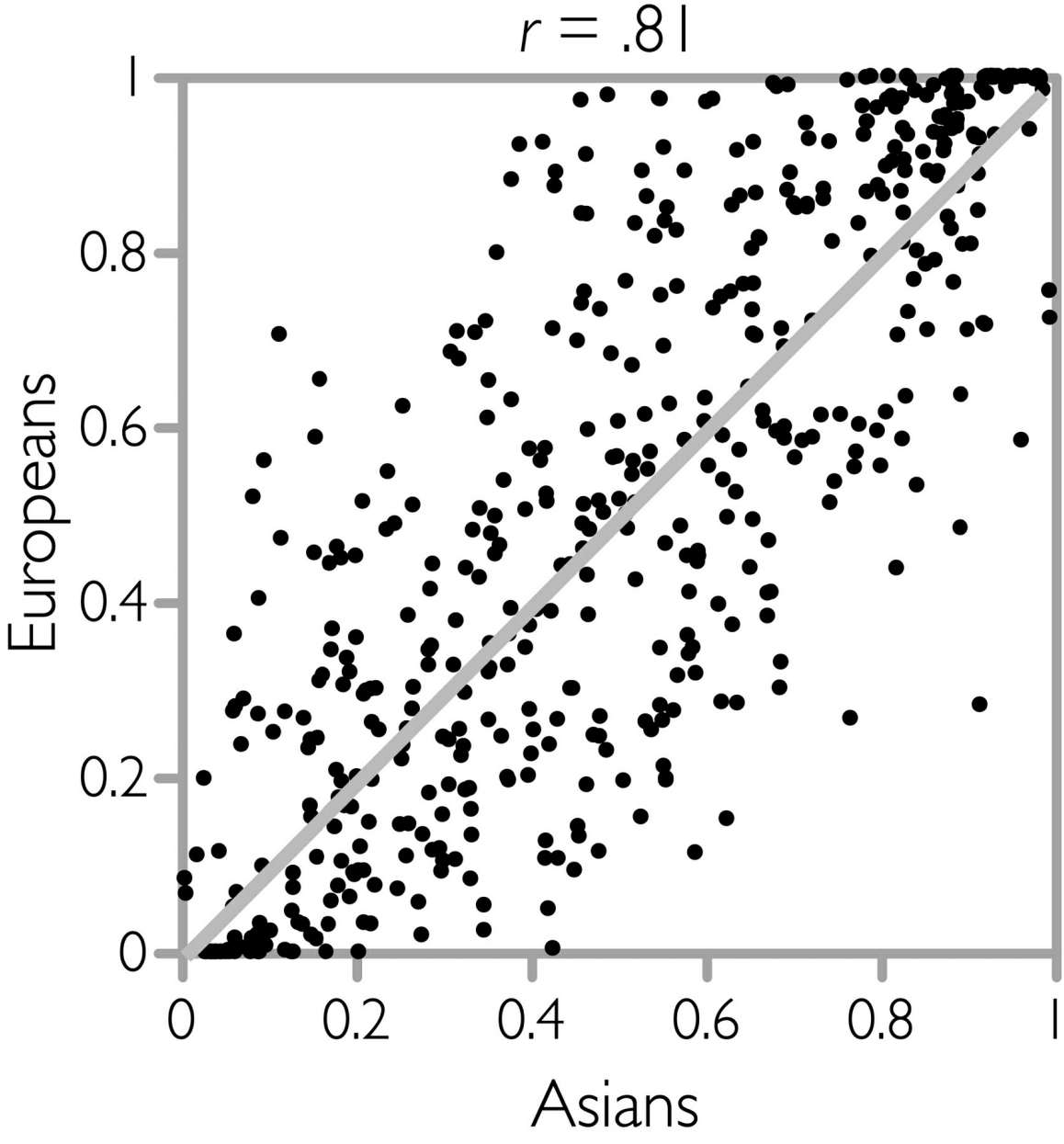
The Traits Related to Cognitive Repertoires

The table here presents information on 22 traits related to personality, abilities, and social behavior that have at least 100 SNPs associated with them. Unlike the previous table, this one combines SNPs that are given different labels in the GWAS Catalog but are associated with the same trait. For example, the trait labeled "well-being" in the table combines SNPs from studies in the GWAS Catalog that were for traits labeled "eudaimonic well-being" and "subjective well-being." The note gives additional information about the traits.[15]

The table provides more information than most readers need. Its purpose is to enable skeptical readers to look at the results from a variety of perspectives. Suppose, for example you think that an absolute difference of .20 is not sufficiently big. The table also shows you the percentage of allele differences that met a threshold of .25, which exceeds 99.9 percent of the within-continent differences. It also shows you the between-continent correlations and the mean allele difference for those traits that met the .20 threshold. Proposition #7 claims that "Continental population differences in variants associated with personality, abilities, and social behavior are common." In effect, the table says that the data confirm that proposition no matter how you look at them.

The Inevitability of Interesting Questions to Ask

Even though we don't know what analyses of these data will show, the existence of so many differences in target allele frequencies will raise interesting questions for a simple reason: We already know that the target alleles for two populations seldom balance out. Look again at the Asian-European scatter plot for schizophrenia.



Source: Author's analysis, GWAS Catalog and Phase 1 of the 1000 Genomes Project.

DIFFERENCES IN TARGET ALLELE FREQUENCIES FOR TRAITS RELATED TO COGNITIVE REPERTOIRES

		Correlations of Target Allele Frequencies						SNPs Meeting the Two Thresholds						Mean Difference in Target Allele Frequencies					
		Within Continents			Across Continents			Africa-Asia		Europe-Africa		Asia-Europe		All SNPs*			SNPs Meeting the .20 Threshold**		
Trait	Unique SNPs	Africa	Asia	Europe	Africa-Asia	Europe-Africa	Asia-Europe	>=.20	>=.25	>=.20	>=.25	>=.20	>=.25	Africa-Asia	Europe-Africa	Asia-Europe	Africa-Asia	Europe-Africa	Asia-Europe
Cognitive Performance																			
Cognitive ability	1,926	0.98	0.98	0.96	0.66	0.72	0.78	42%	33%	39%	27%	33%	22%	0.18	0.16	0.20	0.37	0.31	0.30
Highest math course	1,345	0.98	0.98	0.96	0.68	0.73	0.76	39%	31%	36%	26%	32%	21%	0.20	0.17	0.17	0.36	0.34	0.31
Educational attainment	2,368	0.98	0.98	0.97	0.71	0.78	0.78	37%	29%	34%	22%	31%	21%	0.16	0.16	0.18	0.36	0.31	0.31
Self-reported math ability	1,012	0.98	0.98	0.97	0.70	0.75	0.78	37%	29%	35%	25%	32%	21%	0.19	0.17	0.16	0.37	0.32	0.31
Neurocognitive function	134	0.98	0.99	0.98	0.78	0.78	0.84	28%	20%	26%	17%	22%	13%	0.15	0.15	0.13	0.34	0.35	0.31
Personality/Temperament																			
Depression	850	0.98	0.98	0.96	0.73	0.73	0.79	37%	27%	37%	26%	33%	22%	0.17	0.16	0.18	0.34	0.31	0.30
Neuroticism	932	0.98	0.99	0.96	0.66	0.70	0.77	39%	31%	40%	27%	35%	25%	0.18	0.17	0.20	0.38	0.32	0.32
Worry	226	0.98	0.98	0.93	0.49	0.63	0.74	40%	33%	44%	30%	34%	22%	0.20	0.16	0.22	0.43	0.32	0.30
Risk tolerance	420	0.98	0.98	0.96	0.63	0.73	0.74	44%	34%	39%	30%	33%	23%	0.18	0.17	0.21	0.38	0.33	0.32
Adventurousness	162	0.98	0.98	0.97	0.65	0.76	0.78	42%	33%	36%	27%	30%	21%	0.18	0.16	0.21	0.38	0.32	0.32
Positive affect	191	0.97	0.98	0.95	0.69	0.69	0.74	35%	28%	35%	23%	32%	21%	0.17	0.16	0.18	0.36	0.31	0.31
Life satisfaction	176	0.97	0.98	0.95	0.66	0.69	0.72	45%	36%	38%	24%	39%	26%	0.17	0.17	0.20	0.34	0.30	0.31
Well-being	408	0.98	0.98	0.96	0.68	0.71	0.73	38%	29%	36%	26%	36%	25%	0.17	0.17	0.19	0.37	0.33	0.32
Cognitive Disorder																			
Schizophrenia	962	0.98	0.98	0.97	0.70	0.74	0.81	39%	29%	35%	27%	29%	21%	0.17	0.15	0.19	0.36	0.34	0.32
Bipolar disorder	134	0.98	0.97	0.97	0.73	0.80	0.77	33%	25%	28%	17%	31%	22%	0.15	0.15	0.17	0.34	0.30	0.30
Autistic traits	215	0.98	0.98	0.97	0.80	0.82	0.87	24%	20%	23%	17%	18%	12%	0.14	0.11	0.15	0.35	0.32	0.31
ADHD	126	0.99	0.99	0.98	0.84	0.87	0.87	28%	18%	23%	14%	26%	12%	0.13	0.13	0.14	0.30	0.29	0.29
Social Behavior																			
Conduct disorder	147	0.98	0.98	0.97	0.75	0.80	0.80	31%	22%	30%	20%	27%	15%	0.15	0.13	0.16	0.35	0.31	0.30
Alcohol consumption	284	0.98	0.99	0.96	0.66	0.71	0.71	42%	33%	42%	31%	37%	25%	0.19	0.18	0.20	0.37	0.33	0.32
Alcohol dependence	124	0.97	0.98	0.98	0.73	0.78	0.79	44%	30%	39%	28%	29%	23%	0.17	0.15	0.19	0.33	0.31	0.34
Brain-related																			
Brain volumes	160	0.98	0.99	0.98	0.77	0.79	0.86	38%	30%	31%	26%	24%	17%	0.16	0.14	0.17	0.34	0.34	0.30
Cerebrospinal fluid	121	0.98	0.99	0.98	0.76	0.79	0.86	27%	21%	29%	21%	20%	16%	0.14	0.12	0.15	0.36	0.32	0.30

*Within-continent mean differences in target allele frequencies are not shown because they were all from .03 to .06 for all subpopulations and all traits.

**Mean within-continent differences that met the .20 threshold are not shown because there were so few of them.

If the investigator's ambition is to identify a role for natural selection in creating population differences, there's no telling whether anything interesting lies in that plot. Getting from raw differences in target allele frequencies to evidence of natural selection is a long and torturous process, and even then the results should be treated provisionally.[16] For that matter, proof of the role of natural selection for many genetic differences will remain unobtainable without methodological breakthroughs. Recall from chapter 8 that one of the most commonly used tools doesn't work for adaptations that occurred more than 30,000 years ago.

But while proving natural selection is difficult, the differences in target allele frequencies across populations can be analyzed without knowing what caused the differences. Such analyses can't be done now with any confidence because of the problems of population stratification, but they will become feasible within a few years when large databases from the major ancestral populations are available.

For purposes of illustration, let's jump ahead to that time and suppose that the schizophrenia scatter plot for Asians and Europeans is free of contamination by population stratification and that target allele frequencies and the weights associated with them can be taken at face value (very big suppositions). The 500 SNPs shown in the scatter plot do not reveal an obvious imbalance between the target allele frequencies above and below the diagonal, but it turns out a modest one does exist. In the full sample of 962 SNPs associated with schizophrenia in the GWAS Catalog, Asians have the higher target allele frequency for 513 SNPs compared to 449 for Europeans. This opens the possibility—only a possibility—that Asians are genetically more susceptible to schizophrenia than Europeans. Whether it is true depends on the magnitude of the differences in target allele frequencies, the effect sizes associated with the SNPs, and a variety of other considerations. Even if population stratification is no longer a problem, the raw difference is useful only for deciding whether it is worthwhile to curate the sample of SNPs to cleanse it of contaminating factors and to analyze polygenic scores for Europeans and Asians. Perhaps the imbalance of 513 to 449 in the raw data will turn out to be meaningful; perhaps it won't.

The imbalance of 513 to 449 for schizophrenia amounts to a 53:47 split per hundred SNPs. The table of 22 traits related to cognitive repertoires presented above has a total of 66 cross-continent pairs. The imbalance is at least 53:47 for 48 of those 66 pairs. It is 55:45 or greater for 33 of them. It is

60:40 or greater for 8 of them. No matter what (reasonable) criterion for a large-enough imbalance you might adopt, many imbalances qualify as large enough to warrant investigation. In time, they will in fact be investigated. It is implausible to expect that *none* of the imbalances will yield evidence of significant genetic differences related to phenotypic differences across continental populations.

The results will often be complex. The same SNPs that affect the trait under investigation will typically be correlated with many other traits as well, which may sometimes mean that SNPs beneficial for a desirable trait also increase vulnerability to undesirable ones, as in the case of the tradeoff between protection against malaria versus the risk of sickle cell anemia. Some analyses may reveal that different populations get to similar end points via different processes, as in the case of sex differences in cognitive toolboxes discussed in chapter 3. But simple or complex, the results are in my view bound to be interesting.

Whether my forecast is reasonable depends on the outcome of a larger debate about establishing genetic causation. One side of that debate holds that my optimism is dead wrong. The results cannot possibly be interesting, because causation cannot possibly be established even after the problems of population stratification have been solved. That debate is given a full discussion in chapter 14. It seems fair to say at least this: It is fundamentally wrong to think of the study of genetic population differences as an exercise in ranking populations from top to bottom. The questions to be explored are far more interesting, complex, and potentially more rewarding than filling out an ethnic scorecard.

WHAT ABOUT UNCOMMON AND RARE VARIANTS?

Most single nucleotide variants (SNVs) are found in fewer than 1 percent of chromosomes and therefore do not qualify as common SNPs. The proportion of SNVs with allele frequencies of less than .01 is currently estimated at 74 percent, but that's going to increase as more rare variants are discovered.[17] The closest to a complete inventory as I write is a 2016 sample based on 10,545 genomes that used deep sequencing techniques and identified 150 million variants in the human genome. This is by no means the total. The study reported that each individual

added to the sample contributed 8,579 variants not previously identified, leading the authors to estimate that a sample of 100,000 genomes would identify 500 million variants.[18] How much additional variance rare variants explain is still uncertain—a few articles have reported that they explain much of the missing heritability in GWAS analysis, but most analyses show minor effects.[19]

In addition to constituting the bulk of all variants, rare variants are also overwhelmingly confined to a single continental population, on the order of 90 percent or more.[20] However, the importance of rare variants to population differences is uncertain. By one line of argument, they should be minor. By definition, a rare SNV has not spread widely through a population. It is either a new mutation or one that has been only weakly selected if at all. Mutations are random events. They don't happen because there's a need for them (e.g., a mutation giving protection against a disease does not occur because the person was living where the disease was endemic). Thus there is no reason to believe that new mutations occurring in two separated populations will be systematically different with regard to their effects on a given trait. But the literature contains a variety of other perspectives on the role of rare variants.[[21]] The short story is that comparatively little is known about the role of rare variants, both generally and with regard to population differences. The action for now is with standing variation in common SNPs.

Known Genetic Continental Population Differences

Our expectations for the future should take into account that many genetic population differences are already established.

Hiding in Plain Sight

We have known for years that biologically complex differences in continental

populations have evolved since humans left Africa. It is an unlikely assertion on its face—how can "race is a social construct" continue to be the received elite wisdom if such differences are already known? But it's true. Two examples of significant genetic differences across populations have been sitting in plain sight for decades: lactase persistence and susceptibility to sickle cell anemia. Details on both of these adaptations are given in the note.[22] Both of these are major adaptations involving many biological systems. For that matter, lightening of skin pigmentation, passed off as trivial because it is only "skin deep," is genetically more complicated than "skin deep" implies.[23] Why, given these examples of complex adaptation that obviously occurred after the Africa exodus, should it ever have been assumed that they were the only ones?

Continental Differences Discovered Through Genome-Wide Analysis

Even though the documentation of continental differences has had a low priority among most genetics researchers, several have been found.

Susceptibility to inflammatory and immune-related diseases. In 2014, Jessica Brinkworth and Luis Barreiro examined the GWA results for three chronic inflammatory diseases (celiac disease, Crohn's disease, and ulcerative colitis) and five autoimmune diseases (type 1 diabetes, multiple sclerosis, rheumatoid arthritis, psoriasis, and systemic lupus erythematosus). They found evidence "that at least some of the present-day autoimmune risk loci have been adaptive and conferred some sort of functional benefit to Europeans in the past."[24] The authors hypothesized that a large sample of Africans would yield evidence that "the genetic determinants of susceptibility to chronic inflammatory and autoimmune diseases in individuals of African descent are distinct from those found among Europeans."[25] A 2016 study conducted by a large international team (first author was Yohann Nédélec) found evidence that differences in immune function arose from natural selection rather than genetic drift: "More specifically," the authors wrote,

"our results suggest that a significant fraction of population differences in transcriptional responses to infection are a direct consequence of local adaptation driven by regulatory variants."[26]

A study of psoriasis using samples of Europeans (called Caucasians in the study) and Chinese identified European-specific loci that had a cumulative effect that "could explain up to 82.83 percent of the prevalence difference of psoriasis between the Caucasian and the Chinese populations."[27] Overall, the authors concluded, "This study not only provides novel biological insights into the involvement of immune and keratinocyte development mechanism, but also demonstrates a complex and heterogeneous genetic architecture of psoriasis susceptibility across ethnic populations."[28]

Respiratory adaptation to high altitudes. Adaptation to high altitudes has occurred among peoples living on the Qinghai Plateau in Tibet, the Andean Altiplano in Peru, and the Semien Plateau in Ethiopia involving changes in pulmonary function, arterial oxygen saturation, hemoglobin concentration, and maternal physiology during pregnancy. The evolutionary routes taken by each population have involved different genes and produced different responses.[29] Resting ventilation among the Andeans is normal for humans in general; among Tibetans, it is 50 percent higher. Arterial oxygen saturation is elevated for Andeans and Ethiopians; not for Tibetans. Hemoglobin concentration is elevated among Andeans, shows a minimal increase among Ethiopians, and is actually lowered in Tibetans.[30] An exotic complication in the case of the Tibetans is that some of the mutations that helped adapt them to high altitude now appear to have come from introgression with the mysterious Denisovans.[31]

Genetic disorders among Ashkenazi Jews. As early as the 1880s, it was noted that Tay-Sachs disease occurred almost exclusively among Ashkenazi Jews. Over the years, several other genetic disorders have been found to be far more prevalent among Ashkenazi Jews than in any other population. The causes of the difference in prevalence are still unresolved. One possibility is a population bottleneck around a thousand years ago, as argued in a 2018 study that analyzed 5,685 Ashkenazi Jewish exomes. The alleles in question included ones for Tay-Sachs.[32]

Another possibility is that natural selection has been at work. In 2009, before access to GWA, Gregory Cochran and Henry Harpending argued that

case, observing that the Jewish genetic disorders are oddly grouped:

> Imagine a fat biochemistry textbook, where each page describes a different function or condition in human biochemistry. Most of the Ashkenazi diseases would be described on just two of those pages. The two most important genetic disease clusters among the Ashkenazim are the sphingoloid storage disorders (Tay-Sachs disease; Goucher's disease; Niemann-Pick disease; and mucolipidosis, type IV) and the disorders of DNA repair (BRCA1 and BRCA2; Fanconi anemia, type C; and Bloom syndrome).[33]

If a population bottleneck were the sole explanation, they calculated that the odds of finding four disorders that affect sphingolipid metabolism would have been about 1 in 100,000.[[34]] The authors concluded instead that we are looking at recently evolved differences across populations. While the explanation remains unclear, this much is undisputed: The disorders are genetic, and so are population differences separating Ashkenazi Jews from everyone else.

Prostate cancer. In 2018, a team of geneticists (first author was Joseph Lachance) studied the genetic sources of the differential rates of prostate cancer in Europeans and Africans. They used SNPs from the GWAS Catalog, Phase 3 of 1000 Genomes, and the large database of African genomes assembled by Sarah Tishkoff of the University of Pennsylvania. They found that a small proportion of SNPs with large target allele frequency differences and large effect sizes make a disproportionate contribution to population differences in the risk of prostate cancer. "Both neutral and selective evolutionary mechanisms appear to have contributed to disparities in the genetic risk of CaP. These mechanisms include founder effects due to the out-of-Africa migration and genetic hitchhiking of disease susceptibility alleles with locally adaptive alleles."[35]

Evidence of natural selection in height, schizophrenia, and body mass index. A team of geneticists (first author was Jing Guo) examined height, body mass index, waist-hip ratio adjusted for BMI, HDL cholesterol, LDL cholesterol, coronary artery disease, type 2 diabetes, Alzheimer's disease, schizophrenia, and educational attainment for evidence of natural selection. The Guo study tapped a variety of databases instead of limiting itself to the GWAS Catalog. The authors found evidence that SNPs associated with

height, schizophrenia, and waist-to-hip ratio have undergone natural selection.[36] They did not find such evidence for the other seven traits in the study.

Blood pressure. A study by a team of Japanese geneticists (first author was Fumihiko Takeuchi) used Europeans and East Asian samples to study continental population differences in blood pressure. They found evidence for two remarkable phenomena: "(1) the colocalization of distinct ancestry-specific variants that are not rare and can exert mutually inverted genetic effects between the ethnic groups and (2) the potential involvement of natural selection in the occurrence of ancestry-specific association signals."[37] They argued that "we have discovered a new model in which genetic effects for transethnic SNPs that form a shared haplotype at a locus are driven by causal variants that are ancestry-specific but are not rare, which can be called a common ancestry-specific variant association model."[38]

And more. Greenlandic Inuits are genetically adapted to a marine diet rich in omega-3 polyunsaturated fatty acids, increasing fitness in a cold and dark environment.[39] The population of San Antonio de los Cobres in Argentina has adapted to high levels of arsenic in the groundwater through positive selection on SNPs involved in the arsenic methylation pathway.[40]

It's early days yet, but the results of the limited genome-wide analyses of differences in continental populations to date point in the same direction: Many continental population differences are out there.

Recapitulation

The story of the raw material for studying continental population differences applies to SNPs related to physiological parameters, diseases, and cognitive repertoires. Substantial between-continent differences in target allele frequencies are common. Around a third of all differences meet a plausible definition of "large." The limited amount of sophisticated genetic analysis of between-continent differences done to date suggests that these extensive differences observed in the raw material will frequently yield productive

results about genuine continental population differences.

A Personal Interpretation of the Material in Part II

Part II has described a parallel universe. In the universe inhabited by the elite media and orthodox academia, it has been settled for decades that race is a social construct. In that universe, the lessons taught by Richard Lewontin and Stephen Jay Gould back in the 1970s and early 1980s still apply.

In the universe inhabited by geneticists who study human populations, the 1990s saw glimpses of a new perspective, and the new century opened up fascinating stories that had previously been closed.

The new understandings about the peopling of the Earth have been the most dramatic. New roles in the evolution of *Homo sapiens* were discovered for Neanderthals and previously unknown hominins. Access to ancient DNA enabled the reconstruction of successive human migrations across Eurasia that have revolutionized our knowledge of prehistory.

The understanding of recent evolution that prevailed as recently as the 1990s has also been overturned. Human evolution does not always proceed at a glacial pace dictated by random mutations. Sometimes changes in standing variation can occur quickly in response to environmental selection pressures. Those environmental pressures have typically been confined to populations in specific geographic areas.

Most recently, the task of assembling the genetic story for specific phenotypic traits has begun. It is still in its early stages, but progress is accelerating nonlinearly. Hence the nervousness that has prevented open discussion of what's going on in the geneticists' parallel universe: the fear that we will discover scary population differences in what I have called cognitive repertoires.

That fear accounts for the taboo that has been attached to discussions of genetics and race. It's no wonder. White Americans' justified guilt about their history of discrimination against blacks, native Americans, and immigrants from Latin America and East Asia gives them reason to worry

that white supremacists will use genetics to rationalize that history.

Let me suggest an alternative way of thinking about ethnic differences. Many of the people in elite circles who honor the taboo are also cosmopolitan. They have had professional colleagues of many ethnicities and have traveled extensively, observing the endless variety of ways in which people in different cultures think and behave. They have no trouble believing from personal experience that Chinese think and behave somewhat differently from Saudi Arabians. So do Saudi Arabians and Senegalese, Senegalese and Norwegians, Norwegians and Italians, northern Italians and southern Italians. Viewed from that perspective, ethnic differences in cognitive repertoires are neither to be doubted nor feared. They exist, and everyone who has seen anything of the world knows it. The mix of nature and nurture? That's not the issue. The differences themselves are facts. People around the world are similar in the basics and different in the details. We connect through the basics. We live with and often enjoy the differences.

The material in Part II does not foreshadow discovery of genetically-grounded population differences in the basics. Rather, I hope I have persuaded you that genetically-grounded differences in the details are to be expected. Some of these genetic differences may consist of alternative routes for getting to similar ends, just as has been found with many cognitive sex differences. Many others will be differences that are neither better nor worse, but just differences. Probably some will lend themselves to value judgments, but even those will cut both ways. No population is free of defects nor possessed of all the virtues.

We can expect most of the genetic differences to range from small to moderate and to explain just a portion of the phenotypic differences we already live with. Every population will be represented from one extreme to the other on every trait. There will be no moral or legal justification for treating individuals differently because of the population to which they belong.

I doubt that these assurances will do much good. The prospect of genetic differences across ancestral populations is still too sensitive for calm discussion. But perhaps this will provide perspective:

We already know of a genetically-grounded population difference on a highly sensitive trait that is far, far larger than any ancestral population difference we are going to find. The populations in question are males and females. The highly sensitive trait is the commission of physical violence

against other humans. The undoubted genetic source of the difference is the Y chromosome. How big is the difference? Judge it by this: About 90 percent of all homicides are committed by males.[41]

If we can live with a population difference that huge on such an important behavioral trait, we can easily live with the smaller differences in continental populations that are likely to be found. The differences that will be documented during the coming years should be greeted with "That's interesting." I fear that the orthodoxy's insistence that population differences in cognitive repertoires *cannot* exist ensures that they initially won't be greeted that way.[42] But they should be.