REVIEW ARTICLE

GENOMIC MEDICINE
W. Gregory Feero, M.D., Ph.D., and Alan E. Guttmacher, M.D., *Editors*

Ancestry and Disease in the Age of Genomic Medicine

Charles N. Rotimi, Ph.D., and Lynn B. Jorde, Ph.D.

Human genetic data are accumulating at an ever-increasing pace, and whole genome sequences of individuals from multiple populations are now publicly available.[1-3] The growing inventory of human genetic variation is facilitating an understanding of why susceptibility to common diseases varies among individuals and populations. In addition, we are gaining insights that may improve the efficacy and safety of therapeutic drugs. Such knowledge is relevant to fundamental questions about our origins, differences, and similarities. Here, we provide a brief review of the current knowledge of human genetic variation and how it contributes to our understanding of human evolutionary history, group identity, and health disparities.

From the Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD (C.N.R.); and the Department of Human Genetics, University of Utah Health Sciences Center, Salt Lake City (L.B.J.). Address reprint requests to Dr. Rotimi at the National Institutes of Health Center for Research on Genomics and Global Health, 12 South Dr., MSC 5635, Bethesda, MD 20851-5635, or at rotimic@mail.nih.gov.

N Engl J Med 2010;363:1551-8.
Copyright © 2010 Massachusetts Medical Society.

Genetic Variation in Individuals and Populations

Human genetic variation contributes substantially to the physical variation we observe among individuals. The two most important components of human genetic variation are single-nucleotide polymorphisms (SNPs) and copy-number variants. SNPs are single DNA nucleotide base pairs that differ among individual DNA sequences. (For example, one individual will have an A-T at a specific position in the haploid DNA sequence, whereas another will have a C-G.) Copy-number variants are larger contiguous blocks of DNA sequence (usually exceeding 1000 bp) that vary in copy number among individuals. (For example, a block of DNA could be duplicated in one person but deleted in another.) Millions of SNPs have been discovered in humans, and at the single-nucleotide level any two humans differ at about 1 in 1000 bp.[4] Copy-number variants are less common, but because they comprise much larger segments of DNA, they contribute an additional 0.4% difference in DNA sequence between any two individuals.[5,6] Any two humans would be about 99.5% identical in DNA sequence, reflecting a relatively recent, common origin of our species.

Humans have about half the amount of genetic variation that is observed in central African chimpanzees[7] and gorillas[8] and about one tenth the amount of variation seen in the fruit fly *Drosophila pseudoobscura*.[9] The fact that we have less variation than many gorilla and chimpanzee populations, despite our much larger current population size, reflects prehistoric bottlenecks in human population size.[10] Some mammals, such as cheetahs, have also undergone population bottlenecks and have even less genetic diversity than humans.[11]

Most common SNPs (i.e., those for which the prevalence of the less common allele is greater than 5%) are shared among populations from different continents.[12] This commonality reflects continued migration and gene flow among human populations throughout history, in addition to our recent common origin.

The New England Journal of Medicine
Downloaded from nejm.org at SEOUL NATIONAL UNIVERSITY on May 13, 2012. For personal use only. No other uses without permission.
Copyright © 2010 Massachusetts Medical Society. All rights reserved.

Many studies have shown that the great majority of genetic variation (around 85 to 90%) can be found within any human population (e.g., samples of persons from Great Britain and from Ghana); only an additional 10 to 15% of variation is gained when considering the entire human population.[13,14] This is another measure of the genetic similarity among persons of different ancestries.

Occasionally, a SNP or a copy-number variant is relatively common in one population but absent (or nearly so) in another. This is sometimes due to the recent emergence of a variant that has not yet had time to spread to other populations — such as a SNP that causes hereditary hemochromatosis and is common in Europe but very rare elsewhere.[15] In other cases, the difference in prevalence is caused by natural selection in a specific local environment. For example, hereditary lactase persistence is unusual in most human populations, but it is prevalent in European and African pastoral populations, where milk consumption beyond childhood has had a selective advantage.[16] Convincing evidence of positive selection has also been seen in genes that affect skin pigmentation,[17] resistance to malaria,[18,19] and resistance to Lassa fever.[20]

In comparisons of the prevalence of a SNP or copy-number variant among populations, a strong correlation between geographic location and genetic similarity is generally observed.[12,13,21] This is hardly surprising, because populations that are geographically close to one another are likely to share a common ancestry and to exchange migrants. Population-level studies have also shown that African populations have relatively greater genetic diversity than other populations and that the variation found outside Africa tends to be a subgroup of African genetic variation.[22] These patterns, which are observed even when comparing a relatively small number of SNPs or copy-number variants (100 or so), support a common African origin of our species.[23]

If larger numbers of polymorphisms are analyzed, individual ancestry can be approximately inferred. Early studies succeeded in correctly classifying most sampled individuals into their continent of origin.[24-26] More recently, by analyzing up to 1 million SNPs at a time, investigators have been able to infer individual ancestry more precisely, often assigning ancestry to specific geographic regions.[12,27,28] Because the prevalence of most SNPs differs by only a little among populations, fine-scaled inferences about ancestry are possible only when large numbers of SNPs are assayed or when SNPs that are already known to show large frequency differences among populations are selected.[29] Few if any SNP variants are present in all members of one population but absent in all members of another. Thus, it is possible to make approximate conclusions about an individual's ancestry from a panel of SNPs (with some cautions[30-32]), but it is not possible to determine with any certainty an individual SNP genotype on the basis of one's ancestry.[33]

Although individual ancestry can be inferred approximately from genetic data, our growing understanding of these data shows that population categories are not discrete, readily discernible entities. Many populations, such as African Americans and Hispanic Americans, have complex recent ancestral histories. African Americans, on average, are estimated to have approximately 20% European ancestry,[34] but this proportion varies substantially among different African-American populations within North America.[35-37] More important, genetic analyses of individual ancestry show that some self-identified African Americans have large proportions (more than 50%) of European genetic ancestry, whereas some self-identified European Americans have substantial recent African genetic ancestry.[37,38]

Moreover, perception of population differences partly depends on the human populations that have been genetically sampled. The current extent of population sampling is, at best, incomplete. The greater the number of human populations genotyped and the smaller the average level of between-population variation,[39] the more difficult it is to delineate population boundaries. Definitions of populations or ethnic groups are complex and sometimes overlap or are mutually contradictory, incorporating cultural, linguistic, biologic, and geopolitical factors.[40] These definitions vary within countries, as they have throughout time, with racial or ethnic categories having changed every 10 years in the U.S. Census.[41]

Population categories, including race and ethnic group, are clearly inadequate to describe fully the pattern and range of variation among individuals. A more accurate assessment of disease risk may be obtained by genotyping disease-associated genetic variants in individuals, rather than using population affiliation as a surrogate.

The New England Journal of Medicine
Downloaded from nejm.org at SEOUL NATIONAL UNIVERSITY on May 13, 2012. For personal use only. No other uses without permission.
Copyright © 2010 Massachusetts Medical Society. All rights reserved.

GENETIC VARIATION AND DISEASE DISPARITIES

Disease prevalence, severity, and resistance vary considerably among ethnic groups as a consequence of inherited factors and noninherited causes, such as poverty, unequal access to care, lifestyle, and health-related cultural practices.[42-44] Well-known examples include the elevated prevalence of Tay–Sachs disease among persons of Ashkenazi Jewish ancestry[44] and sickle cell disease, thalassemia, and glucose-6-phosphate dehydrogenase deficiency in some populations of African descent.[18,44-46] However, Tay–Sachs disease is observed in non-Jewish populations and has a relatively high prevalence in parts of French-speaking Canada.[47] Similarly, hemoglobinopathies are found in many global populations.[44] Labeling of these diseases as occurring only in black populations can reduce the likelihood of diagnosis in other populations.

Because diseases vary in prevalence among populations, it is to be expected that the frequencies of genetic variants that contribute to their causation also vary. A comparison of the prevalence of risk variants associated with susceptibility to 26 diseases in 11 populations from across the world in the International HapMap Project showed significant differences in the prevalence of such variants among populations.[48] Some variants showed a difference in prevalence by factors of 20 to 40 (Table 1). This variation is important in the context of genomewide association studies, in which the SNPs in the genomes of thousands of affected case subjects and unaffected control subjects are genotyped, with the aim of identifying SNPs that are associated with disease — and ultimately, identifying variants that directly confer susceptibility to disease. Such studies have identified more than 800 unique SNP-trait associations, significant at $P<5\times10^{-8}$, that contribute to common diseases, such as age-related macular degeneration, diabetes, and Crohn's disease.[49] However, nearly 90% of all genomewide association studies have been carried out in populations of European ancestry. Remarkably, only one genomewide association study (the MalariaGen Project) has involved a large collection of samples from African subjects,[50] and only a handful have involved African Americans and other groups of non-European ancestry.[51]

Table 1. Median and Range of Minor-Allele Frequency for Loci Identified by Genomewide Association Studies in 11 HapMap Populations Worldwide.*

Disease	No. of Loci	Minor-Allele Frequency	
		Median	Range†
Type 2 diabetes	41	0.365	0.099–0.564
Crohn's disease	45	0.381	0.137–0.769
Systemic lupus erythematosus	21	0.338	0.106–0.896
Psoriasis	15	0.336	0.096–0.681
Breast cancer	17	0.243	0.120–0.707
Prostate cancer	25	0.362	0.144–0.915

* The 11 HapMap global populations were persons of African ancestry in the Southwest region of the United States; Utah residents with northern and western European ancestry; Han Chinese in Beijing; Chinese in metropolitan Denver; Gujarati Indians in Houston; Japanese in Tokyo; Luhya in Webuye, Kenya; persons of Mexican ancestry in Los Angeles; Masai in Kinyawa, Kenya; Tuscans in Italy; and Yoruba in Ibadan, Nigeria.

† Data show considerable differences in population frequency of loci for the six selected diseases.

Clearly, the catalogue of worldwide variation in disease-causing SNPs is incomplete, and it will be important to conduct genomewide association studies in many additional populations. Such studies may be particularly informative in African populations because of their complex history and high level of genetic diversity. Because of this diversity, most available SNP microarrays may not adequately capture disease-related variation, as shown in two studies describing the complete sequences of the genomes of two groups (Khoisan and Bantu) in southern Africa.[3] It has been estimated that a genomewide association study of an African population would require approximately 1.5 million SNPs to achieve the same resolution as a study of a European population using 0.6 million SNPs.[50,52] Furthermore, high levels of diversity within and among African populations increase the likelihood of false positive associations, owing to population stratification.[50,52] Perhaps more important, heterogeneity in haplotype structure among various African ethnic groups is likely to lead to reduced statistical power to detect true positive signals by means of a genomewide association study, especially in multisite projects such as MalariaGen.[50]

Whole-genome sequence analysis avoids the difficulties associated with SNP ascertainment in specific populations. Whole-genome sequences will reveal collections of rare variants that may account for substantial disease susceptibility[53] but are more likely to be population-specif-

The New England Journal of Medicine
Downloaded from nejm.org at SEOUL NATIONAL UNIVERSITY on May 13, 2012. For personal use only. No other uses without permission.
Copyright © 2010 Massachusetts Medical Society. All rights reserved.

ic. This further emphasizes the need for whole-genome data collection in a wide variety of human populations, an effort that has been initiated as part of the ongoing 1000 Genomes Project.[54]

MAPPING OF COMMON DISEASE GENES

The use of ancestry in mapping of genes that contribute to common diseases, such as prostate cancer, chronic kidney disease, and hepatitis C virus (HCV) infection, illustrates how genomic science may inform our understanding of the causes of disease and health disparities. It is well established that Americans of African ancestry have a significantly higher risk of chronic kidney disease and kidney failure than do European Americans.[55,56] An increase in the risk of kidney failure[56] by a factor of 3.6 can be attributed in part to socioeconomic factors, such as inadequate access to health care, and clinical factors, such as a higher incidence of hypertension and diabetes (which may, in turn, be influenced by socioeconomic factors). Even after adjustment for these factors, however, substantial risk remains,[57,58] suggesting a genetic influence. Two independent studies have tested this hypothesis with the use of a gene-mapping technique that incorporates genetic admixture in African Americans. Multiple common SNPs in the gene that encodes nonmuscle myosin heavy chain type II isoform A (*MYH9*) were associated with an increase in the risk of focal segmental glomerulosclerosis and end-stage renal disease by a factor of 2 to 4 in persons without diabetes.[58-60] Genetic variation in *MYH9* accounted for a large proportion of the excess risk of both disorders that was observed in African Americans, as compared with European Americans.[58-60]

Prostate cancer is a second example of a common disease that varies considerably among populations; its prevalence in the United States is greatest among African-American men.[61,62] The rate of death from prostate cancer is increased by a factor of 2.4 among African-American men, as compared with European-American men.[61] Again, disparities in treatment and access to adequate health care contribute to this difference,[62] but there is increasing evidence that genetic variation makes a significant contribution. Several independent studies have identified and replicated a locus on chromosome 8q24 that is associated with prostate cancer in men from different ancestral backgrounds, including European, African, Latino, and Japanese men.[63-67] The effect of SNPs associated with prostate cancer on chromosome 8q24 varies according to ancestry, with population attributable risks ranging from 32% among European Americans to 46% among Hispanic Americans and 68% among African Americans.[67] (The population attributable risk is defined as the proportion of cases that would not occur if the risk factor were eliminated.) These SNPs, if shown in follow-up studies to be causal, may explain a substantial proportion of the observed disparity in prostate-cancer incidence.

On the one hand, these results provide compelling reasons to use information about individual and group ancestry in investigations of the causes of disease. On the other hand, continental ancestry, race, and ethnic group may be inadequate descriptors for the reasons we have discussed. We present two examples in which direct genetic testing may provide a better and more accurate alternative than such traditional descriptors.

The first example concerns population screening for the HLA-B*5701 allele to prevent the abacavir hypersensitivity syndrome (AHS). About 4 to 8% of patients with human immunodeficiency virus infection who are treated with abacavir have serious, potentially life-threatening AHS.[68,69] A double-blind, prospective, randomized study involving 1956 patients from 19 countries showed that screening for the HLA-B*5701 allele substantially reduced the incidence of AHS.[70,71] However, there is considerable interpopulation variation in the frequency of this allele, ranging from 0 to approximately 20% (Fig. 1).[72] As a result, the prevalence of AHS varies considerably, a variation that has implications for the utility and cost-effectiveness of screening in different population groups.[68,73] For example, with an overall AHS incidence of 5 to 8% in populations of European ancestry, only about 14 European patients would have to be screened, on average, to prevent one case of AHS. The incidences of AHS and HLA-B*5701 are much lower in some other populations, leading some observers to recommend screening for HLA-B*5701 in European populations but not in some Asian or African populations.[74,75]

The New England Journal of Medicine
Downloaded from nejm.org at SEOUL NATIONAL UNIVERSITY on May 13, 2012. For personal use only. No other uses without permission.
Copyright © 2010 Massachusetts Medical Society. All rights reserved.

A better understanding of the global distribution of the HLA-B*5701 variant leads to the conclusion that such recommendations may be problematic. Among five U.S. ethnic groups that participated in the International HapMap Project, the highest prevalence of HLA-B*5701 (17.6%) was seen not among Americans of European ancestry but among Gujarati Indians in Houston (Fig. 1). A similar estimate of 20% was observed in a sample of persons from India.[72] Substantial differences in allele prevalence were seen in two African HapMap populations: the prevalence in the Kenyan Masai group was 13.6%, a proportion that is more than double that in European samples (3.4% among Tuscans and 5.8% among residents of Utah with European ancestry); the frequency of the allele was zero among the Yoruba in Nigeria. This range in prevalence in the samples from African populations overlaps the ranges in Europeans and for the most part those in other groups. Labels such as "black" or "African" therefore obscure biomedically relevant variation and could lead to less vigilance among physicians for AHS in susceptible patients. For these reasons, the Food and Drug Administration's recommendation to screen all groups, regardless of population identity, for HLA-B*5701 before abacavir administration is prudent and justified.[76]

A second example is provided by population differences in the spontaneous clearance of HCV infection and the response to treatment of the virus. HCV infection may spontaneously resolve (viral clearance), persist without complications, or cause cirrhosis and hepatocellular carcinoma.[77,78] Spontaneous clearance of the infection has been reported to occur in 36.4% of persons of non-African ancestry, as compared with only 9.3% in persons of African ancestry.[78] A genomewide association study of persons chronically infected with hepatitis C and treated with pegylated interferon (peginterferon) alfa and ribavirin showed a strong association between SNP rs12979860 on chromosome 19q13 and a sustained virologic response.[79] This SNP, close to the gene encoding interferon lambda-3 (*IL28B*), accounted for more than a doubling in the difference in response to drug treatment for hepatitis, and the effect was similar in patients with European ancestry and in those with African ancestry.[79] A follow-up study showed that the C/C genotype of SNP rs12979860 was strongly associated with resolution of HCV infection among persons of either European ancestry or African ancestry.[80] The frequency of the advantageous allele varies considerably among global populations, reaching more than 90% in East Asia but having an intermediate prevalence in Europe and a low prevalence in Africa (Table 2).[79,80] This difference in allele frequency explains about half the variation in response rates between patients of African-American ancestry and those of European ancestry.[79,80]

These examples indicate how an understand-

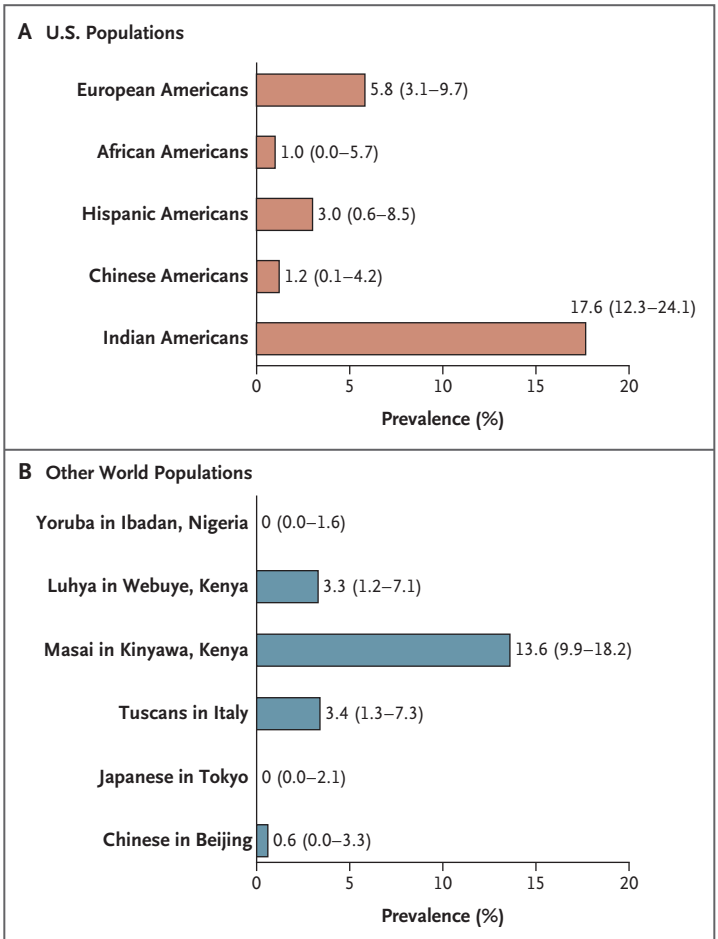

Figure 1. Variation in the HLA-B*5701 Locus in 11 HapMap Samples.

Screening for the HLA-B*5701 (rs2395029) allele substantially reduces the incidence of abacavir hypersensitivity in patients being treated for human immunodeficiency virus infection. Shown are the prevalences (with 95% confidence intervals) of the HLA-B*5701 variant in five U.S. ethnic groups (Panel A) and in six global ethnic groups (Panel B). This variant has a prevalence of 13.6% among the Masai in Kenya but a prevalence of 0% among the Yoruba in Nigeria, which indicates that the use of racial or continental labels such as "black" or "African" can sometimes obscure important, biomedically relevant variation.

The New England Journal of Medicine
Downloaded from nejm.org at SEOUL NATIONAL UNIVERSITY on May 13, 2012. For personal use only. No other uses without permission.
Copyright © 2010 Massachusetts Medical Society. All rights reserved.

Table 2. Mean and Range of the Frequency of the C Allele for Single-Nucleotide Polymorphism (SNP) rs12979860 in African, European, and East Asian Populations.*

Population Group	No. of Individuals	No. of Populations	Frequency	
			Mean	Range
Africa	428	10	36.2	23.1–54.8
Europe	761	13	68.35	52.8–85.7
East Asia	380	8	94.93	90.0–100.0

* SNP rs12979860 is 3 kb upstream from *IL28B*, which encodes interferon lambda-3. The C/C genotype of SNP rs12979860 strongly enhances resolution of hepatitis C virus infection among individuals of either European or African ancestry.[78,79]

ing of the disease-related effects of specific genetic variants provides the basis for direct genetic testing in individuals and alleviates reliance on population categories to improve disease diagnosis and treatment.

REDUCING DISPARITIES IN THE GENOMIC ERA

Genomics, like most human tools, is a double-edged sword. This is especially true in the context of understanding and eliminating health disparities.[43,46,81] Unquestionably, genomics provides novel insights into the causes of and susceptibility to disease and adverse reactions to drugs.[68-73,76,82,83] If used carelessly, however, genomics and associated technologies may exacerbate disparities at multiple levels because of unequal application among human populations; reinforcement of racial stereotyping, identified by the Institute of Medicine as a primary factor in the unequal treatment of minority patients[42]; overshadowing of perhaps more important determinants of disparities, including poverty, racism, unequal access to health care, and cultural practices; exacerbation of existing disparities in access to care owing to increased costs that are likely to be associated with genetic tests and procedures; and conflation of genetic variation at the population level with that among individuals. However, it is worth mentioning that most of these potential drawbacks do not rest on genomic and related technologies themselves but on the interpretation of genomic data within the sociopolitical and economic context of societies.

Genomic data present a new array of opportunities and challenges. Data collection should be extended to as many diverse populations as possible. To achieve this goal, innovative funding mechanisms and additional commitments from genome scientists will be required. It is also critical to assess nongenetic factors, which vary substantially among populations and may interact in important ways with genetic risk factors. Analyses of these effects and interactions can be especially powerful in the context of large, long-term prospective studies.[84] Finally, to make the best use of genomic data, it will be essential to educate students, practitioners, and the public about the beneficial applications of genomics, as well as its limitations. Such measures may help to reduce disparities in health.

Supported by the Intramural Research Program of the Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health (Dr. Rotimi), and by grants from the National Science Foundation and the National Institutes of Health (to Dr. Jorde).

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Dr. Adebowale Adeyemo for the original version of the figure, and Drs. Ed Ramos, Daniel Shrine, and Roland Ikhide Ikheloa for their editorial comments.

REFERENCES

1. Drmanac R, Sparks AB, Callow MJ, et al. Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. Science 2010;327:78-81.
2. Yngvadottir B, Macarthur DG, Jin H, Tyler-Smith C. The promise and reality of personal genomics. Genome Biol 2009; 10:237.
3. Schuster SC, Miller W, Ratan A, et al. Complete Khoisan and Bantu genomes from southern Africa. Nature 2010;463: 943-7.
4. Sachidanandam R, Weissman D, Schmidt SC, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 2001;409:928-33.
5. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. Nature 2006;444:444-54.
6. Sebat J, Lakshmi B, Troge J, et al. Large-scale copy number polymorphism in the human genome. Science 2004;305: 525-8.
7. Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 2005;437: 69-87.
8. Yu N, Jensen-Seaman MI, Chemnick L, Ryder O, Li WH. Nucleotide diversity in gorillas. Genetics 2004;166:1375-83.
9. Li W-H, Sadler LA. Low nucleotide diversity in man. Genetics 1991;129:513-23.
10. Amos W, Hoffman JI. Evidence that two main bottleneck events shaped modern human genetic diversity. Proc Biol Sci 2009;277:131-7.
11. Menotti-Raymond M, O'Brien SJ. Dating the genetic bottleneck of the African cheetah. Proc Natl Acad Sci U S A 1993; 90:3172-6.
12. Xing J, Watkins WS, Witherspoon DJ, et al. Fine-scaled human genetic structure

The New England Journal of Medicine
Downloaded from nejm.org at SEOUL NATIONAL UNIVERSITY on May 13, 2012. For personal use only. No other uses without permission.
Copyright © 2010 Massachusetts Medical Society. All rights reserved.

revealed by SNP microarrays. Genome Res 2009;19:815-25.
13. Li JZ, Absher DM, Tang H, et al. Worldwide human relationships inferred from genome-wide patterns of variation. Science 2008;319:1100-4.
14. Jorde LB, Watkins WS, Bamshad MJ, et al. The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. Am J Hum Genet 2000;66:979-88.
15. Merryweather-Clarke AT, Pointon JJ, Shearman JD, Robson KJ. Global prevalence of putative haemochromatosis mutations. J Med Genet 1997;34:275-8.
16. Tishkoff SA, Reed FA, Ranciaro A, et al. Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet 2007;39:31-40.
17. Sturm RA. Molecular genetics of human pigmentation diversity. Hum Mol Genet 2009;18:R9-R17.
18. Tishkoff SA, Varkonyi R, Cahinhinan N, et al. Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance. Science 2001;293:455-62.
19. Ayodo G, Price AL, Keinan A, et al. Combining evidence of natural selection with association analysis increases power to detect malaria-resistance variants. Am J Hum Genet 2007;81:234-42.
20. Sabeti PC, Varilly P, Fry B, et al. Genome-wide detection and characterization of positive selection in human populations. Nature 2007;449:913-8.
21. Jakobsson M, Scholz SW, Scheet P, et al. Genotype, haplotype and copy-number variation in worldwide human populations. Nature 2008;451:998-1003.
22. Conrad DF, Jakobsson M, Coop G, et al. A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. Nat Genet 2006;38:1251-60.
23. Fagundes NJ, Ray N, Beaumont M, et al. Statistical evaluation of alternative models of human evolution. Proc Natl Acad Sci U S A 2007;104:17614-9.
24. Rosenberg NA, Pritchard JK, Weber JL, et al. Genetic structure of human populations. Science 2002;298:2381-5.
25. Bamshad MJ, Wooding S, Watkins WS, Ostler CT, Batzer MA, Jorde LB. Human population genetic structure and inference of group membership. Am J Hum Genet 2003;72:578-89.
26. Tang H, Quertermous T, Rodriguez B, et al. Genetic structure, self-identified race/ethnicity, and confounding in case-control association studies. Am J Hum Genet 2005;76:268-75.
27. Novembre J, Johnson T, Bryc K, et al. Genes mirror geography within Europe. Nature 2008;456:98-101. [Erratum, Nature 2008;456:274.]
28. Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. Nature 2009;461:489-94.
29. Kosoy R, Nassir R, Tian C, et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. Hum Mutat 2009;30:69-78.
30. Bolnick DA, Fullwiley D, Duster T, et al. The science and business of genetic ancestry testing. Science 2007;318:399-400.
31. Lee SS, Bolnick DA, Duster T, Ossorio P, Tallbear K. The illusive gold standard in genetic ancestry testing. Science 2009; 325:38-9. [Erratum, Science 2009;325: 946.]
32. Royal CD, Novembre J, Fullerton SM, et al. Inferring genetic ancestry: opportunities, challenges, and implications. Am J Hum Genet 2010;86:661-73.
33. Jorde LB, Wooding SP. Genetic variation, classification and "race." Nat Genet 2004;36:Suppl:S28-S33.
34. Tishkoff SA, Reed FA, Friedlaender FR, et al. The genetic structure and history of Africans and African Americans. Science 2009;324:1035-44.
35. Parra EJ, Marcini A, Akey J, et al. Estimating African American admixture proportions by use of population-specific alleles. Am J Hum Genet 1998;63:1839-51.
36. Halder I, Yang BZ, Kranzler HR, Stein MB, Shriver MD, Gelernter J. Measurement of admixture proportions and description of admixture structure in different U.S. populations. Hum Mutat 2009; 30:1299-309.
37. Bryc K, Auton A, Nelson MR, et al. Genome-wide patterns of population structure and admixture in West Africans and African Americans. Proc Natl Acad Sci U S A 2010;107:786-91.
38. Shriver MD, Parra EJ, Dios S, et al. Skin pigmentation, biogeographical ancestry and admixture mapping. Hum Genet 2003;112:387-99.
39. Watkins WS, Rogers AR, Ostler CT, et al. Genetic variation among world populations: inferences from 100 Alu insertion polymorphisms. Genome Res 2003; 13:1607-18.
40. Sankar P, Cho MK. Toward a new vocabulary of human genetic variation. Science 2002;298:1337-8.
41. Racial and ethnic classification used in census 2000 and beyond. Washington, DC: Census Bureau, 2008. (http://www.census.gov/population/www/socdemo/race/racefactcb.html.)
42. Smedley BD, Stith AY, Nelson AR, eds. Unequal treatment: confronting racial and ethnic disparities in health care. Washington, DC: National Academy Press, 2003.
43. Sankar P, Cho MK, Condit CM, et al. Genetic research and health disparities. JAMA 2004;291:2985-9.
44. Molnar S. Human variation: race, types and ethnic groups. 5th ed. Upper Saddle River, NJ: Prentice-Hall, 2001.
45. Luzzatto L. Glucose 6-phosphate dehydrogenase deficiency: from genotype to phenotype. Haematologica 2006;91: 1303-6.
46. Braun L. Race, ethnicity, and health: can genetics explain disparities? Perspect Biol Med 2002;45:159-74.
47. Scriver CR. Human genetics: lessons from Quebec populations. Annu Rev Genomics Hum Genet 2001;2:69-101.
48. Adeyemo A, Rotimi C. Genetic variants associated with complex human diseases show wide variation across multiple populations. Public Health Genomics 2010;13:72-9.
49. Hindorff LA, Sethupathy P, Junkins HA, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A 2009;106:9362-7.
50. Jallow M, Teo YY, Small KS, et al. Genome-wide and fine-resolution association analysis of malaria in West Africa. Nat Genet 2009 May 24 (Epub ahead of print).
51. Need AC, Goldstein DB. Next generation disparities in human genomics: concerns and remedies. Trends Genet 2009; 25:489-94.
52. A global network for investigating the genomic epidemiology of malaria. Nature 2008;456:732-7.
53. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747-53.
54. Kuehn BM. 1000 Genomes Project promises closer look at variation in human genome. JAMA 2008;300:2715.
55. Freedman BI. Susceptibility genes for hypertension and renal failure. J Am Soc Nephrol 2003;14:Suppl 2:S192-S194.
56. Renal Data System. USRDS 2007 annual data report: atlas of end-stage renal disease in the United States. Bethesda, MD: National Institute of Diabetes and Digestive and Kidney Diseases, 2007.
57. Klag MJ, Whelton PK, Randall BL, Neaton JD, Brancati FL, Stamler J. End-stage renal disease in African-American and white men: 16-year MRFIT findings. JAMA 1997;277:1293-8.
58. Rao M, Balakrishnan VS. The genetic basis of kidney disease risk in African Americans: MYH9 as a new candidate gene. Am J Kidney Dis 2009;53:579-83.
59. Kao WH, Klag MJ, Meoni LA, et al. MYH9 is associated with nondiabetic end-stage renal disease in African Americans. Nat Genet 2008;40:1185-92.
60. Kopp JB, Smith MW, Nelson GW, et al. MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. Nat Genet 2008;40:1175-84.
61. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. CA Cancer J Clin 2008;58:71-96.

The New England Journal of Medicine
Downloaded from nejm.org at SEOUL NATIONAL UNIVERSITY on May 13, 2012. For personal use only. No other uses without permission.
Copyright © 2010 Massachusetts Medical Society. All rights reserved.

62. Shavers VL, Brown ML. Racial and ethnic disparities in the receipt of cancer treatment. J Natl Cancer Inst 2002;94:334-57.
63. Cheng I, Plummer SJ, Jorgenson E, et al. 8q24 And prostate cancer: association with advanced disease and meta-analysis. Eur J Hum Genet 2008;16:496-505.
64. Freedman ML, Haiman CA, Patterson N, et al. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. Proc Natl Acad Sci U S A 2006;103:14068-73.
65. Robbins C, Torres JB, Hooker S, et al. Confirmation study of prostate cancer risk variants at 8q24 in African Americans identifies a novel risk locus. Genome Res 2007;17:1717-22.
66. Witte JS. Multiple prostate cancer risk variants on 8q24. Nat Genet 2007;39:579-80.
67. Haiman CA, Patterson N, Freedman ML, et al. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet 2007;39:638-44.
68. Ingelman-Sundberg M. Pharmacogenomic biomarkers for prediction of severe adverse drug reactions. N Engl J Med 2008;358:637-9.
69. Phillips E, Mallal S. Successful translation of pharmacogenetics into the clinic: the abacavir example. Mol Diagn Ther 2009;13:1-9.
70. Hetherington S, Hughes AR, Mosteller M, et al. Genetic variations in HLA-B region and hypersensitivity reactions to abacavir. Lancet 2002;359:1121-2.
71. Mallal S, Nolan D, Witt C, et al. Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir. Lancet 2002;359:727-32.
72. Nolan D, Gaudieri S, Mallal S. Pharmacogenetics: a practical role in predicting antiretroviral drug toxicity? J HIV Ther 2003;8:36-41.
73. Hughes DA, Vilar FJ, Ward CC, Alfirevic A, Park BK, Pirmohamed M. Cost-effectiveness analysis of HLA B*5701 genotyping in preventing abacavir hypersensitivity. Pharmacogenetics 2004;14:335-42.
74. Abel S, Paturel L, Cabié A. Abacavir hypersensitivity. N Engl J Med 2008;358:2515.
75. Park WB, Choe PG, Song KH, et al. Should HLA-B*5701 screening be performed in every ethnic group before starting abacavir? Clin Infect Dis 2009;48:365-7.
76. Information for healthcare professionals: abacavir (marketed as Ziagen) and abacavir-containing medications. Silver Spring, MD: Food and Drug Administration, 2008. (http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm123927.htm.)
77. Seeff LB. Natural history of chronic hepatitis C. Hepatology 2002;36:Suppl 1:S35-S46.
78. Thomas DL, Astemborski J, Rai RM, et al. The natural history of hepatitis C virus infection: host, viral, and environmental factors. JAMA 2000;284:450-6.
79. Ge D, Fellay J, Thompson AJ, et al. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. Nature 2009;461:399-401.
80. Thomas DL, Thio CL, Martin MP, et al. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. Nature 2009;461:798-801.
81. Cooper RS, Kaufman JS, Ward R. Race and genomics. N Engl J Med 2003;348:1166-70.
82. Goldstein DB. Common genetic variation and human traits. N Engl J Med 2009;360:1696-8.
83. Hirschhorn JN. Genomewide association studies — illuminating biologic pathways. N Engl J Med 2009;360:1699-701.
84. Manolio TA, Bailey-Wilson JE, Collins FS. Genes, environment and the value of prospective cohort studies. Nat Rev Genet 2006;7:812-20.

Copyright © 2010 Massachusetts Medical Society.

IMAGES IN CLINICAL MEDICINE

The *Journal* welcomes consideration of new submissions for Images in Clinical Medicine. Instructions for authors and procedures for submissions can be found on the *Journal*'s Web site at **NEJM.org**. At the discretion of the editor, images that are accepted for publication may appear in the print version of the *Journal*, the electronic version, or both.

The New England Journal of Medicine
Downloaded from nejm.org at SEOUL NATIONAL UNIVERSITY on May 13, 2012. For personal use only. No other uses without permission.
Copyright © 2010 Massachusetts Medical Society. All rights reserved.