Genes and Environments: Moving toward Personalized Medicine in the Context of Health Disparities

Rick Kittles, PhD

Understanding gene-environment interactions that may influence disease is the cornerstone of a personalized medicine approach built on diagnostics, risk assessment/ risk modification, pharmacogenetics and biology. Although genetic and personalized medicine can influence clinical decision making, currently most genetic information is based on populations of European ancestry. Additional human genome research must include diverse populations in order to assess the impact DNA sequence variation and environmental influences have on human disease risk. Within this article, we present a brief overview of human genome variation and discuss how epigenomics may influence gene expression. Examples of gene-environment interactions are explored and linked to several health disparities and health outcomes among communities of color. With the Jackson Heart Study poised to take the next steps in examining genes and the environment in ways that other cohorts cannot, we will be closer to a more inclusive personalized medicine goal that transforms medicine from curative to preemptive for all. (*Ethn Dis*. 2012; 22[Suppl 1]:S1-43–S1-46)

Key Words: Genomics, Epigenomics, Genetic Ancestry

From the Department of Medicine, University of Illinois at Chicago, Chicago, Illinois.

Address correspondence to Rick Kittles, PhD; Section of Hematology/Oncology (M/C 767); Department of Medicine; University of Illinois at Chicago; 900 S Ashland Ave. Suite 3302; Chicago, IL 60607-4067; 312-355-5956; 312-413-4131(fax); rkittles@uic.edu

Introduction

The Human Genome Project (HGP) was an international, collaborative research program designed to map the human genome, the largest genome to be sequenced with the identification and location of >20,500 human genes.[1,2] This map provides details about the structure, location, and organization of human genes. Once efforts were completed in determining the sequence of more than 3 billion nucleotides, focus began on the identification and cataloging of tens of millions of single nucleotide polymorphisms (SNPs).[3] Early during this phase of SNP discovery, efforts were focused on identifying common SNPs assuming a common variant/ common disease model where common SNPs were seen to impact complex diseases such as type 2 diabetes and hypertension. Recent work however suggests that multiple rare variants may play a much more significant role than first thought.[4] One current and fascinating challenge is that many of these SNPs, appear to be important in identifying risk of disease, but are located in areas that are not labeled as genes or expressed regions in the genome.

The International HapMap Project also began the process of indexing SNPs for genetic studies on complex diseases.[3] The HapMap public database contains common variation in the human genome, and provides insight on the block-like structure of linkage disequilibrium across the genome and haplotype diversity across major population groups. This information can be used to guide the design and analysis of genetic association studies and has helped identifying loci that are linked to specific disease or risk factors.[3,5]

Genetic variation is structured across human populations with African descent populations exhibiting higher genetic diversity than non-African populations. Resequencing studies have revealed that African descent populations have a much higher prevalence of rare variants (allele frequencies >.01) than non-African populations. Many of these variants are within genes or regulatory regions that impact important physiological and drug pathways and thus may be important in disease risk and drug efficacy.

Because these variants may impact disease risk, research needs to continue exploring genetic variation across human populations. This is particularly important for genome wide association (GWA) studies where millions of SNPs across the genome of patients and healthy individuals are examined to find signatures that might reveal evidence of association of disease and markers in the genome.[6,7] Most importantly, risk regions discovered in one population may not necessarily present risk in other populations or, if they do, they may not necessarily confer the same magnitude of risk in other populations.

Emergence of Personalized Medicine

Genetics continues to be integrated into medicine, emerging toward new directions by using genetic risk to build personalized or individualized medicine plans. Thus, understanding how genetic variation is structured within and between populations is important in genetic medicine. Personalized medicine is utilizing genetic information more and more in clinical practice and has improved diagnostics, pharmacogenetics, the development of new drugs, risk assessment and ultimately risk modification.[8] Although there is promise in

personalized medicine, it is still an expensive concept and, historically, studies mainly have focused on individuals of European ancestry.

Will personalized medicine work for all populations? While research is pointing to its benefits, how will clinicians be best positioned to make use of the tools of personalized medicine? Clinical decision-making informed by genetic tests may help improve patient outcomes. However, if care is not taken to have participation across all communities, genetic medicine could actually increase health disparities. Currently, most genetic information and mapping is based on populations of European ancestry, meaning that most diagnostic tests and information on risk assessment are specific for individuals of European ancestry. We know relatively little about risk assessment in populations of color or populations that have historically been underserved in medicine. Fortunately, the Jackson Heart Study is in a position to provide insightful genetic and environmental data on African Americans, particularly relating to cardiovascular disease, which may aid in eliminating health disparities.

UNDERSTANDING HUMAN GENOME SEQUENCE

Within the human genome sequence, evolutionary conserved regions are important: approximately 5% are conserved and 1.5% are protein-coded.[9] Independent and unbiased experiments are needed to identify all of these functional regions. Some GWA studies are finding that regulatory regions for genes may be located far away from the affected genes.[10] The ENCODE Project (ENCylcopedia Of DNA Elements) was launched by the National Human Genome Research Institute (NHGRI) in 2003 to identify all functional elements of the human genome sequence. Results of the pilot phase provided identification and analysis of functional elements in 1% of the human genome.[10] This and continuing data from ENCODE will enhance our understanding of the regulation of gene expression on a spatial, temporal and quantitative level. It will help us identify the important gene expressions, how they interact, how polymorphisms affect gene expression, and if/how we can predict and manipulate gene expression. It will expand our knowledge of genetic basis of disease by understanding the SNPs in non-coding regions. Finally, through the rapidly growing field of epigenomics and the ENCODE data, we will have a greater understanding of histone modifications, DNA methylation, Dnase hypersensitive sites, micro RNAs (miRNAs), and small non-coding RNSs (ncRNAs) and the impact of environment (eg, exposures such as alcohol, smoking, sexual behaviors, etc) on their functions.

Epigenetic mechanisms, such as gene modification (DNA methylation which activates or represses genes) occurs when a DNA strand is exposed to an external factor, such as a methyl group (an epigenetic factor found in some diets). The DNA, a very long, tightly packed chromosomal structure can be wound around histones (proteins), which are affected by epigenetic mechanisms. Certain areas of the genome have many gene modifications, while other areas have limited, if any, gene modification. With the binding of epigenetic factors to histones and the resulting activation of the DNA, certain factors have been linked to health endpoints such as cancer, autoimmune disease, mental disorders and diabetes.

An example of the utility of understanding these epigenetic mechanisms is illustrated in the recent study by Jia and colleagues that showed how the 8q24 region is associated with risk of prostate cancer.[11] In this study, Jia profiled a 5-megabase chromatin segment in a conserved region where genes were not present but that encompassed all risk regions for RNA expression, histone modifications and locations occupied by RNA polymerase II and the androgen receptor (AR). They found that two enhancers in one risk region were occupied by AR and responded to androgen treatment; one contained a SNP (rs11986220) that resides with a FoxA1 binding site, with the prostate cancer risk allele facilitating both stronger FoxA1 binding and stronger androgen responsiveness.

From this and other studies, we understand that even with the sequencing of genome of a person with disease, genetics alone will not provide the answers needed to improve health of that individual. Research, such as that being conducted by the Jackson Heart Study, must also continue to explore the social determinants of health and other environmental exposures, that, when utilized in the context and in relation to genetics, we will develop a more defined picture of how genes interact with the environment.

GENE-ENVIRONMENT INTERACTIONS

Our genes and our environments work together to produce almost all of the phenotypes in the human species. For geneticists, understanding how environmental factors (physical, social, behavioral) affects genetic variation, also known as natural selection, may provide clues to gene function and other pathways to disease. Although environmental effects may vary across populations, gene-environment interactions may be informative for treatment and modifying effects of genes. In fact, strong environmental risk factors are likely easier to modify than genes.

Gene-environment interactions can mask the effects or the detection of the genetic effects in disease.[12] Environmental factors vary considerably across populations and may also modify genetic effects within populations. For example, using data from NHANES, Scragg et al looked at vitamin D levels (25 hydroxy vitamin D) in non-Hispanic

Whites, non-Hispanic Blacks, and Mexican Americans.[13] They found vitamin D levels among Whites to be almost twice that of Blacks (79.6 nmol/L vs 49.1 nmol/L, respectively).[13] In fact, data suggest that a majority of Blacks are vitamin D insufficient.[14] With the known health benefits of vitamin D, often derived from sunlight, the lower levels of vitamin D places Blacks at greater risk for cancer, and cardiovascular and immune system diseases. Increasing vitamin D to levels known to be beneficial may impact cancer and cardiovascular risk.[15]

Gene-environment interactions will become increasingly more important for our research efforts and, in particular, for the next phases of the Jackson Heart Study. For example, research by Cooper et al illustrated that, although African descent populations in North America, Caribbean and West Africa, have similar genetic background, body mass index (BMI) levels, and ultimately the prevalence of hypertension, vary greatly among the three groups, and are positively correlated with increasing BMI as populations go from west Africa to the Caribbean to north America .[16]

As we seek reasons for the poorer health outcomes (heart disease, diabetes, asthma, cancer, HIV/AIDS, infant mortality) in communities of color, we must continue to study diet, lifestyle, socioeconomic status and other environmental exposures (stress, discrimination, medical literacy, etc) independently of race to understand the gene-environment interaction that affects disease. Although we know that African Americans have a high genetic heterogeneity, we must also understand the geographic distribution of African Americans in the United States to fully understand gene-environmental interactions that may vary between Black populations in the South, North or other parts of the United States.

Ancestry informative markers (AIMs) and the use of genetic ancestry are also important in identifying reasons for poor health outcomes. We use these markers to estimate effects of genetic ancestry since all populations show some level of variation in genetic ancestry. For example, Figure 1 shows the broad distribution of genetic ancestry for African Americans (AAs) living in Washington, D.C. and European Americans (EAs) living in State College, Pennsylvania.[17] Similarly, broad patterns are seen when looking at the distribution of ancestry among Hispanic populations (eg, Mexican Americans in Colorado and Puerto Ricans from New York).

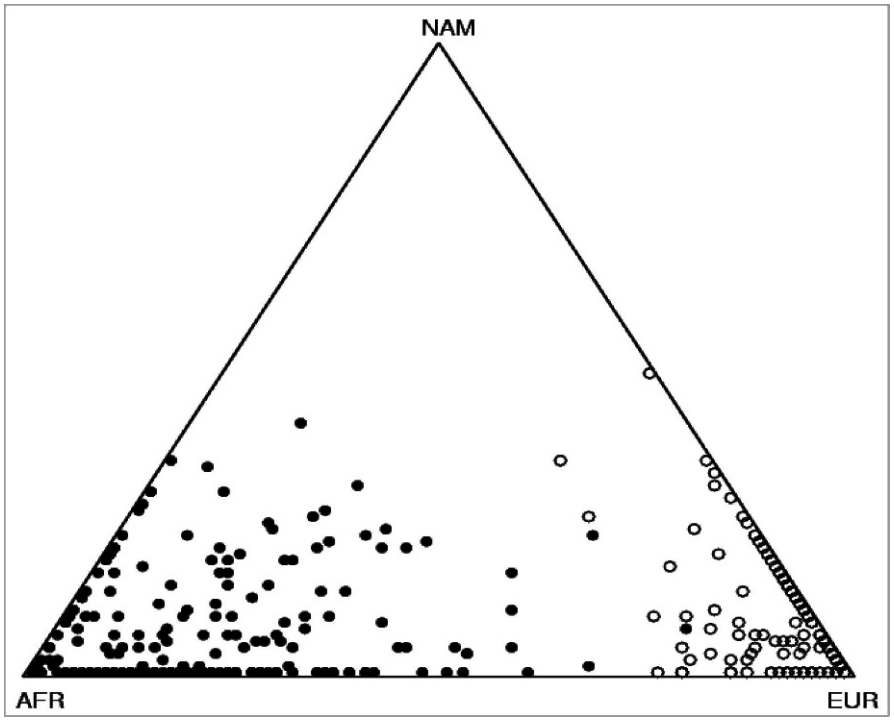

Fig 1. Individual ancestry of African American sample (*n*=221) from Washington, DC compared to European American sample (*n*=193) from State College, Penn. Open circle = European American; Closed circle = African American

A large number of AIMs have now been identified and ancestry can be estimated across chromosomal regions and can assist in mapping genes for disease.[18] (Table 1) However, genetic composition is not the sole contributor to disparity in these diseases (multiple sclerosis, lung cancer, stroke, end-stage renal disease, prostate cancer and hypertensive heart disease). For example, admixture mapping for prostate cancer risk genes, reveals a strong genetic ancestral component irrespective of the environment.[19] However, for

Table 1. Candidate disease for admixture mapping in African Americans

Disease	Relative Risk (95% CI)*
Multiple sclerosis	0.45 (0.35–0.55)
Lung cancer	1.48 (1.38–1.67)
Stroke	1.57 (1.27–1.94)
End-stage renal disease	1.87 (1.47–2.39)
Prostate cancer	2.73 (2.03–3.86)
Hypertensive heart disease	2.80 (2.00–4.93)

* Relative risks are for African Americans vs European Americans.

hypertension, genetic ancestry does not appear to contribute much when you control for social economic status (SES); and, as of yet, we have been unsuccessful in using admixture mapping to explain the higher rates of diabetes among African Americans. This would suggest that there are strong environmental factors that contribute to diabetes disparities among African Americans.

CONCLUSION

Over the next decade, the equitable benefits from genetic medicine will depend on population genetics, market economics, and historical and cultural identities. Gene-environment interactions will be an exciting area of research focus, particularly as it relates to health disparities. However, we must prioritize, based on research, which health disparities have strong genetic influences due to genetic ancestry. We must study the effects from the environment and develop interventions to address these disparities. We must also increase trans-disciplinary research efforts by engaging more social scientists in the conversation to gain better appreciation and understanding of critical non-genetic variables that should also be explored. Thankfully, the Jackson Heart Study is poised to take the next steps in examining genes and the environment in ways that other cohorts cannot. With this research and that from other institutions, we will be able to transform medicine from curative to preemptive.

REFERENCES

1. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409:860–921.
2. McPherson JD, Marra M, Hillier L, et al. A physical map of the human genome. *Nature*. 2001;409:934–41.
3. Altshuler DM, Gibbs RA, Peltonen L, et al. Integrating common and rare genetic variation in diverse human populations. *Nature*. 2010;467:52–8.
4. Cirulli ET, Goldstein DB. Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Reviews Genetics*. 2010;11:415–25.
5. Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES. High-resolution haplotype structure in the human genome. *Nature Genetics*. 2001;29:229–32.
6. Bustamante CD, Burchard EG, De la Vega FM. Genomics for the world. *Nature*. 2011; 475:163–5.
7. Manolio TA. Genomewide association studies and assessment of the risk of disease. *N Engl J Med*. 2010;363:166–76.
8. Chan IS, Ginsburg GS. Personalized medicine: progress and promise. *Ann Rev Genom Hum Gen*. 2011;12:217–44.
9. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science.* 2004;306:636–40.
10. Birney E, Stamatoyannopoulos JA, Dutta A, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*. 2007;447: 799–816.
11. Jia L, Landan G, Pomerantz M, et al. Functional enhancers at the gene-poor 8q24 cancer-linked locus. *PLoS Genetics*. 2009;5: e1000597.
12. Cornelis MC, Agrawal A, Cole JW, et al. The Gene, Environment Association Studies consortium (GENEVA): maximizing the knowledge obtained from GWAS by collaboration across studies of multiple conditions. *Gen Epidemiol*. 2010;34:364–72.
13. Scragg R, Sowers M, Bell C. Serum 25-hydroxyvitamin D, diabetes, and ethnicity in the Third National Health and Nutrition Examination Survey. *Diabetes Care*. 2004;27: 2813–8.
14. Murphy AB, Kelley B, Nyame YA, et al. Predictors of serum vitamin D levels in African American and European American men in Chicago. *Am J Men Health*. 2012.
15. Grant WB, Boucher BJ. Requirements for Vitamin D across the life span. *Biol Res Nurs*. 2011;13:120–33.
16. Cooper R, Rotimi C, Ataman S, et al. The prevalence of hypertension in seven populations of west African origin. *Am J Public Health*. 1997;87:160–8.
17. Kittles RA, Santos ER, Oji-Njideka NS, Bonilla C. Race, skin color and genetic ancestry: implications for biomedical research on health disparities. *Calif J Health Promo*. 2007;5:9–23.
18. Winkler CA, Nelson GW, Smith MW. Admixture mapping comes of age. *Annu Rev Genomics Hum Genet*. 2010;11:65–89.
19. Bock CH, Schwartz AG, Ruterbusch JJ, et al. Results from a prostate cancer admixture mapping study in African-American men. *Hum Genet*. 2009;126:637–42.