Racial Differences and the Genetics of Hypertension

Richard S. Cooper, MD, and Xiaofeng Zhu, PhD

Address

Department of Preventive Medicine and Epidemiology, Loyola University Stritch School of Medicine, 2160 S. First Avenue, Maywood, IL 60153, USA. E-mail: rcooper@luc.edu

Current Hypertension Reports 2001, 3:19–24 Current Science Inc. ISSN 1522–6417 Copyright © 2001 by Current Science Inc.

Variation in the prevalence of hypertension among human populations has been used to examine a wide range of research questions. The best known example has always been the twofold greater prevalence among blacks, which has provoked a debate about the relative importance of environmental and genetic factors. This problem is complicated because it requires an understanding of both the genetics of hypertension and the genetic basis of variation among populations. Molecular data accumulated over the past year are beginning to provide new insights into this old question. Modest progress has now been made in the genetics of hypertension. Important advances have been made in understanding rare hypertension syndromes, and large-scale studies of the general population are under way, although the results to date are inconclusive. Markers used to search for hypertension genes can also provide estimates of between-population variability. These data, which are based on genetic variants used for etiologic research, confirm previous evidence of group-specific variation, while underscoring the limited magnitude of this variation. Despite rapid progress, this work is still in its infancy and raises more questions than it answers.

Introduction

In the United States, understanding the causes and consequences of population variation in blood pressure has been a major preoccupation of researchers and public health practitioners. Along with the well-established social class gradient, the finding of more frequent and more severe high blood pressure among blacks has been among the most consistent epidemiologic observations in this field [1]. While it is well accepted that the social class gradient results from environmental factors, the extent to which the racial differences reflect genetic variation remains an unanswered question. The molecular revolution has created the expectation that this question can be tested directly, and the technological advances in genetic sequencing and analysis over the past year suggest to some that the possibility may actually be close at hand.

The requirement for an answer to the question posed in this paper—whether genes play a role in the higher rate of hypertension among blacks compared with whites—is twofold. First, it will be necessary to know the important genes and variants that underlie susceptibility to hypertension. Second, we will need to understand race on the genetic level. In this brief review, we consider recent developments relevant to both these questions.

Genes and Hypertension

Three basic types of studies are being used to investigate the genetic factors that contribute to hypertension (Table 1).

Rare syndrome studies

Positional cloning, the method used so successfully to identify the genes that underlie monogenetic or Mendelian disorders such as cystic fibrosis and Huntington's chorea, has now been applied to several familial types of hypertension $[2 \cdot \cdot, 3, 4]$. The advantage of this approach is that when a positive result is obtained, it provides an unambiguous explanation of the cause of the syndrome under investigation and often yields basic insights into pathophysiology. While the number of individuals with these syndromes will always be small, the implications of the findings can apply to much larger groups.

For example, uncovering the basis of familial hypercholesterolemia opened an enormously productive field of research on hepatic cholesterol metabolism and the development of the statin drugs. Perhaps as a parallel, an important insight to emerge from the studies of singlegene blood pressure disorders has been the consistency with which abnormalities in the pathways handling sodium excretion have been identified. For example, Liddle syndrome, a rare condition that leads to excess sodium reabsorption in the kidney, has been shown to have a precise genetic and physiologic origin [3]. Most recently, Geller et al. [2••] identified a mutation in the mineralocorticoid receptor that results in hyperfunction. During pregnancy, when circulating levels of mineralocorticoids are high, this abnormality leads to severe hypertension. Furthermore, biochemical studies among these patients identified the molecular basis for the altered function, thus illuminating a generalizable process in receptor function as related to blood pressure regulation.

While acknowledging the apparent success of this use of molecular genetics, it is not likely that studies of rare syndromes will be helpful in explaining population variation in hypertension.

Candidate gene studies

Case-control studies—or what geneticists call "association" studies-have been widely used because they are more powerful than linkage analysis. If we reason from physiologic pathways, specific genes are targeted as candidates for regulators of blood pressure, and, by inference, hypertension risk. The most prominent among this category are the genes of the renin-angiotensin system, and an enormous amount of literature now exists on this topic. The marker that is associated with variation in the angiotensin 1-converting enzyme (ACE) activity—the insertion/ deletion (I/D) marker—has been studied in almost every ethnic and racial group. This marker is of ancient origin and is similar in frequency in most human populations. While the strong relationship to ACE level is present across populations, a consistent relationship to cardiovascular conditions has not been demonstrated [5].

The results from small studies or the analyses of subgroups have generally been used to promote the view that the ACE I/D polymorphism confers susceptibility to cardiovascular disease. Larger studies over the past few years have tended to be entirely negative, however. For example, our group recently examined 13 ACE markers in 1400 individuals and found no relationship to blood pressure [6]. A study of 11,000 individuals was likewise required to put to rest the notion that ACE contributes to risk of myocardial infarction [7].

In theory, candidate gene studies hold the potential to address why hypertension risk varies among populations. With the availability of the sequence of the human genome, and as more chromosomal regions are identified with linkage analysis from genome-wide scans, candidate gene studies are likely to become increasingly important. Just as the frequencies of certain disease-associated genotypes must vary between cases and controls, variation in the frequencies of these "at risk" genotypes between populations will underlie group difference in risk.

To test this hypothesis, however, it will first be necessary to define the genetic variation that predisposes to hypertension. On the basis of early studies, a genetic variant in the angiotensinogen gene was thought to be associated with hypertension, and this variant is more common in black persons [8]. A recent large study could not replicate the association with hypertension in either blacks or whites, however [9•].

There will be many more chapters to this story. Two recent papers from the United Kingdom have reported associations of new gene variants with hypertension in blacks. A variant in a gene influencing the reabsorption of

Table 1. Genes and hypertension: where are we?

Rare syndromes

- Substantial progress; insights into pathophysiology Candidate genes
- Few universally accepted candidates; several interesting stories (eg, the renin-angiotensin system, beta-2, adducin) Genome-wide scan
- No consistent findings; a "work in progress"

sodium in the kidney was found to be more common in hypertensives [10]. The same group also reported that a polymorphism in the G protein, which is involved in signal transduction, was significantly increased in black hypertensives [11]. In both cases the variant associated with higher risk was more common in blacks. As expected, the higher frequency of the "at risk" allele suggested to the authors that these genes might explain part of the putative predisposition to hypertension.

The preceding examples of ACE and angiotensinogen both serve, however, as a clear warning about how difficult it will be to find consistent relationships for hypertension genes. It will next be necessary to demonstrate that the markers being studied are true causal mutations or that they have the same effect in different populations in which the background of other genes may differ. Likewise, certain mutations could be important only in certain environments that may vary among populations. On the basis of studies of average size, however, it has been the rule that the associations of candidate hypertension genes are inconsistent in different populations [10,12–14]. At this stage, it is impossible to know whether that variation is a result of sampling variation or true differences in effect among populations.

Genome-wide scans

In the absence of an appropriate candidate gene, a search can be made over the entire genome [15-17]. Using "anonymous markers" from the parts of the chromosomes that are not involved in coding proteins, one can test the similarity of the inheritance pattern of phenotypes and the marker genotypes observed in a pedigree. Several studies have used this method as a first step to localize DNA regions that might contain functional mutations [16,17].

This method was applied successfully by Bray *et al.* [18] in a sample of white families from Rochester, Minnesota. Sequencing of the β_2 -adrenergic receptor gene that was located in the linked region revealed two variants that were associated with hypertension risk. Again, however, the complexity of this task must be emphasized. While statistically significant, the associations identified in the β_2 adrenergic receptor study accounted for only 2% of the variation observed in blood pressure, or a 1- to 2-mm Hg difference between persons with and those without the mutation [18]. Clearly such subtle findings will be missed in some studies and may be difficult to compare across populations. However, a recent study conducted among black patients in the United Kingdom also reported that the β_2 -adrenergic receptor gene variants were associated with blood pressure [19]. If confirmed, this finding would suggest that the effect is similar in the two populations.

In sum, practical tools are being developed that make it possible to examine whether genetic variation influences risk of hypertension. While the application of these tools has met with limited success, the tools have not reached the level of sophistication that would allow their direct application to a study of the genetic determinants of blood pressure variation between populations.

Genetic Variation Among Populations: What Is This Thing Called Race?

Few concepts in biology are more problematic than race. Although population variation in physical traits is undeniable, the extent to which this variation can be used to define groups remains very much an open question [1]. At the same time, for the traits of greatest interest to medical science, including hypertension, risk is heavily influenced by environmental factors, which also vary systematically among populations. In technical terms, therefore, confounding can be said to exist between genes and the social or environmental exposures. To solve this problem, one needs either to use better measures of the environment or a clear definition of what constitutes race on a genetic level. Environmental exposures that accumulate over a lifetime are notoriously difficult to measure; better epidemiologic methods are needed. To make progress on the question being discussed here, a better understanding of the concept of race is also important.

Defining what constitutes a race, in genetic terms, must begin with the general problem of whether our species aggregates naturally into subunits of any size. We now know, for example, that with as few as five "tandem repeat" or "microsatellite" markers, we can unambiguously identify all individual members of our species. This method, known popularly as "DNA fingerprinting," has brought about the near collapse of the US system of assigning guilt for homicide and demonstrates the power of molecular methods.

This success, however, follows from the fundamental biological principle that all organisms are unique. The challenge of classifying individuals into groups—that is, creating meaningful categories out of naturally occurring populations, as opposed to distinguishing individuals from each other—is very different. Classification implies that structure exists in the naturally occurring distribution. If one incorrectly forces structure on the data by classifying continuous variation into categories, the result may be to reduce, rather than increase, the chances of understanding the phenomenon of interest. Significant advances in biology have occurred when a misconceived categorical way of thinking is replaced with a "system approach," based on the recognition that the data fit a graded and continuous distribution. For example, the Darwinian perspective, based on the concept of "descent with variation," is based on a break with the earlier system that used "typologies" or categories to classify organisms, and it opened up a fundamentally new way of understanding biology.

Describing the genetic basis of race, therefore, is tantamount to asking whether the collective human genome is a single whole or whether it aggregates into subgroups. To frame the question adequately, it will be necessary to represent the human genome in a manner that provides a comprehensive and unbiased assessment of variation at the level of the individual, the population, and the species. This representation should furthermore communicate our understanding of the human genome conceptually, mathematically, and graphically, and provide a reference point from which the role of genetic variation among populations can be evaluated. It is toward this goal that population geneticists have been applying the new tools of molecular research.

By analogy, one might first consider the problem of defining what constitutes the basis of gender or species differences at the level of the genome. In descriptive terms we know that the "X" and "Y" chromosomes throw large "switches" that influence every cell. While this presumably works through hormonal regulation, the precise genetic basis for this effect is still unknown. Nonetheless, classification by gender is not a problem. In terms of species, however, the message being learned from whole-genome sequencing projects is the remarkable degree of overlap across species. For example, we share approximately 30% sequence identity with distant relatives in the biological world, including plants, while with our nearest relatives among the primates we share 99% DNA identity [20••].

The question that naturally follows from these observations is, What is the molecular basis of specification? Efforts have now been launched to define the genetic meaning of species. Rearrangement of chromosomes has been the most prominent hypothesis for the divergence of species; however, this hypothesis has been recently challenged by molecular evidence [21]. Accumulation of sequence variation must be the underlying process, where chromosomal rearrangement does not occur. In studying race, a parallel yet obviously much more difficult task will be faced—identifying the specific sets of variants that could be used to form the basis of a genetic definition.

A method to study race must therefore be premised on the availability of appropriate data on sequence variation. Substantial progress has been made over the past year in methods that can define genomic variation. Mass genotyping of both "microsatellite" markers and single nucleotide polymorphisms is now possible [16– 19,20••,22,23]. These techniques can be used to generate dense sets of markers in specific regions or in linkage analysis involving the genome as a whole, as described above. These methods, however, identify only "markers" that differentiate segments of DNA that have a different evolutionary history; they do not provide the information necessary to relate genetic variation to phenotypic variation or variation in the organism. This form of genetic data should therefore be seen as an indirect description of genetic variation because it is based on markers that have no functional significance. Because these data have no direct implication for the physical state of the organism, they do not tell us why there are population differences in hypertension risk.

Estimating Population Genetic Variation

In terms of genetic analysis, defining races would require the construction of categories that could be used to classify people on the basis of their DNA. These categories would, by necessity, correspond to genetically related populations that had intermarried over long periods of time. These populations would then obviously have to share some physical traits that differentiated them in a meaningful way from other human populations. While these premises can be accepted, the hazards of this approach must be acknowledged. As stated above, a classification system that assumes the human genome consists of separate units may rest on an erroneous premise, and thereby reduce rather than enhance our ability to understand the biological processes.

While it is not difficult to identify specific traits that aggregate within large human populations, most obviously the degree of skin pigmentation, classifying people on the basis of a single trait (and its genetic correlate once that has been defined), would not provide a meaningful classification scheme. The purpose of a classification scheme must be to define naturally aggregating groups—what 19th-century biologists referred to as "carving nature at its joints."

The arrangement forced on the data by classification must therefore have some meaning at a higher organizational or conceptual level. Gender, for example, is clearly a meaningful classification, and a simple genetic marker the "Y" chromosome—can be defined, even if its full effect remains obscure. While all species evolved from a common ancestor, classification of contemporary variation also creates information. Thus, while humans and chimpanzees are identical for many genes, they vary as organisms in a meaningful way, and research projects are under way to identify the essential genetic basis of this species difference. However, for subspecies, or races, it is far from clear that either requirement for classification can be met, that is, meaning is being created and a genetic basis can be defined.

If the individuals designated by the proposed scheme resembled each other on only a single trait (*eg*, skin color in Australian aborigines and Cameroonian pygmies), then nothing would have been accomplished. If the attempt is to acknowledge similarity of shared traits, then the scheme is no more than a recapitulation of geography

Table 2. Genetic distance between US blacks and whites, estimated by heterozygosity of microsatellite markers

Within blacks Within whites	0.799
Between blacks and whites	0.808
"Distance"	0.024

and history without separate biological meaning. The classification process must tie together a bundle of genetic traits that make one population distinct from the others. In essence, the classification process must organize and give meaning to the variation across multiple regions in the genome, since that is the unit on which classification depends. In effect, one would need to undertake a cluster analysis for a large representative set of variants that were directly related to physical traits. While that goal remains a long way off, and may well be unattainable in the end, it is possible to use existing data to probe this question more deeply.

Estimates of genetic diversity based on microsatellite markers

Microsatellite markers, as noted previously, are short tandem repeat sequences that occur in portions of the genome that are not translated. Because they have no functional consequences, mutations accumulate at a high rate in these regions and are useful as measures to discriminate one person from another. Many epidemiologic investigations of hypertension have now used microsatellite markers to perform a genome-wide scan [16]. In a recent collaborative study, 2800 individuals were examined as part of the Family Blood Pressure Program [24]. Since several population groups were involved, these data can also be used as one way to summarize variation. After exclusion of the "X" and "Y" chromosomes, data on 368 markers were available from a genome-wide scan performed by the Mammalian Genotyping Service at the Marshfield Medical Foundation. Samples of whites and blacks were used to estimate the amount of variation that exists between these two groups.

The simplest measure used to compare groups is the relative amount of variation within populations compared with between populations, usually based on an estimation of heterozygosity. First, the probability that two alleles selected at random from within a group will be different is calculated; next, the probability that an allele selected at random from one group will differ from an allele randomly selected from the other population group is calculated; finally, the difference between those probabilities is determined. As shown in Table 2, for US blacks, the within-population probability that two alleles are different is 80%, while for whites it is 77%. A comparable result, when the alleles are selected from the two different populations, is in the order of 81%. The



Figure 1. Frequency and population distribution of single nucleotide polymorphisms (SNPs) in European-origin and African populations.

"genetic distance" is therefore 1% to 3%, while the degree of overlap in genetic sharing is therefore 98%.

Estimates of genetic diversity based on single nucleotide polymorphisms

In addition to microsatellites, variation in single nucleotides themselves can now be measured on a large scale $[20 \cdot ,23]$. These variants occur both in coding regions of genes, some of which then alter proteins, and in those regions that do not result in any coding changes. These data can be used to provide an "upper bound" on the amount of variation that occurs between populations. As would be expected, in most cases when a common variant is found, it is shared among all human populations. This result occurs because sufficient time must have passed for this mutation to increase in frequency; the older the mutation, the more likely it was to be present in the original founder population in Africa, and thereafter distributed throughout the world.

Halushka et al. [20..] studied a sample of 75 genes known to play a role in blood pressure regulation in samples of persons of European and African ancestry. All single nucleotide polymorphisms were identified using a DNA chip. In Figure 1, the frequency of the less common variant (ie, "minor allele") is plotted in Europeans and Africans. The alleles that have a frequency exceeding 10% are more likely to be shared (ie, present in both groups). Only the rare alleles occur in one group or the other. This distribution reflects the relatively young age of our species and underscores the finding that almost all variation is found within populations, with very little between groups. Expressed in terms of similarity and difference, the frequency of alleles that occur in only one population will be around 0.01%. This estimate is smaller than that obtained with microsatellites because less variation occurs for these single nucleotide polymorphisms, especially those in coding regions.



Figure 2. Schematic representation of population genetic variation.

The schematic representation in Figure 2 suggests how the two groups could have diverged over time. The "trunk" of the human tree is rooted in Africa, and the largest amount of diversity continues to be present in African populations [25]. As population subgroups left Africa, they may have taken with them a subset of genetic variation (represented as the area above the dashed line), and they would have developed additional unique DNA variants in the time since separation.

Framework for interpreting data on genetic variation

Although DNA marker data are helpful because they set bounds on the amount of variation to be expected across populations, they fall short of answering the question of interest. Ultimately what one wants to know is, How much impact does the genetic variation have on the organism? As noted previously, the gene variants that influence blood pressure are still unknown for the most part and the markers used for all the studies we discussed are effectively "silent."

The conventional measures of population variation that have been used until now are therefore not likely to provide the basis for an adequate test of whether the human genome can be divided into meaningful clusters. Now that the sequence data for the genome of many organisms have become available, a new understanding of "genomics" will soon be available. What we learn from the new field of genomics may well improve our ability to summarize the information over the whole genome and to address these questions more effectively.

Conclusions

This review has dealt primarily with the technical developments in the study of race and the genetics of hypertension. Clearly this field can be thought of only as a "work in progress" since genetic variation known to be important in disease susceptibility has not yet been described for hypertension. For the sake of argument in this discussion, we have temporarily accepted the hypothesis that race can be defined as a biological category. Of course, this may well not be the case. It should go without saying that race is already known to have enormous social meaning, and that may account entirely for the hypertension patterns that are observed in the United States [1,26,27]. For example, the rates of hypertension among black populations vary widely, being much lower in Africa and the Caribbean, demonstrating the primary influence of the environment [28]. In fact, the difficulties facing a genetic answer to the problems discussed here lead to at least an interim conclusion that on a molecular basis we may never be able to dissect our species into subgroups. Conceptually, that result reinforces the view of those who, like Darwin, argue that we "have no right to give names to objects we cannot define" [29].

Posing more detailed questions will help deepen this debate, however, and the problem of population variation in blood pressure is an important practical application of the notion of race. As population genetics takes advantage of the "molecular revolution," our understanding of these questions will undoubtedly deepen over the next several years.

References and Recommended Reading

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- •• Of major importance
- 1. Cooper RS, Rotimi C: Hypertension in populations of West African origin. Is there a genetic predisposition? *Am J Hypertens* 1994, **12**:215–227.
- 2.•• Geller DS, Farhi A, Pinkerton N, et al.: Activating mineralocorticoid receptor mutation in hypertension exacerbated by pregnancy. *Science* 2000, 289:119–123.

This report elegantly demonstrates the power of molecular genetics to elucidate familial hypertension.

- Warnock DG: Liddle syndrome: an autosomal dominant form of human hypertension. *Kidney Int* 1998, 53:18–24.
- 4. Lifton RP: Molecular genetics of blood pressure regulation. *Science* 1996, **272:**676–680.
- Agerholm-Larsen B, Nordestgaard BH, Tyhjaerg-Hansen A: ACE gene polymorphism in cardiovascular disease: meta-analyses of small and large studies in whites. Arterioscler Thromb Vasc Biol 2000, 2:484–492.
- 6. Zhu X, McKenzie C, Forrester T, *et al.*: Localization of a small genomic region associated with elevated ACE. *Am J Hum Genetics* 2000, **67**:1144–1153.
- Keavney B, McKenzie C, Parish S, et al.: Large-scale test of hypothesized associations between the angiotensinconverting enzyme insertion/deletion polymorphism and myocardial infarction in about 5000 cases and 6000 controls. International Studies of Infarct Survival (ISIS) Collaborators. Lancet 2000, 355:434–442.
- Rotimi C, Cooper R, Ogunbiyi O, et al.: Hypertension, serum angiotensinogen and molecular variants of the angiotensinogen gene among Nigerians. Circulation 1997, 95:2348–2350.

9.• Province MA, Boerwinkle E, Chakravarti A, et al.: Lack of association of the angiotensinogen-6 polymorphism with blood pressure levels in the comprehensive NHLBI Family Blood Pressure Program. J Hypertens 2000, 18:867–875.
Because the literature is filled with small studies, this cooperative

multicenter project demonstrates what will be required to find consistent results.

- Williams SM, Addy JH, Phillips JA, et al.: Combinations of variations in multiple genes are associated with hypertension. Hypertension 2000, 36:2–6.
- 11. Dong Y, Zhu H, Sagnella GA, et al.: Association between the C825T polymorphism of the G protein beta3-subunit gene and hypertension in blacks. *Hypertension* 1999, 34:1193–1196.
- 12. Schork NJ, Chakravarti A, Theil B, et al.: Lack of association between a bi-allelic polymorphism in the adducin gene and blood pressure in Caucasians and African Americans. Am J Hypertens 2000, 13:693–698.
- Ranade K, Hsiung AC, Wu KD, et al.: Lack of evidence for an association between alpha-adducin and blood pressure regulation in Asian populations. Am J Hypertens 2000, 13:704-709.
- 14. Larson N, Hutchinson R, Boerwinkle E: Lack of association of three functional gene variants with hypertension in African Americans. *Hypertension* 2000, 6:12970–12300.
- 15. Risch N: Searching for the genetic determinants in a new millennium. *Nature* 2000, **405**:847–856.
- Krushkal J, Ferrell R, Mockrin S, et al.: Genome-wide linkage analyses of systolic blood pressure using highly discordant siblings. Circulation 1999, 99:1407–1410.
- 17. Xu X, Yang J, Rogus J, et al.: Mapping of a blood pressure quantitative trait locus to chromosome 15q in a Chinese population. Hum Mol Genet 1999, 8:2551–2555.
- Bray MS, Krushkal J, Li L, et al.: Positional genomic analysis identifies the beta2-adrenergic receptor gene as a susceptibility locus for human hypertension. *Circulation* 2000, 101:2877–2822.
- 19. Kotanko P, Binder A, Tasker J, *et al.*: Esssential hypertension in African Caribbean associates with a variant of the beta2 adrenoreceptor. *Hypertension* 1999, **4**:773–776.
- 20.•• Halushka M, Fan J-B, Bentley K, *et al.*: Patterns of singlenucleotide polymorphisms in candidate genes for blood pressure homeostasis. *Nat Genet* 1999, 22:239–247.

This paper presents information on all DNA variation in many hypertension genes in blacks and whites.

- 21. Fischer G, James SA, Roberts IN, et al.: Chromosomal evolution in Saccharomyces. Nature 2000, 405:451-454.
- Zhu X, Cooper R, Adeyemo A, et al.: Linkage and association analysis of angiotensin I-converting enzyme (ACE) gene polymorphisms with ACE activity and blood pressure. *Circulation* 2000, submitted.
- Cargill M, Altshuler D, Ireland J, et al.: Characterization of single-nucleotide polymorphisms in coding regions of human genes. Nature Genet 1999, 22:231–238.
- 24. The FBPP Investigators. The Family Blood Pressure Program (FBPP): design and objectives. Am J Hypertens 2000, in press.
- 25. Relethford JH: Genetics of modern human origins and diversity. Annu Rev Anthropol 1998, 27:1–23.
- Kaufman JS, Cooper RS: Considerations for use of racial/ ethnic classification in etiologic research. Am J Epidemiol 2000, in press.
- 27. Cooper RS: Health and the social status of blacks in the United States. Ann Epidemiol 1993, **3:**137–144.
- 28. Cooper RS, Rotimi CN, Ward R: **The puzzle of hypertension in** African Americans. *Sci American.* 1999, **280**:56–63.
- 29. Darwin C: *The Descent of Man.* Princeton, NJ: Princeton University Press; 1981.