Workshop

Hypertension Genetics, Single Nucleotide Polymorphisms, and the Common Disease: Common Variant Hypothesis

Peter A. Doris

Abstract—The investigation of heritable susceptibility to disease is an effort to associate disease phenotype with underlying genotype. Such genotype: phenotype associations have been demonstrated for a large number of monogenic disorders. The usual strategy has been to use linkage mapping in affected families to identify chromosomal loci from which candidate genes and genotypes can be tested for association with disease. This strategy has not been similarly successful for common heritable disease susceptibilities including hypertension that involve multiple genes and gene-environment interactions. Development of extensive collections of single nucleotide polymorphisms (SNPs) raises the possibility that these SNPs can be used as markers in genome-wide association mapping studies to identify hypertension susceptibility loci. In this approach, large numbers of markers are typed in cases and controls with the expectation that markers interrogating SNPs that are involved in inheritance of disease susceptibility will emerge through their association with this trait in the affected population. Essential hypertension is a common disorder. The term “common” has 2 implications: first, that the disease is prevalent; and, second, that it is widespread. Such frequency and distribution characteristics could arise if the susceptibility alleles for hypertension were prevalent in the founding population of contemporary human beings and became distributed with human global dispersal. This common disease: common variant concept is attractive because it suggests that the genetic heterogeneity underlying hypertension susceptibility could be relatively small. It also allows the possibility that nonrandom association of alleles (linkage disequilibrium, LD) can be used to reduce the number of SNP markers required to identify disease susceptibility alleles because a single marker can act as a surrogate for variation flanking it. The influence of a number of important factors on the detectability of hypertension susceptibility alleles by SNP mapping approaches is not yet fully defined. These factors include the locus and allelic diversity of hypertension, the weaker relationship (compared with Mendelian traits) between genotype and phenotype, the accuracy of high throughput genotyping techniques, the extensive role of nongenetic factors, and the extent and heterogenous nature of LD across the genome. (Hypertension. 2002;39[part 2]:323-331.)

Key Words: hypertension, essential □ genes □ epidemiology □ polymorphism

Considerable progress has been made in understanding and characterizing variation in the human genome. There has also been progress in uncovering the role of evolution, recombination and mutation in shaping existing human genetic variation, in defining the genetic history of contemporary humans and in recognizing the possible relationship between ancestral variants in genes and common, polygenic diseases. This progress has lead to the common disease: common variant hypothesis which speculates that the gene variation underlying susceptibility to common heritable diseases existed within the founding population of contemporary humans.1–3 As this population expanded outwards from Africa, the ancient gene variation of these founders has been distributed globally. The disease susceptibility alleles included among this variation have persisted at moderate frequency presumably because they have been selectively neutral, at least until the recent emergence of technological advancements creating the environment required for disease manifestation. The contrasting notion is that disease susceptibility arose independently (and presumably as a result of geographically distinct sets of gene variants) in various dispersed populations. The former hypothesis is attractive because environmental changes that occurred as technology advanced are associated with the emergence of essential hypertension regardless of the geographic distribution of a population. This characteristic is consonant with the occurrence of a shared, preexisting set of genetic predispositions in the global population. It is also attractive because it proposes a susceptibility model that is tractable to investigation, or will soon be as the resources required rapidly become available.

A dense collection of genetic markers mapped across the genome may allow sequence variants causing common disease to be directly identified. Most stable variation in the genome occurs in the form of single nucleotide polymorphism. Single-nucleotide polymorphisms (SNPs) represent about 90% of the common variation in the genome (insertion/
deletion polymorphisms and variable repeat elements providing the rest). Common SNPs, by definition, have a minor allele frequency greater than 1%. This variation arises through a single mutation event in the history of the population. The likelihood of recurrent mutation at the same site is low. Consequently, SNPs are stable. They are di-allelic, the 2 alleles represent the “ancestral” or “wild-type” and the variant form.

The human genome comprises between 3 and 3.5 billion bases. Of these bases, ~3% encode genes and their adjacent regulatory sequence. It is likely that the genetic variation relevant to hypertension susceptibility lies in these coding and closely associated regulatory regions. Common sequence variants occur in about 1 in every 1000 bases of coding/regulatory sequence.4–7 If inherited hypertension susceptibility is due to common coding/regulatory sequence variants, the relevant variation must lie among the ~100,000 such variants. Several large-scale projects to identify and catalog human SNP variation have begun. Most often, SNPs are identified by electronic “mining” of sequence data. Alignment of sequence data from different individuals can reveal stretches of similar sequence that are of high sequencing quality and which differ at single bases. This work has resulted in the recent creation of a database containing 1.5 million SNPs.8 About 60,000 of these lie in coding and adjacent regulatory regions. To what extent they represent rare, rather than common, SNPs is not directly known. In the common disease:common variant model, SNPs that are rare or exclusive to a subpopulation will have no value because they cannot account for the distribution or prevalence of disease.

The advancing availability of tools to exploit sequence variation to identify polymorphisms contributing to disease susceptibility creates new opportunities for progress in hypertension research. This paper will outline some of the opportunities provided by SNP-based mapping studies to characterize the genetic contribution to hypertension susceptibility. It will describe important areas of recent progress that may soon make these studies feasible and will discuss some of the obstacles that must be considered if this new approach to resolving the heritable elements of essential hypertension is to succeed.

Human Evolutionary Genetics

The common disease:common variant hypothesis draws on growing, but incomplete, insights into human evolution and the history of human global dispersion. This hypothesis is supported by evidence that the extant global population is the progeny of a recent (~100,000 years) dispersal of a single sub-Saharan African founding population of relatively small size that replaced all other hominids. The strongest genetic evidence in favor of this recent African origin/replacement (RAOR) scenario arises from studies of sequence variation in the mitochondrial DNA (mtDNA) and in the non-recombining region of the Y chromosome (NRY).

NRY polymorphisms have provided a rich resource to unravel the relatedness of males in various geographically dispersed groups comprising the global population.9 Because this male-specific chromosomal region is not subject to recombination, its accumulated mutations catalog a history of genetic relationship among existing males. If the contemporary population is largely the result of the simple tree and branch phylogeny that the RAOR model proposes, then each branch point represents a bifurcation in which preceding NRY variation is transmitted and onto which new branch-specific variation accumulates. Because the rate of mutation is relatively constant (about 10⁻⁸ mutations per nucleotide per generation9,10), NRY variation provides an opportunity to describe population dispersal and estimate the approximate time and geographic origin of ancestors common to all human males. Investigation of NRY variation indicates a single phylogenetic tree down which contemporary humans are descended from common ancestors in Africa about 35 to 200,000 years (~2000 to 10,000 generations) ago.9,11 Mitochondrial DNA reflects matrilineal inheritance. Variation patterns in mtDNA result in a set of conclusions similar to those derived from NRY.12,13 Because these 2 sources reflect separate genetic lineages, concordance between them has provided high credibility for the RAOR model of contemporary human origins.

In marked contrast to the cohesive picture which has emerged from mtDNA and NRY studies, a major debate in anthropology centers on whether the modern human species (Homo sapiens sapiens) emerged in Africa recently and completely replaced other contemporaneous humans or whether it is the product of multiregional evolution occurring among a polytypic species widely distributed in the Old World.14–16 The multiregional model accommodates intermingling of modern with archaic humans and implies a much greater genetic diversity and a much more complex phylogeny than the RAOR model. Intermediate models can also be conceived in which major gene flows out of Africa became globally distributed, but also mingled with existing, dispersed populations.

Principal arguments countering the RAOR theory center on whether the simple branching phylogenetic hierarchy supported by mtDNA and NRY data may also be congruent with the multiregional model (see Relethford17 for a review). There is also evidence that the deepest known mtDNA lineage (from ancient DNA extracted from a ~60,000-year-old fossil of a morphologically “modern” human) exists in Australia.18 Such ancient lineages with this geographic distribution are not compatible with the RAOR model. Anthropomorphogenic evidence also suggests gene flow between archaic and modern humans.14 Finally, in contrast with evidence from NRY and mtDNA, analysis of genetic diversity in several autosomal loci indicates an origin of modern humans that is more complex than can be explained by the RAOR model.19,20

Because older populations would have longer to accumulate mutations than younger ones, the greater degree of genetic diversity which exists in sub-Saharan Africans strengthens the RAOR model.21 Estimates of the effective size of the ancestral population from which contemporary humans are descended indicate a number of about 10,000.22,23 This small number is simpler to accommodate under the narrow RAOR phylogenetic tree than the multiregional model. It is also a number which bodes favorably for the
common disease:common variant hypothesis because recent expansion of the human population to its current size of near 6 billion from a small group of founders implies substantial sharing of alleles, including those responsible for common diseases, among living humans.

Although there has been progress in describing human genetic demography, considerable complexity in the genetic origins of modern humans is conceivable. Furthermore, the tools to examine this complexity are capable of only sampling a miniscule part of the data, much of which, by virtue of its historic nature, is lost to study. In spite of the attractiveness of the RAO model, it is premature to conclude that the common disease:common variant hypothesis is fully supported by existing insight into human evolutionary genetics.

**Linkage Disequilibrium, Recombination, and Gene Conversion**

Human hypertension loci have been mapped by several linkage analysis studies. Linkage analysis uses collections of related individuals with members who manifest a trait (eg, hypertension, elevated blood pressure) to examine the co-inheritance with the trait of widely distributed markers in order to infer the genomic position of alleles contributing to the trait. In contrast, the genome-wide association studies which are now being contemplated examine populations, not related individuals, and compare the genotypes of a large number of polymorphic markers in subjects who manifest a trait (eg, hypertension cases) compared with controls. In order to directly identify the polymorphisms contributing to disease susceptibility very large collections of markers will be required to ensure that markers interrogating the disease-causing variants are represented. Genotyping many individuals for ≈100 000 SNPs is a daunting and expensive task. However, such genome-wide association studies may be made more feasible if the number of markers to be typed can be reduced by exploiting the preservation of nonrandom association between a polymorphisms contributing to a trait and adjacent markers (linkage disequilibrium).

Linkage disequilibrium (LD) is the increased sharing of alleles above that expected by random sharing. If a locus A, contains 2 alleles, A1 and A2, that are equally frequent while another locus Z, on the same chromosome, contains 2 alleles, Z1 and Z2, also equally frequent, then, assuming the independence of these loci, the haplotypes A1Z1, A1Z2, A2Z1, and A2Z2 should occur with equal frequency. However, if the mutation that created the Z2 allele occurred on a chromosome that also contained the A1 allele, then the A1Z2 haplotype originated as a nonrandom feature. This haplotypic association will tend to be preserved as the chromosome containing it is transmitted to subsequent generations. However, it can be interrupted by a recombination event resulting in crossing over. The likelihood of interruption will decrease if locus A and locus Z are close to one another because the chances of a recombination event separating them are low.

Linkage disequilibrium reflects human population history because its magnitude is influenced by recombination, by the historical size and pattern of expansion and contraction of the population, by migration followed by admixture, by selection, and by random drift. The magnitude and positional pattern of distribution of LD in the human population influences the total number of test markers required to identify disease loci in LD association studies. If LD is well preserved, then a single marker may be a reliable reporter for extensive adjacent polymorphism. However, if it is low, little or no reduction in marker density may be possible. If it is a positionally heterogenous mixture of regions of high and low LD, then successful exploitation of LD will require a modified approach encompassing greater knowledge of the regional pattern and extent of LD distribution across the genome.

Estimation of the extent of LD in the human population has been obtained by sampling LD in various limited genomic regions. The average distance in which useful LD is preserved (eg, the distance over which 50% of pair-wise markers demonstrate association) is ≈50 to 60 kilobases (kb). With a uniformly distributed marker set, LD of this magnitude means that somewhere around 50 000 markers per subject would be required to identify disease alleles using genome-wide LD association. However, in some chromosomal loci, LD extends as far as 500 kb, whereas in other loci no useful disequilibrium exists. Because linkage disequilibrium is highly variable across the genome the use of average distances to direct marker selection means that disease susceptibility variation residing in loci containing low LD would usually not be identified.

Disruption of allelic association by crossing over during recombination is not the only mechanism that can erode LD. Gene conversion is an additional meiotic process of sequence alteration in which one allele directs the conversion of a partner allele to its own form. The Holliday 4-way DNA recombination junction is the physical element by which both crossing over and gene conversion occur. Gene conversion may be explained by repair of heteroduplex DNA formed in short tracts (up to 1 kb) adjacent to the Holliday junction. Gene conversion represents a further problem for association mapping which seeks to exploit LD. Gene conversion disrupts LD over short distances.

Thus, chromosomal regions which, on the basis of allelic association over long range, appear to retain high levels of LD, may in fact harbor short regions where LD has been disrupted by gene conversion. A recent study by Ardlie and colleagues reports that LD between tightly linked markers in humans is much less than expected. A plausible explanation for this observation is that LD has been reduced by gene conversion. This phenomenon represents an important impediment to LD mapping studies because LD markers selected to survey extensive regions of LD that also harbor regions affected by gene conversion may fail to detect allelic association.

Evenly distributed SNP markers do not exploit fully the advantage of existing knowledge of the structure of genetic variation. Sequence variation in genes results in the existence within the population of multiple forms of genes (gene haplotypes) that represent specific combinations of variation in a gene. The haplotype structure of 319 human genes sampled from a collection of globally diverse subjects has recently been reported. More than 4000 haplotypes of these 319 genes were identified. Only 16% of these haplotypes (≈650) were common to the diverse populations represented...
in the sample. Although these “cosmopolitan” haplotypes represent a small portion of the total range of haplotypes encountered, they are highly frequent and account for over 80% of the total haplotypes observed in the subjects. Therefore, a little more than 2 markers per gene (650/319) may capture the large majority of common haplotypes in these genes. The selection of markers to represent not only genes, but also the common haplotypes of genes provides a more powerful approach to detect disease association. However, existing knowledge of the common haplotypes of human genes is very limited.

**Genotyping Technology**

Even in the most optimistic appraisal, it is clear that genome-wide disease association studies will require very large amounts of genotyping. The use of thousands of markers per individual shifts the cost of such studies toward those associated with genotyping. Typing of SNP markers is made less complex and more automatable because of their diallelic nature. A large effort has been invested to develop high throughput SNP genotyping technologies in anticipation of genome-wide association studies using these markers. The emerging technologies that are likely to be successful are those that combine low cost with accuracy. Enzyme reagents remain the single largest cost component but can be reduced to some degree by reduction of reaction volumes and by amplifying and analyzing multiple SNPs in single reactions (multiplexing). However, these economies require detection techniques that are highly sensitive, that are capable of resolving multiplexed genotyping reactions without signal cross talk, and that retain their accuracy while sensitivity is exploited. At present, the cost to call a single SNP genotype using high throughput, commercialized technologies is between $0.50 and $1. To embark on a project requiring typing of tens of thousands of markers in many individuals requires both high levels of confidence and financial support.

Accurate genotyping is essential in linkage disequilibrium mapping studies. Array-based hybridization formats tend to have high, parallel throughput, but poor accuracy. Other methods employ mini-sequencing reactions in which the SNP is typed by annealing an oligonucleotide to a PCR product containing a SNP and extending the oligo over the polymorphic base and determining which base is incorporated at the SNP. Mass spectrometry provides an attractive approach to genotyping by oligo extension because of the extremely high sensitivity of the technique and the possibility to perform multiple reactions (both PCR amplification of template and extension reactions) in a single tube. The exonuclease/fluorescence resonance energy transfer method (Taq-Man, LightCycler) requires no post-PCR processing, but multiplexing and scale-down potential is limited by fluorescence detection constraints.

At present, there is very little independent assessment of the accuracy of most SNP genotyping techniques. Most of the publicly available analyses of accuracy have been generated by the companies that have developed commercial implementations of SNP genotyping technologies. Accuracy can be lost either by random miscalculation or as the result of nonrandom errors that cause one allele to be consistently over- or under-called.

Errors in genotyping breakdown allelic association and consequently reduce the power to detect direct or LD association with disease. The prospect of high throughput SNP genotyping is on the horizon, but important issues of cost and accuracy remain to be fully defined.

**Heterogeneity**

The most optimistic view of the genetics of essential hypertension is that a small number of genes contained polymorphisms capable of increasing susceptibility to hypertension in a recent (~100 000 years), small (~10 000 individuals), founding population that is now globally dispersed and that completely replaced preceding human forms. These polymorphisms have been selectively neutral (at least until recently) and remain common in essentially all contemporary subpopulations. Recent changes in environment that have accompanied the emergence of technologically advanced cultures have afforded contemporary humans a sedentary lifestyle with high sodium and calorie intake. These and other environmental changes may have provided the necessary interactions to allow preexisting susceptibility polymorphisms to present the hypertension phenotype.

The history of genetic disease, however, is that such relative simplicity is far from usual. Mendelian disorders can have extraordinary levels of locus and allelic heterogeneity. A good example of allelic heterogeneity is provided by the several hundreds of rare allelic variants of the low density lipoprotein receptor that are associated with familial hypercholesterolemia, whereas none of the common variants of this gene are associated with disease. Retinitis pigmentosa and its associated forms of hereditary retinal blindness (macular degeneration and Usher syndrome) provide an extreme example of locus heterogeneity in which over 100 loci can create a similar disease phenotype.

The genetic architecture of blood pressure homeostasis also suggests that a simple view of hypertension genetics may be optimistic. Blood pressure is maintained by a well-characterized, complex network of integrated and redundant systems including renal, neuronal, endocrine, and vascular mechanisms. Within each of these systems multiple genes are expressed to contribute to the specialized functions by which blood pressure regulation is effected. At present, it is not known whether inheritance of hypertension susceptibility in humans is attributable to variation in a narrow subset of these genes, or whether large numbers of genes can contribute to hypertension. Several Mendelian forms of hypertension have now been solved (Table 1). Mendelian hypertension in humans involves at least 10 different genes. Numerous mouse knockout strains have been produced in which the targeted gene deletion produces elevated blood pressure or increases susceptibility to elevated blood pressure in the presence of environmental factors (i.e., nongenetic influences such as elevated sodium intake, reduced renal mass, etc.). The diversity of such genes (Table 2) again accents the possibly large number of genes that possess alleles that may affect blood pressure in the human population and emphasizes that the degree of genetic heterogeneity in susceptibility to hypertension in humans may also be high.
Bardet-Biedl Syndrome Types 2 and 4 with Severe Exacerbation in Pregnancy

Early Onset, Autosomal Dominant Hypertension

WNK1

Pseudohypoaldosteronism. Type II

Syndrome

Pseudohypoaldosteronism. Type I, Liddle's Syndrome

17-alpha-hydroxylase deficiency

Mineralocorticoid receptor gene

BBS2 and BBS4

At the opposite, also unlikely extreme, the most pessimistic view of the genetics of hypertension is that common variants play no important role in this disease and that it is the result of many rare mutations (high allelic heterogeneity) in many of the large number of genes which participate in normal mechanisms of blood pressure regulation (high locus heterogeneity). An important role for rare variants and high allelic heterogeneity in complex diseases is the conclusion of a recent modeling study by Pritchard.46 This model proposes that the common presumption that selection against complex disease variations is so weak that it can be ignored is incorrect because such selectively neutral variations are likely to have drifted to fixation or elimination. Consequently, it proceeds to examine the implications of weak purifying selection against susceptibility alleles. The main conclusion is that most genetic variance underlying complex disease is likely attributable to loci where susceptibility mutations are mildly deleterious and where the overall mutation rate (and allelic heterogeneity) is relatively high.

Modeling studies of the role of common or rare variants in complex disease are shaped by parameter values about which there is little direct knowledge. However, it is important to view them as a guide to indicate what range of possibility must be considered in tackling the challenge of complex disease genetics. There is little basis to know at which point between the 2 extreme examples given above the real genetic susceptibility alleles. The task of answering this question falls to further

### TABLE 1. Mendelian Forms of Hypertension in Humans

<table>
<thead>
<tr>
<th>Syndrome</th>
<th>Gene(s)</th>
<th>Normal Gene Function/Disease Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apparent mineralocorticoid excess</td>
<td>11-beta-Ketoreductase</td>
<td>Loss of the normal action of this gene allows activation of renal mineralocorticoid receptors by cortisol.</td>
</tr>
<tr>
<td>Glucocorticoid-remediable aldosteronism</td>
<td>CYP11B2 and CYP11B1</td>
<td>Fusion of the CYP11B2 and CYP11B1 genes places aldosterone production under control of ACTH.</td>
</tr>
<tr>
<td>17-alpha hydroxylase deficiency</td>
<td>17-alpha-hydroxylase</td>
<td>17-hydroxylase is necessary for both cortisol synthesis. Deficiency results in increased in ACTH. Excessive synthesis of deoxycorticosterone and corticosterone produces hypertension.</td>
</tr>
<tr>
<td>Pseudohypoaldosteronism. Type I, Liddle’s Syndrome</td>
<td>SCNN1B and SCNN1G</td>
<td>Excessive renal sodium reabsorption arises from constitutive activation of the renal epithelial sodium channel because of truncation mutations in either the beta subunit or the gamma subunit of this channel.</td>
</tr>
<tr>
<td>Pseudohypoaldosteronism. Type II</td>
<td>WNK1 and WNK4</td>
<td>These genes encode proteins that may affect paracellular ion reabsorption in the distal nephron.</td>
</tr>
<tr>
<td>Early Onset, Autosomal Dominant Hypertension with Severe Exacerbation in Pregnancy</td>
<td>Mineralocorticoid receptor gene</td>
<td>Mechanism incompletely defined.</td>
</tr>
<tr>
<td>Bardet-Biedl Syndrome Types 2 and 4</td>
<td>BBS2 and BBS4</td>
<td>Affected genes are novel chaperonins, mechanism of disease unknown.</td>
</tr>
</tbody>
</table>

### TABLE 2. Targeted Gene Deletion and Blood Pressure in Mice

<table>
<thead>
<tr>
<th>Gene (Symbol)</th>
<th>Gene-Environment Interaction</th>
<th>Blood Pressure Phenotype Compared to +/+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endothelial nitric oxide synthase (Nos3)</td>
<td></td>
<td>-/- = +18 mm Hg systolic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+/- = +4 mm Hg systolic</td>
</tr>
<tr>
<td>Insulin receptor substrate (irs1)</td>
<td></td>
<td>-/- = +13 mm Hg systolic</td>
</tr>
<tr>
<td>Dopamine receptor (Drd3)</td>
<td></td>
<td>-/- = +23 mm Hg systolic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+/- = +20 mm Hg systolic</td>
</tr>
<tr>
<td>11-beta-hydroxysteroid dehydrogenase type II, (Hsd11b2)</td>
<td></td>
<td>-/- = +25 mm Hg mean</td>
</tr>
<tr>
<td>Apolipoprotein E (Apoe)</td>
<td></td>
<td>-/- = +22 mm Hg systolic</td>
</tr>
<tr>
<td>Beta-1 subunit of calcium-dependent potassium channel (Kcnmb1)</td>
<td></td>
<td>-/- = +20 mm Hg mean</td>
</tr>
<tr>
<td>Natriuretic peptide receptor/Guanylyl cyclase-A (Npr1)</td>
<td></td>
<td>-/- = +34–39 mm Hg systolic</td>
</tr>
<tr>
<td>Angiotensin type 2 receptor (Apt2)</td>
<td></td>
<td>Enhanced hypertension with DOCA/salt</td>
</tr>
<tr>
<td>Glomerular epithelial protein 1/protein tyrosine phosphatase type O (Ptpro)</td>
<td></td>
<td>Hypertension related to concomitant renal ablation</td>
</tr>
<tr>
<td>Bradykinin receptor type 2 (Bdkrb2)</td>
<td></td>
<td>Dietary sodium loading</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-/- = -20 mm Hg mean</td>
</tr>
<tr>
<td>Dopamine and cyclic AMP regulated phosphoprotein (Ppp1r18)</td>
<td></td>
<td>-/- = +12 mm Hg mean</td>
</tr>
<tr>
<td>Cytochrome p450 monooxygenase 4A14 (Cyp4a14)</td>
<td></td>
<td>Males -/- = +30 mm Hg systolic, +30 mm Hg mean, and +25 mm Hg diastolic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females -/- = +17 mm Hg systolic, +14 mm Hg mean, and +11 mm Hg diastolic</td>
</tr>
<tr>
<td>Heme oxygenase 1 (Hmox1)</td>
<td></td>
<td>1 kidney, 1 clip</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-/- = +24 mm Hg systolic</td>
</tr>
</tbody>
</table>
characterization of human genetic variation, to careful, well-designed experimentation, and to cautious interpretation. A reasonable approach to try to disprove the null hypothesis that common hypertension variants can be identified by association mapping is to focus such studies on hypertension chromosomal loci identified by linkage mapping.

**Problems in the Genotype-Phenotype Relationship**
Progress in complex disease genetics has been hampered by the low detectance power of mapping approaches. This problem reflects heterogeneity in both genetic and environmental influences. One response to this problem is to classify the affected population according to intermediate phenotypes that may indicate shared underlying genetic mechanisms. The identification of such phenotypes, along with efforts to understand the biology underlying them, offers the important benefit of reducing the potential confounding effects of genetic heterogeneity and variable environmental influences. This approach can yield another important benefit: it increases the likelihood that a relationship between genotype and phenotype will be sufficiently strong to be detected because the path from genotype to intermediate phenotype is presumably more direct than the path to the composite phenotype of elevated blood pressure.

An important component of the low detectance power of complex disease mapping studies may result from differing degrees of genetic participation in young versus old subjects. Studies of coronary artery disease indicate that genetic determination in young individuals is much stronger than in old individuals, an observation that might be explained by increasing participation of nongenetic factors in disease etiology with aging. Systolic hypertension has a strong age association with prevalence levels reported in the National Health and Nutrition Examination Survey (NHANES) III, ranging from about 10% in white and Mexican-American females, aged between 40 and 49 years, to above 60% in the same groups aged between 70 and 79 years. Therefore, it is important to consider that the role of heredity in determining susceptibility to hypertension may also decline with age. Consequently stratification of population genetics studies to consider this possible age-dependent participation of genetic susceptibility may increase the capacity to detect loci involved.

These considerations result in one simple message: successful SNP-based studies will require maximum attention to subject characterization and selection. If a genotype:phenotype relationship is to be found, then much of the background noise from non-genetic factors as can be eliminated should be, and disease classification to reduce genetic heterogeneity should be sought by use of well-defined intermediate phenotypes.

**Insight from Model Systems**
It is reasonable to expect that model organisms can provide some representation of the currently unknown pattern of genetic diversity involved in complex quantitative human traits. *Drosophila melanogaster* has been used to define the genetic basis of quantitative traits such as abdominal and sternopleural neurosensory bristle number and adult longevity. The inbred laboratory strains of drosophila used in these studies probably under-represent the allelic and locus diversity determining these traits in natural populations. In these strains, over 20 quantitative trait loci (QTLs) affect bristle number, a few have large effects, while most QTLs correspond to loci of small effect. A few of these QTLs have been resolved to the level of the gene responsible or contain candidate genes predicted to influence the trait. Among the QTLs identified, sex-linked and epistatic features have been observed. Seventeen QTLs affecting life span have been identified. These QTLs have sex dependency, gene-environment interactions and non-additive gene-gene interactions. These studies emphasize that even presumably simple traits like bristle density can be affected by large numbers of loci and be subject to complicating influences determined by sex and environmental interactions. However, the identification of a role for oligogenes (genes small in number, but having relatively large influence on a complex trait) provides some optimism that oligogenic influences on hypertension susceptibility may also exist to be identified.

**Progress in Identifying Complex Disease Susceptibility Alleles**
Some optimism can also be derived from successful efforts to identify susceptibility alleles for common diseases. Alzheimer’s dementia (AD) is a heterogenous disorder that, like hypertension has both rare Mendelian susceptibility as well as common polygenic inheritance patterns. Linkage mapping studies suggested a susceptibility locus on the proximal long arm of chromosome 19. Apolipoprotein E maps in this region and 1 of the 4 allelic variants of this gene, ApoE epsilon 4, was found to be associated with the disease. Although many other loci contribute to AD susceptibility, APOE e4 is a significant risk factor which increases the odds ratio in heterozygotes and further increases it in homozygotes.

Thus, while possession of the e4 allele of apolipoprotein E is neither necessary nor sufficient to cause AD, it represents a single, well defined and substantiated common genetic risk factor identified by positional approaches and confirmed by association studies.

Hypertension research may be at a similar threshold of progress. Several high quality, genome-wide linkage mapping projects have been completed in a range of subject populations (Table 3) yielding some concordance in the susceptibility loci that have emerged. Positional candidate genes from linked loci have been selected and large-scale association studies are underway. One interesting association between polymorphism in the beta 2 adrenergic receptor, a positional candidate gene present in a locus identified by linkage analysis, and blood pressure level has been confirmed, illustrating the principal that association studies of positional candidate gene sequence variation may yield further progress in defining sequence variation relevant to hypertension.
TABLE 3. Genome-Wide Linkage Surveys for Blood Pressure Loci in Human Populations: Chromosomes for Which Linkage Has Been Identified

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Population(s)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Quebec</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>Amish, Quebec, Finnish, US white</td>
<td>24–27</td>
</tr>
<tr>
<td>3</td>
<td>Chinese</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>Quebec, US white</td>
<td>24,27</td>
</tr>
<tr>
<td>6</td>
<td>US African-American and US white</td>
<td>26,27</td>
</tr>
<tr>
<td>7</td>
<td>Quebec</td>
<td>24</td>
</tr>
<tr>
<td>8</td>
<td>Quebec</td>
<td>24</td>
</tr>
<tr>
<td>11</td>
<td>Chinese, Unspecified</td>
<td>29,30</td>
</tr>
<tr>
<td>15</td>
<td>Chinese, US white</td>
<td>27,30</td>
</tr>
<tr>
<td>16</td>
<td>Chinese</td>
<td>30</td>
</tr>
<tr>
<td>17</td>
<td>US white</td>
<td>28</td>
</tr>
<tr>
<td>18</td>
<td>US white</td>
<td>26,28</td>
</tr>
<tr>
<td>19</td>
<td>Quebec</td>
<td>24</td>
</tr>
<tr>
<td>22</td>
<td>Finnish</td>
<td>25</td>
</tr>
<tr>
<td>X</td>
<td>Finnish</td>
<td>25</td>
</tr>
</tbody>
</table>

Independence of Inheritance of Hypertension Susceptibility and Susceptibility to Injury from Hypertension

Hypertension is a key determinant of risk for morbidity and mortality for a number of diseases including myocardial hypertrophy, infarction, heart failure, stroke and renovascular disease (adverse outcomes of hypertension, AOH). However, increasing evidence suggests that susceptibility to these outcomes may be influenced by inheritance and that these susceptibilities are separate and distinct from the inheritance of susceptibility to hypertension. Animal models of genetic hypertrophy have been developed that have resulted in the fixation of alleles that confer resistance to the AOH. These animal models may offer insights into the genetic basis of susceptibility to AOH in humans. Such insights are difficult to obtain directly because the disease phenotypes in humans (stroke, heart failure, etc.) are frequently late-onset, life-shortening outcomes that limit access by the investigator to materials from the affected subjects and from affected subjects in their families.

The clearest example of an animal model of genetic susceptibility to AOH is the spontaneously hypertensive rat (SHR) in which both stroke-resistant and stroke-susceptible strains have been identified.63 In these strains the level of susceptibility to AOH is the spontaneously hypertensive rat. Other examples of genetically influenced susceptibility to AOH include the fawn-hooded rat which demonstrates heritable susceptibility to renovascular injury in hypertension.72,73 A locus on rat chromosome 1, which influences susceptibility to renovascular injury in the fawn-hooded rat, has been identified, and this locus has been confirmed in SHR using congenic approaches.74

Work to map the loci that confer a risk of AOH in humans lags behind the mapping of hypertension loci. However, it is possible that genetic factors determining such heritability may reflect a less complex genetic architecture and so may be more readily identified and confirmed by genome-wide and LD association studies.

Conclusions

The scientific and technical resources to systematically pursue the identification of gene variation that causes heritable susceptibility to hypertension using genome-wide SNP association studies are maturing. These include the availability of a large number of SNP markers, tools for high throughput, rapid and low cost genotyping of these markers and definition of the extent and nature of LD in the human population. Current knowledge suggests that identification of the common haplotypes of human genes and the use of haplotype, in favor of single marker, association mapping may avert some of the important obstacles to progress resulting from discontinuity in the pattern of LD. Although such work is not yet feasible on a genome-wide scale, the development of haplotype information in genes mapping to hypertension linkage loci may provide a conservative strategy for determining the feasibility of using association approaches to achieve fine mapping and disease susceptibility gene identification.

Acknowledgment

Work in the author’s laboratory is supported by National Institutes of Health grant DK45538.

References


12. Penny D, Steel M, Waddell PJ, Hendy MD. Improved analyses of human 

11. Hammer MF. A recent common ancestry for human Y chromosomes. 


The low density lipoprotein receptor (LDLR) gene in familial hypercholesterolemia. http://www.ucl.ac.uk/lhr/. 


Nuzhdin SV, Dilda CL, Mackay TF. The genetic architecture of selection response. Inferences from fine-scale mapping of bristle number quanti-


Hypertension Genetics, Single Nucleotide Polymorphisms, and the Common Disease: Common Variant Hypothesis
Peter A. Doris

Hypertension. 2002;39:323-331
doi: 10.1161/hy0202.104087

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://hyper.ahajournals.org/content/39/2/323

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Hypertension can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Hypertension is online at:
http://hyper.ahajournals.org//subscriptions/