European Society of Hypertension Scientific Newsletter:
Update on Hypertension Management
2008; 9: No. 33

DISCOVERING THE GENETIC DETERMINANTS OF HYPERTENSION

Sandosh Padmanabhan[1], Olle Melander[2], Claire Hastie[1], Christian Delles[1], Anna F. Dominiczak[1]
[1]BHF Glasgow Cardiovascular Research Centre, University of Glasgow, Glasgow, United Kingdom
[2]University Hospital Malmö, Lund University, Malmö, Sweden

Hypertension is a common complex disease which arises as a result of the interplay between multiple genetic and environmental factors. The determination of the genetic variants involved in hypertension should provide new insight into susceptibility to the disease, its progression and severity, leading to novel pharmaceutical targets, with the ultimate goal of improving prevention, diagnosis and treatment [1, 2]. There is an inverse relationship between the magnitude of genetic effect and allele frequency, indicating that few variants of clinical consequence will be common [3]. Genetic variants with large effect sizes manifest as Mendelian or single-gene disorders with mutations in at least 10 genes known to cause hypertension or hypotension, primarily by affecting renal tubular electrolyte transport functions [4, 5]. However, these rare alleles account for less than 1% of human hypertension and have not been associated with common forms. Evidence from most replicated associations of other complex diseases indicate that the effect sizes associated with hypertension will be less than 1.5 for each causative variant. While individually these effect sizes are minor, the combination of even a few common polymorphisms imparting small relative risks can have substantial population attributable risks (an estimate of the percentage of cases of disease that would be avoided if the exposure were removed, combining information about the relative risk and the prevalence of the genetic variant). Therefore, the downstream functions of such common genetic variants will become potential targets for lifestyle and pharmacological interventions.

Gene mapping studies in hypertension

Linkage and candidate gene analysis of hypertension have produced inconclusive results so far, but there are valid arguments that favour gene mapping of hypertension in large populations. These have been borne out by the recent successes in genome-wide association (GWA) studies of other complex diseases [6–8]. Linkage mapping is a method of studying genetic markers of known chromosomal location that are co-inherited with the disease in a pedigree. It is a powerful tool for finding Mendelian disease genes, but has produced weak, and inconsistent, signals in complex disease studies. Association mapping, commonly used in candidate gene studies, is based on linkage disequilibrium (LD). LD is the non-random association of alleles at two or more loci on a chromosome and results in the greater co-occurrence of two genetic markers in a population than would be expected for independent markers (Figure 1). In practical terms, LD results in single nucleotide polymorphisms (SNPs) that are in close proximity and travel together in a block when passed from parent to child, allowing one SNP in a block to serve as a surrogate for the other SNPs in the block, thus obviating the need for testing all the SNPs individually. A new mutation that arises within a block travels along with the members of the block for hundreds of generations. In short, linkage measures the co-segregation between a genetic marker and a disease affection status in a pedigree as a result of meiotic recombination events in the last two or three generations, while LD measures co-segregation in a population (a very large pedigree extending back to the founders), resulting from much earlier ancestral meiotic recombination events.

Linkage analysis

The genome-wide linkage approach with microsatellites has shown evidence for the existence of several chromosomal regions that are linked to blood pressure or essential hypertension on almost all chromosomes [9–12]. Interestingly, some human chromosomes — 1, 2, 3, 17, 18 — contain multiple linkage loci for hypertension, often with overlapping confidence intervals. This suggests that hypertension may be linked to fewer genomic regions with no single genomic region having a uniformly large effect on predisposition to hypertension. However, linkage mapping has been difficult because of the polygenic nature of hypertension, possibly involving multiple pleiotropic variants of low penetrance, epistasis, the ethnic diversity of human populations, phenotypic heterogeneity and the inability to control for environmental factors. In addition, linkage analysis has poor power for detecting common alleles that have low penetrance. All these factors hinder the replication of results and decrease the likelihood of success in linkage studies in general. The Medical Research Council funded British Genetics of Hypertension (MRC BRIGHT) study [9] attempted to increase detectance (the probability of any particular susceptibility genotype being carried given that the individual has a particular phenotype) of the disease locus by ascertaining disease severity and including only non-diabetic, non-obese hypertensives, whose blood pressure was in the top 5% of blood pressure distribution in the UK. However, this study was able to identify only a single locus suggestive of linkage on chromosome 5q [13]. Subsequently we demonstrated a successful attempt at reducing heterogeneity by using antihypertensive drug response to partition different pathways of hypertension [11]. In the BRIGHT population, hypertensive sib-pairs who were non-responsive to ACE inhibitors, ARBs or β-blockers showed significant linkage on chromosome 2p (LOD = 4.84 at 90.68 Kosambi cM) [11]. This susceptibility locus co-localises to a region found in African-American hypertensives in the Family Blood Pressure Program who showed evidence of linkage with hypertension status at 93 cM with a LOD score of 2.84 [14]. Thus the chromosomal 2p locus independently identified in different populations may contain a gene or genes for the salt-sensitive form of hypertension, which is common among Africans, and the same mechanism may be operative in a subset of white European hypertensives identified by unresponsiveness to β-blockers and ACE inhibitors.

Candidate gene analysis

SNPs are the most common genetic variant studied in candidate gene studies. SNPs are DNA sequence variation, occurring when a single nucleotide in the sequence is altered. A variation must occur in at least 1% of the population to be considered an SNP. In the human genome consisting of three billion nucleotide bases SNPs occur at a frequency of 1 in every 1000 bases. As the coding region comprises only 5% of the genome, most SNPs will be found in regions that do not affect protein structure. Coding region SNPs usually tend to be synonymous (coding for the same protein without any change in amino acid sequence), and thus most of the common SNPs might not be so informative. Amino acid-altering non-synonymous coding-region SNPs are rare and harder to find because of expected selection against them in human evolution.

In the hypertension disease model the candidate genes fall into five broad categories: the renin-angiotensin-aldosterone system, the adrenergic system, vascular-related genes, metabolism-related genes and those with other miscellaneous roles. The selection of candidate genes has generally been based on a mechanistic understanding of the roles of the encoded proteins in blood pressure regulation. Novel genetic pathways that may determine familial susceptibility to hypertension include loci that encode growth factors, molecules of oxidative stress and inflammatory response [15]. However, this strategy has not been very successful, and the failure of these studies to identify the genetic basis of the common forms of hypertension suggests that there are several limitations to this approach. First, the choice of candidate genes may be inappropriate. Second, the genes in which heritable variations that affect blood pressure occur might be involved in events that take place either upstream of the points of action of the selected candidates or in the downstream signalling events. Finally, candidate gene studies rely on

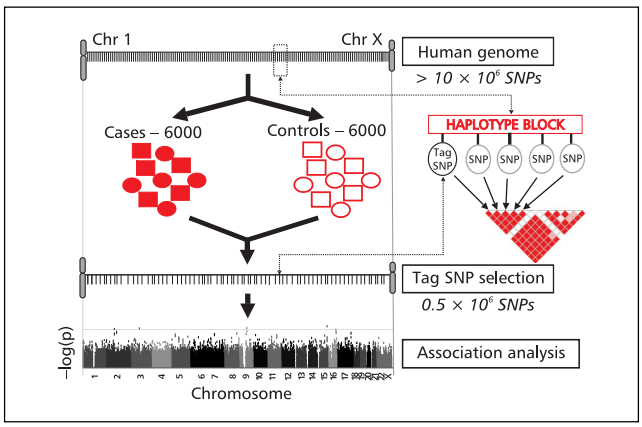

Figure 1. Flow plan for genome wide association study for hypertension using linkage disequilibrium and tag SNPs

prior hypotheses about disease mechanisms, so that discovery of genetic variants in previously unknown pathways is precluded.

Genome Wide Association Study (Figure 1)

GWA studies are large-scale association mapping using SNPs that make no assumptions about the genomic location or function of the causal variant and are a comprehensive approach to testing the hypothesis that common alleles contribute to heritable phenotype variation [16]. Although it is not yet technically feasible to resequence every base on the genome or genotype the ~11 million currently known SNPs, the availability of genome-wide SNP arrays that genotype 100,000 to 1,000,000 SNPs per sample and the availability of LD patterns on a genome-wide scale make these studies possible. Such studies typically measure sets of special DNA tag SNPs selected from the catalogue of common human genetic variations provided by the HapMap project [17], enriched with non-synonymous SNPs, as well as SNPs in evolutionarily conserved regions of the genome. The recent surge in GWA studies for complex diseases with replicated results of SNP association shows that complex diseases are finally yielding their secrets. Most of these studies are case-control studies, where SNP frequencies are compared between the two groups, and those that differ significantly are then validated in independent samples. Despite these successes, the Wellcome Trust Case Control Consortium (WTCCC) study showed an absence of any positive signals for hypertension using 2000 cases and 3000 common population controls [6]. Nor did the study detect any genes previously implicated by candidate gene association studies. Among the reasons for this result are poor coverage of genes by the Affymetrix chip, inadequate power to detect genotypic relative risks < 1.3, phenotypic heterogeneity among the cases and misclassification of unphenotyped controls (*see below*). It is pertinent to point out that the SNP chips capture a substantial proportion of the known common variations, but not rare variants and non-SNP variants such as deletions and insertions.

Study Design Issues

To succeed in finding complex disease genes a study must detect a relatively weak statistical signal, and hence study design flaws can potentially have a dramatic impact on the probability of success [18]. The key factor which determines the success of a given study design is the allelic architecture, which refers to the number of distinct alleles that impact disease susceptibility at a given disease locus, their frequencies and penetrances. Penetrance is the likelihood, or probability, that a particular genotype will be expressed in the phenotype.

Phenotyping of cases and controls

Most of the successful replicated association study results are in diseases where phenotyping specificity can reduce the source of heterogeneity. Strict criteria to define a phenotype have to apply; in hypertension this may involve ambulatory 24-hour blood pressure measurement as opposed to office blood pressure readings. Intermediate phenotypes such as endothelial function and vascular stiffness may help to understand the physiological link between gene variants and established disease. Of note, cases and controls should undergo the same phenotyping protocol to establish both the presence of disease in cases but also the absence of disease in controls. The value of phenotyping controls is all the more evident in the case of the negative results from the WTCCC GWA study of hypertension. The use of common controls in this experiment may have caused some attrition in power because of misclassification of some of the controls. It was estimated that a 5% misclassification of controls would result in a loss of power equivalent to a sample size reduction of 10%. However, the high prevalence of hypertension might have led to a 30% and not a 5% misclassification bias [6].

Although there is a general thrust towards case-control GWA studies, family studies should not be forgotten. These are robust to population stratification and family information is important to refine the genetic model and risk estimates, control for the effects of shared environment and detect heterogeneity, imprinting, and epigenetic phenomena [19].

Sample size and power

At least 6000 cases and 6000 controls are required to have adequate power to detect causative SNPs with a minor allele frequency (MAF) > 10% and a genotypic relative risk < 1.3. One way of increasing power would be to select cases and controls from the extreme ends of the distribution of blood pressure in the population, thus increasing the odds ratio. Choice of significance threshold is important when simultaneously considering many SNPs as in GWA studies. Significance thresholds of $p < 10^{-6}$ have been proposed for GWA to allow for the very small prior probability that any given locus or region is truly associated with disease. Replication studies attempted for SNPs reaching a significance level of 5×10^{-7} in the WTCCC study have all been confirmed [6, 18, 20].

Gene environment interactions

There is growing recognition that environmental and behavioural changes, in interaction with a genetic predisposition, have produced most of the recent increases in chronic diseases including essential hypertension [21, 22]. However, the most important implication of gene-environment interactions is that they can suggest approaches for modifying the effects of deleterious genes by avoiding environmental exposure [22]. Despite much information on both genetic and environmental risk factors, there are relatively few robust, replicated gene-environment interaction studies [23–25].

The successful detection of gene-gene or gene-environment interactions is dependent on recruiting adequate sample sizes in the prospective cohort studies. In contrast to case-control studies, which typically begin when disease cases have occurred, prospective cohort studies avoid this inherent bias by investigating a representative sample of the population before disease onset. In terms of genetic main effects, 5000 incident cases are needed for a minimum detectable odds ratio (MDOR) of 1.5 with 80% power and MAF = 5%. For genotypic and environmental prevalences of 10% and above, 10,000 cases are needed to provide adequate power for interaction effects with an MDOR greater than 2 [22].

The road ahead

It is now appreciated that all genetic studies must be rigorous about the standardisation, precision and accuracy of phenotyping, quality control analysis and interpretation. The challenge is to translate the small effect size into clinically useful diagnostic or therapeutic tools with an impact on public health. There is also growing evidence for heritable changes in gene function without changes in DNA sequence (epigenetics). Future work should focus on the interplay between epigenetic mechanisms and the environment in disease causation. Another type of genetic variation overlooked in all association studies are the structural variations, specifically sub-microscopic rearrangements between 500 bp and 5Mb in size, commonly called copy number variation (CNV). Identification of these validated gene variants should help in the understanding of disease biology, but their relevance to clinical practice and public health will depend on whether they can improve diagnosis, prevention or treatment strategies. All this will need more efficient epidemiological studies enrolling prospective population cohorts with informed consent, along with collection and storage of biological samples, active collaboration between basic scientists and clinicians, development of novel computational and statistical approaches and the incorporation of large-scale genomic, proteomic, metabolomic and epigenetic analyses into a systems biology framework.

References

1. Merikangas KR, Risch N. Genomic priorities and public health. Science 2003; 302: 599–601.
2. Palmer LJ, Cardon LR. Shaking the tree: mapping complex disease genes with linkage disequilibrium. Lancet 2005; 366: 1223–1234.
3. Morton NE. Significance levels in complex inheritance. Am J Hum Genet 1998; 62: 690–697.
4. Mongeau JG, Biron P, Sing CF. The influence of genetics and household environment upon the variability of normal blood pressure: the Montreal Adoption Survey. Clin Exp Hypertens 1986; 8: 653–660.
5. Lifton RP. Genetic dissection of human blood pressure variation: common pathways from rare phenotypes. Harvey Lect 2004; 100: 71–101.
6. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447: 661–678.
7. Saxena R, Voight BF, Lyssenko V, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007; 316: 1331–1336.
8. Scott LJ, Mohlke KL, Bonnycastle LL, et al. A Genome-wide association study of type 2 diabetes in finns detects multiple susceptibility variants. Science 2007; 316: 1341–1345.
9. Caulfield M, Munroe P, Pembroke J, et al. Genome-wide mapping of human loci for essential hypertension. Lancet 2003; 361: 2118–2123.
10. Cowley AW, Jr. The genetic dissection of essential hypertension. Nat Rev Genet 2006; 7: 829–840.
11. Padmanabhan S, Wallace C, Munroe PB, et al. Chromosome 2p shows significant linkage to antihypertensive response in the British Genetics of Hypertension Study. Hypertension 2006; 47: 603–608.
12. Koivukoski L, Fisher SA, Kanninen T, et al. Meta-analysis of genome-wide scans for hypertension and blood pressure in Caucasians shows evidence of susceptibility regions on chromosomes 2 and 3. Hum Mol Genet 2004; 13: 2325–2332.
13. Munroe PB, Wallace C, Xue MZ, et al. Increased support for linkage of a novel locus on chromosome 5q13 for essential hypertension in the British Genetics of Hypertension Study. Hypertension 2006; 48: 105–111.
14. Barkley RA, Chakravarti A, Cooper RS, et al. Positional identification of hypertension susceptibility genes on chromosome 2. Hypertension 2004; 43: 477–482.
15. Dominiczak AF, Graham D, McBride MW, et al. Corcoran Lecture. Cardiovascular genomics and oxidative stress. Hypertension 2005; 45: 636–642.
16. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 2005; 6: 95–108.
17. Frazer KA, Ballinger DG, Cox DR, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007; 449: 851–861.
18. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996; 273: 1516–1517.
19. Clerget-Darpoux F, Elston RC. Are linkage analysis and the collection of family data dead? Prospects for family studies in the age of genome-wide association. Hum Hered 2007; 64: 91–96.
20. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004; 96: 434–442.
21. Chakravarti A, Little P. Nature, nurture and human disease. Nature 2003; 421: 412–414.
22. Manolio TA, Bailey-Wilson JE, Collins FS. Genes, environment and the value of prospective cohort studies. Nat Rev Genet 2006; 7: 812–820.
23. Chen J, Giovannucci EL, Hunter DJ. MTHFR polymorphism, methyl-replete diets and the risk of colorectal carcinoma and adenoma among U.S. men and women: an example of gene-environment interactions in colorectal tumorigenesis. J Nutr 1999; 129: 560S–564S.
24. Bloemenkamp KW, Rosendaal FR, Helmerhorst FM, Buller HR, Vandenbroucke JP. Enhancement by factor V Leiden mutation of risk of deep-vein thrombosis associated with oral contraceptives containing a third-generation progestagen. Lancet 1995; 346: 1593–1596.
25. Lehtimaki T, Moilanen T, Porkka K, et al. Association between serum lipids and apolipoprotein E phenotype is influenced by diet in a population-based sample of free-living children and young adults: the Cardiovascular Risk in Young Finns Study. J Lipid Res 1995; 36: 653–661.

Printed by Via Medica, www.viamedica.pl