Novel Integrative Approaches to the Identification of Candidate Genes in Hypertension

Norbert Hubner, Chana Yagil, Yoram Yagil

Abstract—Hypertension, diabetes, and obesity are common diseases that are genetically expressed as complex traits. The clinical expression of these diseases, which run in families but have no clear pattern of inheritance, has been attributed to the interaction between multiple genes and the environment. Identifying the underlying genes, a crucial step in understanding the molecular pathogenesis of complex diseases, has had limited success so far, stressing the need for novel strategies to move this process forward. Innovative integrative genetic–genomic approaches have been proposed recently for the identification of new high-priority candidate genes. These strategies, which combine expression profiling with genetic linkage in different ways, may represent a breakthrough in the search for the genes involved in complex diseases. (Hypertension. 2006;47:1-5.)

Key Words: genes ■ gene expression ■ hypertension ■ genetics

Research on complex diseases has benefited greatly from the complete sequencing of several mammalian genomes and from high-throughput functional genomic technologies. The completed human, mouse, and rat genomic sequences have facilitated the identification of a few candidate genes underlying susceptibility loci for complex diseases, such as diabetes and asthma. Identifying the genes responsible for complex diseases, such as essential hypertension, has been less successful, possibly in part because, aside from being polygenic, these diseases are under strong environmental influence.

Several genetic strategies have been used over the past two decades in the investigation of hypertension and other complex diseases, but with only limited success. The candidate gene strategy has yielded many dozens of candidate genes in hypertension. Genetic linkage studies have yielded a large number of quantitative trait loci (QTL) that harbor a large number of genes. Gene expression profiling studies in hypertensive and normotensive models have yielded differences in the expression levels of many genes. But none of these strategies, when applied as stand-alone, has been successful in limiting the very large number of candidate genes and focusing the attention of investigators on a defined set of highly likely candidate genes. In contrast, the combined use of two or more genetic strategies has had better success in guiding the focus on high-priority genes. Several straightforward examples of combining investigative approaches have been reported in fields other than hypertension. One example is the use of genetic linkage to guide gene expression profiling, which led investigators to identify that variation in the atrial natriuretic factor precursor may have a bearing on stroke. Another example is the use of gene expression profiling in congeneric strains that led to the identification of the gene encoding the CD36 fatty acid transporter that contributes to features of insulin resistance. Nevertheless, the success of these combined genetic strategies in identifying high-priority candidate genes has been sporadic, at best.

Recently, two reports appeared that applied innovative schemes, which integrate in different ways genome-wide linkage analysis with genome-wide expression profiling to the search for the genes underlying complex diseases. In the first report, the investigators focused on genes that were differentially expressed between the parental strains on salt loading and that also mapped within a region linked to the regulation of blood pressure. They integrated the data stemming out of differential expression studies with data resulting from linkage analysis of a physiological phenotype. Their strategy is depicted in Figure 1. In the second report, the investigators integrated gene expression with genetic linkage data from segregating populations by treating gene expression as a quantitative trait and mapping expression QTL (eQTL) for those traits. Their strategy is summarized in Figure 2. In this review, we will highlight the principles underlying the two strategies, which have the common goal of identifying candidate genes for complex diseases. We will use the two recent reports by Yagil et al and Hubner et al as examples of how these strategies can be successfully applied to the search for the genes underlying hypertension and the metabolic syndrome.
Integrating Data From Differential Expression With Genetic Linkage Data

This strategy is based on combining data generated during genome-wide linkage analysis and differential gene expression profiling, two commonly used investigative schemes. Genetic linkage maps physiological QTL (pQTL) to a specific place (locus) on the genome, answering the question “where” in the genome the culprit genes are located. Differential gene expression profiling determines which genes are differentially expressed among contrasting strains that express contrasting phenotypes (hypertension versus normotension), answering the question “which” genes are differentially expressed in relation to the expressed phenotype (hypertension). Genetic linkage has yielded a large number of pQTL for the various complex diseases and dozens of pQTL in the search for the genetic basis of hypertension both in humans and in experimental models. Differential gene expression profiling has generated a long list of several thousands of genes that are differentially expressed in hypertension.

In the study by Yagil et al, the investigators applied the linkage analysis strategy in the male Sabra rat, a model of salt-sensitive hypertension, and detected two pQTL on chromosome 1. They were able to functionally confirm the presence of the pQTL on that chromosome by means of congenic strains. Using the genetic linkage approach alone, they produced a long list of 1102 annotated genes within the pQTL and were unable to additionally focus their investigation. The investigators applied next differential gene expression profiling in kidneys, contrasting parental Sabra hypertension prone (SBH/y) and resistant (SBN/y) rats. They made two fundamental assumptions. The first was that any relevant gene would be differentially expressed between the parental strains, particularly after salt loading. This approach resulted in a list of genes that were not only influenced by the genetic background but also by salt loading, which reflects the nature of hypertension in this model. The second assumption was that any relevant genes would be differentially expressed in the kidneys. They found that of the 15,442 genes that were expressed in their microarray, 2470 genes were differentially expressed in the kidneys of the two strains. Using a 2×2 factorial ANOVA and cluster analysis, they determined that 739 of the differentially expressed genes were of direct relevance to genetic hypertension and reflected an interaction between salt loading and allele status. This number of genes was still too large and did not achieve the needed scaling down of the number of candidate genes. In the third step of their study, the investigators integrated the data generated from genetic linkage with the data resulting from differential gene expression profiling. They reasoned that differentially expressed genes that also mapped within the physiological QTL would be the strongest possible candidate genes for salt-sensitive hypertension, thus answering the question regarding which of the differentially expressed genes were the most likely to be directly involved in hypertension in their model. Of the 739 genes that were differentially expressed between the strains and that were of relevance to hypertension, 19 mapped to rat chromosome 1. Eight of the 19 genes fell within the blood pressure–related pQTL that they had defined previously. Differences in the renal expression of 7 of the 8 genes were confirmed by real-time PCR.

The study by Yagil et al exemplifies the scheme through which the investigators were able, by integrating the data resulting from genetic linkage studies and differential gene expression profiling, to reduce the number of candidate genes from 1102 within the pQTL or from 739 genes that were differentially expressed and relevant to hypertension to just 7 high-priority candidate genes.

Integrating Genetic Linkage With Expression Profiling

This strategy is aimed at identifying genetic variants that influence gene expression, ultimately answering the question which transcripts regulate the expression of known candidate genes. The strategy, which has been elegantly reviewed by Li
Comparison and Feasibility of the 2 Integrative Approaches

Combining gene expression with other genetic tools provides an additional dimension to tackle complex genetic disease. Which of the two integrative strategies described above is preferable, and which should investigators adopt?

The strengths of the first integrative approach, which combines linkage with differential gene expression profiling, are that it is simple and straightforward, relatively inexpensive, and widely applicable. Furthermore, because abundant data are already available from genetic linkage and differential expression studies, investigators should be able to apply this strategy not only de novo but also post hoc to the data they have already generated. The weakness of this strategy is that two assumptions must be made, namely that the genes must be differentially expressed to be significant and that the genes must be expressed in the organ or tissues that are sampled.

The strength of the second integrative approach is that it uses linkage analysis of both physiological phenotypes and of expression data. A second major strength of this approach is that it is the one and only strategy that allows us, on a genome-wide scale, to identify allele-specific transcripts/genes that regulate gene expression. The disadvantage of this comprehensive integrative approach is its high cost. To use expression data as quantitative phenotypes, a large number of samples must be tested. The current high cost of microarray chips would be a major limiting factor. An additional disadvantage is the relative lack of experience with this relatively novel strategy. Several recent studies suggest, nonetheless, that this approach can, indeed, successfully identify candidate genes for complex traits.

Can the two strategies substitute one for the other? No, because the two approaches are complementary. The first approach generates a list of novel candidate genes, some of which would be “causative” genes and others “regulatory” genes, with no way to differentiate between the two. The second approach clearly identifies the regulatory transcripts or genes. Both approaches are necessary, because they add important information and constitute a significant step forward in the quest to elucidate the genetic basis of complex diseases in general and cardiovascular disease and hypertension in particular.
Perspectives

The integrative approach combining several strategies at the genome level is necessary to break through the impasse confronting many investigators in the search for the genetic basis of complex diseases, including hypertension. The recently reported successes in integrating genetic–genomic strategies, two examples of which are provided in this review, lend additional impetus to the contention that combining and integrating strategies produce a higher yield than using any genetic strategy as stand-alone. Integration of genetic strategies may constitute the long-sought key to simplifying genetic strategy as stand-alone. Integration of genetic–genomic strategies may lend additional impetus to the contention that combining and integrating strategies produce a higher yield than using any genetic strategy as stand-alone.

Acknowledgments

The German National Genome Research Network (NGFN II) supported N.H., and D-Cure, Binational Science Foundation, Chief Scientist’s Office of the Ministry of Health, and the Israel Science Foundation supported Y.C. Y.C. thanks Friedrich C. Luft for his helpful comments.

References


Huber et al. Integrative Genetics 5


Novel Integrative Approaches to the Identification of Candidate Genes in Hypertension
Norbert Hubner, Chana Yagil and Yoram Yagil

Hypertension. 2006;47:1-5; originally published online December 12, 2005;
doi: 10.1161/01.HYP.0000197951.82190.c4
Hypertension is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2005 American Heart Association, Inc. All rights reserved.
Print ISSN: 0194-911X. Online ISSN: 1524-4563

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/47/1/1

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/