Discovering the Genetics of Complex Disorders Through Integration of Genomic Mapping and Transcriptional Profiling

Robert S. Danziger[1]* Ming You[2] and Huda Akil[3]

[1]Department of Medicine, University of Illinois (Chicago, IL), [2]Division of Surgery, Washington Univeristy (St. Louis, MO), [3]Mental Health Research Institute, University of Michigan (Ann Arbor, MI)

Abstract: Discovering the genes that contribute to complex polygenic diseases represents a significant challenge. Investigating the structural genetic variations associated with these disorders may not be sufficiently informative about vulnerability genes modulated by the environment. Characterizing gene expression patterns does not identify the primary differences in gene structure. The convergence of global screening of gene expression patterns with extensive structural genomic information may be necessary to identify the gene clusters that contribute to these pervasive diseases. These integrative efforts, though promising, are in their early phases and will require further refinement.

INTRODUCTION

It is now generally accepted that many of the most pervasive illnesses are rarely inherited in a Mendelian fashion and that they are polygenic in etiology, with significant gene–environment interactions. As such, understanding the genetics of complex diseases is a focus and challenge for medicine. Identifying the genes that contribute to or modify complex inherited diseases will serve as a basis for both the pathophysiological understanding of these diseases and the discovery of novel drug targets.

Two general approaches have been applied to the discovery of genes for complex diseases - the functional approach, including the candidate-gene approach, and the positional approach. For "functional cloning" approaches, the identification of candidate genes is based on a series of 'educated' guesses without reference to chromosomal map position. The major shortcomings of these methods is the requirement for *a priori* knowledge of gene function and the expectation that critical genes have a detectable impact on the phenotype. "Positional cloning" approaches locate the gene solely on the basis of map position and obviate the need for *a priori* knowledge of gene function. In positional cloning, the identification of the culprit gene within the relevant region is a major challenge. Recently, the "positional candidate" method [1] - which is based on the combination of the identification of a chromosomal sub-region or quantitative trait locus (QTL), usually by linkage analysis, with an examination of the region for attractive candidate genes - has been used. Most prominently, this has successfully identified point mutations in the *RET* proto-oncogene, which cause multiple endocrine neoplasia type 2 [2], and mutS, hPMS1, hPMS2 and mutL, which cause hereditary non-polyposis colon cancer [3-5]. However, the weakness of this method is that it relies on the density of the transcript map and on a priori knowledge of gene function to identify good candidates within the region of interest. Moreover, the structural genetic approach does not address the biological path(s) that leads from a normal allelic variation to vulnerability to an illness.

While the above strategies focus on discovering differences in gene structure, recent approaches have evolved to globally assess gene function, including transcriptional profiling and alterations in protein levels. These methods are still in their early phases and continue to require technical and analytical refinements. They offer the significant advantage of examining function specifically in the affected tissue, thereby defining a profile of altered activity that results from the interaction of genes and environment. By the same token, the observed alterations in gene function are, in the majority of cases, not likely to be primary or directly attributable to genetic variations. Moreover, many might also be secondary to the disease process rather than its cause. Thus, transcriptional profiling alone cannot define the genetic etiology of complex disorders.

Bringing the global functional screening methods, such as transcriptional profiling, together with genomic mapping and positional cloning could theoretically combine the strengths of both approaches and overcome their individual limitations. However, this is a demanding process requiring both genetic expertise and specific knowledge of the organ or physiological system involved in the disorder. Its progress depends either on the availability of appropriate human tissue or on an informative animal model of the disease. It will also depend on the power of the individual technologies and our ability to analyze, mine and interpret the resulting information.

Here we review how the integration of transcriptional profiling and genomics is being used to investigate the genetics of three common, complex phenotypes/disorders: blood-pressure control and hypertension; cell growth and cancer; and central nervous system (CNS) function and neuropsychiatric disorders. The reviewed approaches are diverse - they use either animal models or direct human studies, and they are at various phases of progress. The fact that these three areas sometimes lack direct parallels is reflective of the state of the field where each class of disorders offers unique challenges and may require

*Address correspondence to this author at the Department of Medicine, University of Illinois at Chicago, 840 S. Wood St., Chicago, IL 60612, USA; Tel: 312-569-6606; Fax: 708-594-2907; E-mail: Rdanziger@aol.com

1573-4021/05 $50.00+.00 © 2005 Bentham Science Publishers Ltd.

somewhat different strategies. But from the combined picture emerging from these areas we can glimpse the range of possibilities that this intersection of approaches can provide. We can also point to the limitations of this strategy that need to be overcome in order to ensure significant progress.

BLOOD-PRESSURE CONTROL AND HYPERTENSION

Blood pressure is a complex quantative phenotype. The greatest insights into the genetics of hyptension have been in the Mendelian forms [6]. So far, the identification of candidate genes has been most successful for genes for which there is *a priori* knowledge of relevance to the pharmacology and physiology of blood-pressure control, e.g., those in the renin-angiotensin-aldosterone (RAS) and adrenergic signaling pathways [7].

Salt-Sensitive Hypertension

Although salt is considered a major environmental risk factor for hypertension, there is significant inter-individual variability in the blood-pressure response to salt, which indicates that it has polygenic underpinnings. So far, a few genes (including aldosterone synthase and 11-beta hydroxylase, the epithelial sodium channel (ENaC), the mineralocorticoid receptor and the serine-threonine kinases, WNK1 and WNK4) have been linked to rare Mendelian forms of human hypertension that show some salt sensitivity [8-15]. However, the candidate genes that have been tested for salt-sensitivity and evaluated to date do not fully explain salt-sensitive hypertension. The need to identify novel genes that have not been considered and to highlight specific genes within regions that have been identified by physical-mapping strategies led to the use of the integrative approach of transcriptional profiling in conjunction with QTL analysis in animal models.

Studies of salt-sensitive hypertension have capitalized on several artificially selected, inbred hypertensive strains of rats, in which congenic constructs and genome-wide scans have been used to identify QTLs for blood pressure. One of the most commonly studied strains is the Dahl salt-sensitive (SS) rat strain, which is an inbred strain that was artificially selected from progenitor Sprague Dawley rats to obtain the maximum blood-pressure response when given a high-salt diet [16]. This is considered as a "renal model" of salt-sensitive hypertension, as cross-transplantation" experiments have shown that the kidney confers salt sensitivity [17]. One of the genes for which there is evidence for linkage of mutations with altered steroid biosynthesis and blood pressure in the Dahl rat of is 11 beta-hydroxylase [18].

Congenic strains and genome-wide scans of the Dahl rat have identified up to 16 blood-pressure QTLs [19-21]. A major challenge is to identify the genes within these QTLs that are associated with salt-sensitivity. Genes for which the renal transcript abundance changes during normal adaptation to a high-salt diet will include those that are important for salt adaptation, that are induced by hypertension, and that are unrelated to salt-adaptation and blood pressure control. The working hypothesis was that genes that are most likely to be important in the pathogenesis of salt sensitivity are those that do not undergo the same change in abundance in animals/strains that develop hypertension when given a high salt diet. Gene transcripts, identified by transcriptional profiling, that had different abundances in the Sprague Dawley rats, a normotensive, salt-resistant strain, on two different salt diets included three that are elements of the RAS system: renin; serum and glucocorticoid kinase 1 (SGK1), which activates the epithelial Na channel (eNaC); and aminopeptidase N, which metabolizes angiotensin III to IV (a blood-flow regulator). Two of these transcripts, SGK1 and aminopeptidase N, did not undergo the same change in abundance in Dahl SS rats, suggesting that they may be good candidate genes for salt-sensitive hypertension. SGK1 and aminopeptidase N genes were mapped, using radiation hybrid mapping, with respect to reported blood-pressure QTLs in the Dahl SS rat [22-23]. Each mapped to QTLs, determined using congenic Dahl strains, on chromosome 1. A polymorphism of the aminopeptidase N gene that is preferentially expressed in the Dahl SS versus Sprague Dawley and Dahl salt-resistant (SR) (the parallel strain selected for salt-resistance from progenitor Sprague Dawley) rats has recently been identified (R.D., unpublished observations). Physiological validation of these genes is now being pursued with protein and activity measurements [23].

In an analogous strategy, Liang *et al.* [24] compared the transcriptional profiles of the renal medulla from Dahl SR rats with Dahl SS rats that were consomic for chromosome 13 from Brown Norway (BN) rats, another salt-resistant strain, on high and low salt diets. Using this four-way comparison, 60 genes were identified. The fact that only a few of these genes (such as 11 beta-HSD type 1, NGF-inducible anti-proliferation protein, 1-Cys peroxiredoxin, kynurene 3, hydroxylase), which have distinct temporal patterns of expression, map to rat chromosome 13 could be caused by several mechanisms. Genomic differences on chromosome 13 and the products that result could be regulatory factors that directly influence the expression of genes on many chromosomes. These genome-wide gene interactions could also be indirect; through downstream signaling pathways, or as a result of feedback regulation by functional alterations. In addition, the possibility of residual BN alleles on chromosomes other than chromosome 13 in SS-13BN/Mcw rats (bred at the Medical College of Wisconsin, Milwaukee) cannot be excluded.

Lee *et al.* [25] sought analysis of differential gene expression patterns in the kidneys of a panel of eight congenic strains, each of which carries a different low-BP QTL allele with a genetic composition that is otherwise similar to that of the hypertensive Dahl salt-sensitive (S) rat strain. Seven out of 37 differentially expressed genes were mapped to congenic regions carrying BP QTLs, but only one of these genes, L-2 hydroxy acid oxidase (Hao2), showed the congenic strain-specific pattern of differential kidney gene expression predicted by its chromosomal location.

Thus, transcriptional profiling has highlighted a number of genes that encode proteins both within and outside the RAS pathway. In each study, genomic mapping data was used to filter the genes by selecting only those which are most likely to co-segregate with hypertension and thereby provide a small group of candidate genes for further testing.

Salt-Independent Hypertension

The identification of candidate genes for salt-independent hypertension by contrasting transcriptional profiles of spontaneously hypertensive (SHR) and normotensive (WKY) rat strains, has led to the identification of 46 differentially expressed genes. Two of these, i.e., monocarboxylate transporter 1 and glutathione S-transferase Y(b) subunit, significantly impact diastolic blood pressure in F2 crosses of the rats [26]. This points to the potential value of integrating transcriptional profiling with other forms of genetic analysis, e.g., backcrosses, to narrow down the number of genes identified by transcriptional profiling for further consideration. The fact that none of the genes mapped to QTLs for systolic hypertension demonstrates one of the limitations of this method (see below).

Stroke Related to Hypertension

Fornage *et al.* [27] used an analogous strategy to determine the genes that are associated with stroke in hypertension. They compared the transcriptional profiles of brains from spontaneously hypertensive rats (SHR) and stroke-prone SHR rats (SHRSP), a strain that was developed by selective breeding for the development of cerebrovascular injury at a high frequency from the SHR rat [28]. Out of 117 differentially expressed unique genes or expressed sequence tags (ESTs), five - DOC2A, natriuretic peptide precursor type A, Kcnab2, CaMK2 inhibitor, and casein kinase 2 - mapped to stroke QTL regions. Other genes that encode multiple kinases of the MAPK/AKT signaling pathways, including JNK2, AKT2, and P13K, did not map to reported QTLs. Thus, by integrating transcriptional profiling with mapping data, 2 of 117 genes have been highlighted as candidates. Furthermore, other genes among the 117 may also precipitate cosegregation analysis of regions to which they map, especially since *a priori* mapping in the SHR is incomplete.

In a study by McBride et al [29] a congenic strain that exhibited a reduced systolic and diastolic blood pressures compared to the SHRSP rat in which a region of chromosome 2 from the normotensive Wistar Kyoto (WKY) strain was introgressed into the genetic background of the SHRSP rat, glutathione S-transferase μ–type 2, a gene mapping to the congenic region and involved in defense against oxidant stress, was differentially expressed.

These varying strategies, all dependent on various animal models of hypertension have yielded novel candidates that now can be tested in various types of human hypertension.

CANCERS AND TUMOR MODIFIER GENES

The risk of cancer is associated with both environmental exposure and genetic predisposition, i.e. gene-environment interaction. For example, lung cancer has frequently been cited as an example of a malignancy that is solely determined by the environment, and the risks associated with cigarette smoking, and certain occupations, such as mining, shipbuilding and petroleum refining, are well established [17–24]. However, there is also strong evidence for genetic susceptibility. Studies of familial aggregation of lung cancer suggest that genetic factors are involved in human lung tumor development [30-35]. Specifically, segregation analyses of lung cancer proband families indicate that a Mendelian codominant inheritance of a rare major autosomal gene is involved [36]. This locus has been shown to account for 69%, 47%, and 22% of the cumulative incidence of lung cancer in patients at age 50s, 60s, and 70s, respectively [36]. Since genetic heterogeneity and enormous variation in exposure levels to tobacco and other environmental agents make it difficult to identify lung cancer susceptibility loci in humans, inbred mouse models offer an effective means of identifying candidate lung cancer susceptibility loci.

Numerous lung tumor-related QTL have been found in various animal models [37,38]. Susceptibility to chemical induction of lung tumors in mice varies according to the strain [39]. Genetic linkage analyses using various mouse crosses have revealed a series of pulmonary adenoma susceptibility (Pas) genes, e.g. Pas1 (chromosome 1), Pas2 (chromosome 17), and Pas3 (chromosome 19), and pulmonary adenoma resistance (Par) genes, eg. Par1 (chromosome 11), Par2 (chromosome 18), Par3 (chromosome 12), and Par4 (chromosome 4) [37-39]. In the past decade, several research groups have been actively pursuing the identification of the genes that underlie these lung tumor QTL [37-39]. However, the identification of these candidate genes has proven rather difficult. One obstacle is the fact that several hundred genes can lie in a 20-30 cM QTL region. Fine-mapping studies are typically time-consuming and labor-intensive. Evaluation of differential gene expression and nucleotide polymorphism of such a large number of genes can also be a significant challenge using a conventional approach. The useof microarray technology in combination with the recently completed mouse/human genome sequences has greatly facilitated the identification of modifier genes in a mapped QTL [40,41]. Indeed, candidate lung tumor modifier genes whose expression varies between parental strains have been identified using transcriptional profiling in two independent studies [42,43].

Recent studies undertook the profiling of lungs from mouse strains with variable susceptibility to lung tumor development as a means to identify, within known QTLs, candidate genes responsible for susceptibility or resistance to lung cancer [42]. The work focused on eight mapped mouse chromosomal regions linked with lung tumor susceptibility (Pas1-4) or resistance (Par1-4). High-density oligonucleotide arrays were used to measure the relative expression levels of >36,000 genes and ESTs in lung tissues of A/J, BALB/cJ, SM/J, C3H/HeJ, and C57BL/6J mice. In total, 50 differentially expressed genes, whose pattern of expression in the strains was also consistent with that of the reported association with cancer, were found in the lung cancer susceptibility QTLs. As A/J is the susceptible strain and C57BL/6J the resistant one, tumor suppressor candidates will have shown higher expression in C57BL/6J, while oncogene candidates will have shown higher expression in A/J for Pas1-4 loci. Genes that show differential expression are Cyclin D2, high mobility group protein 2A4, Cyclophilin H, Protein tyrosine phosphatase BK, RecQ protein-like, ECA39 protein, and Hes related protein. Pas2 QTL is located at the H-2 locus whose haplotypes correlate with the incidence and multiplicity of mouse lung tumor induction. Candidates that show differential expression are Tapasin, KE2, H2-K region

expressed gene 6, Notch4, Regulator of cullins 1, MHC psoriasis candidate protein, Histocompatibility 2M region locus 9, Eclonucleotide pyrophosphatase/ phosphodiesterase 4, Heterogeneous nuclear ribonucleoprotein K, and CDC5-like. Pas3 candidates include cdc25 homologue A (S cerevisiae), semopharin 4G, golgi specific brefeldin A resistance factor 1, adrenergic receptor (beta 1). Pas4 candidates are tyrosine kinase adaptor protein 1, parathyroid hormone receptor, Wiskott-Aldrich syndrome homologue binding protein, SMARC D3, cyclin dependent kinase 5, G(i) alpha 2, stromal antigen 1, and similar to topoIIBP.

The identificationof these candidate genes is an important first step towards the determinationof the genes that are responsible for a given lung tumor QTL. Real-time RT-PCR is usually used as the second step to corroborate the candidate genes that emerge from oligonucleotide arrays. The definitive confirmation of the real gene(s) will require both *in vitro* and *in vivo* studies. *In vitro* studies involve the characterization of allelic differences in cell proliferation, cell cycle, and apoptosis, which will provide important clues as to which candidates may potentially be important in lung tumorigenesis. The ultimate proof that a candidate gene is actually the gene for the tumor QTL may require targeted allelic substitution (knock-in mice). In the resulting mice, only the putative allelic variant from a donor (inbred strain) replaces the homologous allele in a recipient inbred strain and results in a change in lung tumor susceptibility in the recipient strain.

Another approach is to first identify all differentially expressed genes in strains of mice with known lung tumor phenotype, followed by mapping of the identified genes to their chromosomal locations of the lung cancer modifier loci. Using this approach, Gariboldi *et al.* [43] analyzed gene expression in the lungs of 16 inbred strains of known susceptibility/resistance to lung tumorigenesis using cDNA microarrays containing 19K full- cDNAs (all of the genes and all of the isoforms of genes in the mouse genome were not represented on these microarrays and thus the candidates may not be in the set that was profiled) [44]. At least 10 genes have shown functional role in cancer growth, differentiation or progression. The remaining genes show potential involvement in cancer. For example, myeloid cell leukemia sequence 1, major vault protein, alpha-methylacyl-CoA racemase, plastin, and Arhc ras homolog gene family - member C are tumor markers, the altered expression of which is associated with tumor development and progression [44]. Among the 91 genes identified, the genes mapping closer to lung cancer modifier loci were: Hypothetical protein FLJ11088 at the Pas1 locus, Stard3 START domain containing 3, Krt1-10 Keratin complex 1 (acidic) gene 10, and Homolog to high sulfur protein B2E at the Par1 locus. These results suggest that the gene expression profile of normal lung tissue can identify highly relevant candidates for the genetic predisposition to lung tumorigenesis in mice without the information from prior QTL mapping studies.

As shown in Table **2**, candidates from the above study [43] are part of several cellular signaling pathways, including apoptosis, cytoskeletal organization, chromatin modeling and the cell cycle. As such, susceptibility and resistance to cancer involves a constellation of factors, some interacting, and failure of any one can lead to carcinogenesis.

However, there is an intrinsic problem in these two experiments of using gene expression profiling to identify candidate lung tumor susceptibility genes in mice. Normal lung tissue contains more than 30 different cell types. The cells that may develop into tumor cells, i.e. the alveolar type II cells and Clara cells, represent a very small fraction of the cells in normal lung. Thus, the bulk of the expression measured comes from cells not involved in cancer development. There is a possibility that some of the differentially expressed genes specific for tumor susceptibility are not readily detectable due to the apparent dilution of the cells related to tumorigenesis. Future studies should use the pure Type II cells or Clara cells from lungs of the parental strains of mice to overcome this problem. Finally, it remains to be seen how many of the mouse susceptibility genes will really have human counterparts that will be useful for diagnostic or therapeutic purposes.

BRAIN-RELATED GENES: FOCUS ON NEUROPSYCHIATRIC DISORDERS

Brain-related complex genetic disorders offer unique challenges, as brain function depends on the activity of neural circuits that include multiple cell types interconnecting across brain regions and requiring the interplay of numerous signaling pathways. Moreover, the brain is programmed to change structurally and functionally as a function of experience. This “neural plasticity” allows experience-based learning and memory and is associated with significant changes in gene expression patterns across the relevant circuit. These characteristics make it difficult to draw a one-to-one connection between the sequence variations of a single gene and a given function or disorder, and underscore the importance of considering integrated brain activity as the central intermediary between genetics and behavior.

These considerations are particularly important in certain psychiatric and neurological disorders such as schizophrenia, bipolar illness, Parkinson and Alzheimer Disease that are mediated by specific neural circuits, are highly dependent on the environment and show a complex lifetime course. But in spite of the likely heterogeneity of the susceptibility genes for these disorders, a common “neural phenotype” likely exists that represents the signature of the disease on the brain. This would represent an overall pattern of gene expression within a relevant circuit as a result of altered activity of both the vulnerability genes and their downstream targets. Thus, it is reasonable to use a convergent approach that combines functional genomic strategies with classical genetic strategies to uncover this “neural phenotype” that defines a neuropsychiatric disorder, and devise appropriate strategies for disease treatment and prevention.

Another challenge is that many brain disorders (particularly psychiatric) are difficult to model in rodents. While there are animals that may be prone to hypertension or cancer or seizures, there are no known rat or mouse strains that are prone to psychosis, an essentially human disorder. However, some animal models can mimic specific aspects of

Table 1. Candidate Genes Identified by the Combination of Transcriptional Profiling and Physical Mapping

Hypertension				
Type	*Rat Chromosome*	*Linked region*	*Candidates from microarray studies*[b]	
			Gene Name (Gene Symbol)	*Position (Mb)**
Salt-sensitive	1	D1rat35(124.7)-D1rat131(145.6) [96]	Aminopeptidase N [23]	135.7
		D1uia8-D1rat18(43.7) [96]	SGK1 [97]	23.5
	2	D2N35 (169.4)-Nep[30]	Hao2 [30]	193.5
	13	Chomosome 13 (consomic Dahl SS/BNw 13 rats) [24]	beta-HSD type 1 [98]	109.2
			NGF-inducible anti-proliferation protein [98]	47
			1-Cys peroxiredoxin [98]	76.8
			kynurenine 3-hydroxylase [98]	91.6
Salt-resistance	2	F2 cross (SHR x WKY) LOD 3.5 [26] Diastolic BP	monocarboxylate transporter [26]	200
			glutathione S-transferase μ type 2 [26]	203.5
Stroke-prone	5	Anf (169.4) LOD 4.7 [99]	ANP [27]	165.1
			Kcnab2 [27]	169.6
			CaMK2 inhibitor [27]	142.5
	3	D3Mit2 (143.0) [99]	Casein kinase 2 [27]	256.2
	1	CYP2b2 (81.4)-DiMit7(169.4) LOD 7.4 [99e]	Doc2A [27]	186.2
			sulfotransferase A1 [27]	186
Cancer				
Lung tumor susceptibility /resistance QTL	*Mouse Chromosome*	*Candidate region / Linked locus with the peak LOD score or least P-value (Mb)*[a, c]	*Candidates from microarray studies*[b]	
			Gene Name (Gene Symbol)	*Position (Mb)*[c]
Pas1	6	D6Osu6 (145.03)---D6Osu12 (145.53) [100] D6Int24 (145.01)---D6Int47 (145.48) [101]	Branched chain aminotransferase 1, cytosolic (Bcat1) [42]	145.1
			Hypothetical protein FLJ11088 [42]	148.3
Pas2	17	D17Mit23(27.0)--- D17Mit231(34.6)/ D17Mit16 (33.4) [102]	Trefoil factor 3, intestinal (Tff3) [42]	29
			Solute carrier family 37 (glycerol-3-phosphate transporter), member 1 (Slc37a1) [42]	29.2
			U2 small nuclear ribonucleoprotein auxiliary factor (U2AF) 1 (U2af1) [42]	29.6
			SNF1-like kinase (Snf1lk) [42]	29.7
			Cytochrome P450, family 4, subfamily f, polypeptide 16 (Cyp4f16) [42]	30.4
			Zinc finger protein 297(Zfp297) [42]	33.7
			H2-K region expressed gene 6 (H2-Ke6) [42]	33.7
			Procollagen, type XI, alpha 2 (Col11a2) [42]	33.8

Table 1 Contd…

Hypertension				
Type	*Rat Chromosome*	*Linked region*	*Candidates from microarray studies*[b]	
			Gene Name (Gene Symbol)	*Position (Mb)**
			Transporter 1, ATP-binding cassette, sub-family B (Tap1) [42]	33.9
			Histocompatibility 2, O region beta locus (H2-Ob) [42]	34
			Notch gene homolog 4, Drosophila (Notch4) [42]	34.6
Par1	11	D11Mit41(95.7)--- D11Mit126 (112.7)/Rara (105.6) [103]	Arachidonate 12-lipoxygenase, 12R type (Alox12b) [42]	73.1
			Unc-119 homolog (C. elegans) (Unc119) [42]	84.9
			Chemokine (C-C motif) ligand 7 (Ccl7) [42]	88.7
		D11Mit4(72.5)---D11Mit14(105.3)/ D11Mit70 (100.8) [104]	Chemokine (C-C motif) ligand 1 (Ccl1) [42]	88.8
			Chemokine (C-C motif) ligand 4 (Ccl4) [42]	90.3
			Speckle type POZ protein (Spop) [42]	102
			similar to 60S RIBOSOMAL PROTEIN L29 [42]	103.7
			ProSAPiP2 protein [42]	103.8
			Proteasome (prosome, macropain) subunit, beta type, 3 (Psmb3) [42]	104.4
			START domain containing 3 (Stard3) [42]	105
			Keratin complex 1, acidic, gene 10 (Krt1-10) [42]	107
			RIKEN cDNA 2310040M23 gene (2310040M23Rik) [42]	107.2
			Keratin complex 1, acidic, gene 24 (Krt1-24) [42]	107.7
			Kelch-like 10 (Drosophila) (Klhl10) [42]	108.1
			Receptor (calcitonin) activity modifying protein 2 (Ramp2) [42]	109.5
			Granulin (Grn) [42]	111.1
			Similar to Proteasome subunit beta type 5 precursor [42]	111.4
Par2	18	D18Mit103 (68.5)---D18Mit188 (70.8) [105]	Methyl-CpG binding domain protein 2 (Mbd2) [42]	68.9
Par3	12	D12Mit52 (75.1)---D12Mit6 (89.2) / D12Mit5 (79.4) [104]	Placental growth factor (Pgf) [42]	83.6

Table 1 Contd…

Hypertension				
Type	*Rat Chromosome*	*Linked region*	*Candidates from microarray studies*[b]	
			Gene Name (Gene Symbol)	*Position (Mb)**
Par4	4	D4Mit39 (31.8)---D4Mit 77 (84.8) [106, d]	Insulin-like growth factor binding protein-like 1 (Igfbpl1) [42]	42.8
			Polymerase (DNA directed), epsilon 3 (p17 subunit) (Pole3) [42]	58.2
			Alpha 1 microglobulin/bikunin (Ambp) [42]	58.8
			Stathmin-like unknown gene [42]	65.5
			EGF-like-domain, multiple 5 (Egfl5) [42]	66
			Protein-tyrosine-phosphatase receptor delta [42]	71.4
			Tyrosinase-related protein 1 (Tyrp1) [42]	76.4
			Interferon alpha family, gene B (Ifnab) [42]	84.2
			Cyclin-dependent kinase inhibitor 2A (Cdkn2a) [42]	84.7

Brain				
Type	*Human Chromosome*	*Linked region/testing*	*Candidates from microarray studies*[b]	
			Gene Name (Gene Symbol)	*Position (Mb)**
Schizophrenia	1	10 kb region containing gene (linkage analysis) [98]	RGS4 [107]	160.2
	7	Mutation screening (association study) [99]	NPY [108]	24.1
Bipolar Disorder	22	Disequilibrium of promotor region [71]	GRK3 [70]	24.3
Alzheimer's and Parkinson's Diseases	10	10 q [79]	GSTO1 and GSTO 2 [74]	105.7
Alcoholism	Rat chromosome 4	20-100 cM region (LOD 9.2) [103]	Alpha-Synuclein [109]	89.9

a. Candidate region of each QTL was based on results of interval mapping studies. The position information for each QTL is added based on the previous interval mapping results (representing 95% confidence intervals based on the 2 LOD score significant threshold and 1-LOD-score-fall supporting interval.

b. Candidates were generated by the recent microarray studies [42,43]. The Pas3 (chr. 19) and Pas 4 (chr. 9) datasets [42] were not included in this review due to their undefined QTL positions.

c. All of chromosome positions were based on the latest Celera mouse genome database (released on April 4, 2004). or rat University of California Santa Cruz (UCSC) Genome Browser.

d. The Peak of LOD score for the Par4 is located in the middle of interval and no marker was designated.

e. Genes to not map to QTL based on UCSC browser even though originally reported to do so,

the human disease. For example, the increased activity associated with mania is modeled in animals with psychostimulants like amphetamine [45]. Thus, while functional genomics in animal models may be suggestive, direct studies in humans are essential.

Clinical neuroscience has recently entered a new phase of relying on convergent approaches to implicate specific genes in the major neurological and psychiatric disorders. Here we focus on a few recent findings that rely primarily on genomic data from human postmortem tissue, which exemplify this approach. However, we will include an example from

bipolar illness research that started with an animal model and used functional genomics to identify a novel candidate that was then pursued in human association studies.

New Genes for Schizophrenia

Schizophrenia affects approximately 1% of the population and has a strong genetic component, with estimates of heritability exceeding 50% [46]. But it is only during the last three years that any susceptibility genes have been implicated in schizophrenia. Genome wide scans had linked numerous chromosomal loci to the disease. A meta-analysis by Badner and Gershon [47] found the strongest evidence of schizophrenia susceptibility loci on 13q and 22q (shared with bipolar illness) and on 8p. Recent association studies of genes within the linked regions have yielded exciting discoveries. Prominent among them is the identification of Neuregulin (NRG1) as a susceptibility gene that has been replicated in a number of cohorts [48-51]. Also receiving much attention is Catecholamine-O-Methyl tansferase (COMT), a functional candidate that has recently been associated with schizophrenia in an Ashkenazi population [52]. A specific COMT allelic variation associated with schizophrenia has been shown to alter the activity of the enzyme, to be correlated with decreased cognitive function [53] and with the induction of another critical enzyme (tyrosine hydroxylase) in a circuit implicated in the *psychotic symptoms* of schizophrenia [54]. A handful of other schizophrenia genes have recently been identified through these classical genetic methods [55].

Recent years have ushered in the use of genomics to describe the "neural phenotype" that is associated with schizophrenia. Although gene arrays offer a number of challenges for studying the neural systems [56], and while there are many difficulties associated with the study of postmortem human brains, such as the effects of agonal state and tissue pH [57], the approach holds promise for brain disorders especially if based on careful characterization of control brains and combined with validation strategies such as Real-Time PCR and *in situ* hybridization [58]. A number of novel candidate genes have emerged from the body of genomic studies focusing on schizophrenia [59-66]. Results reveal genes expression changes relating to oligodendrocytes and myelination, cytoskeletal organization, energy metabolism, synaptic function and plasticity and intracellular signaling pathways [67], and some of these genes and pathways have been replicated by independent laboratories [63,68,69]. Among the neurotransmission/signaling genes, two have been pursued from gene expression to gene structure: RGS4 and NPY (Boxes **1** and **2**).

A Novel Candidate in Bipolar Illness

Although few animal models capture the full extent of major psychiatric disorders, some models focus on specific but central symptoms as a means of investigating underlying mechanisms. This affords the opportunity to discover novel genes in animal models and test them in humans. Niculescu *et al.* [45] used a single injection of methamphetamine in rats to model the hyperactivity that is associated with mania. These authors then carried out oligonucleotide microarray studies in brain regions that are associated with emotional reactivity- contrasting treated and untreated animals. Genes that were significantly altered were cross-matched against human genomic loci associated with either bipolar disease or schizophrenia. A G protein-coupled receptor kinase (GRK3) emerged as the gene most significantly induced by methamphetamine in rat and is present in one of the strongly implicated human loci (22q12). The same group then studied GRK3 in transmission disequilibrium analysis in bipolar families. They identified several SNPs in the promoter region of this gene and provided evidence for a significant association between one of these SNPs (P-5) and bipolar illness. The authors suggest that GRK3, given its involvement in the desensitization of G protein receptor, may be one of the genes that can alter signaling in the brain of bipolar subjects [70]. Whether the SNP in question is indeed associated with altered expression of GRK3 or downstream molecules in human brain, and whether altered GRK3 activity can impact animal behavior remains to be determined.

Genomics for the Study of Alzheimer and Parkinson Disease

A number of genes have been implicated in the increased risk for Alzheimer Disease (AD) and Parkinson Disease (PD) [71,72]. But beyond the risk of getting the disease, the age-at-onset of these neurological disorders can vary widely and is also genetically determined. Previous work demonstrated that a common locus (10q) controls the age-at-onset for both AD and PD 74**.** However, this 15-cM region includes a large number of genes. Thus, Hauser *et al.* [73] used what they term 'genomic convergence' - the combination of gene arrays with genetic association data - to select the genes of interest. They carried out transcriptional profiling on hippocampi obtained from AD patients and controls. Four genes that were significantly altered in expression between cases and controls were found to map to the 10q linkage region. However, only one of these genes was significantly associated with the onset of the disease in the follow-up association studies. Allelic association studies for age-at-onset effects in AD and PD revealed a significant association for glutathione S-transferase, omega-1 (GSTO1)] and uncovered an association with a second transcribed member of same class, GSTO2 located next to GSTO1 [74] . Interestingly, these genes were only associated with age-at-onset of both diseases but not with overall risk of developing either disorder. While the functions of GSTO1 and GSTO2 are not well understood, there is evidence that they may play a role in the biotransformation and disposition of drugs and products of oxidative stress, as well as in the post-translational modification of interleukin-1 {1742}.Thus this novel discovery is in line with pre-existing hypotheses about the possible role of oxidative stress and inflammation in neurodegenerative disorders.

These early results with neuropsychiatric diseases exemplify the promiseof combining functional genomics with QTL analysis to uncover novel molecular candidates for the susceptibility to these illnesses. These studies have face validity in that they are consistent with our broad notions of the illnesses by revealing genes related to signaling or neuronal plasticity. Yet, they open new frontiers for research in that the specific genes they reveal are not necessarily the

Box 1: *RGS4*

This regulator of G-protein signaling acts across many receptors as a GTPase accelerating the dissociation of G-alpha bound GTP and shortening the signaling. Decreased expression of RGS4 in the dorsolateral prefrontal cortex of schizophrenic subjects was first discovered by microarray and confirmed by *in situ* hybridization [60]. Since this gene maps to 1q21-q22, a locus linked to schizophrenia, it was tested as a candidate gene in two US populations and one Indian population using both linkage analysis and a population case/control approach [107]. Four SNPs located in the first intron showed significant association in the family-based tests. Moreover, families with affected siblings showed greater allele sharing and identity by descent. However, the case-control population-based comparisons did not yield significant results. While the global association for all 4 SNPS was highly significant, the transmitted alleles and haplotypes differed across populations even within the United States. Thus, whether this truly reflects a vulnerability gene in schizophrenia remains to be fully established.

Box 2: *NPY*

Neuropeptide Y (NPY), a peptidergic neurotransmitter, may represent the best example of implicating a gene in schizophrenia by moving from gene expression to genotype. In some ways, NPY can be construed as a functional candidate since alterations in NPY function in schizophrenic cortex was first discovered by examining peptide content [110] and other findings have suggested potential alterations in NPY release in schizophrenia. However, it was the more recent genomic findings that propelled its examination as a susceptibility gene. Kuromitsu et al [111] used DNA microarrays and reported that NPY mRNA levels are significantly decreased in the prefrontal cortex of subjects suffering from both schizophrenia and bipolar disorder. An association study in a Japanese cohort [108] led to the discovery of a novel polymorphism in the promoter region of the NPY gene, whereby the –485 C to T variation results in a significant decrease in promoter activity. The –485T allele was overly represented in the disease group (P=0.004), suggesting that it may decrease NPY expression in brain and confer susceptibility to schizophrenia.

obvious candidates that might be selected based on prior knowledge.

So, as previously suggested by neuroscientists, the psychiatric disorders are diseases of neurotransmission probably involving several transmitter families including monoaminergic, peptidergic, GABAergic and Glutamatergic mechanisms, as well the modulation of the duration or efficacy of synaptic transmission. However, the specific molecules relevant to the illness (NPY, NRG1, GRK3, RGS4) may not have been the first in line for testing if it were not for the genomic studies. Similarly, while there was evidence of inflammatory, oxidative, and apoptotic mechanisms involved in neurodegenerative disorders, the specific genes uncovered (e.g. GSTO1 and GSTO2) would not have been obvious candidates. However, it is important to note that we have not defined clear criteria to ascertain whether we have yet captured the most important candidates or combinations thereof.

Genomic Studies in Brain Still have Many Untapped Directions

The genes discovered so far appear to each explain a small percentage of the variance or likelihood of expressing a particular disorder. Knowledge of neural circuitry can be combined with genomic strategies to further focus on the critical genes. For example, if a particular splice variant of a gene is present in a circuit implicated in the illness, it would be this particular variant that might become the focus of the genomic and genetic investigation. Moreover, genomic strategies can reveal ensembles of altered genes, and these may be seen in anatomically distinct components of the neural circuit (e.g. alterations in a neurotransmitter gene expressed in the presynaptic neurons of one brain region and in its receptor expressed in the postsynaptic neurons of the target region). Such combinations of altered gene expression involving multiple genes within a relevant circuit would be informative in the search for novel susceptibility genes that are altered at the structural level. The role of selected and inbred rodent models for the study of neural disorders is also critical both to identify novel genes in the animal models or to follow up on the discoveries made in human. Thus, the convergence of neurobiology, genomics and genetics is likely to prove very powerful in pinpointing the susceptibility genes and in establishing the neural phenotype associated with the emergence and progression of brain disorders.

Comparison with other Approaches and Future Directions

The examples described above broadly capture the current status of three fields using the integration of genomics and transcriptional profiling to discover genes that contribute to complex disorders. They point to both strengths and limitations of this approach and underscore not only the promise but the technical and conceptual challenges that face the field.

Strengths

The work to date suggests that combining transcriptional profiling with genetic linkage analysis in order to identify novel candidates genes is likely to be more powerful than previous strategies that rely on *either approach alone*. Integrating profiling data with physical mapping data not only removes genes that are not good candidates, but also circumvents the fine-mapping or interval sequencing that is needed to identify candidate genes within QTLs.

Expression profiling can potentially capture the effect of interactions that alter gene expression, either among the vulnerability genes or between them and the environment. In doing so, expression profiling may point to some genes that would have never been investigated otherwise, However, as

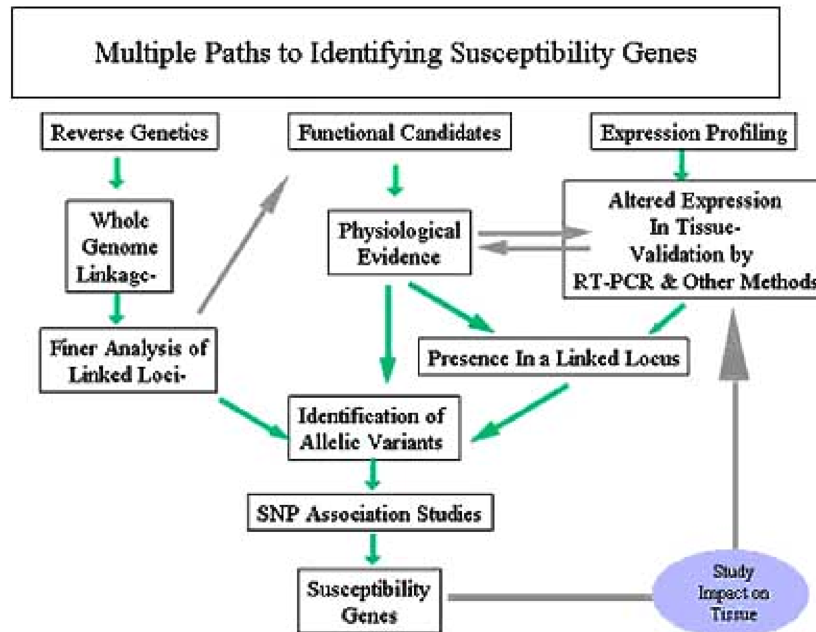

Fig. (1). Schematic Diagram of Relationship Between Expression Profiling and Genetic Approaches in identifying susceptibility genes associated with complex disorders.

discussed below, sophisticated informatics and statistical analysis strategies are needed to define the subsets of genes worthy of such pursuit.

Given the amount of information that is obtained using an integrated strategy of transcriptional profiling and positional cloning, this approach to studying the genetics of complex disorders may be fairly economical in terms of efficiency and cost, specifically as it can circumvent the need for individual gene-by-gene testing to determine disease-related changes. In addition, the expense of transcriptional profiling continues to drop, which makes it both more accessible and feasible for the study of larger populations. For the most part, the first level of validation of candidate genes that are identified by transcriptional profiling is no different than it is for candidate genes that are identified by other methods - requiring, in the case of animals, the development of subcongenics, transgenics, and knockouts; and in the case of humans, identification of polymorphisms, linkage and SNP analysis. However, understanding how the interaction of multiple genes results in disease vulnerability is likely to be more demanding.

Technical Limitations

An important limitation of transcriptional profiling is false positives (10-20% in our experience; R.D., unpublished observations), which necessitate the validation of transcriptional profiling results. Managing false positives can be accomplished in several ways, including performing experiments in triplicate and confirmation by traditional approaches, e.g., RT-PCR or Northern blot analysis or *in situ* hybridization, or by newer approaches that compare transcriptional profiles of DNA and oligonucleotide microrarrays or microarrays from different manufacturers.

The greatest limitation of approaches using transcriptional is lack of capture of all the genes that may be important (false negatives). The value of the transcriptional profiling is contingent on the number of genes and the number of different isoforms of genes, e.g., splice variants, represented on the microarray. Differences in gene expression levels might not be detected due to a lack of sensitivity for genes with low levels of expression, especially in a highly heterogeneous tissue such as the brain, in which many critical genes are expressed at low levels or in a very region- and cell-specific manner [75]. In addition, gene transcripts may be relevant to a phenotype in a tissue or cell-type not profiled. Moreover, polymorphisms not affecting transcript abundance, including those that lead to functional change in protein, alternative splicing, post-translational changes, and translational efficiency, will not be selected or highlighted by transcriptional profiling. Cell sorting or cell capture strategies focusing on the tissue types of interest, Improved gene arrays and new profiling technologies that allow the identification of SNPs and sequence differences may provide power in these areas in the future.

A limitation unique to postmortem human tissue is that the mode of death can affect gene expression, not randomly but in a systematic manner that can bias the results [57]. The effects are biologically interesting in pointing to the impact of certain agonal factors on gene function in particular tissues, but they can be dramatic and can therefore mask the gene expression changes associated with the disease.

Finally, analysis of data deriving from microarray technology represents an ongoing challenge. One needs to use the appropriate statistical tools to manage the issues of false positives and false negatives, to normalize data across samples and experiments, and to address problem of detection sensitivity and reliability [76]. These issues become particularly thorny when one is interested in changes of small magnitude, especially for genes that are near the limits of detection sensitivity, as is often the case for brain

studies. However, there are also some increasingly sophisticated strategies that analyze the coordinate changes in functionally related genes. This allows the detection of alterations in gene ensembles, even when the individual changes are modest, and allow one to assess whether the overall ensemble is statistically significant [77,78]. However, for these more powerful statistical approaches to be truly effective, they need to rely on increasingly more sophisticated ontological annotations of what constitutes ensembles of genes, not only in classical metabolic pathways, or even in intracellular or extracellular pathways, but also in physiological systems including neural circuits.

Conceptual Challenges

While the technical concerns are still significant, they are likely surmountable in whole or in part. The largest challenges facing the field of complex genetics are conceptual or strategic, and many begin after the putative novel candidates have been discovered. For Mendelian disorders the identification of the relevant gene represents a starting point for investigating the function of that gene, its role in the pathophysiology of the illness and its potential as a direct or indirect target for treatment. While each of these endeavors can be arduous, the overall research strategy is conceptually straightforward. By contrast, for complex genetic disorders, there are significant uncertainties at every step.

Firstly, even if we use the integrative approach discussed here, how do we know that we have identified all the key genetic players? For instance, when is it legitimate to say that the critical genes for susceptibility to Alzheimer's disease have been identified and the field can move on to study their mechanisms of action?. The answer might be that we will know it when we have accounted for a significant percentage of the variance-i.e. when we can statistically explain say over 90% of the cases based on combinations of identified vulnerability genes. Since we do not have yet a case of a complex genetic disorder where this has been reached, it is hard to determine how readily achievable this goal is. Some allelic variations may well contribute to a large proportion of the variance. But there will likely be others that may not be detected by standard strategies and that are nevertheless critical to discover, either for a subset of the population, or in combination with some other genes where they play a pivotal modifier role.

The other key issue is addressing the central characteristic of complex disorders- gene-gene interactions. Here again, thoughtful statistical analyses will be required.. Increasingly sophisticated mathematical strategies are being used for determining epistasis [79-81]. A particular challenge arises in discovering interactions when the main effects of the potential interacting genes or loci are not significant in themselves. This is where the gene expression profile, by pointing to ensembles of activity, with participating genes located in different loci, may help in the selection of combinations to test for epistasis.

It will also be critical to consider the animal models that will be required for testing the function of the multiple genes. The strategy for Mendelian disorders is to create knockout animals, or knock-in animals that harbor that mutation and use them as models. For studying complex disorders, animal models would be required that also address the issue of interactions. It may be possible to compare inbred strains, e.g., hypertensive rats, to examine differences in transcript abundance that may arise from epistatic interactions, both within and between regions of interest [82], and to study allele fixing in the process of breeding. Many rat strains are available and may have undergone genetic shift with outbreeding, e.g., the Sprague Dawley rat, and different inbreeding, e.g., Lewis rat. Further strategies, e.g., sub-congenics, double congenics, can be used to address these issues [83-86]. This is especially important when genome-wide scans are used due to the imprecision and influence of factors such as age, sex, and diet on QTLs mapped solely by linkage analysis [87].

The most direct proof that the combination of identified genes is necessary and sufficient for causing vulnerability to the disorder may necessitate the creation of transgenic animals with alterations of multiple genes (complex transgenics). Ideally, these animals would express the relevant allelic variations in more than one combination, representing the combinations most commonly seen in human populations. One would then ascertain how the combinations of all or a subset of those genes affects gene expression and function in the relevant tissue, and how this in turn results in vulnerability to specific environmental factors.

Testing of the role of individual or combinations of genes in disease vulnerability in humans will also require novel strategies, including better characterization of the clinical phenotype [88] and testing under different environmental conditions that would uncover the interaction of that environment with a given gene or set of genes [89,90].

In summary, the integration of transcriptional profiling with physical mapping data is proving useful in identifying candidate genes and gene patterns for some of the most challenging complex diseases. In addition to hypertension, cancer and brain disorders, ongoing applications of this strategy include other polygenic disorders, including arthritis, asthma, and inflammatory bowel disease [91-95]. As global profiling techniques become more accurate, as linkage analysis and genome-wide scanning improve, as the informatics and data analysis strategies become more sophisticated and as the conceptual challenges are confronted, these approaches should yield significant breakthroughs in understanding the biology of complex genetic disorders.

ACKNOWLEDGEMENTS

We would like to thank Dr. Margit Burmeister and Dr. David Burke (U. of Michigan) for very helpful discussions. This work was supported by NIH 1R21DK065628-01A1, American Heart Assoc Grant-In-Aid, and Phillip Morris Research Institute for RSD, and 5 P01 MH042251, L99 MH60398 and the Pritzker Consortium for Severe Psychiatric Disorders for HA.

REFERENCES

[1] Collins FS. Positional cloning moves from perditional to traditional. Nat Genet 1995; 9: 347-350.

[2] Mulligan LM, Kwok JB, Healey CS, *et al.* Germ-line mutations of the RET proto-oncogene in multiple endocrine neoplasia type 2A. Nature 1993; 363: 458-460.

[3] Leach FS, Nicolaides NC, Papadopoulos N, *et al.* Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell 1993; 75: 1215-1225.

[4] Nicolaides NC, Papadopoulos N, Liu B, *et al.* Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 1994; 371: 75-80.

[5] Papadopoulos N, Nicolaides NC, Wei YF, *et al.* Mutation of a mutL homolog in hereditary colon cancer. Science 1994; 263: 1625-1629.

[6] Lifton RP, Gharavi, AG, & Geller, DS Molecular mechanisms of human hypertension. Cell 2001; 104: 545-556.

[7] Carluccio M, Soccio, M, & De Caterina R. Aspects of gene polymorphisms in cardiovascular disease: the renin-angiotensin system. Eur J Clin Invest 2001; 31: 476-488.

[8] Hansson JH, Nelson-Williams C, Suzuki H, *et al.* - Hypertension caused by a truncated epithelial sodium channel gamma subunit: genetic heterogeneity of Liddle syndrome. Nat Genet 1995; 11(1): 76-82.

[9] Stewart PM, Krozowski ZS, Gupta A, *et al.* Hypertension in the syndrome of apparent mineralocorticoid excess due to mutation of the 11 beta-hydroxysteroid dehydrogenase type 2 gene. Lancet 1996; 347: 88-91.

[10] Mune T, Rogerson FM, Nikkila H, *et al.* Human hypertension caused by mutations in the kidney isozyme of 11 beta-hydroxysteroid dehydrogenase. Nat Genet 1995; 10: 394-399.

[11] Busjahn A, Aydin A, Uhlmann R, *et al.* Serum- and glucocorticoid-regulated kinase (SGK1) gene and blood pressure. Hypertension 2002; 40: 256-260.

[12] Wilson FH, Disse-Nicodeme S, Choate KA, *et al.* Human hypertension caused by mutations in WNK kinases. Science 2001; 293: 1107-1112.

[13] Lifton RP, Dluhy RG, Powers M, *et al.* The molecular basis of glucocorticoid-remediable aldosteronism, a Mendelian cause of human hypertension. Trans Assoc Am Physicians 1992; 105: 64-71.

[14] Lifton RP, Dluhy RG, Powers M, *et al.* A chimaeric 11 beta-hydroxylase/aldosterone synthase gene causes glucocorticoid-remediable aldosteronism and human hypertension. Nature 1992; 355: 262-265.

[15] Cover CM, Wang JM, St Lezin E, *et al.* Molecular variants in the P450c11AS gene as determinants of aldosterone synthase activity in the Dahl rat model of hypertension. J Biol Chem 1995; 270: 16555-16560.

[16] Churchill PC, Churchill MC, & Bidani A.K, *et al.* Kidney cross transplants in Dahl salt-sensitive and salt-resistant rats. Am J Physiol 1992; 262: H1809-H1817.

[17] Dahl LK, Knudsen KD, Heine, M, *et al.* Effects of chronic excess salt ingestion. Genetic influence on the development of salt hypertension in parabiotic rats: evidence for a humoral factor. J Exp Med 1967; 126: 687-699.

[18] Cicila GT, Rapp JP, Wang JM, *et al.* Linkage of 11 beta-hydroxylase mutations with altered steroid biosynthesis and blood pressure in the Dahl rat. Nat Genet 1993; 3: 346-353.

[19] Garrett MR, Saad, Y, Dene, H, *et al.* Blood pressure QTL that differentiate Dahl salt-sensitive and spontaneously hypertensive rats. Physiol Genomics 2000; 3: 33-38.

[20] Garrett MR, Dene H, Walder R, *et al.* Genome scan and congenic strains for blood pressure QTL using Dahl salt-sensitive rats. Genome Res 1998; 8: 711-723.

[21] Garrett MR, Joe B, Dene H, *et al.* Identification of blood pressure quantitative trait loci that differentiate two hypertensive strains. J Hypertens 2002; 20: 2399-2406.

[22] Rapp JP - Genetic analysis of inherited hypertension in the rat. Physiol Rev 2000; 80: (1), 135-172.

[23] Farjah, M, Washington TL, Roxas BP, *et al.* Dietary NaCl regulates renal aminopeptidase N: relevance to hypertension in the Dahl rat. Hypertension 2004; 43: 282-285.

[24] Liang M, Yuan B, Rute E, *et al.* Renal medullary genes in salt-sensitive hypertension: a chromosomal substitution and cDNA microarray study. Physiol Genomics 2002; 8: 139-149.

[25] Lee SJ, Liu J, Qi N, *et al.* Use of a panel of congenic strains to evaluate differentially expressed genes as candidate genes for blood pressure quantitative trait loci. Hypertens Res 2003; 26: 75-87.

[26] Okuda T, Sumiya T, Mizutani K, *et al.* Analyses of differential gene expression in genetic hypertensive rats by microarray. Hypertens Res 2002; 25: 249-255.

[27] Fornage M, Swank MW, Boerwinkle E, *et al.* Gene expression profiling and functional proteomic analysis reveal perturbed kinase-mediated signaling in genetic stroke susceptibility. Physiol Genomics 2003; 15: 75-83.

[28] Yamori Y, Tomimoto K, Ooshima A, *et al.* Proceedings: Developmental course of hypertension in the SHR-substrains susceptible to hypertensive cerebrovascular lesions. Jpn Heart J 1974; 15: 209-210.

[29] McBride MW, Carr FJ, Graham D, *et al.* Microarray analysis of rat chromosome 2 congenic strains. Hypertension 2003; 41: 847-853.

[30] Goffman, TE, Hassinger, DD, & Mulvihill, JJ, *et al.* Familial respiratory tract cancer. Opportunities for research and prevention. JAMA 1982; 247: 1020-1023.

[31] Lynch, HT, Kimberling WJ, Markvicka SE, *et al.* Genetics and smoking-associated cancers. A study of 485 families. Cancer 1986; 57: 1640-1646.

[32] McDuffie HH. Clustering of cancer in families of patients with primary lung cancer. J Clin Epidemiol 1991; 44: 69-76.

[33] Sellers TA, Ooi WL, Elston RC, *et al.* Increased familial risk for non-lung cancer among relatives of lung cancer patients. Am J Epidemiol 1987; 126: 237-246.

[34] Shaw GL, Falk, RT, Pickle LW, *et al.* Lung cancer risk associated with cancer in relatives. J Clin Epidemiol 1991; 44: 429-437.

[35] Tokuhata GK & Lilienfeld, AM. Familial aggregation of lung cancer in humans. J Natl Cancer Inst 1963; 30: 289-312.

[36] Sellers TA, Chen PL, Potter JD, *et al.* Segregation analysis of smoking-associated malignancies: evidence for Mendelian inheritance. Am J Med Genet 1994; 52: 308-314.

[37] Deman P. Cancer susceptibility in the mouse: genetics, biology and implications for human cancer. Nat Rev Genet 2003; 4: 721-734.

[38] Dragani TA. 10 years of mouse cancer modifier loci: human relevance. Cancer Res 2003; 63: 3011-3018.

[39] Herzog CR, Lubet RA, & You M. Genetic alterations in mouse lung tumors: implications for cancer chemoprevention. J Cell Biochem 1997; Suppl 28-29, 49-63.

[40] Aitman TJ, Glazier AM, Wallace CA, *et al.* Identification of Cd36 (Fat) as an insulin-resistance gene causing defective fatty acid and glucose metabolism in hypertensive rats. Nat Genet 1999; 21: 76-83.

[41] Karp CL, Grupe A, Schadt E, *et al.* Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma. Nat Immunol 2000; 1: 221-226.

[42] Lemon WJ, Bernert H, Sun H, *et al.* Identification of candidate lung cancer susceptibility genes in mouse using oligonucleotide arrays. J Med Genet 2002; 39: 644-655.

[43] Gariboldi M, Spinola M, Milani S, *et al.* Gene expression profile of normal lungs predicts genetic predisposition to lung cancer in mice. Carcinogenesis 2003; 24: 1819-1826.

[44] Miki R, Kadota K, Bono H, *et al.* Delineating developmental and metabolic pathways in vivo by expression profiling using the RIKEN set of 18,816 full-length enriched mouse cDNA arrays. Proc Natl Acad Sci USA 2001; 98: 2199-2204.

[45] Niculescu A B III, Segal DS, Kuczenski R, *et al.* Identifying a series of candidate genes for mania and psychosis: a convergent functional genomics approach. Physiol Genomics 2000; 4: 83-91.

[46] Baron M. Manic-depression genes and the new millennium: poised for discovery. Mol Psychiatry 2002; 7: 342-358.

[47] Badner JA & Gershon ES. Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. Mol Psychiatry 2002; 7: 405-411.

[48] Stefansson H, Sigurdsson E, Steinthorsdottir V, *et al.* Neuregulin 1 and susceptibility to schizophrenia. Am J Hum Genet 2002; 71: 877-892.

[49] Stefansson H, Sarginson J, Kong A, *et al.* Association of neuregulin 1 with schizophrenia confirmed in a Scottish population. Am J Hum Genet 2003; 72: 83-87.

[50] Williams NM, Preece A, Spurlock G, *et al.* Support for genetic variation in neuregulin 1 and susceptibility to schizophrenia. Mol Psychiatry 2003; 8: 485-487.

[51] Yang JZ, Si TM, Ruan Y, *et al.* Association study of neuregulin 1 gene with schizophrenia. Mol Psychiatry 2003; 8: 706-709.

[52] Shifman S, Bronstein M, Sternfeld M, *et al*. A highly significant association between a COMT haplotype and schizophrenia. Am J Hum Genet 2002; 71: 1296-1302.

[53] Egan MF, Goldberg TE, Kolachana BS, *et al.* Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. Proc Natl Acad Sci USA 2001; 98: 6917-6922 (2001).

[54] Akil M, Bhaskar S, Kolachana DA, *et al.* Catechol-O-methyltransferase genotype and dopamine regulation in the human brain. J Neurosci 2003; 23: 2008-2013.

[55] Harrison PJ & Owen MJ. Genes for schizophrenia? Recent findings and their pathophysiological implications. Lancet 2003; 361: 417-419.

[56] Evans SJ, Datson NA, Kabbaj M, *et al.* Evaluation of Affymetrix Gene Chip sensitivity in rat hippocampal tissue using SAGE analysis. Serial Analysis of Gene Expression. Eur J Neurosci 2002; 16: 409-413.

[57] Li YJ, Scott WK, Hedges DJ, *et al.* Age at onset in two common neurodegenerative diseases is genetically controlled. Am J Hum Genet 2002; 70: 985-993.

[58] Bunney WE, Bunney, BG, Vawter MP, *et al.* Microarray technology: a review of new strategies to discover candidate vulnerability genes in psychiatric disorders. Am J Psychiatry 2003; 160: 657-666.

[59] Mirnics K, Middleton FA, Lewis DA, *et al.* Analysis of complex brain disorders with gene expression microarrays: schizophrenia as a disease of the synapse. Trends Neurosci 2001; 24: 479-486.

[60] Mirnics K, Middleton FA, Stanwood, GD, *et al.* Disease-specific changes in regulator of G-protein signaling 4 (RGS4) expression in schizophrenia. Mol Psychiatry 2001; 6: 293-301.

[61] Mirnics K, Middleton FA, Marquez A, *et al.* Molecular characterization of schizophrenia viewed by microarray analysis of gene expression in prefrontal cortex. Neuron 2000; 28: 53-67.

[62] Middleton FA, Mirnics, K, Pierri, JN, *et al.* Gene expression profiling reveals alterations of specific metabolic pathways in schizophrenia. J Neurosci 2002; 2718-2729.

[63] Vawter MP, Crook JM, Hyde TM, *et al.* Microarray analysis of gene expression in the prefrontal cortex in schizophrenia: a preliminary study. Schizophr Res 2002; 58: 11-20.

[64] Vawter MP, Evans S, Choudary P, *et al.* Gender-specific gene expression in post-mortem human brain: localization to sex chromosomes. Neuropsychopharmacology 2004; 29: 373-384.

[65] Mimmack ML, Ryan, M, Baba H, *et al.* Gene expression analysis in schizophrenia: reproducible up-regulation of several members of the apolipoprotein L family located in a high-susceptibility locus for schizophrenia on chromosome 22. Proc Natl Acad Sci USA 2002; 99: 4680-4685.

[66] Hakak Y, Walker JR, Li C, *et al.* Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. Proc Natl Acad Sci USA 2001; 98: 4746-4751.

[67] Lehrmann E, Hyde TM, Vawter MP, *et al.* The use of microarrays to characterize neuropsychiatric disorders: postmortem studies of substance abuse and schizophrenia. Curr Mol Med 2003; 3: 437-446.

[68] Tkachev D, Mimmack ML, Ryan MM, *et al.* Oligodendrocyte dysfunction in schizophrenia and bipolar disorder. Lancet 2003; 362: 798-805.

[69] Davis KL & Haroutunian V. Global expression-profiling studies and oligodendrocyte dysfunction in schizophrenia and bipolar disorder. Lancet 2003; 362: 758.

[70] Barrett TB, Hauger RL, Kennedy JL, *et al.* Evidence that a single nucleotide polymorphism in the promoter of the G protein receptor kinase 3 gene is associated with bipolar disorder. Mol Psychiatry 2003: 8: 546-557.

[71] Corder EH, Saunders AM, Strittmatter WJ, *et al.* Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 1993: 261: 921-923.

[72] Martin ER, Scott WK, Nance MA, *et al.* Association of single-nucleotide polymorphisms of the tau gene with late-onset Parkinson disease. JAMA 2001: 286: 2245-2250.

[73] Hauser MA, Li YJ, Takeuchi S, *et al.* Genomic convergence: identifying candidate genes for Parkinson's disease by combining serial analysis of gene expression and genetic linkage. Hum Mol Genet 2003: 12: 671-677.

[74] Li YJ, Oliveira SA, Xu P, *et al.* Glutathione S-transferase omega-1 modifiesage-at-onset of Alzheimer disease and Parkinson disease. Hum Mol Genet 2003: 12: 3259-3267.

[75] Evans SJ, Choudary PV, Vawter MP, *et al.* DNA microarray analysis of functionally discrete human brain regions reveals divergent transcriptional profiles. Neurobiol Dis 2003: 14: 240-250.

[76] Irizarry RA, Hobbs B, Collin F, *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003: 4: 249-264.

[77] Beissbarth T. & Speed TP. GOstat: find statistically overrepresented Gene Ontologies within a group of genes. Bioinformatics 2004; 20: 1464-1465.

[78] Mootha VK, Lindgren CM, Eriksson, *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003: 34: 267-273.

[79] Cordell H J. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Hum Mol Genet 2002; 11: 2463-2468.

[80] Hoh J. & Ott J. Mathematical multi-locus approaches to localizing complex human trait genes. Nat Rev Genet 2003: 4: 701-709.

[81] Yi N, Xu S, & Allison DB. Bayesian model choice and search strategies for mapping interacting quantitative trait Loci. Genetics 2003; 165: 867-883.

[82] Dracopoli NC. Pharmacogenomic applications in clinical drug development. Cancer Chemother Pharmacol 2003; 52 Suppl 1: S57-S60.

[83] Rapp JP, Garrett MR, & Deng AY. Construction of a double congenic strain to prove an epistatic interaction on blood pressure between rat chromosomes 2 and 10. J Clin Invest 1998; 101: 1591-1595.

[84] Pravenec M, Landa V, Zidek V, *et al.* Transgenic rescue of defective Cd36 ameliorates insulin resistance in spontaneously hypertensive rats. Nat Genet 2001; 27: 156-158.

[85] Hubner N, Lee YA, Lindpaintner K, Ganten D, & Kreutz R. Congenic substitution mapping excludes Sa as a candidate gene locus for a blood pressure quantitative trait locus on rat chromosome 1. Hypertension 1999; 34: 643-648 (1999).

[86] Saad Y, Garrett MR, & Rapp JP. Multiple blood pressure QTL on rat chromosome 1 defined by Dahl rat congenic strains. Physiol Genomics 2001; 4: 201-214.

[87] Kato N, Mashimo T, Nabika T, *et al.* Genome-wide searches for blood pressure quantitative trait loci in the stroke-prone spontaneously hypertensive rat of a Japanese colony. J Hypertens 2003; 21: 295-303.

[88] Merikangas KR. & Risch N. Will the genomics revolution revolutionize psychiatry? Am J Psychiatry 2003; 160: 625-635.

[89] Caspi A, Sugden K, Moffitt TE, *et al.* Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science 2003; 301: 386-389.

[90] Caspi A, McClay J, Moffitt *et al.* Role of genotype in the cycle of violence in maltreated children. Science 2002; 297: 851-854.

[91] Wester L, Koczan D, Holmberg J, *et al.* Differential gene expression in pristane-induced arthritis susceptible DA versus resistant E3 rats. Arthritis Res Ther 2003; 5: R361-R372.

[92] Heller RA, Schena M, Chai A, *et al.* Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. Proc Natl Acad Sci USA 1997; 94: 2150-2155.

[93] Zou J, Young S, Zhu F, *et al.* Microarray profile of differentially expressed genes in a monkey model of allergic asthma. Genome Biol 2002; 3: research0020.

[94] Zou J, Young S, Zhu F, *et al.* Identification of differentially expressed genes in a monkey model of allergic asthma by microarray technology. Chest 2002; 121: 26S-27S.

[95] Sheppard D, Roger S. Mitchell lecture. Uses of expression microarrays in studies of pulmonary fibrosis, asthma, acute lung injury, and emphysema. Chest 2002; 121: 21S-25S.

[96] Saad Y, Garrett M R, & Rapp J P. - Multiple blood pressure QTL on rat chromosome 1 defined by Dahl rat congenic strains. Physiol Genomics 2001; 4(3): 201-214.

[97] Farjah M, Roxas BP, Geenen DL, & Danziger RS. Dietary salt regulates renal SGK1 abundance: relevance to salt sensitivity in the Dahl rat. Hypertension 2003; 41: 874-878.

[98] Liang M, Yuan B, Rute E, *et al.* Insights into Dahl salt-sensitive hypertension revealed by temporal patterns of renal medullary gene expression. Physiol Genomics 2003; 12: 229-237.

[99] Rubattu S, Volpe M, Kreutz R, *et al.* Chromosomal mapping of quantitative trait loci contributing to stroke in a rat model of complex human disease. Nat Genet 1996; 13: 429-434.
[100] Zhang Z, Futamura M, Vikis HG, *et al.* Positional cloning of the major quantitative trait locus underlying lung tumor susceptibility in mice. Proc Natl Acad Sci USA 2003; 100: 12642-12647.
[101] Manenti G, Galbiati F, Gianni-Barrera R, *et al.* Haplotype sharing suggests that a genomic segment containing six genes accounts for the pulmonary adenoma susceptibility 1 (Pas1) locus activity in mice. Oncogene 2004.Apr 5.
[102] Wang M, Lemon WJ, Liu G, *et al.* Fine mapping and identification of candidate pulmonary adenoma susceptibility 1 genes using advanced intercross lines. Cancer Res 2003; 63: 3317-3324.
[103] Manenti G, Gariboldi M, Elango R, *et al.* Genetic mapping of a pulmonary adenoma resistance (Par1) in mouse. Nat Genet 1996; 12: 455-457.
[104] Pataer A, Nishimura M, Kamoto, *et al.* Genetic resistance to urethan-induced pulmonary adenomas in SMXA recombinant inbred mouse strains. Cancer Res 1997; 57: 2904-2908.
[105] Lee GH, Matsushita H, & Kitagawa T. Fine chromosomal localization of the mouse Par2 gene that confers resistance against urethane-induction of pulmonary adenomas. Oncogene 2001; 20: 3979-3985.
[106] Festing MF, Lin L, Devereux TR, *et al.* At least four loci and gender are associated with susceptibility to the chemical induction of lung adenomas in A/J x BALB/c mice. Genomics 1998; 53: 129-136.
[107] Chowdari KV, Mirnics K, Semwal P, *et al.* Association and linkage analyses of RGS4 polymorphisms in schizophrenia. Hum Mol Genet 2002; 11: 1373-1380.
[108] Itokawa M, Arai M, Kato S, *et al.* Association between a novel polymorphism in the promoter region of the neuropeptide Y gene and schizophrenia in humans. Neurosci Lett 2003; 347: 202-204.
[109] Liang T, Spence J, Liu L, *et al.* alpha-Synuclein maps to a quantitative trait locus for alcohol preference and is differentially expressed in alcohol-preferring and -nonpreferring rats. Proc Natl Acad Sci USA 2003; 100: 4690-4695.
[110] Frederiksen SO, Ekman R, Gottfries CG, Widerlov E, & Jonsson S. Reduced concentrations of galanin, arginine vasopressin, neuropeptide Y and peptide YY in the temporal cortex but not in the hypothalamus of brains from schizophrenics. Acta Psychiatr Scand 1991; 83: 273-277.
[111] Kuromitsu J, Yokoi A, Kawai T, *et al.* Reduced neuropeptide Y mRNA levels in the frontal cortex of people with schizophrenia and bipolar disorder. Gene Expr Patt 2004; 1: 17-21.

Received: September 09, 2004 Accepted: September 21, 2004