Background: Tissue gene expression profiling with arrays measures the transcription of thousands of genes. However, this approach cannot be readily used to guide clinical neurologic practice.

Objectives: To determine whether clinical neurologic diseases are associated with unique patterns of up- and down-regulated genes in whole blood and to explore the possibility of using peripheral blood as a surrogate tissue in these diseases.

Design: Case-control study.

Setting: University-based pediatric and adult neurology clinics.

Participants: Patients with neurofibromatosis type 1, epilepsy, or Tourette syndrome diagnosed using traditional clinical criteria; controls without disease; and controls with neurologic disease.

Main Outcome Measure: Blood gene expression levels of greater than 12,000 genes, measured using U95A arrays.

Results: Neurofibromatosis type 1 and childhood epilepsy treated with carbamazepine or valproic acid are associated with distinct patterns of blood gene expression. Patients with valproic acid–responsive vs valproic acid–refractory epilepsy formed distinct subclusters. Tourette syndrome was characterized by several gene expression clusters. In 1 cluster, 6 genes— all associated with immune cell function—were overexpressed.

Conclusion: Blood gene expression profiling can provide surrogate markers for neurologic diseases without obvious blood phenotypes.
lepsy. Because many genes may contribute to an inherited basis of drug response, we reasoned that a high-throughput approach might yield important insights.

Our final aim was to explore whether blood gene expression patterns distinguish complex, heritable neurologic diseases for which no causative gene has been identified, which we tested in Tourette syndrome (TS). This syndrome seems to follow an autosomal dominant inheritance pattern, but linkage analyses have been unsuccessful, and multiple genetic or environmental factors may contribute to the phenotype.

METHODS

STUDY PARTICIPANTS

A total of 129 individuals participated in the study: control subjects without neurologic disease (n=13; 7 males and 6 females; mean±SD age, 27.7±6.9 years), children with headache (n=15; 7 boys and 8 girls; mean±SD age, 12.7±3.0 years), children with epilepsy (n=24; 12 boys and 12 girls; mean±SD age, 10.6±3.6 years), patients with NF1 (n=12; 7 males and 5 females; mean±SD age, 28.0±19.4 years), patients with bipolar disorder or schizophrenia (n=17; 15 men and 2 women; mean±SD age, 54.4±8.4 years), adults with epilepsy (n=20; 7 men and 13 women; mean±SD age, 35.2±10.5 years), adults with Parkinson disease or progressive supranuclear palsy (n=12; 5 men and 7 women; mean±SD age, 65.6±12.6 years), and patients with TS (n=16; 12 males and 4 females; mean±SD age, 17.4±13.9 years).

Diagnoses of NF1 were based on established clinical criteria. Children with epilepsy were treated with valproic acid (n=11) or carbamazepine (n=6) or were newly diagnosed and untreated (n=7). Patients with TS had to meet Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, criteria for TS and have at least 1 affected first-degree family member. Protocols were approved by the institutional review boards at the Cincinnati Children's Hospital Medical Center and the University of Cincinnati, and all participants or their guardians provided informed consent.

BLOOD SAMPLE COLLECTION AND PROCESSING

Ten- to 15-mL blood samples were collected in tubes containing EDTA and then were mixed with Trizol LS reagent (Invitrogen, Carlsbad, Calif) within 15 minutes. Total RNA was isolated and purified using the RNeasy mini kit (Qiagen Inc, Valencia, Calif). The quality of total RNA was assessed using a bioanalyzer (Agilent 2100; Agilent Technologies Inc, Palo Alto, Calif) and was quantified by spectrophotometry.

MICROARRAY PROCESSING AND DATA ANALYSIS

Sample labeling, hybridization to U95A arrays (>12000 sequences), and image scanning were performed as described in the Affymetrix Expression Analysis Technical Manual. Arrays were normalized using the invariant set normalization method. Expression values were calculated using the perfect match–only model-based expression index and data analysis software (dChip version 1.2; Harvard University, Cambridge, Mass).

INTERSUBJECT AND INTRASUBJECT VARIATION

The variation of gene expression in human blood has been related to many factors, which would tend to obscure important correlations (type II error) or create spurious correlations (type I error) when comparing disease cases with controls. We...
therefore analyzed 266 highly variably expressed genes in 14 samples taken on separate days from 7 healthy donors using expression data analysis software (GeneSpring version 6.0; Silicon Genetics, Redwood City, Calif), which performs “unsupervised cluster analysis,” pairing samples by the highest correlation of gene expression.13 If the intraindividual gene expression patterns are stable, reflecting the genetic and environmental factors unique to each individual, the program clusters each individual’s samples.

**RESULTS**

**VARIATION OF BLOOD GENE EXPRESSION IN CONTROLS WITHOUT DISEASE**

Unsupervised cluster analysis of the 266 genes with the highest variation paired individuals’ samples (data not shown), indicating that variation in blood gene expression related to the temporal and technical reasons is smaller than the interindividual variation.

**BLOOD GENOMIC EXPRESSION PATTERN OF NF1**

For the comparisons between patients with NF1 and the 3 sets of controls, 24, 55, and 58 genes were differentially expressed in NF1 blood compared with that of controls, and the probability of generating the same number of genes by chance was 2.3%, 0.2%, and 0.7%,
respectively. Hierarchical cluster analysis using these gene lists demonstrated that the NF1 samples clustered separately from each set of controls (Figure 1). In contrast, the 3 comparisons between controls produced 1, 3, and 0 genes, and the probability of generating the same number of genes by chance was 70%, 29%, and 100%, respectively. Of pathophysiologic interest, many dysregulated genes in NF1 blood are related to tissue remodeling and tumorigenesis.5

**BLOOD GENOMIC EXPRESSION PATTERNS IN CHILDREN WITH EPILEPSY**

The global permutation-based test showed that the expression patterns in blood caused by valproic acid and carbamazepine are significant ($P = .005$ for valproic acid vs no treatment and $P = .02$ for carbamazepine vs no treatment). Cluster analysis segregated the valproic acid samples into 2 subclusters. One subcluster included all 3 valproic acid–resistant patients, and the other included all 8 valproic acid–responsive patients (Figure 2). Many mitochondrial genes are overexpressed in valproic acid responders, pointing to the possible involvement of mitochondria in the determination of valproic acid efficacy.6

**BLOOD GENOMIC EXPRESSION IN TS**

Permutation analysis showed that the blood gene expression pattern associated with TS has a $P = .20$. In other words, 20% of random permutations of the class label...
generated the same number of up- and down-regulated genes. Thus, we did not find evidence that the clinical diagnosis of TS is associated with a single, unique gene expression profile in whole blood.

Subgroup analysis showed that there were 6 up-regulated genes and 1 down-regulated gene in TS ($P < .05$ for each of 8 comparisons) (Figure 3). Granzyme B is involved in the target-killing process of cytotoxic T cells or natural killer cells. NKG2E encodes a lectin-like receptor, which plays a role in the recognition of the major histocompatibility complex molecules by natural killer cells and some cytotoxic T cells. CD94 is also preferentially expressed by natural killer cells and forms heterodimers with NKG subunits. NK-p46 participates in natural killer cell–mediated lysis of cells infected with intracellular bacteria. IMPA2, expressed in caudate, a region shown in neuroimaging studies to be involved in TS and obsessive-compulsive disorder, was down-regulated.

Post hoc, K-means cluster analysis for the 6 cytotoxic T-cell/natural killer cell genes yielded cluster A (low expressers) and cluster B (high expressers). The proportion of patients with TS in the higher expression group was significantly greater than the proportions in the control groups ($P < .05$ by $\chi^2$ test). Within TS, permutation-based analysis showed that less than 4% of random permutations generated the same number of differentially expressed genes as those in clusters A and B (Figure 4).

Global gene expression profiling holds potential for classifying diseases and predicting clinical outcomes based on molecular criteria. Because blood is the most accessible tissue, blood gene expression profiling has been used to explore hematologic malignancies and autoimmune and infectious disorders. Our studies suggest that this approach can be extended to brain diseases.

Brain tissue is not accessible in vivo for most neurologic and psychiatric diseases. Our findings in NF1 support the notion that monogenetic neurologic or multiorgan disorders can be identified by distinct blood gene expression profiles. The results of our pediatric epilepsy and TS studies suggest that this approach may be extended to complex disorders. Blood genomic profiling might provide an accessible platform to categorize a polygenic condition into meaningful molecular subtypes. As exemplified in the anticonvulsant study, part of the interindividual variation in valproic acid efficacy may be related to an identifiable, efficacy-related drug effect at the blood transcription level.

We did not identify a unique, important blood expression pattern in TS. However, because TS may not be a monogenetic disorder, our findings may be consistent with the hypothesis that TS is the result of multiple genetic or environmental factors, creating multiple phenotypes. The most useful application of unbiased, high-throughput technologies like gene expression profiling might be to identify profiles that distinguish molecular subtypes of complex disorders, from which the association of genomic pattern and various clinical features can be probed. Our findings relating to the functioning of immune cells is particularly intriguing, given studies that suggest that autoimmune mechanisms triggered by infection with group A $\beta$-hemolytic
streptococci are involved in the pathogenesis of some patients with TS.

The results of these and future blood gene expression studies in neurologic diseases must be interpreted cautiously. Most genes regulated in blood have low fold changes and high variability. The low fold change is not unexpected considering the absence of obvious blood phenotypes. The variation in gene expression patterns in peripheral blood come from multiple non–disease-related sources. RNA quality is also critical.

Finally, the internal statistical validation techniques, for example, permutation analysis, used in this pilot study have limitations. Among genes that seem to be up- and down-regulated, there are disease-related and artificial changes. When internal validation methods demonstrate that the number of abnormally expressed genes exceeds what would be expected by chance, this does not confirm that an individual gene’s expression level is disease related. Selectively using quantitative reverse transcription polymerase chain reaction (data not shown) confirms the gene expression levels measured by the array but does not discriminate between spurious statistical differences that result from multiple comparisons vs real differences associated with disease. The predictive value of any set of dysregulated genes ultimately requires external validation, for example, cosegregation of gene expression profiles with differential survival or treatment response, comorbidity, or inheritance patterns, or abnormal expression of genes in disease-causing pathways. A final test is to apply the identified genes as a prediction rule in a second cohort to determine whether the genomic profile accurately distinguishes diseased individuals from controls.

Accepted for publication: July 1, 2004.

Correspondence: Donald L. Gilbert, MD, MS, Division of Neurology, Cincinnati Children’s Hospital Medical Center, ML2015, 3333 Burnet Ave, Cincinnati, OH 45229-3039 (d.gilbert@cchmc.org).

Author Contributions: Study concept and design: Tang, Gilbert, Glauser, and Sharp. Acquisition of data: Tang, Gilbert, Glauser, and Hershey. Analysis and interpretation of data: Tang and Gilbert. Drafting of the manuscript: Tang, Gilbert, and Sharp. Critical revision of the manuscript for important intellectual content: Tang, Gilbert, Glauser, and Hershey. Statistical analysis: Tang. Obtained funding: Sharp. Administrative, technical, and material support: Gilbert, Glauser, and Hershey. Study supervision: Sharp.

Funding/Support: This study was supported by grants NS41920 (Dr Gilbert); NS28167, AG19561, NS38084, NS42774, and NS43252 and an American Heart Association Bugher Award (Dr Sharp); NS040261 and NS044956 (Dr Glauser); and NS045752 (Dr Hershey) from the National Institutes of Health, Bethesda, Md.

Previous Presentation: This study was presented in part at the 32nd Annual Child Neurology Society Meeting; October 2, 2003; Miami, Fla.

REFERENCES


