American Journal of Medical Genetics Part B (Neuropsychiatric Genetics) 136B:12–25 (2005)

Gene Expression Analysis of Peripheral Blood Leukocytes From Discordant Sib-Pairs With Schizophrenia and Bipolar Disorder Reveals Points of Convergence Between Genetic and Functional Genomic Approaches

Frank A. Middleton,[1,2,3]* Carlos N. Pato,[1,2,4,5] Karen L. Gentile,[1,3] Lindsay McGann,[1,3] Andrea M. Brown,[1,3] Marco Trauzzi,[1,3] Heba Diab,[1,3] Christopher P. Morley,[1,2] Helena Medeiros,[1,2] Antonio Macedo,[6] M. Helena Azevedo,[6] and Michele T. Pato[1,2,4,5]

[1]*Center for Neuropsychiatric Genetics, Upstate Medical University, Syracuse, New York*
[2]*Department of Psychiatry, Upstate Medical University, Syracuse, New York*
[3]*Department of Neuroscience and Physiology, Upstate Medical University, Syracuse, New York*
[4]*Department of Psychiatry, Georgetown University, Washington, DC*
[5]*Veterans Administration Medical Center, Washington, DC*
[6]*Psicologia Medica, Universidade de Coimbra, Coimbra, Portugal*

We performed global RNA transcript analysis and comprehensive gene group analysis of peripheral blood leukocyte (PBL) RNA from two groups of matched sib-pairs that were discordant for either schizophrenia (n = 33 sib-pairs) or bipolar disorder (n = 5 sib-pairs). The pairs chosen for these analyses were selected from families with known patterns of genetic linkage (5q for schizophrenia and 6q for bipolar disorder). At the single gene level, we obtained lists of the transcripts with the most significant changes in expression and from these lists determined those with the highest degree of predictive power for classifying subjects according to diagnosis in these samples. At the gene group level, we comprehensively analyzed pairwise expression changes of more than 4,000 functional groups and cytogenetic locations, and present a novel method of displaying these data that we term "cytogenomic" mapping. Verification of selected changes in expression was performed using quantitative real-time RT-PCR. Our results provide compelling evidence for the utility of analyzing PBL RNA for changes in expression in neuropsychiatric disorders.
© 2005 Wiley-Liss, Inc.

KEY WORDS: microarray; GeneChip; white blood cell; RNA

The work was performed at Center for Neuropsychiatric Genetics, Upstate Medical University, Syracuse, NY 13210; Department of Psychiatry, Upstate Medical University, Syracuse, NY 13210; and Department of Neuroscience and Physiology, Upstate Medical University, Syracuse, NY 13210.

Grant sponsor: VA Merit Award; Grant sponsor: NIMH; Grant numbers: MH52618, MH058693.

*Correspondence to: Dr. Frank A. Middleton, Department of Neuroscience and Physiology, 3281 Weiskotten Hall, Upstate Medical University, 750 East Adams Street, Syracuse, NY 13210. E-mail: middletf@upstate.edu

Received 20 September 2004; Accepted 2 December 2004

DOI 10.1002/ajmg.b.30171

© 2005 Wiley-Liss, Inc.

INTRODUCTION

We have recently reported strong suggestive genome-wide linkages for both schizophrenia and bipolar disorder in pedigrees from the Portuguese Island Collection [PIC; see Pato et al., 2004; Sklar et al., 2004]. In a follow-up study of bipolar disorder using a denser SNP-based genotyping method, we achieved genome-wide significance for linkage at 6q22 [Middleton et al., 2004]. In schizophrenia, the peak linkage signals were obtained on chromosome 5q31-35. In order to help expedite the selection of candidate genes in these (and other) regions, we have implemented a strategy involving the measurement of changes in gene expression in peripheral blood leukocytes (PBLs) in discordant sib-pairs from the two pedigree sets.

The analysis of mRNA transcript levels in PBLs from living subjects offers several advantages compared with studies involving only end point postmortem tissue specimens. Such advantages include the ability to completely match subject characteristics such as age, gender, family background, time of blood draw, geographical/environmental variables, diet, and the cellular composition of the samples. Moreover, it also becomes feasible to design studies that examine expression profiles of PBLs during the progression of the disease, or in response to drug treatments. These gene expression patterns, when obtained in well-controlled studies, have increased the power to help refine candidate gene selection for mutational screening, as well as obtain lists of genes with predictive power for classifying different diseases and their treatment response. Indeed, in the cancer field, it has become standard practice at some clinics that specialize in childhood leukemias to compare the expression profiles of PBLs or lymphocytes on all patients. These expression patterns not only correctly subtype the leukemia, but also serve as highly accurate predictors of the disease course [e.g., Yeoh et al., 2002].

In neuropsychiatric diseases, no studies to date have been published on the potential utility of gene expression profiling of PBLs for either diagnosis or disease characterization. There has been, however, at least one study to date examining the gene expression profile of transformed lymphoblasts from a small number of subjects with schizophrenia in a single pedigree, which showed promising results [Vawter et al., 2004]. In the present, study, we wished to determine the changes in gene expression in a larger sample involving 33 sib-pairs from families segregating for schizophrenia from the PIC population where we had detected strong linkage to 5q.

TABLE I. Largest Significant Pairwise Expression Changes in Schizophrenia Sib Pairs

Probe ID	Fold change	*P*-value	Gene title	Symbol	Location
Increased expression					
201743_at	2.53	0.00108	CD14 antigen	CD14	5q31.1
202437_s_at	2.36	0.00864	Cytochrome P450, family 1, subfamily B, polypeptide 1	CYP1B1	2p21
204614_at	2.33	0.00331	Serine (or cysteine) proteinese inhibitor, clade B (ovalbumin), member 2	SERPINB2	18q21.3
201109_s_at	2.16	0.00990	Thrombospondin 1	THBS1	15q15
221731_x_at	2.08	0.01777	Chondroitin sulfate proteoglycan 2 (versican)	CSPG2	5q14.3
212681_at	2.04	0.02058	Erythrocyte membrane protein band 4.1-like 3	EPB41L3	18p11.32
210111_s_at	2.02	0.00043	KIAA0265 protein	KIAA0265	7q32.3
205098_at	1.97	0.00330	Chemokine (C-C motif) receptor 1	CCR1	3p21
217996_at	1.96	0.00407	Pleckstrin homology-like domain, family A, member 1	PHLDA1	12q15
202435_s_at	1.96	0.03702	Cytochrome P450, family 1, subfamily B, polypeptide 1	CYP1B1	2p21
201110_s_at	1.90	0.04771	Thrombospondin 1	THBS1	15q15
20436_s_at	1.90	0.04004	Cytochrome P450, family 1, subfamily B, polypeptide 1	CYP1B1	2p21
218559_s_at	1.86	0.00091	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	MAFB	20q11.2-q13.1
217997_at	1.85	0.01743	Pleckstrin homology-like domain, family A, member 1	PHLDA1	12q15
212636_at	1.82	0.00120	Quaking homolog, KH domain RNA binding (mouse)	QKI	6q26-27
211776_s_at	1.80	0.01556	Erythrocyte membrane protein band 4.1-like 3	EPB41L.3	18p11.32
206710_s_at	1.80	0.01838	Erythrocyte membrane protein band 4.1-like 3	EPB41L.3	18p11.32
204470_at	1.77	0.02088	Chemokine (C-X-C motif) ligand 2	CXCL2	4q21
212993_at	1.77	0.01313	Sin3-associated polypetide, 18 kDa	–	9q34.3
218195_at	1.76	0.00527	Chromosome 6 open reading frame 211	C6orf211	6q25.1
208892_a_at	1.75	0.01232	Dual specificity phosphatase 6	DUSP6	12q22-q23
204620_s_at	1.75	0.03852	Chondroitin sulfate proteoglycan 2 (versican)	CSPG2	5q14.3
201694_s_at	1.74	0.01116	Early growth response 1	EGR1	5q31.1
213836_s_at	1.74	0.00421	Hypothetical protein FLJ10055	FLJ10055	17q24.3
206343_s_at	1.73	0.01010	Neuregulin 1	NRG1	8p21-p12
204049_s_at	1.72	0.02217	Phosphatase and actin regulator 2	PHACTR2	6q24.1
204619_s_at	1.69	0.03953	Chondroitin sulfate proteoglycan 2 (versican)	CSPG2	5q14.3
201108_s_at	1.69	0.02084	Thrombospondin 1	THBS1	15q15
203973_s_at	1.67	0.00375	CCAAT/enhancer binding protein (C/EBP), delta	CEBPD	8p11.2-p11.1
205863_at	1.67	0.03650	S100 calcium binding protein A12 (calgranulin C)	S100A12	1q21
207719_x_at	1.66	0.02342	KARP-1 binding protein	KAB	1q44
222243_s_at	1.65	0.00014	Transducer of ERBB2, 2	TOB2	22q13.2-q13.31
205922_at	1.65	0.00371	Vanin 2	VNN2	6q23-q24
222028_at	1.65	0.00093	Zinc finger protein 45 (a Kruppel-associated box (KRAB) domain polypeptide)	ZNF45	19q13.2
220088_at	1.64	0.03512	Complement component 5 receptor 1 (C5a ligand)	C5R1	19p13.3-q13.4
205495_s_at	1.63	0.01893	Granulysin	GNLY	2p12-q11
210844_x_at	1.63	0.00142	Catenin (cadherin-associated protein), alpha 1, 102 kDa	CTNNA1	5q31
208891_at	1.63	0.01859	Dual specificity phosphatase 6	DUSP6	12q22-q23
208716_s_at	1.63	0.00689	Putative membrane protein	LOC54499	1q22-q25
200765_x_at	1.62	0.00056	Catenin (cadherin-associated protein), alpha 1, 102 kDa	CTNNA1	5q31
Decreased expression					
209170_s_at	−1.89	0.01378	Glycoprotein M6B	GPM6B	Xp22.2
213797_at	−1.65	0.02433	Viperin	cig5	2p25.2
212621_at	−1.63	0.01017	KIAA0286 protein	KIAA0286	12q13.2
214059_at	−1.61	0.03891	Interferon-induced protein 44	IFI44	1p31.1
219863_at	−1.61	0.01801	Cyclin-E binding protein 1	CEB1	4q22.1-q23
210797_s_at	−1.60	0.00820	2′-5′-oligoadenylate synthetase-like	OASL	12q24.2
214453_s_at	−1.55	0.04213	Interferon-induced protein 44	IFI44	1p31.1
216252_x_at	−1.50	0.02649	Tumor necrosis factor receptor superfamily, member 6	TNFRSF6	10q24.1
205660_at	−1.47	0.00608	2′-5′-oligoadenylate synthetase-like	OASL	12q24.2
210676_x_at	−1.45	0.01189	RAN binding protein 2-like 1	RANBP2L1	2q13
204780_s_at	−1.44	0.00725	Tumor necrosis factor receptor superfamily, member 6	TNFRSF6	10q24.1
210425_x_at	−1.44	0.03848	Golgin-67	GOLGIN-67	15q11.2
204747_at	−1.42	0.02484	Interferon-induced protein with tetratricopeptide repeats 4	IFIT4	10q24
215831_at	−1.42	0.02860	PRO1621 protein	PRO1621	11
203992_s_at	−1.41	0.00763	Ubiquitously transcribed tetratricopeptide repeat, X chromosome	UTX	Xp11.2
220104_at	−1.41	0.01244	Zinc finger CCCH type, antiviral 1	ZC3HAV1	7q34
213703_at	−1.36	0.00483	Hypothetical protein LOC150759	LOC150759	2q11.2
204083_s_at	−1.35	0.02416	Tropomyosin 2 (beta)	TPM2	9p13.2-p13.1
204369_at	−1.34	0.04573	Phosphoinositide-3-kinase, catalytic, alpha polypeptide	PIK3CA	3q26.3
216358_at	−1.33	0.01088	SWI/SNF related, matrix associated actin dependent regulator of chromatin, subfamily e1	SMARCE1	17q21.2
217506_at	−1.30	0.00600	Transcribed sequence with moderate similarity to hypothetical protein FLJ20378	—	—

(*Continued*)

TABLE I. (*Continued*)

Probe ID	Fold change	*P*-value	Gene title	Symbol	Location
208931_s_at	−1.30	0.00297	Interleukin enhancer binding factor 3, 90 kDa	ILF3	19p13.2
214982_at	−1.30	0.00598	U5 snRNP-specific protein, 200-KD	U5-200KD	2q11.2
210232_at	−1.30	0.02568	Cell division cycle 42 (GTP binding protein, 25 kDa)	CDC42	1p36.1
216110_x_at	−1.30	0.04868	*Homo sapiens* cDNA FLJ14080 fis, clone HEMBB1002152	—	2q31.1
222307_at	−1.30	0.02200	Hypothetical protein LOC282997	LOC282997	10q25.3
219209_at	−1.30	0.01517	Melanoma differentiation associated protein-5	MDA5	2p24.3-q24.3
207115_x_at	−1.29	0.02557	mbt domain containing 1	MBTD1	17q21.33
220809_at	−1.28	0.03526	Hypothetical protein FLJ14327	FLJ14327	16q23.2
210095_s_at	−1.28	0.00346	Insulin-like growth factor binding protein 3	IGFBP3	7q13q-12
221728_x_at	−1.28	0.01621	X (inactive)-specific transcript	XIST	Xq13.2
217104_at	−1.28	0.02709	Hypothetical protein LOC283687	LOC283687	15q24.3
211115_x_at	−1.28	0.00056	Survival of motor neuron protein interacting protein 1	SIP1	14q13
209321_s_at	−1.28	0.00881	Adenylate cyclase 3	ADCY3	2p24-p22
209387_s_at	−1.28	0.03185	Transmembrane 4 superfamily member 1	TM4SF1	3q21-q25
209314_s_at	−1.28	0.00716	HBS1-like (*S. cerevisiae*)	HBS1L	6q23-q24
214487_s_at	−1.27	0.01701	RAP2B, member of RAS oncogene family	RAP2B	3q25.2
202861_at	−1.27	0.03941	Period homolog 1 (*Drosophila*)	PER1	17p13.1-p12
218706_s_at	−1.27	0.01122	HCV NS3-transactivated protein 2	NS3TP2	5q23.3
220704_at	−1.27	0.01071	Zinc finger protein, subfamily 1A, 1 (Ikaros)	ZNFN1A1	7p13-p11.1

Shading indicates more than one probe set identified a change in this transcript among those listed in this table; these can be considered independent validations.

In addition, we performed a highly focused analysis of gene expression alterations in five sib-pairs from specific PIC families segregating for bipolar disorder that had linkage to 6q22. Our aims were to determine if this approach would have diagnostic utility and also help identify candidate genes with abnormal expression patterns in the regions that display significant linkage. Importantly, we have not assumed a priori that the changes in transcript levels that occur in PBLs necessarily indicate that similar patterns of expression alterations will be evident in the brains of subjects with schizophrenia or bipolar disorder. Rather, we merely hypothesized that if underlying genetic abnormalities altered transcript expression in a consistent manner, we would be able to detect such effects if the transcript was expressed in PBLs. We suggest that such changes, when present, may reflect primary pathogenetic mechanisms, or simply conserved pathophysiological features of the illness or its treatment.

METHODS

Subject Ascertainment

Methods for subject ascertainment and classification are the same as previously described [Pato et al., 2004]. Families with two or more affected individuals were ascertained from systematic screening of all treating clinicians, treatment facilities, social services, and extensive family interviews. In the Azores, all four psychiatric hospitals and the two general hospitals participated in the study. Similarly in Madeira, both psychiatric hospitals and the general hospital participated. On the mainland, families were identified by our collaborators at the University of Coimbra. Informed consent was obtained in writing from all subjects for participation in the genetic and family studies. Collection of blood and family history information was approved by all of the appropriate Institutional Review Boards. Best estimate diagnoses according to DSM-IV were made by two independent blinded researchers. All cases, where there was disagreement, were reviewed by a third senior psychiatrist blind to the status of the case (MT Pato, MD).

The specific subjects for this study were selected as gender- and age-matched discordant sibs from families that participated in our linkage studies. For studies of schizophrenia, 40 such sib-pairs were selected from all families segregating for schizophrenia from the PIC population where we had detected strong linkage to 5q and had cell samples available for RNA purification. For bipolar disorder, we performed a highly focused analysis of gene expression alterations in five sib-pairs from specific PIC families segregating for bipolar disorder with linkage to 6q22. These 45 age- and gender-matched sib-pairs were screened for alterations in white blood cell composition (see below), which eliminated five schizophrenia sib-pairs from further consideration. After processing all arrays, a priori quality control criteria (excessive 3′/5′ ratios for beta actin and GAPDH, and/or scale factors exceeding 10.0 for any subject) led us to eliminate two additional sib-pairs from our analyses. Thus, in total, 33 sib-pairs were used for our schizophrenia studies and five sib-pairs for our bipolar studies. To achieve the highest level of subject matching and quality control, we found it helpful to examine cellular composition using differential white blood cell counts (Wright's stain method). All of the samples included in this report had normal blood count differentials. The specific age and blood cell composition values for the schizophrenic group are the same as previously described (Petryshen et al., 2004; 22 female pairs, 11 male pairs; mean ages $\pm$ SD of affecteds and unaffecteds $= 44.8 \pm 12.4$ and 42.8 ± 12.9, respectively; mean neutrophil counts $= 58.9 \pm 4.8\%$ and $56.8 \pm 4.7\%$; mean lymphocyte counts $= 34.1 \pm 7.0\%$ and $33.9 \pm 4.2\%$). For the bipolar subject pairs, the mean values $\pm$ SD for affecteds and unaffecteds was age $= 41.2 \pm 5.8$ and 40.8 ± 11.4; neutrophils $= 55.4 \pm 4.4$, 59.2 ± 2.9; lymphocytes $= 38.4 \pm 4.7$, 35.2 ± 2.8. There were no significant pairwise or unpaired differences between the subject groups and the matched controls for any of these values.

Microarray Gene Expression Sample Preparation

Total RNA was extracted from leukocyte cell preparations from 66 siblings selected from the 40 schizophrenia pedigrees used in a previous linkage study [see Sklar et al., 2004] and 10 siblings selected from the 25 bipolar pedigrees used in a previous linkage study [see Middleton et al., 2004]. Total RNA quality and quantity was assessed using UV spectrophotometry and comparison of 28S:18S ratios with the Bioanalyzer RNA Nano Chip (Agilent). Microarray samples were labeled and processed according to standard protocols,

TABLE II. Largest Significant Pairwise Expression Changes in Bipolar Sibling Pairs

Probe ID	Fold change	*P*-value	Gene title	Symbol	Location
Increased expression					
203556_at	2.08	0.03908	Zinc fingers and homeoboxes 2	ZHX2	8q24.13
209332_s_at	1.94	0.01483	MAX protein	MAX	14q23
202351_at	1.88	0.01839	Integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)	ITGAV	2q31-q32
201648_at	1.87	0.01542	Janus kinase 1 (a protein tyrosine kinase)	JAK1	1p32.3-p31.3
211994_at	1.78	0.02617	Protein kinase, lysine deficient 1	PRKWNK1	12p13.3
208638_at	1.76	0.00846	Thioredoxin domain containing 7 (protein disulfide isomerase)	TXNDC7	2p25.1
208856_x_at	1.76	0.03823	Ribosomal protein, large, PO	RPLP0	12q24.2
214737_x_at	1.74	0.02115	Heterogeneous nuclear ribonucleoprotein C (C1/C2)	HNRPC	14q11.2
212626_x_at	1.67	0.03681	Heterogeneous nuclear ribonucleoprotein C (C1/C2)	HNRPC	14q11.2
221491_x_at	1.64	0.01463	Major histocompatibility complex, class II, DR beta 3	HLA-DRB3	6p21.3
209835_x_at	1.64	0.04319	CD44 antigen (homing function and Indian blood group system)	CD44	11p13
200745_s_at	1.61	0.00057	Guanine nucleotide binding protein (G protein), beta polypeptide 1	GNB1	1p36.33
207616_s_at	1.60	0.03284	TRAF family member-associated NFKB activator	TANK	2q24-q31
209331 _s_at	1.59	0.00476	MAX protein	MAX	14q23
200746_s_at	1.56	0.00108	Guanine nucleotide binding protein (G protein), beta polypepide 1	GNB1	1p36.33
209586_s_at	1.55	0.02552	TcD37 homolog	HTCD37	1q21
222294_s_at	1.54	0.03711	RAB27A, member RAS oncogene family	RAB27A	15q15-q21.1
200084_at	1.50	0.03447	Small acidic protein	SMAP	11p15.2
204396_s_at	1.50	0.00192	G protein-coupled receptor kinase 5	GRK5	10q24-qter
204373_s_at	1.48	0.01338	Centrosome-associated protein 350	CAP350	1p36.13-q41
201061_s_at	1.48	0.03523	Stomatin	STOM	9q34.1
218284_at	1.48	0.04873	DKFZP586N0721 protein	DKFZP586N0721	15q22.31
202698_x_at	1.46	0.04149	Cytochrome-*c* oxidase subunit IV isoform 1	COX4I1	16q22-qter
201560_at	1.44	0.04648	Chloride intracellular channel 4	CLIC4	1p36.11
206544_x_at	1.43	0.03723	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a2	SMARCA2	9p22.3
202061_s_at	1.42	0.00889	sel-1 suppressor of lin-12-like (*C. elegans*)	SEL1L	14q24.3-q31
202840_at	1.41	0.01641	TAF15 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 68 kDa	TAF15	17q11.1-q11.2
204735_at	1.40	0.00917	Phosphodiesterase 4A, cAMP-specific (phosphodiesterase E2 dunce homolog, *Drosophila*)	PDE4A	19p13.2
204070_at	1.39	0.01290	Retinoic acid receptor responder (tazarotene induced) 3	RARRES3	11q23
208938_at	1.39	0.00493	Papillary renal cell carcinoma (translocation-associated)	PRCC	1q21.1
201144_s_at	1.38	0.02715	Eukaryotic translation intiation factor 2, subunit 1 alpha, 35 kDa	EIF2S1	14q24.1
205049_s_at	1.38	0.03527	CD79A antigen (immunoglobulin-associated alpha)	CD79A	19q13.2
200605_s_at	1.37	0.01436	Protein kinase, cAMP-dependent regulatory, type 1, alpha (tissue specific extinguisher 1)	PRKAR1A	17q23-q24
200885_at	1.36	0.04139	Ras homolog gene family, member C	RHOC	1p13.1
204489_s_at	1.33	0.01981	CD44 antigen (homing function and Indian blood group system)	CD44	11p13
214271_x_at	1.33	0.02639	Ribosomal protein L12	RPL12	9q34
212014_x_at	1.33	0.02870	CD44 antigen (homing function and Indian blood group system)	CD44	11p13
212352_s_at	1.31	0.04363	Transmembrane trafficking protein	TMP21	14q24.3
214836_x_at	1.31	0.04004	Kappa-immunoglobulin germline pseudogene (Chr22.4) variable region (subgroup V kappa II)	—	2p11.2
221737_at	1.31	0.03167	Guanine nucleotide binding protein (G protein) alpha 12	GNA12	7p22-p21
Decreased expression					
214022_s_at	−1.68	0.01385	Interferon induced transmembrane protein 1 (9–27)	IFITM1	11p15.5
201662_s_at	−1.64	0.03130	Acyl-CoA synthetase long-chain family member 3	ACSL3	2q34-q35
208800_at	−1.60	0.04517	Signal recognition particle 72 kDa	SRP72	4q11
201523_x_at	−1.54	0.01262	Ubiquitin-conjugating enzyme E2N (UBC13 homolog yeast)	UBE2N	12q22
205099_s_at	−1.47	0.02248	Chemokine (C-C motif) receptor 1	CCR1	3p21
219557_s_al	−1.45	0.03653	Nuclear receptor interacting protein 3	NRIP3	11p15.3
218852_at	−1.44	0.04879	Chromosome 14 open reading frame 10	C14orf10	14q13.2
201524_x_at	−1.42	0.03537	Ubiquitin-conjugating enzyme E2N (UBC13 homolog, yeast)	UBE2N	12q22
206689_s_at	−1.41	0.04480	Transcription factor binding to IGHM enhancer 3	TFE3	Xp11.22

(*Continued*)

TABLE II. (*Continued*)

Probe ID	Fold change	*P*-value	Gene title	Symbol	Location
205382_s_at	−1.38	0.03956	D component of complement (adipsin)	DF	19p13.3
212414_s_at	−1.37	0.03062	Septin 6	SEPT6	Xq25
218543_s_at	−1.35	0.04712	Zinc finger CCCH type domain containing 1	ZC3HDC1	7q34
212658_at	−1.34	0.04834	Lipoma HMGIC fusion partner-like 2	LHFPL2	5q14.1
215424_s_at	−1.33	0.00794	SKI interacting protein	SKIIP	14q24.3
217960_s_at	−1.33	0.03531	Translocase of outer mitochondrial membrane 22 homolog (yeast)	TOMM22	22q12-q13
203077_s_at	−1.28	0.04364	SMAD, mothers against DPP homolog 2 (*Drosophila*)	SMAD2	18q21.1
202164_s_at	−1.28	0.03718	CCR4-NOT transcription complex, subunit 8	CNOT8	5q31-q33
218753_at	−1.28	0.03344	Hypothetical protein FLJ10307	FLJ10307	1p35.2
203055_s_at	−1.28	0.02857	Rho guanine nucleotide exchange factor (GEF) 1	ARHGEF1	19q13.13
212397_at	−1.28	0.02452	Radixin	RDX	11q23
221490_at	−1.28	0.04365	Ubiquitin associated protein 1	UBAP1	9p22-p21
212766_s_at	−1.26	0.04957	Hypothetical protein FL12671	FLJ12671	1q23.1
201083_s_at	−1.26	0.02615	BCL2-associated trancription factor	BCLAF1	6q22-q23
209704_at	−1.25	0.04701	Likely ortholog of mouse metal response element binding transcription factor 2	M96	1p22.1
218214_at	−1.25	0.04535	Hypothetical protein FLJ11773	FLJ11773	12q13.13
201978_s_at	−1.25	0.04803	KIAA0141 gene product	KIAA0141	5q31.3
208811_s_at	−1.24	0.02519	DnaJ (Hsp40) homolog, subfamily B, member 6	DNAJB6	7q36.3
218403_at	−1.24	0.03064	Hypothetical protein HSPC132	HSPC132	12q24.31
201856_s_at	−1.21	0.03464	Zinc finger RNA binding protein	ZFR	5p13.3
201712_s_at	−1.21	0.03245	RAN binding protein 2	RANBP2	2q12.3
220349_s_at	−1.21	0.01209	Endo-beta-N-acetylglucosaminidase	FLJ21865	17q25.3
217956_s_at	−1.20	0.00290	E-1 enzyme	MASA	4q21.3
217208_s_at	−1.19	0.04560	Discs, large homolog 1 (*Drosophila*)	DLG1	3q29
202947_s_at	−1.19	0.02718	Glycophorin C (Gerbich blood group)	GYPC	2q14-q21
204777_s_at	−1.18	0.00702	Mal, T-cell differentiation protein	MAL	2cen-q13
212757_s_at	−1.17	0.04972	Calcium/calmodulin-dependent protein kinase (CaM kinase) II gamma	CAMK2G	10q22
213560_at	−1.17	0.04758	Growth arrest and DNA-damage-inducible, beta	GADD45B	19p13.3
218615_s_at	−1.16	0.00929	Hypothetical protein FLJ10902	FLJ10902	3q13.33
201966_at	−1.15	0.04461	NADH dehydrogenase (ubiquinone) Fe-S protein 2, 49 kDa (NADH-coenzyme Q reductase)	NDUFS2	1q23
216333_x_at	−1.15	0.02708	Tenascin XB	TNXB	6p21.3

hybridized to the Human Genome U133A GeneChip© (Affymetrix), washed and stained on a fluidics station (Affymetrix) according to the EukGE-WS2 protocol, and scanned using the G2500A Gene Array Scanner. The scan files were normalized using the Gene Chip Robust Multichip Analysis method [Irizarry et al., 2003].

Statistical changes in gene expression were determined in pairwise fashion using the Significance Analysis algorithm (GeneTraffic, Iobion). All of the transcript probes that were significantly changed ($P < 0.05$) were ranked by mean pairwise fold change. This was applied to both the entire genome for schizophrenia and bipolar disorder (Tables I and II, respectively) and each of the candidate genome regions (5q and 6q) (Table III). Notably, because of the relatively small number of sib-pairs used in the bipolar study, we present the top fold changes in the 6q locus using a threshold of $P < 0.10$ for this disorder.

Prediction Classification

From the lists of significantly changed genes, we also sought to derive a preliminary list of genes whose expression patterns might be useful as predictors of diagnosis. For this analysis, we used the Class Predictor algorithm (GeneSpring) to predict the value, or "class," of individual parameters in the set of samples. This was done using both the Euclidian nearest neighbor method and support vector machine (SVM) method. The genes with the highest predictive power in our dataset were ranked by their predictive strength (Table IV). Importantly, we point out that because we used the methods that require defining a training set, the list we present still requires validation in an independent cohort.

Gene Group Analysis

We determined the relevant biological pathway information in our expression data using custom-written software [PathStat; see Middleton et al., 2004] to extract distributions of differential expression ratios for functionally related groups of transcripts using publicly curated databases. These databases included the different groups of the Gene Ontology database (http://us.expasy.org/cgi-bin/enzyme-search-cl), the Kyoto Encyclopedia of Genes and Genomes database (http://www.genome.ad.jp/kegg/pathway.html), the Enzyme Commission database (http://us.expasy.org/cgi-bin/enzyme-search-cl), and the Protein Family (pFam) database (http://www.sanger.ac.uk/Software/Pfam/search.shtml). Lists of the specific probes on the Affymetrix U133A GeneChip that belong to each of these groupings are available in a single annotated file from the Affymetrix NetAffx website (http://www.affymetrix.com/support/technical/byproduct.affx?product=hgu133-20). To perform the pathway analyses, the scaled and normalized gene expression level was first calculated using robust multi-chip analysis (RMA), and then the expression level of each gene in the affected pairs (BP or SCZ subjects) was compared to the corresponding expression level of the same gene in the matched

TABLE III. Changed Individual Probesets in Loci of Interest

Probe ID	Fold change	P-value	Gene title	Symbol	Location
Schizophrenia 5q linkage region ($P < 0.05$)					
201743_at	2.53	0.00108	CD14 antigen	CD14	5q31.1
201694_s_at	1.74	0.01116	Early growth response 1	EGR1	5q31.1
210844_x_at	1.63	0.00142	Catenin (cadherin-associated protein), alpha 1, 102 kDa	CTNNA 1	5q31
200765_x_at	1.62	0.00056	Catenin (cadherin-associated protein), alpha 1, 102 kDa	CTNNA 1	5q31
203218_at	1.59	0.00311	Mitogen-activated protein kinase 9	MAPK9	5q35
217840_at	1.48	0.01072	DEAD (Asp-Glu-Ala-Asp) box polypeptide 41	DDX41	5q35.3
214658_at	1.46	0.03732	CGI-109 protein	CGI-109	5q23.1
201506_at	1.46	0.02373	Transforming growth factor, beta-induced, 68 kDa	TGFBI	5q31
205896_at	1.44	0.00904	Solute carrier family 22 (organic cation transporter), member 4	SLC22A4	5q31.1
202360_at	1.36	0.00170	Mastermind-like 1 (*Drosophila*)	MAML1	5q35
202227_s_at	1.36	0.00387	Bromodomain containing 8	BRD8	5q31
220495_s_at	1.34	0.01182	Chromosome 5 open reading frame 14	C5orf14	5q31.2
212900_at	1.33	0.03403	SEC24 related gene family, member A (*S. cerevisiae*)	SEC24A	5q31.2
212137_at	1.31	0.01715	Likely ortholog of mouse Ia related protein	LARP	5q33.2
Bipolar 6q linkage region ($P < 0.10$)					
208623_s_at	2.57	0.08125	Villin 2 (ezrin)	VIL2	6q25.2-q26
210105_s_at	2.01	0.08039	FYN oncogene related to SRC, FGR, YES	FYN	6q21
216033_s_at	1.66	0.09324	FYN oncogene related to SRC, FGR, YES	FYN	6q21
212265_at	1.41	0.09018	Quaking homolog, KH domain RNA binding (mouse)	QKI	6q26-27
201915_at	1.17	0.02738	SEC63-like (*S. cerevisiae*)	SEC63	6q21
221311_x_at	1.17	0.09196	Hypothetical protein dJ122O8.2	DJ12208.2	6q14.2-q16.1
210156_s_at	1.16	0.02719	Protein-L-isoaspartate (D-aspartate) O-methyltransferase	PCMT1	6q24-q25
204207_s_at	1.14	0.07257	RNA guanylyltransferase and 5′-phosphatase	RNGTT	6q16
205116_at	1.11	0.05556	Laminin, alpha 2 (merosin, congenital muscular dystrophy)	LAMA2	6q22-q23
215904_at	1.09	0.06463	Myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	MLLT4	6q27
206005_s_at	1.08	0.07908	Chromosome 6 open reading frame 84	C6orf84	6q15
217399_s_at	1.03	0.03835	Forkhead box O3A	FOXO3A	6q21
209608_s_at	−1.19	0.09159	Acetyl-coenzyme A acetyltransferase 2 (acetoacetyl Coenzyme A thiolase)	ACAT2	6q25.3-q26
201083_s_at	−1.26	0.02615	BCL2-associated transcription factor 1	BCLAF1	6q22-q23

control subject. This ratio was generated only if the expression level in at least one of the samples was greater than the median RMA normalized signal intensity of the dataset (computed signal ~12 or higher). PathStat then compiled the gene-by-gene expression ratios for each functional group, as well as all the genes on the chip, and calculated a t statistic for the group that was proportional to the number of standard deviations that group's expression had shifted in the affected subject pair, adjusted for the number of transcripts in each group. Gene group ratios were only generated if an average of at least three transcript ratios were computed in each group [Middleton et al., 2002, 2004]. For each disease, we present the omnibus data (all 5 or 33 subject pairs combined) for this report. This highly conservative method enables one to detect large-scale coordinated changes in functional gene group expression in complex treatment paradigms or disease states. In the present analysis, PathStat mapped our expression data to a total of more than 3,000 separate publicly curated pathways. Tables of the top 25 increased and decreased gene group alterations in each disorder were generated.

Cytogenomics

In addition to the functional gene groups, we also utilized a new approach for mining and mapping expression changes within the context of cytogenetic loci, an approach we term "cytogenomics." For example, single megabase bins generally contain approximately 8–12 gene probes of which approximately half will be expressed in the blood or brain of human subjects. Each of these megabase bins was treated as a gene group cluster and the pairwise and genewise differences in expression used to create high-density cytogenomic Z score plots (Fig. 1). These Z score plots were compared directly with the NPL Z score plot and a Chi Square plot derived from the linkage and association analyses of schizophrenia and bipolar disorder. Two examples of the utility of this approach are provided (Fig. 2). To perform the genome-wide association analysis, we used Varia (Silicon Genetics) to construct a haplotype map (E-M algorithm) of the complete set of 25 bipolar families genotyped with the Affymetrix Human Mapping Assay Xba 141 [see Middleton et al., 2004]. A preliminary genome-wide family-based Transmission Disequilibrium Test (TDT) was performed and the most significant results for the entire genome were listed (Table VIII). Moreover, the Chi Square values from this analysis were overlaid with the expression and linkage plots for chromosome 6q (Fig. 2).

Microarray Validation With Quantitative Real-Time RT-PCR (qRT-PCR)

Validation of selected changes in expression was performed using quantitative real-time RT-PCR (qRT-PCR), using RNA from 19 of the 33 schizophrenia sibpairs and all 5 bipolar sibpairs for whom sufficient RNA was available. Primer sequences for each gene of interest were designed using Primer3 software, and are available upon request. For the RT reactions, equal amounts of total RNA (250 ng) from each sample were reverse transcribed (Superscript II protocol) with an oligo dT primer prior to quantitative PCR according to the

TABLE IV. Genes With Most Power for Distinguishing Affected and Control Subjects

Probe ID	Pred strength	Map	Product	LOI
201253_s_at	34.51	16q13	CDP-diacylglycerol–inositol 3-phosphatidyltransferase (phosphatidylinositol synthase)	
207850_at	34.51	4q21	Chemokine (C-X-C motif) ligand 3	
205098_at	33.18	3p21	Chemokine (C-C motif) receptor 1	
204524_at	30.9	16p13.3	3-Phosphoinositide dependent protein kinase-1	
221732_at	30.9	17q25.3	Ectonucleoside triphosphate diphosphohydrolase 8	
203333_at	30.9	1q22	Kinesin-associated protein 3	
50374_at	30.9	17q25.3	Hypothetical protein LOC339229	
202682_s_at	30.9	3p21.3	Ubiquitin specific protease 4 (proto-oncogene)	
221203_s_at	30.9	3q27.3	Hypothetical protein FLJ10201	
220046_s_at	30.9	3q26.1	Cyclin L1	
213708_s_at	29.97	17q21.1	Transcription factor-like 4	
200765_x_at	29.97	5q31.2	Catenin (cadherin-associated protein), alpha 1, 102 kDa	←
200919_at	29.36	1p34.3	Polyhomeotic-like 2 (*Drosophila*)	
221983_at	29.36	2q36.2	Chromosome 2 open reading frame 17	
212265_at	29.25	6q26-q27	Quaking homolog, KH domain RNA binding (mouse)	←
221355_at	29.25	2q33-q34	Cholinergic receptor, nicotinic, gamma polypeptide	
200745_s_at	29.25	1p36.33	Guanine nucleotide binding protein (G protein), beta polypeptide 1	
204446_s_at	29.25	10q11.2	Arachidonate 5-lipoxygenase	
222028_at	29.25	19q13.2	Zinc finger protein 45 (a Kruppel-associated box (KRAB) domain polypeptide)	
222103_at	29.25	12q13.2	Activating transcription factor 1	
208715_at	29.25	1q22-q25	Putative membrane protein	
218652_s_at	29.25	4p16.3	Hypothetical protein FLJ20265	
205424_at	28.06	17q21.2	ProSAPiP2 protein	
208732_at	28.06	8q11.23	RAB2, member RAS oncogene family	
204208_at	28.06	6q16	RNA guanylyltransferase and 5′-phosphatase	←
204336_s_at	28.06	20q13.3	Regulator of G-protein signaling 19	
211572_s_at	28.06	20p13	Solute carrier family 23 (nucleobase transporters), member 2	
216042_at	28.06	1p36.2	Tumor necrosis factor receptor superfamily, member 25	
201632_at	28.06	12q24.31	Eukaryotic translation initiation factor 2B, subunit 1 alpha, 26 kDa	
211574_s_at	28.06	1q32	Membrane cofactor protein (CD46, trophoblast-lymphocyte cross-reactive antigen)	
212626_x_at	27.61	14q11.1	Heterogeneous nuclear ribonucleoprotein C (C1/C2)	
207943_x_at	27.61	6q24-q25	Pleiomorphic adenoma gene-like 1	←
218195_at	27.61	6q25.1	Chromosome 6 open reading frame 211	←
221204_s_at	11.9	10q22	Cartilage acidic protein 1	
208289_s_at	11.2	11q24	Etoposide induced 2.4 mRNA	
210574_s_at	10.06	1p35-p34	Nuclear distribution gene C homolog (*A. nidulans*)	
217414_x_at	9.77	16p13.3	Hemoglobin, alpha 2	
206649_s_at	9.572	Xp11.22	Transcription factor binding to IGHM enhancer 3	
201083_s_at	9.572	5q23.3	BCL2-associated transcription factor 1	←
202711_at	9.572	Xq12	Ephrin-B1	
222048_at	9.572	22q11.23	Adrenergic, beta, receptor kinase 2	
218753_at	9.572	1p35.3	Hypothetical protein FLJ10307	
209331_s_at	9.572	14q23	MAX protein	
209889_at	9.13	10q24.2	SEC31-like 2 (S. cerevisiae)	
204482_at	9.13	22q11.21	Claudin 5 (transmembrane protein deleted in velocardiofacial syndrome)	
214487_s_at	9.13	3q25.2	RAP2B, member of RAS oncogene family	
201382_at	9.13	1q24-q25	Siah-interacting protein	
209902_at	9.13	3q22-q24	Ataxia telangiectasia and Rad3 related	
217826_s_at	9.13	6q16.1	Ubiquitin-conjugating enzyme E2, J1 (UBC6 homolog, yeast)	←
204070_at	9.13	11q23	Retinoic acid receptor responder (tazarotene induced) 3	

LOI, locus of interest based on this population for schizophrenia or bipolar disorder.

standard TaqMan protocol (Applied Biosystems) using SYBR-Green I dye for amplicon detection with an ABI-7000 Real Time Sequence Detection System (Applied Biosystems). Statistical analysis was performed using a pairwise repeated measures ANOVA comparing the difference in the number of cycles to threshold (ΔCt) between each target gene and alpha tubulin. Group differences were calculated by determining the mean pairwise difference in the delta Ct values per subject group (the $\Delta\Delta$Ct), and a fold change calculated according to the formula, Fold change $= 2^{-\Delta\Delta Ct}$.

RESULTS

Sib-Pair Analyses

Global significance analysis. Of the 22,283 probe sets on the array, in a strict pairwise analysis with the RMA normalized data, more than 2,000 genes showed significant changes in expression in our schizophrenia sib-pairs and 248 genes showed significant changes in expression in our much smaller set of bipolar sib-pair data ($P < 0.05$). Given the larger number of samples used in the schizophrenia analysis,

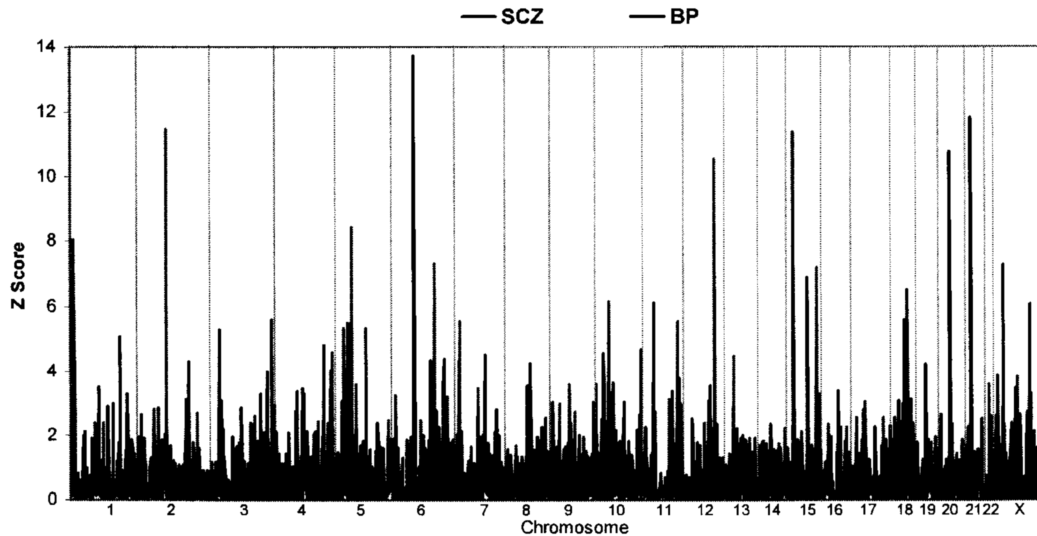

Fig. 1. Cytogenomic plot of expression abnormalities by MB. To generate this map, the pairwise differential expression ratio distribution of all expressed transcripts within 1 MB proximity was determined, and the relative shift in this ratio distribution that compared the entire genome was calculated (as a Z score). For simplification, only the absolute value of the Z score is shown for each disorder. *We note that SCZ and BP data do not display much similarity in their cytogenomic patterns*.

implementing a multiple testing correction (Benjamini-Hochberg) still produced a list of approximately 300 genes with significant changes in expression. However, because such a correction in *P*-values was not possible with the much smaller bipolar dataset, we chose to present purely uncorrected *P*-values in this report.

The 40 genes with the most robust increases and decreases in expression among those that were significantly affected are shown for each disorder (Tables I and II). While the limited sample size does not allow us to make specific comments on each of the genes in these tables, we have decided to present these data in order for other groups that may be working on similar efforts to be able to compare (and potentially combine) their findings with ours. We do point out that several of the genes with the most significant changes in expression did not exhibit large enough alterations to be included in these tables, which are based purely on fold changes.

Targeted significance analysis. We identified 729 transcript probes on the U133A GeneChip that were localized to 5q for use in analyzing our SCZ sib-pairs and 431 transcripts on 6q that were used to analyze expression in our BP sib-pairs. The 14 genes showing the largest fold changes (increases or decreases) at each locus were shown using a *P*-value threshold of 0.05 for schizophrenia and 0.10 for bipolar disorder (Table III).

Prediction classification. We used the Class Predictor Algorithm to identify those genes with the greatest potential diagnostic utility in schizophrenia and bipolar disorder in this population. After optimization, the nearest neighbor algorithm produced a list of 35 transcripts with approximately 95% accuracy. Overall, 70 of 76 subjects were correctly classified as a control, BP or SCZ subject based on sib-pair expression differences, with four subjects misclassified and two subjects not classified. The SVM method obtained 100% accuracy with as few as 10 genes. The list of transcripts in both of these respective classifier lists was nearly identical. For simplicity, we present the top 50 predictor genes, ranked by their predictive strength according to the SVM algorithm (Table IV).

Gene group analysis. We assessed the expression patterns of all publicly curated functional pathways that could be mapped to the U133A GeneChip content. This analysis of more than 3,000 unique transcript collections revealed a set of gene groups with robust changes in expression in each disorder (Table V). Extending this analysis to include groups of genes located at the same physical position (in 1-MB bins), indicated fairly strong expression changes in a number of key loci in each disorder (Table VI). Importantly, we note that the Z scores for the top affected loci in each disorder were not correlated in the BP and SCZ datasets.

Cytogenomic mapping of expression abnormalities. The binning of the data used in the Gene Group Analysis allowed us to create Z score maps of the pairwise expression changes in our datasets, which we plotted for comparative purposes in a manner similar to a whole genome screen (Fig. 1). We point out that this figure revealed considerably different cytogenomic profiles for these two disorders. For illustrative purposes, we have also chosen to display the potential overlap of genetic and functional genomic signals on chromosomes 5q for schizophrenia and 6q for bipolar disorder (Fig. 2). This side by side analysis revealed a fair amount of overlap in the genetic linkage, association, and functional genomic results as was suggested by the data (Table VI).

Family-based association. Analysis of haplotype-based TDT performed by Varia using the 25 bipolar family dataset revealed a number of genomic regions with nominally significant linkage disequilibrium (Table VII). Despite the limits of the resolution of the SNP map used to generate haplotypes ($\sim$210 KB), it is notable that several of the haplotype blocks highlighted are located in proximity to both the 6q linkage peak and the gene expression Z score peaks. We have begun to use higher density SNP genotyping arrays to achieve greater resolution and explore these intriguing findings in more detail.

Validation of selected genes with real-time quantitative RT-PCR. In our PBL sib-pairs, we examined the expression of multiple individual transcripts with an independent technique to confirm the accuracy and reliability of the microarray data. Due to limited RNA, only 19 of the original 33 schizophrenic sib-pairs (but all 5 bipolar sib-pairs) were used in these studies. Several genes with increased and decreased expression, including the Sensory and Motor Neuron Derived Factor (SMDF) variant of Neuregulin 1 (NRG1), Transcription factor-like 4 (TCFL4), Serotonin Receptor Type 4 (5HT4), and A Disintegrin and Matrix Metalloproteinase 19 (ADAM19) were all confirmed as showing the same (or greater) amount of

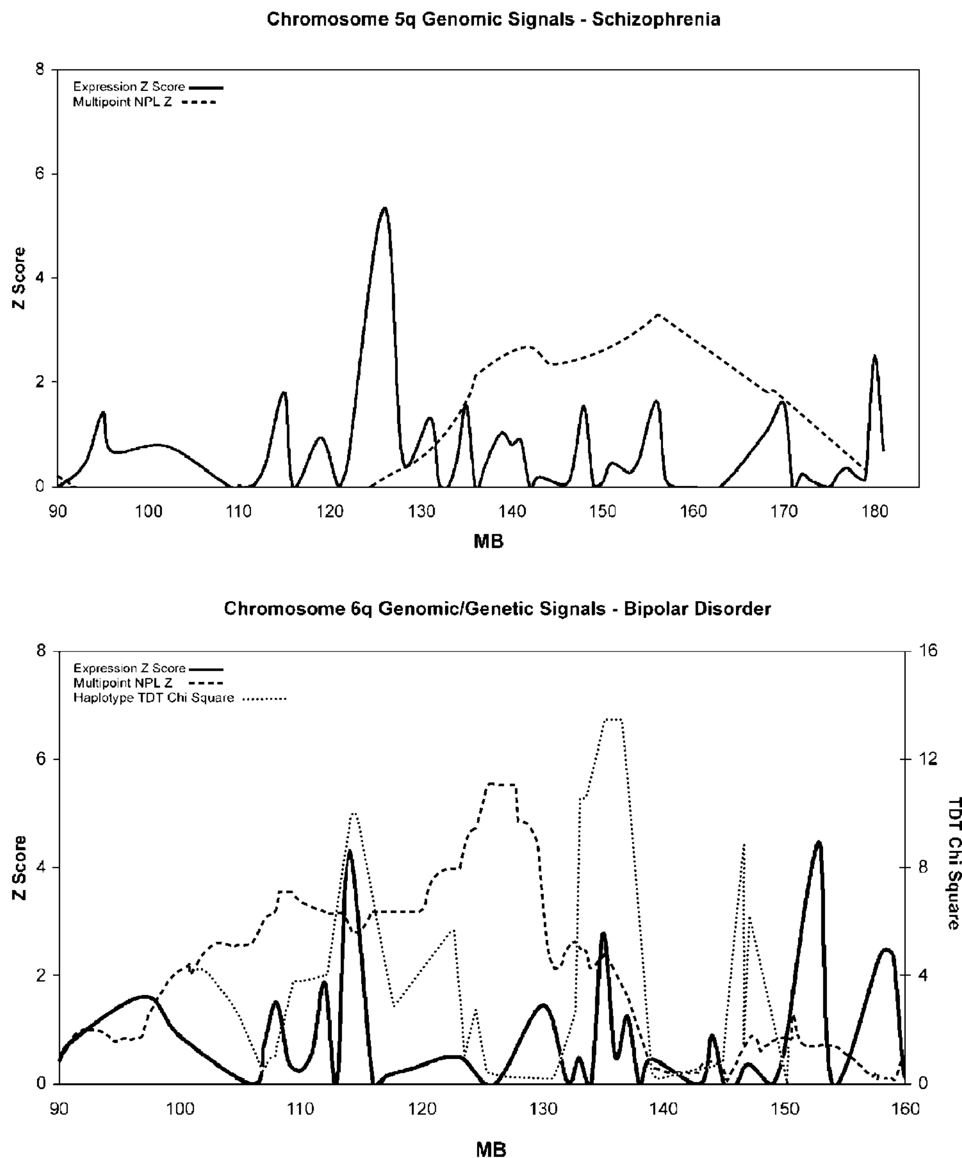

Fig. 2. Regions of putative genetic and functional genomic convergence. **Top**, Nonparametric linkage results in families used in a medium resolution microsatellite scan of chromosome 5q [Sklar et al., 2004]. Data have been replotted in physical (MB) coordinates (dashed line). The pairwise gene expression data from 33 discordant sib-pairs from these families were used to generate an expression Z score plot through the same region (solid line). **Bottom**, Families with 6q22 linkage from the 25 used in a previous genome-wide screen [dashed line; Middleton et al., 2004] were analyzed with a high-density SNP genotyping array. The data from this same dataset were used to create a whole genome haplotype map and test for TDT. The Chi Square values obtained in the TDT analysis for chromosome 6q are illustrated (dotted line). Five age- and gender-matched discordant sib-pairs from these families were used for leukocyte expression analysis in the current study (expression Z score as solid line). We note that the peak expression Z scores in both schizophrenia and bipolar disorder are located near peak linkage regions, although the positions of the TDT peaks show even more similarity with the expression Z score peaks.

change seen in the array data (Table VIII). A fifth transcript (Synaptobrevin 2, or VAMP2) was confirmed as showing the same direction of change as the array data, but the *P*-value did not attain significance (Table VIII). Moreover, other genes that did not change in the array data were also confirmed, including at least two housekeeping genes—alpha tubulin and RNA Polymerase II (Table VIII). Overall, the mean pairwise fold changes for these seven transcripts were highly correlated in the array and PCR datasets ($R = 0.951$).

Genes involved in presynaptic function in the brain are also altered in PBLs. Expression profiling of postmortem brain tissue in subjects with schizophrenia has revealed some commonly affected intracellular cascades, particularly those involving presynaptic secretory function [PSYN genes; see Mirnics et al., 2000]. In order to begin to address whether PBLs also share some of the same expression changes as postmortem brain, we searched for significant ($P < 0.05$) pairwise expression changes in the PSYN genes in our schizophrenia data. Interestingly, some 35 unique transcripts were found with such changes. Most of these involved small differences (less than 20% mean pairwise changes). However, among the transcripts with 20% or greater increases in expression were synaptophysin-like protein (SYPL, +1.4-fold), synaptosomal-associated protein, 23 kDa (SNAP23, +1.4-fold), N-ethylmaleimide-sensitive factor (NSF, +1.4-fold), dynactin 3 (DCTN3, +1.4-fold), sorting nexin 2 (SNX2,

TABLE V. Top Functional Gene Group Alterations in Schizophrenia and Bipolar Disorder

Category	Gene group	Ave N	Fold change	t	Category	Gene group	Ave N	Fold change	t
Bipolar disorder					Schizophrenia				
Increased expression					Increased expression				
GO-MOL	Structural protein of ribosome	78.4	1.40	10.12	GO-BIO	Eye morphogenesis	3.0	2.02	2.80
GO-BIO	Protein biosynthesis	103.6	1.29	9.02	GO-MOL	Glycosaminoglycan binding	4.0	1.61	2.50
GO-MOL	RNA binding	209.6	1.16	8.17	GO-MOL	Hyaluronic acid binding	4.1	1.54	2.18
GO-CELL	Cytosolic large ribosomal (60S)-subunit	45.2	1.35	7.42	PFAM	Thrombospondin type 3 repeat	3.0	1.90	2.18
GO-CELL	Cytosolic small ribosomal (40S)-subunit	28.6	1.41	5.64	GO-MOL	Polypeptide N-acetylgalactosaminyltransferase	5.0	1.26	2.05
GO-CELL	Hemoglobin	8.2	2.26	4.32	PFAM	Glyceraldehyde 3-phosphate dehydrogenase	3.0	1.26	2.04
PFAM	Ribosomal L10	8.4	1.44	3.44	GO-BIO	Beta-tubulin folding	4.5	1.30	2.04
GO-CELL	Ribosome	10.2	1.33	3.33	GO-BIO	Induction of apoptosis by p53	3.0	1.18	2.03
GO-CELL	Microvilli	4.2	1.99	3.30	PFAM	Pyridine nucleotide-disulfide	4.5	1.30	2.02
GO-BIO	Eye morphogenesis	3.0	1.49	3.27	PFAM	CDP-alcohol phosphatidyltransferase	3.0	1.27	2.00
GO-BIO	Golgi to secretory vesicle transport	4.0	1.60	3.09	EC	2.1.1.77 Protein-L-isoaspartate O-methyltransferase	3.0	1.23	1.95
PFAM	Ribosomal_L3	5.0	1.53	3.04	GO-MOL	Poly-pyrimidine tract binding	4.0	1.26	1.91
PFAM	MHC_I	33.6	1.12	3.00	PFAM	Vacuolar protein sorting	3.0	1.20	1.88
GO-BIO	mRNA processing	58.8	1.11	2.91	GO-MOL	Proteoglycan	5.5	1.37	1.83
GO-BIO	Olfaction	5.0	1.45	2.89	KEGG	O-Glycans biosynthesis	6.0	1.22	1.80
PFAM	Ezrin or radixin or moesin	9.2	1.39	2.88	GO-MOL	GPI-anchored membrane-bound receptor	3.9	1.48	1.76
PFAM	N-Myristoyltransferase, C-terminal domain	3.2	1.15	2.88	GO-MOL	Antibacterial response protein	7.1	1.44	1.75
GO-MOL	Retinoic acid receptor	6.0	1.09	2.86	GO-MOL	Sterol carrier	3.0	1.32	1.75
PFAM	tRNA synthetases class I 4	4.0	1.35	2.86	GO-BIO	Cell recognition	6.5	1.34	1.73
PFAM	Proteasome	23.6	1.14	2.84	GO-MOL	KDEL receptor	3.0	1.23	1.73
GO-MOL	Double-stranded DNA binding	17.2	1.16	2.83	PFAM	Cysteine-rich repeat	3.0	1.20	1.72
PFAM	Ku70 or Ku80 beta-barrel/C-terminal	3.0	1.33	2.80	GO-CELL	DNA directed RNA polymerase III	3.3	1.19	1.71
PFAM	PAS domain	18.0	1.12	2.76	EC	EC 3.5.1 Hydrolases acting on carbon-nitrogen bonds	6.1	1.34	1.62
GO-BIO	protein kinase C activation	5.0	1.31	2.74	GO-BIO	Protein methylation	4.2	1.18	1.57
PFAM	HECT-domain (ubiquitin-transferase)	20.2	1.14	2.65	PFAM	RasGAP C-terminus	4.0	1.24	1.56
Decreased expression					Decreased expression				
PFAM	BTB or POZ domain	51.2	−1.09	−3.61	EC	4.6.1.1 Adenylate cyclase	3.0	−1.21	−3.71
GO-CELL	Extracellular space	123.6	−1.05	−3.12	PFAM	CCAAT-binding transcription factor subunit B	3.2	−1.12	−2.46
GO-MOL	Sterol transporter	4.0	−1.53	−3.08	GO-MOL	Fibroblast growth factor receptor	3.3	−1.15	−2.25
PFAM	Bcl-2	11.2	−1.23	−3.02	PFAM	UDP-glucoronosyl and UDP-glucosyl transferase	3.7	−1.15	−2.17
PFAM	SNF7	7.0	−1.19	−2.98	PFAM	7 transmembrane receptor (rhodopsin family)	3.5	−1.15	−2.14
EC	6.2.1.3 Long-chain-fatty-acid−CoA ligase	4.0	−1.45	−2.96	GO-MOL	Transforming growth factor beta receptor ligand	4.2	−1.13	−2.12
PFAM	Hemopexin	12.4	−1.07	−2.91	PFAM	MOCO sulfurase C-terminal domain	3.0	−1.09	−2.09
PFAM	Matrixin	13.2	−1.07	−2.88	PFAM	Fibrillar collagen C-terminal domain	6.6	−1.13	−2.05
GO-BIO	Cholesterol metabolism	8.2	−1.21	−2.83	GO-MOL	Leukemia inhibitory factor receptor ligand	3.0	−1.08	−2.05
GO-CELL	Spindle	7.8	−1.21	−2.77	PFAM	Connexin	5.0	−1.09	−2.03
GO-MOL	Protein kinase inhibitor	10.0	−1.08	−2.75	PFAM	Receptor family ligand binding region	8.7	−1.14	−2.03
PFAM	S-100 or ICaBP type calcium binding protein	11.6	−1.27	−2.75	GO-BIO	Blood pressure regulation	4.4	−1.11	−2.00
GO-MOL	Blood group antigen	3.8	−1.11	−2.72	GO-BIO	Cartilage condensation	3.5	−1.14	−1.96
GO-MOL	Defense or immunity protein	16.2	−1.12	−2.72	PFAM	Ligand-gated ion channel	3.2	−1.14	−1.93
GO-MOL	Anion transporter	3.2	−1.49	−2.70	PFAM	MyTH4 domain	3.3	−1.10	−1.92
GO-BIO	O-linked glycosylation	8.2	−1.15	−2.70	GO-MOL	Hormone	3.2	−1.11	−1.91
PFAM	SNF	3.4	−1.12	−2.68	PFAM	Astacin (peptidase family M12A)	3.9	−1.11	−1.89
PFAM	AMP-binding	11.6	−1.16	−2.68	PFAM	Metallothionein	6.0	−1.31	−1.89
PFAM	7 transmembrane receptor (rhodopsin family)	6.2	−1.11	−2.67	PFAM	MAM domain	3.9	−1.11	−1.88
PFAM	pfkB family carbohydrate	3.8	−1.28	−2.67	GO-BIO	Phototransduction	5.2	−1.13	−1.87
PFAM	Matrix metalloprotease	11.2	−1.06	−2.64	GO-MOL	Calcium channel	6.2	−1.17	−1.86
GO-BIO	Iron homeostasis	10.6	−1.19	−2.59	PFAM	Apolipoprotein A1/A4/E family	3.3	−1.07	−1.86
EC	3.4.21 Serine endopeptidases	16.0	−1.07	−2.59	PFAM	Zn-finger in Ran binding 1	3.0	−1.23	−1.85
PFAM	Major intrinsic protein	8.4	−1.09	−2.56	KEGG	Glutamate metabolism	3.3	−1.10	−1.83
GO-BIO	Phagocytosis, engulfment	5.2	−1.29	−2.50	EC	5.3.99.2 Prostaglandin D synthase	3.0	−1.04	−1.79

TABLE VI. Top Loci Expression Alterations in Schizophenia and Bipolar Disorder

Bipolar disorder					Schizophrenia				
		Z score					Z score		
Chr	MB	BP	SCZ	NPL >2 ±15 MB	Chr	MB	BP	SCZ	NPL >2 ±15 MB
6	52	13.72	0.08		2	72	1.48	11.48	
21	31	11.84	1.05		15	37	2.31	11.38	
20	23	10.78	0.03	Y	12	102	0.64	10.58	
X	29	7.29	0.77		12	103	0.07	8.58	
15	68	6.87	2.39		5	73	0.74	8.45	
11	37	6.13	2.65	Y	1	7	1.19	8.08	
18	45	5.59	0.62		6	130	1.45	7.32	Y
3	187	5.59	0.60		15	86	0.80	7.19	
10	36	5.07	6.15		18	53	3.84	6.48	
1	182	5.05	0.09		10	36	5.07	6.15	
4	160	4.83	0.00		X	128	0.75	6.07	
10	134	4.70	1.77		7	28	0.98	5.56	
7	104	4.49	1.42		11	123	4.04	5.55	
6	153	4.37	1.18		5	62	1.64	5.53	
6	114	4.32	2.53	Y	5	126	0.25	5.34	Y
2	162	4.28	0.70	Y	5	44	0.74	5.33	
8	74	4.24	1.46		3	24	2.31	5.30	
19	39	4.20	0.17	Y	5	0	3.09	4.61	
11	123	4.04	5.55		10	27	0.75	4.55	Y
3	174	3.98	0.15		13	28	0.00	4.46	
18	53	3.84	6.48		5	68	1.73	4.13	
11	129	3.77	1.35		4	186	2.02	4.03	
6	152	3.68	2.82		X	14	0.51	3.85	
8	67	3.56	0.38	Y	X	71	0.22	3.82	
1	86	3.53	0.95		4	157	2.64	3.69	

Loci within 15 megabases (15 MB) of an NPL Z score >2.0 in this population are indicated for each disorder.

+1.3-fold), vacuolar protein sorting 33B (VPS33B, +1.3-fold), tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta (YWHAB, +1.2-fold), and vesicle-associated membrane protein 4 (VAMP4, +1.2-fold). On the other hand, far fewer PSYN transcripts showed pairwise decreases in expression of this magnitude, with the notable exceptions of clathrin, heavy polypeptide (CLTC, -1.2-fold) and synaptojanin 2 (SYNJ2, −1.2-fold). These findings of increased expression of NSF, synaptophysin, and other related genes are actually opposite the findings reported in previous microarray studies of postmortem brain tissue. Taken together, these observations reinforce the potential power of PBL expression data to detect significant expression alterations of some of the same genes implicated by postmortem brain tissue studies, although the direction of change is not necessarily the same.

DISCUSSION

In this very preliminary study, we have explored the potential utility of gene expression data acquired from PBLs of subjects with schizophrenia or bipolar disorder and their discordant age- and gender-matched siblings. Specifically, we sought to: (1) identify those genes and functional gene groups that are the most affected in the disease (pathophysiogenomics); (2) evaluate candidate genes and gene regions for abnormal expression patterns that are correlated with known locations of genetic linkage or association in the same population; (3) identify genes with the greatest degree of diagnostic utility; (4) search for candidate genomic loci that might be involved in the primary disease pathogenesis; and (5) attempt to correlate specific SNP haplotypes with altered expression patterns. We briefly highlight some of our progress in each of these aims.

Most Significantly Changed Genes in Schizophrenia and Bipolar Disorder

It is not possible for us to review all of the single gene findings in our dataset within the context of this preliminary report. Thus, we have chosen to highlight a few of the observations that we believe have particular novelty or relate to the preexisting literature on schizophrenia, bipolar disorder, or brain function.

Schizophrenia. There were a number of genes among those with the most significant increases in expression that are known to be involved in immune and/or inflammatory function (e.g., CD14 antigen, chemokine receptor 1; Table I). Interestingly, however, we also detected a number of expression changes in genes that are known to be involved in brain development (e.g., alpha catenin, neuregulin 1 (SMDF variant)). The single most significantly affected gene was TCFL4, which was 1.42-fold decreased, $P = 0.0003$), but did not achieve a magnitude large enough to be listed in Table I. This gene is also known as MAX-like protein (or MLX). We also highlight the increase in expression of transducer of ERBB2, 2 (TOB2, 1.65-fold increased, $P = 0.00014$) (Table I). The change in TOB2 is noteworthy because this gene is located in close proximity to a schizophrenia susceptibility locus (22q13) and because of the alteration in expression we observed for neuregulin 1, SMDF (sensory and motor neuron derived factor) variant. TOB2 is a transducer of the tyrosine kinase receptors from the ErbB family, which also exert control over NRG signaling. The role of NRG is critical to early development of the central nervous system and includes stimulation of Schwann cell growth as well as generation of acetylcholine receptors at the neuromuscular synapse [reviewed in Falls, 2003]. Of the seven splice variants of NRG on the array, only

TABLE VII. Location of TDT Peaks Relative to Expression Peaks in Bipolar Disorder

	Haplotype TDT						Expression
SNP Ids	Chi square	*P*-value	dof	Chr	Alleles	MB	Z (MB)
rs1392096	13.42	0.0198	5	1	C/T	205.3	3.31 (203)
rs1166868					C/T	205.5	
rs4130547					A/G	205.8	
rs1450344	10.43	0.0153	3	3	A/G	151.5	3.30 (153)
rs1388622					A/G	152.4	
rs2210798	13.29	0.0013	2	6	C/T	8.5	1.93 (6)
rs1591454					C/G	8.5	
rs2225765					A/G	8.5	
rs2225766					A/G	8.5	
rs2327577	13.33	0.0098	4	6	G/T	135.1	2.78 (135)
rs728030					G/T	135.2	
rs2327578					C/T	135.3	
rs724875					A/G	136.7	
rs720565					C/G	136.7	
rs721123	9.23	0.0264	3	7	G/T	146.3	2.80 (136)
rs721124					C/T	146.3	
rs1406288					C/T	146.4	
rs719311					A/T	147.0	
rs719312					C/G	147.0	
rs967306	11.33	0.0452	5	9	A/G	11.6	3.02 (2)
rs958842					G/T	11.7	
rs1816823					C/T	11.7	
rs2009463					C/G	11.7	
rs956094					C/T	11.7	
rs317155	11.33	0.0452	5	11	C/T	89.3	—
rs1986670					A/G	89.5	
rs1986671					C/G	89.5	
rs2167050	9.31	0.0023	1	11	A/C	132.1	3.77 (129)
rs921268					A/G	132.1	
rs1512981	9.04	0.0287	3	12	A/G	71.1	—
rs1398562					G/T	71.2	
rs2118087	9.83	0.0200	3	13	A/G	57.5	—
rs719193					G/T	57.6	
rs2183493					A/G	58.0	
rs2183492	9.83	0.0200	3	13	C/T	58.0	—
rs1330754					C/T	58.0	
rs3901894	9.03	0.0289	3	13	C/G	95.2	—
rs544080					A/C	95.5	
rs2899589	12.00	0.0025	2	15	A/G	53.4	1.62 (46)
rs725150					C/T	53.4	
rs1382859	10.81	0.0045	2	15	C/G	90.4	2.16 (97)
rs1382860					A/G	90.4	
rs717788	12.50	0.0285	5	17	C/T	11.7	—
rs1519252					A/G	12.5	
rs1401539					G/T	12.5	
rs952785	11.63	0.0088	3	18	C/T	57.3	3.84 (53)
rs582970					C/T	57.9	

All expression peaks noted exceeded a Z score of 1.6 and were within 10 MB of the TDT peak.

the SMDF variant has altered gene expression in schizophrenic subjects versus controls. A full report on these findings was recently reported elsewhere [Petryshen et al., 2004]. Together, our data support further examination of the role of NRG signaling in schizophrenia.

Bipolar disorder. Among the genes with the most consistent and significant changes in expression was one not previously reported in bipolar disorder, termed MAX (1.94-fold increase, $P = 0.015$; Table II). The finding regarding increased MAX expression in bipolar disorder is particularly interesting in light of the increased expression of TCFL4/MLX in schizophrenia. The MAX gene encodes for a member of the basic region-helix-loop-helix-zipper proteins [Gilladoga et al., 1992]. The MAX protein has been shown to associate with N-, L-, and c-myc proteins [Gilladoga et al., 1992] and other proteins and transcription factors, such as TCFL4. It has been found that TCFL4 works in conjunction with ChREBP (carbohydrate response element-binding protein) to regulate the expression of genes responsive to glucose [Stoeckman et al., 2004]. Interestingly, we also observed a number of transcripts involved in G protein signaling to exhibit altered expression in bipolar disorder (e.g., guanine nucleotide binding protein (G protein), beta polypeptide 1; G protein-coupled receptor kinase 5; Table II). Collectively, these data suggest there are prominent changes in cellular transduction mechanisms in this illness.

Functional Gene Group Alterations

Through the use of custom-written software, we were able to analyze the expression patterns of entire groups of genes that perform the same function in cells. The basis and utility of

TABLE VIII. Representative Real-Time Quantitative RT-PCR Validation Assays

Disease	Gene	Location	Array Finding (n = 33 pairs)			Real-time (n = 19 pairs)		
			Fold Chg	*P*-value	# Probes	Fold Chg	*P*-value	# Repl
BP	RNA polymerase II, A	17p13.1	−1.09	0.497	1	−1.02	0.485	2
SCZ	Alpha tubulin, ubiquitous	12q13.12	1.04	0.640	5	1.02	0.170	4
SCZ	ADAM 19	5q33.3	−1.11	0.180	2	−1.28	0.026	3
SCZ	Serotonin receptor 4	5q33.1	−1.14	0.053	3	−1.31	0.018	3
SCZ	Synaptobrevin 2 (VAMP2)	17p13.1	−1.08	0.030	1	−1.23	0.074	3
SCZ	TCFL4	17q21	1.27	0.010	3	1.18	0.048	2
SCZ	Neuregulin 1, SMDF variant	8p21-p12	1.72	0.012	1	3.80	0.014	3

Correlation in reported differences ($\log_2$ scale) for seven transcripts. Array versus PCR: R = 0.951.
P-value superscripts indicate # tails.

this approach has been reviewed previously [Middleton et al., 2002, 2004]. In the present dataset, we have obtained evidence that several major biological pathways related to lipid and fatty acid metabolism are decreased in bipolar disorder (e.g., sterol transporter, long-chain-fatty-acid–CoA ligase, and cholesterol metabolism; Table VI). In the schizophrenia dataset, these same gene groups were not changed, although the group apolipoprotein A1/A4/E family was similarly decreased (Table VI). Interestingly, among the most increased gene groups in bipolar disorder was the group containing transcripts of the Ezrin family. Ezrin interacts with actin to stabilize uptake process (including the uptake of cholesterol), and has been shown to be important in neurite outgrowth during cortical formation. Ezrin itself is not apparently expressed in neurons, but is abundant within radial glia and migrating cells in the intermediate zone [Johnson et al., 2002]. Together, these data suggest that subjects with schizophrenia and bipolar disorder both exhibit alterations of molecules involved in fatty acid and lipid metabolism that are vital to normal brain function, although the specific genes showing the greatest changes are distinct for each disorder.

Classification by Gene Expression

We tested the ability of our dataset to correctly distinguish subjects with schizophrenia or bipolar disorder from their unaffected discordant sibs using the nearest neighbor class predictor algorithm (GeneSpring). An iterative approach was used to determine the number of genes and gene neighbors that produced the highest accuracy. The use of 35 potential candidates and 7 nearest neighbors successfully classified 70 of 76 subjects in accordance with their true diagnosis (i.e. unaffected, BP or SCZ).

Of the six misclassified or unclassified subjects, three SCZ patients were classified as BP, one SCZ was misclassified as unaffected, and two subjects were not classified. Thus, in its initial training set, the algorithm had an accuracy of classification of 87% for schizophrenia (27 of 31 correct diagnoses) and 89% for the combined SCZ and BP subject sets (32 of 36 correct diagnoses). We wish to stress that the small number of BP subjects used in this study likely contributed to a less distinct predictor gene set being obtained for this disease. Interestingly, only one patient was misclassified as an unaffected subject. Moreover, two of the subjects with schizophrenia that were classified as probable bipolar subjects actually had family histories of bipolar disorder and/or carried a Best Estimate Diagnosis of schizoaffective disorder. By convention, schizoaffective disorder, depressed has been considered affected in schizophrenia linkage studies and schizoaffective disorder, bipolar has been considered affected in bipolar linkage studies. This convention may promote confounds that would limit the ability of gene expression profiling. For example, if we eliminate those schizoaffective subjects that were classified as schizophrenics in our linkage studies, three of the four classification errors would be removed. This would increase the diagnostic accuracy to 96% (27 of 28 correct). Although preliminary, this observation highlights the need for careful evaluation of the nature of schizoaffective disorder and the use of these patients as part of the core phenotype definition. In addition to the nearest neighbor method, we also obtained 100% accurate classification using a SVM model with our gene expression data. Most of the genes with the highest predictive strength were the same for these two methods, and are listed in Table IV. Notably, several of these predictor genes are located on the loci of interest (e.g., alpha catenin) or have been previously mentioned (e.g., TCFL4).

Cytogenomic Mapping of Expression Data Compared to Linkage/LD Data

Our method of analyzing the expression of groups of genes in close physical proximity (cytogenomics) has provided clear examples where the abnormalities converge with the loci identified through linkage and/or association screening. Despite the clear limitations of this preliminary study, our results support the potential utility of this multifaceted approach to the study of neuropsychiatric illness. Based on this approach, however, we have already extended these results to specifically test candidate genes within the 5q linkage region in schizophrenia that displayed abnormal expression patterns and found significant relationships to haplotype-linkage disequilibrium in the same region [Petryshen et al., 2004].

CONCLUSIONS

Overall, we find that screening of gene expression patterns in PBLs of subjects with schizophrenia or bipolar disorder shows potential both in terms of diagnostic utility as well as revealing new biological insights into these disorders. Much of the data we are accumulating point toward alterations of specific loci and specific biological pathways in each illness. We recognize the need for follow-up studies of our results. Nonetheless, the data we present provide a framework for our ongoing research and allow other groups to test specific hypotheses in their datasets.

ACKNOWLEDGMENTS

We are very grateful to the families and individuals who participated in these studies. We also thank Xin Zhao for assistance in software development, and Celia Carvalho, Ana Dourado, Isabel Coelho, M.J. Soares, Jose Valente, and Carlos Paz Ferreira for assistance in the clinical ascertainment

of subjects. Support for this work was derived, in part, from a VA Merit Award (to M.T. Pato) and NIMH grants MH52618 and MH058693 (to C.N. Pato and M.T. Pato).

REFERENCES

Falls DL. 2003. Neuregulins: Functions, forms, and signaling strategies. Exp Cell Res 284(1):14–30.

Gilladoga AD, Edelhoff S, Blackwood EM, Eisenman RN, Disteche CM. 1992. Mapping of MAX to human chromosome 14 and mouse chromosome 12 by in situ hybridization. Oncogene 7(6):1249–1251.

Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. 2003. Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31(4):e15.

Johnson MW, Miyata H, Vinters HV. 2002. Ezrin and moesin expression within the developing human cerebrum and tuberous sclerosis-associated cortical tubers. Acta Neuropathol (Berl) 104:188–196.

Middleton FA, Mirnics K, Pierri JN, Lewis DA, Levitt P. 2002. Gene expression profiling reveals alterations of specific metabolic pathways in schizophrenia. J Neurosci 22(7):2718–2729.

Middleton FA, Ramos EJ, Xu Y, Diab H, Zhao X, Das UN, Meguid M. 2004a. Application of genomic technologies: DNA microarrays and metabolic profiling of obesity in the hypothalamus and in subcutaneous fat. Nutrition 20:14–25.

Middleton FA, Pato MT, Gentile KL, Morley CP, Zhao X, Eisener AF, et al. 2004b. Genomewide linkage analysis of bipolar disorder by use of a high-density single-nucleotide—Polymorphism (SNP)genotyping assay: A comparison with microsatellite marker assays and finding of significant linkage to chromosome 6q22. Am J Hum Genet 74:886–897.

Mirnics K, Middleton FA, Marquez AM, Lewis DA, Levitt P. 2000. Molecular characterization of schizophrenia revealed by microarray analysis of gene expression in prefrontal cortex. Neuron 28:53–67.

Pato CN, Pato MT, Kirby A, Petryshen TL, Medeiros H, Carvalho C, Macedo A, Dourado A, Coelho I, Valente J, Soares MJ, Ferreira CP, Lei M, Verner A, Hudson TJ, Morley CP, Kennedy JL, Azevedo MH, Daly MJ, Sklar P. 2004. Genome-wide scan in Portuguese Island families implicates multiple loci in bipolar disorder: Fine mapping adds support on chromosomes 6 and 11. Am J Med Genet 127B(1):30–34.

Petryshen TL, Middleton FA, Kirby A, Aldinger KA, Purcell S, Tahl AR, Morley CP, McGann L, Gentile KL, Rockwell GN, Medeiros HM, Carvalho C, Macedo A, Dourado A, Valente J, Ferreira CP, Patterson NJ, Azevedo MH, Daly MJ, Pato CN, Pato MT, Sklar P. 2005. Support for involvement of neuregulin 1 in schizophrenia pathophysiology. Mol Psychiatry 10:328.

Sklar P, Pato MT, Kirby A, Petryshen TL, Medeiros H, Carvalho C, Macedo A, Dourado A, Coelho I, Valente J, Soares MJ, Ferreira CP, Lei M, Verner A, Hudson TJ, Morley CP, Kennedy JL, Azevedo MH, Lander E, Daly MJ, Pato CN. 2004. Genome-wide scan in Portuguese Island families identifies 5q31-5q35 as a susceptibility locus for schizophrenia and psychosis. Mol Psychiatry 9:213–218.

Stoeckman AK, Ma L, Towle HC. 2004. Mlx is the functional heteromeric partner of the carbohydrate response element-binding protein in glucose regulation of lipogenic enzyme genes. J Biol Chem 279:15662–15669.

Vawter MP, Ferran E, Galke B, Cooper K, Bunney WE, Byerley W. 2004. Microarray screening of lymphocyte gene expression differences in a multiplex schizophrenia pedigree. Schizophr Res 67:41–52.

Wagner AJ, LeBeau MM, Diaz MO, Hay N. 1992. Expression, regulation, and chromosaomal localization of the Max gene. Proc Natl Acad Sci 89(7):3111–3115.

Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, Behm FG, Raimondi SC, Relling MV, Patel A, Cheng C, Campana D, Wilkins D, Zhou X, Li J, Liu H, Pui CH, Evans WE, Naeve C, Wong L, Downing JR. 2002. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell 1:133–143.