American Journal of Medical Genetics Part B (Neuropsychiatric Genetics) 141B:653–662 (2006)

Rapid Publication

PhenoChipping of Psychotic Disorders:
A Novel Approach for Deconstructing and Quantitating Psychiatric Phenotypes

Alexander B. Niculescu,[1]* Len L. Lulow,[2] Corey A. Ogden,[2] Helen Le-Niculescu,[1] Daniel R. Salomon,[3] Nicholas J. Schork,[2] Michael P. Caligiuri,[2] and James B. Lohr[2]**

[1]*Laboratory of Neurophenomics, Institute of Psychiatric Research, Indiana University School of Medicine, Indianapolis, Indiana*
[2]*Department of Psychiatry, UC San Diego, La Jolla, California*
[3]*Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, California*

Psychiatric phenotypes as currently defined are primarily the result of clinical consensus criteria rather than empirical research. We propose, and present initial proof of principle for, a novel approach to characterizing psychiatric phenotypes. We have termed our approach PhenoChipping, by analogy with, and borrowing paradigms and tools from, gene expression microarray studies (GeneChipping). A massive parallel profiling of cognitive and affective state is done with a PhenoChip composed of a battery of existing and new quantitative psychiatric rating scales, as well as hand neuromotor measures. We present preliminary data from 104 subjects, 72 with psychotic disorders (bipolar disorder—41, schizophrenia—17, schizoaffective disorder—14), and 32 normal controls. Microarray data analysis software and visualization tools were used to investigate: 1. relationships between phenotypic items ("phenes"), including with objective motor measures, and 2. relationships between subjects. Our analyses revealed phenotypic overlap among, as well as phenotypic heterogeneity within, the three major psychotic disorders studied. This approach may be useful in helping us move beyond current diagnostic classifications, and suggests a combinatorial building-block (Lego-like) structure underlies psychiatric syndromes. The adaptation of microarray informatic tools for phenotypic analysis readily facilitates direct integration with gene expression profiling of lymphocytes in the same individuals, a strategy for molecular biomarker identification. Empirically derived clusterings of (endo)phenotypes and of patients will better serve genetic, pharmacological, and imaging research, as well as clinical practice.

KEY WORDS: psychiatric profiling; endophenotypes; bipolar; schizophrenia; microarrays; biomarkers

Please cite this article as follows: Niculescu AB, Lulow LL, Ogden CA, Le-Niculescu H, Salomon DR, Schork NJ, Caligiuri M, Lohr JB. 2006. PhenoChipping of Psychotic Disorders: A Novel Approach for Deconstructing and Quantitating Psychiatric Phenotypes. Am J Med Genet Part B 141B:653–662.

INTRODUCTION

Psychiatric phenotypes are currently characterized by consensus criteria derived primarily from clinical experience, as embodied in DSM-IV TR [American Psychiatric Association, Task Force on DSM-IV, 2000] and ICD-10 [Janca et al., 1996]. While this is an advance over the fairly recent past, when categorization relied on somewhat vague clinical descriptions, there is significant room for improvement. The current criteria have, as a strength, provided a common language for psychiatrists across different sites and different countries, with good inter-rater reproducibility [Spitzer and Williams, 1994; Spitzer, 2001; Spitzer and First, 2005]. The major weaknesses are that they are categorical rather than dimensional, and not empirically derived on a consistent basis. As such, they may not entirely and accurately reflect the phenomenological reality, or have a direct correspondence with the underlying biology. There is a need for more quantitative, empirical approaches to psychiatric phenotyping, for both research and clinical purposes. The broad nature of current psychiatric phenotypical constructs is a rate-limiting step for precise and reproducible genetic research, clinical trials, and clinical practice.

The clinical overlap of phenotypes associated with major psychotic disorders such as bipolar disorder, schizoaffective disorder, and schizophrenia, on the one hand, along with the complexity of these psychiatric disorders on the other hand, points to overlapping (shared) mechanisms between disease classes, as well as heterogeneous mechanisms within a disease class. Besides overlap in clinical symptomatology and genetic studies [Berrettini, 2000; Craddock et al., 2006], another body of evidence that supports the existence of shared mechanisms is that various pharmacological treatments are often successful in relieving symptoms across disorders. Conversely, certain pharmacological treatments may only be efficacious for a subgroup of people within a disease class, consistent with the existence of heterogeneity within these disorders [Harrison and Weinberger, 2005; Tamminga and Holcomb, 2005].

Kraepelin and Bleuler, the forefathers of modern psychiatry, had anticipated some of the issues we are facing in terms of overlap among, and heterogeneity within, psychotic disorders.

*Correspondence to: Alexander B. Niculescu, Laboratory of Neurophenomics, Institute of Psychiatric Research, Indiana University School of Medicine, Indianapolis, IN, 46202. E-mail: anicules@iupui.edu

**Correspondence to: James B. Lohr, Department of Psychiatry, UC San Diego, La Jolla, CA 92093. E-mail: jlohr@ucsd.edu

Received 21 May 2006; Accepted 23 June 2006

DOI 10.1002/ajmg.b.30404

Nevertheless, it is expected that the next version of the Diagnostic and Statistical Manual for Mental Disorders (DSM), DSM-V, will be an incremental improvement along current categorizations and criteria [Kupfer, 2005], rather than a complete overhaul to include more insights derived from genetics, imaging, and neurophysiological studies. Concerted attempts to be more empirical about the phenomenology of psychotic disorders, particularly bipolar disorders, have been undertaken in the past [Craddock et al., 2004], albeit not yet as part of a concerted integration with genetic work. Personality and temperament measures developed by Cloninger and Akiskal have pointed to dimensional aspects of psychopathology and the existence of a continuum between normality and psychopathology [Maremmani et al., 2005]. There is also the growing realization that endophenotypes [Gottesman and Gould, 2003] are shared in a modular fashion among various psychiatric disorders [Niculescu and Akiskal, 2001; Lenox et al., 2002; Hyman and Fenton, 2003], and an increasing appreciation of the need for approaches that would better define the phenotypic structure of psychiatric disorders [Kelsoe and Niculescu, 2002; Hyman, 2003; Hasler et al., 2004; Krishnan, 2005].

Classifying psychiatric phenotypes based on empirical data analysis may help clarify and quantify the issues of overlap and heterogeneity, and thus place the field on a more biologically relevant footing. Admittedly, this is a large undertaking, with multiple caveats. Nevertheless, if new subtypes can be reliably identified from empirical data analysis of patients profiled on a variety of phenotypic and genetic measures, that may be a starting point towards unraveling their different neurobiological etiologies.

Gene expression profiling with microarrays (GeneChipping) is an empirical, discovery-based approach that has generated new insights in multiple fields, as well as new methodological paradigms. A microarray generally consists of thousands of nucleic acid probes attached to a glass slide. Labeled messenger RNAs, the product of gene transcription (gene expression) from a tissue that is being interrogated, are hybridized with the microarray, and the type and numbers of transcripts that stick to the chip are quantified using a specialized scanner. The readout from the scanner gives a quantitative profile of gene expression in the tissue sample analyzed. A widely used pattern recognition method in microarray analysis is unsupervised hierarchical clustering, in which the similarity between genes determined by expression profiles across multiple conditions is measured. This approach has led to notable successes in cancer biology in terms of improved classification of tumor types, subtypes, and staging, compared to classic histopathological methodologies [Bittner et al., 2000; Lu et al., 2005]. Early attempts at using hierarchical clustering in psychiatry [Mezzich, 1978] have not been consistently pursued over the ensuing years.

We reasoned that a similar conceptualization and approach to the one used in cancer biology could be used in psychiatry for empirically studying phenotypes in a massively parallel, quantitative, fashion ("phene" expression). Continuing the analogy with cancer, phene expression may provide advantages compared to classical psychopathologic approaches, similar to those gene expression has provided for tumor classifications compared to classical histopathologic approaches. We have termed the approach PhenoChipping, by analogy to GeneChipping.

It is of interest to have as part of a comprehensive phenotypic profiling (PhenoChipping) approach both subjective measures (quantitative answers to questions about mood, anxiety, cognition) and objective measures (neurophysiology, imaging, gene expression, biochemical assays). New correlations and biomarkers may be revealed by data mining of integrated datasets. Objective phenotypic measurements frequently used include neurophysiology (EEG, neuromotor measures) and brain imaging (fMRI, PET). Hand neuromotor measures, in particular, are easy to administer and deploy, which makes them attractive for large-scale field studies. They engage fundamental fronto-striatal circuits regulating limbic and neuromotor behavior, which may have been recruited also for higher mental functions by evolution. Correlations between motor measures and clinical parameters have been reported in both bipolar disorders and schizophrenia, including in never medicated schizophrenia; moreover, looking at right hand versus left-hand measures may provide a window into brain hemispheric lateralization of pathology [Cortese et al., 2005]. The relationship between cognitive impairment and motor abnormalities remains an important area for further research. Moreover quantitative hand neuromotor measures have been shown, in affective disorders, to be predictive of antidepressant non-response [Caligiuri et al., 2003].

The PhenoChip used in this report consists of a battery of existing psychiatric rating scales (for psychosis, well-being and mood) and one newly developed affective scale, together with right and left hand neuromotor measures, all quantitative in nature. We were particularly interested how responses to questionnaires that reflect an internal subjective experience might correlate with objective neuromotor measures.

Affective abnormalities are an integral part of major psychotic disorders, yet they are often overlooked and not tested for in patients with psychosis, as opposed to patients with mood disorders. We have developed a simple-minded, quantitative, visual analog scale to assess affective state (Total Affective State Scale—TASS), based on combining and placing on a continuum the DSM-IV criteria for depression, mania, and anxiety. This scale has already demonstrated its usefulness in imaging studies of bipolar disorder, as recently reported by us [Caligiuri et al., 2006]. In our current report, we were particularly interested to see if we could detect with it novel differences in individual measures of affect in different psychotic disorders diagnostic categories. From a pragmatic standpoint, we reasoned that we have a higher likelihood of uncovering new phenomenology in an area that has been less explored (mood in psychosis). More generally, we wanted to look at the interdependence of cognition and mood.

With that in mind, we placed on our PhenoChip each of the 11 individual items in TASS, along with the Total Mood subscale, Total Anxiety subscale, and overall Total Affect composite scale (for a total of 14 probes), and placed on it only the composite scales for the 11 other rating instruments (including 4 neuromotor measures, 2 for each hand). Thus, our prototype PhenoChip had 25 probes in total. While being comprehensive, it is purposefully biased towards the finer grained detection of affective phenomenology. Moreover, while we chose to have on the chip the current standard rating scales for psychosis and mood (such as PANSS, HAM-D, YMRS—see below), a variety of other rating scales could have been chosen or added. Thus, we are very cognizant that our results, as with custom microarrays in the gene expression field, may reflect not only the limited number of subjects we sampled, but also the particular choice of probes included on the PhenoChip.

MATERIALS AND METHODS

Demographics and Subject Enrollment

We have collected a sample of 104 subjects, consisting of 41 subjects with bipolar disorder, 17 with schizophrenia, 14 with schizoaffective disorder, and 32 without significant psychiatric illness (normal controls), determined by the Structured Clinical Interview for the DSM-IV Axis I Disorders, Clinician Version (SCID-I).

Subjects consisted of men and women over 18 years of age. A demographic breakdown is shown in Table I. Subjects were

TABLE I. Demographic Data

	Controls	Bipolar	Schizophrenia	Schizoaffective
Number of subjects	32	41	17	14
Gender				
Male: female	24: 8	24: 17	14: 3	8: 6
Age				
Mean years (SD) range	48.4 (8.5) 32–64	44.3 (10.1) 21–65	47.3 (7.6) 27–59	38.6 (7.5) 29–49
Illness duration				
Mean years (SD) range		18.3 (11.1) 1–47	22.8 (11.3) 3–39	14.9 (8.7) 5–34

recruited from the general population, the patient population at the Veterans Affairs San Diego Healthcare System, and the University of California at San Diego, as well as various facilities that serve people with mental illnesses in San Diego County. The subjects were recruited largely through referrals from care providers, through the use of brochures left in plain sight in public places and mental health clinics, and through word of mouth. Subjects were excluded if they had significant medical or neurological illness or had evidence of active substance abuse or dependence. All subjects understood and signed informed consent forms before assessments began.

Administration of the PhenoChip

Subjects completed diagnostic assessments (SCID), and then were PhenoChipped. The PhenoChip we used consisted of a battery of: (1) existing psychiatric rating scales: Positive and Negative Symptoms Scale (PANSS) (with a positive symptom subscale—PANSSPOS, a negative symptom subscale—PANSSNEG, and a disorganization subscale—PANSSGEN) [Kay et al., 1987], Hamilton Rating Scale for Depression (HAM-D 17 and HAM-D 28) [Hamilton, 1960, 1980], Young Mania Rating Scale (YMRS) [Young et al., 1978], Medical Outcomes Study Short Form-36 (SF-36) [Ware et al., 1996]; (2) a new visual-analog scale: TASS [Caligiuri et al., 2006], as well as (3) hand neuromotor measures: VS-velocity scaling, FI-force instability [Caligiuri et al., 1998].

The battery was administered in one of three predetermined counterbalanced orders. Subjects were paid for their participation. Testers were not blind to the subject's diagnosis, but were not aware of the study hypotheses or the approach that would be used for empirical data analysis.

Visual Analog Scale—Total Affective State Scale (TASS)

The newly developed visual analog scale, the TASS, quantifies mood and anxiety symptoms at the time of administration (Fig. 1). It has a mood subscale and an anxiety subscale. The seven-item mood subscale (Simplified Mood State Subscale—SMS) is based on: (a) combining the DSM-IV criteria for depression and mania, and (b) placing the items on a continuum. The four-item anxiety subscale (Simplified Anxiety State Subscale—SAS), quantifies feelings of uncertainty, fear, and anger. The advantages of TASS, and the reasons for using it, are that: (1) it quantifies state, (2) it measures phenotypes on a continuum, from normal to pathology, (3) it is self-rated, which facilitates administration and ease of use.

We have enrolled more bipolar patients than schizophrenia and schizoaffective patients (41 vs. 17 vs. 14) (Table I) specifically to have a larger sample of patients with known affective symptomatology for the purpose of validating our scale. Besides the face validity of using DSM-IV items for its creation, TASS has internal consistency demonstrated by a high degree of correlation between items, as well as external consistency, demonstrated by the high degree of inverse correlation with HAM-D28, a scale measuring depression (Fig. 1c). Moreover, for the purposes of our studies, we are more interested in the scores of the individual items in TASS, in a modular endophenotypic fashion or as probes on the PhenoChip, rather than how TASS fits as a classic diagnostic measurement scale.

Data Analysis

To determine which phenes had significantly different scores between each disease group and normal controls, a student's *t*-test for independent samples was used. This analysis was performed using Statistica (version 6.1). The average value of the raw scores for each phene is used in our *t*-test calculation. A *P*-value <0.05 is considered significant (Fig. 1b and Table II). If, however, we apply a conservative Bonferroni correction for multiple testing, as there are 25 probes on our PhenoChip and three diagnostic groups, the threshold for significance would change to $P < 0.00066$.

To analyze the relationships between phenes, a standardization of the data is necessary because of the varying dynamic ranges in which the various psychiatric rating scales measures and neurophysiological hand motor functions are quantified. For example, the HAM-D28 has a score range of 0–82, while the TASS has a range of 0–1,100. The Cohen's d effect size [Cohen, 1988] was used as our method of standardizing scores for the diagnostic groups, in which Cohen's d effect size $= M_1 - M_2/\sigma_{pooled}$, where M_1 is the average score of the disease group for the phene of interest, and M_2 is the average score of the control group for that same phene. σ_{pooled} is the standard deviation of all of the scores that went into calculating both M_1 and M_2.

To keep the calculations consistent, we used a modified Z score (an individual "effect size") to calculate the scores for individual subjects, in which Z score $= X_1 - M_{2/}\sigma_{pooled}$, where X_1 is the individual score for the phene of interest and M_2 the average score of the control group for that same phene. σ_{pooled} is the standard deviation of all of the scores that went into calculating both M_1 and M_2.

Clustering Analysis Using GeneSpring

We have adapted GeneSpring (Agilent Technologies, Palo Alto, CA) the most widely used, commercially available, microarray gene expression analysis software, for the novel use of analyzing and visualizing phenotypic data. We have inputted the scores on phenotypic items numbers in lieu of the usual use of gene expression intensity numbers. All the subsequent analyses were carried out using the same tools as for gene expression datasets, per the manufacturer's instructions (www.chem.agilent.com). A "genome" (phenome) was created in the program, consisting of the 25 items on the PhenoChip—each item acting as an individual "gene" (phene). The Z scores for each phene in all samples were imported into GeneSpring. No further normalization was applied to the data

a Example of TASS item- Thinking Activity

How high is the amount of mental energy and thinking activity going on in your mind right now? Compare to the most slowed down you ever remember your thinking being, and compared to the most alert and fast you ever remember your thinking being.

[----------]

lowest **highest**

0 100 mm

b Results of measurements using TASS

	Distance in Spectrum (mm)							
	Normal Controls (N=32)		Bipolar (N=41)		Schizophrenia (N=17)		Schizoaffective (N=14)	
Total Affective State Scale (TASS)	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Simplified Mood Scale (SMS)								
Mood	69.25	16	57.68	16.21	54.82	20.18	54.5	29.35
Motivation to do things (Motivdo)	70.16	16.39	60.20	21.72	55.53	22.36	56.43	28.70
Movement activity (Mvmtactv)	65.19	20.16	55.12	21.35	54.53	18.33	55	28.99
Thinking activity (Thnkactv)	69.13	16.35	60.68	18.24	54	15.92	57.71	24.74
Self-esteem (Selfestm)	69.34	18.33	53.32	25.09	55.24	26.59	53.93	31.95
Interest in pleasurable activities (Interest)	69.88	17.07	57.78	22.78	54.24	15.98	48.14	33.08
Appetite	70.06	19.18	52.71	20.21	53.35	24.34	51.64	32.61
Total Mood (TotMood)	483	92.25	397.49	103.90	381.71	98.82	377.5	166.68
Simplified Anxiety State Scale (SAS)								
Anxiety	42.59	22.53	41.27	21.52	46.06	23.81	44.71	25.68
Uncertainty (Uncertnt)	39.78	21.49	55.24	24.54	53.35	22.37	52.07	31.02
Fear	25.25	19.56	39.39	24.12	34.18	22.66	32.29	24.11
Anger	21.47	20.02	32.05	23.34	23.41	17.92	22.79	26.43
Total Anxiety (TotAnxty)	129.1	65.81	167.95	67.91	157	58.56	151.86	77.95
Total Affect (Totaffect)	612.1	91.54	565.44	96.99	538.71	82.75	529.36	173.11

c Total Affective State Scale – correlation with HAM-D28

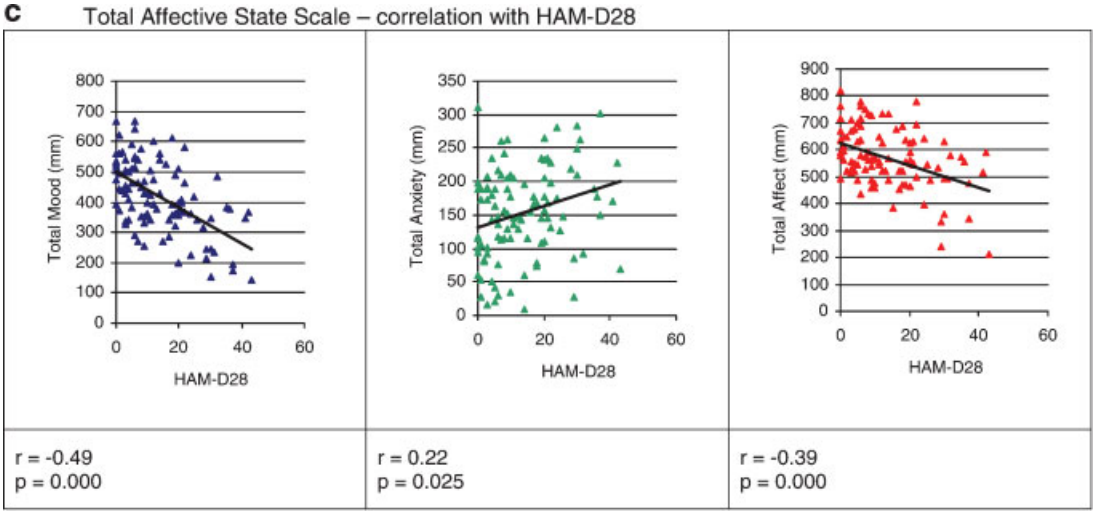

Fig. 1. TASS. **a**: Example of TASS item - Thinking Activity. **b**: Results of measurements using TASS. **c**: TASS-correlation with HAM-D28.

TABLE II. Psychotic Disorders Compared to Normal Controls

	Bipolar subjects (n = 41)			Schizophrenic subjects (n = 17)			Schizoaffective subjects (n = 14)		
	Student's *t*-test			Student's *t*-test			Student's *t*-test		
Phenes	Effect size	*t*-value (df = 71)	*P*-value	Effect size	t-value (df = 47)	*P*-value	Effect size	t-value (df = 44)	*P*-value
Motor measures									
FIL	0.159	0.671	0.505	0.440	1.484	0.144	***0.579***	1.854	0.0704
FIR	0.296	1.261	0.211	0.391	1.314	0.195	0.449	1.418	0.163
LVS	−0.487	−2.111	*0.0383*	***−0.528***	−1.799	*0.0785*	−0.132	−0.409	0.685
RVS	−0.426	−1.833	0.0709	***−0.606***	−2.088	*0.0423*	−0.360	−1.128	0.266
SFGEN									
SF-36	***−0.768***	−3.501	*0.000805*	***−0.899***	−3.286	*0.00193*	***−1.227***	−4.611	*0.000034*
Simplified Mood Scale (SMS)									
Mood	***−0.680***	−3.043	*0.00329*	***−0.772***	−2.741	*0.00863*	***−0.670***	−2.175	*0.0350*
Motivdo	−0.496	−2.158	*0.0343*	***−0.741***	−2.615	*0.0119*	***−0.638***	−2.060	*0.0453*
Mvmtactv	−0.473	−2.048	*0.0443*	***−0.532***	−1.816	0.0758	−0.436	−1.375	0.176
Thnkactv	−0.474	−2.053	*0.0438*	***−0.859***	−3.110	*0.00317*	***−0.578***	−1.853	0.0705
Selfestm	***−0.678***	−3.035	*0.00336*	***−0.632***	−2.187	*0.0338*	***−0.641***	−2.074	*0.0440*
Interest	***−0.570***	−2.503	*0.0146*	***−0.861***	−3.120	*0.00309*	***−0.874***	−2.950	*0.00507*
Appetite	***−0.809***	−3.723	*0.000392*	***−0.748***	−2.641	*0.0112*	***−0.732***	−2.401	*0.0207*
TotMood	***−0.798***	−3.663	*0.000477*	***−0.960***	−3.570	*0.000835*	***−0.826***	−2.763	*0.00834*
Simplified Anxiety Scale (SAS)									
Anxiety	−0.061	−0.256	0.799	0.152	0.503	0.618	0.091	0.282	0.780
Uncertnt	***0.635***	2.819	**0.00624**	***0.602***	2.075	**0.0435**	0.490	1.553	0.128
Fear	***0.610***	2.695	**0.00879**	0.427	1.439	0.157	0.335	1.045	0.302
Anger	0.472	2.043	**0.0448**	0.101	0.335	0.739	0.0602	0.186	0.853
TotAnxty	***0.561***	2.459	**0.0164**	0.435	1.466	0.149	0.327	1.020	0.313
TOTAFFECT									
SMS + SAS	−0.482	−2.090	*0.0402*	***−0.776***	−2.758	*0.00825*	***−0.656***	−2.125	*0.0392*
PANSS items									
PANSSPOS	***0.834***	3.865	**0.000243**	***1.220***	4.968	**0.000009**	***1.429***	5.906	**0.000000**
PANSSNEG	***0.342***	1.460	0.149	***1.236***	5.068	**0.000007**	***0.945***	3.247	**0.00224**
PANSSGEN	***1.392***	8.152	**0.000000**	***1.377***	6.058	**0.000000**	***1.530***	6.719	**0.000000**
Depression scales									
HAM-D17	***1.383***	8.054	**0.000000**	***1.209***	4.901	**0.000012**	***1.511***	6.554	**0.000000**
HAM-D28	***1.416***	8.435	**0.000000**	***1.233***	5.046	**0.000007**	***1.661***	8.080	**0.000000**
Mania rating scale									
YMRS	***0.850***	3.952	**0.000181**	0.467	1.582	0.120	0.453	1.431	0.160

The effect sizes and the independent *t*-test *P*-values for each phene in a comparison between disease groups and the normal controls are shown. Numbers in bold text represent phenes that are significantly increased compared to normal controls and numbers in italic text represent phenes that are significantly decreased compared to normal controls (Student's *t*-test, $P \leq 0.05$). All values that have a Cohen's d effect size greater than 0.50 are bold-italic.

inside GeneSpring. Two-way hierarchical clustering analysis was applied to the Z scores to investigate relationships between samples and relationships between phenes. Standard correlation is used as the similarity metric. Hierarchical clustering was performed in two ways: clustering by the average scores (effect sizes) of each diagnostic group (three samples-bipolar (BAD), schizophrenia (SZ), and schizoaffective (SZA)) (Fig. 3), and clustering across the individual scores (Z scores) of all subjects (104 samples) (Fig. 4).

RESULTS

Phenes that were significantly different between each disease group and normal controls are shown in Table II, both with effect size data and *t*-test data. An effect size of greater than 0.50 is considered medium to high, and significant. We identified 22 phenes in bipolar subjects (11 increased and 11 decreased), 16 phenes in schizophrenic subjects (10 increased and 6 decreased), and 13 phenes in schizoaffective subjects (8 increased and 5 decreased) that were significantly changed ($P < 0.05$).

Venn Diagram Analysis

Venn diagrams based on the differentially changed phenes in bipolar disorder, schizophrenia, and schizoaffective disorder, compared with controls, are shown in Figure 2. Figure 2a represents the phenes that were significantly increased, and Figure 2b represents the phenes that were significantly decreased. Several of the differentially expressed phenes were shared between the three psychotic disorders. These shared phenes included four that were increased (PANSSPOS, PANSSGEN, HAM-D17, HAM-D28) and eight that were decreased (SF-36, Mood, Motivdo, Selfestem, Interest, Appetite, Totmood, Totaffect). These results speak to the fact that the three major psychotic disorders share phenotypic characteristics. Interestingly, bipolar disorder had six uniquely changed phenes: Fear, Anger, Totanxty, and YMRS were increased; LVS and Mvmtactv were decreased.

We divided the phenes into three categories, from less specific to more specific. Category I phenes are changed in all three psychotic disorders in our sample, compared to normal controls. Category II phenes are changed in two out of the three psychotic disorders, compared to normal controls. Category III phenes are just changed in one disorder, compared to controls.

The Category I phenes increased in all three psychotic disorder groups are: PANSSPOS, PANSSGEN, HAM-D17, and HAM-D28. They have to do with positive symptoms psychosis, disorganization, and depression. The Category I phenes, decreased in all three psychotic disorders groups, are SF-36, Mood, Motivdo, Selfestm, Interest, Appetite, TotMood,

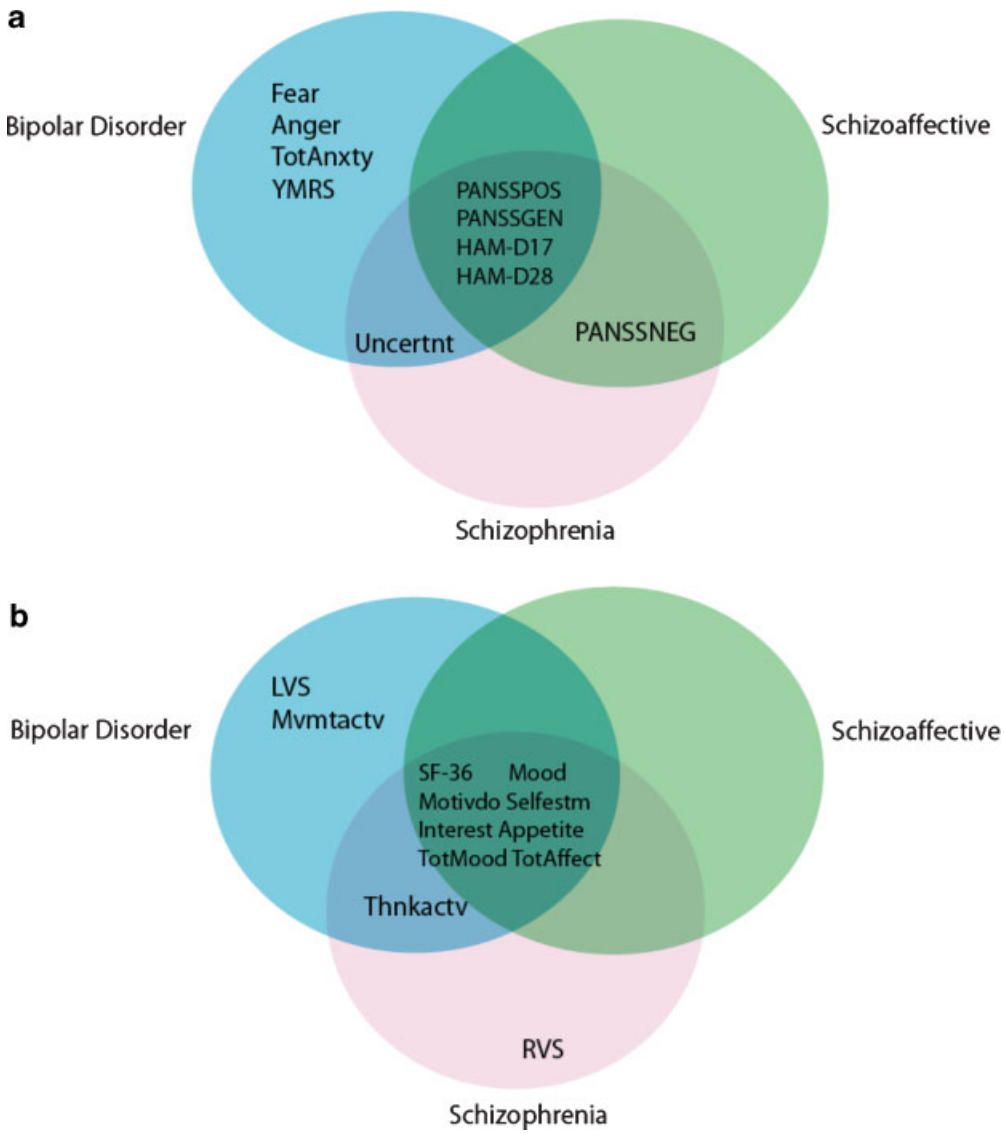

Fig. 2. Venn diagrams of the differentially changed phenes in bipolar disorder, schizophrenia, and schizoaffective, compared with controls. A *t*-test was used to determine significance (*P*-value <0.05). **a**: Representation of the phenes that were significantly increased compared to normal controls. **b**: Representation of the phenes that were significantly decreased compared to normal controls.

TotAffect. They have to do with well being and mood. These results suggest that the three groups of patients, at the time of PhenoChipping, were overall in a more depressed, psychotic, low well being state compared to normal controls. Furthermore, the results suggest that the areas of endophenotypic and neurobiological overlap common to all three psychotic disorders have to do with both cognition and mood.

The Category II phene increased in common in bipolar disorder and schizophrenia is Uncertnt (Uncertainty), and decreased in common in these two disorders is Thnkactv (Thinking Activity). These results suggest that these two groups of patients, at the time of PhenoChipping, were overall in a state characterized by slow thinking, perhaps in part as a paralyzing consequence of high uncertainty. Furthermore, they suggest that an area of endophenotypic and neurobiological overlap between bipolar disorder and schizophrenia has to do with thinking activity and decision-making [Minassian et al., 2004]. The Category II phene increased in common in schizophrenia and schizoaffective disorder is PANSSNEG. This result suggests that these two groups of patients, at the time of the PhenoChipping, were experiencing more negative symptoms than normal controls, and that negative symptoms may be a core endophenotypic and neurobiological feature of schizophrenia spectrum disorders- or a medication side-effect of typical antipsychotics, which are used preponderantly in these two groups of psychotic disorders, compared to bipolar disorder.

The Category III phenes increased only in bipolar disorder patients were Fear, Anger, TotAnxty, MRS. They have to do with anxiety, irritability, and activation. The Category III phenes decreased only in bipolar disorder patients were Left velocity scaling (LVS) and Movement Activity (Mvmtactv). They have to do with right hemisphere activity, and overall energy to move. These results suggest that the bipolar patients, at the time of the PhenoChipping, were in an irritable, psychomotorly retarded state, having to do preferentially with their right hemisphere. Furthermore, they suggest that an area of endophenotypic and neurobiological specificity for bipolar disorders compared to schizophrenia spectrum disorders has to do with anxiety and irritability. Last

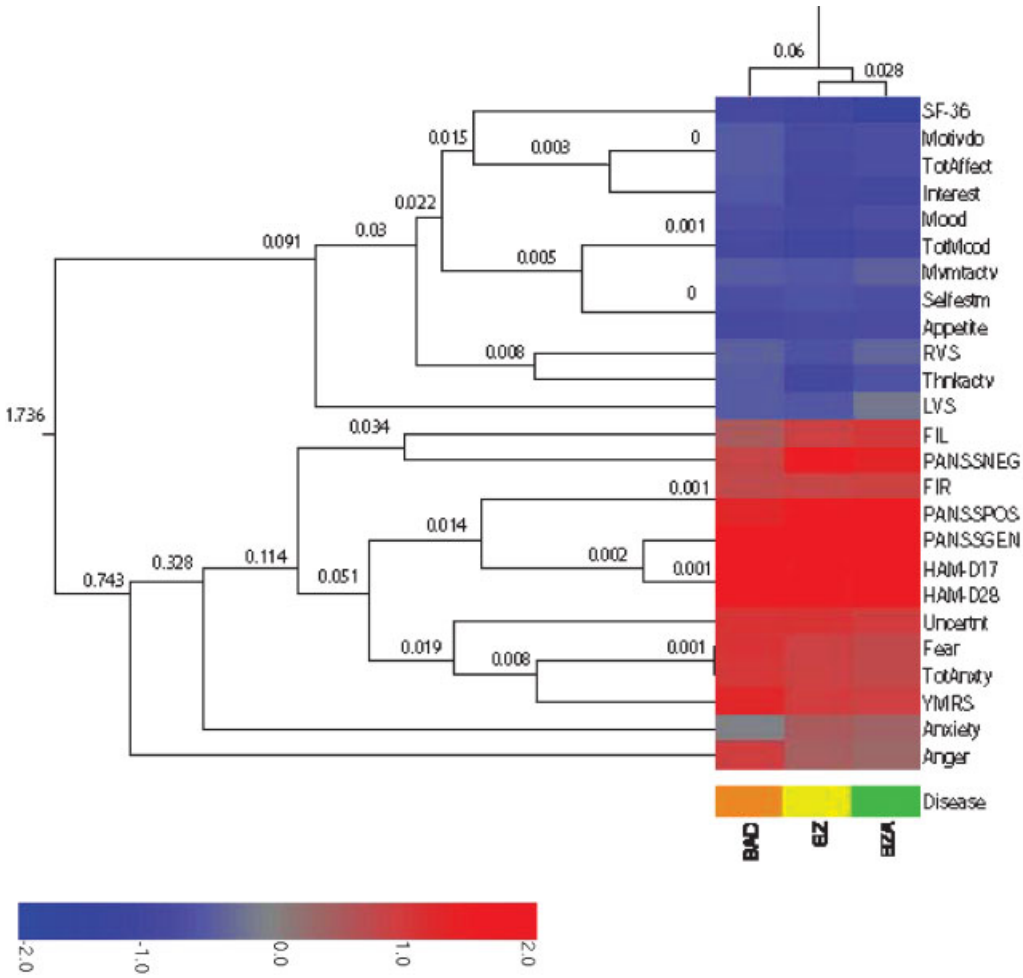

Fig. 3. Clustering of phenes: *overlap across psychotic disorders*. Two-way hierarchical clustering of the disease groups and 25 phenes based on the Cohen's d effect size values for each phene. All effect sizes were calculated comparing the individual disease groups with the normal controls. Each row represents a phene, while each column represents a disease group. Red and blue indicate effect sizes (expression levels) respectively above and below zero, according to the color scale shown at the bottom. Values that are shown on the dendrogram, represent the branch distance, which was determined by the standard correlation similarity measure feature in GeneSpring. Disease groups are listed as BAD for bipolar affective disorder; SZ for schizophrenia; and SZA for schizoaffective.

but not least, an objective neuromotor measure, LVS, having to do with right hemisphere activity, could potentially be used as a behavioral biomarker for bipolarity and to monitor treatment response. This is consistent with recent results in the field using fMRI [Caligiuri et al., 2004].

The Category III phene decreased only in schizophrenia patients is Right hand Velocity Scaling (RVS). It has to do with left hemisphere activity. This result suggests that the schizophrenic patients, at the time of the PhenoChipping, were in a psychomotorly retarded state, having to do preferentially with their left hemisphere. Furthermore, they suggest that an area of endophenotypic and neurobiological specificity for schizophrenia, compared to psychotic disorders with a major affective component, has to do with left hemisphere function. Last but not least, an objective neuromotor measure, RVS, having to do with left hemisphere activity, could potentially be used as a behavioral biomarker for schizophrenia and to monitor treatment response.

Clustering of Phenes

Two-way unsupervised hierarchical clustering of the three diagnostic groups was first applied, based on the average effect size for all phenes across the three groups. Results are displayed in a color-coded "heat map" (Fig. 3), where diagnostic groups are ordered on the horizontal axis and phenes on the vertical axis on the basis of similarity of their effect sizes. Of interest, expression patterns are fairly similar across the three diagnostic groups, with schizophrenia and schizoaffective more similar to each other than to bipolar disorder.

The phenes grouped into two main clusters: phenes that increased in expression compared to normal controls (FIL, PANSSNEG, FIL, PANSSPOS, PANSSGEN, HAM-D17, HAM-D28, Uncertnt, Fear, TotAnxty, YMRS, Anxiety, Anger) and phenes that decreased in expression compared to normal controls (SF-36, Motivdo, TotAffect, Interest, Mood, TotMood, Mvmtactv, Selfestm, Appetite, RVS, Thnkactv, LVS). Notably, all of the well-being and mood measures, with the exception of YMRS, were found to be decreased across all three disorders. However, HAMD, Fear, and Anger were increased. Taken together, this suggests that at the time of PhenoChipping, the subjects were overall in a state of irritable dysphoria. The score on YMRS may be measuring the activation aspect of this state rather than true (hypo) mania.

Examples of phenes that clustered together most closely across all three psychotic disorders groups, in our preliminary

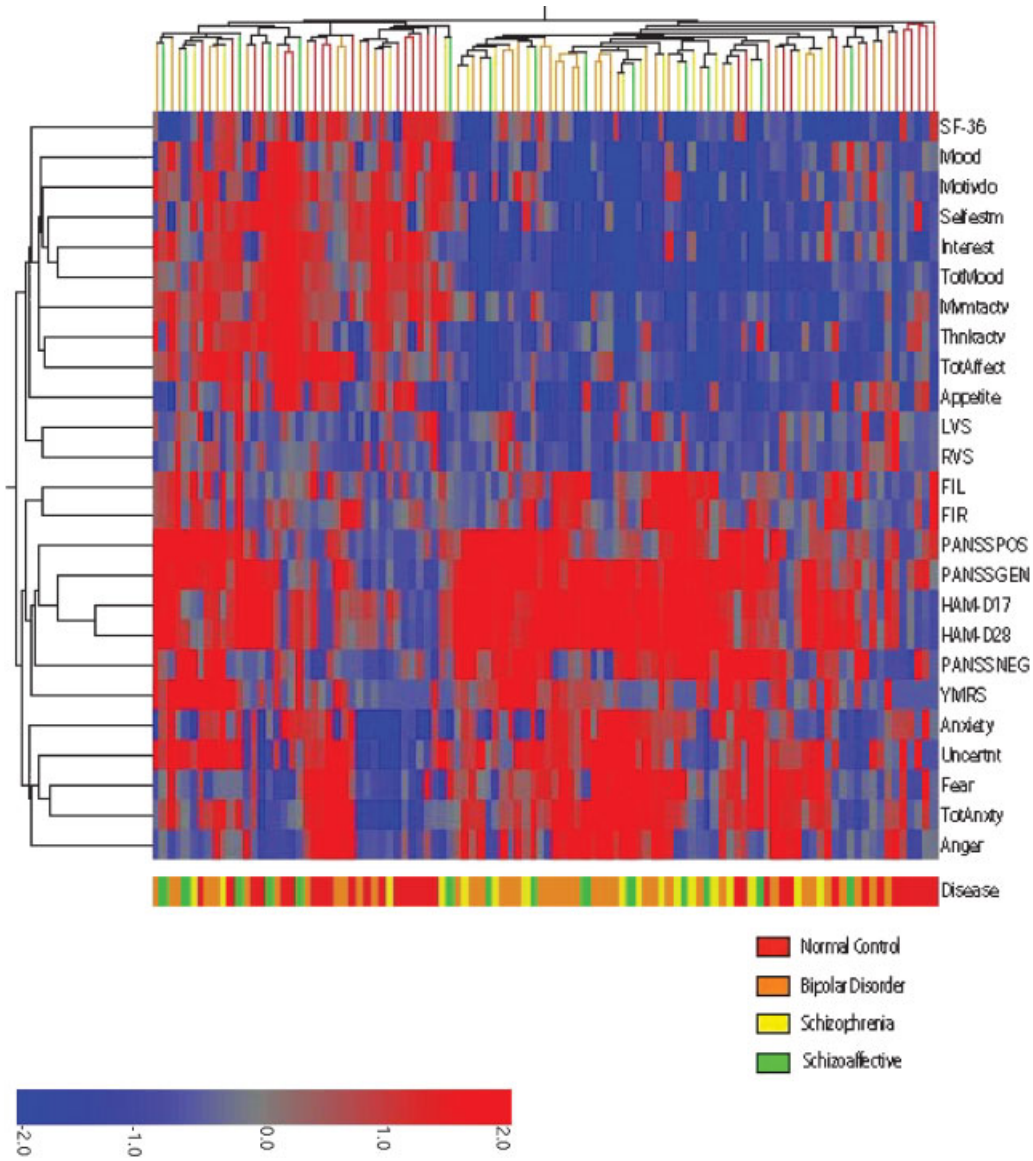

Fig. 4. Clustering of subjects: *heterogeneity within individual psychotic disorders*. Two-way hierarchical clustering of all subjects and 25 phenes based on the Z score for each phene. All individual effect sizes were calculated by comparing each individual subject's phenes with the averages of the normal controls.

results so far, are: Motivation and Total Affect, Self-esteem and Appetite, Fear and Total Anxiety, RVS, and Thinking Activity, Force Instability of Right hand (FIR), and PANSSPOS, Force Instability Left hand (FIL), and PANSSNEG.

Our non-hypothesis driven, discovery-based approach thus uncovers new empirical relationships between phenotypic items, which are of high neurobiological interest. One such result is the relationship between motivation to do things and affective state. Another on is the relationship between self-esteem and appetite, with clinical implications for abnormal weight changes in these and related disorders.

Our approach also uncovers relationships between objective phenes (hand neuromotor measures) and subjective phenes. One such result is the relationship between RVS and Thinking Activity, suggesting a possible left hemispheric dominance of the neurobiological correlate of this measured phenotype. This may have clinical implications for using left-hemisphere stimulation, through methods such as Transcranial Magnetic Stimulation (TMS), for patients with sluggish thinking, as seen in depression or negative symptoms schizophrenia.

Clustering of Subjects

Next, unsupervised two-way hierarchical clustering was applied to all of the 104 subjects, based on the Z scores for all the phenes across all of the subjects. Results are displayed in Figure 4 as a color-coded heat map, where the subjects are ordered on the horizontal axis and phenes are ordered on the vertical axis on the basis of similarity of their individual effect sizes. Interestingly, the four major diagnostic groups as established by SCID (normal controls, bipolar, schizophrenia, schizoaffective) fail to cluster together in four distinct groups. The fact that subjects from different diagnostic groups are interspersed speaks to the overlap among current diagnostic classifications (including normal controls), as well as to their internal heterogeneity. More profoundly, it points to the fact that the somewhat artificial boundaries between control and

affected subjects may become blurred when dimensional rather than categorical approaches are used, and should provide a rationale and impetus for population QTL studies in psychiatric genetics.

The clustering of this set of individual subjects leads to pairs of highly similar subjects (pseudo-twins) from different diagnostic groups that share more characteristics with each other than with subjects in their own diagnostic group. It has not escaped our attention that this methodology may prove useful in pairing subjects for genetic, pharmacological, and imaging studies. Moreover, clinically, it identifies subjects that may respond similarly to treatments, and should be treated psychiatrically in the same way.

DISCUSSION

We have developed, and present initial proof of principle for, an empirical approach to characterizing psychiatric phenotypes, termed PhenoChipping. The approach consists of a massive parallel sampling of cognitive and affective state, employing paradigms and analysis tools from the microarray gene expression field. Our preliminary results revealed overlap among, as well as heterogeneity within, the three major psychotic disorders studied: bipolar disorder, schizophrenia, and schizoaffective disorder. Moreover, the use of hand neuromotor measures has provided preliminary evidence supportive of hemispheric lateralization of cognition and mood, as well as leads for objective behavioral biomarker development.

Limitations and Confounds

First, our sample, while well characterized and providing some interesting insights, is limited in size. It is likely that a larger sample size would increase the statistical power of our results, and the discriminatory power of the PhenoChipping approach.

Second, our preliminary data is based on cross-sectional analysis at a single timepoint, which may reflect predominantly state rather than trait factors. Multiple PhenoChipping measurements, at different timepoints, would likely permit addressing state versus trait issues, by looking at how phenes change over time.

Third, as all subjects with major psychotic disorders that were enrolled were on a psychopharmacological regimen, our results may reflect, at least in part, a combination of medication (side) effects and underlying disease phenomenology. This may be especially true for hand motor measures in patients on antipsychotic medication (for the Velocity Scaling measure), or mood stabilizing medications (for the Force Instability measure). It remains for future work to address the issue of factoring out possible medication effects by an intra-subject analysis of repeated measurements, while the subjects are on the same pharmacology. Moreover, PhenoChipping of first degree relatives who do not have overt clinical illness, are unmedicated, but may have (endo)phenotypic abnormalities, is an important area of future research.

Fourth, in the studies described in this report, we did not collect blood for genetic and genomic studies. This should be an integral part of future PhenoChipping work, in view of building a blood sample bank with comprehensive phenotypic information attached to it, for studies integrating phenotype with genotype and gene expression. DNA extracted from whole blood can be used for studies of polymorphisms in candidate genes of interest related to bipolar disorder and schizophrenia [Niculescu et al., 2000; Chowdari et al., 2002; Barrett et al., 2003; Hattori et al., 2003; Geller et al., 2004; Kohn et al., 2004; Ogden et al., 2004; Green et al., 2006]. Whole blood (predominantly lymphocyte) RNA can be extracted for microarray gene expression studies. Lymphocyte gene expression profiling has emerged as a particularly interesting area of research in the quest for peripheral molecular biomarkers of mental illness [Vawter et al., 2004; Glatt et al., 2005; Middleton et al., 2005; Segman et al., 2005; Tsuang et al., 2005]. Fresh blood, with phenotypic state information gathered at time of harvesting, may be more informative than immortalized lymphocytes, and avoid some of the caveats of Epstein–Barr virus (EBV) immortalization and cell culture passaging.

Conclusions and Future Directions

Our objective was to develop and provide proof of principle for a novel approach, PhenoChipping, as a way to better understand the phenotypic structure of major psychotic disorders. Our preliminary data documents both overlap among, and heterogeneity within, the three major psychotic disorder studied, and suggests a combinatorial building-block (Lego-like) structure underlies these psychiatric syndromes. It is hoped that our approach, boot-strapped and integrated with genetics, genomics, and imaging studies, will help move the field of psychiatry forward, beyond current categorical diagnostic classifications. The integration of phenotypic data with genotype and gene expression data will be particularly facilitated by the importation of normalized data inside the same analysis software, GeneSpring or another similar program. This was one of the key reasons, we adapted its use for phenotypic data analysis. Applying a methodology developed in one discipline (genomics) can facilitate understanding of another discipline (phenomics) [Kelsoe and Niculescu, 2002], especially if one is trying to connect the two together.

An immediate practical application for our integrative strategy would be in pharmacogenomics; a second would be the identification of peripheral behavioral and molecular biomarkers of illness. A better understanding of major psychotic disorders such as bipolar disorder, schizophrenia, and schizoaffective disorder, will lead to more targeted treatments, with improved efficacy and decreased side-effects. This will have an impact on patient health, well-being, quality of life, and independent functioning. Moreover, early diagnosis and intervention may prevent the full-blown development of illness in genetically susceptible individuals. Ultimately, we propose that our work is an initial step in the direction of developing more individualized diagnosis and treatments for psychiatric patients-personalized psychiatry as a component of personalized medicine [Gould and Manji, 2004], where patients' individual profiles, rather than broad categorical diagnostic constructs, are the targets of therapeutic intervention.

ACKNOWLEDGMENTS

This work was supported by an APA/AstraZeneca Young Minds in Psychiatry award to ABN, as well as VISN-22 MIRECC funds and a VA Merit Award to JBL. We thank Dr. John I. Nurnberger, Jr. (Indiana University School of Medicine) for useful discussions and comments.

REFERENCES

American Psychiatric Association, Task Force on DSM-IV. 2000. Diagnostic and statistical manual of mental disorders: DSM-IV-TR. Washington, DC: American Psychiatric Association. xxxvii, 943 p.

Barrett TB, Hauger RL, Kennedy JL, Sadovnick AD, Remick RA, Keck PE, McElroy SL, Alexander M, Shaw SH, Kelsoe JR. 2003. Evidence that a single nucleotide polymorphism in the promoter of the G protein receptor kinase 3 gene is associated with bipolar disorder. Mol Psychiatry 8(5):546–557.

Berrettini WH. 2000. Are schizophrenic and bipolar disorders related? A review of family and molecular studies. Biol Psychiatry 48(6):531–538.

Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, Hendrix M, Radmacher M, Simon R, Yakhini Z, Ben-Dor A, et al. 2000. Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature 406(6795):536–540.

Caligiuri MP, Lohr JB, Ruck RK. 1998. Scaling of movement velocity: A measure of neuromotor retardation in individuals with psychopathology. Psychophysiology 35(4):431–437.

Caligiuri MP, Gentili V, Eberson S, Kelsoe J, Rapaport M, Gillin JC. 2003. A quantitative neuromotor predictor of antidepressant non-response in patients with major depression. J Affect Disord 77(2):135–141.

Caligiuri MP, Brown GG, Meloy MJ, Eyler LT, Kindermann SS, Eberson S, Frank LR, Lohr JB. 2004. A functional magnetic resonance imaging study of cortical asymmetry in bipolar disorder. Bipolar Disord 6(3): 183–196.

Caligiuri MP, Brown GG, Meloy MJ, Eberson S, Niculescu AB, Lohr JB. 2006. Striatopallidal regulation of affect in bipolar disorder. J Affect Disord 91(2–3):235–242.

Chowdari KV, Mirnics K, Semwal P, Wood J, Lawrence E, Bhatia T, Deshpande SN, BK T, Ferrell RE, Middleton FA, et al. 2002. Association and linkage analyses of RGS4 polymorphisms in schizophrenia. Hum Mol Genet 11(12):1373–1380.

Cohen J. 1988. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: L. Erlbaum Associates. xxi, 567 p.

Cortese L, Caligiuri MP, Malla AK, Manchanda R, Takhar J, Haricharan R. 2005. Relationship of neuromotor disturbances to psychosis symptoms in first-episode neuroleptic-naive schizophrenia patients. Schizophr Res 75(1):65–75.

Craddock N, Jones I, Kirov G, Jones L. 2004. The Bipolar Affective Disorder Dimension Scale (BADDS)–a dimensional scale for rating lifetime psychopathology in bipolar spectrum disorders. BMC Psychiatry 4:19.

Craddock N, O'Donovan MC, Owen MJ. 2006. Genes for schizophrenia and bipolar disorder? Implications for psychiatric nosology. Schizophr Bull 32(1):9–16.

Geller B, Badner JA, Tillman R, Christian SL, Bolhofner K, Cook EH Jr. 2004. Linkage disequilibrium of the brain-derived neurotrophic factor Val66Met polymorphism in children with a prepubertal and early adolescent bipolar disorder phenotype. Am J Psychiatry 161(9):1698–1700.

Glatt SJ, Everall IP, Kremen WS, Corbeil J, Sasik R, Khanlou N, Han M, Liew CC, Tsuang MT. 2005. Comparative gene expression analysis of blood and brain provides concurrent validation of SELENBP1 up-regulation in schizophrenia. Proc Natl Acad Sci USA 102(43):15533–15538.

Gottesman II, Gould TD. 2003. The endophenotype concept in psychiatry: Etymology and strategic intentions. Am J Psychiatry 160(4):636–645.

Gould TD, Manji HK. 2004. The molecular medicine revolution and psychiatry: Bridging the gap between basic neuroscience research and clinical psychiatry. J Clin Psychiatry 65(5):598–604.

Green EK, Raybould R, Macgregor S, Hyde S, Young AH, O'Donovan MC, Owen MJ, Kirov G, Jones L, Jones I, et al. 2006. Genetic variation of brain-derived neurotrophic factor (BDNF) in bipolar disorder: Case-control study of over 3000 individuals from the UK. Br J Psychiatry 188:21–25.

Hamilton M. 1960. A rating scale for depression. J Neurol Neurosurg Psychiatry 23:56–62.

Hamilton M. 1980. Rating depressive patients. J Clin Psychiatry 41(12 Pt 2): 21–24.

Harrison PJ, Weinberger DR. 2005. Schizophrenia genes, gene expression, and neuropathology: On the matter of their convergence. Mol Psychiatry 10(1):40–68; image 5.

Hasler G, Drevets WC, Manji HK, Charney DS. 2004. Discovering endophenotypes for major depression. Neuropsychopharmacology 29(10):1765–1781.

Hattori E, Liu C, Badner JA, Bonner TI, Christian SL, Maheshwari M, Detera-Wadleigh SD, Gibbs RA, Gershon ES. 2003. Polymorphisms at the G72/G30 gene locus, on 13q33, are associated with bipolar disorder in two independent pedigree series. Am J Hum Genet 72(5):1131–1140.

Hyman SE. 2003. Diagnosing disorders. Sci Am 289(3):96–103.

Hyman SE, Fenton WS. 2003. Medicine. What are the right targets for psychopharmacology? Science 299(5605):350–351.

Janca A, Kastrup MC, Katschnig H, Lopez-Ibor JJ, Mezzich JE, Sartorius N. 1996. The ICD-10 multiaxial system for use in adult psychiatry: Structure and applications. J Nerv Ment Dis 184(3):191–192.

Kay SR, Fiszbein A, Opler LA. 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 13(2):261–276.

Kelsoe JR, Niculescu AB 3rd. 2002. Finding genes for bipolar disorder in the functional genomics era: From convergent functional genomics to phenomics and back. CNS Spectr 7(3):215–226.

Kohn Y, Danilovich E, Filon D, Oppenheim A, Karni O, Kanyas K, Turetsky N, Korner M, Lerer B. 2004. Linkage disequlibrium in the DTNBP1 (dysbindin) gene region and on chromosome 1p36 among psychotic patients from a genetic isolate in Israel: Findings from identity by descent haplotype sharing analysis. Am J Med Genet Part B Neuropsychiatr Genet 128(1):65–70.

Krishnan KR. 2005. Psychiatric disease in the genomic era: Rational approach. Mol Psychiatry 10(11):978–984.

Kupfer DJ. 2005. Dimensional models for research and diagnosis: A current dilemma. J Abnorm Psychol 114(4):557–559.

Lenox RH, Gould TD, Manji HK. 2002. Endophenotypes in bipolar disorder. Am J Med Genet 114(4):391–406.

Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, et al. 2005. MicroRNA expression profiles classify human cancers. Nature 435(7043):834–838.

Maremmani I, Akiskal HS, Signoretta S, Liguori A, Perugi G, Cloninger R. 2005. The relationship of Kraepelian affective temperaments (as measured by TEMPS-I) to the tridimensional personality questionnaire (TPQ). J Affect Disord 85(1–2):17–27.

Mezzich JE. 1978. Evaluating clustering methods for psychiatric diagnosis. Biol Psychiatry 13(2):265–281.

Middleton FA, Pato CN, Gentile KL, McGann L, Brown AM, Trauzzi M, Diab H, Morley CP, Medeiros H, Macedo A, et al. 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. Am J Med Genet Part B Neuropsychiatr Genet 136B(1):12–25.

Minassian A, Paulus MP, Perry W. 2004. Increased sensitivity to error during decision-making in bipolar disorder patients with acute mania. J Affect Disord 82(2):203–208.

Niculescu AB 3rd, Akiskal HS. 2001. Proposed endophenotypes of dysthymia: Evolutionary, clinical and pharmacogenomic considerations. Mol Psychiatry 6(4):363–366.

Niculescu AB 3rd, Segal DS, Kuczenski R, Barrett T, Hauger RL, Kelsoe JR. 2000. Identifying a series of candidate genes for mania and psychosis: A convergent functional genomics approach. Physiol Genomics 4(1):83–91.

Ogden CA, Rich ME, Schork NJ, Paulus MP, Geyer MA, Lohr JB, Kuczenski R, Niculescu AB. 2004. Candidate genes, pathways and mechanisms for bipolar (manic-depressive) and related disorders: An expanded convergent functional genomics approach. Mol Psychiatry 9(11):1007–1029.

Segman RH, Shefi N, Goltser-Dubner T, Friedman N, Kaminski N, Shalev AY. 2005. Peripheral blood mononuclear cell gene expression profiles identify emergent post-traumatic stress disorder among trauma survivors. Mol Psychiatry 10(5):500–513 , 425.

Spitzer RL. 2001. Values and assumptions in the development of DSM-III and DSM-III-R: An insider's perspective and a belated response to Sadler, Hulgus, and Agich's "On values in recent American psychiatric classification". J Nerv Ment Dis 189(6):351–359.

Spitzer RL, First MB. 2005. Classification of psychiatric disorders. JAMA 294(15):1898–1899; author reply 1899–1900.

Spitzer RL, Williams JB. 1994. American psychiatry's transformation following the publication of DSM-III. Am J Psychiatry 151(3):459–460.

Tamminga CA, Holcomb HH. 2005. Phenotype of schizophrenia: A review and formulation. Mol Psychiatry 10(1):27–39.

Tsuang MT, Nossova N, Yager T, Tsuang MM, Guo SC, Shyu KG, Glatt SJ, Liew CC. 2005. Assessing the validity of blood-based gene expression profiles for the classification of schizophrenia and bipolar disorder: A preliminary report. Am J Med Genet Part B Neuropsychiatr Genet 133(1):1–5.

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(1):41–52.

Ware JE Jr, Bayliss MS, Rogers WH, Kosinski M, Tarlov AR. 1996. Differences in 4-year health outcomes for elderly and poor, chronically ill patients treated in HMO and fee-for-service systems. Results from the Medical Outcomes Study. JAMA 276(13):1039–1047.

Young RC, Biggs JT, Ziegler VE, Meyer DA. 1978. A rating scale for mania: Reliability, validity and sensitivity. Br J Psychiatry 133:429–435.