Published in final edited form as:
Curr Pharmacogenomics Person Med. 2009 September ; 7(3): 164–188.

Alternatively Spliced Genes as Biomarkers for Schizophrenia, Bipolar Disorder and Psychosis: A Blood-Based Spliceome-Profiling Exploratory Study

S.J. Glatt[1,*], **S.D. Chandler**[2,#], **C.A. Bousman**[2,#], **G. Chana**[2], **G.R. Lucero**[2], **E. Tatro**[2], **T. May**[2], **J.B. Lohr**[2], **W.S. Kremen**[2,3], **I.P. Everall**[2], and **M.T. Tsuang**[2,3,4]

[1] Department of Psychiatry and Behavioral Sciences, and Medical Genetics Research Center; SUNY Upstate Medical University; 750 East Adams Street; Syracuse, NY, 13210; USA

[2] Center for Behavioral Genomics; Department of Psychiatry; University of California, San Diego; 9500 Gilman Drive; La Jolla, CA 92039; USA

[3] Veterans Affairs San Diego Healthcare System; 3350 La Jolla Village Drive; San Diego, CA 92161; USA

[4] Harvard Institute of Psychiatric Epidemiology and Genetics; Harvard Departments of Epidemiology and Psychiatry; 25 Shattuck Street; Boston, MA 02115; USA

Abstract

Objective—Transcriptomic biomarkers of psychiatric diseases obtained from a query of peripheral tissues that are clinically accessible (e.g., blood cells instead of post-mortem brain tissue) have substantial practical appeal to discern the molecular subtypes of common complex diseases such as major psychosis. To this end, spliceome-profiling is a new methodological approach that has considerable conceptual relevance for discovery and clinical translation of novel biomarkers for psychiatric illnesses. Advances in microarray technology now allow for improved sensitivity in measuring the transcriptome while simultaneously querying the “exome” (all exons) and “spliceome” (all alternatively spliced variants). The present study aimed to evaluate the feasibility of spliceome-profiling to discern transcriptomic biomarkers of psychosis.

Methods—We measured exome and spliceome expression in peripheral blood mononuclear cells from 13 schizophrenia patients, nine bipolar disorder patients, and eight healthy control subjects. Each diagnostic group was compared to each other, and the combined group of bipolar disorder and schizophrenia patients was also compared to the control group. Furthermore, we compared subjects with a history of psychosis to subjects without such history.

Results—After applying Bonferroni corrections for the 21,866 full-length gene transcripts analyzed, we found significant interactions between diagnostic group and exon identity, consistent with group differences in rates or types of alternative splicing. Relative to the control group, 18 genes in the bipolar disorder group, eight genes in the schizophrenia group, and 15 genes in the combined bipolar disorder and schizophrenia group appeared differentially spliced. Importantly,

[*]Address correspondence to this author at the Department of Psychiatry and Behavioral Sciences, and Medical Genetics Research Center; SUNY Upstate Medical University; 750 East Adams Street; Syracuse, NY, 13210, USA; Tel: (315) 464-7742; Fax: (315) 464-7744; glatts@upstate.edu.
[#]These authors contributed equally to this work.

DUALITY/CONFLICT OF INTERESTS
None declared/applicable.

SUPPLEMENTARY MATERIAL
Supplementary material is available on the publishers Web site along with the published article.

thirty-three genes showed differential splicing patterns between the bipolar disorder and schizophrenia groups. More frequent exon inclusion and/or over-expression was observed in psychosis. Finally, these observations are reconciled with an analysis of the ontologies, the pathways and the protein domains significantly over-represented among the alternatively spliced genes, several of which support prior discoveries.

Conclusions—To our knowledge, this is the first blood-based spliceome-profiling study of schizophrenia and bipolar disorder to be reported. The battery of alternatively spliced genes and exons identified in this discovery-oriented exploratory study, if replicated, may have potential utility to discern the molecular subtypes of psychosis. Spliceome-profiling, as a new methodological approach in transcriptomics, warrants further work to evaluate its utility in personalized medicine. Potentially, this approach could also permit the future development of tissue-sampling methodologies in a form that is more acceptable to patients and thereby allow monitoring of dynamic and time-dependent plasticity in disease severity and response to therapeutic interventions in clinical psychiatry.

Keywords

Gene expression; spliceome-profiling; blood-based biomarkers; psychosis; CNS disease diagnostics

INTRODUCTION

Schizophrenia and bipolar disorder are two of the most prevalent mental health disorders, collectively affecting approximately 1.5% of the population at some point in their lifetime [1,2]. These are also among the most severe and debilitating disorders with substantial public health importance, typically developing in late adolescence or early adulthood and often imposing near-lifelong disability on affected individuals, as well as considerable strain on their caretakers and society at large. One way to combat schizophrenia and bipolar disorder would be to discover biological markers - or “biomarkers” - for these illnesses, which potentially could revolutionize their rational diagnosis and management. Conceivably, biomarkers could expedite and standardize the process of primary and differential diagnosis, which presently involves considerable time, effort, and uncertainty. Biomarkers also might allow for earlier identification of affected individuals, in turn hastening their receipt of effective (perhaps personalized) treatment and improving prognoses. Furthermore, biomarkers could form the basis for early intervention and prevention efforts targeting at- risk individuals, which might reduce the morbidity and prevalence of these crippling disorders. Collectively, these advances would translate into an enormous improvement in global public health.

Given such immense promise, biomarkers for schizophrenia and bipolar disorder have been pursued through many decades and approaches. As both disorders are thought to have relatively high heritability (~80%) [3,4], biomarkers tied to specific candidate genes have often formed the basis for these pursuits. For example, both schizophrenia and bipolar disorder have been the subject of hundreds of candidate-gene association analyses [5–7] and, most recently, a handful of genome-wide association studies [8–14]. Through meta-analysis, our group and others have identified several genetic polymorphisms that influence the risk for one or the other disorder [6,15–21]. Collectively, these polymorphisms may explain a small portion of the heritability of each disorder; however, the majority of heritable variance in susceptibility remains unexplained by such variants, as does the considerable amount of risk attributable to non-heritable factors. Furthermore, genetic polymorphisms, as (inherited or acquired) fixed factors, lack flexibility as biomarkers due to their inability to fluctuate over time within individuals and to reflect the course of illness, including periods of risk,

prodrome, first episode, chronic illness, remission, and treatment-responsiveness or non-responsiveness. As such, comprehensive biomarker profiles of complex psychiatric disorders may require the integration of markers from multiple “omic” domains [22], including static genomic factors (*i.e.*, DNA polymorphisms) as well as dynamic factors reflected in the transcriptome [*i.e.*, the total messenger RNA (mRNA) expressed in a cell or tissue at a given time [23].

Over the past four years, we and others [24–28] have documented the potential utility of blood-based transcrip-tomic profiling of mRNA abundances by microarray as a source of biomarkers for schizophrenia and bipolar disorder. We first used mRNA expression patterns in circulating peripheral blood mononuclear cells (PBMCs) to identify a large number of genes whose expression levels distinguished patients with schizophrenia or bipolar disorder from each other and from unaffected control subjects based on liberal significance criteria ($p<0.05$) [29]. We later re-analyzed those data adopting a more conservative permutation-based approach toward the control of type-I errors, which reduced the number of genes identified as differentially expressed in the blood of schizophrenia patients (relative to control subjects) from 567 to 123. Six of these genes were differentially expressed in both PBMCs and postmortem brain tissue from a separate sample of schizophrenia patients, further supporting their candidacy as biomarkers [30]. Of these, *SELENBP1* (which codes for selenium binding protein 1) subsequently emerged as the strongest putative biomarker based on the similar magnitude of its up-regulation in both PBMCs and brain, validation of this effect at the protein level, and replication of *SELENBP1* dysregulation in an independent series of postmortem brain tissue samples [31]. Interestingly, the strongest result in our replication study was observed in the comparison of tissue from individuals with a history of psychosis (including all schizophrenia patients and a portion of bipolar disorder patients) and individuals without such a history (including the remaining non-psychotic bipolar disorder patients and all unaffected control subjects). This result in particular suggested that future pursuits of schizophrenia and bipolar disorder biomarkers in PBMCs might also profit from a focus on psychosis as a common feature of the disorders, which may be more strongly linked than either diagnosis to changes in the transcriptome. In fact, this conceptualization mirrors the emerging recognition of partially overlapping genetic contributions to the etiology of the two disorders [4,32–35].

Toward facilitating the identification of mRNA bio- markers of disease, technology has been developed recently that exponentially increases the sensitivity and specificity of existing transcriptome-profiling systems (and eliminates 3′ bias) by yielding estimates of exon-level mRNA abundance. In addition to the ability to summate exon-level data into a more accurate measure of full-length gene expression, these microarrays allow for the identification and measurement of different splice variants of all expressed genes as well as the detection of novel splicing events. Given prior evidence of alternative splicing of select candidate genes in schizophrenia [36–40] and bipolar disorder [41], it was of keen interest to determine if a preliminary survey of the entire human “spliceome” in PBMCs could reveal useful biomarkers for these disorders, as well as their sometimes-shared clinical feature, psychosis. Additionally, the present study presents a critical overview on the feasibility of spliceome-profiling as a new methodological approach in transcriptomics, and in personalized medicine research more generally.

MATERIALS AND METHODS

Ascertainment

Subjects with schizophrenia (SCZ; $n=13$) or bipolar disorder (BPD; $n=11$) were recruited from the University of California, San Diego (UCSD) Psychopharmacology Research Initiatives Center for Excellence (PRICE) participant network. Healthy control subjects

(CNT; *n*=10) were recruited from the same catchment area through the use of flyers and print advertisements. All participants underwent a brief initial phone screening to assess their appropriateness for possible inclusion in the study. During this screening, information was gathered related to the study's inclusion and exclusion criteria, as well as basic demographics (*e.g.*, age, sex, ancestry).

Inclusion criteria required participants to: 1) be between the age of 18 and 55 years; 2) have at least an eighth-grade education; 3) speak English as their first language; and 4) have no documented evidence of mental retardation. Subjects in the two patient groups (SCZ and BPD) were further required to have met criteria for their primary diagnosis (schizophrenia or bipolar disorder) for at least two years. Exclusion criteria were: 1) substance abuse or dependence in the past year; 2) neurologic problems (*e.g.*, stroke, meningitis); 3) systemic medical illnesses (*e.g.*, heart disease, diabetes); 4) history of head injury with documented loss of consciousness lasting longer than 10 minutes; 5) pregnancy; or 6) physical disabilities. Subjects in the CNT group were also excluded if they had a personal or family history of a psychotic disorder, bipolar disorder, major depressive disorder, or a cluster-A (schizotypal, schizoid, or paranoid) personality disorder. All participants gave written consent prior to enrollment in the study, and all study procedures were approved by the Institutional Review Board at UCSD.

Clinical Assessment and Data Analyses

Individuals satisfying inclusion and exclusion criteria during the initial phone screening were scheduled for a two-hour in-person clinical assessment using the Diagnostic Interview for Genetic Studies (DIGS) [42], which was administered by a trained Masters-level research assistant. The DIGS was used to verify information obtained from the phone screening, collect additional psychiatric data, and ultimately assist in accurate diagnosis and classification. To ensure accurate diagnosis and classification of participants, each DIGS was reviewed by two independent doctoral-level clinicians. When discrepancies in diagnoses occurred, an attempt was made to resolve them and come to a consensus diagnosis. If a consensus could not be reached, the participant was excluded from the study. Participants were also excluded if one or more of the inclusion or exclusion criteria were found not to be satisfied following review of the DIGS data.

The DIGS interview also yielded information on important covariates and potential confounding variables of relevance to the analyses of gene expression and alternative splicing. Age was measured continuously in years, whereas sex (male/female), ancestry (European, African, Hispanic, or Asian), current smoking status (yes/no), and history of psychosis (yes/no) were coded as categorical variables. Finally, each participant's current medication regimen was reviewed and coded as a binary categorical variable (yes/no) for each of the predominant classes of medication used by subjects in the sample, including antipsychotic and mood-stabilizing drugs.

Continuously distributed demographic and clinical variables were compared between diagnostic groups by analyses of variance (ANOVAs), while categorical demographic and clinical variables were compared between the groups by χ^2-tests. These analyses were conducted in Stata SE software, version 9.2 (StataCorp; College Station, TX)

mRNA Sample Acquisition, Stabilization, Isolation, and Storage

After completing the DIGS interview, subjects were scheduled to provide a 10-ml sample of blood on a subsequent visit. To control for potential environmental (*e.g.*, diet) and biological (*e.g.*, circadian) influences on gene expression, all blood draws were performed in the morning after subjects fasted overnight. Each blood sample was collected into an EDTA-

coated collection tube and immediately transferred to an RNase-free laboratory, where all subsequent procedures took place. The blood sample was passed over a LeukoLOCK™ filter, which was flushed with PBS and then fully saturated with RNAlater® [43]. Each LeukoLOCK™ filter, containing bound, isolated, stabilized, and purified white blood cells, was sealed and stored in a sterile box at -20°C. Once mRNA samples were acquired from all 34 subjects, the entire batch of samples was processed to isolate mRNA. Eluted mRNA samples were stored at -20°C until transferred to the GeneChip™ Microarray Core (San Diego, CA) for quality assurance and microarray hybridization. LeukoLOCK™ filters, RNAlater®, and TRI reagent® were obtained from Applied Biosystems, Inc. (Foster City, CA), while all other reagents and supplies were obtained from VWR International, LLC (West Chester, PA) unless otherwise specified.

mRNA Quantitation and Quality Assurance

The concentration of mRNA in each DNA-free sample was quantified by the absorption of ultraviolet light at two wavelengths (260 and 280 nm), which was measured on a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific; Wilmington, Delaware). The quantity of mRNA in each of the 34 samples far exceeded the minimally sufficient amount required for microarray hybridization. The purity of each mRNA sample was estimated by the 260:280 nm absorbance ratio, with an acceptable range designated *a priori* as 1.7–2.1. The sample from one BPD subject had a value below this range and thus was excluded from further analyses. The quality of each mRNA sample was quantified by the RNA Integrity Number (RIN) [44], which was determined on an RNA 6000 Labchip Kit on an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.; Santa Clara, CA). According to convention [44], a RIN of 6.0 or greater was deemed to be indicative of acceptable quality. The samples from one BPD subject and two CNT subjects had values below 6.0 and thus were excluded from further analyses. A total of 30 subjects (SCZ, n=13; BPD, n=9; CNT, n=8) provided samples with acceptable levels of mRNA quantity, purity, and quality, which were then assayed on GeneChip® Human Exon 1.0 ST Arrays (Affymetrix, Inc.; Santa Clara, CA) per the “Whole Transcript (WT) Sense Target Labeling Assay” protocol [45] using 1 μg of total RNA from each sample.

Microarray Data Import, Normalization, Transformation, and Summarization

Partek® Genomics Suite software, version 6.3 © 2008 (Partek Incorporated; St. Louis, MO), was utilized for all analytic procedures performed on microarray scan data. First, interrogating probes from the most conservative “Core” probe set were imported. According to the manufacturer, probes in this set “are supported with the most reliable evidence from RefSeq and full length mRNA GenBank records containing complete CDS (coding sequence) information” [46]. Corrections for background signal were applied using the robust multi-array average (RMA) method [47]. The set of 30 GeneChips was standardized using quan-tile normalization, and expression levels of each probe underwent log-2 transformation to yield distributions of data that more closely approximated normality. As each exon was typically measured by multiple probe sets, summari- zation of redundant probe sets was obtained by median polish. According to convention [48], probe sets with a maximum signal:noise ratio of less than 3.0 were excluded from subsequent analyses.

Microarray Data Analyses

Five sets of comparisons of diagnostic groups were performed, as follows: 1) BPD *vs*. CNT; 2) SCZ *vs*. CNT; 3) BPD *vs*. SCZ; 4) BPD+SCZ *vs*. CNT; and 5) subjects with a history of psychosis [PSYCH(+), which included all SCZ subjects and six of the nine BPD subjects] *vs*. subjects with no history of psychosis [PSYCH(−), which included all CNT subjects and the remaining three BPD subjects].

The principal analyses of these data were designed to detect expressed genes (*i.e.*, full-length mRNA transcripts) that were significantly different in constitution between diagnostic groups, which might indicate different rates or types of alternative splicing events between them. These analyses compared groups of interest on the mean expression level of all exons in each gene on a gene-by-gene basis through analyses of covariance (ANCOVAs) and inspection of interaction terms. Four classes of factors were included in each ANCOVA model. First, we included diagnostic group as the primary factor of interest. Next, we included potentially important background biological characteristics of the subjects, including age, sex, and ancestry. Third, we included variables that accounted for expected (and ultimately observed) differences between the groups in their current rates of smoking and use of prescribed psychotropic medications. Comparisons involving the SCZ group included a variable indexing current use of an antipsychotic medication, while comparisons involving the BPD group included a variable to indicate whether or not each subject was currently on a mood-stabilizer; analyses that involved both patient groups included both medication indicator variables. Lastly, we included a series of factors that allowed for the detection of alternate splicing events. Since not all exons in a gene express at the same level, exon identity (ID) was added to the model to account for exon-to-exon differences. Since multiple measurements (on the multiple exons) come from the same subject, subject ID was added to the model to accommodate the assumption of independence that is fundamental to ANCOVA. The last, most critical term included was the interaction of exon ID with diagnostic group, which allowed for the detection of differences in the expression of any exons in the different diagnostic groups [49].

Once a gene was identified as being influenced by a significant diagnostic group-by-exon ID interaction (indicative of alternative splicing), *post-hoc* comparisons were made between the groups to identify dysregulated exons in these genes by simply comparing diagnostic groups on the mean expression level of each individual exon using ANCOVAs. These analyses included the same factors and interaction terms described above for whole-gene analyses, except for those factors that enabled the detection of alternate splicing (*i.e.*, exon identity and its interaction with diagnostic group). These analyses were useful for identifying which exons contributed to or accounted for the significance of alternative splicing events identified in the principal omnibus analyses.

After all quality-control procedures were executed, 21,866 full-length gene transcripts and 232,448 exons were included in the analyses. Due to the large number of statistical tests to be performed, the probability of committing type-I errors (*i.e.*, finding false-positive results) in this study was greatly inflated. We addressed the threat of inflation of the type-I-error rate in three ways. The first method was statistical, as we controlled family-wise error rates (FWERs) at 5% using Bonferroni corrections and estimated the false-discovery rate (FDR) for all comparisons (expressed as *q*-values, or the proportion of findings at a given significance level that are expected to be false discoveries). In general, this approach followed the guidelines provided by Mirnics *et al*. [50], with the exception that we did not filter genes based on any fold-change criterion, since it is not empirically known (nor would it be expected) that any fold-change criterion is either universally applicable to all genes or is biologically meaningful. Secondly, we required a significant interaction of diagnostic group and exon ID in omnibus analyses of full-length gene transcripts before comparing the expression levels of individual exons between groups. The third method was by data reduction through secondary analyses of groups of genes: after performing each type of analysis, the generated lists of significantly alternatively spliced genes were subjected to the DAVID algorithm [51] to determine if they were enriched for genes that aggregated in the same biological pathways (defined by BioCarta- or KEGG [52]), represented similar ontologies (defined by GOC [53]), or exhibited common protein domains (defined by InterPro [54] or PIR [55]).

RESULTS

Demographic, Substance Use, and Clinical Variables

The three ascertained diagnostic groups were comparable on demographic variables (Table 1). No significant differences were observed between SCZ, BPD, and CNT groups on age (p=0.788), sex (p=0.788), or ancestry (p=0.264). The number of current smokers differed significantly between the groups (p=0.003), with the CNT group having significantly fewer current smokers than either the SCZ group ($\chi^2_{(1)}$=11.748, p=0.001) or the BPD group ($\chi^2_{(1)}$= 6.296, p=0.012); however, notably, the rate of smoking did not differ between SCZ and BPD groups ($\chi^2_{(1)}$=1.119, p=0.290).

As expected, the groups also differed in their rates of use of antipsychotic (p<0.001) and mood-stabilizing medications (p=0.001). The rate of use of antipsychotic medication was higher in the SCZ group than in either the BPD group ($\chi^2_{(1)}$=4.090, p=0.043) or, obviously, the CNT group ($\chi^2_{(1)}$= 17.231, p<0.001), and the BPD group also used this class of medications at a higher rate than did the CNT group ($\chi^2_{(1)}$=6.296, p=0.012). Mood stabilizers were used at a higher rate in the BPD group than in either the SCZ group ($\chi^2_{(1)}$=8.564; p=0.003) or the CNT group ($c^2_{(1)}$=10.578, p=0.001), but the SCZ and CNT groups did not differ significantly in their rates of use of these medications ($\chi^2_{(1)}$=1.360, p=0.243).

Lastly, the rate of lifetime psychotic symptoms differed significantly by diagnostic group (p<0.001). As a function of our ascertainment scheme and exclusion criteria, the CNT group contained no subjects with a history of psychosis, which was a significant difference from both the SCZ group ($\chi^2_{(1)}$=21.000; p<0.001) and the BPD group ($\chi^2_{(1)}$=8.242; p=0.004). All subjects in the SCZ group had a history of psychosis, whereas only two-thirds of the subjects in the BPD group reported such symptoms ($\chi^2_{(1)}$=5.018; p=0.025).

Microarray Analyses: Alternatively Spliced Genes and Dysregulated Exons

Tables 2–6 provide information on genes that were found to be influenced by a Bonferroni-corrected significant interaction of diagnostic group and exon ID, indicative of differential patterns and/or rates of alternative splicing between the groups. Supplementary Tables 1–5 provide the individual significantly dysregulated exons of these genes which may have contributed to or accounted for the significant alternative splicing event(s) detected in each gene. In each table, the genes are sorted by p-values in ascending order, with the most significantly alternatively spliced gene at the top.

BPD vs. CNT—When comparing the BPD and CNT groups, the expression levels of 18 different genes (out of 21,866 full-length transcripts surveyed) surpassed a stringent, Bonferroni-corrected threshold for significance of the interaction between diagnostic group and exon ID (Table 2). Supplementary Table 1 identifies the 33 individual exons of these 18 genes that were nominally significantly dysregulated (p<0.05) between the two groups; however, due to the large number of comparisons made at the exon level (232,448) and the consequent severity of the adjustment for multiple testing, the dysregulation of no individual exon nor any of those discussed in subsequent comparisons remained significant after Bonferroni correction. Relative to the CNT group, 19 of these exons were significantly down-regulated in the BPD group while the remaining 14 were up-regulated; this proportion of up- and down-regulated exons did not differ from the ratio that might be expected by chance (binomial p=0.095).

A prototypical example of differential splicing between the two groups is illustrated in Fig. (1), which shows each group's pattern of expression of all exons of *PTK2B*, the gene with

the smallest *p*-value for the interaction of diagnostic group (BPD *vs*. CNT) and exon ID. BPD and CNT groups exhibited highly comparable raw levels of expression of most exons of *PTK2B*, but intensities of some exons appeared to diverge, especially in the vicinity of two known splicing sites. When controlling for all covariates, this divergence was statistically significant at exons 2 (p=0.010), 31 (p=0.007), and 33 (p=0.027), while the apparent decrease in expression of exon 29 (a known alternatively spliced exon) in the BPD group was not significant (p=0.299).

Results such as that observed for exons 2, 31, and 33 suggest several possibilities, including: 1) subjects in the BPD group, on average, express less of the known splice variants of *PTK2B* that include this exon (this is supported by the significantly lower mean expression level of this exon); 2) some subjects in the BPD group do not express much or any of the splice variants containing this exon, while other individuals in the BPD group express such splice variants at normal levels (this is supported by the larger-than-average standard error of expression of this exon); 3) the groups do not differ in their rates or levels of expression of different splice variants, but this exon somehow is selectively inhibited in its expression; 4) the BPD group expresses less of some novel, unrecognized splice variant(s) that include this exon; or, conceivably, 5) a type-I error has occurred. Against the last option, the *q*-values for *PTK2B* and the other 17 genes listed in Table 2 were found to range from $1.26e^{-11}$ to $5.74e^{-06}$, suggesting that these 18 results have a very low probability of representing false-discoveries. The genes on this list did not represent an enrichment of any particular ontology, pathway, or protein domain.

SCZ vs. CNT—Only eight genes surpassed the Bonferroni-corrected threshold for significance of the diagnostic group-by-exon ID interaction when comparing SCZ and CNT groups (Table 3). The expected FDRs for these genes ranged from $4.18e^{-10}$ to $7.57e^{-06}$. The 20 individual significantly dysregulated exons of these eight genes are presented in Supplementary Table 2. In contrast to the comparison of BPD and CNT groups, in which the slight majority (19/33) of significantly dysregulated exons were down-regulated in BPD, almost all (18/20) of the exons that were significantly dysregulated in the SCZ group were down-regulated. The chance of observing this ratio of down-regulated to up-regulated genes was extremely low (binomial p=$1.81e^{-4}$).

Notably, despite the small number of genes emerging as significant in this comparison, several ontologies were significantly over-represented by these eight genes, including the biological processes “phosphorylation” (p=0.025), “phosphorus metabolism” (p=0.036), “phosphate metabolism” (p=0.036), and “cellular metabolism” (p=0.044), as well as the molecular function “catalytic activity” (p=0.047).

BPD vs. SCZ—The comparison of BPD and SCZ groups revealed 33 genes that were influenced by a Bonferroni-adjusted significant interaction of diagnostic group and exon ID (Table 4). Expected FDRs for these 33 genes were between $9.28e^{-12}$ and $5.31e^{-06}$. Consistent with the observation of more frequent down-regulation of alternatively spliced exons in the SCZ group than in the BPD (each relative to the CNT group), 75/103 significantly dysregulated exons in this comparison (Supplementary Table 3) were down-regulated in the BPD group relative to the SCZ group; this ratio of down-regulated to up-regulated genes was significantly different from chance expectation (binomial p=$4.00e^{-6}$). As with the comparison of BPD and CNT groups described above, no ontologies, pathways, or protein domains were significantly over-represented in this gene list.

BPD+SCZ vs. CNT—When the BPD and SCZ groups were combined and jointly contrasted with the CNT group, 15 genes were found to exhibit significantly different patterns of alternative splicing between the groups after correcting for multiple comparisons

using the Bonferroni method (Table 5). As with the prior comparisons of diagnostic groups, the expected FDRs among these 15 genes were exceedingly low, ranging from $9.07e^{-10}$ to $5.21e^{-06}$. Of the 24 exons that were significantly differentially expressed between groups in this comparison (Supplementary Table 4), 15 were up-regulated in the combined BPD+SCZ group relative to the CNT group while the remaining nine exons were significantly down-regulated; however, this ratio did not deviate from chance expectation (p=0.078). The gene list for this comparison was most significantly enriched with genes representing biological processes such as "phosphorus metabolism" (p=0.004) and "phosphate metabolism" (p=0.004), but numerous other broad ontologies were over-represented as well, including the biological processes "development" (p=0.007), "phosphorylation" (p=0.018), "protein modification" (p=0.029), and "metabolism" (p=0.021), and the molecular functions "protein binding" (p=0.020) and "nucleotide binding" (p=0.050).

PSYCH(+) vs. PSYCH(−)—By far, the greatest disparity in patterns of expression of alternatively spliced genes was seen when comparing groups with a different history of psychosis. A total of 156 genes surpassed the Bonferroni-adjusted significance threshold for the interaction of diagnostic group with exon ID when comparing the PSYCH(+) and PSYCH(−) groups (Table 6). These 156 genes had associated FDRs ranging from $8.21e^{-27}$ to $5.28e^{-06}$. Of the 16,555 nominally significantly dysregulated exons of these 156 genes (Supplementary Table 5), only 64 were down-regulated in the PSYCH(+) group relative to the PSYCH(−) group, while the remaining 16,491 exons were up-regulated; this represented a highly significant departure from the ratio expected by chance ($p<1.00e^{-10}$).

Many ontologies, pathways, and protein domains were significantly over-represented by alternatively spliced genes in the comparison of PSYCH(+) and PSYCH(−) groups (Table 7). Each of the five ontological terms significantly over-represented by alternatively spliced genes in the comparison of SCZ *vs*. CNT groups were also significantly over-represented in the comparison of PSYCH(+) *vs*. PSYCH(−) groups; these included "phosphorylation" (p=0.042), "phosphorus metabolism" (p=0.015), "phosphate metabolism" (p=0.015), "cellular metabolism" (p=0.021), and "catalytic activity" (p=0.009). Of note, genes in the *NOTCH* signaling pathway were also present in this list at a significantly higher rate than would be expected by chance (p=0.015), as were genes that bind to (p=0.005) or are expressed in (p=0.035) the cytoskeleton, genes involved in the ubiquitin cycle (p=0.047) or having ubiquitin-protein ligase activity (p=0.009), and genes that activate ($p=7.50e^{-04}$) or regulate ($p=5.20e^{-05}$) GTPase activity.

DISCUSSION

Transcriptomic biomarkers of psychiatric diseases obtained from a query of peripheral tissues that are clinically accessible (e.g., blood cells instead of post-mortem brain tissue) have substantial practical appeal to discern the molecular subtypes of common complex diseases such as major psychosis. To our knowledge, this is the first blood-based spliceome-profiling study of schizophrenia and bipolar disorder to be reported. The present study has relevance both as an original exploratory study as well as a new methodological approach in the study of genetic factors that may contribute to disease susceptibility. Beyond these qualities, the study suggests other potential applications of the methodology, such as the study of dynamic host responses to environmental factors/perturbations such as drug treatment or other exposures.

The chief result from this pilot study is the demonstration that exomic and spliceomic profiling can identify transcripts that reliably differentiate groups of psychiatric patients from each other and from non-mentally ill individuals. Second, this pilot study yielded information on numerous specific exons, alternatively spliced genes, and functionally or

structurally related groups of those genes that are expressed in varying amounts in the peripheral blood of patients with schizophrenia, bipolar disorder, or psychosis. Third, from a methodological standpoint, a focus on spliceome-profiling may offer a renewed interest and deeper insights on peripheral tissue markers of central nervous system diseases, an area of scientific inquiry that has thus far lacked from a paucity of transcriptomics-based biomarkers in tissues that are clinically accessible in a form that is also acceptable to patients. Seen in this light, blood-based biomarkers deserve further exploration particularly by taking into account human variation in the spliceome. However, we underscore that the findings presented in this report are exploratory in nature and intended to serve as a baseline inquiry for spliceome-based biomarkers. These results will undoubtedly require replication and the test of triangulation by other independent biomarker technology platforms in different tissues before they can be considered solid leads in the pursuit of biomarkers for these disorders.

With these caveats in mind, several conclusions from this discovery-oriented exploratory study can be distilled. First, a relatively small number of genes display differential patterns of expression of alternative splice variants between diagnostic groups that are reliable after applying stringent corrections for multiple testing. The comparisons producing the fewest such differences were between the CNT group and either of the other ascertained diagnostic groups (SCZ and BPD), which is notable given the vast differences between the groups not just in diagnoses but in variables associated with the disorders (*e.g.*, comorbid conditions, pharmacologic treatment) and potential confounding variables (*e.g.*, smoking). In contrast, a fair number of genes showed differential patterns of expression of alternatively spliced variants between BPD and SCZ groups, which is considerable given their much higher degree of comparability on clinical and potential confounding variables. The largest number of significant differences in splicing patterns was observed when groups were contrasted on their history of psychosis without regard to diagnostic boundaries. Disentangling the source(s) of this enhancement is difficult. For some genes, it may be that the larger sample size of the PSYCH(+) and PSYCH(−) groups relative to any two individual ascertained diagnostic groups allowed some small effects to achieve statistical significance in the former contrast where they otherwise might not in the latter. Alternatively, the phenotype of psychosis simply may be more strongly linked than any particular diagnostic entity to the expression of alternatively spliced genes in PBMCs. Of course, both situations also may be operating concurrently.

A second conclusion can be drawn based on our analyses of biological pathways, ontologies, and protein domains represented by the alternatively spliced genes. Many of the structural or functional categories over-represented by the alternatively spliced genes were very broad in definition (*e.g.*, cellular metabolism or catalytic activity), but some were quite specific and potentially informative. For example, the SCZ and combined BPD+SCZ groups (relative to the CNT group) showed differential splicing of genes linked to phosphorylation, phosphorus metabolism, and phosphate metabolism. Relative to the PSYCH(−) group, the PSYCH (+) group also showed an over-abundance of alternatively spliced genes representing these processes. In addition, the PSYCH(+) group expressed more differentially spliced variants of genes linked to very specific biological processes such as GTPase activity, as well as *Notch* and ubiquitin pathways, each of which independently has been linked to schizophrenia and/or psychosis previously [20,56–59]. Thus, we conclude that the genes that show differential expression of alternatively spliced variants in schizophrenia or psychosis are not randomly distributed, but aggregate in mechanistically meaningful pathways, some of which are supported by prior work and some of which can foster new hypotheses.

A final (and perhaps most striking) general conclusion to be drawn from this study is that psychosis appears to be marked by a global up-regulation of exons in transcripts expressed

in PBMCs. Of the 16,555 exons (in 156 genes) that were significantly differentially expressed in the PSYCH (+) group, 99.6% were up-regulated while just 0.4% were down-regulated. The strength of this result suggests that a systematic process (*e.g.*, down-regulation of splicing factors) may be operating in psychosis, which leads to widespread over-expression of selected exons or more frequent inclusion of those exons in expressed transcripts. Aside from stimulating work to unravel the biological basis for the phenomenon, this finding may encourage the advancement of *unidirectional* hypotheses in subsequent analyses of splicing patterns in psychosis, which in turn would warrant the use of one-tailed statistical tests and the conservation of inferential power.

Beyond these conclusions, this exploratory study can serve as a useful comparator for future spliceome-wide or candidate-gene analyses of alternative splicing, as well as a replication study for prior reports of alternative splicing in these disorders. Significant changes in the expression of alternatively spliced transcripts in blood or brain tissue samples from schizophrenia patients have been reported for a number of genes, including *CAMK2A* [60], *CTNNA2* [61], *DRD3* [62], *ERBB4* [36,38], *GRIN1* (*NMDARI*/*NR1*) [39], *GRM3* [63], and *QKI* [64]. Further supporting the conceptualization of mRNA expression as perhaps the most basic endophenotype, dysregulated expression of exons in some of these alternatively spliced variants has been shown to be strongly governed by specific *cis*-acting polymorphisms in the same genes [36,38,41,63]. Conversely, few significant findings have been reported in bipolar disorder; in fact, we could only find one study examining the issue, which reported a non-significant difference in the expression of ISYNA1 splice variants in blood from bipolar disorder patients relative to control subjects [65]. We too failed to find a significant interaction of diagnostic group (BPD *vs*. CNT) and exon ID for *ISYNA1* (p=0.107), but we also failed to find evidence supporting the differential expression of splice variants of any of the genes previously found to be alternatively spliced in schizophrenia as well. Since none of these genes exhibited a Bonferroni-corrected (or even nominally) significant interaction of diagnostic group and exon ID, we did not interrogate specific exons of these genes in our primary analyses reported above. However, *post-hoc* inspection of the data did reveal nominally significant evidence for dysregulation of one exon in each of *CTNNA2* (p=0.030), *DRD3* (p=0.005), *ERBB4* (p=0.040), and *QKI* (p=0.035).

We caution the reader that several caveats must be considered in comparing our results to those described above. First, in many instances different tissues are being compared (PBMCs and postmortem brain). Although expression of many full-length gene transcripts is known to be reasonably well correlated [Spearman's rho(ρ)=0.73] between blood and brain tissue [25], it is not known if this correspondence extends to individual exons or all splice variants of each gene. In addition, while the overall correlation between blood and brain gene expression is reasonably high, the correlation varies widely for individual genes, with some showing near perfect correspondence and others showing little or no correlation. Although this has yet to be demonstrated across the whole human spliceome, similar results may be expected, and thus it is possible that the discrepancies between prior studies and our own are driven by tissue-specificity of splicing and expression of the evaluated genes. Second, our results for these genes were derived by microar-ray whereas the original discoveries were typically made by quantitative reverse-transcription PCR; thus, it is possible that our failure to provide independent replication of these discoveries was due to a lack of sensitivity of our chosen platform. Third, and of particular relevance to the results of Sartorius *et al*. [63], Silberberg *et al*. [38], Atz *et al*. [41] and Law *et al*. [36], is the fact that we did not genotype the DNA polymorphisms that these investigators showed to regulate gene expression and splicing; thus, it is possible that we may have replicated their findings if we controlled for the governing DNA polymorphisms these authors identified. Fourth, our sample size was relatively small, which may have prohibited us from detecting effects that may have attained statistical significance in a larger sample, and this problem was likely

compounded by the stringency of the corrections for multiple testing that we applied. Fifth, we emphasize that many factors (and combinations of factors) would be expected to influence gene expression or alternate splicing levels in studies of chronic illnesses where patient samples often reflect cumulative exposures from drug treatment, diet, smoking, and co-morbid disease states in addition to those conferred by diagnostic categories. In particular, there was a significant disparity in smoking status between our various diagnostic groups, with no present smokers in the CNT group compared to greater than 50% of present smokers in both the BPD and SCZ groups. If smoking status significantly influences the rates and types of splicing events that occur in peripheral blood cells, then it is possible that some of the observed differences we ascribed to diagnoses may actually be type-I errors, and instead should be attributed solely to the diagnostic-group differences in smoking rates. The present study was limited in our ability to model all potentially influential covariates by constraints on our sample size. On the other hand, even with a large study sample, confounding variables such as differences in comedication use between diagnostic groups and healthy controls are difficult to avoid. Hence, future spliceome-profiling studies in drug-naive (and non-smoking) patients or individuals who present with a first episode of psychosis could conceivably provide invaluable additional insights to further extend the observations from the present exploratory investigation. Lastly, we note however that our findings in the present study were reconciled with an analysis of the ontologies, the biological pathways and the protein domains significantly over-represented among the alternatively spliced genes, several of which support prior discoveries.

Other alternatively spliced genes identified in our study do not necessarily replicate prior findings of altered expression of splice variants of candidate genes, but in another manner provide support for involvement of these genes in the disorder. For example, we found significantly differential expression of alternatively constituted *CHI3L1* transcripts in the SCZ and CNT groups, which was accounted for by decreased expression of 11 different exons in the SCZ group. This result is highly consistent with results from Zhao *et al*., who first discovered that schizophrenia-associated risk haplotypes of *CHI3L1* were associated with lower trans- criptional activity and lower expression of the gene [66]; however, the effect of the implicated polymorphisms of this gene on the expression of particular splice variants remains to be determined. Polymorphisms in *BRD1* [67], *IL3* [68,69], *JARID2* [70], *MAPK14* [71], and *SNAP29* [72,73] have previously been associated with risk for schizophrenia and also appear to exhibit alternative splicing in the SCZ group of the present study relative to either the CNT or BPD groups. Again, whether these genetic associations and patterns of alternative splicing reflect a common mechanistic link to the disorder is an empirical question to be addressed subsequently.

In addition to pursuing such unanswered questions, future work should proceed along several other trajectories. First, based on sequence analysis of the 230 genes listed in Tables 2–6, we have determined that 216 of these genes harbor DNA polymorphisms in putative exonic splicing enhancers, 151 harbor polymorphisms in putative exonic splicing silencers, 36 harbor polymorphisms in canonical splice sites, and 1 gene harbors a polymorphism in a predicted novel splice site at an intron/exon boundary. These polymorphisms are prime candidates for genotyping and association analysis with the disorders under study as well as the expression levels of the alternatively spliced transcripts that these polymorphisms produce. Second, previously completed transcriptomic studies of schizophrenia and bipolar disorder may warrant re-examination in light of the patterns of differential alternate splicing we observed. Failure to replicate the specific (*i.e*., gene-level) results of transcriptomic studies of these disorders is not uncommon. Frequently cited causes of such discrepancies include tech- nical and methodological variables, or the potential that networks of genes (rather than specific transcripts) are more likely to generalize across subjects and studies [50]. These are indeed potentially valid explanations for the phenomenon; however, it is also

possible that variable patterns of expression of splice variants (or differing prevalences of influential polymorphisms [74]) across samples have given the appearance of incomparable full-length gene expression, particularly when coupled with the 3′ bias inherent in previous generations of expression microarrays. Third, further control of potentially influential covariates (*e.g.*, diet and exercise, dosage and duration of medication usage, length of illness, *etc*.) should be attempted; however, this will require much larger sample sizes so that the effects of these many factors (and their interactions) can be simultaneously modeled. Lastly, our results should be verified using more sensitive mRNA quantification methods, and verified independently in either spliceome-wide or targeted replication efforts, after which some of the identified candidate exons and alternatively spliced genes may be validated as useful biomarkers for these conditions.

An important lesson learned from prior efforts to develop biomarkers for psychiatric disorders is that no single tissue, molecule, or marker is likely to yield sufficient power for improving existing behavioral classification schemes; rather, integration of the best markers from multiple domains (*e.g.*, DNA, brain mRNA, blood mRNA, protein, *etc*.) may lead to the most reliable profiles [22], which then can be used to begin appropriate, group-tailored interventions. Beyond this working model, it also may be necessary to introduce an iterative component to deal with the abundant phenotypic heterogeneity of the major psychoses, which potentially may map onto heterogeneity in etiologic and biomarker profiles. Thus, as broadly influential etiologic factors and their associated biomarkers are identified, segments of the larger schizophrenia and bipolar disorder phenotypes may be “carved out”, some of which [*e.g.*, PSYCH(+)] are influenced by that factor and some of which [*e.g.*, PSYCH(−)] are not. The subsequent detection of both environmental and biological influences on-and biomarkers of-such subgroups will be facilitated by their relative homogeneity; however, additional phenotypic “cleavage points” should be anticipated until groups with a highly similar etiologic and biomarker profile are obtained.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by the UCSD Center for AIDS Research Genomics Core, and National Institutes of Health grants P30MH062512 (Igor Grant), R01MH079881, R25MH074508, R25MH081482, and R41MH079728 (I.P.E.), R01AG018386, R01AG022381, and R01AG022982 (W.S.K.), R01DA012846, R01DA018662, R01MH065562, R01MH 071912, and R21MH075027 (M.T.T.), P50MH081755 (Eric Courchesne/S.J.G.), and a NARSAD Young Investigator Award (S.J.G.).

LIST OF ABBREVIATIONS

ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
BPD	Bipolar disorder
CDS	Coding sequence
CNT	Control
DIGS	Diagnostic Interview for Genetic Studies
FDR	False-discovery rate
FWER	Family-wise error rate

ID	Identity
mRNA	Messenger ribonucleic acid
PBMC	Peripheral blood mononuclear cell
PRICE	Psychopharmacology Research Initiatives Center for Excellence
PSYCH	Psychosis
QRT-PCR	Quantitative reverse-transcription PCR
RIN	RNA Integrity Number
RMA	Robust multi-array average
SCZ	Schizophrenia
UCSD	University of California, San Diego
WT	Whole transcript

References

1. Saha S, Chant D, McGrath J. Meta-analyses of the incidence and prevalence of schizophrenia: conceptual and methodological issues. Int J Methods Psychiatr Res. 2008; 17(1):55–61. [PubMed: 18286464]
2. Merikangas KR, Akiskal HS, Angst J, et al. Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbid-ity Survey replication. Arch Gen Psychiatry. 2007; 64(5):543–52. [PubMed: 17485606]
3. Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. Arch Gen Psychiatry. 2003; 60(12):1187–92. [PubMed: 14662550]
4. Cardno AG, Marshall EJ, Coid B, et al. Heritability estimates for psychotic disorders: the Maudsley twin psychosis series. Arch Gen Psychiatry. 1999; 56(2):162–8. [PubMed: 10025441]
5. Kato T. Molecular genetics of bipolar disorder and depression. Psychiatry Clin Neurosci. 2007; 61(1):3–19. [PubMed: 17239033]
6. Allen NC, Bagade S, McQueen MB, et al. Systematic meta-analyses and field synopsis of genetic association studies in schizophrenia: the SzGene database. Nat Genet. 2008; 40(7):827–34. [PubMed: 18583979]
7. Smoller JW, Gardner-Schuster E. Genetics of bipolar disorder. Curr Psychiatry Rep. 2007; 9(6): 504–11. [PubMed: 18221632]
8. Sklar P, Smoller JW, Fan J, et al. Whole-genome association study of bipolar disorder. Mol Psychiatry. 2008; 13(6):558–69. [PubMed: 18317468]
9. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature. 2007; 447(7145):661–78. [PubMed: 17554300]
10. Baum AE, Akula N, Cabanero M, et al. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. Mol Psychiatry. 2008; 13(2):197–207. [PubMed: 17486107]
11. Lencz T, Morgan TV, Athanasiou M, et al. Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. Mol Psychiatry. 2007; 12(6):572–80. [PubMed: 17522711]
12. Sullivan PF, Lin D, Tzeng JY, et al. Genomewide association for schizophrenia in the CATIE study: results of stage 1. Mol Psychiatry. 2008; 13(6):570–84. [PubMed: 18347602]
13. Shifman S, Johannesson M, Bronstein M, et al. Genome-wide association identifies a common variant in the reelin gene that increases the risk of schizophrenia only in women. PLoS Genet. 2008; 4(2):e28. [PubMed: 18282107]
14. Kirov G, Zaharieva I, Georgieva L, et al. A genome-wide association study in 574 schizophrenia trios using DNA pooling. Mol Psychiatry. 2008 [Epub ahead of print].

15. Chen C, Glatt SJ, Tsuang MT. The tryptophan hydroxylase gene influences risk for bipolar disorder but not major depressive disorder: results of meta-analyses. Bipolar Disord. in press.
16. Lasky-Su JA, Faraone SV, Glatt SJ, et al. Meta-analysis of the association between two polymorphisms in the serotonin transporter gene and affective disorders. Am J Med Genet B Neuropsy-chiatr Genet. 2005; 133B(1):110–5.
17. Glatt SJ, Faraone SV, Tsuang MT. Association between a functional catechol O-methyltransferase gene polymorphism and schizophrenia: meta-analysis of case-control and family-based studies. Am J Psychiatry. 2003; 160(3):469–76. [PubMed: 12611827]
18. Glatt SJ, Faraone SV, Tsuang MT. Meta-analysis identifies an association between the dopamine D2 receptor gene and schizophrenia. Mol Psychiatry. 2003; 8(11):911–5. [PubMed: 14593428]
19. Glatt SJ, Jonsson EG. The Cys allele of the DRD2 Ser311Cys polymorphism has a dominant effect on risk for schizophrenia: evidence from fixed- and random-effects meta-analyses. Am J Med Genet B Neuropsychiatr Genet. 2006; 141(2):149–54. [PubMed: 16402354]
20. Glatt SJ, Wang RS, Yeh YC, et al. Five NOTCH4 polymorphisms show weak evidence for association with schizophrenia: evidence from meta-analyses. Schizophr Res. 2005; 73(2–3):281–90. [PubMed: 15653273]
21. Baum AE, Hamshere M, Green E, et al. Meta-analysis of two genome-wide association studies of bipolar disorder reveals important points of agreement. Mol Psychiatry. 2008; 13(5):466–7. [PubMed: 18421293]
22. Ozdemir V, Williams-Jones B, Glatt SJ, et al. Shifting emphasis from pharmacogenomics to theragnostics. Nat Biotechnol. 2006; 24(8):942–6. [PubMed: 16900136]
23. Duffus JH, Nordberg M, Templeton DM. IUPAC Glossary of Terms Used in Toxicology, 2nd Edition. Pure Applied Chemistry. 2007; 79(7):1153–344.
24. Vawter MP, Ferran E, Galke B, et al. Microarray screening of lymphocyte gene expression differences in a multiplex schizophrenia pedigree. Schizophr Res. 2004; 67(1):41–52. [PubMed: 14741323]
25. Sullivan PF, Fan C, Perou CM. Evaluating the comparability of gene expression in blood and brain. Am J Med Genet B Neuropsy-chiatr Genet. 2006; 141B(3):261–8.
26. Zvara A, Szekeres G, Janka Z, et al. Over-expression of dopamine D2 receptor and inwardly rectifying potassium channel genes in drug-naive schizophrenic peripheral blood lymphocytes as potential diagnostic markers. Dis Markers. 2005; 21(2):61–9. [PubMed: 15920292]
27. Middleton FA, Pato CN, Gentile KL, et al. 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 B Neuropsychiatr Genet. 2005; 136(1):12–25. [PubMed: 15892139]
28. Bowden NA, Weidenhofer J, Scott RJ, et al. Preliminary investigation of gene expression profiles in peripheral blood lymphocytes in schizophrenia. Schizophr Res. 2006; 82(2–3):175–83. [PubMed: 16414245]
29. Tsuang MT, Nossova N, Yager T, et al. Assessing the validity of blood-based gene expression profiles for the classification of schizophrenia and bipolar disorder: a preliminary report. Am J Med Genet B Neuropsychiatr Genet. 2005; 133(1):1–5. [PubMed: 15645418]
30. Glatt SJ, Everall IP, Kremen WS, et al. Comparative gene expression analysis of blood and brain provides concurrent validation of SELENBP1 up-regulation in schizophrenia. Proc Natl Acad Sci USA. 2005; 102(43):15533–8. [PubMed: 16223876]
31. Kanazawa T, Chana G, Glatt SJ, et al. The utility of SELENBP1 gene expression as a biomarker for major psychotic disorders: Replication in schizophrenia and extension to bipolar disorder with psychosis. Am J Med Genet B Neuropsychiatr Genet. in press.
32. Owen MJ, Craddock N, Jablensky A. The genetic deconstruction of psychosis. Schizophr Bull. 2007; 33(4):905–11. [PubMed: 17551090]
33. Potash JB. Carving chaos: genetics and the classification of mood and psychotic syndromes. Harv Rev Psychiatry. 2006; 14(2):47–63. [PubMed: 16603472]
34. Badner JA, Gershon ES. Meta-analysis of whole-genome linkage scans of bipolar disorder and schizophrenia. Mol Psychiatry. 2002; 7(4):405–11. [PubMed: 11986984]

35. Berrettini W. Bipolar disorder and schizophrenia: convergent molecular data. Neuromolecular Med. 2004; 5(1):109–17. [PubMed: 15001816]
36. Law AJ, Kleinman JE, Weinberger DR, et al. Disease-associated intronic variants in the ErbB4 gene are related to altered ErbB4 splice-variant expression in the brain in schizophrenia. Hum Mol Genet. 2007; 16(2):129–41. [PubMed: 17164265]
37. Zhao C, Xu Z, Chen J, et al. Two isoforms of GABA(A) receptor beta2 subunit with different electrophysiological properties: Differential expression and genotypical correlations in schizophrenia. Mol Psychiatry. 2006; 11(12):1092–105. [PubMed: 16983389]
38. Silberberg G, Darvasi A, Pinkas-Kramarski R, et al. The involvement of ErbB4 with schizophrenia: association and expression studies. Am J Med Genet B Neuropsychiatr Genet. 2006; 141B(2):142–8. [PubMed: 16402353]
39. Le Corre S, Harper CG, Lopez P, et al. Increased levels of expression of an NMDARI splice variant in the superior temporal gyrus in schizophrenia. Neuroreport. 2000; 11(5):983–6. [PubMed: 10790868]
40. Huntsman MM, Tran BV, Potkin SG, et al. Altered ratios of alternatively spliced long and short gamma2 subunit mRNAs of the gamma-amino butyrate type A receptor in prefrontal cortex of schizophrenics. Proc Natl Acad Sci USA. 1998; 95(25):15066–71. [PubMed: 9844016]
41. Atz ME, Rollins B, Vawter MP. NCAM1 association study of bipolar disorder and schizophrenia: polymorphisms and alternatively spliced isoforms lead to similarities and differences. Psychiatr Genet. 2007; 17(2):55–67. [PubMed: 17413444]
42. Nurnberger JI Jr, Blehar MC, Kaufmann CA, et al. Diagnostic interview for genetic studies. Rationale, unique features, and training. Arch Gen Psychiatry. 1994; 51:849–59. [PubMed: 7944874]
43. Gonzales J, Kemppainen J, Latham G, et al. Isolate RNA from white blood cells captured by a novel filter system. Ambion Tech-Notes. 2005; 12(4):24–5.
44. Schroeder A, Mueller O, Stocker S, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol. 2006; 7:3. [PubMed: 16448564]
45. Affymetrix I. GeneChip® Whole Transcript (WT) Sense Target Labeling Assay Manual, Version 4. P/N 701880. 2006 [Accessed April 14, 2009]. Available from : http://www.affymetrix.com/support/downloads/manuals/wt_sensetarget_label_manual.pdf
46. Affymetrix I. Exon and Gene Array Glossary. 2008 [Accessed April 14, 2009]. Available from http://www.affymetrix.com/support/help/exon_glossary/index.affx
47. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics. 2003; 4(2):249–64. [PubMed: 12925520]
48. Handran, S.; Pickett, S.; Verdick, D. Key Considerations for Accurate Microarray Scanning and Image Analysis. In: Kamberova, G.; Shah, S., editors. DNA Array Image Analysis: Nuts & Bolts. DNA Press, LLC; Salem, MA: 2002. p. 83-98.
49. Partek® software On-line Help, version 6.3. Copyright © 2008 Partek Inc., St. Louis, MO, USA. [Accessed April 14, 2009].
50. Mirnics K, Levitt P, Lewis DA. Critical appraisal of DNA microar-rays in psychiatric genomics. Biol Psychiatry. 2006; 60(2):163–76. [PubMed: 16616896]
51. Dennis G Jr, Sherman BT, Hosack DA, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003; 4(5):p3. [PubMed: 12734009]
52. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000; 28(1):27–30. [PubMed: 10592173]
53. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000; 25(1):25–9. [PubMed: 10802651]
54. Mulder NJ, Apweiler R, Attwood TK, et al. InterPro, progress and status in 2005. Nucleic Acids Res. 2005; 33(Database issue):D201–5. [PubMed: 15608177]
55. Wu CH, Yeh LS, Huang H, et al. The Protein Information Resource. Nucleic Acids Res. 2003; 31(1):345–7. [PubMed: 12520019]
56. Wang Z, Wei J, Zhang X, et al. A review and re-evaluation of an association between the NOTCH4 locus and schizophrenia. Am J Med Genet B Neuropsychiatr Genet. 2006; 141B(8): 902–6. [PubMed: 16894623]

57. Thomas EA. Molecular profiling of antipsychotic drug function: convergent mechanisms in the pathology and treatment of psychiatric disorders. Mol Neurobiol. 2006; 34(2):109–28. [PubMed: 17220533]
58. Hashimoto R, Yoshida M, Kunugi H, et al. A missense polymorphism (H204R) of a Rho GTPase-activating protein, the chimerin 2 gene, is associated with schizophrenia in men. Schizophr Res. 2005; 73(2–3):383–5. [PubMed: 15653288]
59. Pickard BS, Malloy MP, Clark L, et al. Candidate psychiatric illness genes identified in patients with pericentric inversions of chromosome 18. Psychiatr Genet. 2005; 15(1):37–44. [PubMed: 15722956]
60. Novak G, Seeman P, Tallerico T. Increased expression of calcium/calmodulin-dependent protein kinase IIbeta in frontal cortex in schizophrenia and depression. Synapse. 2006; 59(1):61–8. [PubMed: 16247765]
61. Mexal S, Berger R, Pearce L, et al. Regulation of a novel alphaN-catenin splice variant in schizophrenic smokers. Am J Med Genet B Neuropsychiatr Genet. 2007
62. Schmauss C. Enhanced cleavage of an atypical intron of dopamine D3-receptor pre-mRNA in chronic schizophrenia. J Neurosci. 1996; 16(24):7902–9. [PubMed: 8987818]
63. Sartorius LJ, Weinberger DR, Hyde TM, et al. Expression of a GRM3 splice variant is increased in the dorsolateral prefrontal cortex of individuals carrying a schizophrenia risk SNP. Neuropsychopharmacology. 2008
64. Aberg K, Saetre P, Jareborg N, et al. Human QKI, a potential regulator of mRNA expression of human oligodendrocyte-related genes involved in schizophrenia. Proc Natl Acad Sci USA. 2006; 103(19):7482–7. [PubMed: 16641098]
65. Shamir A, Shaltiel G, Mark S, et al. Human MIP synthase splice variants in bipolar disorder. Bipolar Disord. 2007; 9(7):766–71. [PubMed: 17988368]
66. Zhao X, Tang R, Gao B, et al. Functional variants in the promoter region of Chitinase 3-like 1 (CHI3L1) and susceptibility to schizophrenia. Am J Hum Genet. 2007; 80(1):12–8. [PubMed: 17160890]
67. Severinsen JE, Bjarkam CR, Kiaer-Larsen S, et al. Evidence implicating BRD1 with brain development and susceptibility to both schizophrenia and bipolar affective disorder. Mol Psychiatry. 2006; 11(12):1126–38. [PubMed: 16924267]
68. Sun S, Wang F, Wei J, et al. Association between interleukin-3 receptor alpha polymorphism and schizophrenia in the Chinese population. Neurosci Lett. 2008; 440(1):35–7. [PubMed: 18547720]
69. Chen X, Wang X, Hossain S, et al. Interleukin 3 and schizophrenia: the impact of sex and family history. Mol Psychiatry. 2007; 12(3):273–82. [PubMed: 17179997]
70. Pedrosa E, Ye K, Nolan KA, et al. Positive association of schizophrenia to JARID2 gene. Am J Med Genet B Neuropsychiatr Genet. 2007; 144B(1):45–51. [PubMed: 16967465]
71. Olsen L, Hansen T, Jakobsen KD, et al. The estrogen hypothesis of schizophrenia implicates glucose metabolism: association study in three independent samples. BMC Med Genet. 2008; 9:39. [PubMed: 18460190]
72. Saito T, Guan F, Papolos DF, et al. Polymorphism in SNAP29 gene promoter region associated with schizophrenia. Mol Psychiatry. 2001; 6(2):193–201. [PubMed: 11317222]
73. Wonodi I, Hong LE, Avila MT, et al. Association between polymorphism of the SNAP29 gene promoter region and schizophrenia. Schizophr Res. 2005; 78(2–3):339–41. [PubMed: 15908182]
74. Benovoy D, Kwan T, Majewski J. Effect of polymorphisms within probe-target sequences on olignonucleotide microarray experiments. Nucleic Acids Res. 2008; 36(13):4417–23. [PubMed: 18596082]

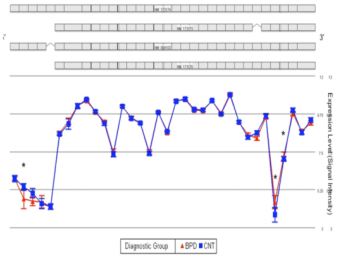

Fig. (1). Alternative Splicing of PTK2B: BPD vs. CNT

The figure illustrates a prototypical example of differential expression of alternatively spliced gene variants between two groups; in this case, BPD and CNT groups. The illustrated gene, *PTK2B*, produced the smallest *p*-value ($8.19e^{-08}$) for the interaction of diagnostic group (BPD *vs*. CNT) and exon ID. The top panel shows the four known splice variants of the gene, while the bottom panel plots the raw microarray expression levels (signal intensities, unadjusted for covariates) of individual exons in the BPD and CNT groups (red triangles and blue squares, respectively). The lines representing signal intensity in each diagnostic group closely correspond to each other over most of the length of the gene, but intensities of some exons appeared to diverge, especially in the area of two known splicing sites. When controlling for all covariates, this divergence was statistically significant (*) at exons 2 (p=0.010), 31 (p=0.007), and 33(p=0.027), while the apparent decrease in expression of exon 29 (a known alternatively spliced exon) in the BPD group was not significant (p=0.299).

Table 1

Descriptive Statistics by Diagnostic Group

	Diagnostic Group			Omnibus	
Variable	**BPD (n=9)**	**SCZ (n=13)**	**CNT (n=8)**	**Test Statistic**	***p***
Age: mean years (s.d.)	42 (8)	44 (9)	45 (7)	$F_{(2,27)}$=0.240	0.788
Sex: female n (%)	2 (22)	4 (31)	3 (38)	$\chi^2_{(2)}$=0.477	0.788
Ancestry: n (%)				$\chi^2_{(6)}$=7.654	0.264
European	7 (78)	5 (39)	5 (63)		
African	1 (11)	6 (46)	1 (12)		
Hispanic	0 (0)	2 (15)	1 (12)		
Asian	1 (11)	0 (0)	1 (12)		
Current Smoker: n (%)	5 (56)	10 (77)	0 (0)	$\chi^2_{(2)}$=11.880	0.003
Current Medication Use: n (%)					
Antipsychotic	5 (56)	12 (92)	0 (0)	$\chi^2_{(2)}$=17.191	<0.001
Mood Stabilizer	7 (78)	2 (15)	0 (0)	$\chi^2_{(2)}$=14.534	0.001
History of Psychosis: n (%)	6 (67)	13 (100)	0 (0)	$\chi^2_{(2)}$=21.388	<0.001

Table 2

Alternatively Spliced Genes: Bipolar Disorder (BPD) *vs*. Control (CNT)

Gene Symbol	Gene Product	Chromosomal Locus	Accession Number	Transcript ID	Probe Sets (*n*)	Diagnostic Group-x-Exon ID Interaction			
						F	*p*	Adjusted *p*	*q*
PTK2B	protein tyrosine kinase 2 beta	8p21.1	NM_173174	3091301	34	4.396	$3.75e^{-12}$	$8.19e^{-08}$	$1.26e^{-11}$
NCF4	neutrophil cytosolic factor 4, 40kDa	22q13.1	NM_013416	3944543	11	10.114	$9.07e^{-11}$	$1.98e^{-06}$	$3.05e^{-10}$
UBR5	ubiquitin protein ligase E3 component n-recognin 5	8q22	NM_015902	3147321	64	2.822	$1.50e^{-10}$	$3.26e^{-06}$	$5.03e^{-10}$
COG4	component of oligomeric golgi complex 4	16q22.1	NM_015386	3697125	24	4.894	$2.60e^{-10}$	$5.67e^{-06}$	$8.75e^{-10}$
ROD1	regulator of differentiation 1 (*S. pombe*)	9q32	NM_005156	3220977	18	5.428	$4.18e^{-09}$	$9.13e^{-05}$	$1.41e^{-08}$
C8orf41	chromosome 8 open reading frame 41	8p12	BC066935	3130823	12	7.534	$5.58e^{-09}$	$1.22e^{-04}$	$1.88e^{-08}$
ACTR10	actin-related protein 10 homolog (*S. cerevisiae*)	14q23	NM_018477	3537813	17	5.429	$1.11e^{-08}$	$2.43e^{-04}$	$3.74e^{-08}$
C18orf10	chromosome 18 open reading frame 10	18q12.2	BC022199	3804358	10	7.891	$5.26e^{-08}$	$1.15e^{-03}$	$1.77e^{-07}$
FLOT2	flotillin 2	17q11-q12	NM_004475	3751121	18	4.594	$1.56e^{-07}$	$3.41e^{-03}$	$5.25e^{-07}$
SFRS2IP	splicing factor, arginine/serine-rich 2, interacting protein	12q13.11	NM_004719	3452145	26	3.620	$1.77e^{-07}$	$3.86e^{-03}$	$5.95e^{-07}$
FAM83H	family with sequence similarity 83, member H	8q24.3	NM_198488	3157722	11	6.667	$1.93e^{-07}$	$4.20e^{-03}$	$6.47e^{-07}$
SPATA3	spermatogenesis associated 3	2q37.1	BC047704	2531648	13	5.706	$2.49e^{-07}$	$5.44e^{-03}$	$8.37e^{-07}$
LIG1	ligase I, DNA, ATP-dependent	19q13.2-q13.3	NM_000234	3866898	29	3.346	$2.68e^{-07}$	$5.86e^{-03}$	$9.02e^{-07}$
FCAR	Fc fragment of IgA, receptor for	19q13.2-q13.4	NM_002000	3841862	12	5.909	$4.10e^{-07}$	$8.95e^{-03}$	$1.38e^{-06}$
GK5	glycerol kinase 5 (putative)	3q23	NM_001039547	2698693	19	4.160	$5.53e^{-07}$	$1.21e^{-02}$	$1.86e^{-06}$
MAMDC4	MAM domain containing 4	9q34.3	NM_206920	3194832	29	3.224	$6.49e^{-07}$	$1.42e^{-02}$	$2.18e^{-06}$
RIOK1	RIO kinase 1 (yeast)	6p24.3	NM_031480	2893721	20	3.807	$1.64e^{-06}$	$3.58e^{-02}$	$5.51e^{-06}$
FBXO34	F-box protein 34	14q22.3	NM_017943	3536786	12	5.396	$1.71e^{-06}$	$3.73e^{-02}$	$5.74e^{-06}$

Table 3

Alternatively Spliced Genes: Schizophrenia (SCZ) *vs.* Control (CNT)

Gene Symbol	Gene Product	Chromosomal Locus	Accession Number	Transcript ID	Probe Sets (*n*)	Diagnostic Group-x-Exon ID Interaction			
						F	*p*	Adjusted *p*	*q*
IRAK4	interleukin-1 receptor- associated kinase 4	12q12	NM_016123	3412296	16	6.333	$1.10e^{-10}$	$2.41e^{-06}$	$4.18e^{-10}$
CYC1	cytochrome c-1	8q24.3	NM_001916	3120051	8	11.177	$6.44e^{-10}$	$1.41e^{-05}$	$2.44e^{-09}$
CHI3L1	chitinase 3-like 1 (cartilage glycoprotein-39)	1q32.1	NM_001276	2451593	13	6.702	$1.77e^{-09}$	$3.86e^{-05}$	$6.71e^{-09}$
FLJ46321	FLJ46321 protein (FLJ46321), mRNA	9q21.32	ENST00000344803	3176711	9	7.076	$2.57e^{-07}$	$5.60e^{-03}$	$9.73e^{-07}$
ATXN3	ataxin 3	14q24.3-q32.2	NM_004993	3576889	12	5.492	$3.54e^{-07}$	$7.74e^{-03}$	$1.34e^{-06}$
DENND1A	DENN/MADD domain containing 1A	9q33.2	NM_020946	3224650	20	3.898	$3.55e^{-07}$	$7.75e^{-03}$	$1.35e^{-06}$
S100A12	S100 calcium binding protein A12	1q21	NM_005621	2435981	3	28.983	$3.97e^{-07}$	$8.67e^{-03}$	$1.51e^{-06}$
ARAF	v-raf murine sarcoma 3611 viral oncogene homolog	Xp11.4-p11.2	NM_001654	3976299	17	3.946	$2.00e^{-06}$	$4.36e^{-02}$	$7.57e^{-06}$

Table 4

Alternatively Spliced Genes: Bipolar Disorder (BPD) *vs*. Schizophrenia (SCZ)

Gene	Gene	Chromosomal	Accession	Transcript	Probe	Diagnostic Group-x-Exon ID Inter−action			
Symbol	**Product**	**Locus**	**Number**	**ID**	**Sets (*n*)**	***F***	***p***	**Adjusted *p***	***q***
SLC44A2	solute carrier family 44, member 2	19p13.1	NM_020428	3820612	24	5.217	$3.80e^{-12}$	$8.30e^{-08}$	$9.28e^{-12}$
DDX24	DEAD (Asp-Glu- Ala-Asp) box polypeptide 24	14q32	NM_020414	3577513	15	7.001	$3.02e^{-11}$	$6.60e^{-07}$	$7.38e^{-11}$
FIG4	FIG4 homolog (*S. cerevisiae*)	6q21	NM_014845	2920962	27	4.340	$1.33e^{-10}$	$2.90e^{-06}$	$3.24e^{-10}$
CTNNB1	catenin (cadherin- associated protein), beta 1, 88kDa	3p21	NM_001904	2618940	18	5.072	$3.62e^{-09}$	$7.90e^{-05}$	$8.83e^{-09}$
ATG16L1	ATG16 autophagy related 16-like 1 (*S. cerevisiae*)	2q3	NM_030803	2532793	30	3.477	$1.86e^{-08}$	$4.07e^{-04}$	$4.55e^{-08}$
EX-OC7	exocyst complex component 7	17q25.1	NM_001013839	3771336	28	3.566	$2.68e^{-08}$	$5.85e^{-04}$	$6.53e^{-08}$
MAPK14	mitogen-activated protein kinase 14	6p21.3-p21.2	NM_139012	2904877	17	4.832	$3.00e^{-08}$	$6.56e^{-04}$	$7.32e^{-08}$
LTBP1	latent transforming growth factor beta binding protein 1	2p22-p21	NM_206943	2476510	42	2.820	$6.38e^{-08}$	$1.39e^{-03}$	$1.55e^{-07}$
JMJD2A	jumonji domain containing 2A	1p34.1	NM_014663	2333429	27	3.513	$6.91e^{-08}$	$1.51e^{-03}$	$1.68e^{-07}$
LE−PRE1	leucine proline-enriched proteoglycan (leprecan) 1	1p3	NM_022356	2409004	29	3.374	$7.22e^{-08}$	$1.58e^{-03}$	$1.76e^{-07}$
MCM4	minichromosome maintenance complex component 4	8q11.2	NM_005914	3097152	23	3.800	$9.47e^{-08}$	$2.07e^{-03}$	$2.31e^{-07}$
EDEM1	ER degradation enhancer, mannosidase alpha-like 1	3p26.	NM_014674	2608801	21	4.008	$9.70e^{-08}$	$2.12e^{-03}$	$2.36e^{-07}$
VPS13D	vacuolar protein sorting 13 homolog D (*S. cerevisiae*)	1p36.22	NM_015378	2320762	77	2.172	$1.15e^{-07}$	$2.50e^{-03}$	$2.79e^{-07}$
TMEM120−A	transmembrane protein 120A	7q11.23	NM_031925	3057520	13	5.490	$1.17e^{-07}$	$2.56e^{-03}$	$2.85e^{-07}$
ATG2B	ATG2 autophagy related 2 homolog B (*S. cerevisiae*)	14q3	NM_018036	3578278	30	3.233	$1.40e^{-07}$	$3.07e^{-03}$	$3.42e^{-07}$
LEPROT	leptin receptor overlapping transcript	1p31.3	NM_017526	2340433	28	3.277	$2.51e^{-07}$	$5.49e^{-03}$	$6.12e^{-07}$
PI4KA	phosphatidylinositol 4-kinase, catalytic, alpha	22q11.2	NM_058004	3953724	40	2.721	$3.64e^{-07}$	$7.95e^{-03}$	$8.85e^{-07}$
SASH1−	SAM and SH3 domain containing 1	6q24.3	NM_015278	2930243	29	3.127	$5.19e^{-07}$	$1.13e^{-02}$	$1.26e^{-06}$
ENTPD6	ectonucleoside triphosphate diphosphohydrolase 6 (putative function)	20p11.2-p11.22	NM_001247	3880706	22	3.614	$5.52e^{-07}$	$1.21e^{-02}$	$1.34e^{-06}$
BSDC1	BSD domain containing 1	1p35.1	NM_018045	2405036	24	3.405	$7.17e^{-07}$	$1.57e^{-02}$	$1.74e^{-06}$
TFR2	transferrin receptor 2	7q22	NM_003227	3064158	18	4.016	$7.22e^{-07}$	$1.58e^{-02}$	$1.75e^{-06}$
TTC15	tetratricopeptide repeat domain 15	2p25.3	NM_016030	2467691	20	3.767	$7.35e^{-07}$	$1.61e^{-02}$	$1.79e^{-06}$
CIRBP	cold inducible RNA binding protein	19p13.3	NM_001280	3815649	8	7.210	$9.48e^{-07}$	$2.07e^{-02}$	$2.30e^{-06}$
EI−F4EBP3	eukaryotic translation initiation factor 4E binding protein 3	5q31.3	NM_003732	2831719	64	2.178	$1.08e^{-06}$	$2.36e^{-02}$	$2.63e^{-06}$

Gene Symbol	Gene Product	Chromosomal Locus	Accession Number	Transcript ID	Probe Sets (*n*)	Diagnostic Group-x-Exon ID Inter–action			
						F	*p*	Adjusted *p*	*q*
FLJ13611	hypothetical protein FLJ13611	5q12.3	NM_001093756	2812315	14	4.607	$1.15e^{-06}$	$2.50e^{-02}$	$2.78e^{-06}$
COQ2	coenzyme Q2 homolog,prenyltransferase (yeast)	4q21.23	NM_015697	2775965	13	4.846	$1.15e^{-06}$	$2.52e^{-02}$	$2.80e^{-06}$
C14orf118	chromosome 14 open reading frame 118	14q22.1-q24.3	NM_017926	3544905	15	4.378	$1.26e^{-06}$	$2.74e^{-02}$	$3.05e^{-06}$
UTRN	utrophin	6q24	NM_007124	2929168	74	2.063	$1.27e^{-06}$	$2.77e^{-02}$	$3.08e^{-06}$
JA-RID2	jumonji, AT rich interactive domain 2	6p24-p23	NM_004973	2896177	21	3.531	$1.59e^{-06}$	$3.48e^{-02}$	$3.86e^{-06}$
BRD1	bromodomain containing 1	22q13.33	NM_014577	3965314	17	3.980	$1.70e^{-06}$	$3.71e^{-02}$	$4.12e^{-06}$
C18orf10	chromosome 18 open reading frame 10	18q12.2	BC022199	3804358	10	5.706	$1.88e^{-06}$	$4.11e^{-02}$	$4.57e^{-06}$
TNKS	tankyrase, TRF1-interacting ankyrin-related ADP-ribose polymerase	8p23.1	NM_003747	3085270	29	2.960	$1.94e^{-06}$	$4.23e^{-02}$	$4.70e^{-06}$
SN–AP29	synaptosomal-associated protein, 29kDa	22q11.21	NM_004782	3937755	11	5.255	$2.19e^{-06}$	$4.78e^{-02}$	$5.31e^{-06}$

Table 5

Alternatively Spliced Genes: Bipolar Disorder (BPD) + Schizophrenia (SCZ) *vs.* Control (CNT)

Gene Symbol	Gene Product	Chromosomal Locus	Accession Number	Transcript ID	Probe Sets (*n*)	Diagnostic Group-x-Exon ID Interaction			
						F	*p*	Adjusted *p*	*q*
VCL	vinculin	10q22.1-q23	NM_014000	3820612	37	3.392	$2.99e^{-10}$	$6.53e^{-06}$	$9.07e^{-10}$
PARP10	poly (ADP-ribose)polymerase family, member 10	8q24.3	NM_032789	3577513	20	4.021	$5.95e^{-08}$	$1.30e^{-03}$	$1.81e^{-07}$
IL3	interleukin 3 (colony-stimulating factor, multiple)	5q31.1	NM_000588	2920962	7	8.607	$8.71e^{-08}$	$1.91e^{-03}$	$2.64e^{-07}$
PERQ1	PERQ amino acid rich, with GYF domain 1	7q22	NM_022574	2618940	30	3.107	$1.67e^{-07}$	$3.64e^{-03}$	$5.05e^{-07}$
FLJ46321	FLJ46321 protein (FLJ46321), mRNA	9q21.32	ENST00000344803	2532793	9	6.632	$1.78e^{-07}$	$3.90e^{-03}$	$5.41e^{-07}$
PTPRC	protein tyrosine phosphatase, receptor type, C	1q31-q32	NM_002838	3771336	38	2.754	$2.24e^{-07}$	$4.90e^{-03}$	$6.79e^{-07}$
LIG1	ligase I, DNA, ATP-dependent	19q13.2-q13.3	NM_000234	2904877	29	3.112	$2.52e^{-07}$	$5.50e^{-03}$	$7.62e^{-07}$
CYC1	cytochrome c-1	8q24.3	NM_001916	2476510	8	6.919	$4.52e^{-07}$	$9.89e^{-03}$	$1.37e^{-06}$
SMAD2	SMAD family member 2	18q21.1	NM_005901	2333429	16	4.200	$4.75e^{-07}$	$1.04e^{-02}$	$1.44e^{-06}$
IDH3A	isocitrate dehydrogenase 3 (NAD+) alpha	15q25.1-q25.2	NM_005530	2409004	15	4.202	$1.05e^{-06}$	$2.29e^{-02}$	$3.17e^{-06}$
ADRBK1	adrenergic, beta, receptor kinase 1	11q13.1	NM_001619	3097152	27	3.037	$1.14e^{-06}$	$2.49e^{-02}$	$3.45e^{-06}$
SRCAP	Snf2-related CBP activator protein	16p11.2	NM_006662	2608801	52	2.294	$1.22e^{-06}$	$2.67e^{-02}$	$3.70e^{-06}$
PLK4	polo-like kinase 4 (*Drosophila*)	4q28	NM_014264	2320762	19	3.620	$1.31e^{-06}$	$2.87e^{-02}$	$3.97e^{-06}$
KIAA0460	KIAA0460	1q21.2	BC045623	3057520	22	3.312	$1.63e^{-06}$	$3.57e^{-02}$	$4.94e^{-06}$
FLOT2	flotillin 2	17q11-q12	NM_004475	3578278	18	3.682	$1.72e^{-06}$	$3.77e^{-02}$	$5.21e^{-06}$

Table 6

Alternatively Spliced Genes: History of Psychosis [PSYCH(+)] *vs.* No History of Psychosis [PSYCH(−)]

Gene Symbol	Gene Product	Chromosomal Locus	Accession Number	Transcript ID	Probe Sets (*n*)	Diagnostic Group-x-Exon ID Interaction			
						F	*p*	Adjusted *p*	*q*
SYNE1	spectrin repeat containing, nuclear envelope 1	6q25	NM_033071	2979871	167	2.788	$3.41e^{-27}$	$7.45e^{-23}$	$8.21e^{-27}$
KIAA0460	KIAA0460	1q21.2	BC045623	2358221	22	7.306	$2.12e^{-18}$	$4.63e^{-14}$	$5.11e^{-18}$
VPS13D	vacuolar protein sorting 13 homolog D (*S. cerevisiae*)	1p36.22	NM_015378	2320762	77	2.958	$3.67e^{-15}$	$8.02e^{-11}$	$8.84e^{-15}$
ZYX	zyxin	7q32	NM_003461	3029129	15	8.215	$9.83e^{-15}$	$2.15e^{-10}$	$2.37e^{-14}$
CLASP1	cytoplasmic linker associated protein 1	2q14.2-q14.3	NM_015282	2573641	53	3.503	$9.92e^{-15}$	$2.17e^{-10}$	$2.39e^{-14}$
JAK1	Janus kinase 1 (a protein tyrosine kinase)	1p32.3-p31.3	NM_002227	2416522	25	5.425	$2.41e^{-14}$	$5.28e^{-10}$	$5.81e^{-14}$
UBR4	ubiquitin protein ligase E3 component n-recognin 4	1p36.13	NM_020765	2399409	115	2.370	$1.31e^{-13}$	$2.87e^{-09}$	$3.16e^{-13}$
DYNC1H1	dynein, cytoplasmic 1, heavy chain 1	14q32.3-qter	NM_001376	3552847	84	2.657	$2.97e^{-13}$	$6.50e^{-09}$	$7.16e^{-13}$
SPG11	spastic paraplegia 11 (autosomal recessive)	15q14	NM_025137	3621948	53	3.138	$3.34e^{-12}$	$7.31e^{-08}$	$8.05e^{-12}$
ESF1	ESF1, nucleolar pre-rRNA processing protein, homolog (*S. cerevisiae*)	20p12.1	NM_016649	3898224	17	6.048	$1.17e^{-11}$	$2.57e^{-07}$	$2.82e^{-11}$
LRP1	low density lipoprotein-related protein 1 (alpha-2-macroglobulin receptor)	12q13-q14	NM_002332	3417842	105	2.283	$1.50e^{-11}$	$3.28e^{-07}$	$3.61e^{-11}$
UBR2	ubiquitin protein ligase E3 component n-recognin 2	6p21.1	NM_015255	2907190	53	3.007	$2.58e^{-11}$	$5.65e^{-07}$	$6.22e^{-11}$
KIAA0082	KIAA0082	6p21.2	ENST00000373451	2905512	26	4.423	$2.72e^{-11}$	$5.94e^{-07}$	$6.54e^{-11}$
HD	huntingtin (Huntington disease)	4p16.3	NM_002111	2715820	89	2.364	$7.20e^{-11}$	$1.57e^{-06}$	$1.73e^{-10}$
ZZEF1	zinc finger, ZZ-type with EF-hand domain 1	17p13.2	NM_015113	3741875	70	2.600	$7.47e^{-11}$	$1.63e^{-06}$	$1.80e^{-10}$
UTRN	utrophin	6q24	NM_007124	2929168	74	2.521	$1.10e^{-10}$	$2.41e^{-06}$	$2.65e^{-10}$
EIF4EBP3	eukaryotic translation initiation factor 4E binding protein 3	5q31.3	NM_003732	2831719	64	2.675	$1.18e^{-10}$	$2.59e^{-06}$	$2.84e^{-10}$
DSC2	desmocollin 2	18q12.1	NM_004949	3802980	22	4.557	$3.40e^{-10}$	$7.43e^{-06}$	$8.17e^{-10}$
HIPK2	homeodomain interacting protein kinase 2	7q32-q34	NM_022740	3075778	19	4.993	$3.87e^{-10}$	$8.45e^{-06}$	$9.29e^{-10}$
HERC1	hect d domain and RCC1 (CHC1)-like domain (RLD) 1	15q22	NM_003922	3628650	87	2.303	$4.59e^{-10}$	$1.00e^{-05}$	$1.10e^{-09}$
MLL3	myeloid/lymphoid or mixed-lineage leukemia 3	7q36.1	NM_170606	3080033	69	2.514	$5.09e^{-10}$	$1.11e^{-05}$	$1.22e^{-09}$
MDN1	MDN1, midasin homolog (yeast)	6q15	NM_014611	2964350	119	2.054	$6.32e^{-10}$	$1.38e^{-05}$	$1.52e^{-09}$
DIP2A	DIP2 disco-interacting protein 2 homolog A (*Drosophila*)	21q22.3	NM_015151	3924674	42	3.077	$1.09e^{-09}$	$2.39e^{-05}$	$2.62e^{-09}$
FAM62B	family with sequence similarity 62 (C2 domain containing)	7q36.3	NM_020728	3082248	31	3.581	$1.25e^{-09}$	$2.72e^{-05}$	$2.99e^{-09}$

Gene Symbol	Gene Product	Chromosomal Locus	Accession Number	Transcript ID	Probe Sets (*n*)	Diagnostic Group-x-Exon ID Interaction			
						F	*p*	Adjusted *p*	*q*
C1orf198	chromosome 1 open reading frame 198	1q42.13-q43	AK096166	2460325	17	5.135	$1.48e^{-09}$	$3.24e^{-05}$	$3.56e^{-09}$
SMG5	Smg-5 homolog, nonsense mediated mRNA decay factor (*C. elegans*)	1q21.2	NM_015327	2438042	31	3.562	$1.50e^{-09}$	$3.28e^{-05}$	$3.60e^{-09}$
EIF2C4	eukaryotic translation initiation factor 2C, 4	1p34.3	NM_017629	2330002	28	3.766	$1.51e^{-09}$	$3.30e^{-05}$	$3.62e^{-09}$
SPEN	spen homolog, transcriptional regulator (*Drosophila*)	1p3	NM_015001	2322103	35	3.321	$1.85e^{-09}$	$4.04e^{-05}$	$4.43e^{-09}$
ARHGAP25	Rho GTPase activating protein 25	2p14	NM_014882	2486927	13	6.088	$2.65e^{-09}$	$5.78e^{-05}$	$6.34e^{-09}$
ATXN7	ataxin 7	3p21.1-p12	NM_000333	2627390	23	4.099	$3.71e^{-09}$	$8.11e^{-05}$	$8.89e^{-09}$
VPS13B	vacuolar protein sorting 13 homolog B (yeast)	8q22.2	NM_017890	3108901	74	2.341	$3.88e^{-09}$	$8.48e^{-05}$	$9.29e^{-09}$
ZFX	zinc finger protein, X-linked	Xp21.3	NM_003410	3971923	7	10.235	$3.94e^{-09}$	$8.61e^{-05}$	$9.43e^{-09}$
FCHSD1	FCH and double SH3 domains 1	5q31.3	NM_033449	2878778	24	3.972	$4.48e^{-09}$	$9.79e^{-05}$	$1.07e^{-08}$
MMP27	matrix metallopeptidase 27	11q24	NM_022122	3388730	10	7.326	$4.55e^{-09}$	$9.95e^{-05}$	$1.09e^{-08}$
ARHGAP27	Rho GTPase activating protein 27	17q21.31	NM_199282	3759778	18	4.727	$4.95e^{-09}$	$1.08e^{-04}$	$1.19e^{-08}$
DOCK5	dedicator of cytokinesis 5	8p21.2	NM_024940	3090512	67	2.401	$6.79e^{-09}$	$1.48e^{-04}$	$1.62e^{-08}$
PLEC1	plectin 1, intermediate filament binding protein 500kDa	8q24	NM_201380	3157901	60	2.505	$7.19e^{-09}$	$1.57e^{-04}$	$1.72e^{-08}$
MYCBP2	MYC binding protein 2	13q22	NM_015057	3518496	97	2.096	$7.76e^{-09}$	$1.70e^{-04}$	$1.86e^{-08}$
ISCU	iron-sulfur cluster scaffold homolog (*E. coli*)	12q24.1	NM_014301	3430776	9	7.769	$8.77e^{-09}$	$1.92e^{-04}$	$2.10e^{-08}$
BIRC6	baculoviral IAP repeat-containing 6 (apollon)	2p22-p21	NM_016252	2476219	85	2.187	$8.83e^{-09}$	$1.93e^{-04}$	$2.11e^{-08}$
MAP3K5	mitogen-activated protein kinase kinase kinase 5	6q22.33	NM_005923	2975867	34	3.215	$9.09e^{-09}$	$1.99e^{-04}$	$2.17e^{-08}$
TRERF1	transcriptional regulating factor 1	6p21.1-p12.1	NM_033502	2954025	32	3.281	$1.28e^{-08}$	$2.79e^{-04}$	$3.05e^{-08}$
NOTCH2	Notch homolog 2 (Drosophila)	1p13-p11	NM_024408	2431112	37	3.043	$1.43e^{-08}$	$3.12e^{-04}$	$3.41e^{-08}$
FAM48A	family with sequence similarity 48, member A	13q13.3	NM_017569	3509910	31	3.322	$1.44e^{-08}$	$3.15e^{-04}$	$3.44e^{-08}$
SP100	SP100 nuclear antigen	2q37.1	NM_001080391	2531377	28	3.503	$1.45e^{-08}$	$3.18e^{-04}$	$3.47e^{-08}$
C7orf26	chromosome 7 open reading frame 26	7p22.1	BC005121	2989141	10	6.904	$1.60e^{-08}$	$3.49e^{-04}$	$3.81e^{-08}$
PIP5K3	phosphatidylinositol-3-phosphate/phosphatidylinositol5-kinase, type III	2q33.3	NM_015040	2525272	47	2.712	$1.72e^{-08}$	$3.77e^{-04}$	$4.11e^{-08}$
CTTN	cortactin	11q13	NM_005231	3338552	27	3.535	$1.99e^{-08}$	$4.34e^{-04}$	$4.74e^{-08}$
YEATS2	YEATS domain containing 2	3q27.1	NM_018023	2655168	39	2.938	$1.99e^{-08}$	$4.35e^{-04}$	$4.75e^{-08}$
PTK2	PTK2 protein tyrosine kinase 2	8q24-qter	NM_005607	3156307	35	3.088	$2.12e^{-08}$	$4.63e^{-04}$	$5.05e^{-08}$
INTS3	integrator complex subunit 3	1q21.3	NM_023015	2359817	35	3.080	$2.30e^{-08}$	$5.03e^{-04}$	$5.49e^{-08}$

Gene Symbol	Gene Product	Chromosomal Locus	Accession Number	Transcript ID	Probe Sets (*n*)	Diagnostic Group-x-Exon ID Interaction			
						F	*p*	Adjusted *p*	*q*
EIF4G2	eukaryotic translation initiation factor 4 gamma, 2	11p15	NM_001418	3362719	26	3.579	$2.50e^{-08}$	$5.47e^{-04}$	$5.96e^{-08}$
HMGCR	3-hydroxy-3-methylglutaryl- Coenzyme A reductase	5q13.3-q14	NM_000859	2815965	27	3.500	$2.64e^{-08}$	$5.77e^{-04}$	$6.29e^{-08}$
LYRM1	LYR motif containing 1	16p11.2	NM_020424	3651588	9	7.344	$2.68e^{-08}$	$5.87e^{-04}$	$6.40e^{-08}$
NFKB2	nuclear factor of kappa light polypeptide gene enhancer	10q24	NM_001077493	3261643	24	3.731	$2.69e^{-08}$	$5.88e^{-04}$	$6.42e^{-08}$
ASXL2	additional sex combs like 2 (*Drosophila*)	2p24.1	NM_018263	2544925	23	3.819	$2.75e^{-08}$	$6.02e^{-04}$	$6.56e^{-08}$
CR1	complement component (3b/4b) receptor 1 (Knops blood group)	1q32	NM_000651	2377332	19	4.272	$2.81e^{-08}$	$6.15e^{-04}$	$6.70e^{-08}$
SRRM2	serine/arginine repetitive matrix 2	16p13.3	NM_016333	3645253	40	2.873	$2.84e^{-08}$	$6.21e^{-04}$	$6.77e^{-08}$
PTPN1	protein tyrosine phosphatase, non-receptor type 1	20q13	NM_002827	3888721	20	4.120	$3.21e^{-08}$	$7.01e^{-04}$	$7.64e^{-08}$
TPR	translocated promoter region (to activated MET oncogene)	1q25	NM_003292	2448232	51	2.563	$3.79e^{-08}$	$8.29e^{-04}$	$9.02e^{-08}$
MACF1	microtubule-actin crosslinking factor 1	1p32-p31	NM_012090	2331213	115	1.923	$3.93e^{-08}$	$8.59e^{-04}$	$9.35e^{-08}$
JARID1C	jumonji, AT rich interactive domain 1C	Xp11.22-p11.21	NM_004187	4009062	31	3.208	$4.15e^{-08}$	$9.08e^{-04}$	$9.88e^{-08}$
ITSN2	intersectin 2	2pter-p25.1	NM_006277	2544238	50	2.571	$4.58e^{-08}$	$1.00e^{-03}$	$1.09e^{-07}$
DFFB	DNA fragmentation factor, 40kDa, beta polypeptide	1p36.3	NM_004402	2317512	13	5.311	$5.97e^{-08}$	$1.31e^{-03}$	$1.42e^{-07}$
TTC15	tetratricopeptide repeat domain 15	2p25.3	NM_016030	2467691	20	4.020	$5.98e^{-08}$	$1.31e^{-03}$	$1.42e^{-07}$
NPAL3	NIPA-like domain containing 3	1p36.12-p35.1	NM_020448	2325410	18	4.259	$6.90e^{-08}$	$1.51e^{-03}$	$1.64e^{-07}$
ABCA1	ATP-binding cassette, sub-family A (ABC1), member 1	9q3	NM_005502	3218528	55	2.446	$6.90e^{-08}$	$1.51e^{-03}$	$1.64e^{-07}$
EIF4G3	eukaryotic translation initiation factor 4 gamma, 3	1p36.12	NM_003760	2400373	47	2.603	$7.31e^{-08}$	$1.60e^{-03}$	$1.74e^{-07}$
RBM26	RNA binding motif protein 26	13q31.1	NM_022118	3519119	23	3.680	$7.36e^{-08}$	$1.61e^{-03}$	$1.75e^{-07}$
PRKAR2A	protein kinase, cAMP-dependent, regulatory, type II, alpha	3p21.3-p21.2	NM_004157	2673730	16	4.559	$7.91e^{-08}$	$1.73e^{-03}$	$1.88e^{-07}$
ADRBK1	adrenergic, beta, receptor kinase 1	11q13.1	NM_001619	3336801	27	3.364	$8.09e^{-08}$	$1.77e^{-03}$	$1.92e^{-07}$
NUP214	nucleoporin 214kDa	9q34.1	NM_005085	3191900	48	2.563	$9.17e^{-08}$	$2.00e^{-03}$	$2.18e^{-07}$
EPHB4	EPH receptor B4	7q22	NM_004444	3064293	23	3.649	$9.17e^{-08}$	$2.01e^{-03}$	$2.18e^{-07}$
SYNE2	spectrin repeat containing, nuclear envelope 2	14q23.2	NM_182914	3539724	134	1.809	$9.59e^{-08}$	$2.10e^{-03}$	$2.28e^{-07}$
DDX3X	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked	Xp11	NM_001356	3974838	20	3.939	$9.88e^{-08}$	$2.16e^{-03}$	$2.34e^{-07}$
EP400	E1A binding protein p400	12q24.33	NM_015409	3438617	63	2.295	$1.05e^{-07}$	$2.31e^{-03}$	$2.50e^{-07}$
HERC2	hect domain and RLD 2	15q13	NM_004667	3614901	57	2.384	$1.06e^{-07}$	$2.31e^{-03}$	$2.51e^{-07}$
SMG7	Smg-7 homolog, nonsense mediated mRNA decay factor (*C. elegans*)	1q25	NM_173156	2371255	31	3.106	$1.07e^{-07}$	$2.34e^{-03}$	$2.54e^{-07}$

Gene Symbol	Gene Product	Chromosomal Locus	Accession Number	Transcript ID	Probe Sets (*n*)	Diagnostic Group-x-Exon ID Interaction			
						F	*p*	Adjusted *p*	*q*
PRKDC	protein kinase, DNA-activated, catalytic polypeptide	8q11	NM_001081640	3134034	99	1.968	$1.13e^{-07}$	$2.47e^{-03}$	$2.68e^{-07}$
CENTB2	centaurin, beta 2	3q29	NM_012287	2712040	26	3.387	$1.14e^{-07}$	$2.49e^{-03}$	$2.71e^{-07}$
HSPH1	heat shock 105kDa/110kDa protein 1	13q12.3	NM_006644	3508330	26	3.387	$1.15e^{-07}$	$2.51e^{-03}$	$2.72e^{-07}$
NFAT5	Nuclear factor of activated T-cells 5, tonicity-responsive	16q22.1	NM_138714	3666779	39	2.781	$1.17e^{-07}$	$2.55e^{-03}$	$2.77e^{-07}$
CCNT1	cyclin T1	12pter-qter	NM_001240	3453218	12	5.431	$1.23e^{-07}$	$2.68e^{-03}$	$2.91e^{-07}$
LRBA	LPS-responsive vesicle trafficking, beach and anchor containing	4q31.3	NM_006726	2789266	65	2.259	$1.26e^{-07}$	$2.75e^{-03}$	$2.98e^{-07}$
HERC3	hect domain and RLD 3	4q21	NM_014606	2735459	27	3.296	$1.40e^{-07}$	$3.07e^{-03}$	$3.33e^{-07}$
RAP1GDS1	RAP1, GTP-GDP dissociation stimulator 1	4q23-q25	NM_021159	2736853	16	4.440	$1.44e^{-07}$	$3.14e^{-03}$	$3.40e^{-07}$
NPL	N-acetylneuraminate pyruvate lyase (dihydrodipicolinate synthase)	1q25	NM_030769	2370926	18	4.126	$1.45e^{-07}$	$3.18e^{-03}$	$3.44e^{-07}$
SLC30A6	solute carrier family 30 (zinc transporter), member 6	2p22.3	NM_017964	2476116	17	4.229	$1.83e^{-07}$	$4.00e^{-03}$	$4.33e^{-07}$
AFF1	AF4/FMR2 family, member 1	4q21	NM_005935	2734784	33	2.947	$2.03e^{-07}$	$4.44e^{-03}$	$4.80e^{-07}$
NRD1	nardilysin (N-arginine dibasic convertase)	1p32.2-p32.1	NM_002525	2412529	36	2.829	$2.07e^{-07}$	$4.52e^{-03}$	$4.89e^{-07}$
ST3GAL6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	3q	NM_006100	2633256	15	4.549	$2.08e^{-07}$	$4.54e^{-03}$	$4.91e^{-07}$
NUP133	nucleoporin 133kDa	1q42.13	NM_018230	2459866	32	2.984	$2.15e^{-07}$	$4.70e^{-03}$	$5.08e^{-07}$
ADAR	adenosine deaminase, RNA-specific	1q21.1-q21.2	NM_001111	2436754	27	3.234	$2.33e^{-07}$	$5.10e^{-03}$	$5.51e^{-07}$
GOLGA2	golgi autoantigen, golgin subfamily a, 2	9q34.11	NM_004486	3226431	16	4.337	$2.40e^{-07}$	$5.25e^{-03}$	$5.67e^{-07}$
DVL3	dishevelled, dsh homolog 3 (*Drosophila*)	3q27	NM_004423	2655438	24	3.432	$2.42e^{-07}$	$5.30e^{-03}$	$5.73e^{-07}$
KIAA0319L	KIAA0319-like	1p34.2	NM_024874	2406139	31	3.016	$2.44e^{-07}$	$5.32e^{-03}$	$5.75e^{-07}$
COL13A1	collagen, type XIII, alpha 1	10q22	NM_005203	3250486	25	3.352	$2.55e^{-07}$	$5.57e^{-03}$	$6.02e^{-07}$
SAPS3	SAPS domain family, member 3	11q13	NM_018312	3337618	26	3.279	$2.68e^{-07}$	$5.86e^{-03}$	$6.33e^{-07}$
HOOK3	hook homolog 3 (*Drosophila*)	8p11.21	NM_032410	3096368	20	3.778	$2.68e^{-07}$	$5.86e^{-03}$	$6.33e^{-07}$
GPBP1	GC-rich promoter binding protein 1	5q11.2	NM_022913	2810458	16	4.311	$2.73e^{-07}$	$5.98e^{-03}$	$6.45e^{-07}$
UBE3C	ubiquitin protein ligase E3C	7q36.3	NM_014671	3033924	34	2.871	$2.87e^{-07}$	$6.28e^{-03}$	$6.78e^{-07}$
BAZ2B	bromodomain adjacent to zinc finger domain, 2B	2q23-q24	NM_013450	2583014	46	2.506	$3.37e^{-07}$	$7.37e^{-03}$	$7.95e^{-07}$
PDDC1	Parkinson disease 7 domain containing 1	11p15.5	NM_182612	3358361	14	4.646	$3.39e^{-07}$	$7.41e^{-03}$	$8.00e^{-07}$
TSC2	tuberous sclerosis 2	16p13.3	NM_000548	3644375	42	2.601	$3.41e^{-07}$	$7.46e^{-03}$	$8.05e^{-07}$
ZDHHC17	zinc finger, DHHC-type containing 17	12q21.2	NM_015336	3423184	21	3.634	$3.50e^{-07}$	$7.65e^{-03}$	$8.25e^{-07}$

Gene Symbol	Gene Product	Chromosomal Locus	Accession Number	Transcript ID	Probe Sets (*n*)	Diagnostic Group-x-Exon ID Interaction			
						F	*p*	Adjusted *p*	*q*
PXK	PX domain containing serine/threonine kinase	3p14.3	NM_017771	2626167	20	3.732	$3.55e^{-07}$	$7.75e^{-03}$	$8.36e^{-07}$
CHD4	chromodomain helicase DNA binding protein 4	12p13	NM_001273	3442054	43	2.568	$3.73e^{-07}$	$8.16e^{-03}$	$8.79e^{-07}$
COL19A1	collagen, type XIX, alpha 1	6q12-q13	NM_001858	2912649	32	2.914	$4.15e^{-07}$	$9.08e^{-03}$	$9.78e^{-07}$
ANKRD28	ankyrin repeat domain 28	3p24.3	NM_015199	2664452	34	2.831	$4.26e^{-07}$	$9.31e^{-03}$	$1.00e^{-06}$
PPT1	palmitoyl-protein thioesterase 1	1p32	NM_000310	2408189	10	5.797	$4.50e^{-07}$	$9.83e^{-03}$	$1.06e^{-06}$
USP34	ubiquitin specific peptidase 34	2p15	NM_014709	2555277	83	2.008	$4.83e^{-07}$	$1.06e^{-02}$	$1.14e^{-06}$
STXBP2	syntaxin binding protein 2	19p13.3-p13.2	NM_006949	3819016	17	4.041	$4.93e^{-07}$	$1.08e^{-02}$	$1.16e^{-06}$
MLL	myeloid/lymphoid or mixed-lineage leukemia	11q23	NM_005933	3351385	48	2.434	$4.95e^{-07}$	$1.08e^{-02}$	$1.16e^{-06}$
WDFY3	WD repeat and FYVE domain containing 3	4q21.23	NM_014991	2776372	87	1.973	$5.42e^{-07}$	$1.19e^{-02}$	$1.28e^{-06}$
MAPK14	mitogen-activated protein kinase 14	6p21.3-p21.2	NM_139012	2904877	17	4.021	$5.49e^{-07}$	$1.20e^{-02}$	$1.29e^{-06}$
TRIM33	tripartite motif-containing 33	1p13.1	NM_015906	2429069	28	3.072	$5.53e^{-07}$	$1.21e^{-02}$	$1.30e^{-06}$
ECOP	EGFR-coamplified and overexpressed protein	7p11.2	NM_030796	3051655	10	5.715	$5.77e^{-07}$	$1.26e^{-02}$	$1.36e^{-06}$
LMTK2	lemur tyrosine kinase 2	7q21.3	NM_014916	3014159	27	3.121	$5.80e^{-07}$	$1.27e^{-02}$	$1.36e^{-06}$
MGAM	maltase-glucoamylase (alpha-glucosidase)	7q34	NM_004668	3028011	36	2.729	$5.85e^{-07}$	$1.28e^{-02}$	$1.37e^{-06}$
DLL1	delta-like 1 (*Drosophila*)	6q27	NM_005618	2986350	15	4.310	$6.34e^{-07}$	$1.39e^{-02}$	$1.49e^{-06}$
ARHGAP26	Rho GTPase activating protein 26	5q31	NM_015071	2833286	29	3.002	$6.42e^{-07}$	$1.40e^{-02}$	$1.51e^{-06}$
HCP5	HLA complex P5	6p21.3	NM_006674	2902326	9	6.127	$6.96e^{-07}$	$1.52e^{-02}$	$1.64e^{-06}$
VPS39	vacuolar protein sorting 39 homolog (*S. cerevisiae*)	15q	NM_015289	3620457	31	2.898	$7.10e^{-07}$	$1.55e^{-02}$	$1.67e^{-06}$
IVNS1ABP	influenza virus NS1A binding protein	1q25.1-q31.1	NM_006469	2448073	21	3.517	$7.44e^{-07}$	$1.63e^{-02}$	$1.75e^{-06}$
KIAA0746	KIAA0746	4p15.2	ENST00000264868	2764192	32	2.851	$7.44e^{-07}$	$1.63e^{-02}$	$1.75e^{-06}$
ITGA2B	integrin, alpha 2b	17q21.32	NM_000419	3759137	31	2.892	$7.54e^{-07}$	$1.65e^{-02}$	$1.77e^{-06}$
NPEPL1	aminopeptidase-like 1	20q13.32	NM_024663	3891048	18	3.820	$8.00e^{-07}$	$1.75e^{-02}$	$1.88e^{-06}$
FCER1G	Fc fragment of IgE, high affinity I, receptor for; gamma	1q23	NM_004106	2363562	6	8.586	$8.52e^{-07}$	$1.86e^{-02}$	$2.00e^{-06}$
ZFYVE26	zinc finger, FYVE domain containing 26	14q24.1	NM_015346	3569441	51	2.337	$8.61e^{-07}$	$1.88e^{-02}$	$2.02e^{-06}$
XYLT1	xylosyltransferase I	16p12.3	NM_022166	3682445	15	4.243	$8.67e^{-07}$	$1.90e^{-02}$	$2.03e^{-06}$
TMEM87A	transmembrane protein 87A	15q15	ENST00000389834	3620515	18	3.778	$1.01e^{-06}$	$2.21e^{-02}$	$2.37e^{-06}$
CSNK1G2	casein kinase 1, gamma 2	19p13.3	NM_001319	3816153	12	4.857	$1.03e^{-06}$	$2.26e^{-02}$	$2.42e^{-06}$

Gene Symbol	Gene Product	Chromosomal Locus	Accession Number	Transcript ID	Probe Sets (*n*)	Diagnostic Group-x-Exon ID Interaction			
						F	*p*	Adjusted *p*	*q*
KIAA0372	KIAA0372	5q15	ENST00000358746	2867693	46	2.416	$1.05e^{-06}$	$2.29e^{-02}$	$2.45e^{-06}$
DDX58	DEAD (Asp-Glu-Ala-Asp) box polypeptide 58	9p12	NM_014314	3203086	22	3.345	$1.32e^{-06}$	$2.88e^{-02}$	$3.09e^{-06}$
ERBB2IP	erbb2 interacting protein	5q12.3	NM_018695	2812435	30	2.872	$1.32e^{-06}$	$2.89e^{-02}$	$3.09e^{-06}$
ZNF638	zinc finger protein 638	2p13.2-p13.1	NM_014497	2488114	39	2.561	$1.32e^{-06}$	$2.89e^{-02}$	$3.10e^{-06}$
CPD	carboxypeptidase D	17q11.2	NM_001304	3716411	32	2.786	$1.36e^{-06}$	$2.96e^{-02}$	$3.17e^{-06}$
GGA2	golgi associated, gamma adaptin ear containing, ARF binding protein 2	16p12	NM_015044	3685183	28	2.958	$1.41e^{-06}$	$3.08e^{-02}$	$3.30e^{-06}$
DPP9	dipeptidyl-peptidase 9	19p13.3	NM_139159	3846926	24	3.185	$1.45e^{-06}$	$3.17e^{-02}$	$3.40e^{-06}$
CREBBP	CREB binding protein (Rubinstein-Taybi syndrome)	16p13	NM_004380	3677795	43	2.451	$1.50e^{-06}$	$3.28e^{-02}$	$3.51e^{-06}$
CDC2L5	cell division cycle 2-like 5	7p13	NM_003718	2998536	30	2.857	$1.51e^{-06}$	$3.30e^{-02}$	$3.53e^{-06}$
INPP5D	inositol polyphosphate-5-phosphatase, 145kDa	2q37.1	NM_005541	2532699	25	3.112	$1.55e^{-06}$	$3.39e^{-02}$	$3.62e^{-06}$
PIGT	phosphatidylinositol glycan anchor biosynthesis, class T	20q12-q13.12	NM_015937	3886889	14	4.287	$1.62e^{-06}$	$3.54e^{-02}$	$3.79e^{-06}$
JMJD2A	jumonji domain containing 2A	1p34.1	NM_014663	2333429	27	2.992	$1.63e^{-06}$	$3.56e^{-02}$	$3.81e^{-06}$
NCAPD3	non-SMC condensin II complex, subunit D3	11q25	NM_015261	3399545	39	2.538	$1.69e^{-06}$	$3.69e^{-02}$	$3.95e^{-06}$
VPS41	vacuolar protein sorting 41 homolog (*S. cerevisiae*)	7p1	NM_014396	7385683	27	2.972	$1.90e^{-06}$	$4.16e^{-02}$	$4.44e^{-06}$
CRAT	carnitine acetyltransferase	9q34.1	NM_000755	3226844	19	3.548	$1.99e^{-06}$	$4.35e^{-02}$	$4.65e^{-06}$
PLXNC1	plexin C1	12q23.3	NM_005761	3426502	32	2.741	$2.05e^{-06}$	$4.47e^{-02}$	$4.78e^{-06}$
ACSL3	acyl-CoA synthetase long-chain family member 3	2q34-q35	NM_004457	2529546	20	3.442	$2.10e^{-06}$	$4.60e^{-02}$	$4.91e^{-06}$
PDS5A	PDS5, regulator of cohesion maintenance, homolog A (*S. cerevisiae*)	4p14	NM_015200	2766588	41	2.466	$2.13e^{-06}$	$4.65e^{-02}$	$4.97e^{-06}$
NNT	nicotinamide nucleotide transhydrogenase	5p13.1-cen	NM_012343	2808438	27	2.956	$2.16e^{-06}$	$4.73e^{-02}$	$5.05e^{-06}$
TBC1D5	TBC1 domain family, member 5	3p24.3	NM_014744	2664891	23	3.193	$2.22e^{-06}$	$4.86e^{-02}$	$5.19e^{-06}$
EML4	echinoderm microtubule associated protein like 4	2p22-p2	NM_019063	2478748	30	2.811	$2.25e^{-06}$	$4.91e^{-02}$	$5.25e^{-06}$
MLL2	myeloid/lymphoid or mixed-lineage leukemia 2	12q12-q14	NM_003482	3453592	80	1.949	$2.26e^{-06}$	$4.93e^{-02}$	$5.27e^{-06}$
PTPRJ	protein tyrosine phosphatase, receptor type, J	11p11.2	NM_002843	3329983	29	2.854	$2.26e^{-06}$	$4.94e^{-02}$	$5.27e^{-06}$
TOP3A	topoisomerase (DNA) III alpha	17p12-p11.2	NM_004618	3748262	25	3.061	$2.26e^{-06}$	$4.95e^{-02}$	$5.28e^{-06}$

Table 7

Ontologies, Pathways, and Protein Domains Significantly Over-Represented Among Alternatively Spliced Genes: History of Psychosis [PSYCH(+)] *vs*. No History of Psychosis [PSYCH(−)]

Domain	Category	Term	*n* (%) of Genes on List	*p*
Ontology	Biological Process	biopolymer modification	29 (5.2)	$1.20e^{-04}$
		macromolecule metabolism	57 (10.0)	$1.40e^{-04}$
		protein metabolism	46 (8.3)	$1.80e^{-04}$
		cellular macromolecule metabolism	44 (7.9)	$1.80e^{-04}$
		cellular protein metabolism	43 (7.7)	$2.70e^{-04}$
		biopolymer metabolism	39 (7.0)	$3.90e^{-04}$
		protein modification	27 (4.8)	$4.50e^{-04}$
		cellular physiological process	98 (18.0)	$4.90e^{-03}$
		primary metabolism	74 (13.0)	$7.50e^{-03}$
		protein amino acid phosphorylation	12 (2.2)	$1.00e^{-02}$
		cell organization and biogenesis	21 (3.8)	$1.30e^{-02}$
		phosphate metabolism	15 (2.7)	$1.50e^{-02}$
		phosphorus metabolism	15 (2.7)	$1.50e^{-02}$
		cellular metabolism	74 (13.0)	$2.10e^{-02}$
		organ development	10 (1.8)	$2.30e^{-02}$
		cell cycle	12 (2.2)	$2.30e^{-02}$
		regulation of translational initiation	3 (0.5)	$2.60e^{-02}$
		organ morphogenesis	6 (1.1)	$2.70e^{-02}$
		endocytosis	5 (0.9)	$2.90e^{-02}$
		Notch signaling pathway	3 (0.5)	$3.00e^{-02}$
		nuclear organization and biogenesis	2 (0.4)	$3.00e^{-02}$
		organelle organization and biogenesis	13 (2.3)	$3.10e^{-02}$
		metabolism	77 (14.0)	$3.80e^{-02}$
		phosphorylation	12 (2.2)	$4.20e^{-02}$
		protein localization	10 (1.8)	$4.30e^{-02}$
		ubiquitin cycle	9 (1.6)	$4.70e^{-02}$
	Cellular Component	intracellular	88 (16.0)	$6.10e^{-05}$
		endomembrane system	10 (1.8)	$1.80e^{-03}$
		eukaryotic translation initiation factor 4F complex	3 (0.5)	$2.20e^{-03}$
		cytoplasm	43 (7.7)	$6.20e^{-03}$
		nuclear envelope	5 (0.9)	$7.30e^{-03}$
		intracellular organelle	70 (13.0)	$9.60e^{-03}$
		organelle	70 (13.0)	$9.70e^{-03}$
		nucleus	44 (7.9)	$2.30e^{-02}$

Domain	Category	Term	*n* (%) of Genes on List	*p*
		cytoskeleton	14 (2.5)	$3.50e^{-02}$
		cell	108 (19.0)	$4.00e^{-02}$
	Molecular Function	protein binding	67 (12.0)	$8.20e^{-10}$
		GTPase regulator activity	11 (2.0)	$5.20e^{-05}$
		enzyme regulator activity	15 (2.7)	$4.40e^{-04}$
		GTPase activator activity	7 (1.3)	$7.50e^{-04}$
		binding	103 (19.0)	$9.80e^{-04}$
		zinc ion binding	29 (5.2)	$1.00e^{-03}$
		enzyme activator activity	8 (1.4)	$1.40e^{-03}$
		microtubule binding	4 (0.7)	$3.30e^{-03}$
		protein kinase activity	13 (2.3)	$3.70e^{-03}$
		adenyl nucleotide binding	21 (3.8)	$4.30e^{-03}$
		cytoskeletal protein binding	9 (1.6)	$4.70e^{-03}$
		phosphotransferase activity, alcohol group as acceptor	14 (2.5)	$5.50e^{-03}$
		tubulin binding	4 (0.7)	$6.30e^{-03}$
		ATP binding	20 (3.6)	$6.40e^{-03}$
		ubiquitin-protein ligase activity	8 (1.4)	$8.80e^{-03}$
		catalytic activity	57 (10.0)	$9.40e^{-03}$
		ligase activity	10 (1.8)	$1.20e^{-02}$
		small GTPase regulator activity	6 (1.1)	$1.30e^{-02}$
		transition metal ion binding	29 (5.2)	$1.40e^{-02}$
		cation binding	37 (6.6)	$1.60e^{-02}$
		kinase activity	15 (2.7)	$1.60e^{-02}$
		ion binding	39 (7.0)	$1.60e^{-02}$
		metal ion binding	39 (7.0)	$1.60e^{-02}$
		protein serine/threonine kinase activity	9 (1.6)	$1.70e^{-02}$
		ligase activity, forming carbon-nitrogen bonds	8 (1.4)	$2.10e^{-02}$
		purine nucleotide binding	22 (3.9)	$2.50e^{-02}$
		nucleotide binding	24 (4.3)	$3.60e^{-02}$
		actin binding	6 (1.1)	$3.70e^{-02}$
		transcription cofactor activity	6 (1.1)	$4.20e^{-02}$
		transferase activity	22 (3.9)	$4.60e^{-02}$
		transferase activity, transferring phosphorus-containing groups	15 (2.7)	$4.80e^{-02}$
Pathway	KEGG Pathway	notch signaling pathway	4 (0.7)	$1.50e^{-02}$
Protein Domain	InterPro Name	Zinc finger, PHD-type	8 (1.4)	$1.10e^{-05}$
		Actin-binding, actinin-type	5 (0.9)	$4.50e^{-05}$

Domain	Category	Term	*n* (%) of Genes on List	*p*
		Bromodomain	5 (0.9)	$2.70e^{-04}$
		Zinc finger, ZZ-type	4 (0.7)	$3.20e^{-04}$
		Spectrin repeat	7 (1.3)	$4.10e^{-04}$
		FY-rich, C-terminal	3 (0.5)	$7.50e^{-04}$
		FY-rich, N-terminal	3 (0.5)	$7.50e^{-04}$
		Involucrin repeat	3 (0.5)	$7.50e^{-04}$
		Regulator of chromosome condensation, RCC1	4 (0.7)	$1.00e^{-03}$
		HECT	4 (0.7)	$1.30e^{-03}$
		Calponin-like actin-binding	5 (0.9)	$1.90e^{-03}$
		Protein kinase	12 (2.2)	$2.20e^{-03}$
		Armadillo-like helical	8 (1.4)	$2.90e^{-03}$
		Tyrosine protein kinase	10 (1.8)	$8.10e^{-03}$
		SET-related region	3 (0.5)	$8.90e^{-03}$
		Serine/threonine protein kinase	10 (1.8)	$1.00e^{-02}$
		Cell surface receptor IPT/TIG	3 (0.5)	$2.60e^{-02}$
		Src homology-3	6 (1.1)	$2.70e^{-02}$
		HEAT	4 (0.7)	$3.00e^{-02}$
		Zinc finger, FYVE-type	3 (0.5)	$3.50e^{-02}$
		Klarsicht/ANC-1/syne-1 homology	2 (0.4)	$3.50e^{-02}$
		Initiation factor eIF-4 gamma, MA3	2 (0.4)	$4.20e^{-02}$
		DEAD/DEAH box helicase, N-terminal	4 (0.7)	$4.30e^{-02}$
		Tyrosine protein kinase, active site	4 (0.7)	$4.60e^{-02}$
		Helicase, C-terminal	4 (0.7)	$4.80e^{-02}$
		Concanavalin A-like lectin/glucanase, subgroup	4 (0.7)	$4.90e^{-02}$
		Plectin repeat	2 (0.4)	$4.90e^{-02}$
		eIF4-gamma/eIF5/eIF2-epsilon	2 (0.4)	$4.90e^{-02}$
	PIR Superfamily	SF010337:acute lymphoblastic leukemia protein, ALR type	2 (0.4)	$1.00e^{-02}$
		SF002662:plectin	2 (0.4)	$2.00e^{-02}$