Genome Medicine

Post-mortem molecular profiling of three psychiatric disorders

Ryne C. Ramaker[1,2†], Kevin M. Bowling[1†], Brittany N. Lasseigne[1†], Megan H. Hagenauer[3], Andrew A. Hardigan[1,2], Nicholas S. Davis[1,8], Jason Gertz[1,9], Preston M. Cartagena[4], David M. Walsh[4], Marquis P. Vawter[4], Edward G. Jones[^], Alan F. Schatzberg[5], Jack D. Barchas[6], Stanley J. Watson[3], Blynn G. Bunney[4], Huda Akil[3], William E. Bunney[4], Jun Z. Li[7], Sara J. Cooper[1] and Richard M. Myers[1*]

Abstract

Background: Psychiatric disorders are multigenic diseases with complex etiology that contribute significantly to human morbidity and mortality. Although clinically distinct, several disorders share many symptoms, suggesting common underlying molecular changes exist that may implicate important regulators of pathogenesis and provide new therapeutic targets.

Methods: We performed RNA sequencing on tissue from the anterior cingulate cortex, dorsolateral prefrontal cortex, and nucleus accumbens from three groups of 24 patients each diagnosed with schizophrenia, bipolar disorder, or major depressive disorder, and from 24 control subjects. We identified differentially expressed genes and validated the results in an independent cohort. Anterior cingulate cortex samples were also subjected to metabolomic analysis. ChIP-seq data were used to characterize binding of the transcription factor EGR1.

Results: We compared molecular signatures across the three brain regions and disorders in the transcriptomes of post-mortem human brain samples. The most significant disease-related differences were in the anterior cingulate cortex of schizophrenia samples compared to controls. Transcriptional changes were assessed in an independent cohort, revealing the transcription factor *EGR1* as significantly down-regulated in both cohorts and as a potential regulator of broader transcription changes observed in schizophrenia patients. Additionally, broad down-regulation of genes specific to neurons and concordant up-regulation of genes specific to astrocytes was observed in schizophrenia and bipolar disorder patients relative to controls. Metabolomic profiling identified disruption of GABA levels in schizophrenia patients.

Conclusions: We provide a comprehensive post-mortem transcriptome profile of three psychiatric disorders across three brain regions. We highlight a high-confidence set of independently validated genes differentially expressed between schizophrenia and control patients in the anterior cingulate cortex and integrate transcriptional changes with untargeted metabolite profiling.

Keywords: Schizophrenia, Bipolar disorder, Major depressive disorder, RNA sequencing, Metabolomics, *EGR1*

* Correspondence: rmyers@hudsonalpha.org
†Equal contributors
^Deceased
[1]HudsonAlpha Institute for Biotechnology, 601 Genome Way, Huntsville, AL 35806, USA
Full list of author information is available at the end of the article

Ramaker *et al. Genome Medicine* (2017) 9:72

Page 2 of 12

Background

Schizophrenia (SZ), bipolar disorder (BPD), and major depressive disorder (MDD) are multigenic diseases with complex etiology and are large sources of morbidity and mortality in the population. All three disorders are associated with high rates of suicide, with ~90% of the ~41,000 people who commit suicide each year in the US having a diagnosable psychiatric disorder [1]. Notably, while clinically distinct, these disorders also share many symptoms, including psychosis, suicidal ideation, sleep disturbances, and cognitive deficits [2–4]. This phenotypic overlap suggests potential common genetic etiology, which is supported by recent large-scale genome-wide association studies (GWAS) [5–8]. However, this overlap has not been fully characterized with functional genomic approaches. Current therapies for these psychiatric disorders are ineffective in many patients and often only treat a subset of an individual patient's symptoms [9]. Approaches targeting the underlying molecular pathologies within and across these types of disorders are necessary to address the immense burden of psychiatric disease around the world and improve care for the millions of people diagnosed with these conditions.

Previous studies [10–14] analyzed brain tissue with RNA sequencing (RNA-seq) in SZ and BPD, and identified altered expression of GABA-related genes in the superior temporal gyrus and hippocampus, as well as differentially expressed genes related to neuroplasticity and mammalian circadian rhythms. Our study focused on the anterior cingulate cortex (AnCg), dorsolateral prefrontal cortex (DLPFC), and nucleus accumbens (nAcc) regions, which are often associated with mood alterations, cognition, impulse control, motivation, reward, and pleasure—all behaviors known to be altered in psychiatric disorders [15, 16]. To assess gene expression changes associated with psychiatric disease in these three brain regions, we performed RNA-seq on macro-dissected post-mortem tissues in four well-documented cohorts of 24 patients each with SZ, BPD, and MDD and 24 controls (CTL) (96 individuals total). Additionally, we conducted metabolomic profiling of AnCg tissue from the same subjects. RNA-seq analysis revealed common expression profiles in SZ and BPD patients, supporting the notion that these disorders share a common molecular signature. Transcriptional changes were most pronounced in the AnCg, with SZ and BPD exhibiting strongly correlated differences from CTL samples. Differentially expressed genes were associated with cell-type composition, with BPD and SZ samples showing decreased expression of neuron-specific genes. We validated this result with RNA-seq data from an independent cohort of 35 cases each of SZ and BPD and CTL post-mortem cingulate cortex samples from the Stanley Neuropathology Consortium Integrative Database (SNCID; http://sncid.stanleyresearch.org) Array Collection. We present a set of validated genes differentially expressed between SZ and CTL patients, perform an integrated analysis of metabolic pathway disruptions, and highlight a role for the transcription factor *EGR1*, whose down-regulation in SZ patients may drive a large portion of observed transcription changes.

Methods

See Additional file 1: Supplemental methods for additional detail.

Patient sample collection and preparation

Sample collection, including human subject recruitment and characterization, tissue dissection, and RNA extraction, was described previously [17, 18] as part of the Brain Donor Program at the University of California, Irvine, Department of Psychiatry and Human Behavior (Pritzker Neuropsychiatric Disorders Research Consortium) under institutional review board approval. In brief, coronal slices of the brain were rapidly frozen on aluminum plates that were previously frozen to –120 °C and dissected as described previously [19]. All samples were diagnosed by psychological autopsy, which included collection and analyses of medical and psychiatric records, toxicology, medical examiners' reports, and 141-item family interviews. Agonal state scores were assigned based on a previously published scale [20]. Controls were selected based on absence of severe psychiatric disturbance and mental illness within first-degree relatives.

We obtained fastq files from RNA-seq experiments for our validation cohort from the SNCID (http://sncid.stanleyresearch.org) Array Collection comprising 35 cases each of SZ, BPD, and CTL of post-mortem cingulate cortex with permission on 30 June 2015. For our analysis, we included the 27 SZ, 26 CTL, and 25 BPD SNCID samples that were successfully downloaded and represented unique samples. SNCID RNA-seq methodology and data processing are described in detail in a previous publication that makes use of the data [10].

RNA-seq and data processing

To extract nucleic acid, 20 mg of post-mortem brain tissue was homogenized in Qiagen RLT buffer + 1% BME using an MP FastPrep-24 and Lysing Matrix D beads for three rounds of 45 s at 6.5 m/s (FastPrep homogenizer, lysing matrix D, MP Bio). Total RNA was isolated from 350 μL of tissue homogenate using the Norgen Animal Tissue RNA Purification Kit (Norgen Biotek Corporation). We made RNA-seq libraries from 250 ng total RNA using poly(A) selection (Dynabeads mRNA DIRECT kit, Life Technologies) and transposase-based non-stranded library construction (Tn-RNA-seq) as described previously [21]. To mitigate potentially confounding batch affects in sample preparation we randomly assigned

Ramaker *et al. Genome Medicine* (2017) 9:72

Page 3 of 12

samples from all brain regions and disorders into batches of 24 samples. We used KAPA to quantify the library concentrations and pooled four samples in order to achieve equal concentration of the four libraries in each lane. Pools were determined by random from the 291 samples. Samples were also randomly selected for pooling in an effort to limit potentially confounding sequencing batch effects. The pooled libraries were sequenced on an Illumina HiSeq 2000 sequencing machine using paired-end 50-bp reads and a 6-bp index read, resulting in an average of 48.2 million reads per library. To quantify the expression of each gene in both Pritzker and SNCID datasets, RNA-seq reads were processed with aRNApipe v1.1 using default settings [22]. Briefly, reads were aligned and counted with STAR v2.4.2a to all genes annotated in GRCh37_E75 [23]. All alignment quality metrics were obtained from the picard tools module (http://broadinstitute.github.io/picard/) available in aRNApipe. Genes expressed from the X and Y chromosomes were omitted from the study.

Quantitative PCR (qPCR) was performed on ten SZ and ten CTL patients to validate *EGR1* RNA-seq measurements. RNA was extracted as described above from tissue lysates a second time. Reverse transcription was performed on 250 ng of input RNA with the Applied Biosystems high capacity cDNA reverse transcription kit. Validated Taqman assays for *EGR1* (Hs00152928_m1) and the housekeeper genes *GAPDH* (Hs02758991_g1) and *ACTB* (Hs01060665_g1) were used for qPCR. cDNA was diluted by a factor of 10 before use as input for the Taqman assay. The qPCR reaction was performed on an Applied Biosystems Quant Studio 6 Flex system using the recommended amplification protocol for Taqman assays.

Sequencing data analysis
All data analysis in R was performed with version 3.1.2.

Differential expression analysis and normalization
To examine gene expression changes, we employed the R package DESeq2 [24] (version 1.6.3), using default settings, but employing likelihood ratio test (LRT) hypothesis testing, and removing non-convergent genes from subsequent analysis. Genes differentially expressed between each disorder and CTL samples, by brain region, were identified with DESeq2 (adjusted p value <0.05), including age, brain pH, post-mortem interval (PMI), and percentage of reads uniquely aligned (PRUA) as covariates (full model, ~Age + PMI + pH + PRUA + Disorder; reduced model, ~Age + PMI + pH + PRUA). For downstream heatmap visualization, principal component analysis (PCA), and cell-type analysis, genes underwent a log-like normalization using DESeq2's varianceStabilizingTransformation function and were corrected for PRUA by computing residuals to a linear model regressing PRUA on normalized gene expression level with

the R lm function unless otherwise specified. DESeq2's default independent filtering method was used to remove genes with an insufficient expression level from further analysis.

PCA and hierarchical clustering
PCA analysis was performed in R on normalized data using the prcomp() command. Hierarchical clustering of normalized gene expression data was done in R with the hclust command (method = "ward", distance = "Euclidean")

Pathway enrichment analysis
Pathway analysis was conducted using the web-based tool LRPath [25] using all gene ontology (GO) term annotations, adjusting to gene read count with RNA-Enrich, including directionality and limiting maximum GO term size to 500 genes. GO term visualization was performed using the Cytoscape Enrichment Map plug-in [26]. The Genesetfile (.gmt) GO annotations from 1 February 2017 were downloaded from http://download.baderlab.org/EM_Genesets/. The LRPath output was parsed and used as an enrichment file with all upregulated pathways colored red and all downregulated pathways colored blue, regardless of degree of upregulation. Mapping parameters were set as p value cutoff = 0.005, false discovery rate (FDR) cutoff = 0.1, and Jaccard coefficient >0.3. Resulting networks were exported as PDFs. Summary terms were added to the plot based on the GO terms in those clusters. In order to assess overlap between significant GO terms in our analysis and the GWAS described by the Psychiatric Genomics Consortium [5], we downloaded the p values reported for SZ hits from their Supplemental Table 4, which contained 424 significant GO terms. We used a chi-squared test to assess significant overlap between the two groups. We report the p values measured in SZ based on this study along with those calculated in our analysis.

EGR1 ChIP-seq peak analysis
Narrow peak bed files filtered to optimal Irreproducible Discovery Rate (IDR) peaks were obtained from the ENCODE data portal (https://www.encodeproject.org/) for *EGR1* ChIP-seq data in GM12878, H1-hESC, and K562 cell lines (ENCODE file IDs ENCFF002CIV, ENCFF002CGW, ENCFF002CLV). Consensus *EGR1* peaks were identified by intersecting peaks from all three cell lines, which resulted in a final list of 4121 peaks common to all cell lines (minimum overlap of 1 bp). The distance from each annotated transcription start site (TSS) to the nearest consensus *EGR1* peak was computed based on TSSs annotated in the ENSEMBL gene transfer format (GTF) file from the Ensembl data release 75 (GRCh37_E75).

Ramaker *et al. Genome Medicine* (2017) 9:72

Page 4 of 12

Cell-specific enrichment analysis

Sets of genes uniquely expressed by several brain cell types were obtained from Fig. 1b in Darmanis et al. [27]. An index for each cell type was created by calculating the median normalized expression value for each set of cell type-associated genes. Index values were compared across patient clusters by non-parametric rank sum tests and Spearman correlation with top principal components. To validate our method, we calculated cell type-specific indices from an independent cohort of previously published purified brain cells [28, 29]. FPKM-normalized gene expression data were obtained from Supplemental Table S4 of Zhang et al. (2015) [28] and cell type indices were calculated as described above. To examine index performance in mixed cell populations, we obtained fastq files for neuron and astrocyte-purified brain samples from Gene Expression Omnibus (GEO) accession GSE73721 and generated raw count files as described above. We next mixed expression profiles in silico by performing random down-sampling of neuron and astrocyte count levels and summing the results such that mixed populations containing specific proportions of counts from neuron- and astrocyte-purified tissue were generated. For example, to generate an 80:20 neuron to astrocyte mixture, neuron and astrocyte count columns (which started at an equivalent number of 5,759,178 aligned reads) were randomly down-sampled to 4,607,342 and 1,151,836 counts, respectively, and summed across each gene to result in a proportionately mixed population of aligned count data simulating heterogeneous tissue. Then we calculated a neuron:astrocyte index ratio capable of predicting the in silico mixing weights. Briefly, we assumed index values for mixed cell populations were directly proportional to mixing weights of their respective purified tissue; thus, the predicted cell proportion for a given cell type was simply calculated as Predicted cell proportion = Observed index value/Purified tissue index value.

To ensure cell type predictive power was unique to indices derived from the genes in Darmanis et al. [27], we generated indices from 10,000 randomly sampled gene sets of equivalent size and examined their performance in predicting in silico mixing weights. Mean squared prediction errors (MSE) were calculated for each of the 10,000 null indices and compared to the MSE of Darmanis et al.-derived indices.

Cell type deconvolution analysis was confirmed using a previously published algorithm implemented in the R package deconRNAseq [30]. The "datasets" input to the deconRNAseq function was a normalized count matrix of all AnCg brain samples and the "signatures" input consisted of a normalized count matrix of astrocyte, neuron, microglia, and oligodendrocyte dissected cells from the GEO accession GSE73721 previously described.

Enrichment analysis for extreme fold change was performed by randomly sampling the fold changes of 1000 null gene sets equivalent in size and expression level (allowing 5% error) to the neuron- and astrocyte-specific gene sets. The median fold change of each 1000 null gene set was compared to the observed median fold change for neuron and astrocyte gene sets, respectively.

Metabolomics

Sample preparation

Sections of approximately 100 mg of frozen tissue were weighed and homogenized for 45 s at 6.5 M/s with ceramic beads in 1 mL of 50% methanol using the MP FastPrep-24 homogenizer (MP Biomedicals). A sample volume equivalent to 10 mg of initial tissue weight was dried down at 55 °C for 60 minutes using a vacuum concentrator system (Labconco). Derivatization by methoximation and trimethylsilylation was done as previously described [31].

We analyzed technical replicates of each tissue sample, in randomized order.

GCxGC-TOFMS analysis

All derivatized samples were analyzed on a Leco Pegasus 4D system (GCxGC-TOFMS), controlled by the ChromaTof software (Leco, St. Joseph, MI, USA). Samples were analyzed as described previously [31] with minor modifications in temperature ramp.

Data analysis and metabolite identification

Peak calling, deconvolution, and library spectral matching were done using ChromaTOF 4.5 software. Peaks were identified by spectral match using the NIST, GOLM [32], and Fiehn libraries (Leco) and confirmed by running derivatized standards (Sigma). We used Guineu for multiple sample alignment [33].

Integrated pathway analysis

Altered metabolites and genes were analyzed for enrichment in KEGG pathways containing both metabolite and gene features. A non-parametric, threshold-free pathway analysis similar to that of a previously described method [34] was first performed on metabolite and gene expression data separately. Our method builds on the principle described by Subramanian that implements a one-tailed Wilcox test to identify pathways enriched for low p values. Instead of just accounting for enrichment at the gene level, we use metabolite or gene p value ranks within each pathway compared to remaining non-pathway metabolites or genes with a one-tailed Wilcox test to test the hypothesis that elements of a given pathway may be enriched for lower p value ranks than background elements. Metabolite and gene p values were subsequently combined to provide an integrated enrichment significance p value using Fisher's method. Pathways had to contain greater than five genes and one

Ramaker *et al. Genome Medicine* (2017) 9:72

Page 5 of 12

metabolite measured in our dataset to be included in the analysis. Additional file 2: Table S10 lists *p* values for enriched pathways based on genes, metabolites, or both combined.

Results

Region-specific gene expression in control and psychiatric brain tissue

We collected post-mortem human brain tissue, associated clinical data including age, sex, brain pH, and post-mortem interval (PMI), and cytotoxicology results (Additional file 2: Tables S1 and S2) for matched cohorts of 24 patients each diagnosed with SZ, BPD, or MDD, as well as 24 control individuals with no personal history of, or first-degree relatives diagnosed with, psychiatric disorders. Importantly, to limit the effect of acute patient stress at the time of death as a potential confounder, we included only patients with an agonal factor score of zero and a minimum brain pH of 6.5 [18]. Using RNA-seq [21], we profiled gene expression in three macro-dissected brain regions (AnCg, DLPFC, nAcc). After quality control, we analyzed 57,905 ENSEMBL genes in a total of 281 brain samples (Additional file 2: Table S3).
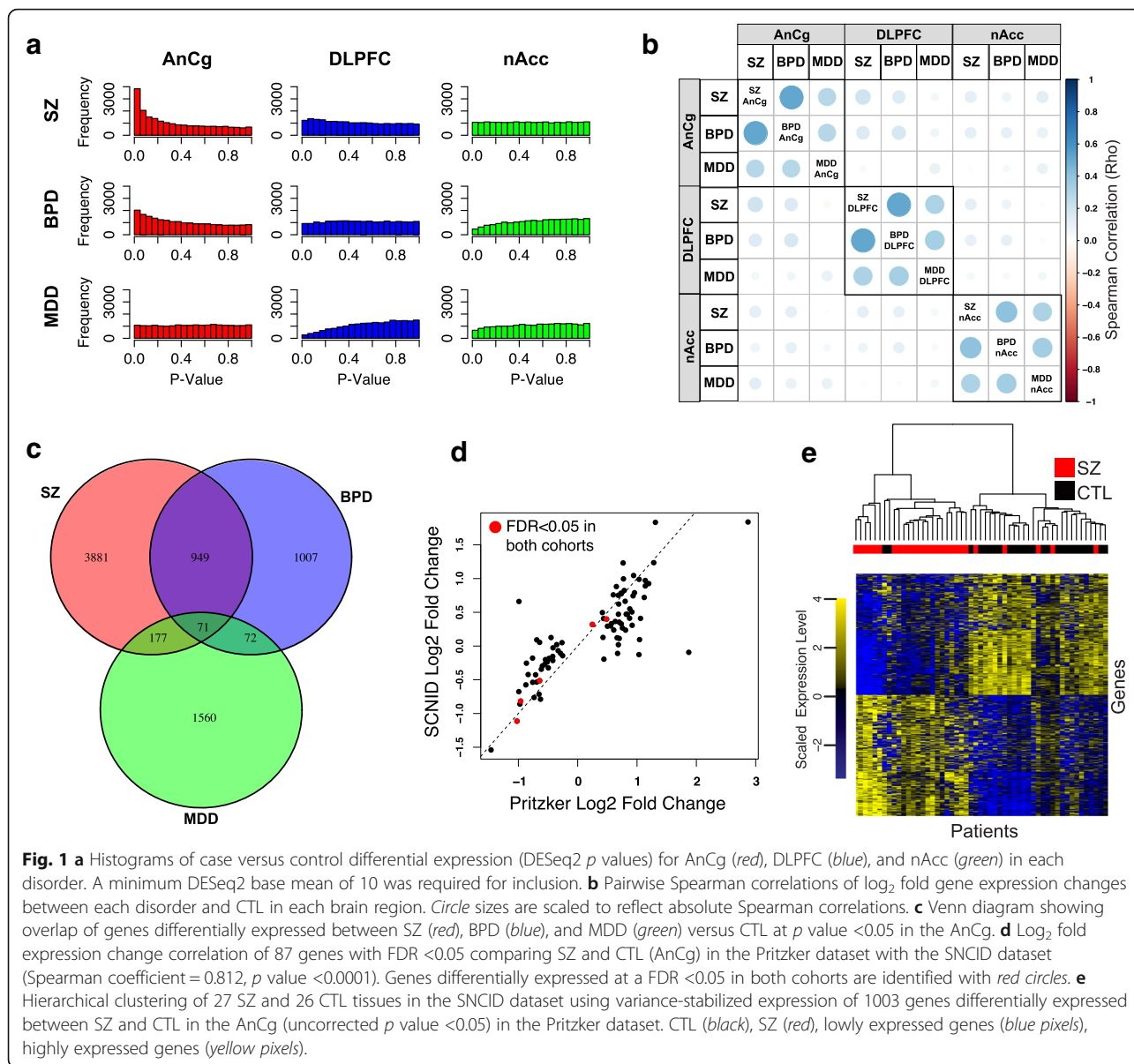
To examine heterogeneity across brain regions and subjects, we performed a principal component analysis (PCA; Additional file 3: Figure S1a) of all genes. The first principal component (PC1, 21.8% of the variation) separates cortical AnCg and DLPFC samples from subcortical nAcc samples. Examination of the first and second principal components for disorder associations reveals a separation of some SZ and BPD samples from all other samples (Additional file 3: Figures S1b and S2a–c). However, in agreement with previously reported post-mortem brain RNA sequencing studies [14], we found several principal components to be highly correlated with quality metrics, including the percentage of reads uniquely aligned and percentage of reads aligned to mitochondrial sequence (absolute Rho >0.5, FDR <1E-16; Additional file 2: Table S4). To reduce the potentially confounding effects of sample quality, we repeated the PCA on expression data normalized to the percentage of reads uniquely aligned for each sample and found that global disease-specific expression differences were significantly reduced and PC1 primarily separated nAcc samples from AnCg and DLPFC brain regions (Additional file 3: Figures S1c and S2d–i).

Disease-specific gene expression in control and psychiatric brains

We next applied DESeq2 [24], a method for analyzing differential sequence read count data, to identify genes differentially expressed across disorders within each brain region after correcting for biological and technical covariates. The largest number of significant expression changes occurred in AnCg between SZ and CTL individuals (87 genes, FDR <0.05; Fig. 1a). Pathway enrichment analysis of differentially expressed genes between SZ and CTL patients revealed 935 GO terms with a FDR <0.05 (Additional file 2: Table S5; 122 Gene Ontology Cell Compartment (GOCC) , 159 Gene Ontology Molecular Function (GOMF), and 654 Gene Ontology Biological Process (GOBP)). Significant GO terms fall into the broad categories of synaptic function and signaling (e.g., neurotransmitter transport, ion transport, calcium signaling; Additional file 3: Figure S3). These terms overlap significantly with those identified by the Psychiatric Genomics Consortium in their analysis of GWAS implicated genes [35], with 68 GO terms meeting a *p* value cutoff of <0.05 in both datasets (*p* value <0.0001, Chi-square test). Additionally, nine genes were differentially expressed between SZ and CTL individuals in DLPFC. Three of these were also identified in AnCg: *SST*, *PDPK2P*, and *KLHL14*. No genes had a FDR <0.05 when comparing BPD or MDD samples to CTLs in any brain region, or comparing SZ and CTL tissues in nAcc (Additional file 2: Table S6). To examine potential common gene expression patterns between the psychiatric disorders, we performed pair-wise correlation calculations of all gene $\log_2$ fold changes for each disorder versus controls in each brain region. Of the nine case-control comparisons (for three regions and three diseases), a particularly strong correlation is observed between BPD and SZ compared to either SZ or BPD and MDD in each brain region (Fig. 1b). In the AnCg, BPD and SZ share 1020 common genes differentially expressed at an uncorrected DESeq2 *p* value <0.05 compared to only 248 and 143 genes shared between MDD and SZ or BPD, respectively (Fig. 1c). This strong overlap between BPD and SZ (Fisher's exact *p* value <1E-16) indicates that although expression changes are weaker in BPD, they follow a trend similar to those identified in SZ.
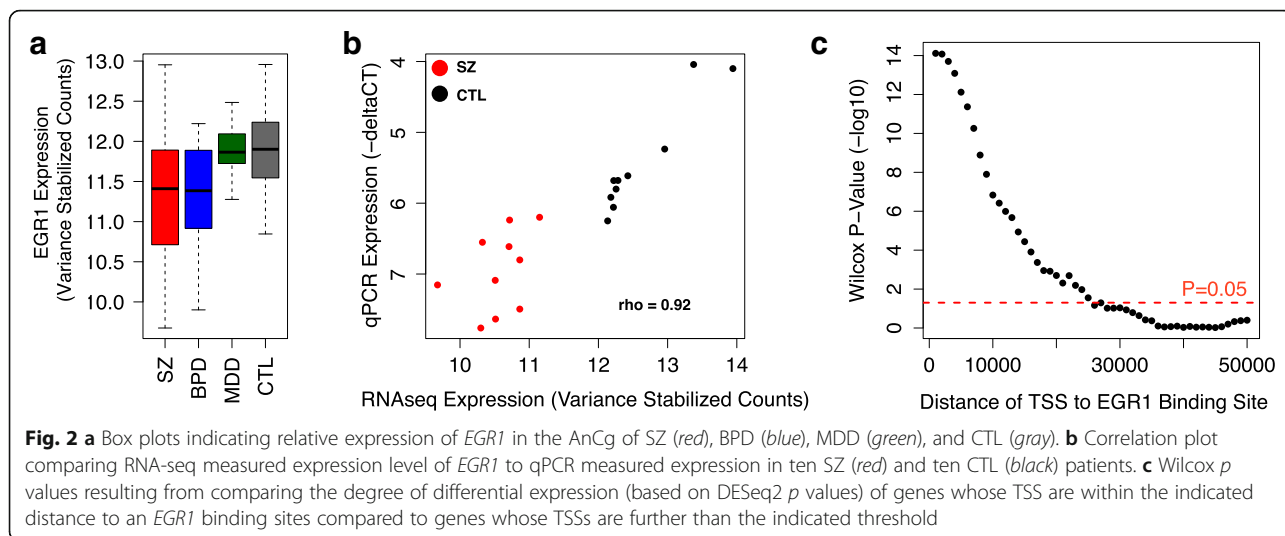
Because previous post-mortem analyses have been limited by, and are particularly vulnerable to, biases inherent to examining a single patient cohort, we sought to generate a robust set of SZ-associated genes by validating our observed expression changes in an independent cohort. To accomplish this, we examined gene expression differences in the AnCg between SZ and CTL samples in the SNCID RNA-seq Array dataset [13], revealing 1003 genes altered (DESeq2 uncorrected *p* value <0.05) in both datasets (Fisher's *p* value <1E-16; Additional file 2: Table S7). The magnitude and direction of change in significant genes in the Pritzker dataset were highly correlated with the SNCID dataset (Rho = 0.202, *p* value <1E-16), particularly in 87 genes that met a cutoff FDR of <0.05 (Rho = 0.812, *p* value <1E-16; Fig. 1d). We

Ramaker *et al. Genome Medicine* (2017) 9:72

Page 6 of 12

Fig. 1 a Histograms of case versus control differential expression (DESeq2 *p* values) for AnCg (*red*), DLPFC (*blue*), and nAcc (*green*) in each disorder. A minimum DESeq2 base mean of 10 was required for inclusion. **b** Pairwise Spearman correlations of log$_2$ fold gene expression changes between each disorder and CTL in each brain region. *Circle* sizes are scaled to reflect absolute Spearman correlations. **c** Venn diagram showing overlap of genes differentially expressed between SZ (*red*), BPD (*blue*), and MDD (*green*) versus CTL at *p* value <0.05 in the AnCg. **d** Log$_2$ fold expression change correlation of 87 genes with FDR <0.05 comparing SZ and CTL (AnCg) in the Pritzker dataset with the SNCID dataset (Spearman coefficient = 0.812, *p* value <0.0001). Genes differentially expressed at a FDR <0.05 in both cohorts are identified with *red circles*. **e** Hierarchical clustering of 27 SZ and 26 CTL tissues in the SNCID dataset using variance-stabilized expression of 1003 genes differentially expressed between SZ and CTL in the AnCg (uncorrected *p* value <0.05) in the Pritzker dataset. CTL (*black*), SZ (*red*), lowly expressed genes (*blue pixels*), highly expressed genes (*yellow pixels*).

performed hierarchical clustering of SZ and CTL samples in the SNCID validation cohort using the 1003 genes differentially expressed, at the less stringent threshold, *p* value <0.05, between SZ and CTL in the Pritzker dataset (Fig. 1e), and found these genes successfully distinguished the two disease groups with only 5 out of 27 SZ and 2 out of 26 CTL samples misclassified.

Of particular interest are the five genes significant at a FDR <0.05 in both cohorts, including a nearly twofold decrease in expression of the transcription factor *EGR1* (Additional file 2: Table S7a; Fig. 2a). qPCR validation confirmed reduced *EGR1* expression in SZ samples (Fig. 2b; Wilcox *p* value = 4.33 × 10E-5). *EGR1*, a zinc finger transcription factor, has been recently implicated in SZ by a GWAS study [5]; thus, we sought to investigate

whether loss of *EGR1* expression might be associated with transcriptional changes observed in the AnCg of SZ patients using publicly available genome-wide occupancy data from the ENCODE consortium (https://www.encodeproject.org). To obtain high confidence *EGR1* binding sites we intersected chromatin immunoprecipitation sequencing (ChIP-Seq) peaks derived from the H1-hESC, K562, and GM12878 cell lines. We found that genes with a transcription start site (TSS) within 1 kb of an *EGR1* binding site had significantly lower DESeq2 *p* values (Wilcox *p* value = 9.68E-5) and reduced expression in SZ versus CTL (Wilcox *p* value = 7.69E-15) compared to genes whose TSSs were greater than 1 kb from an *EGR1* binding site. A monotonic decrease in this effect was observed as the distance threshold used for this comparison was increased from 1 to 50 kb (Fig. 2c).
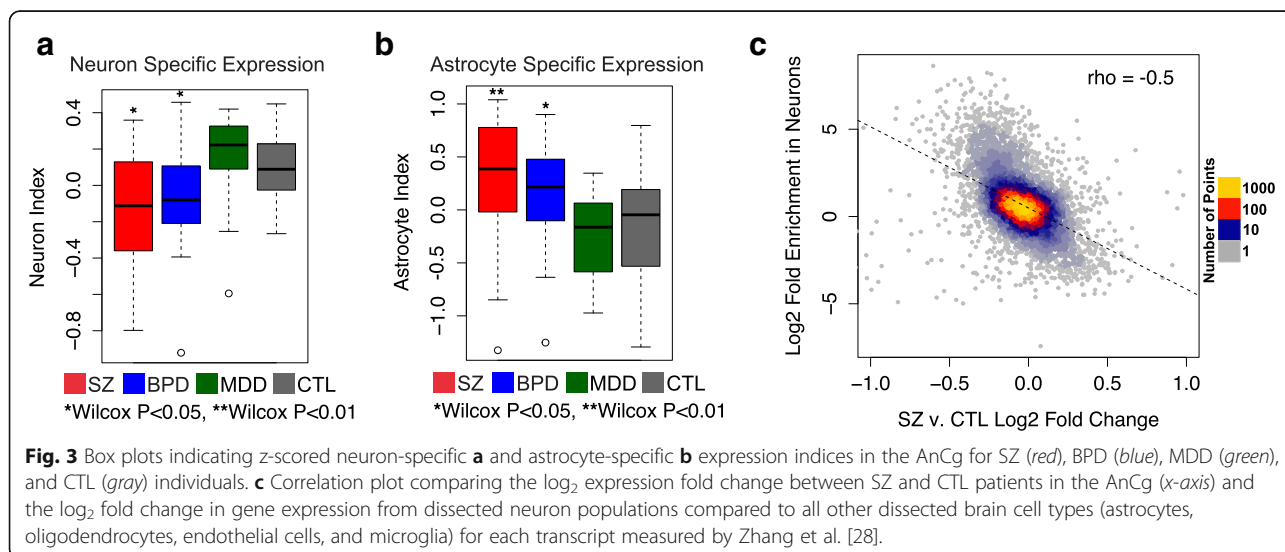
Ramaker *et al. Genome Medicine* (2017) 9:72

Page 7 of 12

Fig. 2 a Box plots indicating relative expression of *EGR1* in the AnCg of SZ (*red*), BPD (*blue*), MDD (*green*), and CTL (*gray*). **b** Correlation plot comparing RNA-seq measured expression level of *EGR1* to qPCR measured expression in ten SZ (*red*) and ten CTL (*black*) patients. **c** Wilcox *p* values resulting from comparing the degree of differential expression (based on DESeq2 *p* values) of genes whose TSS are within the indicated distance to an *EGR1* binding sites compared to genes whose TSSs are further than the indicated threshold

Cell type-specific changes

In addition to dysregulation of broadly acting transcription factors, another mechanism that can drive large-scale transcriptional changes in bulk tissue is alterations in constituent cell type proportions. Previous studies have observed decreases in neuron density and increased glial scarring in psychiatric disorders [36, 37]. To test for signs of changing cell populations in our dataset we applied a method to deconvolute RNA expression data and estimate cell type proportions. Darmanis et al. [27] identified genes capable of classifying cells into the major neuronal, glial, and vascular cell types in the brain based on single cell RNA sequencing. We used these gene sets to generate cell type indices using the median of normalized counts for each cell type-specific gene set. We tested these indices on purified brain cell populations and in silico mixed cell populations from

Zhang et al. [28, 29] to demonstrate their accuracy and specificity (Additional file 3: Figure S4).

Application of these cell type indices to patient AnCg expression data revealed a significant decrease in neuron specific gene expression (Wilcox *p* value <0.05) and a significant increase in astrocyte-specific expression (Wilcox *p* value <0.05) in SZ and BPD patients compared to controls (Fig. 3a, b). Other brain cell type indices were not significantly different between psychiatric patients and controls (Additional file 3: Figure S5). An alternative algorithm for cell type deconvolution, DeconRNASeq, showed similar results (Additional file 3: Figure S6a, b).

Additionally, we showed that neuron-specific genes identified by Darmanis et al. [27] are enriched for decreased expression in SZ compared to controls and astrocyte-specific genes are enriched for increased expression (Additional file 3: Figure S6c). Again, these enrichments are specific to this

Fig. 3 Box plots indicating z-scored neuron-specific **a** and astrocyte-specific **b** expression indices in the AnCg for SZ (*red*), BPD (*blue*), MDD (*green*), and CTL (*gray*) individuals. **c** Correlation plot comparing the log$_2$ expression fold change between SZ and CTL patients in the AnCg (*x-axis*) and the log$_2$ fold change in gene expression from dissected neuron populations compared to all other dissected brain cell types (astrocytes, oligodendrocytes, endothelial cells, and microglia) for each transcript measured by Zhang et al. [28].

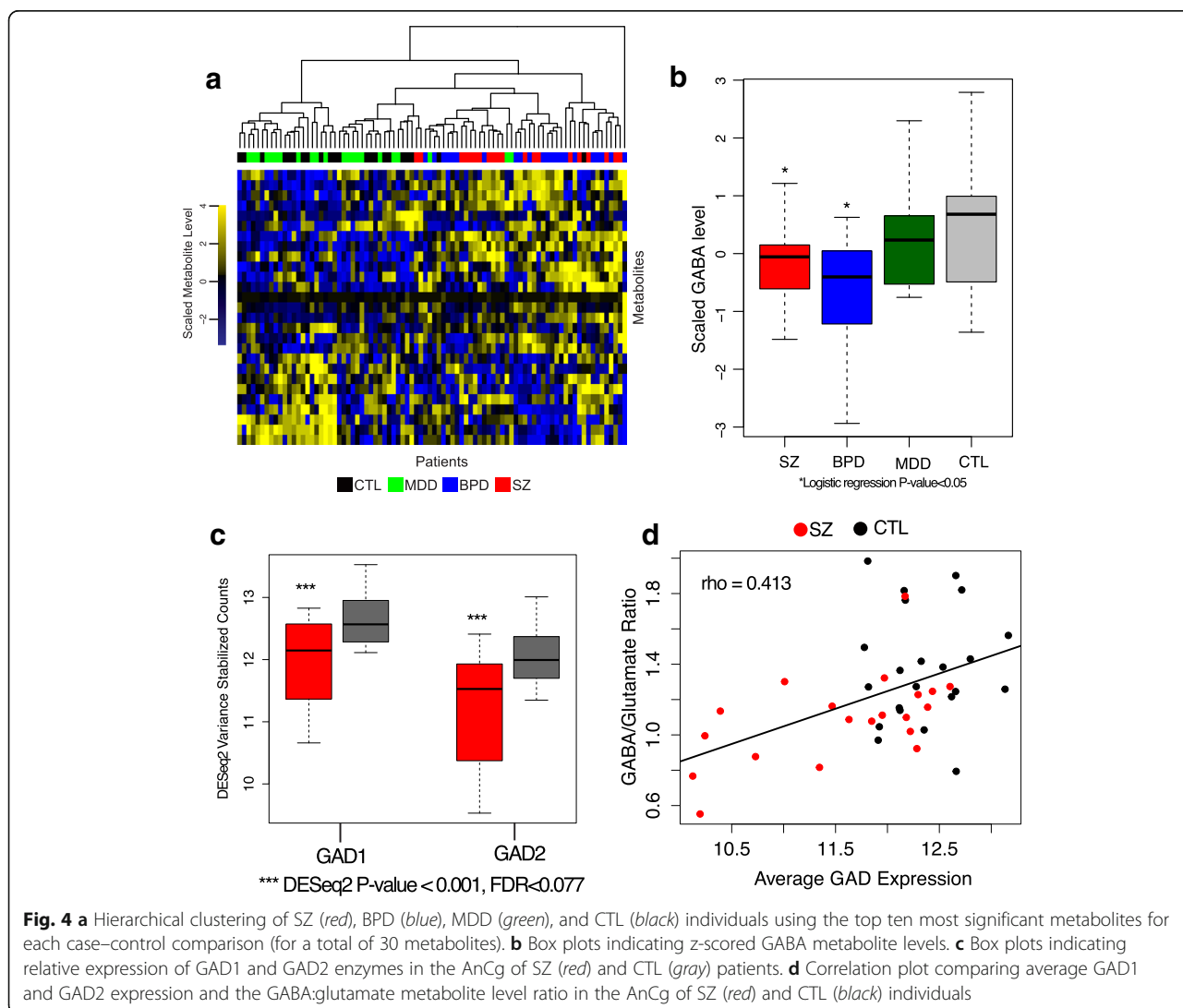Ramaker *et al. Genome Medicine* (2017) 9:72

Page 8 of 12

gene set and are not reproduced by 1000 expression-matched, randomly sampled gene sets (Additional file 3: Figure S6d,e). Further supporting a decrease in neuronal gene expression, we found a significant negative correlation between gene expression changes in patient brains relative to control brains and the degree of neuron-specific transcription (fold enrichment of neuronal gene expression over other cell types; SZ Rho = –0.50 and BPD Rho = –0.41, *p* value <1E-16; SZ shown in Fig. 3c).

Transcriptomic changes reflected in altered metabolomic profiles

To assess the biochemical consequences of expression changes, we used 2D-GCMS to measure metabolite levels in 86 of the AnCg samples (sufficient tissue was unavailable for ten samples). We measured and identified 141 unique metabolites (Additional file 2: Table S8). We found no metabolites reached statistical significance (FDR <0.05); however, eight metabolites had a FDR <0.1 when

comparing SZ to CTL. Similar to our gene expression analysis, metabolite levels (Additional file 2: Table S9) successfully differentiated SZ and BPD patients from CTLs (Fig. 4a), while MDD metabolite profiles were very similar to CTLs. Several of the most significant metabolites, including GABA, are known to be relevant to BPD and SZ (Fig. 4b) [38]. Furthermore, GABA:glutamate metabolite ratios correlate strongly with average *GAD1* and *GAD2* expression levels measured by RNA-seq (Rho = 0.413, *p* value = 0.007; Fig. 4c, d). This metabolite–gene relationship is consistent with previous multi-level phenomic analyses [39] and demonstrates realized biochemical consequences from altered gene expression. Notably, reductions in GABA could coincide with loss of neuron-specific gene expression as suggested by the RNA-seq data. Integrated pathway analyses of metabolite and gene expression data revealed disruption of synaptic and neurotransmitter signaling in SZ compared to CTL (Additional file 2: Table S10 and Additional file 3: Figure S7).

Fig. 4 a Hierarchical clustering of SZ (*red*), BPD (*blue*), MDD (*green*), and CTL (*black*) individuals using the top ten most significant metabolites for each case–control comparison (for a total of 30 metabolites). **b** Box plots indicating z-scored GABA metabolite levels. **c** Box plots indicating relative expression of GAD1 and GAD2 enzymes in the AnCg of SZ (*red*) and CTL (*gray*) patients. **d** Correlation plot comparing average GAD1 and GAD2 expression and the GABA:glutamate metabolite level ratio in the AnCg of SZ (*red*) and CTL (*black*) individuals

Ramaker *et al. Genome Medicine* (2017) 9:72

Page 9 of 12

Discussion

Here, we describe a large transcriptomic dataset across three brain regions (DLPFC, AnCg, and nAcc) in SZ, BPD, and MDD patients, as well as CTL samples matched for agonal state and brain pH. In MDD, we did not identify any genes that meet genome-wide significance for differential expression between cases and controls in any brain region. This finding agrees with previous post-mortem RNA-seq studies [40]; however, sample size and the choice of brain regions examined likely contributed to our inability to replicate results from previous non-transcriptome-wide sequencing-based approaches comparing MDD to CTL in post-mortem brain [41]. One limitation of our study is that females are underrepresented at a rate of about 5:1. This reflects the increased chance of accidental death among males [42], but limits us in our ability to make more general conclusions about these disorders and to address known differences between the sexes as they relate to these disorders. We also do not have information on the smoking status for our cohort, which is an important covariate as smoking rates are higher among patients with psychiatric disorders and smoking has been demonstrated to affect gene expression [43, 44]. Another potential limitation inherent to post-mortem cohort analyses is accounting for patient drug use. As detailed in Additional file 2: Table S2, patient toxicology reports were positive for several prescribed and illicit drugs that were not present in CTL samples. As this is a bias inherent to psychiatric patients, it is impossible to disentangle from non-treatment-related disease patterns in a post-mortem analysis.

Another important limitation of post-mortem RNA-seq studies is RNA quality. We found a significant proportion of variation in our data to be associated with multiple alignment quality metrics. Significant effort went into controlling for potential sources of bias due to differences in RNA quality. We only included tissue from patients with an agonal score of 0 and who had a brain pH of 6.5 or greater. We also controlled for brain pH, post-mortem interval, and alignment quality in all differential expression analyses. Our study, as well as future post-mortem studies, could be improved by directly measuring RNA quality at the time of sample preparation (e.g., RNA integrity number (RIN)). Even with these caveats, we believe our data yield new insights contributing to a growing understanding of these disorders.

The most dramatic gene expression signals we observed were brain region-specific. The majority of disease-associated expression differences were seen in the AnCg of SZ compared to CTL individuals. The AnCg has been associated with multiple disease-relevant functions, including cognition, error detection, conflict resolution, motivation, and modulation of emotion [45–47]. We observed a striking overlap in SZ- and BPD-associated expression changes consistent with previous findings [38, 48].

One of the more intriguing genes significantly down-regulated (FDR <0.05) in both cohorts of SZ patients was the zinc finger transcription factor *EGR1*. We provide evidence that this factor binds upstream of genes with altered expression in SZ and is associated with decreased expression in SZ patients. Down-regulation of *EGR1* has been previously described in the prefrontal cortex of post-mortem brain samples from SZ patients [49, 50]. *EGR1* has also previously been associated with several phenotypes relevant to psychiatric disorders, including neural differentiation [51], emotional memory formation [52], and response to antipsychotics [53], and has recently been described as part of a transcription factor–miRNA co-regulatory network capable of acting as a biomarker in peripheral blood cells for SZ [54]. In mice, loss of *EGR1* has been linked to neuronal loss in a model of Alzheimer's disease [55]. *EGR1* is also important for regulation of the NMDA receptor pathway, which is critical for synaptic plasticity and memory formation and has been implicated in SZ in humans [56]. We believe a more detailed examination of genome-wide *EGR1* occupancy in post-mortem brain tissue or cultured neurons could yield additional information and assessment of the functional consequences of *EGR1* loss is required to confirm this factor's role in SZ pathogenesis.

We also see evidence for depletion of neuron-specific genes and increased levels of astrocyte-specific genes in SZ and BPD patients. This observation is further supported by metabolomic analysis of the AnCg, which found a concordant decrease in GABA levels in BPD and SZ individuals. Neuronal depletion has been previously described in SZ [36, 37]. Insufficient tissue remains from our patient cohort to validate computational cell type predictions immunohistochemically; however, our data strongly suggest that future post-mortem studies should be cognizant of cell type heterogeneity across patient samples. The method for cell type composition estimation is limited in its accuracy to estimating only the major classes of cells present. Genes represented in cell types present at only a small minority could be over- or under-represented using this technique. Based on these results, future studies should consider using robust techniques for assessing tissue composition to examine potential cell type proportion differences between disease cohorts and to identify which transcriptional changes occur in conjunction with, and independent of, those differences.

We observed very few or no significant expression differences in the DLPFC and nAcc, which contradicts several previous studies [14, 57]. We do not intend to claim that no transcriptional changes occur in these

Ramaker *et al. Genome Medicine* (2017) 9:72

Page 10 of 12

brain regions as our study was designed to broadly compare transcriptional alterations across multiple brain regions in multiple psychiatric disorders, thereby sacrificing exceptional sample sizes in any single disorder in any specific brain region. However, our data do suggest that, of the regions we tested, the strongest transcriptional changes occur in the AnCg of SZ patients. Moreover, these data provide a useful resource for future studies facilitating the testing of preliminary hypotheses or validation of significant findings.

Conclusions

Our study provides several meaningful and novel contributions to the understanding of psychiatric disease. We provide a well-annotated data set that has the potential to act as a broadly applicable resource for investigators interested in molecular changes in multiple psychiatric disorders across multiple brain regions. We have conducted an extensive characterization of the molecular overlap between SZ and BPD at the gene expression and metabolite levels across multiple brain regions. We provide a high-confidence set of genes differentially expressed between SZ and CTL individuals utilizing two independent cohorts and highlight down-regulation of *EGR1* as a potential driver of broader scale transcription changes. We also establish that a significant proportion of transcriptome variation within SZ and BPD cohorts is correlated with expression changes in previously identified cell type-specific genes.

Additional files

Additional file 1: Supplemental text with detailed methods. (DOCX 134 kb)

Additional file 2: Supplemental tables S1–S10. (XLSX 349 kb)

Additional file 3: Supplemental figures S1–S7. (PDF 524 kb)

Abbreviations
AnCg: Anterior cingulate gyrus; BPD: Bipolar disorder; ChIP-seq: Chromatin immunoprecipitation with DNA sequencing; CTL: Control; DLPFC: Dorsolateral prefrontal cortex; FDR: False discovery rate; GABA: Gamma-aminobutyric acid; GEO: Gene Expression Omnibus; GO: Gene ontology; GWAS: Genome-wide association study; MDD: Major depression disorder; nAcc: Nucleus accumbens; PCA: Principal component analysis; PMI: Post-mortem interval; qPCR: Quantitative PCR; RNA-seq: RNA sequencing; SNCID: Stanley Neuropathology Consortium Integrative Database; SZ: schizophrenia; TSS: Transcription start site

Authors' contributions
HA, SJW, AFS, WEB, JDB, HK, SJC, and RMM conceived the study. KMB, RCR, BNL, SJC, AAH, MH, JZL, and RMM designed the experiments. EGJ performed brain dissections. PMC procured the brain tissue samples. MPV analyzed pH on all cases and matched the four cohorts. DWM obtained demographic and clinical data on all subjects through analyses of medical records and next-of-kin interviews. NSD, JG, and KMB collected RNAs and performed Tn-RNA-seq library construction. RCR and BNL analyzed the RNA-seq data. RCR and SJC performed and analyzed metabolomics experiments. KMB, RCR, and BNL wrote the first draft of the paper. JZL, BGB, WEB, SJW, SJC, HA, and RMM contributed to the writing of the paper. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Sample collection, including human subject recruitment and characterization, was conducted as part of the Brain Donor Program at the University of California, Irvine, Department of Psychiatry and Human Behavior (Pritzker Neuropsychiatric Disorders Research Consortium). Subject tissue was procured with the recorded informed consent of the decedents' legal next-of-kin. Tissue collection for this study was approved by the Institutional Review Board at the University of California, Irvine (UCI 88-041, UCI 97-74) and was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]HudsonAlpha Institute for Biotechnology, 601 Genome Way, Huntsville, AL 35806, USA. [2]Department of Genetics, The University of Alabama at Birmingham, Birmingham, AL, USA. [3]Mental Health Research Institute, University of Michigan, Ann Arbor, MI, USA. [4]Department of Psychiatry and Human Behavior, College of Medicine, University of California, Irvine, CA, USA. [5]Department of Psychiatry, Stanford University School of Medicine, Stanford, CA, USA. [6]Psychiatry, Weill Cornell Medical College, New York, NY, USA. [7]Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA. [8]Present address: Duke University, Durham, NC, USA. [9]Present address: University of Utah School of Medicine, Salt Lake City, UT, USA.

References
1. National Center for Health Statistics. Health, United States, 2005: With Chartbook on Long-term Trends in Health. Hyattsville, Maryland. 2005.
2. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing. 2013.
3. Caldwell CB, Gottesman II. Schizophrenics kill themselves too: a review of risk factors for suicide. Schizophr Bull. 1990;16:571–89.
4. Siris SG. Suicide and schizophrenia. J Psychopharmacol. 2001;15:127–35.
5. Ripke S, Neale BM, Corvin A, Walters JTR, Farh K-H, Holmans PA, et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature. 2014;511:421–7.
6. Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature. 2009;460:748–52.
7. Cross-disorder Psychiatric Genomics Group. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet. 2013;381:1371–9.

8. Lee SH, Ripke S, Neale BM, Faraone SV, Purcell SM, Perlis RH, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. Nat Genet. 2013;45:984–94.

9. Jääskeläinen E, Juola P, Hirvonen N, McGrath JJ, Saha S, Isohanni M, et al. A systematic review and meta-analysis of recovery in schizophrenia. Schizophr Bull. 2013;39:1296–306.

10. Hwang Y, Kim J, Shin JY, Kim JI, Seo JS, Webster MJ, et al. Gene expression profiling by mRNA sequencing reveals increased expression of immune/inflammation-related genes in the hippocampus of individuals with schizophrenia. Transl Psychiatry. 2013;3, e321.

11. Kohen R, Dobra A, Tracy JH, Haugen E. Transcriptome profiling of human hippocampus dentate gyrus granule cells in mental illness. Transl Psychiatry. 2014;4, e366.

12. Akula N, Barb J, Jiang X, Wendland JR, Choi KH, Sen SK, et al. RNA-sequencing of the brain transcriptome implicates dysregulation of neuroplasticity, circadian rhythms and GTPase binding in bipolar disorder. Mol Psychiatry. 2014;19:1179–85.

13. Darby MM, Yolken RH, Sabunciyan S. Consistently altered expression of gene sets in postmortem brains of individuals with major psychiatric disorders. Transl Psychiatry. 2016;6, e890.

14. Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat Neurosci. 2016;19(11):1442–53.

15. Olsen CM. Natural rewards, neuroplasticity, and non-drug addictions. Neuropharmacology. 2011;61:1109–22.

16. Wenzel JM, Rauscher NA, Cheer JF, Oleson EB. A role for phasic dopamine release within the nucleus accumbens in encoding aversion: a review of the neurochemical literature. ACS Chem Neurosci. 2015;6:16–26.

17. Evans SJ, Choudary PV, Vawter MP, Li J, Meador-Woodruff JH, Lopez JF, et al. DNA microarray analysis of functionally discrete human brain regions reveals divergent transcriptional profiles. Neurobiol Dis. 2003;14:240–50.

18. Li JZ, Vawter MP, Walsh DM, Tomita H, Evans SJ, Choudary PV, et al. Systematic changes in gene expression in postmortem human brains associated with tissue pH and terminal medical conditions. Hum Mol Genet. 2004;13:609–16.

19. Jones EG, Hendry SH, Liu XB, Hodgins S, Potkin SG, Tourtellotte WW. A method for fixation of previously fresh-frozen human adult and fetal brains that preserves histological quality and immunoreactivity. J Neurosci Methods. 1992;44:133–44.

20. Johnston NL, Cervenak J, Shore AD, Torrey EF, Yolken RH, Cerevnak J. Multivariate analysis of RNA levels from postmortem human brains as measured by three different methods of RT-PCR. Stanley Neuropathology Consortium. J Neurosci Methods. 1997;77:83–92.

21. Gertz J, Varley KE, Davis NS, Baas BJ, Goryshin IY, Vaidyanathan R, et al. Transposase mediated construction of RNA-seq libraries. Genome Res. 2012;22:134–41.

22. Alonso A, Lasseigne BN, Williams K, Nielsen J, Ramaker RC, Hardigan AA, et al. aRNApipe: a balanced, efficient and distributed pipeline for processing RNA-seq data in high performance computing environments. Bioinformatics. 2017;33(11):1727–9.

23. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.

24. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.

25. Kim JH, Karnovsky A, Mahavisno V, Weymouth T, Pande M, Dolinoy DC, et al. LRpath analysis reveals common pathways dysregulated via DNA methylation across cancer types. BMC Genomics. 2012;13:526.

26. Isserlin R, Merico D, Voisin V, Bader GD. Enrichment Map—a Cytoscape app to visualize and explore OMICs pathway enrichment results. F1000Research. 2014;3:141.

27. Darmanis S, Sloan SA, Zhang Y, Enge M, Caneda C, Shuer LM, et al. A survey of human brain transcriptome diversity at the single cell level. Proc Natl Acad Sci U S A. 2015;112(23):7285–90.

28. Zhang Y, Sloan SA, Clarke LE, Caneda C, Plaza CA, Blumenthal PD, et al. Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse. Neuron. 2015;89:37–53.

29. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keeffe S, et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. J Neurosci. 2014;34:11929–47.

30. Gong T, Szustakowski JD. DeconRNASeq: a statistical framework for deconvolution of heterogeneous tissue samples based on mRNA-Seq data. Bioinformatics. 2013;29:1083–5.

31. Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, Anderson N, et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. Nat Protoc. 2011;6:1060–83.

32. Hummel J, Strehmel N, Selbig J, Walther D, Kopka J. Decision tree supported substructure prediction of metabolites from GC-MS profiles. Metabolomics. 2010;6:322–33.

33. Castillo S, Mattila I, Miettinen J, Orešič M, Hyötyläinen T. Data analysis tool for comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry. Anal Chem. 2011;83:3058–67.

34. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545–50.

35. Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat Neurosci. 2015;18:199–209.

36. Vita A, De Peri L, Deste G, Sacchetti E. Progressive loss of cortical gray matter in schizophrenia: a meta-analysis and meta-regression of longitudinal MRI studies. Transl Psychiatry. 2012;2, e190.

37. Berretta S, Pantazopoulos H, Lange N. Neuron numbers and volume of the amygdala in subjects diagnosed with bipolar disorder or schizophrenia. Biol Psychiatry. 2007;62:884–93.

38. Thompson M, Weickert CS, Wyatt E, Webster MJ. Decreased glutamic acid decarboxylase(67) mRNA expression in multiple brain areas of patients with schizophrenia and mood disorders. J Psychiatr Res. 2009;43:970–7.

39. Skelly DA, Merrihew GE, Riffle M, Connelly CF, Kerr EO, Johansson M, et al. Integrative phenomics reveals insight into the structure of phenotypic diversity in budding yeast. Genome Res. 2013;23:1496–504.

40. Kim S, Hwang Y, Webster MJ, Lee D. Differential activation of immune/inflammatory response-related co-expression modules in the hippocampus across the major psychiatric disorders. Mol. Psychiatry. 2016;21(3):376-85. doi:10.1038/mp.2015.79.

41. Sequeira A, Morgan L, Walsh DM, Cartagena PM, Choudary P, Li J, et al. Gene expression changes in the prefrontal cortex, anterior cingulate cortex and nucleus accumbens of mood disorders subjects that committed suicide. PLoS One. 2012;7, e35367.

42. Sorenson SB. Gender disparities in injury mortality: consistent, persistent, and larger than you'd think. Am J Public Health. 2011;101:353–8.

43. Aubin H-J, Rollema H, Svensson TH, Winterer G. Smoking, quitting, and psychiatric disease: a review. Neurosci Biobehav Rev. 2012;36:271–84.

44. Wolock SL, Yates A, Petrill SA, Bohland JW, Blair C, Li N, et al. Gene × smoking interactions on human brain gene expression: finding common mechanisms in adolescents and adults. J Child Psychol Psychiatry. 2013;54:1109–19.

45. Weissman DH, Gopalakrishnan A, Hazlett CJ, Woldorff MG. Dorsal anterior cingulate cortex resolves conflict from distracting stimuli by boosting attention toward relevant events. Cereb Cortex. 2005;15:229–37.

46. Paus T. Primate anterior cingulate cortex: where motor control, drive and cognition interface. Nat Rev Neurosci. 2001;2:417–24.

47. Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD. Anterior cingulate cortex, error detection, and the online monitoring of performance. Science. 1998;280:747–9.

48. Woo T-UW, Kim AM, Viscidi E. Disease-specific alterations in glutamatergic neurotransmission on inhibitory interneurons in the prefrontal cortex in schizophrenia. Brain Res. 2008;1218:267–77.

49. Yamada K, Gerber DJ, Iwayama Y, Ohnishi T, Ohba H, Toyota T, et al. Genetic analysis of the calcineurin pathway identifies members of the EGR gene family, specifically EGR3, as potential susceptibility candidates in schizophrenia. Proc Natl Acad Sci U S A. 2007;104:2815–20.

50. Pérez-Santiago J, Diez-Alarcia R, Callado LF, Zhang JX, Chana G, White CH, et al. A combined analysis of microarray gene expression studies of the human prefrontal cortex identifies genes implicated in schizophrenia. J Psychiatr Res. 2012;46:1464–74.

51. Zhang L, Cho J, Ptak D, Leung YF. The role of egr1 in early zebrafish retinogenesis. PLoS One. 2013;8:1–11.

52. Baumgärtel K, Genoux D, Welzl H, Tweedie-Cullen RY, Koshibu K, Livingstone-Zatchej M, et al. Control of the establishment of aversive memory by calcineurin and Zif268. Nat Neurosci. 2008;11:572–8.

53. Bruins Slot LA, Lestienne F, Grevoz-Barret C, Newman-Tancredi A, Cussac D. F15063, a potential antipsychotic with dopamine D(2)/D(3) receptor

Ramaker *et al. Genome Medicine* (2017) 9:72

Page 12 of 12

antagonist and 5-HT(1A) receptor agonist properties: influence on immediate-early gene expression in rat prefrontal cortex and striatum. Eur J Pharmacol. 2009;620:27–35.

54. Xu Y, Yue W, Shugart YY, Li S, Cai L, Li Q, et al. Exploring transcription factors-microRNAs co-regulation networks in schizophrenia. Schizophr Bull. 2016;42:1037–45.

55. Koldamova R, Schug J, Lefterova M, Cronican AA, Fitz NF, Davenport FA, et al. Genome-wide approaches reveal EGR1-controlled regulatory networks associated with neurodegeneration. Neurobiol Dis. 2014;63:107–14.

56. Fromer M, Pocklington AJ, Kavanagh DH, Williams HJ, Dwyer S, Gormley P, et al. De novo mutations in schizophrenia implicate synaptic networks. Nature. 2014;506:179–84.

57. Guillozet-Bongaarts AL, Hyde TM, Dalley RA, Hawrylycz MJ, Henry A, Hof PR, et al. Altered gene expression in the dorsolateral prefrontal cortex of individuals with schizophrenia. Mol Psychiatry. 2014;19:478–85.