Received: 30 June 2022 | Revised: 23 November 2022 | Accepted: 5 January 2023

DOI: 10.1002/ajmg.b.32931

RESEARCH ARTICLE

American Journal of Medical Genetics Part B Neuropsychiatric Genetics WILEY

The relationship between case–control differential gene expression from brain tissue and genetic associations in schizophrenia

Nicholas E. Clifton[1,2] | Anton Schulmann[3] | Schizophrenia Working Group of the Psychiatric Genomics Consortium | Peter A. Holmans[1] | Michael C. O'Donovan[1] | Marquis P. Vawter[3]

[1]MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK

[2]University of Exeter Medical School, University of Exeter, Exeter, UK

[3]Functional Genomics Laboratory, Department of Psychiatry and Human Behavior, School of Medicine, University of California, Irvine, California, USA

Correspondence
Nicholas Clifton, MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Maindy Road, Cardiff, UK.
Email: n.clifton@exeter.ac.uk

Funding information
Medical Research Council, Grant/Award Numbers: MR/L010305/1, G0800509; National Institute of Mental Health, Grant/Award Numbers: U01MH109514, R01MH085801; Takeda Pharmaceutical Company

Abstract

Large numbers of genetic loci have been identified that are known to contain common risk alleles for schizophrenia, but linking associated alleles to specific risk genes remains challenging. Given that most alleles that influence liability to schizophrenia are thought to do so by altered gene expression, intuitively, case–control differential gene expression studies should highlight genes with a higher probability of being associated with schizophrenia and could help identify the most likely causal genes within associated loci. Here, we test this hypothesis by comparing transcriptome analysis of the dorsolateral prefrontal cortex from 563 schizophrenia cases and 802 controls with genome-wide association study (GWAS) data from the third wave study of the Psychiatric Genomics Consortium. Genes differentially expressed in schizophrenia were not enriched for common allelic association statistics compared with other brain-expressed genes, nor were they enriched for genes within associated loci previously reported to be prioritized by genetic fine-mapping. Genes prioritized by Summary-based Mendelian Randomization were underexpressed in cases compared to other genes in the same GWAS loci. However, the overall strength and direction of expression change predicted by SMR were not related to that observed in the differential expression data. Overall, this study does not support the hypothesis that genes identified as differentially expressed from RNA sequencing of bulk brain tissue are enriched for those that show evidence for genetic associations. Such data have limited utility for prioritizing genes in currently associated loci in schizophrenia.

KEYWORDS
FINEMAP, gene expression, genome-wide association study, postmortem tissue, schizophrenia, transcriptomics

Michael C. O'Donovan and Marquis P. Vawter shared senior authors.

1 | INTRODUCTION

The recent and largest genome wide association study (GWAS) of schizophrenia identified 287 loci associated with schizophrenia (Trubetskoy et al., 2022). Due to linkage disequilibrium (LD), most associated loci encompass several genes, and it is not known with certainty which genes are functionally impacted by the associated risk alleles. The prioritization of individual genes within GWAS-associated loci is a major challenge for genomics research and is key to translational outcomes. Aiming to identify the most likely schizophrenia-relevant genes, the Psychiatric Genomics Consortium (PGC) (Trubetskoy et al., 2022) applied various gene prioritization methods to the 287 identified loci. These methods, based on the inference of causal SNPs from positional fine-mapping using FINEMAP (Benner et al., 2016), evidence that the associated alleles colocalize with expression quantitative trait loci (eQTLs) through Summary-based Mendelian Randomization (SMR) (Zhu et al., 2016), and data from chromatin conformation (Hi-C) analysis, highlighted 120 genes as likely to explain associations at some of the loci. However, likely associated genes were not identified at the majority of loci, and for those prioritized genes, the evidence is probabilistic rather than definitive. Thus, there is considerable scope for further attempts at linking the association signals from that study to specific genes.

Given that few common variant associations to schizophrenia can credibly be attributed to nonsynonymous alleles (Trubetskoy et al., 2022), it is widely believed the effects of most are likely to be mediated by effects on gene expression, a hypothesis for which there is some evidence (Bray & O'Donovan, 2019; Huo, Li, Liu, Li, & Luo, 2019; Richards et al., 2012). If this hypothesis is correct, then genes that are differentially expressed in people with schizophrenia should be enriched for those that are functionally affected by the causal alleles underpinning genetic associations. Moreover, if this enrichment is nontrivial, differential gene expression data could inform gene prioritization from GWAS data. Research into other polygenic conditions, such as diabetes, has benefited from using this approach (Chen et al., 2008; Parikh, Lyssenko, & Groop, 2009), but thus far, it has not yet made an important impact on genomic studies of schizophrenia.

Transcriptomics studies of postmortem brain samples have led to the identification of large numbers of differentially expressed genes in the brains of people with schizophrenia (Collado-Torres et al., 2019; Fromer et al., 2016; Gandal, Haney, et al., 2018; Gandal, Zhang, et al., 2018; Hoffman et al., 2019; Jaffe et al., 2018; Ramaker et al., 2017). Analyses of genomic and transcriptomic case–control datasets have shown that genes with similar biological functions are enriched for associations in both types of data, including, for example, genes related to neuronal signaling (Gandal, Zhang, et al., 2018; Jaffe et al., 2018). This overlap further supports the hypothesis of a relationship between genomics and differential expression transcriptomics data, albeit that support is indirect. Here, using the largest schizophrenia GWAS dataset (Trubetskoy et al., 2022), we now directly test the relationship between differential gene expression and genetic association to evaluate the utility of differential gene expression for refining gene prioritization in schizophrenia.

2 | METHODS

2.1 | Transcriptome meta-analyses

RNA-Seq gene-level count data were obtained from synapse.com (Synapse ID: syn12080241), DOI: https://doi.org/10.7303/syn12080241. We included only dorsolateral prefrontal cortex (DLPFC) samples from controls and cases with an age of death ≥17 years (age of youngest SCZ donor) from the CMC, CMC_HBCC, LIBD, and BrainGVEX studies (802 controls, 563 schizophrenia, 221 bipolar disorder). Technical replicates were filtered to include only that with the highest RIN value. Downstream analysis was restricted to expressed genes (>5 counts per million in at least 10 samples). Gene-level counts were normalized using trimmed mean of M-values (TMM) normalization in edgeR (Robinson, McCarthy, & Smyth, 2010), log-transformed, and precision-weights calculated using the limma/voom approach (Law, Chen, Shi, & Smyth, 2014). Surrogate variable analysis was implemented using the sva package (Leek, Johnson, Parker, Jaffe, & Storey, 2012), and the number of surrogate variables (SVs) estimated using the "Leek" method (Leek, 2011) and set to 7. The brain bank of origin (including subcollection for CMC and BrainGVEX) was used as a covariate in the null model. Differential gene expression tests between schizophrenia cases and controls were then performed using a weighted least-squares linear regression in limma with the following model: Gene expression $\sim$ brain bank $+ SV_{[1-7]} +$ Dx, where Dx is control or schizophrenia.

To ensure that downstream analyses of the data derived above were not specific to the procedures used for determining differential expression, we repeated the analyses using a published schizophrenia case–control RNA sequencing analysis performed by the PsychENCODE consortium (Gandal, Zhang, et al., 2018), hereafter referred to as PsychENCODE. The RNA-seq samples used by PsychENCODE overlapped with our own but included additional samples from three studies (UCLA-ASD, Yale-ASD, BrainSpan) (Gandal, Zhang, et al., 2018). A further difference in that study (Gandal, Zhang, et al., 2018) was the set of covariates used: age, age^2, batch, sex, postmortem interval, RNA integrity (RIN), RIN^2, brain bank, brain region, 24 aggregate sequencing metrics (seqPCs), seqPC3^2, and 4 surrogate variables.

Both the present and PsychENCODE datasets were filtered to retain only protein-coding genes.

2.2 | Genomic data

Association and gene prioritization data were taken from the largest available schizophrenia case–control GWAS (Trubetskoy et al., 2022) of up to 76,755 people with schizophrenia and 243,649 controls, which reported common variant associations at 287 distinct genomic loci, referred to herein as the PGC. Genome-wide significant SNPs ($p < 5 \times 10^{-8}$) were assigned to 1,565 unique protein-coding genes

using gene boundaries from Ensembl assembly GRCh37. Fine-mapping of the significant loci, implemented by FINEMAP (Benner et al., 2016), identified a FINEMAP-broad set of 435 protein-coding genes which contained at least one SNP in the FINEMAP 95% credible set for each locus. The FINEMAP 95% credible set for each locus is the minimum set of SNPs for which the sum of the posterior probabilities (PPs) of being causal includes 95% of the locus-wide PP. The PGC further defined a subset of 64 protein-coding genes as FINEMAP-prioritized based on additional criteria, requiring the loci to remain significant after replication in an extended GWAS, and to either contain the entire set of credible SNPs, or at least one nonsynonymous or untranslated region variant with PP ≥ 0.1. The PGC also identified a broad set of 101 SMR genes where there was evidence consistent with the hypothesis that the GWAS association colocalized with an eQTL for that gene and where the HEIDI test suggested LD was not an evident explanation for that colocalization (Zhu et al., 2016). From these, we defined here 53 SMR-broad genes which were protein-coding and derived only from the subset of eQTLs from the adult brain. SMR genes were further prioritized ($n = 35$) by the PGC based on additional criteria, including colocalization of eQTLs with FINEMAP credible SNPs, or evidence from Hi-C chromosomal conformation analysis (Trubetskoy et al., 2022). We removed one FINEMAP-prioritized and two SMR-prioritized genes not represented in our case–control transcriptomic data. Gene sets are presented in Table S1. β values from adult brain SMR analysis were used as a continuous measure of evidence from SMR.

2.3 | Regression modeling

In primary analyses, a linear model was fitted regressing the differential expression *t*-statistic for each gene on either a continuous gene property (for example, gene-wide association statistic from MAGMA, or observed/expected upper bound fraction (LOEUF) score (Karczewski et al., 2020) for loss of function mutation intolerance, see below) or a binary variable denoting prioritization by one or more of the above criteria. Analyses were performed using either the signed *t*-statistic (a positive value indicating increased expression in schizophrenia, a negative value decreased expression in schizophrenia) or, where there was no prior expectation of direction-specific effects, the absolute *t*-statistic. Sets of differentially expressed genes, defined by *p* value thresholding, were tested by logistic regression. To compare the differential expression statistics of prioritized genes with those of the remaining genes in the same GWAS locus, we performed a conditional logistic regression analysis in R, using the loci as strata. In all regression analyses, average gene expression (mean $\log_2$ transcripts per million) was included in the model as a covariate.

2.4 | MAGMA genetic association analyses

Genes and gene sets were tested for enrichment for genetic association with schizophrenia using MAGMA (version 1.08b) (de Leeuw, Mooij, Heskes, & Posthuma, 2015) with summary statistics from the combined European and East Asian datasets provided by the PGC (Trubetskoy et al., 2022). Summary statistics from the GWAS were filtered to include only SNPs with a minor allele frequency ≥ 1% and imputed INFO score ≥ 0.8. SNPs were assigned to genes using a 35 kb upstream and 10 kb downstream window around the longest transcript to include proximal regulatory regions (Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium, 2015). A European and East Asian reference panel of the Haplotype Reference Consortium (McCarthy et al., 2016) was used to estimate LD between SNPs, and gene-wide association statistics were calculated using the SNP-wise mean model. Gene set enrichment analysis was performed in MAGMA using a one-tailed competitive association test, controlling for gene size and SNP density. All protein-coding genes expressed in the transcriptomic dataset were used as the background against which enrichment for association was tested. The relationship between gene differential expression *t*-statistic and enrichment for association with schizophrenia was determined using gene property analysis (linear regression) in MAGMA. Multiple testing of gene sets was adjusted for using the Bonferroni method.

2.5 | Partitioned heritability analysis

We applied LD score regression analysis to estimate the heritability explained by various differentially expressed sets of genes using published methods (Bulik-Sullivan et al., 2015; Finucane et al., 2015). Summary statistics (Trubetskoy et al., 2022) were filtered as described above, and SNPs were assigned to each gene set using a 10 kb window around the transcribed region of each gene (Kim et al., 2019). Each gene set annotation was tested against a baseline LD model containing 97 annotations describing a broad set of coding, conserved and regulatory regions (Gazal et al., 2017; Hujoel, Gazal, Hormozdiari, van de Geijn, & Price, 2019). In primary analyses, annotations were tested for heritability enrichment, each compared to the remaining SNPs in the genome. To evaluate the relative heritability enrichment of an annotation, conditioned on the baseline and any additional annotations, we extracted the partitioned heritability coefficient τ_c, which was converted to a *Z*-score by dividing by the standard error.

3 | RESULTS

3.1 | Correlates of case–control differential expression

The results of differential gene expression analysis are provided in Table S2.

The probability of a gene being reported as differentially expressed can be biased by its transcript abundance (Yoon & Nam, 2017). Consistent with this, the differential expression *t*-statistic was related to the DLPFC expression level (mean $\log_2$ transcripts per million) such that highly expressed transcripts had a greater likelihood

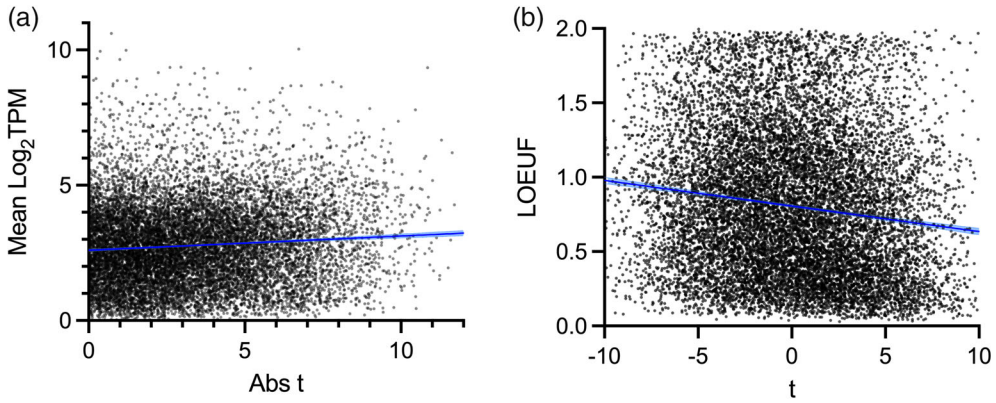

FIGURE 1 Gene property correlates of differential expression. (a) Read count bias in differential gene expression statistics. Shown is the absolute *t*-statistic (Abs *t*) from differential expression analysis and the mean log transcripts per million (TPM) for each protein-coding gene in RNA sequencing analyses comparing schizophrenia cases and controls. Differential expression is more likely to be detected in highly expressed genes. Average expression was used as a covariate in subsequent regression analyses. (b) Imbalanced differential expression of loss-of-function intolerant genes. Shown is the (directional) *t*-statistic from differential expression analysis and the loss-of-function observed/expected upper bound fraction (LOEUF) for each protein-coding gene. Genes less tolerant to loss of function mutations were more likely to be overexpressed in cases compared to controls

of being differentially expressed ($\beta = .13$, $p = 6.4 \times 10^{-23}$) (Figure 1a). Subsequent regression analyses controlled for this effect by covarying for mean tissue expression.

Genes with low tolerance to loss-of-function (LoF) mutations are enriched for genetic association with schizophrenia (Trubetskoy et al., 2022) and have a higher expression on average (Karczewski et al., 2020). We tested if the differential expression was also associated with LoF intolerance, using the LOEUF score while controlling for mean tissue expression. Higher expression in cases vs. controls was associated with LoF mutation intolerance (i.e., lower LOEUF score; $\beta = -1.21$, $p = 1.1 \times 10^{-66}$) (Figure 1b). Thresholding for differential expression at a 0.05 FDR cut-off gave a similar picture: genes with significant upregulation having lower tolerance to LoF mutations (logistic regression analysis: $\beta = -.61$, $p = 1.2 \times 10^{-46}$) while significantly downregulated genes were more tolerant to LoF mutations ($\beta = .50$, $p = 1.5 \times 10^{-36}$).

3.2 | Relationship of differential expression to GWAS statistics

We found no relationship between differential gene expression and measures of the gene-wide significance using MAGMA ($\beta = -.0067$, $p = .14$). Moreover, genes ($n = 1{,}565$) mapping to genome-wide significant loci were not more likely to be differentially expressed than other brain-expressed genes that did not map to the loci ($\beta = -.059$, $p = .38$) (Figure 2a). Repeated analyses in another differential gene expression analysis undertaken by PsychENCODE (Gandal, Zhang, et al., 2018) yielded similar results (Figure 2b).

Genes reported as significant in the differential expression analyses were enriched for heritability. However, they were not more enriched than those that did not show such evidence of differential expression (Table 1), indicating this might reflect the heritability enrichment expected at brain expressed genes. Adding brain-expressed genes to the conditional annotations resulted in none of the gene sets, differentially expressed or not, being significantly enriched for heritability as indicated by the τ_c statistic (Table 1). As a positive control, we repeated these analyses using targets of Fragile X mental retardation protein (FMRP) (Darnell et al., 2011), a gene set with strong prior evidence for genetic association with schizophrenia (Clifton, Rees, et al., 2021; Pardiñas et al., 2018). This set retained strong evidence for enriched SNP heritability (Table 1) even after conditioning for brain expression.

Gene set association analyses in MAGMA mirror these findings; up- or downregulated genes defined in either the present or the earlier study (Gandal, Zhang, et al., 2018) were not enriched for association with schizophrenia after conditioning on the set of brain expressed genes (this study: upregulated $\beta = -.036$, $p = 1.0$; downregulated $\beta = .018$, $p = .73$. PsychENCODE: upregulated $\beta = -.044$, $p = 1.0$; downregulated $\beta = .047$, $p = .15$).

3.3 | Relationship of differential expression to GWAS prioritized genes

The subset of genes at associated loci comprising the FINEMAP-broad set did not show significantly stronger differential expression (i.e., higher absolute *t*-statistic) than other genes in the same locus in a conditional logistic regression ($\beta = .018$, $p = .61$) (Figure 2c). However, contrary to expectation, FINEMAP-prioritized genes had smaller differential gene expression (i.e., had lower absolute *t*-statistic) than other genes in the locus ($\beta = -.22$, $p = .038$) (Figure 2e). Analyses based on PsychENCODE transcriptomic data gave qualitatively similar findings (FINEMAP-broad: $\beta = -.10$, $p = .090$; FINEMAP-prioritized: $\beta = -.63$, $p = .0073$) (Figure 2d,f).

Genes in the SMR-broad or SMR-prioritized sets had a lower signed differential expression *t*-statistic than the remaining genes

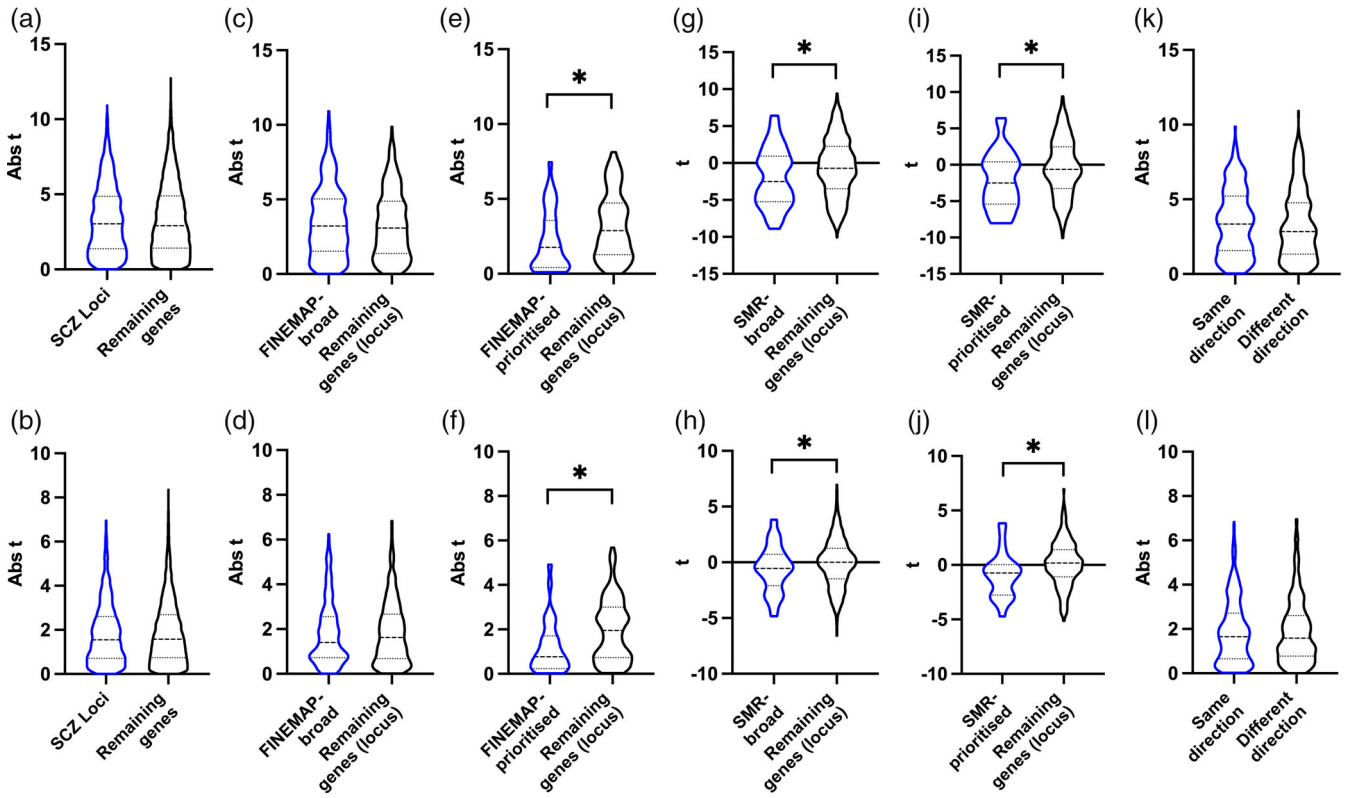

FIGURE 2 Case–control differential expression of GWAS prioritized genes. Results are presented based on findings from case–control differential gene expression analyses performed in the current study (top row) or PsychENCODE (bottom row). (a, b) The absolute *t*-statistic (Abs *t*) from differential expression analysis was compared between all genes overlapping with any of 270 GWAS significant loci and all remaining protein-coding genes. (c–f) The absolute *t*-statistic from differential expression analysis was compared between FINEMAP-broad or FINEMAP-prioritized genes and the remaining genes in the locus. (g–j) The signed *t*-statistic from differential expression analysis was compared between SMR-broad or SMR-prioritized genes and the remaining genes in the locus. (k, l) The absolute β value from SMR analysis was compared between genes with the same, or different, direction of effect in SMR and differential expression analyses. Group differences were analyzed using logistic regression analysis (a, b, k, l) or conditional logistic regression analysis (c–j). Asterisks indicate significant ($p < .05$) differences between groups

TABLE 1 Partitioned heritability analysis

Gene set	Proportion SNPs	h^2	Enrichment	Enrichment *p*	*p* (τ_c)	*p* (τ_c) + brain-expressed
This study—upregulated	0.16	0.23	1.47	3.30E−08	0.0048	0.42
This study—downregulated	0.13	0.19	1.52	6.28E−09	7.2E−04	0.30
This study—not differentially expressed	0.18	0.29	1.56	1.69E−14	1.51E−06	0.20
PsychENCODE—upregulated	0.073	0.11	1.49	6.32E−06	0.0027	0.35
PsychENCODE—downregulated	0.068	0.11	1.60	2.31E−07	3.05E−04	0.19
PsychENCODE—not differentially expressed	0.33	0.47	1.46	1.37E−19	2.49E−05	0.36
FMRP targets	0.049	0.11	2.36	6.3E−17	2.11E−06	6.84E−6

Note: *p* values are presented based on the significance of the enrichment for SNP heritability, after conditioning on 97 baseline annotations (Gazal et al., 2017; Hujoel et al., 2019), and after the addition of all brain-expressed genes to the conditional annotations. Proportion of total heritability, h^2.

in the locus (this study: SMR-broad $\beta = -.15$, $p = .0021$; SMR-prioritized $\beta = -.15$, $p = .011$. PsychENCODE: SMR-broad $\beta = -.21$, $p = .013$; SMR-prioritized $\beta = -.32$, $p = .0029$) (Figure 2g–j), with the tendency being for SMR genes to be under-expressed in cases. However, there was no significant difference in the absolute differential expression statistics for genes where differential expression was in the same direction as that predicted by SMR and those where it was in the opposite direction (this study: $\beta = .0058$, $p = .52$; PsychENCODE: β −.0025, $p = .86$) (Figure 2k,l).

4 | DISCUSSION

Large-scale transcriptomic and genomic studies have identified many genes that show evidence for association with schizophrenia. However, while individual genes have been reported with significant associations from both types of data (Fromer et al., 2016), differential gene expression data from transcriptomics has not yet informed systematic efforts at gene prioritization from GWAS. Moreover, it is unclear whether such information should be so used, and if it is, what weight to assign to it. Aiming to clarify those questions, in this study of coding genes, we sought to quantify the relationships between differential gene expression, genome wide association test statistics, associated loci, and genes prioritized within loci. Overall, we find no evidence for robust relationships between differential gene expression and several measures of genetic association, suggesting that in its current form, the former does not assist with mapping GWAS associations to specific genes.

We observed that genes that are less tolerant to LoF mutation are more likely to be overexpressed in people with schizophrenia than in controls. This finding superficially may have some resonance with earlier reports that LoF mutation intolerant genes are enriched for association with schizophrenia (Pardiñas et al., 2018; Trubetskoy et al., 2022). However, further analyses suggest these are unrelated observations. First, genes that show differential gene expression in schizophrenia showed no evidence of being enriched for genetic associations. Second, genes mapping to schizophrenia-associated loci were not more likely to be differentially expressed than genes outside these loci. Third, surprisingly, genes prioritized by fine-mapping showed evidence for less differential expression than other genes at the associated loci. Fourth, genes showing evidence for differential expression were not more enriched for SNP-based heritability than brain expressed genes that were not differentially expressed. Like our study, a previous analysis found that genes differentially expressed in schizophrenia are enriched for heritability compared with other genes in the genome (Gandal, Zhang, et al., 2018) but did not test whether this could be explained by the property of being brain-expressed (Pardiñas et al., 2018) rather than being differentially expressed. Here, we found that regardless of whether genes were upregulated, downregulated, or neither, they were similarly enriched for SNP heritability. In all cases, the enrichment disappeared after conditioning on brain expression.

Our observation that FINEMAP-prioritized genes were significantly less likely to be differentially expressed than the remaining genes in the locus is surprising. One possible but highly speculative interpretation comes from a report (Gandal, Haney, et al., 2018) that in people with schizophrenia, antipsychotic drugs partially normalize primary disorder-relevant transcriptomic changes. If this is correct, then true susceptibility genes, which are expected to be enriched among those prioritized by fine-mapping, might be less likely than random genes to show differential expression due to normalization by antipsychotics. This potential effect can be tested by comparing the transcriptional patterns of genes associated with schizophrenia and pharmacological treatment.

Genes prioritized for association with schizophrenia through SMR were more likely to be expressed at lower levels in schizophrenia than other genes in the same loci, regardless of whether the putatively causal eQTLs are expected to increase or decrease expression. The absence of an expected relationship between the direction of expression predicted by the genomic-eQTL data and that observed directly by differential expression adds further to our suggestion that differential gene expression in its current form does not provide a useful means to prioritize genomic association findings. However, the observation that, as a class, SMR-associated genes were observed to be expressed at lower levels in schizophrenia than in controls in the differential gene expression study is intriguing. We can only speculate as to a possible mechanism, but one possibility is that regardless of the direction of cis-acting effects, the net effect of causal genetic variation on brain pathology is to reduce the representation of the cellular fraction in bulk tissue in which those genes are expressed, leading to the observation of lower expression. Single-cell rather than bulk tissue case–control expression studies are necessary to evaluate that possibility.

We note that conclusions drawn in the present study concerning concordance between observed and expected gene expression might be limited by the tissue. First, our results are based on studies of bulk brain tissue containing multiple cell types with independent expression dynamics. Genomics and transcriptomics may show stronger patterns of concordance with the expansion of cell-specific case–control studies of differential gene expression and cell-specific eQTLs. Second, the current study was based on RNA samples derived from patients with ages ranging from 17 to more than 90. Since the expression of schizophrenia-associated genes varies across different ages (Clifton et al., 2019), and molecular pathology during brain development is thought to influence predisposition for schizophrenia (Clifton, Collado-Torres, et al., 2021; O'Brien et al., 2018; Owen, O'Donovan, Thapar, & Craddock, 2011; Weinberger, 2017), genetic effects predicted by GWAS might only manifest during specific developmental periods.

Differential gene expression between cases and controls reflects the effects of many factors beyond the cis-acting effects of common genetic variation that are the focus of the present article. Some of these, for example, trans-acting genetic effects, epigenetic modification from exposure to environmental risk factors, and expression changes that are of pathophysiological relevance but are downstream of the primary genetic changes, may be more readily exposed by transcriptomics than by current approaches to genomics. For example, epigenetics may have effects on the transcriptome during neurodevelopment which may be missed in adult transcriptomic studies. In addition, epigenetic-induced alterations in developmental gene expression direct cellular migration in the brain and affect single cell types. Such effects of true pathophysiological relevance, but which are not directly attributable to risk genes highlighted by genetic associations, might explain why despite poor overlap at the gene level, GWAS and differential expression studies of schizophrenia converge on common biological pathways (Collado-Torres et al., 2019; Gandal, Zhang, et al., 2018; Jaffe et al., 2018).

ACKNOWLEDGMENTS

This work was supported by Medical Research Council Centre Grant No. MR/L010305/1, Program Grant No. G0800509 and by the National Institute of Mental Health (USA) under Award Numbers U01MH109514 and R01MH085801. The content is the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Data were generated as part of the PsychENCODE Consortium, supported by: U01DA048279, U01MH103339, U01MH103340, U01MH103346, U01MH103365, U01MH103392, U01MH116438, U01MH116441, U01MH116442, U01MH116488, U01MH116489, U01MH116492, U01MH122590, U01MH122591, U01MH122592, U01MH122849, U01MH122678, U01MH122681, U01MH116487, U01MH122509, R01MH094714, R01MH105472, R01MH105898, R01MH109677, R01MH109715, R01MH110905, R01MH110920, R01MH110921, R01MH110926, R01MH110927, R01MH110928, R01MH111721, R01MH117291, R01MH117292, R01MH117293, R21MH102791, R21MH103877, R21MH105853, R21MH105881, R21MH109956, R56MH114899, R56MH114901, R56MH114911, R01MH125516, and P50MH106934 awarded to: Alexej Abyzov, Nadav Ahituv, Schahram Akbarian, Alexander Arguello, Lora Bingaman, Kristin Brennand, Andrew Chess, Gregory Cooper, Gregory Crawford, Stella Dracheva, Peggy Farnham, Mark Gerstein, Daniel Geschwind, Fernando Goes, Vahram Haroutunian, Thomas M. Hyde, Andrew Jaffe, Peng Jin, Manolis Kellis, Joel Kleinman, James A. Knowles, Arnold Kriegstein, Chunyu Liu, Keri Martinowich, Eran Mukamel, Richard Myers, Charles Nemeroff, Mette Peters, Dalila Pinto, Katherine Pollard, Kerry Ressler, Panos Roussos, Stephan Sanders, Nenad Sestan, Pamela Sklar, Nick Sokol, Matthew State, Jason Stein, Patrick Sullivan, Flora Vaccarino, Stephen Warren, Daniel Weinberger, Sherman Weissman, Zhiping Weng, Kevin White, A. Jeremy Willsey, Hyejung Won, and Peter Zandi. Data were generated as part of the CommonMind Consortium supported by funding from Takeda Pharmaceuticals Company Limited, F. Hoffmann-La Roche Ltd and NIH grants R01MH085542, R01MH093725, P50MH066392, P50MH080405, R01MH097276, RO1-MH-075916, P50M096891, P50MH084053S1, R37MH057881, AG02219, AG05138, MH06692, R01MH110921, R01MH109677, R01MH109897, U01MH103392, and contract HHSN271201300031C through IRP NIMH. Brain tissue for the study was obtained from the following brain bank collections: the Mount Sinai NIH Brain and Tissue Repository, the University of Pennsylvania Alzheimer's Disease Core Center, the University of Pittsburgh NeuroBioBank and Brain and Tissue Repositories, and the NIMH Human Brain Collection Core. CMC Leadership: Panos Roussos, Joseph Buxbaum, Andrew Chess, Schahram Akbarian, Vahram Haroutunian (Icahn School of Medicine at Mount Sinai), Bernie Devlin, David Lewis (University of Pittsburgh), Raquel Gur, Chang-Gyu Hahn (University of Pennsylvania), Enrico Domenici (University of Trento), Mette A. Peters, Solveig Sieberts (Sage Bionetworks), Thomas Lehner, Stefano Marenco, Barbara K. Lipska (NIMH).

CONFLICT OF INTEREST

MOD is supported by a collaborative research grant from Takeda Pharmaceuticals. Takeda played no part in the conception, design, implementation, or interpretation of the present study.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Nicholas E. Clifton https://orcid.org/0000-0003-2597-5253

REFERENCES

Benner, C., Spencer, C. C. A., Havulinna, A. S., Salomaa, V., Ripatti, S., & Pirinen, M. (2016). FINEMAP: Efficient variable selection using summary data from genome-wide association studies. *Bioinformatics*, *32*(10), 1493–1501. https://doi.org/10.1093/bioinformatics/btw018

Bray, N. J., & O'Donovan, M. C. (2019). The genetics of neuropsychiatric disorders. *Brain and Neuroscience Advances*, *2*, 239821281879927. https://doi.org/10.1177/2398212818799271

Bulik-Sullivan, B., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics Consortium, … Neale, B. M. (2015). LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, *47*(3), 291–295. https://doi.org/10.1038/ng.3211

Chen, R., Morgan, A. A., Dudley, J., Deshpande, T., Li, L., Kodama, K., … Butte, A. J. (2008). FitSNPs: Highly differentially expressed genes are more likely to have variants associated with disease. *Genome Biology*, *9*(12), R170. https://doi.org/10.1186/gb-2008-9-12-r170

Clifton, N. E., Collado-Torres, L., Burke, E. E., Pardiñas, A. F., Harwood, J. C., Di Florio, A., … Hall, J. (2021). Developmental profile of psychiatric risk associated with voltage-gated cation channel activity. *Biological Psychiatry*, *90*(6), 399–408. https://doi.org/10.1016/j.biopsych.2021.03.009

Clifton, N. E., Hannon, E., Harwood, J. C., Di Florio, A., Thomas, K. L., Holmans, P. A., … Hall, J. (2019). Dynamic expression of genes associated with schizophrenia and bipolar disorder across development. *Translational Psychiatry*, *9*(1), 74. https://doi.org/10.1038/s41398-019-0405-x

Clifton, N. E., Rees, E., Holmans, P. A., Pardiñas, A. F., Harwood, J. C., Di Florio, A., … Pocklington, A. J. (2021). Genetic association of FMRP targets with psychiatric disorders. *Molecular Psychiatry*, *26*(7), 2977–2990. https://doi.org/10.1038/s41380-020-00912-2

Collado-Torres, L., Burke, E. E., Peterson, A., Shin, J., Straub, R. E., Rajpurohit, A., … Jaffe, A. E. (2019). Regional heterogeneity in gene expression, regulation, and coherence in the frontal cortex and hippocampus across development and schizophrenia. *Neuron*, *103*(2), 203–216.e8. https://doi.org/10.1016/j.neuron.2019.05.013

Darnell, J. C., Van Driesche, S. J., Zhang, C., Hung, K. Y. S., Mele, A., Fraser, C. E., … Darnell, R. B. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell*, *146*(2), 247–261. https://doi.org/10.1016/j.cell.2011.06.013

de Leeuw, C. A., Mooij, J. M., Heskes, T., & Posthuma, D. (2015). MAGMA: Generalized gene-set analysis of GWAS data. *PLoS Computational Biology*, *11*(4), e1004219. https://doi.org/10.1371/journal.pcbi.1004219

Finucane, H. K., Bulik-Sullivan, B., Gusev, A., Trynka, G., Reshef, Y., Loh, P.-R., … Price, A. L. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics*, *47*(11), 1228–1235. https://doi.org/10.1038/ng.3404

Fromer, M., Roussos, P., Sieberts, S. K., Johnson, J. S., Kavanagh, D. H., Perumal, T. M., … Sklar, P. (2016). Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nature Neuroscience*, *19*(11), 1442–1453. https://doi.org/10.1038/nn.4399

Gandal, M. J., Haney, J. R., Parikshak, N. N., Leppa, V., Ramaswami, G., Hartl, C., … Geschwind, D. H. (2018). Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science*, *359*(6376), 693–697. https://doi.org/10.1126/science.aad6469

Gandal, M. J., Zhang, P., Hadjimichael, E., Walker, R. L., Chen, C., Liu, S., … Geschwind, D. H. (2018). Transcriptome-wide isoform-level

dysregulation in ASD, schizophrenia, and bipolar disorder. *Science*, *362*(6420), eaat8127. https://doi.org/10.1126/science.aat8127

Gazal, S., Finucane, H. K., Furlotte, N. A., Loh, P.-R., Palamara, P. F., Liu, X., … Price, A. L. (2017). Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nature Genetics*, *49*(10), 1421–1427. https://doi.org/10.1038/ng.3954

Hoffman, G. E., Bendl, J., Voloudakis, G., Montgomery, K. S., Sloofman, L., Wang, Y.-C., … Roussos, P. (2019). CommonMind consortium provides transcriptomic and epigenomic data for schizophrenia and bipolar disorder. *Scientific Data*, *6*(1), 180. https://doi.org/10.1038/s41597-019-0183-6

Hujoel, M. L. A., Gazal, S., Hormozdiari, F., van de Geijn, B., & Price, A. L. (2019). Disease heritability enrichment of regulatory elements is concentrated in elements with ancient sequence age and conserved function across species. *American Journal of Human Genetics*, *104*(4), 611–624. https://doi.org/10.1016/j.ajhg.2019.02.008

Huo, Y., Li, S., Liu, J., Li, X., & Luo, X.-J. (2019). Functional genomics reveal gene regulatory mechanisms underlying schizophrenia risk. *Nature Communications*, *10*(1), 670. https://doi.org/10.1038/s41467-019-08666-4

Jaffe, A. E., Straub, R. E., Shin, J. H., Tao, R., Gao, Y., Collado-Torres, L., … Weinberger, D. R. (2018). Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. *Nature Neuroscience*, *21*(8), 1117–1125. https://doi.org/10.1038/s41593-018-0197-y

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., … MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, *581*(7809), 434–443. https://doi.org/10.1038/s41586-020-2308-7

Kim, S. S., Dai, C., Hormozdiari, F., van de Geijn, B., Gazal, S., Park, Y., … Price, A. L. (2019). Genes with high network connectivity are enriched for disease heritability. *American Journal of Human Genetics*, *104*(5), 896–913. https://doi.org/10.1016/j.ajhg.2019.03.020

Law, C. W., Chen, Y., Shi, W., & Smyth, G. K. (2014). Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology*, *15*(2), R29. https://doi.org/10.1186/gb-2014-15-2-r29

Leek, J. T. (2011). Asymptotic conditional singular value decomposition for high-dimensional genomic data. *Biometrics*, *67*(2), 344–352. https://doi.org/10.1111/j.1541-0420.2010.01455.x

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., & Storey, J. D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*, *28*(6), 882–883. https://doi.org/10.1093/bioinformatics/bts034

McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., … Haplotype Reference Consortium. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics*, *48*(10), 1279–1283. https://doi.org/10.1038/ng.3643

Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. (2015). Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nature Neuroscience*, *18*(2), 199–209. https://doi.org/10.1038/nn.3922

O'Brien, H. E., Hannon, E., Hill, M. J., Toste, C. C., Robertson, M. J., Morgan, J. E., … Bray, N. J. (2018). Expression quantitative trait loci in the developing human brain and their enrichment in neuropsychiatric disorders. *Genome Biology*, *19*(1), 194. https://doi.org/10.1186/s13059-018-1567-1

Owen, M. J., O'Donovan, M. C., Thapar, A., & Craddock, N. (2011). Neurodevelopmental hypothesis of schizophrenia. *The British Journal of Psychiatry*, *198*(3), 173–175. https://doi.org/10.1192/bjp.bp.110.084384

Pardiñas, A. F., Holmans, P., Pocklington, A. J., Escott-Price, V., Ripke, S., Carrera, N., … Walters, J. T. R. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics*, *50*(3), 381–389. https://doi.org/10.1038/s41588-018-0059-2

Parikh, H., Lyssenko, V., & Groop, L. C. (2009). Prioritizing genes for follow-up from genome wide association studies using information on gene expression in tissues relevant for type 2 diabetes mellitus. *BMC Medical Genomics*, *2*, 72. https://doi.org/10.1186/1755-8794-2-72

Ramaker, R. C., Bowling, K. M., Lasseigne, B. N., Hagenauer, M. H., Hardigan, A. A., Davis, N. S., … Myers, R. M. (2017). Post-mortem molecular profiling of three psychiatric disorders. *Genome Medicine*, *9*(1), 72. https://doi.org/10.1186/s13073-017-0458-5

Richards, A. L., Jones, L., Moskvina, V., Kirov, G., Gejman, P. V., Levinson, D. F., … O'Donovan, M. C. (2012). Schizophrenia susceptibility alleles are enriched for alleles that affect gene expression in adult human brain. *Molecular Psychiatry*, *17*(2), 193–201. https://doi.org/10.1038/mp.2011.11

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Trubetskoy, V., Pardiñas, A. F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T. B., … van Os, J. (2022). Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*, *604*(7906), 502–508. https://doi.org/10.1038/s41586-022-04434-5

Weinberger, D. R. (2017). Future of days past: Neurodevelopment and schizophrenia. *Schizophrenia Bulletin*, *43*(6), 1164–1168. https://doi.org/10.1093/schbul/sbx118

Yoon, S., & Nam, D. (2017). Gene dispersion is the key determinant of the read count bias in differential expression analysis of RNA-seq data. *BMC Genomics*, *18*(1), 408. https://doi.org/10.1186/s12864-017-3809-0

Zhu, Z., Zhang, F., Hu, H., Bakshi, A., Robinson, M. R., Powell, J. E., … Yang, J. (2016). Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature Genetics*, *48*(5), 481–487. https://doi.org/10.1038/ng.3538

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Clifton, N. E., Schulmann, A., Schizophrenia Working Group of the Psychiatric Genomics Consortium, Holmans, P. A., O'Donovan, M. C., & Vawter, M. P. (2023). The relationship between case–control differential gene expression from brain tissue and genetic associations in schizophrenia. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 1–8. https://doi.org/10.1002/ajmg.b.32931