

ORIGINAL ARTICLE

Differential expression of transcriptional regulatory units in the prefrontal cortex of patients with bipolar disorder: potential role of early growth response gene 3

B Pfaffenseller^{1,2}, PV da Silva Magalhães^{1,3}, MA De Bastiani², MAA Castro⁴, AL Gallitano⁵, F Kapczinski^{1,3} and F Klamt²

Bipolar disorder (BD) is a severe mental illness with a strong genetic component. Despite its high degree of heritability, current genetic studies have failed to reveal individual loci of large effect size. In lieu of focusing on individual genes, we investigated regulatory units (regulons) in BD to identify candidate transcription factors (TFs) that regulate large groups of differentially expressed genes. Network-based approaches should elucidate the molecular pathways governing the pathophysiology of BD and reveal targets for potential therapeutic intervention. The data from a large-scale microarray study was used to reconstruct the transcriptional associations in the human prefrontal cortex, and results from two independent microarray data sets to obtain BD gene signatures. The regulatory network was derived by mapping the significant interactions between known TFs and all potential targets. Five regulons were identified in both transcriptional network models: early growth response 3 (*EGR3*), TSC22 domain family, member 4 (*TSC22D4*), interleukin enhancer-binding factor 2 (*ILF2*), Y-box binding protein 1 (*YBX1*) and MAP-kinase-activating death domain (*MADD*). With a high stringency threshold, the consensus across tests was achieved only for the *EGR3* regulon. We identified *EGR3* in the prefrontal cortex as a potential key target, robustly repressed in both BD signatures. Considering that *EGR3* translates environmental stimuli into long-term changes in the brain, disruption in biological pathways involving *EGR3* may induce an impaired response to stress and influence on risk for psychiatric disorders, particularly BD.

Translational Psychiatry (2016) 6, e805; doi:10.1038/tp.2016.78; published online 10 May 2016

INTRODUCTION

Bipolar disorder (BD) is a severe mental illness with a strong genetic component. Heritability is high, as evaluated by monozygotic and dizygotic twin concordance, albeit not perfect.^{1,2} Nevertheless, current molecular genetic studies indicate that no particular locus of large effect is involved in its etiology. The candidate gene approach has not delivered robust results and genome-wide association studies also often fail to show strong positive signals.² Current polygenic analyses are consistent with hundreds or thousands of susceptibility variants of weak effect, with the variation in single-nucleotide polymorphisms explaining 20 to 30% of the heritability shown in family and twin studies.^{3,4}

The generally low yield of pure genetic association studies has generated interest in alternative approaches.⁵ Gene expression studies move the discussion beyond statistical associations into the realm of neurobiology. RNA analysis can be more informative of the status of the cell, as it reflects a functional state not only influenced by genetic polymorphisms, but also by transcriptional modulation. Using high-throughput technologies, such as microarrays, the differential expression of genomic DNA in the form of mRNA has the potential to lead to disease signatures.⁶ One major caveat is that gene expression is tissue specific. For BD studies, this means obtaining brain tissue from postmortem donors, and few brain bank collections exist for this illness.⁷

Given the difficulty of obtaining brain tissue, the data on gene expression in BD is quite limited (compared with cancer research, for example). To get an idea of the sparseness of existing data in BD, a recent systematic review of gene expression studies revealed publicly available data from only 57 unique BD cases.⁸ Furthermore, gene expression varies depending on the area and structure under study.⁹ This underscores the importance of selecting the appropriate brain regions and using a methodological framework to extract biologically meaningful information from large-scale data. Although the neurocircuitry involved in the mood disorders is expected to be complex, there are major areas of interest that could be fruitfully explored in postmortem studies. Overall, interest has focused on the limbic–cortical–striato–pallido–thalamic circuits. The prefrontal cortex is a relevant nexus, where recent research points to areas in the medial prefrontal and orbitofrontal cortex.¹⁰

In the field of systems biology, approaches to identify candidate master regulators (MRs) have focused on transcription factors (TFs) that exert large influences on a phenotype. Recently, a number of computational methods have been developed to identify groups of genes, and even entire pathways, coordinated by a small number of TFs. These approaches have successfully identified gene units that are impaired in diseases such as cancer and diabetes.^{11–15} As in other traditional medical fields, psychiatry is

¹Bipolar Disorder Program, Laboratory of Molecular Psychiatry, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil; ²Laboratory of Cellular Biochemistry, Department of Biochemistry, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil; ³Department of Psychiatry, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil; ⁴Bioinformatics and Systems Biology Laboratory, Federal University of Paraná, Polytechnic Center, Curitiba, Brazil and ⁵Department of Basic Medical Sciences, University of Arizona College of Medicine, Phoenix, AZ, USA. Correspondence: Professor PV da Silva Magalhães, Department of Psychiatry, Universidade Federal do Rio Grande do Sul, 2350 Ramiro Barcelos Street, Porto Alegre 90035 903, Brazil.

E-mail: pedromaga2@gmail.com

Received 16 March 2016; accepted 23 March 2016

currently focusing on the study of biological pathways. Reverse engineering algorithms are used to reconstruct cell type-specific regulatory networks from high-throughput data. This approach efficiently reduces the complexity of the network, allowing the identification of MR TFs with tissue-specific signatures.

Here we query the genetic regulatory signature of BD through a series of steps. First, we analyze a large gene expression data set from healthy human prefrontal cortex across the lifespan¹⁶ to construct a regulatory network of known TFs and all potential targets. As gene expression is tissue specific, this empirical approach has the benefit of being a more realistic representation of prefrontal cortex functioning. Next, we identify genes that are differentially expressed in the prefrontal cortex of patients and healthy controls from two separate data sets of postmortem tissue samples. We reasoned that if the BD gene expression signature in the prefrontal cortical regulatory units is influenced by the activation (or repression) of specific TFs, then both the downstream targets, as well as the upstream regulators, of these TFs should be among the most differentially expressed genes in the BD phenotype. The use of network-based approaches to elucidate biological mechanisms of complex diseases may allow a clearer view of the molecular networks governing the pathophysiology and reveal potential targets for drug design and therapeutic intervention.

MATERIALS AND METHODS

Microarray data

The data used to reconstruct the transcriptional associations in the human prefrontal cortex was obtained from a large-scale microarray study describing an extensive series ($n = 269$) of brain tissue samples from fetal development through aging,¹⁶ downloaded from GEO (accession number GSE30272). Two studies using independent microarray platforms (Affymetrix Human Genome 133A GeneChips and Codelink Expression Assay arrays) were used to obtain the BD gene expression signatures (accession numbers GSE12679 and GSE5388). The two selected data sets were the best and most interesting sets available to our goal. As we aimed to evaluate differential expression of transcriptional regulatory units in the prefrontal cortex of bipolar disorder patients, we sought the most specific samples available (data set 1—data from pyramidal neurons) and other presenting more sensibility (data set 2—data from all cortical cells), both data sets from the same prefrontal area (Brodmann area 9—BA9). As such, the samples and raw data were independently processed and generated, contributing to check the consistency of the regulatory units characterized in our study.

For the first data set, the samples were obtained using laser-capture microdissection¹⁷ and included 1000 pyramidal neurons isolated from region BA9 of prefrontal cortex from five individuals with BD and six control subjects. For the second data set, fresh-frozen prefrontal cortex tissue (region BA9) was obtained from the Neuropathology Consortium of the Stanley brain collection (Stanley Medical Research Institute, Chevy Chase, MD, USA)¹⁸ and included 30 BD and 31 control samples. The demographic variables for the samples have been scrutinized in the original studies that produced the public data. For the first data set, the information available includes age (years): patients— 43.4 ± 6.3 , controls— 41.2 ± 7.2 ; gender (F/M): patients—1/4, controls—2/4; postmortem interval (hours): patients— 28.2 ± 5.4 , controls— 30 ± 18.6 . The study related to the second data set presents more detailed information that is summarized below. Age (years): patients— 44.5 ± 10.7 , controls— 43.8 ± 7.3 ; gender (F/M): patients—14/16, controls—7/24; brain pH: patients— 6.48 ± 0.27 , controls— 6.62 ± 0.27 ; postmortem interval (hours): patients— 37.2 ± 17.7 , controls— 29.1 ± 13.1 ; suicide (Y/N): patients—12/18, controls—0/31; alcohol abuse (none or light/moderate to severe): patients—13/17, controls—27/4. Regarding treatment of patients, lithium (Y/N): 8/22; valproate (Y/N): 10/20; electroconvulsive therapy history (Y/N): 2/28. Some potential confounding issues are not available to be controlled; however, they are addressed in the discussion.

Transcriptional network inference

The transcriptional networks were constructed using the R package RTN.¹⁹ Regarding the code availability used in this study, it is publicly available

from Bioconductor in the R packages RTN (<http://bioconductor.org/packages/RTN/>). Gene probes (complementary sequences to the target mRNAs used in microarray to assay gene expression) were filtered based on their coefficient of variation and mutual information was calculated in the R package *minet*.²⁰ The regulatory structure of the network is derived by mapping the significant interactions between known TFs and all potential targets in the gene expression matrix. The interactions that are below a minimum mutual information threshold are eliminated by permutation analysis. Unstable interactions are additionally removed by bootstrap analysis using 1000 bootstrap samples to create the consensus bootstrap network (that is, the relevance network). In an additional step, the Data Processing Inequality (DPI) algorithm is applied with tolerance = 0.0 to eliminate interactions that are likely to be mediated by another TF.¹¹ As the DPI removes the weakest edge of each network triplet, the vast majority of the interactions that are likely to be indirect are eliminated in this step. The resulting DPI-filtered transcriptional network is subsequently interrogated in the enrichment analysis. Both DPI-filtered and unfiltered transcriptional networks are used to visualize the final results. The analysis pipeline, resampling procedures and methods used to reconstruct the transcriptional networks are summarized in Supplementary Figure 1.

MR and gene set enrichment analysis

The MR analysis is described elsewhere.¹⁵ Briefly, the gene set enrichment analysis (GSEA) is used to assess whether a given transcriptional regulatory unit (regulon) is enriched for genes that are differentially expressed among two classes of microarrays. The GSEA uses a rank-based scoring metric obtained from the differentially expressed signatures to test the association between gene sets and the ranked phenotypic difference. The current analysis treats regulons as gene sets, and the BD signatures as the phenotype, in an extension of the GSEA analysis as previously described.²¹ The GSEA was performed in the R package RTN using 1000 permutations.

Two-tailed GSEA

The two-tailed GSEA assesses the direction of inferred connection between a given MR and the differentially expressed signatures, a proxy for induced or repressed associations. The method is based on the Connectivity Map procedure.²² The regulon is split into two subgroups, positive targets (*A*) and negative targets (*B*) using Pearson's correlation, whereas genes in the phenotype are ranked using the differentially expressed signatures (that is, top-down phenotype). The distribution of *A* and *B* are then tested by the GSEA statistics in the ranked phenotype, producing independent enrichment scores (ES) for each subgroup. A good separation of the two distributions and maximum deviation from zero near opposite extremes is desirable for a clear association. Therefore, an additional step is executed testing the differential enrichment ($ES_A - ES_B$). A high positive differential score indicates that the phenotype induced the regulon, whereas a high negative differential score indicates that the phenotype repressed the regulon. The two-tailed GSEA was performed in R using the function *tni.gsea2* in the RTN package with 1000 permutations.

Analysis of gene expression data

The Bioconductor package *limma*²³ was used to call differentially expressed genes, and the log fold change (logFC) metric was used to obtain the ranked phenotypes required for the GSEA analysis.

RESULTS

A tissue-specific regulatory network for the human prefrontal cortex

We first established a tissue-specific transcriptional network model computed from a large-scale human prefrontal microarray data set (transcriptional network reconstruction summarized in Figure 1a). The microarray data were pre-processed and probes with low variation were removed from the analysis. Two transcriptional networks (*TN1* and *TN2*) were then derived by computing the mutual information between annotated TFs and all potential targets in the data set. *TN1* represents the totality of the 269-microarray samples in the study, whereas *TN2* is derived from a subsample with adult human prefrontal cortex only (see Supplementary Figure 1 for additional details of the resampling

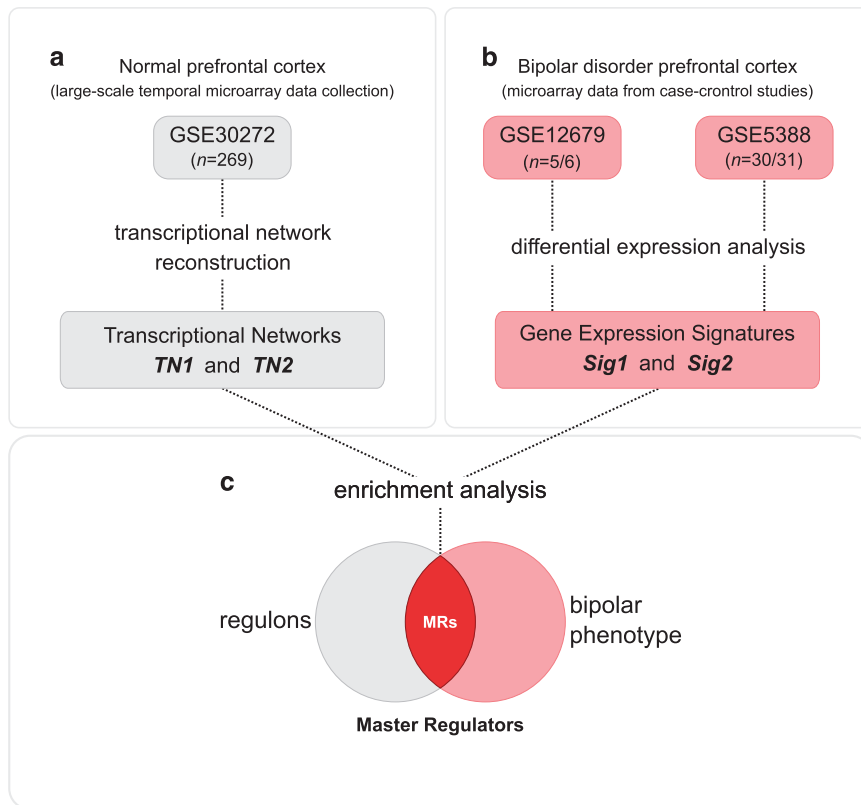


Figure 1. Master regulator (MR) analysis flowchart. **(a)** Data source used to reconstruct the transcriptional regulatory units in normal human prefrontal cortex. A large-scale microarray study (GSE30272) describing an extensive series of brain tissue from fetal development through aging was used to compute the transcription factor-centric regulatory networks (regulons). The transcriptional network *TN1* represents the totality of the 269-microarray samples in the study, whereas the *TN2* derives from a subsample with adult human prefrontal cortex only. The analysis pipeline, resampling procedures and methods used to reconstruct the transcriptional networks are further detailed in the Supplementary Figure 1. **(b)** Flowchart summarizing the microarray data used to obtain two independent bipolar disorder gene expression signatures (bipolar phenotypes). Signature 1 (*Sig1*; GSE12679) is derived from laser-capture microdissected human neurons isolated from postmortem dorsolateral prefrontal cortex, whereas signature 2 (*Sig2*; GSE5388) is derived from human postmortem brain tissue from adult subjects. **(c)** The enrichment analysis aims to identify transcriptional regulatory units associated with the gene expression signatures.

procedures). The association map in Figure 2a summarizes the transcriptional network *TN1* and shows the degree of similarity among the inferred regulatory units (regulons). The node size represents the number of targets in a given regulon, whereas edge width corresponds to the number of common targets between any two regulons assessed by the Jaccard coefficient (JC). In this reference network, each target can be linked to multiple TFs and regulation can occur as a result of both direct (TF–target) and indirect interactions (TF–TF–target). To preserve the dominant TF–target pairs for the subsequent enrichment analyses, we additionally applied the DPI algorithm, which removes the weakest interaction in any eventual triplet formed by two TFs and a common target gene (see methods and computational pipeline summarized in Supplementary Figure 1).

MRs in BD prefrontal cortex

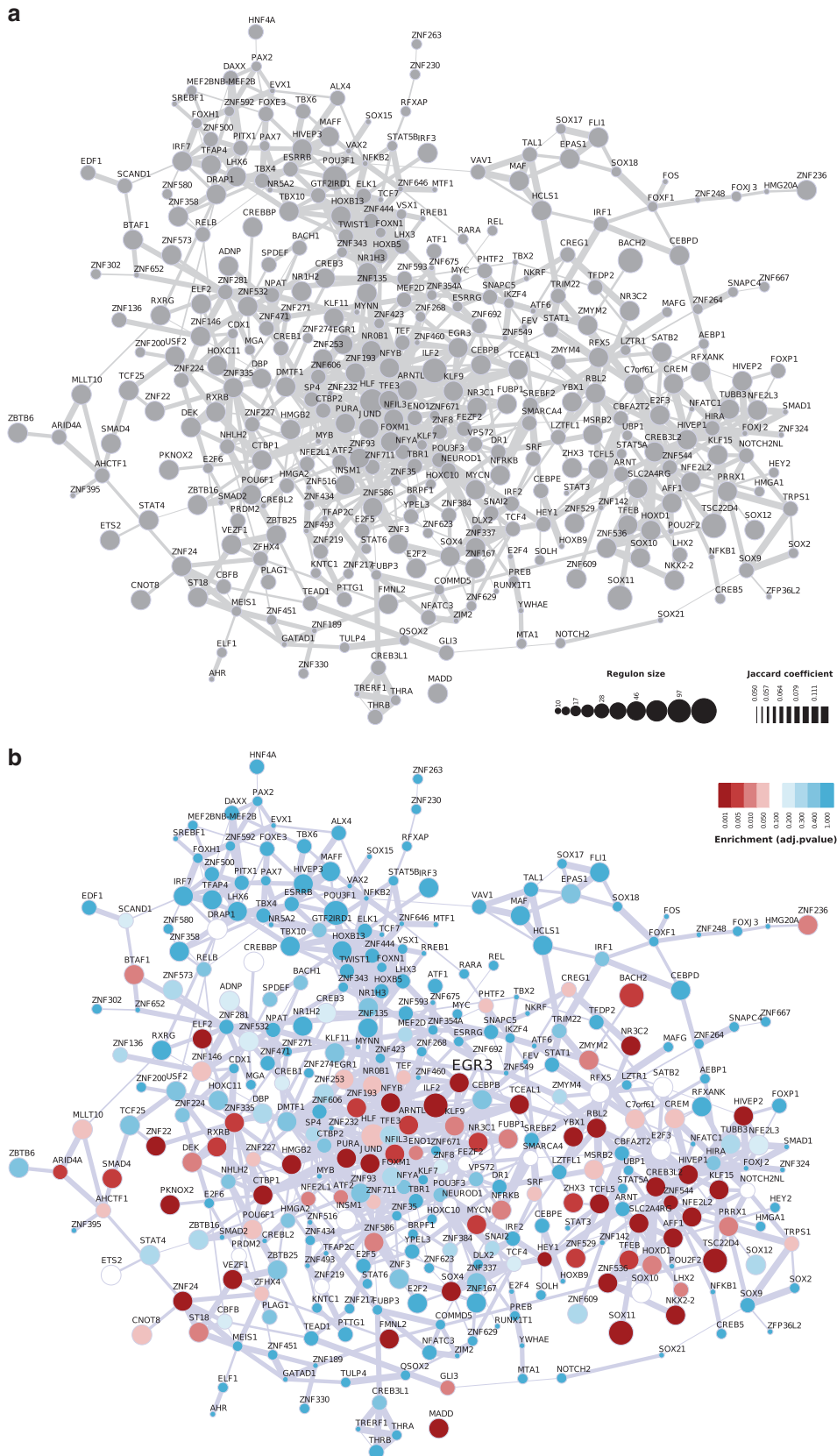
The inferred transcriptional network model was next used to query regulons enriched for the BD gene expression signatures. These signatures were obtained by differential expression analysis using microarray data from the two independent case–control sets (Figure 1b). Signature 1 (*Sig1*) is derived from laser-capture microdissected human neurons isolated from postmortem dorsolateral prefrontal cortex, whereas signature 2 (*Sig2*) is derived from human postmortem brain tissue from adult subjects. The MR analysis¹³ aims to identify regulons associated with the gene

expression signatures (Figure 1c). Therefore, our primary goal here is to generate hypotheses regarding the transcriptional regulation in BD, identifying the MRs responsible for coordinating the activity of the signature genes. We used the GSEA statistics to test the enrichment of the signature genes in each regulon. Figure 2b presents the results of the GSEA analysis using *TN1* and *Sig1* and shows the distribution of the BD phenotype onto the transcriptional association map.

Among the several candidates identified, 10 regulons were significantly enriched for both gene expression signatures (Figure 3a), five of which are consensus in both *TN1* and *TN2* transcriptional network models: early growth response protein 3 (*EGR3*), TSC22 domain family, member 4 (*TSC22D4*), interleukin enhancer-binding factor 2 (*ILF2*), Y-box binding protein 1 (*YBX1*) and MAP-kinase-activating death domain (*MADD*). When a high stringent threshold is applied, the overall consensus across all the tests is only obtained for the regulon of the *EGR3*. The GSEA plots in Figure 3b shows the distribution of the top-five consensus MRs in the BD phenotype.

Mode of action of the computationally defined regulons

To visualize the five MRs, we show in Figure 4a the correlation pattern observed between the TFs (square nodes) and its inferred targets (round nodes) assessed by the Pearson's correlation on *TN1*. This network graph shows all interactions inferred for each



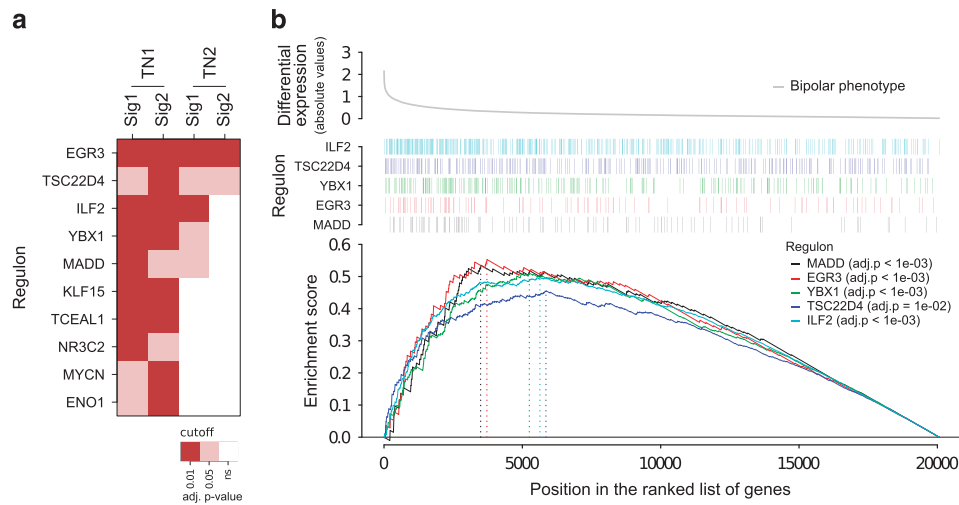


Figure 3. Consensus master regulators enriched for the bipolar disorder signatures. **(a)** Gene set enrichment statistics showing the regulatory units consistently enriched for the expression signatures *Sig1* and *Sig2* on the transcriptional network *TN1*, together with the results obtained for the same regulons on *TN2*. **(b)** Gene set enrichment plots showing the distribution of the top-five master regulators (that is, the consensus regulatory units) across the ranked bipolar phenotype represented by the absolute differential expression values (absolute logFC) derived from *Sig1*. The enrichment score is obtained based on the distribution of the hits: the x axis indicates the position of all genes ranked by the phenotype, and the hits indicate the position of each gene of a given regulon (see methods for additional description on the GSEA statistics). ns, not significant.

regulon, whether positive (red targets) or negative (blue targets). Using this information, we extended the gene set enrichment analysis in order to access how the mode of action of these regulons are connected with the BD gene signatures. We used a two-tailed GSEA statistics on regulons split into positive (red) and negative (blue) targets (Figure 4b), and the resulting distributions were tested against the BD phenotype ranked from the highest (+) to the lowest (–) differential expression values.

Accordingly, *EGR3* and *MADD* negative targets are associated with the positive phenotype (that is, most induced genes), whereas the positive targets are associated with negative phenotype (that is, most repressed genes), providing a high negative differential score with adjusted *P*-value < 0.001. It suggests that both *EGR3* and *MADD* regulons are repressed in the BD gene signature, whereas the other three regulons appear to be increased.

DISCUSSION

The primary goal of this study was to generate hypotheses regarding transcriptional regulation in BD, and to identify putative regulatory units that are dysfunctional in the prefrontal cortex of patients. To that end, we sought differentially expressed signatures that converged from two available gene expression data sets with distinctive strengths and weaknesses. Using MR analysis, our major finding was that the *EGR3* regulon was robustly repressed in both BD gene expression signatures. Four additional MRs showed a lower level of association with the BD phenotype.

EGR3 is a member of the *EGR* gene family of immediate early genes transcription factors. These genes are expressed at basal levels throughout the brain, including the cortex, hippocampus and other limbic areas, and the basal ganglia.²⁴ *EGR* expression is induced at high levels in response to environmental events and stressful stimuli across a range of intensities. In the brain, this activation is triggered by neurotransmitter-receptor stimulation or depolarization.²⁴ Numerous behavioral and electrophysiologic studies in animals have shown that the *EGR* family has a role in memory acquisition and consolidation and hippocampal synaptic plasticity.^{24–29} *EGR3*, in particular, is required for the

normal response to stress as well as in the neuroplasticity induced by this responsivity, ultimately regulating neuronal gene expression.^{27,30,31}

Of particular relevance are studies demonstrating that mice lacking functional *EGR3* display behavioral and physiologic abnormalities consistent with models of mental illness. These include a heightened response to stress (evidenced behaviorally and by elevated release of corticosterone), hyperactivity and failure to habituate to environmental stimuli and social cues.²⁷ The hyperactivity, a rodent psychosis phenotype, is reversible with antipsychotic medications used to treat BD.^{32,33} *EGR3* regulates expression of important plasticity associated genes, such as those encoding the activity regulated cytoskeletal associated gene (*Arc*)^{30,34} and GABA receptor subunit 4 (*GABRA4*),³⁵ and other member of the *EGR* family regulates the synaptic vesicle associated proteins *synapsin 1*³⁶ and *synapsin 2*.³⁷ Thus, requirement of *EGR3* in processes of memory, learning and synaptic plasticity is likely to be mediated by these, and presumably other as-yet unidentified, target effector genes.

The neuronal expression of *EGR3* is regulated by synaptic activity and is coupled to MAPK-ERK signaling.^{24,28} Gallitano-Mendel and colleagues noted that *EGR3* is activated downstream of numerous proteins associated with risk for psychotic illness, including neuregulin 1 (*NRG1*), calcineurin (*CN*) and *N*-methyl-D-aspartate (*NMDA*) receptors.^{27,32,38–41} Moreover, drugs that induce psychosis via serotonin 2A receptors (5-HT₂ARs) regulate expression of *EGR3*.⁴² They have hypothesized that these genes, together with targets of *EGR3*, comprise a pathway of proteins which, when disrupted at any level, increases risk for psychotic illness. In addition, brain-derived neurotrophic factor (*BDNF*) has been shown to induce *EGR3* expression via a PKC/MAPK-dependent pathway.³⁵ These are all interesting links, as *BDNF* has been proposed as a critical factor in the reduced cellular resilience associated to BD.^{43,44} A growing body of data has shown that peripheral *BDNF* levels are decreased during BD episodes and with the illness progression.^{45,46} Despite limitations in the studies and conflicting results in this area, it is intriguing to speculate that reduced peripheral *BDNF* levels, whether related to its decreased expression in the brain, may influence the *EGR3* repression as this

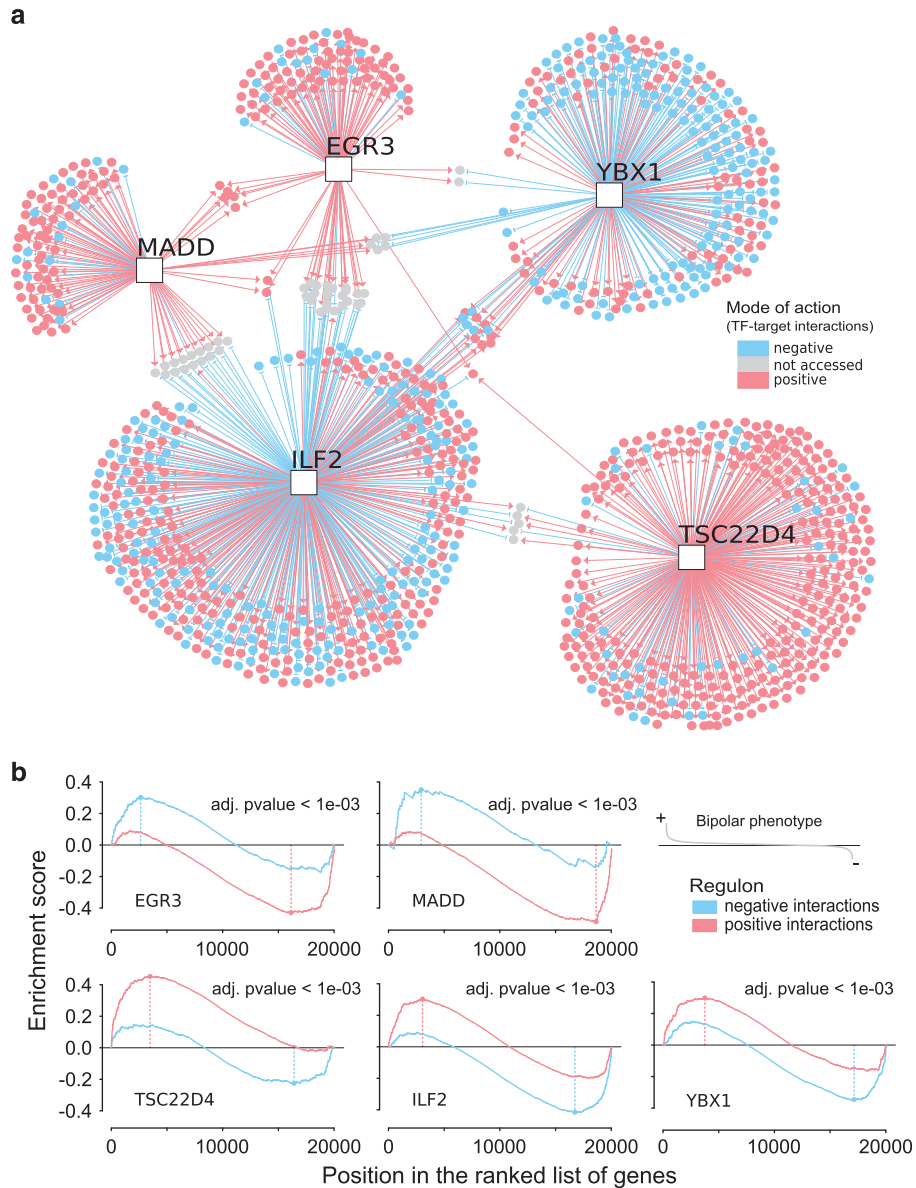


Figure 4. Regulatory units associated with the bipolar disorder phenotype. **(a)** The regulatory network shows the transcription factor (TF)–target interactions of the five master regulators, each one comprising one TF (square nodes) and all inferred targets (round nodes). The mode of action represented in red/blue colors corresponds to the correlation pattern observed between a given transcription factor and its targets, assessed by the Pearson's correlation on *TN1*. **(b)** Two-tailed gene set enrichment analysis. The enrichment plots show the distribution of the genes in each regulon across the ranked phenotype derived from *Sig1*. Regulons are split in positive (red) and negative (blue) targets, whereas the phenotype is ranked from the highest (+) to the lowest (–) differential expression values (logFC), that is, from the most increased to the most decreased gene expression values.

TF is regulated by *BDNF*. Obviously, further research is warranted for substantially improving the knowledge regarding the link between *BDNF* and *EGR3* in a shared biological pathway and their role in BD.

Downstream, *EGR3* targets the promoter region of genes involved in neuroplasticity or stimuli response. So far, experimental studies show effects on NMDA receptor subunits NR1 and NR2B, and type A GABA receptor, and possibly on genes involved in microglia deregulation associated with psychiatric disorders, such as the triggering receptor expressed on myeloid cells 1 (*TREM-1*).^{27,35,47,48} *EGR3* also regulates the expression of NGFR (*p75NTR*),⁴⁹ a receptor for neurotrophins that is involved in the regulation of axonal elongation.⁵⁰ Perhaps most intriguing is the recent findings that the *EGR3* target gene *ARC*, which modifies

synapses in response to environmental stimuli,⁵¹ is implicated in risk for psychotic disorders.^{41,52–54} Altogether, *EGR3* targets trigger different downstream genes and pathways involved in processes such as synaptic plasticity, axon extension, regulation of *BDNF* and receptors expression, among others.⁵⁵

EGRs translate environmental events into long-term changes in neural gene expression. This has led to the hypothesis that dysfunction in EGRs may account for both the genetic and environmental influences on risk for psychiatric illnesses.^{32,41,56} The *EGR* family has been more closely scrutinized in schizophrenia, with fewer and less-consistent studies in BD.^{47,56–58} *EGR3* has been significantly associated with schizophrenia in the Japanese,⁴⁰ Korean and Han Chinese populations,^{59,60} and recently in a population of European descent.⁴¹ Although not all studies have

found significant associations,^{57,61,62} a meta-analysis of the studies in Asian populations supported association between *EGR3* and schizophrenia.⁶⁰ In addition, the AA genotype of the *rs35201266* SNP was recently associated with the hemodynamic state of the prefrontal cortex in both patients with schizophrenia and healthy participants, possibly suggesting a pathway from neurodevelopment to brain function.⁶³ Furthermore, a study involving an entire network of TFs and microRNAs related to schizophrenia identified *EGR3* as the central gene in the regulatory network.⁵⁵

Studies examining a potential role for *EGR3* in BD identified nominal associations that did not meet the threshold for significance following the strict Bonferroni correction for multiple comparisons. In the first study, examining association of genes involved in circadian rhythms with BD, *EGR3* was the sole gene that achieved a significance level of $P < 0.05$.⁵⁷ A second investigation, a family-based association study identified a nominal association of *EGR3* with risk for child with BD I.⁵⁶ These findings suggest that *EGR3* may be a fruitful gene for future genetics studies to identify mechanisms by which environment and genetic predisposition interact to influence BD. Although the statistically significant findings supporting an association between *EGR3* and psychiatric illness have been in schizophrenia, research has increasingly demonstrated that the molecular and genetic processes underlying BD and schizophrenia are highly coincident.⁶⁴

The possible effects of mood stabilizers, psychotropic medications and substance use on the *EGR3* regulon are also an interesting point to be discussed. However, there are few studies on this matter, most of them evaluating other *EGR* genes. To our knowledge, there are no studies showing association between lithium or valproate effects and *EGR3* expression. In this context, studies observed that the expression of *EGR1* was increased by lithium in mouse frontal cortex,⁶⁵ and by valproate in neural stem cells.⁶⁶ Considering that lithium has been associated with neurogenesis, it is conceivable that it induces *EGR* genes as well. Nevertheless, our results point to the repression of this regulon in bipolar disorder, suggesting that lithium treatment did not influence our findings. Studies in rodents have shown that other psychoactive medications induce immediate early genes in the brain. For instance, chronic treatment with aripiprazole induces differential gene expression of *EGR1*, *EGR2* and *EGR4* in the rat frontal cortex;⁶⁷ *EGR1* is differentially expressed also in rat striatum after haloperidol and clozapine treatments.⁶⁸ Though less studied than *EGR1*, expression of *EGR3* is induced by several of the same stimuli of *EGR1*, including antipsychotic medications or drugs that induce psychosis.^{25,42} Other factors that might possibly affect the *EGR3* regulon is alcohol or substance use; a relationship between drug intake or withdrawal and induction of *EGR* genes has been reported. For instance, amphetamine and cocaine increase *EGR1* mRNA expression in the striatum,⁶⁹ cocaine also induce *EGR3* in the striatum²⁵ and amphetamine or alcohol withdrawal induce *EGR1* expression.⁷⁰ Considering these observations, it seems unlikely that antipsychotic treatment, alcohol or other substances are responsible for our findings, as they induce *EGR3* and other growth response genes, whereas our results pointed to the repression of the *EGR3* regulon in BD signatures.

Our data suggest that decreased function of *EGR3* may be involved in BD. As a MR of a network of genes and pathways that mediate critical neurobiological processes, dysfunction in *EGR3* indicates a possible explanation for both the influence of environment, as well as the role of numerous genes in the pathogenesis of BD. The identified network thus provides potential targets for follow-up experimental evaluation and development of novel therapeutics for this severe mental illness. The results presented here are both innovative and exploratory, and are therefore in need of confirmation before more definitive assertions regarding the relevance of *EGR3* in BD can be made. A new generation of bioinformatics methods has been developed to deal with the notorious limitations of functional genomics data.¹²

Nevertheless, further validation through basic science laboratory approaches, including mRNA expression of *EGR3* and key interacting genes in BD postmortem PFC tissue using PCR with reverse transcription, is an important step towards firmly confirming our results. However, studies in *EGR3*-deficient mice demonstrating psychosis-like phenotypes and hyperactivity that can be reversed with antipsychotic medications that are used in treatment of BD already provide important support for our findings.^{27,32,33}

Limitations of this study include the fact that only two microarray sets were used to obtain the signatures. Although we do not intend to perform an exhaustive analysis of all regions and all available data sets, our bioinformatics approach is constrained by the availability of a unique cohort study with a large sample size (to compute the regulatory unities) and the gene expression signatures interrogating the tissue under study. The starting point for our analysis was the public availability of a unique study, which sampled prefrontal cortex from people with bipolar disorder obtained using laser-capture microdissection. As this study was limited by sample size, and there are no other analyses using this technique, we next sought the largest available data set that used brain homogenates from the same prefrontal area (BA9), and we found just one study with these criteria. Hence the two data sets were formed not by all sets, but by the best and most interesting sets available. Other limitation is that the analysis was restricted to prefrontal cortex (BA9) and gene expression profiles might look different in other laminar and brain regions. Future studies should aim at evaluating laminar and regional specificity of our results, and validating our findings with biochemical/molecular analyses in independent biological samples, as well as studying *EGR3* targets, their role in BD and in the mechanisms of action of drugs. When it comes to the means of modulation of *EGR3*, it is likely that genetic and epigenetic mechanisms might be underlying the alterations seen in BD, which warrants further studies on this matter as well.

In conclusion, we have used an innovative approach based on MR analysis to study transcriptional regulation in BD. This method identified the *EGR3* gene as a potential key target, with the *EGR3* regulon robustly repressed in both of the two BD gene expression data sources we examined from postmortem prefrontal cortex. Considering that *EGR3* is activated throughout the brain in response to stressful environmental stimuli, the possible disruption in biological pathways involving *EGR3* may result in an impaired response and adaptation to stress. This could result in reduced neurobiological resilience and ultimately lead to the symptoms of executive and cognitive dysfunction seen in BD. The bioinformatics approach used in this work may give insights for identifying targets possibly involved with the risk for psychiatric disorders and inspire drug-discovery programs that can affect these disorders.

CONFLICT OF INTEREST

Professor Kapczynski has received grant/research support from Astra-Zeneca, Eli Lilly, the Janssen-Cilag, Servier, CNPq, CAPES, NARSAD and the Stanley Medical Research Institute; has been a member of the speakers' boards for Astra-Zeneca, Eli Lilly, Janssen and Servier; and has served as a consultant for Servier. The remaining authors declare no conflict of interest.

ACKNOWLEDGMENTS

This study was supported by the Brazilian funds CNPq/MS/SCTIE/DECIT—Pesquisas Sobre Doenças Neurodegenerativas (466989/2014-8) and MCT/CNPq INCT-TM (573671/2008-7), without interference in the ongoing work. BP received a fellowship from CAPES, Brazil; and MADB received a fellowship from CNPq, Brazil. We thank Gabriel R Fries for helpful discussions and critical reading of the manuscript.

REFERENCES

- Kieseppa T, Partonen T, Haukka J, Kaprio J, Lönqvist J. High concordance of bipolar I disorder in a nationwide sample of twins. *Am J Psychiatry* 2004; **161**: 1814–1821.
- Craddock N, Sklar P. Genetics of bipolar disorder. *Lancet* 2013; **381**: 1654–1662.
- International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
- Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 2011; **88**: 294–305.
- Gershon ES, Alliey-Rodriguez N, Liu C. After GWAS: searching for genetic risk for schizophrenia and bipolar disorder. *Am J Psychiatry* 2011; **168**: 253–256.
- Altar CA, Vawter MP, Ginsberg SD. Target identification for CNS diseases by transcriptional profiling. *Neuropsychopharmacology* 2009; **34**: 18–54.
- McCullumsmith RE, Meador-Woodruff JH. Novel approaches to the study of postmortem brain in psychiatric illness: old limitations and new challenges. *Biol Psychiatry* 2011; **69**: 127–133.
- Seifuddin F, Pirooznia M, Judy JT, Goes FS, Potash JB, Zandi PP. Systematic review of genome-wide gene expression studies of bipolar disorder. *BMC Psychiatry* 2013; **13**: 213.
- Hawrylycz M, Ng L, Page D, Morris J, Lau C, Faber S et al. Multi-scale correlation structure of gene expression in the brain. *Neural Netw* 2011; **24**: 933–942.
- Price JL, Drevets WC. Neural circuits underlying the pathophysiology of mood disorders. *Trends Cogn Sci* 2012; **16**: 61–71.
- Margolin AA, Wang K, Lim WK, Kustagi M, Nemenman I, Califano A. Reverse engineering cellular networks. *Nat Protoc* 2006; **1**: 662–671.
- Castro MA, de Santiago I, Campbell TM, Vaughn C, Hickey TE, Ross E et al. Regulators of genetic risk of breast cancer identified by integrative network analysis. *Nat Genet* 2016; **48**: 12–21.
- Carro MS, Lim WK, Alvarez MJ, Bollo RJ, Zhao X, Snyder EY et al. The transcriptional network for mesenchymal transformation of brain tumours. *Nature* 2010; **463**: 318–325.
- Piao G, Saito S, Sun Y, Liu ZP, Wang Y, Han X et al. A computational procedure for identifying master regulator candidates: a case study on diabetes progression in Goto-Kakizaki rats. *BMC Syst Biol* 2012; **6**(Suppl 1): S2.
- Fletcher MN, Castro MA, Wang X, de Santiago I, O'Reilly M, Chin SF et al. Master regulators of FGFR2 signaling and breast cancer risk. *Nat Commun* 2013; **4**: 2464.
- Colantuoni C, Lipska BK, Ye T, Hyde TM, Tao R, Leek JT et al. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* 2011; **478**: 519–523.
- Harris LW, Wayland M, Lan M, Ryan M, Giger T, Lockstone H et al. The cerebral microvasculature in schizophrenia: a laser capture microdissection study. *PLoS One* 2008; **3**: e3964.
- Ryan MM, Lockstone HE, Huffaker SJ, Wayland MT, Webster MJ, Bahn S. Gene expression analysis of bipolar disorder reveals downregulation of the ubiquitin cycle and alterations in synaptic genes. *Mol Psychiatry* 2006; **11**: 965–978.
- Castro MA, Wang X, Fletcher MN, Meyer KB, Markowitz F. RedeR: R/Bioconductor package for representing modular structures, nested networks and multiple levels of hierarchical associations. *Genome Biol* 2012; **13**: R29.
- Meyer PE, Lafitte F, Bontempi G. minet: A R/Bioconductor package for inferring large transcriptional networks using mutual information. *BMC Bioinformatics* 2008; **9**: 461.
- 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 USA* 2005; **102**: 15545–15550.
- Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 2006; **313**: 1929–1935.
- Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004; **3**: Article3.
- O'Donovan KJ, Tourtellotte WG, Millbrandt J, Baraban JM. The EGR family of transcription-regulatory factors: progress at the interface of molecular and systems neuroscience. *Trends Neurosci* 1999; **22**: 167–173.
- Yamagata K, Kaufmann WE, Lanahan A, Papapavlou M, Barnes CA, Andreasson KI et al. Egr3/Pilot, a zinc finger transcription factor, is rapidly regulated by activity in brain neurons and colocalizes with Egr1/zif268. *Learn Mem* 1994; **1**: 140–152.
- Jones MW, Errington ML, French PJ, Fine A, Bliss TV, Garel S et al. A requirement for the immediate early gene Zif268 in the expression of late LTP and long-term memories. *Nat Neurosci* 2001; **4**: 289–296.
- Gallitano-Mendel A, Izumi Y, Tokuda K, Zorumski CF, Howell MP, Muglia LJ et al. The immediate early gene early growth response gene 3 mediates adaptation to stress and novelty. *Neuroscience* 2007; **148**: 633–643.
- Li L, Yun SH, Keblesh J, Trommer BL, Xiong H, Radulovic J et al. Egr3, a synaptic activity regulated transcription factor that is essential for learning and memory. *Mol Cell Neurosci* 2007; **35**: 76–88.
- Cheval H, Chagneau C, Levasseur G, Veyrac A, Faucon-Biguier N, Laroche S et al. Distinctive features of Egr transcription factor regulation and DNA binding activity in CA1 of the hippocampus in synaptic plasticity and consolidation and reconsolidation of fear memory. *Hippocampus* 2012; **22**: 631–642.
- Li L, Carter J, Gao X, Whitehead J, Tourtellotte WG. The neuroplasticity-associated arc gene is a direct transcriptional target of early growth response (Egr) transcription factors. *Mol Cell Biol* 2005; **25**: 10286–10300.
- Maple AM, Zhao X, Elizalde DI, McBride AK, Gallitano AL. Htr2a expression responds rapidly to environmental stimuli in an Egr3-dependent manner. *ACS Chem Neurosci* 2015; **6**: 1137–1142.
- Gallitano-Mendel A, Wozniak DF, Pehek EA, Milbrandt J. Mice lacking the immediate early gene Egr3 respond to the anti-aggressive effects of clozapine yet are relatively resistant to its sedating effects. *Neuropsychopharmacology* 2008; **33**: 1266–1275.
- Williams AA, Ingram WM, Levine S, Resnik J, Kamel CM, Lish JR et al. Reduced levels of serotonin 2A receptors underlie resistance of Egr3-deficient mice to locomotor suppression by clozapine. *Neuropsychopharmacology* 2012; **37**: 2285–2298.
- Lyford GL, Yamagata K, Kaufmann WE, Barnes CA, Sanders LK, Copeland NG et al. Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. *Neuron* 1995; **14**: 433–445.
- Roberts DS, Hu Y, Lund IV, Brooks-Kayal AR, Russek SJ. Brain-derived neurotrophic factor (BDNF)-induced synthesis of early growth response factor 3 (Egr3) controls the levels of type A GABA receptor alpha 4 subunits in hippocampal neurons. *J Biol Chem* 2006; **281**: 29431–29435.
- Thiel G, Schoch S, Petersohn D. Regulation of synapsin I gene expression by the zinc finger transcription factor zif268/egr-1. *J Biol Chem* 1994; **269**: 15294–15301.
- Petersohn D, Schoch S, Brinkmann DR, Thiel G. The human synapsin II gene promoter. Possible role for the transcription factor zif268/egr-1, polyoma enhancer activator 3, and AP2. *J Biol Chem* 1995; **270**: 24361–24369.
- Hippenmeyer S, Shneider NA, Birchmeier C, Burden SJ, Jessell TM, Arber S. A role for neuregulin1 signaling in muscle spindle differentiation. *Neuron* 2002; **36**: 1035–1049.
- Gerber DJ, Hall D, Miyakawa T, Demars S, Gogos JA, Karayiorgou M et al. Evidence for association of schizophrenia with genetic variation in the 8p21.3 gene, PPP3CC, encoding the calcineurin gamma subunit. *Proc Natl Acad Sci USA* 2003; **100**: 8993–8998.
- 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 USA* 2007; **104**: 2815–2820.
- Huentelman MJ, Muppana L, Corneveaux JJ, Dinu V, Pruzin JJ, Reiman R et al. Association of SNPs in EGR3 and ARC with schizophrenia supports a biological pathway for schizophrenia risk. *PLoS One* 2015; **10**: e0135076.
- González-Maeso J, Yuen T, Ebersole BJ, Wurmbach E, Lira A, Zhou M et al. Transcriptome fingerprints distinguish hallucinogenic and nonhallucinogenic 5-hydroxytryptamine 2A receptor agonist effects in mouse somatosensory cortex. *J Neurosci* 2003; **23**: 8836–8843.
- Manji HK, Quiroz JA, Payne JL, Singh J, Lopes BP, Viegas JS et al. The underlying neurobiology of bipolar disorder. *World Psychiatry* 2003; **2**: 136–146.
- Berk M, Kapczinski F, Andreazza AC, Dean OM, Giorlando F, Maes M et al. Pathways underlying neuroprogression in bipolar disorder: focus on inflammation, oxidative stress and neurotrophic factors. *Neurosci Biobehav Rev* 2011; **35**: 804–817.
- Fernandes BS, Gama CS, Ceresér KM, Yatham LN, Fries GR, Colpo G et al. Brain-derived neurotrophic factor as a state-marker of mood episodes in bipolar disorders: a systematic review and meta-regression analysis. *J Psychiatr Res* 2011; **45**: 995–1004.
- Frey BN, Andreazza AC, Houenou J, Jamain S, Goldstein BI, Frye MA et al. Biomarkers in bipolar disorder: a positional paper from the International Society for Bipolar Disorders Biomarkers Task Force. *Aust N Z J Psychiatry* 2013; **47**: 321–332.
- Weigelt K, Carvalho LA, Drexhage RC, Wijkhuijs A, de Wit H, van Beveren NJ et al. TREM-1 and DAP12 expression in monocytes of patients with severe psychiatric disorders. EGR3, ATF3 and PU.1 as important transcription factors. *Brain Behav Immun* 2011; **25**: 1162–1169.
- Kim JH, Roberts DS, Hu Y, Lau GC, Brooks-Kayal AR, Farb DH et al. Brain-derived neurotrophic factor uses CREB and Egr3 to regulate NMDA receptor levels in cortical neurons. *J Neurochem* 2012; **120**: 210–219.
- Gao X, Daugherty RL, Tourtellotte WG. Regulation of low affinity neurotrophin receptor (p75(NTR)) by early growth response (Egr) transcriptional regulators. *Mol Cell Neurosci* 2007; **36**: 501–514.

- 50 Dechant G, Barde YA. The neurotrophin receptor p75(NTR): novel functions and implications for diseases of the nervous system. *Nat Neurosci* 2002; **5**: 1131–1136.
- 51 Farris S, Lewandowski G, Cox CD, Steward O. Selective localization of arc mRNA in dendrites involves activity- and translation-dependent mRNA degradation. *J Neurosci* 2014; **34**: 4481–4493.
- 52 Kirov G, Pocklington AJ, Holmans P, Ivanov D, Ikeda M, Ruderfer D *et al*. *De novo* CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry* 2012; **17**: 142–153.
- 53 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**: 179–184.
- 54 Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P *et al*. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* 2014; **506**: 185–190.
- 55 Guo AY, Sun J, Jia P, Zhao Z. A novel microRNA and transcription factor mediated regulatory network in schizophrenia. *BMC Syst Biol* 2010; **4**: 10.
- 56 Gallitano AL, Tillman R, Dinu V, Geller B. Family-based association study of early growth response gene 3 with child bipolar I disorder. *J Affect Disord* 2012; **138**: 387–396.
- 57 Mansour HA, Talkowski ME, Wood J, Chowdari KV, McClain L, Prasad K *et al*. Association study of 21 circadian genes with bipolar I disorder, schizoaffective disorder, and schizophrenia. *Bipolar Disord* 2009; **11**: 701–710.
- 58 Byrne EM, Heath AC, Madden PA, Pergadia ML, Hickie IB, Montgomery GW *et al*. Testing the role of circadian genes in conferring risk for psychiatric disorders. *Am J Med Genet B Neuropsychiatr Genet* 2014; **165B**: 254–260.
- 59 Kim SH, Song JY, Joo EJ, Lee KY, Ahn YM, Kim YS. EGR3 as a potential susceptibility gene for schizophrenia in Korea. *Am J Med Genet B Neuropsychiatr Genet* 2010; **153B**: 1355–1360.
- 60 Zhang R, Lu S, Meng L, Min Z, Tian J, Valenzuela RK *et al*. Genetic evidence for the association between the early growth response 3 (EGR3) gene and schizophrenia. *PLoS One* 2012; **7**: e30237.
- 61 Liu BC, Zhang J, Wang L, Li XW, Wang Y, Ji J *et al*. No association between EGR gene family polymorphisms and schizophrenia in the Chinese population. *Prog Neuropsychopharmacol Biol Psychiatry* 2010; **34**: 506–509.
- 62 Kyogoku C, Yanagi M, Nishimura K, Sugiyama D, Morinobu A, Fukutake M *et al*. Association of calcineurin A gamma subunit (PPP3CC) and early growth response 3 (EGR3) gene polymorphisms with susceptibility to schizophrenia in a Japanese population. *Psychiatry Res* 2011; **185**: 16–19.
- 63 Nishimura Y, Takizawa R, Koike S, Kinoshita A, Satomura Y, Kawasaki S *et al*. Association of decreased prefrontal hemodynamic response during a verbal fluency task with EGR3 gene polymorphism in patients with schizophrenia and in healthy individuals. *Neuroimage* 2014; **85**(Pt 1): 527–534.
- 64 Shinozaki G, Potash JB. New developments in the genetics of bipolar disorder. *Curr Psychiatry Rep* 2014; **16**: 493.
- 65 Kim SH, Yu HS, Park HG, Ahn YM, Kim YS, Lee YH *et al*. Egr1 regulates lithium-induced transcription of the Period 2 (PER2) gene. *Biochim Biophys Acta* 2013; **1832**: 1969–1979.
- 66 Almutawaa W, Kang NH, Pan Y, Niles LP. Induction of neurotrophic and differentiation factors in neural stem cells by valproic acid. *Basic Clin Pharmacol Toxicol* 2014; **115**: 216–221.
- 67 Cheng MC, Liao DL, Hsiung CA, Chen CY, Liao YC, Chen CH. Chronic treatment with aripiprazole induces differential gene expression in the rat frontal cortex. *Int J Neuropsychopharmacol* 2008; **11**: 207–216.
- 68 Nguyen TV, Kosofsky BE, Birnbaum R, Cohen BM, Hyman SE. Differential expression of c-fos and zif268 in rat striatum after haloperidol, clozapine, and amphetamine. *Proc Natl Acad Sci USA* 1992; **89**: 4270–4274.
- 69 Moratalla R, Robertson HA, Graybiel AM. Dynamic regulation of NGFI-A (zif268, egr1) gene expression in the striatum. *J Neurosci* 1992; **12**: 2609–2622.
- 70 Beckmann AM, Wilce PA. Egr transcription factors in the nervous system. *Neurochem Int* 1997; **31**: 477–510.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

Supplementary Information accompanies the paper on the Translational Psychiatry website (<http://www.nature.com/tp>)