Title: Gene set enrichment analysis and expression pattern exploration implicate an involvement of neurodevelopmental processes in bipolar disorder

Authors:

Thomas W. Mühleisen[a,b*], Céline S. Reinbold[b,*], Andreas J. Forstner[c,d], Lilia I. Abramova[e], Martin Alda[f,g], Gulja Babadjanova[h], Michael Bauer[i], Paul Brennan[j], Alexander Chuchalin[h], Cristiana Cruceanu[k,l,m], Piotr M. Czerski[n], Franziska Degenhardt[c,d], Sascha B. Fischer[b], Janice M. Fullerton[o,p], Scott D. Gordon[q], Maria Grigoroiu-Serbanescu[r], Paul Grof[s,t,u], Joanna Hauser[n], Martin Hautzinger[v], Stefan Herms[b,c,d], Per Hoffmann[b,c,d], Jutta Kammerer-Ciernioch[w], Elza Khusnutdinova[x,y], Manolis Kogevinas[z], Valery Krasnov[aa], André Lacour[ab], Catherine Laprise[ac,ad], Markus Leber[ae], Jolanta Lissowska[af], Susanne Lucae[ag], Anna Maaser[c,d], Wolfgang Maier[ah], Nicholas G. Martin[q], Manuel Mattheisen[ai,aj], Fermin Mayoral[ak], James D. McKay[al], Sarah E. Medland[q], Philip B. Mitchell[am,an], Susanne Moebus[ao], Grant W. Montgomery[q], Bertram Müller-Myhsok[ag,ap,aq], Lilijana Oruc[ar], Galina Pantelejeva[e], Andrea Pfennig[i], Lejla Pojskic[as], Alexey Polonikov[at,au], Andreas Reif[av], Fabio Rivas[ak], Guy A. Rouleau[k], Lorena M. Schenk[c,d], Peter R. Schofield[o,p], Markus Schwarz[aw], Fabian Streit[ax], Jana Strohmaier[ax], Neonila Szeszenia-Dabrowska[ay], Alexander S. Tiganov[e], Jens Treutlein[ax], Gustavo Turecki[l,m,az], Helmut Vedder[aw], Stephanie H. Witt[ax], Thomas G. Schulze[ba], Marcella Rietschel[ax], Markus M. Nöthen[c,d], Sven Cichon[a,b, c,d,§]

[*]These authors contributed equally to this work.

[a]Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, Germany

[b]Human Genomics Research Group and Division of Medical Genetics, Department of Biomedicine, University and University Hospital Basel, Basel, Switzerland

[c]Institute of Human Genetics, University of Bonn School of Medicine & University Hospital Bonn, Bonn, Germany

[d]Department of Genomics, Life & Brain Research Center, University of Bonn, Bonn, Germany

[e]Russian Academy of Medical Sciences, Mental Health Research Center, Moscow,

Russian Federation

[f]Department of Psychiatry, Dalhousie University, Halifax, Canada

[g]National Institute of Mental Health, Klecany, Czech Republic

[h]Institute of Pulmonology, Russian State Medical University, Moscow, Russian Federation

[i]Department of Psychiatry and Psychotherapy, University Hospital Carl Gustav Carus, TU Dresden, Dresden, Germany

[j]Genetic Epidemiology Group, International Agency for Research on Cancer (IARC), Lyon, France

[k]Montreal Neurological Institute, McGill University, Montreal, Canada

[l]Department of Human Genetics, McGill University, Montreal, Canada

[m]McGill Group for Suicide Studies & Douglas Research Institute, Montreal, Canada

[n]Laboratory of Psychiatric Genetics, Department of Psychiatry, Poznan University of Medical Sciences, Poznan, Poland

[o]Neuroscience Research Australia, Sydney, Australia

[p]School of Medical Sciences Faculty of Medicine, University of New South Wales, Sydney, Australia

[q]Queensland Institute of Medical Research (QIMR), Brisbane, Australia

[r]Biometric Psychiatric Genetics Research Unit, Alexandru Obregia Clinical Psychiatric Hospital, Bucharest, Romania

[s]The International Group for the Study of Lithium-Treated Patients (IGSLI), Berlin, Germany

[t]Mood Disorders Center of Ottawa, Ottawa, Ontario, Canada K1G 4G3

[u]Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada M5T 1R8

[v]Department of Psychology, Clinical Psychology and Psychotherapy, Eberhard Karls University Tübingen, Tübingen, Germany

[w]Center of Psychiatry Weinsberg, Weinsberg, Germany

[x]Institute of Biochemistry and Genetics, Ufa Scientific Center of Russian Academy of

Sciences, Ufa, Russian Federation

[y]Department of Genetics and Fundamental Medicine of Bashkir State University, Ufa, Russian Federation

[z]Center for Research in Environmental Epidemiology (CREAL), Barcelona, Spain

[aa]Moscow Research Institute of Psychiatry, Moscow, Russian Federation

[ab]German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany

[ac]Département des sciences fondamentales, Université du Québec à Chicoutimi, Saguenay, QC, Canada

[ad]Centre intégré universitaire de santé et services sociaux du Saguenay–Lac-Saint-Jean, Saguenay, Québec, Canada

[ae]Department of Psychiatry & Psychotherapy, University of Cologne, Cologne, Germany

[af]Department of Cancer Epidemiology and Prevention, Maria Sklodowska-Curie Memorial Cancer Centre and Institute of Oncology Warsaw, Warsaw, Poland

[ag]Max Planck Institute of Psychiatry, Munich, Germany

[ah]Department of Psychiatry, University of Bonn, Bonn, Germany

[ai]Department of Biomedicine and Centre for integrative Sequencing, iSEQ, Aarhus University, Aarhus, Denmark

[aj]The Lundbeck Foundation Initiative for integrative Psychiatric Research, iPSYCH, Aarhus and Copenhagen, Denmark

[ak]Department of Psychiatry, Hospital Regional Universitario, Biomedical Institute of Malaga, Malaga, Spain

[al]Genetic Cancer Susceptibility Group, International Agency for Research on Cancer (IARC), Lyon, France

[am]School of Psychiatry, University of New South Wales, Randwick, Australia

[an]Black Dog Institute, Prince of Wales Hospital, Randwick, Australia

[ao]Institute of Medical Informatics, Biometry and Epidemiology, University Duisburg-Essen, Essen, Germany

[ap]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany

[aq]University of Liverpool, Institute of Translational Medicine, Liverpool, United Kingdom

[ar]Psychiatric Clinic, Clinical Center University of Sarajevo, Bolnička 25, 71000 Sarajevo, Bosnia and Herzegovina

[as]Institute for Genetic Engineering and Biotechnology, University of Sarajevo, Zmaja od Bosne 8 - Campus, 71000 Sarajevo, Bosnia and Herzegovina

[at]Department of Biology, Medical Genetics and Ecology, Kursk State Medical University, Kursk, Russian Federation

[au]Research Institute for Genetic and Molecular Epidemiology, Kursk State Medical University, Kursk, Russian Federation

[av]Department of Psychiatry, Psychosomatic Medicine and Psychotherapy, University Hospital Frankfurt am Main, Frankfurt am Main, Germany

[aw]Psychiatric Center Nordbaden, Wiesloch, Germany

[ax]Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Germany

[ay]Department of Epidemiology, Nofer Institute of Occupational Medicine, Lodz, Poland

[az]Department of Psychiatry, McGill University, Montreal, Canada

[ba]Institute of Psychiatric Phenomics and Genomics, Ludwig-Maximilians-University Munich, Munich, Germany

[§]**Corresponding author:** Sven Cichon, PhD. Department of Biomedicine, University of Basel, Hebelstrasse 20, CH-4031 Basel, Switzerland. Phone: +41 61 265 23 60. Fax: +41 61 265 36 21. E-mail: sven.cichon@unibas.ch.

Keywords:

- Bipolar disorder
- Pathway analysis
- *GRB2 events in ERBB2 signaling*
- *NCAM signaling for neurite out-growth*
- Neurodevelopmental disorder

INTRODUCTION

Bipolar disorder (BD) is a genetically complex mental illness. During the past ten years, several genome-wide association studies (GWAS) of BD were conducted and have identified 19 loci harboring common genetic susceptibility variants (Sullivan et al., 2017). It is assumed that with growing sample sizes the number of loci will increase, as has been successfully demonstrated for schizophrenia, where GWAS in 61,000 patients found 155 independent loci (Sullivan et al., 2017).

Gene set enrichment analysis (GSEA) is a powerful tool to retrieve more biological information from existing GWAS. Such multi-locus approaches utilize functional frameworks of ontologies or pathways to integrate genes at strongly to moderately associated loci. Using the same sample size, GSEA therefore has greater statistical power to detect a polygenic contribution of individually small effects to overall risk than single-locus analyses (Lee et al., 2012).

Here, we applied GSEA algorithms to a large published GWAS on BD, including approximately 9,700 patients and 14,200 controls (Mühleisen et al., 2014). We found associations between BD and two signaling pathways involved in brain development.

METHODS AND MATERIALS

Phenotype and SNP Data

For GSEA, we used combined data from the Systematic Investigation of the Molecular Causes of Major Mood Disorders and Schizophrenia (MooDS) and Psychiatric Genomics Consortium (PGC) consortia comprising 2,267,487 autosomal single-nucleotide polymorphisms (SNPs) from 9,747 patients with life-time diagnoses of BD and 14,278 controls, as described by Mühleisen et al. (2014). Written informed consent was obtained from all patients and controls before participation in the study.

Gene Set Enrichment Analyses

For discovery, we used Meta-Analysis Gene-set Enrichment of variaNT Associations (MAGENTA; (Segrè et al., 2010)) with its default settings. At genome-wide level, each gene was mapped to the GWAS SNP showing the lowest p-value within gene boundaries (RefSeq definitions), to minimize the effect of a potential confounding factor introduced by overlapping gene boundaries (Sedeño-Cortés and Pavlidis, 2014). P-values of these index SNPs were corrected for confounders such as gene size, SNP density and linkage disequilibrium-related properties in the stepwise multiple linear regression model of MAGENTA. Resulting gene scores were assigned to target gene sets. For each target gene set, the observed number of gene scores above the user-defined threshold (here 95%) is evaluated against the expected number of gene scores above this threshold for gene sets of identical size, randomly sampled from the genome multiple times. A non-parametric test produces the nominal p-value for each tested target gene set. False-discovery rate (FDR) was used to correct for multiple testing (P_{FDR}).

For secondary analysis of the significantly enriched pathways, we applied Gene Set Analysis SNP (GSA-SNP; (Nam et al., 2010)) on the same input data. GSA-SNP uses p-values of SNPs to calculate enrichment scores by using the Z-statistic method. But instead of using the maximum effect per gene as a proxy for the respective gene, we chose the second-best p-value to represent the effect of each gene to avoid spurious associations (Kwon et al., 2012).

For pathways, we used curated target gene sets (pathways) from Reactome as available through the Molecular Signature Database (v6.0; (Subramanian et al., 2005). Their sizes were restricted from 20 to 200 to avoid overly narrow or broad gene sets. This resulted in 397 sets for GSEA.

Gene expression data

BrainScope enables interactive visual exploration of spatial and temporal human brain transcriptomes from the Allen Institute for Brain Science (Huisman et al., 2017). Here we focused on the dataset *Developmental Transcriptome* from the BrainSpan atlas that had been pre-processed and re-analyzed by BrainScope's developers resulting in the dataset *Developing human (comparative explorer)* with RNA expression levels of 18,233 genes (Entrez Gene definitions) that were z-score normalized, to have a zero mean and a standard deviation of 1.

To explore changes of co-expressed genes in brain regions and time windows, we used heat maps of the comparative explorer from BrainScope under default settings. Each square of a heat map displayed the average regional expression of the selected gene(s) across pooled tissue samples (replicates, developmental stages) from donor brains (controls). For BD-associated pathways from GSEA results, heat maps were assembled and annotated using standard graphical software. The brain regions covered neocortical areas including primary cortices (auditory, motor, somatosensory, visual), pre- and orbitofrontal cortices, the temporal cortex (inferolateral, posterior superior), the parietal cortex (posteroventral); principal structures of the diencephalon including parts of the basal ganglia (amygdala, striatum) and limbic system (anterior cingulate, amygdala, hippocampus) coiled around and connected to thalamus and hypothalamus; the hindbrain (cerebellar cortex). The time windows comprised fetal development (from early 2nd trimester to birth), infancy (from birth to one year), childhood (from two to eleven years), adolescence (from 13 to 19 years), and adulthood (from 21 to 40 years). BrainScope, BrainSpan, and Entrez Gene are publicly accessible at www.brainscope.nl, www.brainspan.org, and www.ncbi.nlm.nih.gov/gene.

RESULTS

Discovery and validation of BD-associated pathways

GSEA by MAGENTA on MooDS-PGC data revealed two study-wide significant Reactome pathways when applying the significance criterion of FDR < 0.05 (**Table 1**). The best finding was *GRB2 events in ERBB2 signaling* (P_{FDR}=0.0377), for which 6 genes were associated (*NRAS*, *KRAS*, *EGFR*, *ERBB2*, *MAPK1*, *HBEGF*) out of 21 in the pathway. The second finding was *NCAM signaling for neurite out-growth* (P_{FDR}=0.0451) for which 11 of 62 genes were associated (*NCAN*, *SPTBN2*, *FYN*, *NRAS*, *CREB1*, *KRAS*, *CACNB3*, *COL2A1*, *CACNB2*, *MAPK1*, *SPTBN1*). Three significant genes were common to both pathways (*NRAS*, *KRAS*, *MAPK1*). The associated genes showed a balanced contribution to the total significance of the two target gene sets (**Supplementary Table 1**). The subsequent GSEA by GSA-SNP on the same input data validated the enrichments in the two target gene sets (P=4.80E-06 and P=3.28E-08, respectively; **Table 1**).

Exploration of gene expression in BD-associated pathways

To assess patterns of co-expressed genes from both pathways in the developing and adult brain, we used data from BrainSpan accessed through BrainScope and screened (i) expression of each single associated gene, (ii) expression of the combined set of associated genes (Combined), and (iii) expression of associated genes in context of target gene sets (Whole pathway). We found that five of the six genes enriched in *GRB2 events in ERBB2 signaling* demonstrated expression peaks during fetal development and infancy, while *MAPK1* expression was lower during prenatal stages and higher during postnatal stages. The combined pattern of the six genes emphasized neural development and was similar to the whole pathway pattern. In *NCAM signaling for neurite out-growth*, four of the ten enriched genes (*NCAN*, *FYN*, *NRAS*, *CREB1*) revealed high expression during fetal and early postnatal development. *CACNB2*, *MAPK1*, *SPTBN1*, and *SPTBN2* showed low expression during fetal stages but increased later on, especially in infancy. Overall, most genes showed peaks of co-expression during fetal development (early second to third trimester) and infancy (birth to 18 months) in many neocortical areas and parts of the limbic system. Spatio-temporal expression patterns of genes stratified by pathway are displayed in **Figure 1**.

DISCUSSION

Current disease models of BD suggest a multifactorial etiology resulting from the additive effects of many gene variants at different loci together with the effect of environmental factors. GWAS have demonstrated that the genotype relative risks of the involved common susceptibility variants are small and that large sample sizes are necessary to achieve sufficient statistical power to identify them (Sullivan et al., 2017). In the present analysis, we chose to apply GSEA to our GWAS data because this approach should have greater statistical power to detect a polygenic contribution of individually small effects to overall risk than single-locus analyses (Lee et al., 2012). To further strengthen our findings, we investigated genes within the implicated pathways for expression at milestones of normal brain development to obtain information on their relevance during ontogenetic stages. Biological pathway studies of BD so far have found evidence for genes involved in calcium channels, hormonal regulation, glutamate signaling, neural development, and histone methylation (Nurnberger et al., 2014; O'Dushlaine et al., 2015).

Our strongest finding was *GRB2 events in ERBB2 signaling* which functions to promote cell proliferation, survival, and differentiation, not only in the brain. Biologically, an association with *ERBB2*, *EGFR*, and *HBEGF* is plausible because they form a ligand-activated receptor complex for signaling and thus seem to be key players of that pathway. The importance of *ERBB2* in BD is further supported by a genome-wide significant association finding (Hou et al., 2016) and by the observation of dysregulated *ERBB2* expression in the dorsolateral prefrontal cortex in both BD and schizophrenia (Shao and Vawter, 2008). This expression alteration is significantly related to lifetime antipsychotic exposure, supporting *ERBB2* as target for clinical research. *ErbB2/B4*-deficient mice exhibit elevated aggression and reduced prepulse inhibition that both can be rescued by clozapine treatment, a frequently used antipsychotic medication (Barros et al., 2009). *EGFR* (alias *ERBB1*) is reported to play an essential role in axon myelination during the first postnatal weeks and can therefore be considered as an important regulator of neurodevelopment (Aguirre et al., 2007). The gene was also supported by single SNP and haplotype analysis in a GWAS of BD (Sklar et al., 2008). HBEGF is a EGF-like binding partner of EGFR and mice lacking *Hb-egf* in the ventral forebrain showed abnormalities in psychomotor behavior and neurotransmission which can be ameliorated by typical or atypical antipsychotics (Oyagi et al., 2009).

Our second finding was *NCAM signaling for neurite out-growth* which modulates neural differentiation and synaptic plasticity. Homophilic binding of NCAM1 molecules at the cell-surface induces signaling that leads to cell-cell adhesion and axon elongation. Association with *NCAN* in this pathway is of major importance since experiments in rats have demonstrated that interference of Ncam1-Ncam1 bindings by concurrent Ncan inhibits these cellular processes (Retzler et al., 1996). *NCAN* encodes an extracellular matrix proteoglycan and has been described as important susceptibility gene for BD (Cichon et al., 2011). Furthermore, *NCAN* was reported to be associated with brain development in health and disease, specifically to gray matter loss in central limbic regions and higher folding in the lateral occipital and prefrontal cortex suggesting impairments of emotion perception and regulation and top-down cognitive functioning (Dannlowski et al., 2015). Behavioral abnormalities in *Ncan*-deficient mice show striking similarities with mania symptoms in humans that can be rescued by lithium treatment, an established mood stabilizer (Miró et al., 2012). Association with *CACNB2* and *CACNB3* represents another highlight of this pathway, since abnormal calcium channel activity is considered to be important for BD (Nurnberger et al., 2014). Unexpectedly, *CACNA1C* was not found among enriched pathway genes, despite strong support of this gene from SNP data. Further evaluation revealed that *CACNA1C* was absent from the pathway definition. A possible link to our other finding exists through a gene overlap of *KRAS*, *NRAS* and *MAPK1* as well as binding between *NCAM1* and *EGFR*.

In both pathways, most genes showed high co-expression during fetal development and infancy in many neocortical and subcortical areas indicating co-expression and possibly co-working of encoded protein functions. These observations provide links to brain regions where known pathophysiological changes in BD patients occur, for instance, in the limbic system which is concerned with many aspects of emotion and behavior.

Limitations

Although both pathway findings were technically reproduced by two different approaches and are based on one of the largest GWAS data of BD so far, association replication in independent samples was not attempted. Gene expression was explored in control brains only, which may show co-expression differences compared with BD brains. Follow-up studies are required to further evaluate the relevance of our findings for etiological and clinical aspects of BD.

Conclusion

The present study found evidence for associations between BD and two signaling pathways. Integration of evidence from genetic studies, brain developmental expression patterns and molecular functions of these pathways support the hypothesis that neurodevelopmental processes play an important role in the etiology of BD.

REFERENCES

Aguirre, A., Dupree, J.L., Mangin, J.M., Gallo, V., 2007. A functional role for EGFR signaling in myelination and remyelination. Nature Neuroscience 10, 990-1002.

Barros, C.S., Calabrese, B., Chamero, P., Roberts, A.J., Korzus, E., Lloyd, K., Stowers, L., Mayford, M., Halpain, S., Muller, U., 2009. Impaired maturation of dendritic spines without disorganization of cortical cell layers in mice lacking NRG1/ErbB signaling in the central nervous system. Proceedings of the National Academy of Sciences 106, 4507-4512.

Cichon, S., Mühleisen, T.W., Degenhardt, F.A., Mattheisen, M., Miró, X., Strohmaier, J., Steffens, M., Meesters, C., Herms, S., Weingarten, M., Priebe, L., Haenisch, B., Alexander, M., Vollmer, J., Breuer, R., Schmäl, C., Tessmann, P., Moebus, S., Wichmann, H.E., Schreiber, S., Müller-Myhsok, B., Lucae, S., Jamain, S., Leboyer, M., Bellivier, F., Etain, B., Henry, C., Kahn, J.-P., Heath, S., Hamshere, M., O'Donovan, M.C., Owen, M.J., Craddock, N., Schwarz, M., Vedder, H., Kammerer-Ciernioch, J., Reif, A., Sasse, J., Bauer, M., Hautzinger, M., Wright, A., Mitchell, P.B., Schofield, P.R., Montgomery, G.W., Medland, S.E., Gordon, S.D., Martin, N.G., Gustafsson, O., Andreassen, O., Djurovic, S., Sigurdsson, E., Steinberg, S., Stefansson, H., Stefansson, K., Kapur-Pojskic, L., Oruc, L., Rivas, F., Mayoral, F., Chuchalin, A., Babadjanova, G., Tiganov, A.S., Pantelejeva, G., Abramova, L.I., Grigoroiu-Serbanescu, M., Diaconu, C.C., Czerski, P.M., Hauser, J., Zimmer, A., Lathrop, M., Schulze, T.G., Wienker, T.F., Schumacher, J., Maier, W., Propping, P., Rietschel, M., Nöthen, M.M., 2011. Genome-wide Association Study Identifies Genetic Variation in Neurocan as a Susceptibility Factor for Bipolar Disorder. The American Journal of Human Genetics 88, 372-381.

Dannlowski, U., Kugel, H., Grotegerd, D., Redlich, R., Suchy, J., Opel, N., Suslow, T., Konrad, C., Ohrmann, P., Bauer, J., Kircher, T., Krug, A., Jansen, A., Baune, B.T., Heindel, W., Domschke, K., Forstner, A.J., Nöthen, M.M., Treutlein, J., Arolt, V., Hohoff, C., Rietschel, M., Witt, S.H., 2015. NCAN Cross-Disorder Risk Variant Is Associated With Limbic Gray Matter Deficits in Healthy Subjects and Major Depression. Neuropsychopharmacology, 2510-2516.

Hou, L., Bergen, S.E., Akula, N., Song, J., Hultman, C.M., Landén, M., Adli, M., Alda, M., Ardau, R., Arias, B., Aubry, J.-M., Backlund, L., Badner, J.A., Barrett, T.B., Bauer,

M., Baune, B.T., Bellivier, F., Benabarre, A., Bengesser, S., Berrettini, W.H., Bhattacharjee, A.K., Biernacka, J.M., Birner, A., Bloss, C.S., Brichant-Petitjean, C., Bui, E.T., Byerley, W., Cervantes, P., Chillotti, C., Cichon, S., Colom, F., Coryell, W., Craig, D.W., Cruceanu, C., Czerski, P.M., Davis, T., Dayer, A., Degenhardt, F., Del Zompo, M., DePaulo, J.R., Edenberg, H.J., Étain, B., Falkai, P., Foroud, T., Forstner, A.J., Frisén, L., Frye, M.A., Fullerton, J.M., Gard, S., Garnham, J.S., Gershon, E.S., Goes, F.S., Greenwood, T.A., Grigoroiu-Serbanescu, M., Hauser, J., Heilbronner, U., Heilmann-Heimbach, S., Herms, S., Hipolito, M., Hitturlingappa, S., Hoffmann, P., Hofmann, A., Jamain, S., Jiménez, E., Kahn, J.-P., Kassem, L., Kelsoe, J.R., Kittel-Schneider, S., Kliwicki, S., Koller, D.L., König, B., Lackner, N., Laje, G., Lang, M., Lavebratt, C., Lawson, W.B., Leboyer, M., Leckband, S.G., Liu, C., Maaser, A., Mahon, P.B., Maier, W., Maj, M., Manchia, M., Martinsson, L., McCarthy, M.J., McElroy, S.L., McInnis, M.G., McKinney, R., Mitchell, P.B., Mitjans, M., Mondimore, F.M., Monteleone, P., Mühleisen, T.W., Nievergelt, C.M., Nöthen, M.M., Novák, T., Nurnberger, J.I., Nwulia, E.A., Ösby, U., Pfennig, A., Potash, J.B., Propping, P., Reif, A., Reininghaus, E., Rice, J., Rietschel, M., Rouleau, G.A., Rybakowski, J.K., Schalling, M., Scheftner, W.A., Schofield, P.R., Schork, N.J., Schulze, T.G., Schumacher, J., Schweizer, B.W., Severino, G., Shekhtman, T., Shilling, P.D., Simhandl, C., Slaney, C.M., Smith, E.N., Squassina, A., Stamm, T., Stopkova, P., Streit, F., Strohmaier, J., Szelinger, S., Tighe, S.K., Tortorella, A., Turecki, G., Vieta, E., Volkert, J., Witt, S.H., Wright, A., Zandi, P.P., Zhang, P., Zollner, S., McMahon, F.J., 2016. Genome-wide association study of 40,000 individuals identifies two novel loci associated with bipolar disorder. Human Molecular Genetics 25, 3383-3394.

Huisman, S.M.H.H., van Lew, B., Mahfouz, A., Pezzotti, N., Höllt, T., Michielsen, L., Vilanova, A., Reinders, M.J.T.T., Lelieveldt, B.P.F.F., van Lew, B., Mahfouz, A., Pezzotti, N., Höllt, T., Michielsen, L., Vilanova, A., Reinders, M.J.T.T., Lelieveldt, B.P.F.F., 2017. BrainScope: interactive visual exploration of the spatial and temporal human brain transcriptome. Nucleic acids research, gkx046.

Kwon, J.-s., Kim, J., Nam, D., Kim, S., 2012. Performance Comparison of Two Gene Set Analysis Methods for Genome-wide Association Study Results: GSA-SNP vs i-GSEA4GWAS. Genomics Inform 10, 123-127.

Lee, S.H., Yang, J., Goddard, M.E., Visscher, P.M., Wray, N.R., 2012. Estimation of

pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. Bioinformatics 28, 2540-2542.

Miró, X., Meier, S., Dreisow, M.L., Frank, J., Strohmaier, J., Breuer, R., Schmäl, C., Albayram, O., Pardo-Olmedilla, M.T., Mühleisen, T.W., Degenhardt, F.A., Mattheisen, M., Reinhard, I., Bilkei-Gorzo, A., Cichon, S., Seidenbecher, C., Rietschel, M., Nöthen, M.M., Zimmer, A., 2012. Studies in humans and mice implicate neurocan in the etiology of mania. The American journal of psychiatry 169, 982-990.

Mühleisen, T.W., Leber, M., Schulze, T.G., Strohmaier, J., Degenhardt, F., Treutlein, J., Mattheisen, M., Forstner, A.J., Schumacher, J., Breuer, R., Meier, S., Herms, S., Hoffmann, P., Lacour, A., Witt, S.H., Reif, A., Müller-Myhsok, B., Lucae, S., Maier, W., Schwarz, M., Vedder, H., Kammerer-Ciernioch, J., Pfennig, A., Bauer, M., Hautzinger, M., Moebus, S., Priebe, L., Czerski, P.M., Hauser, J., Lissowska, J., Szeszenia-Dabrowska, N., Brennan, P., McKay, J.D., Wright, A., Mitchell, P.B., Fullerton, J.M., Schofield, P.R., Montgomery, G.W., Medland, S.E., Gordon, S.D., Martin, N.G., Krasnow, V., Chuchalin, A., Babadjanova, G., Pantelejeva, G., Abramova, L.I., Tiganov, A.S., Polonikov, A., Khusnutdinova, E., Alda, M., Grof, P., Rouleau, G.a., Turecki, G., Laprise, C., Rivas, F., Mayoral, F., Kogevinas, M., Grigoroiu-Serbanescu, M., Propping, P., Becker, T., Rietschel, M., Nöthen, M.M., Cichon, S., 2014. Genome-wide association study reveals two new risk loci for bipolar disorder. Nature communications 5, 3339.

Nam, D., Kim, J., Kim, S.-Y., Kim, S., 2010. GSA-SNP: a general approach for gene set analysis of polymorphisms. Nucleic Acids Research 38, W749-W754.

Nurnberger, J.I., Jr., Koller, D.L., Jung, J., Edenberg, H.J., Foroud, T., Guella, I., Vawter, M.P., Kelsoe, J.R., Psychiatric Genomics Consortium Bipolar Group, 2014. Identification of pathways for bipolar disorder: a meta-analysis. JAMA Psychiatry 71, 657-664.

O'Dushlaine, C., Rossin, L., Lee, P.H., Duncan, L., Parikshak, N.N., Newhouse, S., Ripke, S., Neale, B.M., Purcell, S.M., Posthuma, D., Nurnberger, J.I., Lee, S.H., Faraone, S.V., Perlis, R.H., Mowry, B.J., Thapar, A., Goddard, M.E., Witte, J.S., Absher, D., Agartz, I., Akil, H., Amin, F., Andreassen, O.A., Anjorin, A., Anney, R., Anttila, V., Arking, D.E., Asherson, P., Azevedo, M.H., Backlund, L., Badner, J.A., Bailey, A.J., Banaschewski, T., Barchas, J.D., Barnes, M.R., Barrett, T.B., Bass, N., Battaglia, A.,

Bauer, M., Bayés, M., Bellivier, F., Bergen, S.E., Berrettini, W., Betancur, C., Bettecken, T., Biederman, J., Binder, E.B., Black, D.W., Blackwood, D.H.R., Bloss, C.S., Boehnke, M., Boomsma, D.I., Breuer, R., Bruggeman, R., Cormican, P., Buccola, N.G., Buitelaar, J.K., Bunney, W.E., Buxbaum, J.D., Byerley, W.F., Byrne, E.M., Caesar, S., Cahn, W., Cantor, R.M., Casas, M., Chakravarti, A., Chambert, K., Choudhury, K., Cichon, S., Mattheisen, M., Cloninger, C.R., Collier, D.A., Cook, E.H., Coon, H., Cormand, B., Corvin, A., Coryell, W.H., Craig, D.W., Craig, I.W., Crosbie, J., Cuccaro, M.L., Curtis, D., Czamara, D., Datta, S., Dawson, G., Day, R., De Geus, E.J., Degenhardt, F., Djurovic, S., Donohoe, G.J., Doyle, A.E., Duan, J., Dudbridge, F., Duketis, E., Ebstein, R.P., Edenberg, H.J., Elia, J., Ennis, S., Etain, B., Fanous, A., Farmer, A.E., Ferrier, I.N., Flickinger, M., Fombonne, E., Foroud, T., Frank, J., Franke, B., Fraser, C., Freedman, R., Freimer, N.B., Freitag, C.M., Friedl, M., Frisén, L., Gallagher, L., Gejman, P.V., Georgieva, L., Gershon, E.S., Giegling, I., Gill, M., Gordon, S.D., Gordon-Smith, K., Green, E.K., Greenwood, T.A., Grice, D.E., Gross, M., Grozeva, D., Guan, W., Gurling, H., De Haan, L., Haines, J.L., Hakonarson, H., Hallmayer, J., Hamilton, S.P., Hamshere, M.L., Hansen, T.F., Hartmann, A.M., Hautzinger, M., Heath, A.C., Henders, A.K., Herms, S., Hickie, I.B., Hipolito, M., Hoefels, S., Holsboer, F., Hoogendijk, W.J., Hottenga, J.-J., Hultman, C.M., Hus, V., Ingason, A., Ising, M., Jamain, S., Jones, E.G., Jones, I., Jones, L., Tzeng, J.-Y., Kähler, A.K., Kahn, R.S., Kandaswamy, R., Keller, M.C., Kennedy, J.L., Kenny, E., Kent, L., Kim, Y., Kirov, G.K., Klauck, S.M., Klei, L., Knowles, J.A., Kohli, M.A., Koller, D.L., Konte, B., Korszun, A., Krabbendam, L., Krasucki, R., Kuntsi, J., Kwan, P., Landén, M., Längström, N., Lathrop, M., Lawrence, J., Lawson, W.B., Leboyer, M., Ledbetter, D.H., Lencz, T., Lesch, K.-P., Levinson, D.F., Lewis, C.M., Li, J., Lichtenstein, P., Lieberman, J.A., Lin, D.-Y., Linszen, D.H., Liu, C., Lohoff, F.W., Loo, S.K., Lord, C., Lowe, J.K., Lucae, S., MacIntyre, D.J., Madden, P.A.F., Maestrini, E., Magnusson, P.K.E., Mahon, P.B., Maier, W., Malhotra, A.K., Mane, S.M., Martin, C.L., Martin, N.G., Matthews, K., Mattingsdal, M., McCarroll, S.A., McGhee, K.A., McGough, J.J., McGrath, P.J., McGuffin, P., McInnis, M.G., McIntosh, A., McKinney, R., McLean, A.W., McMahon, F.J., McMahon, W.M., McQuillin, A., Medeiros, H., Medland, S.E., Meier, S., Melle, I., Meng, F., Meyer, J., Middeldorp, C.M., Middleton, L., Milanova, V., Miranda, A., Monaco, A.P., Montgomery, G.W., Moran, J.L., Moreno-De-Luca, D., Morken, G., Morris, D.W., Morrow, E.M., Moskvina, V., Muglia, P., Mühleisen, T.W., Muir, W.J., Müller-Myhsok, B., Murtha, M., Myers, R.M., Myin-Germeys, I., Neale, M.C., Nelson, S.F., Nievergelt, C.M., Nikolov, I., Nimgaonkar, V.,

Nolen, W.A., Nöthen, M.M., Nwulia, E.A., Nyholt, D.R., Oades, R.D., Olincy, A., Oliveira, G., Olsen, L., Ophoff, R.A., Osby, U., Owen, M.J., Palotie, A., Parr, J.R., Paterson, A.D., Pato, C.N., Pato, M.T., Penninx, B.W., Pergadia, M.L., Pericak-Vance, M.A., Pickard, B.S., Pimm, J., Piven, J., Potash, J.B., Poustka, F., Propping, P., Puri, V., Quested, D.J., Quinn, E.M., Ramos-Quiroga, J.A., Rasmussen, H.B., Raychaudhuri, S., Rehnström, K., Reif, A., Ribasés, M., Rice, J.P., Rietschel, M., Roeder, K., Roeyers, H., Rothenberger, A., Rouleau, G., Ruderfer, D., Rujescu, D., Sanders, A.R., Sanders, S.J., Santangelo, S.L., Sergeant, J.A., Schachar, R., Schalling, M., Schatzberg, A.F., Scheftner, W.A., Schellenberg, G.D., Scherer, S.W., Schork, N.J., Schulze, T.G., Schumacher, J., Schwarz, M., Scolnick, E., Scott, L.J., Shi, J., Shilling, P.D., Shyn, S.I., Silverman, J.M., Slager, S.L., Smalley, S.L., Smit, J.H., Smith, E.N., Sonuga-Barke, E.J.S., Clair, D.S., State, M., Steffens, M., Steinhausen, H.-C., Strauss, J.S., Strohmaier, J., Stroup, T.S., Sutcliffe, J.S., Szatmari, P., Szelinger, S., Thirumalai, S., Thompson, R.C., Todorov, A.A., Tozzi, F., Treutlein, J., Uhr, M., van den Oord, E.J.C.G., Van Grootheest, G., Van Os, J., Vicente, A.M., Vieland, V.J., Vincent, J.B., Visscher, P.M., Walsh, C.A., Wassink, T.H., Watson, S.J., Weissman, M.M., Werge, T., Wienker, T.F., Wijsman, E.M., Willemsen, G., Williams, N., Willsey, A.J., Witt, S.H., Xu, W., Young, A.H., Yu, T.W., Zammit, S., Zandi, P.P., Zhang, P., Zitman, F.G., Zöllner, S., Devlin, B., Kelsoe, J.R., Sklar, P., Daly, M.J., O'Donovan, M.C., Craddock, N., Kendler, K.S., A Weiss, L., Wray, N.R., Zhao, Z., Geschwind, D.H., Sullivan, P.F., Smoller, J.W., Holmans, P.A., Breen, G., 2015. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nature Neuroscience 18, 199-209.

Oyagi, A., Oida, Y., Kakefuda, K., Shimazawa, M., Shioda, N., Moriguchi, S., Kitaichi, K., Nanba, D., Yamaguchi, K., Furuta, Y., Fukunaga, K., Higashiyama, S., Hara, H., 2009. Generation and characterization of conditional heparin-binding EGF-like growth factor knockout mice. PLoS ONE 4.

Retzler, C., Göhring, W., Rauch, U., 1996. Analysis of Neurocan Structures Interacting with the Neural Cell Adhesion Molecule N-CAM. Journal of Biological Chemistry 271, 27304-27310.

Sedeño-Cortés, A.E., Pavlidis, P., 2014. Pitfalls in the application of gene-set analysis to genetics studies. Trends in Genetics 30, 513-514.

Segrè, A.V., Consortium, D., investigators, M., Groop, L., Mootha, V.K., Daly, M.J., Altshuler, D., 2010. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. PLoS genetics 6, e1001058.

Shao, L., Vawter, M.P., 2008. Shared Gene Expression Alterations in Schizophrenia and Bipolar Disorder. Biological Psychiatry 64, 89-97.

Sklar, P., Smoller, J.W., Fan, J., Ferreira, M.A.R., Perlis, R.H., Chambert, K., Nimgaonkar, V.L., McQueen, M.B., Faraone, S.V., Kirby, A., De Bakker, P.I.W., Ogdie, M.N., Thase, M.E., Sachs, G.S., Todd-Brown, K., Gabriel, S.B., Sougnez, C., Gates, C., Blumenstiel, B., Defelice, M., Ardlie, K.G., Franklin, J., Muir, W.J., McGhee, K.A., MacIntyre, D.J., McLean, A., VanBeck, M., McQuillin, A., Bass, N.J., Robinson, M., Lawrence, J., Anjorin, A., Curtis, D., Scolnick, E.M., Daly, M.J., Blackwood, D.H., Gurling, H.M., Purcell, S.M., 2008. Whole-genome association study of bipolar disorder. Molecular Psychiatry 13, 558-569.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., Mesirov, J.P., 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences 102, 15545-15550.

Sullivan, P.F., Agrawal, A., Bulik, C., Andreassen, O.A., Borglum, A., Breen, G., Cichon, S., Edenberg, H., Faraone, S.V., Gelernter, J., Mathews, C.A., Nievergelt, C.M., Smoller, J., O'Donovan, M., 2017. Psychiatric Genomics: An Update and an Agenda. bioRxiv.

TABLES AND FIGURES

Figure 1 Expression patterns of genes in BD-associated pathways during normal brain development. Each square of a heat map displays the spatio-temporal expression of the selected gene(s) in the indicated regions and stages in control brains. Levels of RNA expression are z-score normalized ranging from blue (low) over white (zero mean) to red (high). Patterns are shown for single enriched genes (gene symbols), the combined set of enriched genes (combined), and the target gene set (whole pathway).

ERBB2 is a member of the EGF receptor family. Since ERBB2 has no ligand-binding domain, it needs a co-receptor to become activated. Upon binding of an EGF ligand, the ERBB2-EGFR heterodimer recruits adaptor protein GRB2 leading to SOS1-mediated guanine-nucleotide exchange on RAS (KRAS, NRAS) and activation of RAF and the MAP kinase cascade (MAPK1).

NCAM1 works on modulation of intracellular signaling, either by activation of FGF receptors or cytoplasmic tyrosine kinases (FYN) that initiate MAP kinase cascades (MAPK1) and a transcription factor (CREB1) which regulates expression of genes for growth and survival of neurites. Spectrins (SPTBN1, SPTBN2) are cytoskeletal molecules and manage to link RPTP-alpha to the cytoplasmic domain of NCAM1. L-type channels (*CACNB2*, *CACNB3*) associate with NCAM1 in growth cones at the sites of NCAM1 clusters leading to processes that promote neurite out-growth.

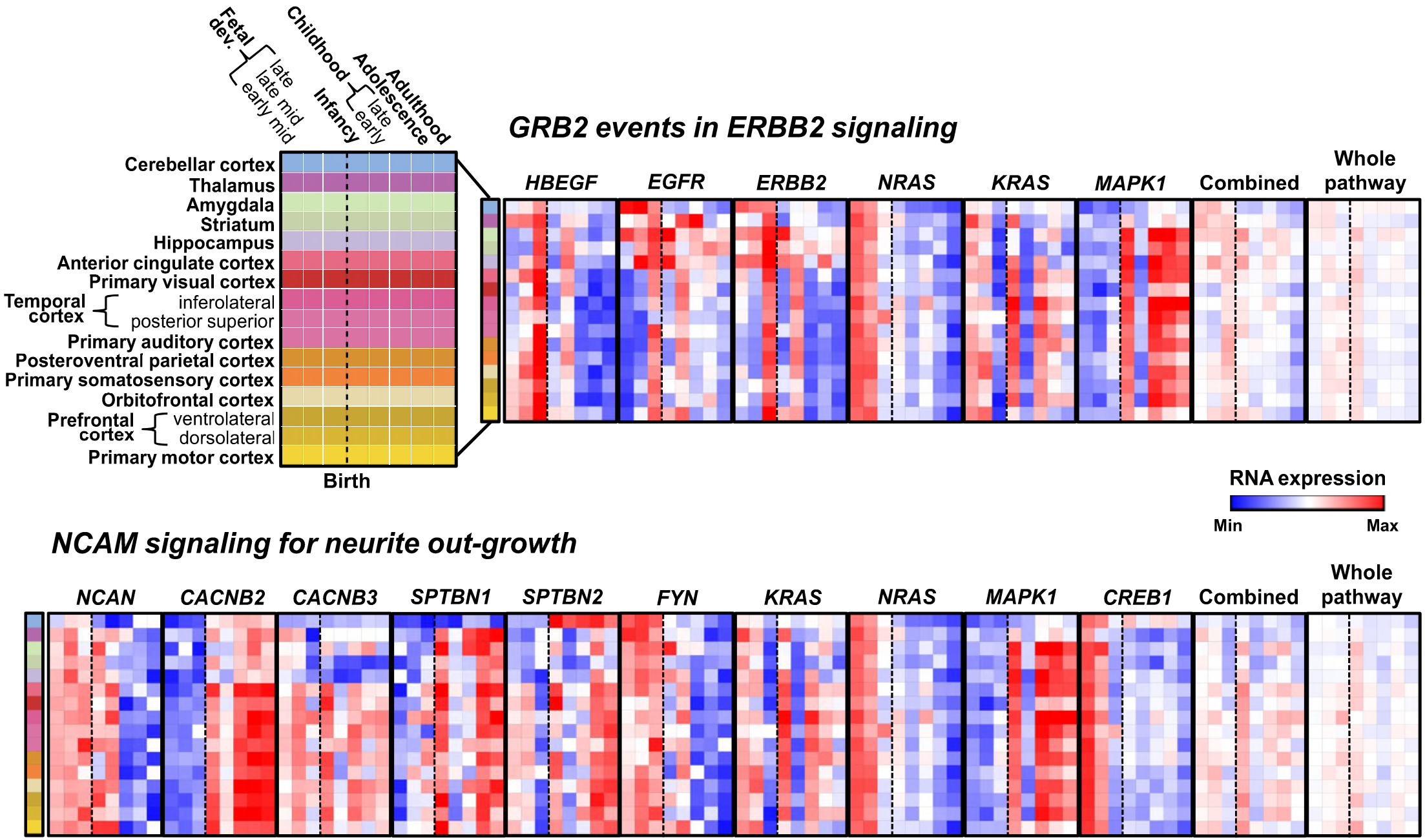
Supplementary Table 1 Study-wide significant pathway results identified in MAGENTA.

Abbreviations: Chr, cytogenetic band of the chromosome; kb, kilobases; * MAGENTA-based gene P value; † GWAS-based results.

Table 1 Association results of the GSEA. MAGENTA and GSA-SNP were used for discovery and validation steps.

			MAGENTA, 95th percentile enrichment cutoff		**GSA-SNP**, 2nd best SNP
Gene set name	**Gene set identifier**	***N* genes**	**P_{FDR}**	**Sign. genes (gene p-value)**	**Empirical p-value**
GRB2 events in ERBB2 signaling	R-HSA-1963640	21	0.0377	*NRAS* (1.94E-03), *KRAS* (2.20E-03), *EGFR* (6.18E-03), *ERBB2* (0.0196), *MAPK1* (0.0222), *HBEGF* (0.0306)	4.80E-06
NCAM signaling for neurite out-growth	R-HSA-375165	62	0.0451	*NCAN* (1.40E-05), *SPTBN2* (6.64E-05), *FYN* (2.75E-04), *NRAS* (1.94E-03), *CREB1* (2.11E-03), *KRAS* (2.20E-03), *CACNB3* (4.86E-03), *COL2A1* (0.0127), *CACNB2* (0.0138), MAPK1 (0.0222), *SPTBN1* (0.0251)	3.28E-08

Abbreviations: *CACNB2*, calcium voltage-gated channel auxiliary subunit beta 2; *CACNB3*, calcium voltage-gated channel auxiliary subunit beta 3; *COL2A1*, collagen type II alpha 1 chain; *CREB1*, cAMP responsive element binding protein 1; *EGF*, epidermal growth factor; *EGFR*, EGF receptor; *ERBB2*, Erb-B2 receptor tyrosine kinase 2; *FYN*, FYN proto-oncogene; *GRB2*, Growth factor receptor-bound protein 2; *HBEGF*, heparin-binding EGF-like growth factor; *KRAS*, KRAS proto-oncogene, GTPase; *MAPK1*, mitogen-activated protein kinase 1; N, number; *NCAM1*, neural cell adhesion molecule 1; *NCAN*, neurocan; *NRAS*, neuroblastoma RAS Viral oncogene homolog; P_{FDR}, FDR-adjusted p-value; *SPTBN1*, spectrin beta, non-erythrocytic 1; *SPTBN2*, spectrin beta, non-erythrocytic 2.


Fetal dev.
late
late mid
early mid
Childhood
Infancy
Adolescence
late
early
Adulthood
Cerebellar cortex
Thalamus
Amygdala
Striatum
Hippocampus
Anterior cingulate cortex
Primary visual cortex
Temporal cortex
inferolateral
posterior superior
Primary auditory cortex
Posteroventral parietal cortex
Primary somatosensory cortex
Orbitofrontal cortex
Prefrontal cortex
ventrolateral
dorsolateral
Primary motor cortex
Birth
GRB2 events in ERBB2 signaling
HBEGF
EGFR
ERBB2
NRAS
KRAS
MAPK1
Combined
Whole pathway
RNA expression
Min
Max
NCAM signaling for neurite out-growth
NCAN
CACNB2
CACNB3
SPTBN1
SPTBN2
FYN
KRAS
NRAS
MAPK1
CREB1
Combined
Whole pathway

DECLARATION OF INTEREST

The authors declare no competing financial interests.

CONTRIBUTORS

MooDS management: M.M.N., M.R., S.C.

Coordinated the work and prepared the manuscript, with feedback from the other authors who approved the manuscript: T.W.M., C.S.R., S.C.

Bioinformatical analysis: C.S.R., T.W.M., A.L.

Statistical analysis: C.S.R., M.L., M.M., F.S.

Sample management, DNA extraction, genotyping, QC in Bonn and Mannheim: F.D., S.B.F., A.J.F., S.H., P.H., A.M., L.M.S., J.S., J.T., S.H.W., M.M.N. (initiation of patient recruitment, supervision of genotyping, obtaining funding), M.R. (initiation and supervision of patient recruitment, obtaining funding)

Recruitment and characterization of patients and controls and generation of control data for Australia: J.M.F., S.D.G., N.G.M., S.E.M., P.B.M., G.W.M., P.R.S.

Recruitment and characterization of patients and controls and generation of control data for Canada: M.A., C.C., C.L., G.A.R., P.G., G.T.

Recruitment and characterization of patients and controls and generation of control data for Germany I and II: M.B., M.G.-S., M.H., J.K.-C., S.L., W.M., S.M., B.M.-M., L.O., A.P., L.R., A.R., T.G.S., M.S., J.S., H.V.

Recruitment and characterization of patients and controls and generation of control data for Poland: P.B., P.M.C., J.H., J.L., J.D.M., N.S.-D.

Recruitment and characterization of patients and controls and generation of control data for Russia: L.I.A., G.B., A.C., E.K., V.K., G.P., A.Po., A.S.T.

Recruitment and characterization of patients and controls and generation of control data for Spain: M.K., F.M., F.R.

ROLE OF THE FUNDING SOURCE

The study was supported by the German Federal Ministry of Education and Research (BMBF) through the Integrated Network IntegraMent (Integrated Understanding of Causes and Mechanisms in Mental Disorders), under the auspices of the e:Med Programme (grant 01ZX1314A to M.M.N., F.D. and S.C., grant 01ZX1314G to M.R.). M.M.N. is a member of the DFG-funded Excellence-Cluster ImmunoSensation. M.M.N. also received support from the Alfried Krupp von Bohlen und Halbach-Stiftung. The study was supported by the German Research Foundation (DFG; grant FOR2107; RI908/11-1 to M.R.; NO246/ 10-1 to M.M.N.). The study was also supported by the Swiss National Science Foundation (SNSF, grant 156791 to S.C.). M.G.S. received grant no. 89/2012 from UEFISCDI, Romania. Canadian patients were genotyped within the ConLiGen project (www.ConLiGen.org), with the support of a grant from the DFG to M.R., M.B., and T.G.S. (RI 908/7-1). Controls for Germany II were drawn from the Heinz Nixdorf Recall Study (HNR) cohort, which was established with the support of the Heinz Nixdorf Foundation. Recruitment of the Australian sample was supported by an Australian National Health and Medical Research Council (NHMRC) program grant (number 1037196). The recruitment of the Canadian patients was supported by a grant from the Canadian Institutes of Health Research #64410 to M.A. The study also used data generated by the GABRIEL consortium (controls for the sample Russia). Funding for the generation of these data was provided by the European Commission as part of GABRIEL contract number 018996 under the Integrated Program LSH-2004-1.2.5-1. Post genomic approaches to understand the molecular basis of asthma aiming at a preventive or therapeutic control and the Wellcome Trust under award 084703. Canadian controls were drawn from the French Canadian study (SLSJ), which was supported in part by the Canada Research Chair Environment and genetics of respiratory diseases and allergy, the Canadian Institutes of Health Research (Operating grant No. MOP-13506), and the Quebec Respiratory Network of the Fonds de rechercheen Santé du Québec (FRQS). Polish controls were recruited by the International Agency for Research on Cancer (IARC)/Centre National de Genotypage (CNG) GWAS Initiative. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We are grateful to all patients and control subjects who contributed to this study. We thank the Bipolar Disorder Working Groups of the MooDS and PGC consortia for providing access to the relevant data. C.L. is the Canada Research Chair in Environment and Genetics of Respiratory Disorders and Allergy, Director of the Asthma Strategic Group of the Respiratory Health Network (RHN) of Fonds de la recherche en santé du Québec (FRSQ) and researcher of the CHILD Study.