HAL
open science

Molecular Pathways involved in Neuronal Cell Adhesion and Membrane Scaffolding contribute to Schizophrenia and Bipolar Disorder Susceptibility.

Aiden Corvin, Colm O'Dushlaine, Elaine Kenny, Eleisa Heron, Gary Donohoe, Derek Morris, Michael Gill, The International Schizophrenia Consortium

To cite this version:

Aiden Corvin, Colm O'Dushlaine, Elaine Kenny, Eleisa Heron, Gary Donohoe, et al.. Molecular Pathways involved in Neuronal Cell Adhesion and Membrane Scaffolding contribute to Schizophrenia and Bipolar Disorder Susceptibility.. Molecular Psychiatry, 2010, 10.1038/mp.2010.7 . hal-00509758

HAL Id: hal-00509758

https://hal.science/hal-00509758v1

Submitted on 16 Aug 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Molecular Pathways involved in Neuronal Cell Adhesion and Membrane Scaffolding contribute to Schizophrenia and Bipolar Disorder Susceptibility.

***Colm O'Dushlaine, PhD;[1] *Elaine Kenny, PhD;[1] Eleisa Heron, PhD;[1] Gary Donohoe, DClinPsyc, PhD;[1] Derek Morris, PhD;[1] Michael Gill, MRCPsych, MD;[1] The International Schizophrenia Consortium[2] & Aiden Corvin, MRCPsych, PhD;[1]**

1. Department of Psychiatry, Trinity College Dublin, Dublin, Ireland

2. The International Schizophrenia Consortium:

Trinity College Dublin Derek W. Morris, Colm T.O'Dushlaine, Elaine Kenny, Emma M. Quinn, Michael Gill, Aiden Corvin; **Cardiff University** Michael C.O'Donovan, George K. Kirov, Nick J. Craddock, Peter A. Holmans, Nigel M.Williams, Lucy Georgieva, Ivan Nikolov, N. Norton, H. Williams, Draga Toncheva,Vihra Milanova, Michael J. Owen; **Karolinska Institutet/University of North Carolina at Chapel Hill** Christina M. Hultman, Paul Lichtenstein, Emma F.Thelander, Patrick Sullivan; **University College London** Andrew McQuillin, Khalid Choudhury, Susmita Datta, Jonathan Pimm, Srinivasa Thirumalai, Vinay Puri, Robert Krasucki, Jacob Lawrence, Digby Quested, Nicholas Bass, Hugh Gurling; **University of Aberdeen** Caroline Crombie, Gillian Fraser, Soh Leh Kuan, Nicholas Walker, David St Clair; **University of Edinburgh** Douglas H. R. Blackwood, Walter J. Muir, Kevin A. McGhee, Ben Pickard, Pat Malloy, Alan W. Maclean, Margaret Van Beck;**Queensland Institute of Medical Research** Naomi R. Wray, Peter M. Visscher, Stuart Macgregor; **University of Southern California** Michele T. Pato, Helena Medeiros, Frank Middleton, Celia Carvalho, Christopher Morley, AymanFanous, David Conti, James A. Knowles, Carlos Paz Ferreira, AntonioMacedo, M. Helena Azevedo, Carlos N. Pato; **Massachusetts General Hospital** Jennifer L. Stone, Douglas M. Ruderfer, Manuel A. R. Ferreira, **Stanley Center for Psychiatric Research and Broad Institute of MIT and Harvard** Shaun M. Purcell, Jennifer L. Stone, Kimberly Chambert, Douglas M. Ruderfer, Finny Kuruvilla, Stacey B. Gabriel, Kristin Ardlie, Mark J. Daly, Edward M. Scolnick, Pamela Sklar.

Correspondence: Dr Aiden Corvin, Dept of Psychiatry, Trinity Centre for Health Sciences, St James's Hospital, Dublin 8, Dublin, Ireland + 353 1 896 2468 (FAX +353 1 896 3405) (acorvin@tcd.ie).

***These authors contributed equally to this work**

Running title: Genetic pathway analysis in Psychotic Disorders

Funding Support: This work was supported by Science Foundation Ireland and the Health Research Board.

The authors report no conflict of interest.

Susceptibility to schizophrenia and bipolar disorder may involve a substantial, shared contribution from thousands of common genetic variants each of small effect. Identifying if risk variants map to specific molecular pathways is potentially biologically informative. We report a molecular pathway analysis using the SNP ratio test (SRT) which compares the ratio of nominally significant ($p<0.05$) to non-significant SNPs in a given pathway to identify 'enrichment' for association signals. We applied this approach to discovery (the International Schizophrenia Consortium (ISC) (n=6,909)) and validation (Genetic Association Information Network (GAIN) (n=2,729)) schizophrenia genome-wide association study (GWAS) datasets. We investigated each of the 212 experimentally validated pathways described in Kyoto Encyclopaedia of Genes and Genomes (KEGG) in the discovery sample. Nominally significant pathways were tested in validation sample, five pathways were significant (p=0.03-0.001); only the Cell Adhesion Molecules (CAM) pathway withstood conservative correction for multiple-testing. Interestingly, this pathway was also significantly associated with bipolar disorder (Wellcome Trust Case Control Consortium (WTCCC) (n=4,847)) (p=0.01). At a gene-level CAM genes associated in all three samples (*NRXN1* and *CNTNAP2*) have previously been implicated in specific language disorder, autism and schizophrenia. The Cell Adhesion Molecules (CAM) pathway functions in neuronal cell adhesion, which is critical for synaptic formation and normal cell signaling. Similar pathways have also emerged from a pathway analysis of autism, suggesting that mechanisms involved in neuronal cell adhesion may contribute broadly to neurodevelopmental psychiatric phenotypes.

Schizophrenia, Bipolar Disorder, pathways, GWAS, neuronal cell adhesion

Introduction

Despite substantial progress, much of the genetic variation in schizophrenia susceptibility remains to be attributed. [1-4] Our recent data suggests a spectrum of risk variation including rare variants and a substantial contribution from thousands of common susceptibility variants of small effect.[4] Additionally, many of these common variants may also contribute to bipolar disorder but not other common complex disease, which accords with epidemiological support for some genetic overlap between these disorders.[5] Identifying and confirming small genetic effects is likely to be challenging and a substantial barrier to translational research in psychoses.

Investigating genomic data at the level of molecular pathways rather than at individual single nucleotide polymorphisms (SNPs) is a potentially promising approach. Pathway-based analyses of genomic data may offer several advantages over traditional genetic association analyses. First, by increasing study power if, as has been suggested, the joint action of common variants within pathways play a major role in predisposing to complex genetic disorders.[6] Second, in being robust to the influence differences in linkage disequilbrium (LD) (e.g. between study populations or SNP arrays) may have on identification of associated variants. Third, by providing additional information relating to function over and above the statistical SNP GWAS data. Namely, implicating a molecular pathway in a disease process may be more biologically informative than interpreting involvement of an anonymous genetic marker.

A number of formal pathway-based analytical methods have been described[6-10, 42] and are reviewed elsewhere.[9,11] In this study we used the SNP-ratio test (SRT)[10] for the following reasons. The SRT is similar to approaches using methods based on gene set enrichment analysis (GSEA)[9] in that it tests for enrichment of statistically associated SNPs in a pathway using empirical p-values. As the SRT uses all SNPs in the pathway it can account for situations where LD, extending beyond the

gene of interest, generates false positives at a SNP-level. The test is robust to allelic or locus heterogeneity and can capture information from multiple signals at a given locus. Because the application of the test is at a pathway rather than gene level, this also precludes the need to adjust for gene size. However, all pathway-based methods are limited by the quality of the pathway annotation available and differences in how pathways are defined. The SRT can be applied flexibly to different pathway resources. For this study we selected only pathways identified in the Kyoto Encyclopaedia of Genes and Genomes (KEGG) database[12] as these have been experimentally validated. Inevitably, pathways will be enriched for well-studied genes, but in a hypothesis-free test this should not affect the type 1 error rate. For this reason we did not weight pathways based on prior evidence for involvement in the aetiology of schizophrenia or psychosis.

The purpose of this study was to apply a pathway-based approach to analysis of GWAS data from two large schizophrenia samples. The initial step was an exploratory analysis in the International Schizophrenia Consortium (ISC) dataset. We investigated 212 experimentally-validated pathways from the KEGG database for evidence that specific pathways were enriched for SNPs associated with schizophrenia ($p \leq 0.05$). Molecular risk pathways identified as meeting this significance threshold were then validated in a large independent dataset ($p \leq 0.05$). A secondary purpose of the study was to investigate the overlap between the validated pathways and bipolar disorder using the same statistical association criteria.

Materials and methods

Participants and genotyping

The International Schizophrenia Consortium was established to promote rapid progress towards the identification of genetic risk factors for schizophrenia. The sample included in the GWAS analysis included 3,322 patients with DSM-IV diagnosis of schizophrenia and 3,587 controls from the same European populations. Single nucleotide polymorphisms (SNPs) were assayed using the Affymetrix

Genome-Wide Human SNP 5.0 and 6.0 arrays.[3] The analysis provided data on 739,995 SNPs, with a genomic inflation factor of 1.09. Full details are provided elsewhere.[4]

The Genetic Association Information Network (GAIN) is a public-private partnership of the Foundation for the National Institutes of Health, Inc. (FNIH) with the goal of finding genetic causes for common diseases including schizophrenia. Details on inclusion criteria and participants are available elsewhere.[13,14] The schizophrenia study (ID phg000013) includes 1351 cases and 1378 controls, representing 2601 European ancestry subjects with GRU (General Research Use) consent and 219 with SARC (schizoaffective) consent. The final dataset consisted of 729,454 SNPs. Details of the participating UK cases and controls in the WTCCC bipolar disorder study (n=4,847) are provided in supplementary material from the original Wellcome Trust Case Control Consortium paper.[15] The final GWAS analysis included 457,796 SNPs (28,629 SNPs and 196 individuals were removed as flagged by WTCCC on quality control grounds).

Data analysis

Pathway analysis methodology

We identified all genic SNPs that mapped to the 212 identified KEGG pathways capturing 4,760 genes (Release 48.0, October 2008). First SNPs were linked to genes via the GenBank sequence database using Human Genome Organisation (HUGO) gene nomenclature, which assigns a unique symbol to every gene. This generated a list of 3,269,098 SNPs known to lie within genes with HUGO gene symbols. A SNP from the GWAS dataset was assigned to a gene only if its coordinates lie within the National Center for Biotechnology Information (NCBI) annotated start and stop coordinates of the gene. We also included the region 5kb upstream and 2kb downstream of each gene to account for variants in potential promoter regions.

SNP ratio test (SRT)

In the discovery dataset, association analyses were performed for each SNP in a given pathway using the Armitage trend test. The SNP ratio test (SRT) than describes the ratio of nominally significant ($p<0.05$) to non-significant SNPs for each pathway. An empirical p-value is generated for each pathway by comparing this ratio to ratios based on simulated datasets (**Figure 1**).

Figure 1 here

These datasets were simulated by randomising the case/control status of individuals in our original dataset 1000 times and then repeating the association analysis. For each simulation, we take the *M* most significant SNPs, where *M* corresponds to the number of SNPs below the cut-off used in the original dataset, e.g. $p<0.05$. The use of *M*, rather than reapplying the p-value threshold should prevent any inflation in empirically significant pathways due to an excess of false positive SNPs in the original GWAS (due to e.g. genotyping error, or other bias). Following this, we counted the number of times a simulated ratio was higher than the original ratio. We did this for each pathway, thus correcting for spurious enrichment of significant SNPs due to factors such as Linkage Disequilibrium (LD). Further details on the SRT methodology are provided in O'Dushlaine *et al* (2009).

Pathways that provided significant evidence ($p<0.05$) for enrichment with associated SNPs, were than tested using the same approach in the validation dataset. Finally, we tested the significant pathways identified in schizophrenia in the WTCCC bipolar disorder dataset. All file manipulations were carried out using PERL within a UNIX framework. Statistical analyses were conducted using STATA 10.

Results

Pathway analysis results

Of the 212 KEGG pathways examined in the ISC dataset, 47 had significant enrichment of associated SNPs when compared to simulated datasets (**Table 1**; full pathway details are provided in Supplementary Table 1). This represented more than we expected by chance (n~11). Five of these 47 pathways were associated in the GAIN schizophrenia sample: Cell adhesion molecules (CAM) (hsa04514),(p=0.001); Cell cycle (hsa04110), (p=0.015); Vesicular trafficking (SNARE)(hsa04130), (p=0.016); Tight junction (hsa04530), (p=0.03) and Glycan structures-biosynthesis 1 (hsa01030), (p=0.03) pathways. Only the CAM pathway exceeded a conservative correction, assuming independence, for the number of pathways taken to the replication phase. Interestingly, the CAM pathway was also significantly associated in the bipolar disorder dataset (p=0.026). However, pathways are not independent (e.g. CAM and Tight junction overlap at a molecular level and are involved in the same functions) and genes may be involved in multiple pathways. We would expect only two of the five schizophrenia pathways to be nominally significant in the GAIN dataset by chance with an α of 0.05 for the empirical p-value. An exploratory analysis of the other four pathways, indicated additional support for Glycan structures-biosynthesis 1 (p=0.0009) and Tight junction (p=0.015) pathways in bipolar disorder (**Table 1**). We did not see the same replication across the six non-psychiatric diseases in the Wellcome Trust Case Control consortium.[15]

Table 1 here

Implicated genes in the schizophrenia risk pathways

We next examined where the SNP-signals were coming from in the different datasets. Specifically, details in the overlap at SNP level between the schizophrenia datasets are provided in Supplementary Table 3. The most significant finding in the GAIN replication set was for the CAM pathway, where there was a substantial enrichment for association in genes involved in neuronal functioning (Supplementary Figure 2). Twenty-eight genes, of 110 contributing to the analysis, had

significantly associated SNPs in both datasets and 14 SNPs across 6 of these genes (*CDH4, GLG1, NRXN1, CNTN1, HLA-DQA2* and *PDCD1LG2*) share common risk alleles in both datasets.

For the Tight Junction pathway, 27 of 128 genes had significantly associated SNPs in both schizophrenia datasets (Supplementary Figure 3). Nine SNPs across 8 genes (*CDC42, CTNNA2, HCLS1, HRAS, PRKCH, MYH11, MYH15,* and *PARD3*) sharing common risk alleles (Supplementary Table 2).

In the Glycan structures-biosynthesis 1 pathway 30 of 111 represented genes were associated in both samples and 6 SNPs across 5 genes (*MAN2A1, GALNT2, GALNT13, OGT* and *XYLT1*) shared common risk alleles (Supplementary Figures 6A-C and Table 2). For the SNARE pathway, 6 of 36 genes have significantly associated SNPs in both schizophrenia datasets with two genes attributed to the same SNPs (Supplementary Figure 5). All 12 significantly associated SNPs in these two genes for both datasets (*STX18* and *TSNARE1*) share common risk alleles (see Supplementary Table 2). The finding for the Cell Cycle pathway is driven entirely by multiple SNPs in the same gene (*MAD1L1*) sharing the same risk alleles in both datasets (Supplementary Figure 4).

Gene overlap between schizophrenia and bipolar disorder

Where molecular pathways overlapped between schizophrenia and bipolar disorder we were interested in establishing if this represented the effect of the same key genes. We examined overlap at a gene-level between the ISC schizophrenia and WTCCC bipolar samples. For the CAM pathway 49 genes had at least one significant SNP in ISC and 22 of these had significant SNPs in WTCCC; for the Tight Junction pathway 47 genes had at least 1 significant SNP in ISC and 16 of these genes were replicated; and for the N-glycan biosynthesis pathway 50 genes had at least 1 significant SNP in ISC and 24 of these replicated. This suggests substantial overlap in key genes in both disorders. However, there was less concurrence at the level of identified risk SNPs and alleles between the

disorders. For example, in the CAM pathway nine genes had the same risk SNPs, including *CDH4* (Supplementary Table 2). For the Tight Junction pathway, seven genes had the same risk SNPs including *PRKCH, CTNNA2* and *MYH15*. In the Glycan structures-biosynthesis 1 pathway seven genes shared common risk SNPs at seven genes, including *MAN2A1*.

Discussion

We report a genomic pathway analysis of psychosis capturing 212 experimentally validated molecular pathways using discovery (ISC) and validation (GAIN) datasets. In the discovery dataset we identified evidence for involvement of a relatively large number of pathways (47 of 212 tested). Five of these pathways replicated in the replication sample, which was more than were expected by chance, one of these (CAM) survived conservative multiple testing correction. This sample may have been underpowered to detect more modest effects; a hypothesis that can be directly tested with the imminent availability of larger GWAS datasets, for example through the Psychiatric GWAS Consortium (PGC).[16]

Three of the five– CAM, Tight junction and SNARE- are involved in processes critical to neurodevelopment and synaptic function. Genes involved in all three have previously been implicated in schizophrenia susceptibility. There is significant overlap between the CAM and Tight junction pathways and both are relevant to synaptic formation and neurotransmission at glutamatergic and GABAergic synapses.[17] There is substantial evidence to support their involvement in cognitive and neuropsychiatric disorders. Genes involved in each of these pathways have been implicated in genetic disorders which impact on cognition (see the Genes to Cognition Database).[18] Interestingly, a recent autism study reported involvement of similar neuronal cell adhesion pathways.[19] A key issue is whether the same genes are involved across neuropsychiatric disorders. Taking the comparision of schizophrenia and bipolar disorder we identified substantial

overlap in nominally associated SNPs between the schizophrenia (ISC) and bipolar disorder (WTCCC) datasets. This suggests that certain genes may be critical in altering pathway function.

Several of the the genes identified as contibuting to risk pathways had previously been identified by genetic association or studies of copy number variation in schizophrenia or other neuropsychiatric disorder phenotypes (https://slep.unc.edu/evidence/). Of these a number had first been reported in the autism literature. We identified shared risk alleles in both schizophrenia samples at SNPs rs9309200 and rs1915220 at *NRXN1* (OMIM: 600565). Disruption of *NRXN1* has been reported as a risk factor for both schizophrenia [20-22] and autism. [23,24] Axonal neurexins form transynaptic complexes with neuroligins on dendrites and are required for the formation of synaptic contacts and for efficient neurotransmission including maintaining normal postsynaptic NMDA receptor function. Significant association involving the gene *CNTNAP2* (OMIM:604569) was detected in all three psychosis samples. *CNTNAP2* is a neurexin superfamily member, which is part of a neurogenetic pathway that is disturbed in different forms of language impairment [25] and autism.[26-28] Friedman and colleagues [29] have reported genomic deletions of varying sizes involving *CNTNAP2* in three non-related patients with schizophrenia and epilepsy but not in 512 healthy controls.

Neurexin directly interacts with the cytoskeleton membrane scaffolding protein CASK (OMIM: 300171), which contains associated SNPs in both schizophrenia datasets. Part of the Tight Junction gene pathway in KEGG, *CASK* may have a role in synaptic plasticity by coupling synaptic vesicle exocytosis to neuronal cell adhesion.[30,31] Tight junctions are composed of at least three types of transmembrane protein- occludin, claudin and junctional adhesion molecules (JAMs) and a cytoplasmic region containing three large protein complexes, the ZO protein complex, the CRB3-Pals1-PATJ protein complex and the PAR-3-aPKC-PAR6 complex. The transmembrane elements mediate cell adhesion and are thought to constitute the intracellular and paracellular diffusion barriers. The CRB3 and PAR transmembrane complexes are involved in the maintenance of cell

polarity and the ZO cytoplasmic complexes organize coupling with other cytoplasmic proteins and to actin microfilaments. We identified a significant enrichment of associated SNPs in genes involving the PAR complex (*CDC42* (OMIM: 116952); *PAR3* (OMIM:606745); *PAR6* (OMIM:607484) and *PRKCZ* (OMIM:176982)). The PAR complex establishes cell polarity for myelination[32] and for normal dendritic spine development.[33] Recent evidence suggests that *CDC42* may have a critical role in *DISC-1* related cell-migration through interaction with the gene *NUDEL*.[34] Reduced expression of *CDC42* has previously been demonstrated in the DLPFC of subjects with SCZ and this was correlated with decreased dendritic spine density in schizophrenia.[35] Incidentally, *GSK3B* (OMIM: 605004) which is also modulated by DISC-1, is part of the Cell Cycle Pathway, although not directly contributing to susceptibility in this study.[36]

The third pathway identified in all three samples, involved carbohydrate structures attached to glycolipids, glycoproteins and proteoglycans. These are known to play a critical role in many biological processes including cell adhesion during development, immune and inflammatory response, molecular trafficking, signal transduction and endocytosis.[37] In mouse models regulation of glycan binding is required for most of the physiological functions of NCAM including brain morphogenesis, axonal trafficking and higher cognitive functioning.[38,39] The diversity of glycan functions is matched by a potentially vast array of hundreds if not thousands of potential structures. None of the overlapping risk alleles identified by this study map to genes previously implicated in psychosis. In fact, where these genes have been linked to human disease previously it has been with connective tissue disorders (e.g. *MAN2A1* (OMIM:154582), *CHST3* (OMIM: 603799), *XYLT1* (OMIM: 608124)) or lipid metabolism (*GALNT2*).[40]

Pathway information should not be taken a face-value, it is important to examine the gene results. For example, inspection of the Cell Cycle Pathway result indicated that this association was driven by the MAD1L1 gene, implicating the gene but not the pathway. Potential confounds in considering

these data are that the methodology may be susceptible to bias in linkage disequilibrium (LD) in the case group, or that the data may reflect increased gene size or greater marker density in the pathways enriched for association. We correct for potential LD bias by randomizing case-control status but maintaining LD structure and repeating association tests to create a distribution of ratios (N=1000 simulations). This process retains the existing LD structure, giving a background estimate. Spurious associations arising from LD would lie within this background range. We found no evidence for increased gene size or greater SNP density for associated versus non-associated pathways (Mann-Whitney p=0.07, n=5 replicated pathways). However, we note that the significance of a pathway is somewhat reliant on the existence of a minimum number of SNPs (~20) which, by extension, relates to factors such as average gene length (Supplementary Figure 2). Specifically, if a pathway has a paucity of SNPs, the likelihood of significant SNPs from a given simulation mapping to this pathway is reduced. Thus, poorly "covered" pathways are difficult to test, a limitation for GWAS in general (Supplementary Figure 2).

By selecting only experimentally-validated pathways we restricted the number of genes included in this analysis. Incorporating other pathway tools to expand the number of genes captured in follow up these findings may be useful, although there is a lack of consensus across methods as to how pathways are defined and classified.[41] It is worth noting that the SRT does not correct for multiple testing at a pathway-level. This is non-trivial because molecular processes can involve more than one pathway, hence our strategy of replicating the exploratory findings. However, the multiplicity problem is greatly reduced by this approach relative to a SNP-level analysis.

This study implicates pathways involved in neuronal cell adhesion and synaptic function in molecular susceptibility to psychosis. Intriguingly, there is now significant convergence using different genetic analyses on key molecules in these pathways being broadly involved in neurodevelopmental disorders including psychosis, autism and language disorders. This may be of

significant diagnostic and therapeutic importance. Our study identified identified multiply associated SNPs at several genes, for example, MAD1L1, CDH4 and TSNARE1. Examining LD-relationships at these loci suggests that this may represent multiple signals. For example, the 26 associated SNPs at MAD1L1 can be captured by 6 tagging SNPs. More definitive interpretation of individual loci will require additional re-sequencing. For further investigation of these pathways in independent datasets, we would propose testing a model that expands promoter and other regulatory regions, includes investigation of other genetic models (e.g. epistasis), allows estimation of effect size/predictive value and investigates whether the pathways identified are associated with age at onset and course of disease. Twinned with ongoing targeted DNA sequencing studies this may be informative in quantifying and specifying the impact of discrete molecular pathways on different clinical outcomes. Further functional work is required in particular to investigate neuronal cell adhesion and membrane scaffolding given the growing convergence across studies of neurodevelopmental disorders on these mechanisms.

ACKNOWLEDGEMENTS

We would like to thank all the participating patients, institutions and medical staff without whose contribution this work would not have been possible. We acknowledge the support of our funders in particular Science Foundation Ireland and the Health Research Board. We appreciate useful comments made by anonymous reviewers.

REFERENCES

1. Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, Steinberg S, *et al*. Large recurrent microdeletions associated with schizophrenia. *Nature* 2008; **455**:232-236.

2. O'Donovan MC, Craddock N, Norton N, Williams H, Peirce T, Moskvina V, *et al*. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet* 2008; **40**: 1053-1055.
3. International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 2008; **455**:237-241.
4. International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia that overlaps with bipolar disorder. *Nature* 2009; **460**: 748-52. press).
5. Lichtenstein P, Yip BH, Bjork C, Pawitan Y, Cannon TD, Sullivan PF, Hultman CM. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: a population-based study. *Lancet* 2009; **373**:234-239.
6. Lesnick TG, Papapetropoulos S, Mash DC, Ffrench-Mullen J, Shehadeh L, de Andrade M, *et al*. A genomic pathway approach to a complex disease: axon guidance and Parkinson disease. *PLoS Genet* 2007; **3**:e98.
7. Askland K, Read C, Moore J. Pathways-based analyses of whole-genome association study data in bipolar disorder reveal genes mediating ion channel activity and synaptic neurotransmission. *Hum Genet* 2008; **125**:63-79.
8. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005; **102**:15545-15550.
9. Wang K, Li M, Bucan M. Pathway-Based Approaches for Analysis of Genomewide Association Studies. *Am J Hum Genet* 2007; **81**: 1278-1283.
10. O'Dushlaine C, Kenny E, Segurado R, Heron E, Morris D, Corvin A. The SNP ratio test: pathway mining of genome-wide association datasets. *Bioinformatics* 2009 (in press).
11. Hong MG, Pawitan Y, Magnusson PK, Prince JA. Strategies and issues in the detection of pathway enrichment in genome-wide association studies. *Hum Genet* 2009; **126**: 1348-54.

12. Kyoto Encyclopaedia of Genes and Genomes: (http://www.genome.ad.jp/kegg/pathway.html)

13. Genetic Association Information Network: (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000021.v2.p1)

14. Suarez BK, Duan J, Sanders AR, Hinrichs AL, Jin CH, Hou C, *et al*. Genomewide linkage scan of 509 European-ancestry and African American families with schizophrenia: suggestige evidence of linkage at 8p23.3-p21.2 and 11p.13.1-q14.1 in the combined sample. *Am J Hum Genet* 2006; **78**: 315-33.

15. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common disease and 3,000 shared controls. *Nature* 2007; **447**: 661-78.

16. Psychiatric GWAS Consortium Coordinating Committee, Cichon S, Craddock N, Daly M, Faraone SV, Gejman PV, *et al*.Genomewide association studies: history, rationale, and prospects for psychiatric disorders. *Am J Psychiatry* 2009; **166**: 540-56.

17. Kang Y, Zhang X, Dobie F, Wu H, Craig AM. Induction of GABAergic postsynaptic differentiation by alpha-neurexins. *J Biol Chem* 2008; **283**:2323-2334.

18. Croning MD, Marshall MC, McLaren P, Armstrong JD, Grant SG. G2Cdb: the Genes to Cognition database. *Nucleic Acids Res.* 2009; **37**:D846-851.

19. Wang K, Zhang H, Deqiong M, Bucan M, Glessner JT, Abrahams BS, *et al*. Common genetic variants on 5q14.1 associate with autistic spectrum disorders. *Nature* (in press).

20. Kirov G, Gumus D, Chen W, Norton N, Georgieva L, Sari M, O'Donovan MC, Erdogan F, Owen MJ, Ropers HH, Ullmann R.. Comparative genome hybridization suggests a role for NRXN1 and APBA2 in schizophrenia. *Hum Mol Genet* 2008; 17:458-465.

21. Rujescu D, Ingason A, Cichon S, Pietilainen OP, Barnes MR, Toulopoulou T, *et al*. Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum Mol Genet* 2009; **18**:988-996.

22. Walsh T, McClellan JM, McCarthy SE, Addington AM, Pierce SB, Cooper GM, *et al*. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008; **320**:539-543.

23. Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Skaug J, *et al***.** Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008; **82**:477-488.

24. Szatmari P, Paterson AD, Zwaigenbaum L, Roberts W, Brian J, Liu XQ, *et al*: Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* 2007; **39**:319-328.

25. Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, *et al*. A functional genetic link between distinct developmental language disorders. *N Engl J Med* 2008; **359**:2337-2345.

26. Alarcon M, Abrahams BS, Stone JL, Duvall JA, Perederiy JV, Bomar JM, *et al*. Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *Am J Hum Genet*. 2008; **82**:150-159.

27. Arking DE, Cutler DJ, Brune CW, Teslovich TM, West K, Ikeda M, *et al*. A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *Am J Hum Genet* 2008; **82**:160-164.

28. Bakkaloglu B, O'Roak BJ, Louvi A, Gupta AR, Abelson JF, Morgan TM, *et al*. Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders. *Am J Hum Genet*. 2008; **82**:165-173.

29. Friedman JI, Vrijenhoek T, Markx S, Janssen IM, van der Vliet WA, Faas BH, *et al*. CNTNAP2 gene dosage variation is associated with schizophrenia and epilepsy. *Mol Psychiatry* 2008; **13**:261-266.

30. Cohen AR, Woods DF, Marfatia SM, Walther Z, Chishti AH, Anderson JM. Human CASK/LIN-2 binds syndecan-2 and protein 4.1 and localizes to the basolateral membrane of epithelial cells. *J Cell Biol.* 1998; **142**:129-138.

31. Setou M, Nakagawa T, Seog DH, Hirokawa N. Kinesin superfamily motor protein KIF17 and mLin-10 in NMDA receptor-containing vesicle transport. *Science* 2000; **288**:1796-1802.

32. Chan JR, Jolicoeur C, Yamauchi J, Elliott J, Fawcett JP, Ng BK, Cayouette M. The polarity protein Par-3 directly interacts with p75NTR to regulate myelination. *Science.* 2006; **314**:832-836.

33. Zhang H, Macara IG. The polarity protein PAR-3 and TIAM1 cooperate in dendritic spine morphogenesis. *Nat Cell Biol* 2006; **8**:227-237.

34. Shen Y, Li N, Wu S, Zhou Y, Shan Y, Zhang Q, Ding C, Yuan Q, Zhao F, Zeng R, Zhu X.. Nudel binds Cdc42GAP to modulate Cdc42 activity at the leading edge of migrating cells. *Dev Cell* 2008; **14**:342-353.

35. Hill JJ, Hashimoto T, Lewis DA. Molecular mechanisms contributing to dendritic spine alterations in the prefrontal cortex of subjects with schizophrenia. *Mol Psychiatry* 2006; **11**:557-566.

36. Mao Y, Ge X, Frank CL, Madison JM, Koehler AN, Doud MK, Tassa C, *et al*. Disrupted in schizophrenia 1 regulates neuronal progenitor proliferation via modulation of GSK3beta/beta-catenin signaling. *Cell.* 2009;**136**(6):1017-31 **.**

37. Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. *Cell* 2006; **126**:855-867.

38. Weinhold B, Seidenfaden R, Röckle I, Mühlenhoff M, Schertzinger F, Conzelmann S, *et al*. Genetic ablation of polysialic acid causes severe neurodevelopmental defects rescued by deletion of the neural cell adhesion molecule. *J Biol Chem* 2005; **280**: 42971-7.

39. Angata K, Chan D, Thibault J, Fukuda M. Molecular dissection of the ST8Sia IV polysialyltransferase. Distinct domains are required for neural cell adhesion molecule recognition and polysialylation. *J Biol Chem*. 2004;**279**(24):25883-90.

40. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, *et al*. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* 2008; **40**:161-169.

41. Elbers C, van Eijk KR, Franke L, Mulder F, van der Schouw YT, Wijmenga C, Onland-Moret NC. Using genome-wide pathway analysis to unravel the aetiology of complex diseases. *Genet Epidemiology* 2009; **33**: 419-31.

42. Holmans P, Green EK, Phawa JS, Ferreira MA, Purcell SM, Sklar P, et al. Gene ontology analysis of GWAS study data sets provides insights into the biology of bipolar disorder. *Am J Hum Genet 2009*; **85**: 13-24.

Figure 1: **Sample SRT results. Example of (a) an SRT non-significant pathway and (b) an SRT significant pathway, using the ISC data. The original ratio of significant to non-significant SNPs is shown as a vertical red line. hsa01030 contains 246 significant and 3151 non-significant SNPs in the original GWAS, hsa04730 contains 196 significant and 3243 non-significant SNPS in the original GWAS dataset.**

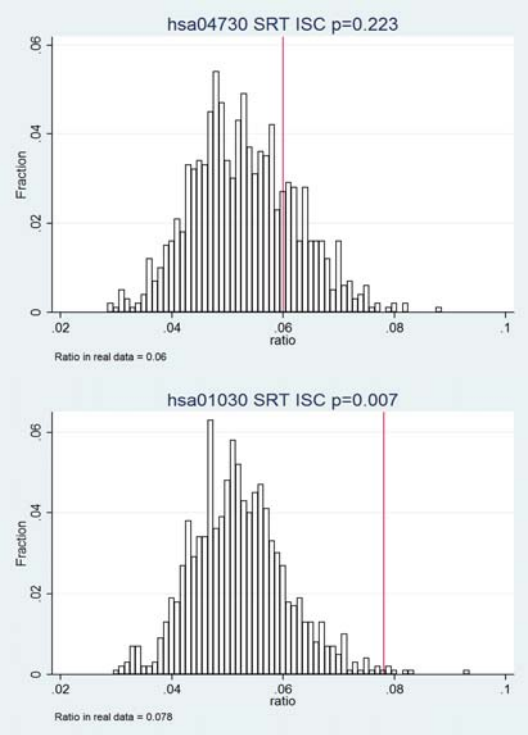

Figure 1: **Sample SRT results. Example of (a) an SRT non-significant pathway and (b) an SRT significant pathway, using the ISC data. The original ratio of significant to non-significant SNPs is shown as a vertical red line. hsa01030 contains 246 significant and 3151 non-significant SNPs in the original GWAS, hsa04730 contains 196 significant and 3243 non-significant SNPS in the original GWAS dataset.**

	ISC			GAIN			WTCCC			
Pathway ID	**s**	**ns**	**p**	**s**	**ns**	**p**	**s**	**ns**	**p**	**Pathway name**
hsa01030	246	3151	0.007	231	3111	0.033	231	1796	0.001	Glycan structures-biosynthesis 1
hsa04110	70	831	0.027	77	799	0.015	84	601	0.14	Cell cycle
hsa04130	47	417	0.017	51	409	0.016	26	297	0.46	SNARE
hsa04514	270	3613	0.022	319	3469	0.001	255	2235	0.03	Cell Adhesion Molecules (CAMs)
hsa04530	332	4375	0.003	310	4315	0.034	308	2633	0.02	Tight Junction

Table 1 SRT results for ISC, GAIN and WTCCC. “s”=number of significant SNPs in pathway, “n”=number of non-significant SNPs in pathway, “p” = empirical p-value from comparing ratio of significant to non-significant SNPs to ratios in simulated datasets.