

Schizophrenia Genetics: Building the Foundations of the Future

Katherine E. Tansey*, Michael J. Owen, and Michael C. O'Donovan

MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, UK

*To whom correspondence should be addressed; MRC Centre for Neuropsychiatric Genetics and Genomics, Institute of Psychological Medicine and Clinical Neurosciences, School of Medicine, Cardiff University, Cardiff, CF24 4HQ, UK; tel: +44-029-20-688-457, fax: +44-029-20-687-068, e-mail: tanseyk@cardiff.ac.uk

In recent years, our understanding of the genetic architecture of schizophrenia, a phrase which denotes the numbers of risk variants, their frequencies and effect sizes, has been transformed. This has come about through advances in technology that have allowed almost the entire human genome to be simultaneously interrogated for the presence of disease-associated genetic variation and allows this to be performed in sample sizes powered for a realistic possibility of success. Another development has been the emergence of international consortia willing to share raw data and their coalescence into super-consortia to achieve sample sizes and bodies of clinical and analytic expertise that was unimaginable a decade ago. These innovations have driven the emergence of statistically robust and replicable genetic findings in schizophrenia, and a rapid escalation in the number of those findings over the last 5 years.

The latest example comes from the Schizophrenia Working Group of the Psychiatric Genomics Consortium (PGC-SCZ) which, at the time of publication, included contributions from around 37 000 individuals with schizophrenia, 302 investigators, 35 countries, and 4 continents.¹ In their recent paper, published in *Nature* in July 2014, the PGC-SCZ group report 128 statistically independent genetic associations, implicating a minimum of 108 conservatively defined schizophrenia-associated genetic loci.¹

Of the identified loci, 83 have not been previously robustly supported as playing a role in schizophrenia, but it is also important to note the findings are consistent with previous literature; 25 loci that had previously been reported as associated with schizophrenia in large samples were again supported in this much larger analysis, confirming that the use of large samples and stringent statistical cut-offs results in reproducible findings. The availability of so many robustly supported findings offers immense opportunities for investigating and advancing our understanding of etiology.

Large Numbers of Alleles Across the Frequency Spectrum

While each individual variant themselves have a small effect size (less than 1.4), it has been estimated that common variation as a whole accounts for around a third to half of the genetic variance in schizophrenia^{2,3} though this may turn out to be an underestimate when heterogeneity and interaction effects are taken into account. Nevertheless, rare variants also play a role. This has been well established for around 20 years with respect to small deletions and duplications known as copy number variants (CNVs)^{4,5} but systematic surveys for other forms of rare genetic variation have had to await the development of high capacity next generation sequencing technology. These studies, particularly the larger ones,^{6,7} have been commented on in a recent review in the journal⁸ and will not be considered in detail here. However, like the early genome-wide association studies (GWASs),^{2,9} rather than striking evidence for individual susceptibility genes, the recent sequencing studies provided evidence that many loci contribute to risk and for enrichment of rare mutations in certain gene sets. This strongly suggests that, as we have seen for GWAS, better powered studies will implicate specific genes.^{10,11}

The evidence that risk variants for schizophrenia occur across the range of allele frequencies is compelling. As noted above, the contribution to risk of schizophrenia arising from the aggregate effect of common variants is not trivial, but even a single common variant can have a similar effect on the variance in the population as rare variants with larger effects simply because they occur much more frequently in the population.¹² This suggests that in order to understand genetic risk mechanisms we will need to explore all parts of the allele frequency spectrum. Moreover, it seems likely that advantages may accrue from a combined approach to gene identification as the PGC-SCZ study pointed to an overlap between genes in schizophrenia GWAS regions and those with de

novo mutations. This suggests that a combination of the superior power of GWAS (via sample size) and the superior precision of sequencing with respect to gene identity might on the one hand be usefully harnessed to identify likely causal genes within GWAS-associated loci and on the other to enhance the power of sequencing by providing targets with enhanced prior probability.

Genetic Associations to Biological Mechanisms

The identity of individual associations, and patterns of enrichment of various types of associations and disease-linked mutations are beginning to shed light on areas of biology that are likely relevant to schizophrenia, although we stress detailed mechanistic conclusions will require other types of research, and that the associations at each locus have not yet been firmly linked to specific genes. Nevertheless, it is notable that within the schizophrenia-associated loci are multiple genes involved in synaptic function and plasticity, particularly genes involved in glutamatergic neurotransmission (*GRM3*, *GRIN2A*, *GRI1A1*, and *SLC38A7*) and neuronal calcium signaling (*CACNA1C*, *CACNB2*, *CAMKK2*, *CACNA1I*, *NRGN*, and *RIMS1*). *DRD2*, which a priori is possibly the strongest of all conceivable candidate genes for schizophrenia based on function, is also associated with the disorder.

Given that most of the common variant associations do not appear to result from DNA changes that affect protein sequence, it is presumed, with some evidence,¹³ that they exert their effects through influencing gene expression. Investigating this hypothesis further, the PGC-SCZ group sought to determine if schizophrenia-associated common variants are concentrated in regulatory elements marked as activating gene expression in particular tissues or cell lines. Importantly, though largely as predicted, associations were enriched in these regulatory elements in various brain tissues and in genes showing high expression in neurons/interneurons. A much more novel and potentially important finding was they were also enriched in these regulatory elements in the immune system, particularly B-lymphocyte cell lineages. This finding is intriguing as it provides some genetic, and therefore etiological, support for the general hypothesis that immune dysregulation plays a role in the development of schizophrenia.¹⁴ However, we need to move beyond general enrichment analyses to identify specific causal variants in specific regulatory elements, understand which genes/proteins are affected by those regulatory elements, and show how genetic variation directly affects immune but not neuronal function, before the immune hypothesis of schizophrenia can be considered to be genetically confirmed.

The limited network analyses conducted by the PGC-SCZ did not identify any generically annotated biological pathways that were enriched for associations. This may reflect the restricted analyses presented, and a more

thorough evaluation of the data is underway. But it may also reflect either high polygenicity and/or limitations in the quality and availability of data upon which to base these bioinformatic-driven analyses. Not only are the functions of many proteins unknown, but even less well documented are the elements that regulate the expression and processing of protein isoforms in specific cellular, developmental, and physiological contexts especially in the brain. Proteomic studies in particular lack the comprehensive scale of transcriptomic studies and are currently limited to targeted approaches. Interestingly, in PGC-SCZ, support was found for one of the most consistently implicated gene sets in schizophrenia, a set comprised of brain-expressed genes that interact with the Fragile X mental retardation (FMRP) protein.^{3,6,7,15} While the biological implications of this are not yet understood, it is likely that improving knowledge of the protein interactome and of the constituent members of sets of proteins involved in brain function will improve our ability to move from patterns of association to pathogenesis. This is likely to require more extensive experimental validation, and iterative refinement, of bioinformatics tools. For example, experimental studies reveal little overlap between genes predicted to be targets of microRNA-137, whose encoding gene lies within a schizophrenia-associated locus, and those whose expression is affected by knockdown or overexpression of the microRNA¹⁶ demonstrating that experimental validation of bioinformatics predictions is essential.

Targets of FMRP are not the only consistently implicated gene set in schizophrenia. Both GWAS^{1,3} and exome sequencing⁷ points to the involvement of multiple calcium channels and multiple genes involved in calcium signaling, a process also implicated in bipolar disorder.^{17,18} Proteins affiliated with the *N*-methyl-D-aspartate receptor and the activity-regulated cytoskeleton-associated protein have been implicated by CNVs analysis¹⁹ and by exome sequencing,^{6,7,20} while as noted above, *GRIN2A* and multiple proteins related to glutamatergic signaling are associated with the disorder in the recent GWAS.¹ These gene sets plausibly converge at the functional level of synaptic plasticity and remodeling,^{6,21} although this hypothesis requires testing through mechanistic experimental studies.

The complexity and inaccessibility of human brain tissue has made it challenging to understand basic disease mechanisms and to translate genetic findings into biology. Discussed more fully in a recent review,⁸ a synthesis of gene discovery with recent advancements in stem cell technology and genome engineering mean that an exciting avenue of research has been opened for psychiatry. This could be either through patient-derived induced pluripotent stem cells²² or through manipulations of cell lines by emerging gene engineering technologies.^{23,24} Further discussion of these developing areas is outside the scope of the current article but have been discussed elsewhere.^{25–33}

Pleiotropy

Pleiotropy occurs when one gene or genetic variant contributes to multiple phenotypes, a phenomenon fast becoming a characteristic of identified genetic risk factors for neuropsychiatric disorders. There is already evidence for extensive sharing of common genetic risk variants between schizophrenia, bipolar disorder major depressive disorder, and attention deficit hyperactivity disorder (ADHD)^{17,34,35} though the evidence for ADHD is somewhat less consistent than that for the other phenotypes. The high genetic correlation is not explained by diagnostic misclassification,³⁶ and instead points to considerable genetic pleiotropy in terms of these categorically defined diagnoses. The same is true for rare genetic risk factors for schizophrenia. It has been clear for several years that the same CNVs that confer risk for schizophrenia also do so for neurodevelopmental disorders including autism spectrum disorders (ASD), intellectual disability, ADHD, and epilepsy.^{4,5,37,38} Given most of these CNVs affect multiple genes, it could conceivably be argued that different genes were involved in each phenotype, but a recent analysis of point mutations supports the view that sharing of risk occurs at the level of genes and types of mutation.⁶ Moreover, the PGC-SCZ study found overlap between the genes in schizophrenia GWAS regions and those with de novo mutations in intellectual disability and ASD providing further support for overlapping genetic risk, and presumably pathophysiology.

The extent of pleiotropy may be surprising, but it is consistent with other disciplines which generally show high rates of comorbidity and lack of specificity in disease associations. Pleiotropy can also facilitate understanding of disease mechanisms by identifying novel intermediate phenotypes on the causal chain. For example, many risk factors for type 2 diabetes are pleiotropic for body mass and are likely to mediate their effects on the former through the latter. Of course in type 2 diabetes, there was a strong hypothesized link with body mass, and exploiting pleiotropy is not going to be quite so simple in psychiatry. Rather, finding pleiotropic links may require deep, and largely speculative, mass phenotyping. Deep phenotyping refers to the collection of large amounts of information on individuals beyond categorical diagnostic status, eg, cognitive and neuroimaging data, or given recent findings (see above) immune function. It is early days with respect to this type of work, but promising results have already been demonstrated for the first GWAS-identified schizophrenia risk gene, *ZNF804A*,⁹ which have been associated with cognition, clinical subdimensions, and brain phenotypes.³⁹⁻⁴² Small samples limit the robustness of the conclusions that have emerged so far from this sort of research.

Exploiting pleiotropy is likely to require very large cohorts, although much smaller samples ascertained for the same relatively high penetrance mutation, eg, a CNV,

are also likely to be highly informative. The considerable genetic overlap between disorders also suggests it will be important that much of this work should be undertaken across current diagnostic boundaries (and also include unaffected individuals given we all carry large numbers of common risk alleles) in order to characterize the impact of genetic variants.

From Genetics to Treatment

Antipsychotics are only partly effective for the positive symptoms of schizophrenia and do little to alleviate the negative symptoms or cognitive deficits. No novel class of drug for schizophrenia has emerged since the 1960s,⁴³ presumably due to the limited understanding and insight into the molecular underpinnings of schizophrenia. It is therefore some cause for optimism that the recent PGC-SCZ publication implicates genetic variants at the locus that contains *DRD2*, the gene that encodes the D2 dopamine receptor, the target of all known effective antipsychotics. The association of genetic variants at this gene can be seen as a “proof of principle” finding in reverse, supporting the hypothesis that pharmacological manipulation of proteins highlighted by common variants can have substantial therapeutic effects regardless of the fact that the genetic effect size on risk is meagre.⁴⁴ It also suggests that other loci implicated by GWAS harbor the potential to guide effective drug discovery for schizophrenia, although we need better knowledge of the actual risk variants, and the proximal and distal functional consequences of those variants in the pathway to disease.

Furthermore, each gene implicated in the development of schizophrenia does not work independently. If the gene itself is an unsuitable drug target, there is the potential for manipulation via interacting proteins or other members participating in the same biological pathway. Therefore, it is important to fully characterize these risk genes and the cellular process, pathways, and phenotypes which they regulate. Translating just one associated locus into an effective treatment for schizophrenia would amply justify the contribution made by tens of thousands of patients with the disorder that have made the recent genetic advances possible, and the investments of effort and resource by researchers and funding bodies, governmental and charitable.

Conclusions

The recent PGC-SCZ publication is a landmark in the process of identifying genetic risk variants for schizophrenia, but it is not one that indicates the end of the journey. As sample sizes and power increases, and with the additional detail provided by sequencing, discovery is likely to accelerate. However, even now, the genetic findings provide the basis for a wealth of research, from molecular and cellular investigations through to defining novel clinical

classifications. It is too early to know the implications for new treatments, although some of the associations may immediately inspire potential therapeutic targets. What would seem inevitable is that if as yet undiscovered treatments for the disorder are possible, and it seems unlikely that the only effective treatment has already been discovered, the opportunities for better understanding of pathogenesis that flow from the genetic data must surely accelerate their discovery. For this to happen, scientists of many disciplines must move away from the comfort of some of the “old favourites,” for which the evidence is much less secure and devote their energies and intellect in pursuit of the new findings that are well grounded in evidence. Looking beyond schizophrenia, the findings definitively demonstrate the power of genetics can be harnessed for psychiatric phenotypes despite their presumed high heterogeneity, the absence of tests to validate diagnosis, and uncertain biological validity.

Funding

The work at Cardiff University was funded by Medical Research Council (MRC) Centre (G0800509) and Program Grants (G0801418); European Community's Seventh Framework Programme (HEALTH-F2-2010-241909, Project EU-GEI).

Acknowledgments

The authors have declared that there are no conflicts of interest in relation to the subject of this study.

References

- Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511:421–427.
- Purcell SM, Wray NR, Stone JL, et al.; International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;460:748–752.
- Ripke S, O'Dushlaine C, Chambert K, et al.; Multicenter Genetic Studies of Schizophrenia Consortium; Psychosis Endophenotypes International Consortium; Wellcome Trust Case Control Consortium 2. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet*. 2013;45:1150–1159.
- Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell*. 2012;148:1223–1241.
- Kirov G, Rees E, Walters JT, et al. The penetrance of copy number variations for schizophrenia and developmental delay. *Biol Psychiatry*. 2014;75:378–385.
- Fromer M, Pocklington AJ, Kavanagh DH, et al. De novo mutations in schizophrenia implicate synaptic networks. *Nature*. 2014;506:179–184.
- Purcell SM, Moran JL, Fromer M, et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*. 2014;506:185–190.
- McCarthy SE, McCombie WR, Corvin A. Unlocking the treasure trove: from genes to schizophrenia biology. *Schizophr Bull*. 2014;40:492–496.
- O'Donovan MC, Craddock N, Norton N, et al.; Molecular Genetics of Schizophrenia Collaboration. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet*. 2008;40:1053–1055.
- Sergi MJ, Rassovsky Y, Widmark C, et al. Social cognition in schizophrenia: relationships with neurocognition and negative symptoms. *Schizophr Res*. 2007;90:316–324.
- Green MF. Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. *J Clinical Psychiatry*. 2006;67(suppl 9):3–8; discussion 36–42.
- Visscher PM, Goddard ME, Derks EM, Wray NR. Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses. *Mol Psychiatry*. 2012;17:474–485.
- Richards AL, Jones L, Moskvina V, et al.; Molecular Genetics of Schizophrenia Collaboration (MGS); International Schizophrenia Consortium (ISC). Schizophrenia susceptibility alleles are enriched for alleles that affect gene expression in adult human brain. *Mol Psychiatry*. 2012;17:193–201.
- Benros ME, Mortensen PB, Eaton WW. Autoimmune diseases and infections as risk factors for schizophrenia. *Ann N Y Acad Sci*. 2012;1262:56–66.
- Szatkiewicz JP, O'Dushlaine C, Chen G, et al. Copy number variation in schizophrenia in Sweden. *Mol Psychiatry*. 2014;19:762–773.
- Hill MJ, Donocik JG, Nuamah RA, Mein CA, Sainz-Fuertes R, Bray NJ. Transcriptional consequences of schizophrenia candidate miR-137 manipulation in human neural progenitor cells. *Schizophr Res*. 2014;153:225–230.
- Cross-Disorder Group of the Psychiatric Genomics C, Genetic Risk Outcome of Psychosis C. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet*. 2013;381:1371–1379.
- Psychiatric GCBWDWG. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet*. 2011;43:977–983.
- Kirov G, Pocklington AJ, Holmans P, et al. De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry*. 2012;17:142–153.
- Timms AE, Dorschner MO, Wechsler J, et al. Support for the N-methyl-D-aspartate receptor hypofunction hypothesis of schizophrenia from exome sequencing in multiplex families. *JAMA Psychiatry*. 2013;70:582–590.
- Shepherd JD, Bear MF. New views of Arc, a master regulator of synaptic plasticity. *Nat Neurosci*. 2011;14:279–284.
- Brennand KJ, Simone A, Jou J, et al. Modelling schizophrenia using human induced pluripotent stem cells. *Nature*. 2011;473:221–225.
- Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339:819–823.
- Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science*. 2013;339:823–826.
- Freedman ML, Monteiro AN, Gayther SA, et al. Principles for the post-GWAS functional characterization of cancer risk loci. *Nat Genet*. 2011;43:513–518.
- Edwards SL, Beesley J, French JD, Dunning AM. Beyond GWAS: illuminating the dark road from association to function. *Am J Hum Genet*. 2013;93:779–797.

27. Wright R, Réthelyi JM, Gage FH. Enhancing induced pluripotent stem cell models of schizophrenia. *JAMA Psychiatry*. 2014;71:334–335.
28. Brennand KJ, Simone A, Tran N, Gage FH. Modeling psychiatric disorders at the cellular and network levels. *Mol Psychiatry*. 2012;17:1239–1253.
29. Sternecker JL, Reinhardt P, Schöler HR. Investigating human disease using stem cell models. *Nat Rev Genet*. 2014;15:625–639.
30. Zhang F, Wen Y, Guo X. CRISPR/Cas9 for genome editing: progress, implications and challenges. *Hum Mol Genet*. 2014;23:R40–R46.
31. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*. 2014;157:1262–1278.
32. Shalem O, Sanjana NE, Hartenian E, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*. 2014;343:84–87.
33. Hsu PD, Zhang F. Dissecting neural function using targeted genome engineering technologies. *ACS Chem Neurosci*. 2012;3:603–610.
34. Cross-Disorder Group of the Psychiatric Genomics C, Lee SH, Ripke S, et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet*. 2013;45:984–994.
35. Hamshere ML, Stergiakouli E, Langley K, et al. Shared polygenic contribution between childhood attention-deficit hyperactivity disorder and adult schizophrenia. *Br J Psychiatry*. 2013;203:107–111.
36. Wray NR, Lee SH, Kendler KS. Impact of diagnostic misclassification on estimation of genetic correlations using genome-wide genotypes. *Eur J Hum Genet*. 2012;20:668–674.
37. Williams NM, Zaharieva I, Martin A, et al. Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a genome-wide analysis. *Lancet*. 2010;376:1401–1408.
38. Hogarty GE, Greenwald DP, Eack SM. Durability and mechanism of effects of cognitive enhancement therapy. *Psychiatr Serv*. 2006;57:1751–1757.
39. Esslinger C, Walter H, Kirsch P, et al. Neural mechanisms of a genome-wide supported psychosis variant. *Science*. 2009;324:605.
40. Walters JT, Corvin A, Owen MJ, et al. Psychosis susceptibility gene ZNF804A and cognitive performance in schizophrenia. *Arch Gen Psychiatry*. 2010;67:692–700.
41. Linden DE, Lancaster TM, Wolf C, et al. ZNF804A genotype modulates neural activity during working memory for faces. *Neuropsychobiology*. 2013;67:84–92.
42. Cummings E, Donohoe G, McDonald C, et al. Clinical symptomatology and the psychosis risk gene ZNF804A. *Schizophr Res*. 2010;122:273–275.
43. Hyman SE. Time for new schizophrenia Rx. *Science*. 2014;343:1177.
44. Collins FS. Reengineering translational science: the time is right. *Science Trans Med*. 2011;3:90cm17.