Genovese *et al.*[13] also provide several noteworthy insights into schizophrenia biology. Notably, these did not stem from the identification of a specific gene but rather from efforts to refine the global excess of ~1,000 dURVs in cases to specific tissues and cell types. They began by showing that dURVs in cases exhibited strong enrichment in brain-specific genes but no evidence for enrichment in 26 other human tissues. They then assessed enrichment of case dURVs in genes specific to different CNS cell types, showing that there was strong enrichment in neurons but none in astrocytes or oligodendrocytes. Finally, they found significant enrichment in sets of genes plausibly expressed at the synapse, including those whose transcripts are bound by FMRP, CELF4 and RBFOX2 (all of which either transport or regulate synaptic RNAs), but no enrichment at other neuronally expressed genes. These findings provide genetic evidence for the idea that synaptic dysfunction in general, and not particular synaptic components (for example, proteins interacting with NMDA-type glutamate receptors or the activity-regulated cytoskeleton-associated protein ARC), is an important pathogenic mechanism in schizophrenia. Deeper insights might be possible if we had a better understanding of gene expression in the human brain. Existing data sets are small, particularly in the context of age-related expression variation across the lifespan, and we lack cell-type-specific data for most CNS cell types. Filling these gaps will be an enormous challenge for the research community, but working toward a more complete enumeration of spatiotemporal variation in gene expression in the human brain, including that of genes whose RNA and protein products localize to the synapse, will be key to future progress in translating genetic findings into understanding of disease mechanisms in schizophrenia and other brain disorders.

COMPETING FINANCIAL INTERESTS
The author declares no competing financial interests.

1. Schizophrenia Working Group of the Psychiatric Genomics Consortium. *Nature* **511**, 421–427 (2014).
2. Fromer, M. *et al. Nature* **506**, 179–184 (2014).
3. Levinson, D.F. *et al. Am. J. Psychiatry* **168**, 302–316 (2011).
4. O'Donovan, M.C. *et al. Nat. Genet.* **40**, 1053–1055 (2008).
5. Purcell, S.M. *et al. Nature* **506**, 185–190 (2014).
6. Purcell, S.M. *et al. Nature* **460**, 748–752 (2009).
7. Ripke, S. *et al. Nat. Genet.* **45**, 1150–1159 (2013).
8. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium *Nat. Genet.* **43**, 969–976 (2011).
9. Singh, T. *et al. Nat. Neurosci.* **19**, 571–577 (2016).
10. Stefansson, H. *et al. Nature* **460**, 744–747 (2009).
11. Walsh, T. *et al. Science* **320**, 539–543 (2008).
12. Xu, B. *et al. Nat. Genet.* **43**, 864–868 (2011).
13. Genovese, G. *et al. Nat. Neurosci.* **19**, 1433–1441 (2016).
14. Power, R.A. *et al. JAMA Psychiatry* **70**, 22–30 (2013).
15. Sullivan, P.F., Kendler, K.S. & Neale, M.C. *Arch. Gen. Psychiatry* **60**, 1187–1192 (2003).

Brains, genes and power

Frank W Albert

Gene expression data from more than 500 human brains shed light on the molecular consequences of genetic variation that contributes to schizophrenia.

Schizophrenia is a neuropsychiatric disorder that afflicts up to 1% of humans and affects diverse brain functions, including perception, emotion and cognition. Despite extensive characterization of the electrophysiological, cell morphological and molecular alterations accompanying the disease, effective treatment remains elusive. This is, in part, because our understanding of the causes of schizophrenia are incomplete.

Schizophrenia is highly heritable: about 80% of the variation in risk among people is due to inherited genetic variation[1]. Thus, large-scale genetic studies offer tremendous promise for the identification of genetic factors that contribute to disease. A study[2] in this issue of *Nature Neuroscience* by the CommonMind Consortium illuminates the molecular consequences of these genetic leads, extending our understanding of how genetic variation contributes to disease pathology.

Tracking down the individual genomic loci that underlie the genetic contribution to schizophrenia has proven challenging. The most recent genome-wide association study (GWAS) required 36,000 schizophrenia case subjects and 110,000 control subjects to identify 108 loci associated with schizophrenia risk[3]. Together, these loci explain only about 4% of the total genetic contribution to schizophrenia liability. The remainder presumably resides in hundreds or thousands of additional loci, each with even smaller effects[4]. Their identification will require even larger study populations. But until then, the known schizophrenia loci provide an excellent opportunity to better understand the disease.

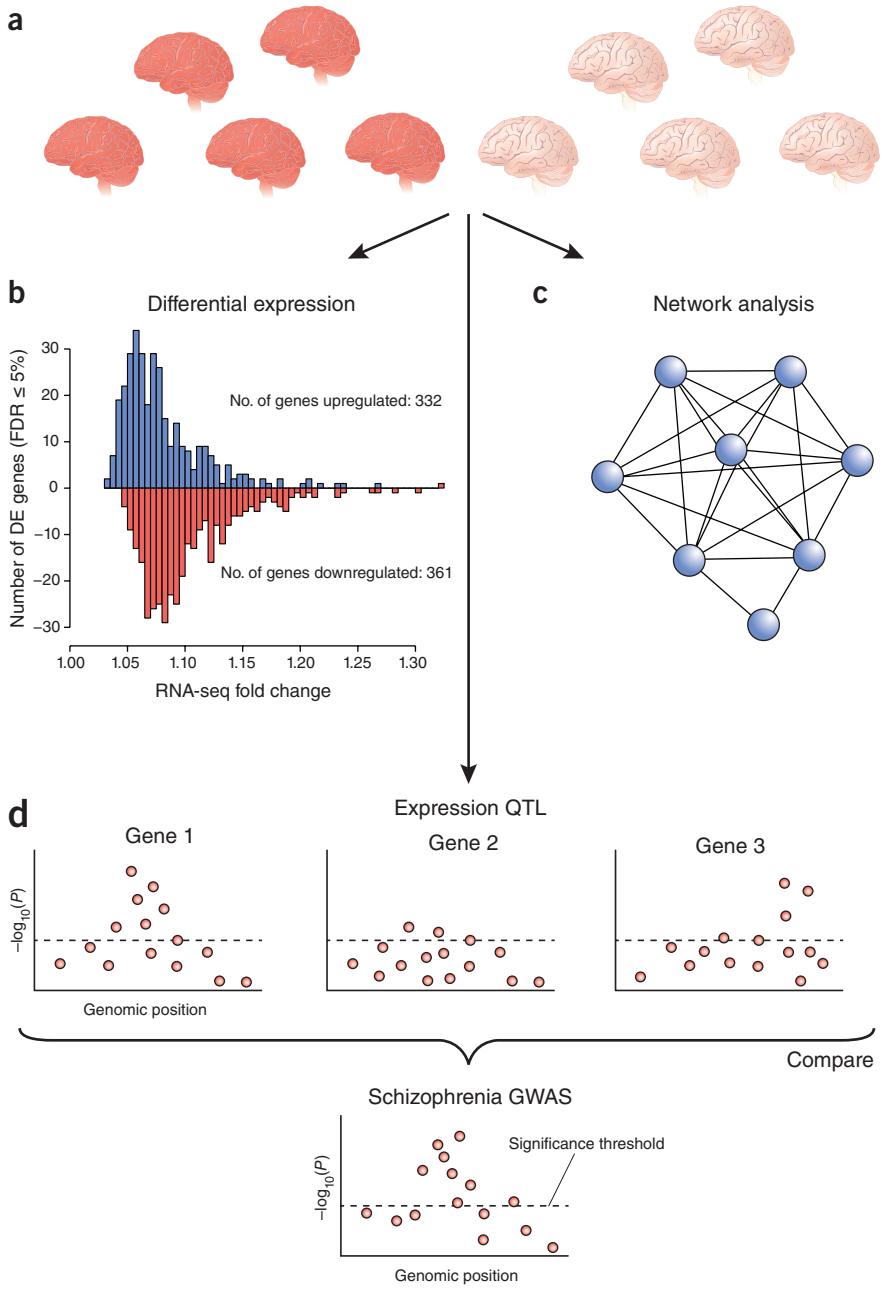
Each GWAS locus points us to a region in the genome that often contains multiple genes and dozens to hundreds of DNA variants. Identifying the responsible genes and variants is one of the most fundamental challenges in modern genetics. Unfortunately, this problem is extremely difficult. Because the variants within GWAS regions are located close to one another, they are rarely separated by recombination. This makes it hard for any one variant to stand out from its neighbors, because their statistical associations spread along the genomic region.

To learn which genes in the schizophrenia GWAS regions might be good candidates, Fromer, Roussos, Sieberts *et al.*[2] examined how the GWAS loci perturb gene expression patterns in the brain (**Fig. 1**). To do this, they measured mRNA levels in over 500 human brains. The samples were also typed for genetic markers along the genome and the expression of each gene was tested for association with these genotypes. Significant associations are called 'expression quantitative trait loci', or eQTL (ref. 5). When a GWAS hit is also an eQTL for a certain gene, this suggests that the altered expression of this gene contributes to the trait.

This approach was first applied on a genome-wide scale in yeast[6] and has since gained popularity in human genetics. It is particularly promising because most genetic risk for many common diseases resides far from protein-coding exons and instead is located in gene regulatory elements[7].

There are two complications with this approach. It is now clear that gene expression itself is genetically complex[8]. A typical gene is influenced by variation all across the genome, both close to the gene (for example, in the promoter) and far from it (for example, in transcription factors that can be located on

Frank W. Albert is in the Department of Genetics, Cell Biology & Development, University of Minnesota, Minneapolis, Minnesota, USA.
email: falbert@umn.edu



Marina Corral Spence/Nature Publishing Group

Figure 1 Brain gene expression in schizophrenia. (**a**) RNA sequencing data from more than 500 brains is used for three sets of analyses. Red or pink shading in brains denotes samples from people with schizophrenia vs. healthy control subjects. (**b**) Differential expression (DE) between brains from patients with schizophrenia and controls pinpoints genes that are misregulated in disease. (**c**) Network analyses reveal molecular pathways with altered disease connectivity. (**d**) Genetic analysis of gene expression levels identifies regulatory genetic variation. These expression associations are compared to genetic associations with schizophrenia. In the examples, expression levels of gene 1 are influenced by genetic variation and are the most similar to the schizophrenia associations, making it a candidate for contributing to the disease.

different chromosomes). The sample sizes that have long been typical for eQTL studies have insufficient statistical power to identify most of these regulatory relationships. As a consequence, many eQTL that are potentially causal for disease are unknown.

At the same time, the available data sets have in fact been large enough to demonstrate that the human genome carries enormous amounts of regulatory variation. Nearly 90% of genes show evidence for association in one tissue or another. Conversely, more than 90% of common human polymorphisms are linked to the expression of at least one gene[9]. This means that colocalization of GWAS hits with eQTL might simply be due to chance. In a way, we have identified both too little and too much human regulatory variation.

Fromer, Roussos, Sieberts *et al.*[2] tackled these complications by making sure that their data set had high statistical power, stemming from a large sample of brains: 258 donated by patients with schizophrenia and 279 from controls without neurological disease. Acquiring this many brains and processing them in a coordinated fashion is an impressive feat. The resulting data set revealed much new regulatory variation: 51% of genes expressed in the brain had an eQTL, drastically extending the set of known brain eQTL.

Notably, 73 of the 108 schizophrenia GWAS hits contained at least one eQTL. How can we be more confident that these eQTL genes are in fact functional intermediates for schizophrenia, rather than overlapping by chance? The authors used an approach[10] that compares the pattern of genetic associations between schizophrenia and gene expression along the genome (**Fig. 1**). This method is one example of a rapidly growing toolkit for the quantitative integration of genetic associations for disease with those for gene expression[11,12].

The model highlighted at least one putatively causal gene in 19 of the GWAS regions. In eight regions, only a single gene remained. The authors experimentally manipulated the expression level of five of these genes in zebrafish embryos and found that three of them indeed had effects on brain development, bolstering a potential role in schizophrenia. To gauge the specificity of this result, it would be interesting to know what fraction of all genes can affect brain development when over- or underexpressed in this manner.

In a second major set of analyses, the authors compared gene expression levels between schizophrenia cases and controls. They found 693 significant expression differences and suggestive evidence that thousands more genes may be misregulated in the brains of patients with schizophrenia. Interestingly, all of these differences had small magnitudes, often considerably smaller than those seen in earlier studies with many fewer samples. In some cases, even the direction of significant expression differences was opposite that of previous studies. How can this be?

The answer most likely lies in the low precision of the effect estimates in the earlier, smaller studies. Gene expression measures are subject to stochastic biological variation and measurement error. When many statistical tests are performed, such as in a genome-wide differential expression comparison, stochastic variation will randomly underestimate some effects and inflate others[13]. To reach statistical significance, an effect must exceed a certain

threshold. Therefore, those effects that happen to be overestimated are more likely to make it into the list of significant results. This effect, often called the winner's curse (see ref. 13 for a review), is stronger in small studies and makes them more prone to 'find' results where there is no difference in reality. The winner's curse can even create apparent effects with opposite sign. It is particularly severe when the true effects are small, which is the case for schizophrenia.

How severe is the winner's curse in schizophrenia? The authors provide a sobering estimate, worked out in detail in the supplementary information. If the ultimate source of differential expression between schizophrenia cases and controls are the eQTL identified in this study, the expected expression difference is a function of the effect size of the eQTL (i.e., how much the two alleles differ in gene expression) and of the difference in allele frequency of the eQTL variant between cases and controls. These quantities are both small. The authors estimate that the true differential expression might be so small that more than 10,000 samples would be required for their reliable detection. Even the current study is underpowered to detect differential expression of this magnitude, and a complete picture of misregulation of expression in schizophrenia remains elusive.

What, then, are we to make of gene expression differences and genetic effects as small as those found in schizophrenia? Without a doubt, it would be convenient if larger effects existed. But small effects are the hand that evolution has dealt us. In genes associated with schizophrenia, any mutations of large effect are likely to be quickly removed by selection. Some of the remaining variants contribute to disease, but not strongly enough to be effectively purged by selection. These small but common effects may subtly influence brain function in the general population[14] and may result in schizophrenia when too many of them happen to come together in an individual.

The functional consequences of small gene-expression changes, such as those caused by genetic variation, are a fundamentally open question in biology. Typical experimental assays of gene function involve strong under- or overexpression that often exceeds variation observed *in vivo*. Exciting pioneering work in yeast suggests that the relationship between finely controlled gene expression and its functional consequences varies dramatically between genes[15]. More quantitative and systematic maps relating expression change to biological consequences will be important for the interpretation of genetic variation.

By expanding our view of the molecular consequences of some of the genetic loci that contribute to schizophrenia, the CommonMind Consortium has taken a big step toward a biological understanding of this disease[2]. This may eventually lead to better, more targeted treatments. Future computational mining of the data by the research community will undoubtedly create further insights. Beyond disease, knowing the genes that can contribute to schizophrenia and other neurological disorders might teach us more about the biology and evolution of uniquely human brain function.

COMPETING FINANCIAL INTERESTS
The author declares no competing financial interests.

1. Sullivan, P.F., Kendler, K.S. & Neale, M.C. *Arch. Gen. Psychiatry* **60**, 1187–1192 (2003).
2. Fromer, M. *et al. Nat. Neurosci.* **19**, 1442–1453 (2016).
3. Schizophrenia Working Group of the Psychiatric Genomics Consortium. *Nature* **511**, 421–427 (2014).
4. Loh, P.-R. *et al. Nat. Genet.* **47**, 1385–1392 (2015).
5. Albert, F.W. & Kruglyak, L. *Nat. Rev. Genet.* **16**, 197–212 (2015).
6. Brem, R.B., Yvert, G., Clinton, R. & Kruglyak, L. *Science* **296**, 752–755 (2002).
7. Gusev, A. *et al. Am. J. Hum. Genet.* **95**, 535–552 (2014).
8. Albert, F.W., Treusch, S., Shockley, A.H., Bloom, J.S. & Kruglyak, L. *Nature* **506**, 494–497 (2014).
9. Aguet, F. *et al.* Local genetic effects on gene expression across 44 human tissues. Preprint at *bioRxiv* http://dx.doi.org/10.1101/074450 (2016).
10. He, X. *et al. Am. J. Hum. Genet.* **92**, 667–680 (2013).
11. Gusev, A. *et al. Nat. Genet.* **48**, 245–252 (2016).
12. Gamazon, E.R. *et al. Nat. Genet.* **47**, 1091–1098 (2015).
13. Button, K.S. *et al. Nat. Rev. Neurosci.* **14**, 365–376 (2013).
14. Robinson, E.B. *et al. Nat. Genet.* **48**, 552–555 (2016).
15. Keren, L. *et al. Cell* **166**, 1282–1294.e18 (2016).

When size matters: *CHD8* in autism

Martin W Breuss & Joseph G Gleeson

Recent models studying loss of the mouse homolog of the autism-associated gene *CHD8* show altered Wnt signaling, cell fate and proliferation. How do these findings shape our understanding of this disease?

Among *de novo* mutations, *CHD8* makes the single biggest known contribution to autism spectrum disorders (ASD). Previous work suggests that the protein functions at the center of a complex network of autism genes[1]. Encoding chromodomain helicase DNA-binding protein 8, *CHD8* acts as a transcriptional repressor by remodeling chromatin structure and recruiting histone H1 to target genes. This supposed function as a 'puppet master' pulling the transcriptional strings of disease-relevant pathways makes it an attractive target in tackling one major question: is there a common underlying mechanism for ASD? In this issue of *Nature Neuroscience*, Durak *et al.*[2] address this question by employing *in utero* knockdown of the mouse homolog *Chd8*. They found that *Chd8* depletion during the neurogenic period shifts the balance toward the production of neurons at the expense of progenitor cells. Adult mice exhibited social behavior deficits reminiscent of those in subjects with *CHD8* mutations. Such abnormalities have also been reported in a recently published study by Katayama and colleagues employing a haploinsufficient engineered mouse model[3].

Durak *et al.* correlated these changes with alteration of the transcriptome that affected a plethora of ASD-risk genes. Two major pathways emerged as altered by *Chd8* depletion: epigenetic regulation via Polycomb repressive complex[2] and canonical Wnt signaling. Whereas the first finding might have important implications for the steady-state suppression of neuronal genes in progenitors, the second yielded the most spectacular results. Forced activation of Wnt signaling not only rescued the precocious neurogenesis but also remedied the long-term social behavior deficits. Notably, the authors show that the connection of Chd8 and this pathway is cell-type specific: in neuronal progenitors, its depletion caused activation; in the human embryo

Martin W. Breuss and Joseph G. Gleeson are in the Department of Neurosciences, Howard Hughes Medical Institute, Rady Children's Institute of Genomic Medicine, University of California, San Diego, La Jolla, California, USA.
e-mail: jogleeson@ucsd.edu