

All in the mind

Findings of an excess of *de novo* structural variants in cases of autism and schizophrenia have raised hopes that neuropsychiatric conditions may yet prove genetically tractable. Past experience suggests that success in finding causative variants will require exceptional rigor and caution.

Many fields of genetics, for example, linkage for mendelian and complex traits, transcriptomics, candidate and genome-wide association studies, medical resequencing, and copy number and structural variant (CNV) association, underwent or are undergoing a process of evolution from biologically inspired guesswork to standardized, adequately powered and statistically grounded methodologies. Each generation of geneticists seems to have rediscovered afresh the problems of multiple testing, ascertainment bias and false positives. Most eventually hit upon multiple experimental replication as an empirical way to deal with a multitude of unexamined confounding factors, and worry that even in this they may be indulging their confirmation bias. The casualty rate in this self-assembled process of discovery is unacceptably high. In this issue, Nicole Allen and colleagues (p 827) describe how they built the SzGene database to carry out a meta-analysis of over 1,000 largely inconsistent genetic association studies of schizophrenia from the candidate gene era. From a field of 3,608 published common variants in 516 different genes, they find just four results with “strong” epidemiological credibility.

Of course, curious enthusiasm drives most investigators. All the historical perspective in the world of science could not prevent us from immediately comparing the genes on this list to the regions in which Walsh *et al.* (*Science* **320**, 539–543; 2008) detected structural variants in their recent case-control analysis of schizophrenia. This comparison will be disappointing, as common structural variants (frequency >1%) were not overrepresented in schizophrenic individuals in this study. Rather, a smattering of rare CNVs was found to be collectively about threefold more frequent in individuals with schizophrenia than in controls (after devoting an equal effort to discovering variants in cases and controls, a procedure now thankfully considered essential by most referees). These variants are notable by their rarity and by being previously undescribed in the Database of Genomic Variants (<http://projects.tcag.ca/variation/>), both marks of uncertain significance given that the discovery of structural variants is very much an ongoing activity.

Unfortunately, the case-control design does not distinguish between a mixture of familial variants (of unknown penetrance and significance) and *de novo* variants (found only in the affected

individual). The distinction is important for indentifying causal variants (McCarroll and Altshuler, *Nat. Genet.* **39**, S37–S42; 2007) because the frequency of mutation creating novel structural variants is expected to be much less than their frequency of transmission. An experimental design genotyping trios is indeed feasible and informative: this is how Sebat *et al.* (*Science* **316**, 445–449; 2007) found *de novo* CNVs in 12 of 118 individuals with sporadic autism, a tenfold higher frequency than the 1% *de novo* CNV found in controls.

Now, Bin Xu and colleagues (p 880) find an elevated frequency of *de novo* structural variants in individuals with sporadic schizophrenia by comparing their genotype to that of both parents. Rare inherited variants were less markedly elevated in sporadic cases. Of particular interest are the recurrent microdeletions at 22q11.2 found in sporadic schizophrenia or schizoaffective disorder.

If many genes can be perturbed to produce a related set of psychiatric phenotypes, how can we establish a causal relationship? What if the visible rearrangements are the byproduct of a mutational process? Sebat *et al.* (*Science* **316**, 445–449; 2007) considered that the hypothesis positing common causation of autism and CNV by a “fragile genome disorder” would predict not single *de novo* CNV but larger numbers clustered in affected individuals. They did not find clusters of CNVs in any individual. However, this conclusion may be open to reevaluation, as existing chip-based methods assess larger CNV and the overwhelming majority of variants smaller than 10 kb may require higher-density arrays or even sequencing to genotype them.

In an excellent review on the genetics of autism, Abrahams and Geschwind (*Nat. Rev. Genet.* **9**, 341–355; 2008) caution that “For some of the very rare, virtually unique, mutations even large sample sizes will not be sufficient to demonstrate statistical association, although the biological significance of the mutation may be clear.” To which we add our standard warning: if these disorders of mind are oligogenic and we do not have quantitative measure of the frequency and circumstances of discovery of the mutations, as well as the genomic background on which they occurred, we have only part of the picture. Incomplete notions of “biological significance”, unsupported by statistical significance, may once again lead us back into the wilderness. ■