

## Asking for more

Because of the usefulness of genome-wide association study (GWAS) data for mapping regulatory variation in the human genome, the journal now asks authors to report the co-location of trait-associated variants with gene regulatory elements identified by epigenetic, functional and conservation criteria. We also ask that authors publish or database the genotype frequencies or association *P* values for all SNPs investigated, whether or not they reached genome-wide significance.

Just as Mendelian genetics has been successful in identifying genes encoding transcription factors that define disease-causing genetic networks, it seems likely that GWAS will succeed in defining a proportion of the binding sites for the same transcription factors in the enhancers and promoters of genes influencing liability to the common forms of related diseases. This was one of the major conclusions from the many discussion topics at our second *Nature* conference in China, the Genomic Analysis of Diseases Workshop (<http://www.natureasia.com/en/events/gad/>). As with last year's meeting (*Nat. Genet.* **43**, 613, 2011), we came away with recommendations for increasing the impact of publications in the journal—particularly by implementing ways in which the large number of new biological hypotheses produced by GWAS can be made more accessible to other investigators who may not yet be aware of their use and importance.

We note that, although basic common-variant GWAS papers reporting large numbers of new loci are still publishable, there is now an increasing emphasis on new biological insights into mechanisms of disease. Last year, it was still a live question whether different populations experience different subtypes of common diseases that might be explained by their particular genetic architecture or gene-environment interactions. Now, after a large number of studies (with most admittedly still in European and Asian populations), this idea has less support than the alternative—that common variants can differ in frequency and linkage disequilibrium structure in ways that influence their detection in GWAS in different populations.

As of June 2 of this year, GWAS has been applied to 666 diseases and traits in a total of 1,271 publications (<http://www.genome.gov/gwastudies/>). Of these publications, 313 have appeared in this journal, reporting nearly half (1,891 out of 3,869) of the SNPs with an association *P* value of  $5 \times 10^{-8}$  or lower. However, we are constantly reminded of the potential

usefulness of SNP associations that do not reach genome-wide levels of significance by researchers wanting data to use in meta-analysis or in combination with nongenetic evidence for hypothesis generation. Indeed, the National Human Genome Research Institute (NHGRI) catalog does contain SNP associations up to a published *P*-value cutoff of  $1 \times 10^{-5}$ , which is a common threshold for initial SNP selection in many studies. There are also some excellent examples where authors have made public the *P* values for all the SNPs tested. We commend these as best practice (in two examples, data from the studies (*Nat. Genet.* **39**, 870–874, 2007, and *Nature* **467**, 832–838, 2010) were deposited at GWAS Central (<https://www.gwascentral.org/>).

It is now feasible to find associations with a greater range of allele frequencies than was previously possible, although both replication and the problems of population stratification become more difficult as variants decrease in frequency. It is good practice to discuss significantly associated rare and common alleles with respect to genes and regulatory elements, as was done, for example, by Gosia Trynka and colleagues (*Nat. Genet.* **43**, 1193–1201, 2011). In this article, the authors very usefully list SNPs, allele states and genotype frequencies in each population by affected status, also mapping SNPs, wherever possible, to gene regulatory elements. Such elements have now been identified for a wide range of cell types by, among others, the ENCODE Project Consortium (*PLoS Biol.* **9**, e1001046, 2011), meaning that it should be straightforward to report the relationship of many GWAS SNPs to such elements.

Not only will the improvements in reporting in this journal increase the impact of papers by engaging users in other research communities, these practices will go a long way toward demonstrating that GWAS is no obscure epidemiological diversion but, rather, the most efficient way we currently have to identify functional regulatory variation in the human genome. ■