

DISEASE GENOMICS

Seeking rarer variants for less

“ Both studies nicely illustrate the different strategies that can be used to bring the benefits of sequencing approaches to large-scale studies ”

Although methods such as exome sequencing have greatly improved our ability to identify low-frequency or rare variants and *de novo* mutations in the human genome, the large sample sizes required to assess their contributions to disease can be prohibitive owing to the high cost of sequencing. Two recent studies highlight different approaches towards making this a more affordable goal.

One area in which there has been particular interest in the role of *de novo* mutations is autism spectrum disorders (ASDs): hundreds of candidate genes for ASDs have been identified by exome sequencing to identify such mutations. However, a key difficulty in this field is the large amount of genetic heterogeneity among people affected by ASDs; this makes it challenging to determine the contribution of particular mutations to the condition. To address this issue, O’Roak and colleagues developed

an ultra-low-cost method for resequencing candidate genes in large numbers of people. They modified an existing method that uses molecular inversion probes combined with massively multiplexed sequencing. They improved the efficiency and affordability of the approach in several ways, including automation of the workflow and cutting reagent costs to less than US\$1 per gene per sample. The authors used this method to resequence 44 ASD candidate genes in 2,494 ASD probands, and they identified 27 *de novo* mutations in 16 of the genes. Fifty-nine per cent of these mutations have predicted severe effects on protein function. They estimate that *de novo* mutations in six of the candidate genes contribute to 1% of sporadic ASD cases.

In another study, Auer and colleagues used imputation to identify low-frequency variants that contribute to blood cell traits, such as counts of platelets and white blood cells. These kinds of traits can be used as intermediate phenotypes for studying the genetics of important common conditions, such as cardiovascular disease. The authors first sequenced the exomes

of a reference panel of 761 African Americans to identify previously uncharacterized rare variants. Then, in a much larger sample of more than 13,000 African Americans, they carried out genome-wide SNP genotyping using an Affymetrix array platform to characterize common SNPs in these individuals. By using imputation — a statistical method that predicts the genotypes of variants that are not directly genotyped — they were able to assess with greater power the low-frequency variants in these individuals and to test them for association with blood cell traits. Several novel associations were identified: for example, between missense variants in the lactase (*LCT*) gene and higher counts of white blood cells.

Both studies nicely illustrate the different strategies that can be used to bring the benefits of sequencing approaches to large-scale studies; these strategies should be applicable to a wide range of traits and conditions.

Louisa Flintoft

ORIGINAL RESEARCH PAPERS O’Roak, B. J. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 15 Nov 2012 (doi:10.1126/science.1227764) | Auer, P. L. et al. Imputation of exome sequence variants into population-based samples and blood-cell-trait-associated loci in African Americans: NHLBI GO exome sequencing project. *Am. J. Hum. Genet.* **91**, 794–808 (2012)

FURTHER READING Bamshad, M. J. et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nature Rev. Genet.* **12**, 745–755 (2011) | Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nature Rev. Genet.* **11**, 499–511 (2010)