nature genetics

Structural variants deconstruct the genome

Common genomic structural variants predispose to deleterious *de novo* genomic rearrangements. Understanding how they do so will require population studies across the continuum of genomic variation and ethical discussion of the nature and uses of human variation.

Structural variants can have stable, heritable effects on gene expression and effects on SNP variation and detection. They also influence mutagenesis and genome evolution. In this issue, David Koolen and colleagues (page 999), Charles Shaw-Smith and colleagues (page 1032) and Andrew Sharp and colleagues (page 1038) report a new 17q21.31 microdeletion syndrome arising from rearrangements of low-copy repeats. In his accompanying News and Views (page 974), James Lupski explains that recurrent *de novo* deletions (and indeed, corresponding duplications) are an expected consequence of this form of genomic variation. What is new, however, are the recently developed tools and concepts that, in principle, permit examination of the genomes of entire populations for the structural variants underlying spontaneous genomic disorders.

Advances in sequencing and genotyping have meant a gradual expansion of genetic analysis of point variation from mendelian mutations via rare variants to common SNPs. In contrast, the revolution in discovery of structural variants has been as abrupt, unpredictable and complex as the variants themselves. Their effects have yet to be fully integrated with the genetics of mutations and standing variation, and many common structural variants, inversions, large duplications and deletions remain to be discovered. While each of the groups searched systematically for structural variation using BAC array hybridization, the progressive sieve-like strategy of Sharp *et al.* initially concentrated attention on regions of segmental duplication. After first identifying common polymorphic variants, they could then screen for rarer *de novo* and familial rearrangements.

One can readily imagine microdeletions resulting from homologous recombination between direct repeats within a chromosome, and a polymorphic variant bearing this configuration may well have a higher probability of rearrangement. Lupski and Sharp *et al.* discuss the observation that rearrangements were predominantly recovered on the inversion-bearing H2 haplotype. Sharp *et al.* add that in three other genomic disorders, microdeletions were found preferentially on the inversion-bearing chromosome. This observation raises the possibility that the less-frequent haplotype usually carries the less-stable configuration of repeats. With chromosomes this complicated, the risk conferred by each structural genotype is unlikely to be simple to assess, and deletions occurring by meiotic recombination in heterozygous combinations will also need to be considered. Shaw-Smith *et al.* analyzed one microdeletion that apparently resulted from meiotic recombination between the H1 and H2 forms of the polymorphic17q21.31 inversion.

The 17q21.31 inversion is carried at a frequency of about 20% in populations of European ancestry, with evidence that it is under positive selection (*Nat. Genet.* **37**, 129–137; 2005). We now know that rearrangements sponsored by this structurally distinct haplotype could account for up to 1% of cases of mental retardation. Between the population genetics of the inversion and the affected individual with a *de novo* microdeletion lies the vast multidimensional territory of human genome variation.

To understand the mechanisms whereby structural variants sponsor rearrangements, it will be necessary to measure, for example, microdeletion frequency in populations that carry the structural variants at different frequencies. On the surface, this exercise is no different than surveying different populations for risk of common and complex disease, or using founder populations for discovery of mendelian mutations. However, populations may differ, not only in their frequency of structural variants, but in their liability to generate novel variants, some of which will be deleterious. Variants within the range of normal variation in one group may produce disease in another sample. So, in drawing contrasts between populations, there is some danger that we might mismeasure not only the genetic risk characteristics of a particular population but also its future evolutionary potential. This is a qualitatively different point that merits some ethical consideration.

To make use of circulating gene variants along with the structure and history of the human population to predict determinants of health and disease demands an analytical effort that is ethical and social as well as scientific. Abdallah Daar and colleagues (*Nat. Rev. Genet.* **7**, 414; 2006) discuss the need for a sufficiently broad conceptual framework within which to integrate our scientific and ethical views of mutations, SNPs, structural variants and other genomic variation. They warn that ethical, legal and social consideration of human genome variation must be able to incorporate new scientific advice along with new discoveries. We add that the ethical framework must be flexible and broad and well-grounded in a scientific understanding of the evolutionary history of the human race (*Nat. Genet.* **36**, S3; 2004). In conclusion, Daar *et al.* cast structural variation as a great challenge with a large incentive for urgent interdisciplinary discussion on the meaning of human genome variation. ■