Nature Reviews Genetics | AOP, published online 30 December 2010; doi:10.1038/nrg2939

GENOMICS

No half measures for haplotypes

Individual human genomes are coming thick and fast, but they lack information on the combination of alleles along single chromosomes — the haplotype. Although common haplotypes can be identified using statistical and family-based approaches, direct methods for readily determining individual haplotypes would be extremely valuable for many areas of research, from population genetics to pharmacogenomics. Methods reported by two groups now offer this possibility.

Although technically distinct, the two approaches address the same problem: how to enable a sequencing or genotyping platform to 'see' haploid rather than diploid information. Fan and colleagues used microfluidics: they developed a device that separates metaphase chromosomes from single cells and amplifies them independently. The amplified products from several chromosomes are then pooled into mixtures that contain only one copy of each homologous pair of chromosomes. This 'haploid' amplified material is then analysed on whole-genome genotyping arrays. The authors applied this strategy — termed direct deterministic phasing (DDP) — to resolve the whole-genome haplotype of a European individual. Over 99% of SNPs were phased in this assay; phasing the SNPs by direct sequencing of the amplified material gave 99.8% concordance with phasing by genotyping.

Kitzman and colleagues separated the genetic material from homologous chromosomes by dividing a fosmid library — constructed from the genomic DNA of a Gujarati Indian individual — into 115 pools. Each pool contained ~5,000 independent clones, each with a ~37 kb insert, and so constituted a random ~3% of the diploid genome. The authors performed high-throughput sequencing on the pools and showed that >99% of reads from each pool are derived from only one chromosome for any genomic single location. Therefore, mapping the reads from individual pools to the reference genome can enable haploid genotype calling. Kitzman *et al.* combined the haploid genotype calls with unphased variants from shotgun whole-genome sequencing to assemble the genome-wide haplotype of this individual. This approach also allows phasing of rare and private variants. Both groups found excellent concordance of their data with HapMap phase predictions, but also found evidence of incorrect phasing when statistical methods rather than experimental methods were used.

These studies also illustrate the potential applications of individual haplotype data. For example, Kitzman *et al.* found evidence of haplotype blocks that might be derived from a poorly ascertained ancestral population, thus demonstrating the value of individual haplotype data to population genetics. It is expected that both approaches would be scalable, so direct determination of phasing could become an integral part of analyses of genetic variation.

Mary Muers

ORIGINAL RESEARCH PAPERS Fan, H. C. *et al.* Whole-genome molecular haplotyping of single cells. *Nature Biotech.* 19 Dec 2010 (doi:10.1038/nbt.1739) | Kitzman, J. O. *et al.* Haplotype-resolved genome sequencing of a Gujarati Indian individual *Nature Biotech.* 19 Dec 2010 (doi:10.1038/nbt.1740)