

TECHNOLOGY

DNA has nowhere to hide

For many years, forensic scientists wanting to determine whether an individual was present at the scene of a crime have had to laboriously test a suboptimal biological sample with a handful of DNA markers. A new report promises to end this challenging state of affairs — it shows that SNP genotyping arrays can detect trace quantities of a suspect's DNA in a complex mixture containing hundreds of other individuals.

High-density SNP genotyping has transformed the identification of disease-associated gene variants, yet its uses were never extended to person identification owing to the belief that individuals within a pooled sample would be impossible to resolve. The authors of this paper have reversed this assumption — first by investigating the feasibility of using hundreds and thousands of SNPs to genotype individuals in a complex mix, then by probing the limitations of the approach through simulations, and finally by testing the idea in a real-life context.

The basic approach involves determining the SNP allele frequencies

of the sample (as inferred from allele probe intensities) and then calculating whether the frequency profile seen in a suspect is closer to the profile seen in the sample than in the reference population. How well does the theory hold? This question was initially addressed in a simulation in which the allele frequencies, the fraction of the suspect's DNA in a sample and experimental noise were modelled on a random sample of over 1,400 individuals of known genotype. The results are striking: by genotyping only 50,000 SNPs, single individuals can be detected even when each one makes up only 0.1% of the sample. By contrast, the amount of noise only marginally affects accuracy. These values were then confirmed experimentally when the framework was applied to real data — on a series of eight mixtures made up of varying proportions of HapMap individuals that were typed using Affymetrix or Illumina platforms.

The ability to identify individuals in a complex mixture promises to transform the field of forensics.

Although the results are impressive, the method comes with a few caveats, not least the need to account for any differences in ancestry between the reference and the experimental populations. And it is not just criminals who risk being unmasked: individuals whose DNA is pooled for research purposes — for example, in genome-wide association studies — are at risk of having their anonymity unveiled by such probing analyses. Although such a risk is currently low, the National Institutes of Health and the Wellcome Trust have responded to this work by removing some of their genetic data from public view.

Tanita Casci

ORIGINAL RESEARCH PAPER Homer, N. *et al.* Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS Genet.* **4**, e1000167 (2008)

FURTHER READING Jobling, M. A. & Gill, P. Encoded evidence: DNA in forensic analysis. *Nature Rev. Genet.* **5**, 739–751 (2004) | Weir, B. S. *et al.* Genetic relatedness analysis: modern data and new challenges. *Nature Rev. Genet.* **7**, 771–780 (2006) | McCartney, C. LCN DNA: proof beyond reasonable doubt? *Nature Rev. Genet.* **9**, 325 (2008)

