

EXPLORING THE FORENSIC APPLICATION OF 700K HIGH-THROUGHPUT WHOLE GENOME ARRAY

LIU Jing¹, LI Cai-xia¹

National Engineering Laboratory for Forensic Science, Beijing Engineering Research Center of Crime Scene Evidence Examination, Institute of Forensic Science

In the field of forensic genetics, the application of STR (short tandem repeat) polymorphisms through capillary electrophoresis (CE) has been used as a routine tool for many years. However, the frequent utilization of single nucleotide polymorphism (SNP) markers for forensic field has been accepted in recent years. Compared with the STRs, the application of SNPs showed great advantages in many sides. Single nucleotide polymorphisms are the changes in the DNA sequence at a particular site, which can be found in the human genome about once in every 1000 bp. Applied in forensic field, they could be used for multiple types of investigation. In the 22nd Congress of the International Society of Forensic Genetics (ISFG) held in Copenhagen, Denmark, six distinguished scientists with experience in SNPs defined 4 basic categories: Individual Identification (IISNPs), Ancestry Inference (AISNPs), Lineage Inference (LISNPs), and Phenotype Inference (PISNPs). The experts discussed the benefits and limitations of SNP. In order to meet each purpose of identification or inference, different group of optimal SNPs need to be identified with discretion. With time passed by, some of the limitations or challenges of SNPs have been solved.

There're all kinds of study to employ SNPs for forensic research in the exist papers. However, most of the works only concentrated on single application, such as ancestry analysis or individual identification. Here we tried to explore the forensic application of Illumina Infinium Global Screening Array for comprehensive purpose, including admixture analysis for ancestry inference, kinship identification for lineage inference, eye color and height prediction for phenotype inference at a time. We applied BeadChip Array Global Screening Array-24+ v1.0 chips, which is designed for global populations, on users collected in a town of western China. Excluding 1 sample with low quality, we made a full process for the left 714 samples. Admixture analysis showed their ancestries focus on populations in Europe, which validated their ancestry sources. Kinship analysis predicted the all the users' relationship from level 1 to level 9. Then we validated the relationship of level 1-3 by investigation locally. After imputation of the original 700,000 SNPs, we obtained above 10,000,000 SNPs for phenotype prediction. Applied IrisPlex system from former paper, we predicted the eye color of the 714 users (table 1). Then we compared the result with their photos.

tab.1 summary of eye color prediction result

probability	brown eyes	blue eyes	Intermediate color eyes
$\geq 90\%$	611	2	0
80% ~ 90%	48	7	0
70% ~ 80%	26	1	0
60% ~ 70%	10	3	0

50% ~ 60%	6	0	0
-----------	---	---	---

We also picked around 200 SNPs for height prediction. The results implied that we could predict height from SNPs with average as 5.43 cm.

With the results above, we successfully validated the application and made some improvement. We also hope to promote and expand the application of SNPs in the forensic field.