

USING SNPs FOR DNA PHENOTYPING AND KINSHIP INFERENCE

Ellen McRae, Parabon NanoLabs, Inc.

DNA phenotyping refers to the prediction of an individual's appearance from his or her DNA. Instead of traditional short-tandem repeats (STRs), which are only useful for identity matching to a known suspect or a database, DNA phenotyping employs single nucleotide polymorphisms (SNPs), which contain the genetic information that determines the physical differences among people. Using a novel approach that includes deep data mining and advanced machine learning, and with funding from the U.S. Department of Defense, we have developed a DNA phenotyping system – Parabon Snapshot – that can, with high accuracy, predict an individual's ancestry, eye color, hair color, skin color, freckling, and face shape using only information from SNP genotypes. Snapshot has been validated on unknown samples and shown to be accurate with as little as 2 ng of DNA.

In addition, we have also developed a novel method for determining kinship between any pair of individuals using genome-wide SNP genotypes. Whereas STRs can only be used for parentage, and mitochondrial and Y-chromosome SNPs depend on a direct maternal or paternal lineage, Snapshot's kinship model can be applied to any pair of subjects and has been shown capable of distinguishing up to 6th-degree relatives (second cousins once-removed) from unrelated pairs of individuals.