Abstract 3

DNA IDENTIFICATION BY PEDIGREE LIKELIHOOD RATIO

Jianye Ge[1, 2] and Ranajit Chakraborty[2]
[1]Department of Biomedical Engineering, University of Cincinnati
[2]Center for Genome Information, Department of Environmental Health, University of Cincinnati College of Medicine, 3223 Eden Avenue, Cincinnati, OH, 45267

Over the past two decades efficient use of DNA forensics in criminal and civil investigations established it as a reliable tool for personal identification. Concerns have recently shifted on developing an infrastructure of DNA-based identification of war victims in mass graves, missing soldiers or military personnel from past wars, missing person from mass disasters caused by natural catastrophes or terrorism acts (e.g. aircraft crashes, World Trade Center tragedy and Southeast Asia Tsunami), etc. When direct reference samples (i.e. antemortem samples) from missing individuals are not available, current identification methods are based on ranking of likelihood ratios that are constructed from comparison of DNA profiles of the remains of suspected missing person with reference sample of family members, a pair at a time. However, when multiple members from the same family are typed as references, this procedure results in loss of information, since family members are not independent.

In this research, we provide a novel method based on the classical Elston-Stewart algorithm for personal identification with autosomal markers, by jointly considering DNA profiles data from all available family references. The method can handle any pedigree structure, even with incest relationships. Persons are identified by ranking the pedigree likelihood ratios with alternative hypotheses (e.g., the missing person is unrelated to the family members of the pedigree) for all putative pedigrees.

In this method, pedigree likelihoods incorporate population substructure effects. A two-phase mutation model is also incorporated for the short tandem repeat loci to avoid possible false exclusions. Simulation study shows that the effect of mutation to pedigree likelihood ratio is moderate, but population substructure can significantly decrease likelihood ratios. Moreover, computational complexity is exponential in terms of the number of untyped family members in a pedigree. In order to reduce the complexity, multiple strategies are introduced, such as collapsing the number of alleles unobserved in the pedigree and evidence sample, separating a pedigree into multiple independent sub-pedigrees, whenever possible, and avoiding reference to the possible genotypes of some untyped family members, etc. Generally, by this method most forensic identification cases can be solved in practically feasible computational time.

This work has been funded by grant NIH-GM-41399 (to RC).