

A SOFTWARE APPLICATION TO MANAGE AND SEARCH DNA PROFILES AND PERFORM FAMILIAL SEARCHES

Rick Staub, Jeff Boschwitz, Sid Sinha, Carlos Ramos, Pedro Reinoso, Sean Dineen, Chad Holovach, Carlos Baena, Dmitry Maschenski

We have developed a multi-function software application capable of performing high-speed searches of large databases of DNA profiles from evidence and reference samples. STR profiles from evidence or reference samples can be uploaded to a database in either batch or manual mode and all samples compared to reference and/or evidence samples already in the databases. Reports are produced following uploads listing all samples that are perfect matches with one another or single-source samples that are possibly contained within mixture profiles. The stringency levels of searches may be adjusted to account for possible drop-out of low-level components of evidence profiles.

Avoidance of adventitious matches can be accomplished by utilizing a search threshold based on calculated match probability rather than number of STR loci typed in the profile. Match probabilities for all selected racial/ethnic groups are produced for all profiles searched in the database. Perfect match searches for single-source profiles can be carried out with blazing speed. Even matches of single-source profiles to mixture profiles can be carried out exceedingly quickly in very large databases.

The software solution also provides the capability of performing familial searches in large databases. These searches are based on kinship likelihood ratios rather than numbers of shared alleles. Searches for parent/child, full-sibling, and avuncular/half-sibling relationships are possible utilizing reports that rank potential relatives by likelihood ratio. To assist in finding true kinship relationships between two iso-patrilineal males, Y-STR profiles are automatically retrieved for all male samples that have been entered into the database with them.

The software provides a web-based interface with several built-in options for ease of use. In particular, all manually entered data are entered only after selecting which of several possible amplification kits was used to type the sample. The loci then appear in the correct order for ease of entry. Additionally, for familial searches, users have the option of selecting the race used for likelihood ratio calculations, one of which is "infer race" and is based on the highest match probability among five major racial groups as an indicator of the most likely race of an unknown profile.

This software solution has a multitude of uses in forensic operations of crime laboratories as well as for military forensics in order to glean more intelligence information from evidence that may not be fit for entry into standard government convicted offender DNA databases, especially if it were used in conjunction with these large databases. Examples of these uses will be discussed and demonstrated in this presentation.