## MATRIX FILES: PROBLEMS & SOLUTIONS (OBTAINING ACCEPTABLE DATA FOR EASY INTERPRETATION)

## Jaime Handelsman, Miriam Narvaez, and Demris Lee

Armed Forces DNA Identification Laboratory, Rockville, MD

## 

A major advantage of fluorescent dye labeling is the ability to multiplex short tandem repeat (STR) loci with different dyes and automate the sequencing process. The ABI 377 is equipped to detect each dye based on its emission spectrum. Currently four different fluorescent dyes are used to detect the bases in an mtDNA sequence and the alleles of STR loci. These include 5-FAM (blue), JOE (green), NED (yellow) and ROX (red). Each of the four dyes emits their maximum fluorescence at different wavelengths with some overlap in the emission range.

A matrix file is a mathematical description of the spectral overlap, which is determined from the automated analysis of dye-labeled DNA fragments (matrix standard samples) for each of the four dyes. With this information, the matrix file virtually instructs the sequencer to filter out the overlap, allowing the sequencer to distinguish between the signals of each dye and display only one color for each base or allele on an electropherogram.

Bad matrix files can result in poor spectral separation or pull-up. This can produce extraneous peaks in the STR electropherograms. The electropherogram will exhibit minor peaks of a different color under the major peak. Likewise, the mitochondrial DNA sequencing electropherograms exhibit high background. Gels analyzed with improper matrix files have the potential to cause significant problems during data analysis and interpretation. Pull-up peaks that fall within STR allelic bins can complicate mixture analysis. Similarly, high background signal can cause difficulties in distinguishing heteroplasmic positions.

These problems can be eliminated with the use of a proper matrix file. A matrix file must be made for each instrument. Our laboratory has ten 377s that are in operation daily. The instruments are used for STR analysis as well as mitochondrial DNA sequencing. After many months of operation our laboratory began experiencing problems with inaccurate matrices. New matrices were created and service technicians optimized the laser alignment. Nevertheless, we continued to experience problems. The STR data repeatedly observed pull-up of the green into the blue and yellow even at low relative fluorescent units. High T background (Rox) was observed with the sequencing data.

The following are various methods our laboratory has tried to help reduce our matrix problems.

**Plates were cleaned in 1M sodium hydroxide.** It is believed that if plates are not washed properly, excess dye can accumulate on the gel plates. During a run, the laser can potentially detect this excess dye, which can cause high levels of background noise. Due to our ongoing matrix problems, we soaked the plates in a 1M solution of Sodium Hydroxide for two hours to remove any excess dyes remaining on the plates. The cleaned plates produced much cleaner data with none to very little background noise and/or pull-up. Currently our lab is in the process of installing dishwashers in order to keep the gel plates clean.

**Volumes of DNA template and reagents were adjusted.** Mitochondrial DNA template volumes used in cycle sequencing reactions were increased. Another method employed has been to decrease the amount of loading buffer in an attempt to increase the concentration of mitochondrial DNA loaded onto sequencing gels.

**New matrices were created for new gel conditions.** Recently, our database section switched from acrylamide/bis gels to Long Ranger el Packs for sequencing gels. The acrylamide/bis gels are 4% acrylamide and the Long Ranger Gel Packs are 5% Long Ranger. The different gel concentrations may affect the matrix file. Therefore, separate matrices are being created for different gel pouring methods.

**Plates and cassettes were marked and tracked.** A matrix file may produce clean data when a specific cassette and set of plates are used, but can produce insufficient data when a different cassette and/or set of plates are used. It may be that ach cassette fits into the sequencer slightly differ, causing the laser to obtain different results. Therefore, the gel plates and cassettes were labeled, and a tracking form was created. Before a run, the scientist would record the plate set number, the cassette number, and the sequencer used on the tracking form. After the run, the scientist would record whether the data obtained was clean or showed pull-up or high background noise. We eventually saw a pattern and were able to pair specific plates, cassettes and sequencers to produce better results.

It is recommended that new matrix files fro each 377 sequencer be created quarterly. In some cases, poor data collected has been successfully re-evaluated using these new matrices.