**IDENTIFICATION AND SECONDARY STRUCTURE ANALYSIS OF A REGION AFFECTING ELECTROPHORETIC MOBILITY OF THE STR LOCUS SE33**

Dennis Y. Wang, Robert L. Green, Robert E. Lagacé, Nicola J. Oldroyd, Lori K. Hennessy, Julio J. Mulero
Life Technologies, Human Identity Group, 850 Lincoln Centre Drive, Foster City, CA 94404

One of the key requirements in the development of any new multiplex STR kit is to ensure that the STR profiles generated by the new kit are concordant with the existing profiles stored on National databases. As such, the ideal approach for developing a new STR kit would be to utilize the existing primer sequences. However, such an approach is not always possible because of spatial limitations in new multiplexes due in large part to an increase in the number of loci. It is then of critical importance for the manufacturer to perform a comprehensive investigation into the impact that new primer sequences might have in order to minimize disconcordances between new and existing kits.

SE33 is one of the most informative markers in forensic use due to its high power of discrimination. During the developmental phases of the AmpFℓSTR® NGM SElect™ PCR Amplification Kit, a number of candidate SE33 primer designs were screened against > 1000 population samples to determine their concordance against commercial kits. One primer pair yielded a high frequency of discordant alleles when compared to the AmpFℓSTR® SEfiler Plus™ PCR Amplification Kit. This discordance was mostly specific to samples of African descent with an estimated frequency of 5.1% and was a result of a mobility shift of approximately +0.84 nt. In an attempt to understand the discordance, we cloned and sequenced all the affected alleles.

The sequencing analysis revealed that the cause of the discordance was a single nucleotide polymorphism (SNP) outside of the SE33 repeat but within the amplicon of this particular set of experimental primers. In total, we identified three different SNPs all within 9 nt of each other, each of which could cause the mobility shift individually. Using the DNA-folding software Mfold, the wild type amplicon was predicted to have a region of strong secondary structure composed of a 5 base-pairing stem-loop. All the SNPs that affected the mobility of the amplicon fell within the stem portion of the structure. Further characterization of this region using site-directed mutagenesis and thermostability measurements strongly suggested that this polymorphic region contains a secondary structure that, when disrupted due to the presence of a variant SNP, resulted in a mobility shift relative to the wild type sequence. To overcome this problem, the SE33 primers used in the final configuration of the NGM SElect™ Kit avoided the amplification of this polymorphic region yielding results that are highly concordant with the SEfiler Plus™ Kit.