GENETIC ANALYSIS OF SOUTHERN BRAZIL SUBJECTS USING THE POWERSEQ^{TM} AUTO/Y SYSTEM FOR SHORT TANDEM REPEAT SEQUENCING

Deborah Silva[1], Fernanda Sawitzki[2], Melissa Scheible[1], Sarah Bailey[1], Clarice Alho[2], Seth Faith[1]
[1]Forensic Sciences Institute, North Carolina State University
[2]Laboratory of Human and Molecular Genetics, Pontificia Universidade Catolica do Rio Grande do Sul

With the advent of Next-Generation Sequencing technology, sequencing of short tandem repeats (STRs) allows for a more detailed analysis when compared to size-based fragment analysis (capillary electrophoresis-CE). Subjects from southern Brazil present a singular genetic background: Majority of subjects with European ancestry (especially Portuguese, Italian, Spanish, and German) and a smaller number of individuals with African and Amerindian ancestry. Both autosomal and Y STRs have been analyzed in individuals from the State of Rio Grande do Sul, which is located in the extreme south of Brazil. However, all analyses published to date encompass data from CE-based fragment analysis. In this study, a genetic analysis of 61 individuals from Southern Brazil was performed on short tandem repeat sequences. Forensically relevant STRs were PCR-enriched using a prototype of the PowerSeq^{TM} AUTO/Y system (Promega Corp.). Next-generation sequencing was performed on an Illumina MiSeq instrument with MiSeq V3 sequencing chemistries. The raw data (FASTQ files) were processed using a custom designed sequence processing tool, *Altius*. Sequences were compared to previously published data on fragment analysis in three different population samples from Southern Brazil. Sequence-based allelic variants were observed in 9 autosomal and in 6 Y STRs. The number of different alleles based on sequence was higher when compared to those based on length: 38.6% higher in autosomal STRs and 14% higher in Y-STRs. The most polymorphic autosomal locus was D12S391, which presented 38 different alleles. Among the loci in the Y chromosome, DYS389_II presented the highest number of isoalleles. In comparison to CE analysis, Observed and Expected Heterozygosity, Polymorphic Information Content (PIC) and Genetic Diversity also presented higher values when the alleles were analyzed based on their sequence. The results obtained in this study demonstrate the importance of genetic analysis based on sequencing and how individuals from South Brazil present a much higher genetic sequence-based diversity when compared to the STRs based on fragment analysis.