Wojciech Łuczak
Laboratorium Technik Biologii Molekularnej
Wydział Biologii
Uniwersytet im. Adama Mickiewicza w Poznaniu

Opracowanie zestawu polimorficznych markerów STR do genotypowania ludzi oraz jego wdrożenie do analiz pokrewieństwa w dalszych relacjach rodzinnych

Streszczenie rozprawy doktorskiej w języku angielskim

The analysis of microsatellite sequences, also known as Short Tandem Repeats (STR), is one of the most important methods used in forensic genetics, including in biological kinship testing. Its widespread use results from its high informativeness, sensitivity, and the ability to analyze degraded DNA. Additionally, this method is characterized by low per-sample cost and a short turnaround time for analysis. Routine STR locus analysis in forensic genetics laboratories relies on the multiplex-PCR technique, which uses fluorescently labeled primers, and capillary electrophoresis, which enables the determination of the lengths of amplified alleles. In forensic investigations and kinship analysis, technology allowing for the simultaneous analysis of twenty CODIS STR loci (Combined DNA Index System) and additional STR loci, such as SE33, PENTA D, and PENTA E, using commercial reagent kits is commonly employed. The analysis of twenty CODIS STR loci is sufficient for relatively simple cases of biological paternity or maternity testing, but it proves to be insufficiently informative for kinship testing in second and third-degree familial relationships.

The aim of this implementation PhD was to develop the KinFinder method for investigating biological kinship in more distant familial relationships. The method was intended to be implemented in the GenMed Molecular Diagnostics Laboratory, and therefore, it needed to be based on the technology of multiplex-PCR and capillary electrophoresis. This approach would allow the company to use the method without incurring additional investment costs for new equipment, laboratory space, or employee training.

The development of the Kinfinder method involved a bioinformatic analysis of the human genome to identify 150-200 of the most polymorphic STR loci, followed by the design of PCR primers for the amplification of each selected locus, and laboratory screening of STR loci to estimate the heterozygosity of these loci in the Polish population. In the next phase, 50 of the most polymorphic loci were selected, and two multiplex-PCR reactions were developed for genotyping human DNA at the selected STR loci. After the design and validation of the method, with the ability to analyze 50 loci in two multiplex-PCR reactions, population

studies were conducted to determine the allele frequencies of each locus in the Polish population. The research described in this dissertation also included sequencing of the loci to link the length of the tandem repeat sequence to the length of the amplicons, as well as the creation of an allelic ladder and a reagent for calibrating the genetic analyzer. Population studies confirmed the very high polymorphism of the 50 developed STR loci. Their average heterozygosity was 88.07%, which is significantly higher compared to the average heterozygosity of CODIS loci in the Polish population, which is 78.95%. The most polymorphic loci in the KinFinder set (D8A26, D15L495, D13S742) had heterozygosity rates of 93.45%, 93.53%, and 93.9%, respectively, comparable to the most polymorphic STR locus used in forensic genetics (SE33), whose heterozygosity in the Polish population ranges from 93.4% to 95.4%, depending on the literature reports.

The results of the study show that the preliminary data used, as well as the newly developed method for estimating STR locus polymorphism, allowed for the verification of the heterozygosity of selected loci in the Polish population. As predicted, it was confirmed that the heterozygosity of some STR loci is higher than that reported in the 1000 Genomes project database and the STRCatalog and WebSTR databases, highlighting the limitations of the short-read sequencing technology used in the 1000 Genomes project for analyzing microsatellite sequence polymorphism in the human genome.

The statistical analysis of kinship tests in various familial relationships confirmed the high informativeness of the KinFinder method and its usefulness, particularly in second and third-degree kinship testing. In line with the objectives of this implementation PhD, the KinFinder method has been incorporated into the standard procedures of the GenMed Molecular Diagnostics Laboratory and has already been applied in over thirty kinship tests, including court cases.