Chapter II

Literature review

Person identification and paternity testing based on blood group substances, red blood cell enzymes and HLA-DQα have a limitation. To circumvent this problem, many other human genetic markers were developed. One of these markers is STR. The polymorphic character of short tandem repeat (STR) loci results from variation in the number of tandemly repeated units from one allele to another (Fig. 4). Different classes of STR loci have been described, including di-, tri-, and tetranucleotides.[24-30]

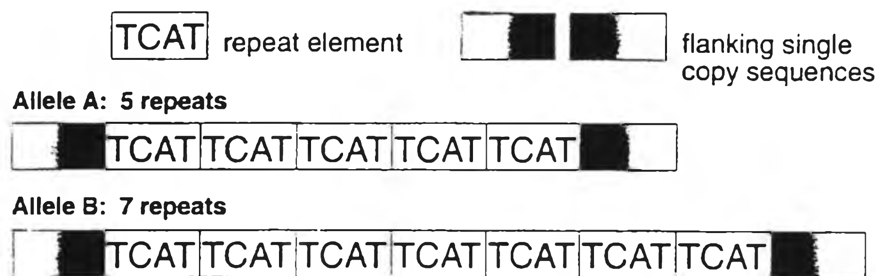

Fig. 4. Organisation and internal structure of short tandem repeats. The core repeat element is usually conserved. However, in complex STR systems, further genetic variation may be due to irregular repeats or short insertions of other elements.[9]

PCR-based STR markers provide several advantages for genetic mapping: (1) ascertainment of genotypic data is easier and faster than with hybridization-based VNTR probes;[4,14,16] (2) STRs often detect highly polymorphic loci;[24-30] (3) STRs are abundant throughout the genome and are easy to develop as genetic markers;[26,29-31] (4) the oligonucleotide-primer sequences flank the repeats.[24-27,30] Tri- and tetranucleotide repeats show better stability during PCR than do dinucleotide repeats, the increased PCR stability facilitates interpretation of genotypes.[23,27-30] The majority of the

markers detected from human genetic maps contain tetranucleotide repeats.[29-30] A lot of tetranucleotide repeat markers have been studied.

In 1994, Hammond et al. studied 13 unlinked STR loci derived from four population groups residing in Houston by a PCR-based DNA typing method and autoradiography for use as a highly discriminatory system of genetic markers in personal identification, specifically for parentage testing, forensic identification and medical applications.[35] They found that 3-5 base repeats amplify more authentically and provide more easily interpretable results than do dinucleotide repeat loci. Multiplex reactions can be performed under identical amplification conditions. Numerous loci are highly polymorphic (heterozygosity 70%-80%) and have a large number of genotypes. All loci show codominant inheritance consistent with the expected pattern for the chromosome to which the locus has been assigned. There is no random association of the alleles. The power of exclusion ranged from 19.01% to 80.72%. The combined average power of exclusion is >99% after typing of the first eight STR loci. Seven STR loci provide a most common genotype frequency of <1 in 10,000 people, for all four population groups. The mutation rate estimates were between $2.36x10^{-5}$ and $1.86x10^{-4}$. They have shown that, between loci, there is no association of the alleles. These STR loci provide an accurate, highly discriminating, sensitive, rapid technique and powerful statistical evidence for DNA typing.

In1995, Huang et al. presented Chinese population data for 3 tetrameric STR loci-CSF1PO,TPOX, TH01- that were typed using multiplex PCR and subsequent electrophoresis in DNA sequencing gels followed by silver staining detection.[36] All loci meet the Hardy-Weinberg expectation. In

addition, there is no evidence for association of alleles among the three loci. However, the data on HUMTH01 allele frequencies are different when compared with the Caucasians and African-Americans. This multiplex analytical procedure is relatively simple, rapid typing method with high resolution which can be implemented into most application - oriented laboratories at minimal cost. In addition, the data demonstrate that these loci can be useful for providing estimates of the frequency of a DNA profile in identity testing cases.

Martin et al.[37] studied Spanish population data on 7 tetrameric short tandem repeat loci by using PCR and subsequent analysis of the PCR products by denaturing polyacrylamide gel electrophoresis followed by silver staining. The loci were HUMFES/FPS, HUMvWA, HUMTH01, HUMF13B, HUMCSF1PO, HUMF13A1 and HUMTPOX. Different STR profiles were obtained by sequential loading of several amplification products corresponding to the same loci and by simultaneous loading of several amplification products corresponding to the different loci. Both typing approaches minimize labor, materials and analysis time. It has been shown that there was no evidence of association between the 7 loci and the allele frequency data can be used for PCR-based DNA profiles in the Spanish population.

In 1996, Lins et al. described the development of multiplex PCR amplification systems to allow the simultaneous amplification of 8 polymorphic tetrameric STR loci including HUMCSF1PO, HUMTPOX, HUMTH01, HUMVWFA31, HUMF13A01, HUMFESFPS, HUMBFXIII and HUMLIPOL.[38] The amplified fragments are separated by denaturing polyacrylamide gel

electrophoresis and detected by silver stain and fluorescence detection, both nonisotopic methods.

The respective silver-staining protocol was described by Bassam et al. [39] It is sensitive with a limit of detection upon visual inspection of the double-stranded DNA at approximately 1 pg/mm^2 band cross-section. This procedure is highly sensitive, avoids unspecified background staining without loss of contrast, uses less silver without oxidizing pretreatments and requires fewer steps. However, it is complex and time consuming and requires the preparation and handling of several solutions, in some cases immediately prior to use. Hence, the silver stain detection method is suitable for routine use in laboratories that desire an inexpensive, high-throughput analysis.

The other nonisotopic method is fluorescence detection. To that end, one of the two primers utilized in the reaction is labeled at the 5' terminus with fluorescein so that only one strand of each allele is visualized. Fluorescence detection offers advantages of speed and convenience, yet with a need for specialized equipment and appropriate software. The fluorescence data may be automatically analyzed, printed and stored in a database. Both silver stain and fluorescein detection methods offer rapid, simple and accurate analysis of multiplex STR systems.

Lins et al. revealed that only two of the eight loci contain variant alleles. The F13A01 locus contains one variant allele, designated allele 3.2, which is two bases shorter than the allele containing 4 repeat units. The TH01 locus displays the relatively frequent allele 9.3, which is one base shorter than allele 10. Sometimes silver stain detection could not resolve the allele 10,9.3 heterozygote sample for the TH01 locus. For fluorescence detection, only one strand of each allele is visualized and the two alleles are resolved.

They also demonstrated that the chance of two individuals to match (Pm) at all eight loci was 3.5×10^{-9} for African-Americans and 5.6×10^{-8} for Caucasian-Americans.[17] The combined typical paternity index (PI) is 497.08 for African-Americans and 186.69 for Caucasian-Americans. The combined power of exclusion (PE) which measures the ability of a genetic test to exclude a falsely accused individual for the eight loci is 0.998 and 0.996 for African-Americans and Caucasian-Americans, respectively. It is simple to obtain statistical analysis of populations, based upon accurate, precise and confident allele determination.

Rostedt et al. have developed a triplex PCR method for D3S1359, HumTH01 and HumTPO tetranucleotide loci and a duplex PCR method for HumFES/FPS and HumvWA31A tetranucleotide loci using high resolution polyacrylamide gel electrophoresis and silver staining.[40] A requirement for this multiplex amplification is that the fragment sizes of the loci to be amplified together must not overlap. Co-amplification of two or more loci effectively reduces the labour and cost necessary and yields more information in a shorter time period. The methods were evaluated for paternity testing and individual identification in the Finnish population. In paternity testing, the combined exclusion power of these five loci is 0.987. For individual identification purposes, the combined power of discrimination for these five loci is 0.999997. These five STR loci were found to be most relevant for forensic analysis and paternity testing.

Because of their properties and advantages, STRs have become ideal markers widely used in many laboratories. They tend towards being the standard markers in the future. Therefore, a Thai population database and

statistical analysis on STR loci are essential, before introducing them into DNA typing.