Chapter IV

Results and Discussion

DNA extraction

DNA extracted by this method exhibits a 260/280 ratio above 1.6. The amount of DNA from 5 – 8 ml blood sample in the individual extract, calculated from OD at 260 nm, ranges from 91.675 to 478.5 μg. Amplification of DNA from the samples was successful. The advantages of the salting-out method are: rapid, inexpensive, avoiding the use of hazardous reagents and requiring few steps. It can be routinely used on fresh blood samples in laboratories for person identification and paternity test based on STR and PCR technology. However, several pipetting steps are required and thus several pipette tips are introduced into the samples. Besides, the steps followed lysis and washing the red blood cells until the pellet appears white, increase the risk of sample contamination thus decreasing the yields. Therefore it should be ensured that cross-contamination has not occurred.

PCR amplifications

Each locus described contains tetrameric short tandem repeats, the polymorphisms of which are due to the variable number of repeats of four nucleotide sequences (see Table 1). All samples were successfully amplified for eight STR loci using the protocol supplied in the GenePrint kit (Promega Corporation) by combining one multiplex reaction with five single-locus PCR. Still, some of these samples failed to be amplified in the first attempt but they were successfully amplified in subsequent attempts. There was no evidence of contamination throughout the whole study. All loci provided robust and easily interpretable signals. Amplification of some STR loci produced PCR

amplification artifacts that could possibly be generated by repeat slippage, primer dimer formation or addition of a terminal nucleotide to the PCR product at less than 100% efficiency (and so on). Some amplification reactions also produced the high molecular weight fragments of undefined origin and these artifacts were undesirable.[23-24,27-28,38] Despite these artifacts being generated, the authentic product yields are strong enough to allow allele determination. Dilution of the amplified products from reactions helps decrease the smearing of silver deposition and minor band components.

For a multiplex reaction of three loci, CSF1PO-TPOX-TH01, the respective size of amplification products is used to group the loci and their alleles that can be coamplified and resolved by sequencing gel separation. The loci for coamplification are chosen based on the size ranges of the alleles, the primer sequence and the primer annealing temperature.[17,19,31-32] The multiplex PCR is known to be vulnerable for two main reasons: dimerisation between primers of different loci and preferential amplification of the smaller alleles within the locus resulting in an 'allelic drop-out' phenomenon.[23,26,40] In order to obtain amplification products with equal band intensity on silver stained gels the optimization of the primer concentration was especially necessary.[23,40] In the present study, all three loci in the multiplex amplified well while generating very few artifacts. As to the intensity of allele bands for this multiplex detected by silver staining, the TH01 locus had the highest intensity and the CSF1PO locus had the lowest (Fig. 8). There was no evidence of preferential amplification within the locus, the heterozygous samples of these three loci had equal band intensity between two alleles in each DNA sample. Hence, it does not cause any confusion in result interpretations. However, it can be anticipated that in some highly degraded samples, the smallest-size

locus, i.e., TH01, may be typeable while the larger loci in the triplex may not be typeable.[56] Co-amplification of multiple loci effectively reduces labour and cost thus yielding more information in a shorter time period.

PCR amplification of the vWA locus was free from artifactual bands, but showed a phenomenon known as repeat slippage of the *Taq* enzyme during the amplification process. This was characterized by the occasional loss of a 4-base or a repeat unit resulting in the display of extra bands below the authentic alleles (Fig. 9).[37,57-58] The presence of repeat slippage products is sometimes referred to as stutter bands. They are more strongly correlated with particular loci than with primer selection or amplification protocols.[23-24,38,57,59] The intensity of the stutter bands is low when compared with the true allele products.[26-27,37] This phenomenon does not interfere with the interpretation of vWA profiles obtained from samples of known origin. Theoretically, the existence of the stutter band below each vWA allele could complicate the interpretation of mixed samples.[37,56-57]

PCR amplification of the F13A01 locus was free from artifactual bands (Fig. 10), but single-locus amplifications of FESFPS, F13B and LPL sometimes generated faint bands of high molecular weight fragments and ladderring artifacts (Fig. 11-12). All of these artifacts generated in the positive control K562 DNA amplification reactions are less intense than those generated in other samples. This probably occurs when too much DNA template is used.[57-58,60] However, the major alleles are strong enough to allow allele determination.

Electrophoresis

PCR products were checked for successful amplification on 2% agarose gels (Fig. 7). Subsequently, they were loaded on a high resolution

polyacrylamide gel. Polyacrylamide gel electrophoresis followed by silver staining is more sensitive than agarose at detecting low yields of the amplicons.[39,46,51,61] It happens that even though some amplicons are not visualized on agarose gel, they can be seen on the polyacrylamide gel. But if excess amplicons (indicated by very intense bands on agarose gel) are loaded on the polyacrylamide gels, the resulting bands will be too broad. And this may lead to typing errors. The dilution of amplified products will decrease the broad bands and give better resolution. The results of sequential loading of several amplicons corresponding to the same loci are shown in Fig. 9-10 and the profile of simultaneous loading of several amplification products corresponding to different loci is depicted in fig. 8 and fig. 12. Both typing approaches minimize labor, materials and analysis time.

Polyacrylamide gel electrophoresis with higher voltage gradients produced faster runs, but led to the gel becoming heated during electrophoresis. Furthermore, it can distort the bands. It is better to run the gel more slowly to allow unambiguous genotyping. The electrophoretic system used in this study can resolve fragments that differ by one base pair thus allowing the detection of microvariant alleles such as TH01 allele 9.3 and F13A01 allele 3.2 as shown in figures 8 and 10.

Silver staining

The detection of denatured amplified products using silver stain sometimes reveals two distinct fragments. The locus FESFPS sometimes displays two bands per allele which appear as closely spaced doublet bands (Fig. 11). Other loci display single visible band per allele but this event does not confuse the typing. STR amplicons visible as separate double bands represent the complementary strands of the PCR products of the same length, yet migrated differentially due to their different nucleotide sequences.[26,28,56-57,62]

On denaturing polyacrylamide gel electrophoresis, the strands containing cytosine and adenine (CA) migrate faster than the guanine and thymine (GT) strands.[26,62] In contrast, only one of the two strands can be observed by fluorescence detection.[38,56] The relative mobilities of the strands are influenced by the concentration of the polyacrylamide gel.[28] The space between two complementary strands can be decreased by decreasing gel concentration or running the gel for a shorter time.

Allele determination

Alleles were determined by comparison with the ladder run on the same electrophoretic gel (Fig. 8-12). The electrophoresis system in this study allowed unambiguous allele designation because of the good resolution between consecutive alleles. All the observed alleles that differed in size by one repeat unit (i.e. 4 basepairs), except for the TH01 allele 9.3 (Fig. 8) and for the F13A01 allele 3.2 (Fig. 10) could be determined. The TH01 allele 9.3 has a single base deletion of a thymidine residue in the 5th of 10 TCAT repeats and the F13A01 allele 3.2 has 2 basepair deletions outsite the repeat region.[37] These microvariant alleles should be included in the allele ladders to provide precision and accuracy in STR typing; example include TH01 allele 9.3 (Fig. 8).[48]

Allele ladders should be used for all STR systems detected by manual electrophoretic systems and should be based on the predominant simple repeat motif of the system in question. All commonly occurring alleles should be present in the ladder.[48,63-64] An allele nomenclature should be based on the number of repeats in accordance with the recommendations of the DNA Commission of the International Society of Forensic Haemogenetics (ISFH).[63] Ideally, the ladder should be placed at regular intervals across the gel, so that every unknown sample was adjacent to a control.[51] This would ensure that

clearly defined points of reference were obtained for the bands in the unknown samples, which either matched a band in the ladder or migrated to a position between two ladder bands.[51] Since the components of the allelic ladder and the sample fragments have the same length and the same sequence, they will migrate the same distance during electrophoresis regardless of the types of environmental parameters changed.[48] Examples of such parameters include the percentage of acrylamide monomers or the cross-linkers in the gel, the ionic strength of the buffer, the electrophoresis voltage, or the coupling of a fluorescent molecule to the sample for detection purposes.[48] Consequently, different laboratories using different or same separation techniques and different detection formats can compare their results with precision and reliability in accordance with the use of same specific-locus allele ladders.[13]

Statistical analysis

The 200 samples of unrelated individuals were studied for 8 STR loci located on different chromosomes. Genotyping was performed by comparison of sample fragments with those of the locus-specific allele ladders.

All eight STR loci were found polymorphic. A total of 10 alleles for F13A01 and vWA, 8 alleles for CSF1PO, 7 alleles for FESFPS, 6 alleles for TH01 and LPL, and 5 alleles for TPOX and F13B could be observed (Table 4-11).

All alleles are different in size by one repeat unit (i.e. 4 basepairs) for all loci, except for the TH01 allele 9.3 and F13A01 allele 3.2.[36-38,66-67] As shown in figure 8 and figure 10, the differences between the TH01 allele 9.3 and 10, and between the F13A01 allele 3.2 and 4 become clearly visible by this electrophoretic system. The 9.3 allele is 1 basepair smaller in size than the 10th allele. The 3.2 allele is 2 basepairs smaller in size than the 4th allele.

For F13A01 (Fig. 10), another allele distinctly longer than allele 16 could be observed by visual comparison with the allele ladder. It was observed in only 1 out of 400 alleles. It could be assumed that it was allele 17 and should be confirmed by DNA sequencing later.

In Figure 9, it could be observed that 2 alleles of the vWA locus migrated more slowly than allele 20. Only 1 of such slow moving alleles was found among 400 alleles studied. From their migration length, it was assumed that this could be allele 21 and allele 22.[66] But this should be further confirmed by DNA sequencing.

The distribution of the observed allele frequencies, the genotypes of each STR locus and the summation of chi-square values from observed and expected numbers under the assumption of the Hardy-Weinberg equilibrium are shown in Table 4-11. The genotypes with the expected number equal or below 5 were pooled together into one group.[65] No significant deviation from the Hardy-Weinberg equilibrium ($p>0.05$) was found for all the systems studied.

All statistical values are shown in Table 12. The unbiased estimate of the expected heterozygosity varied between loci, ranging from 0.8076 for vWA to 0.4971 for LPL.

The vWA, TH01, CSF1PO, F13A01, FESFPS are highly informative (PIC > 0.5) and the TPOX, LPL, F13B are moderately informative (0.5 > PIC > 0.25). The most informative locus is vWA. The least informative locus is F13B.

Evaluation for forensic purposes was carried out by calculating the discrimination power, the probability of matching for identity testing, and the power of exclusion, the paternity index for paternity testing.

The average power of exclusion which measures the ability of a genetic test to exclude a falsely accused individual was calculated for each

locus. The PE_{trio} ranged from 0.6361 for vWA to 0.2627 for F13B. Generally, the PE_{duo} is below the PE_{trio} for all loci (Table 12). The combined PE_{trio} for the CSF1PO-TPOX-TH01 (CTT) and F13A01-FESFPS-vWA (FFv) amounts to 0.8364 and 0.8985, respectively. The combined PE_{trio} for the six loci (CTTFFv) and all eight loci is 0.9834 and 0.9911.

The combined $PI_{typtcal}$ for the CSF1PO-TPOX-TH01 and F13A01-FESFPS-vWA amounts to 3.8249 and 7.4159, respectively. The combined $PI_{typtcal}$ for the six loci (CTTFFv) and all eight loci is 28.3649 and 28.3979. The paternity index reflects how many more times likely it is that the person being tested is the biological father, rather than a randomly selected individual. The typical paternity index is assigned to a locus rather than an individual case.

The combined probability of matching (Pm) values calculated for the CSF1PO-TPOX-TH01 and F13A01-FESFPS-vWA is 3.5913×10^{-3} and 1.2717×10^{-3}. The chance that two individuals match at six loci and all eight loci is 4.5669×10^{-6} and 4.3081×10^{-7}, respectively.

The combined discrimination power (DP) values calculated for the CSF1PO-TPOX-TH01 and F13A01-FESFPS-vWA are 0.9964 and 0.9987. The chance that two individuals chosen at random from the Thai population will have different phenotypes at six loci and all eight loci are 0.999995433 and 0.999999569.

The distribution of alleles of the eight STR loci in the Thai population was either different ($p<0.05$) or similar ($p>0.05$) to those of Caucasians[35], Spanish[37]and Taiwanese Chinese[67] when analyzed by chi-square test of the Statistics/Data Analysis Program (Stata 5.0). This program is supported by the Clinical Epidemiology Unit (CEU.), Faculty of Medicine, Chulalongkorn University. The results are shown in Table 13.

The heterozygosity of each STR locus of the Thai population when compared with other populations is shown in Figure 13.

In Figure 13, it can be seen that the TPOX, F13B and LPL loci did not appear as highly polymorphic as observed in other populations. [35,38]

From this study, the test for six loci - CSF1PO-TPOX-TH01- F13A01-FESFPS-vWA has proven useful for identity testing and paternity test. In the future, if TPOX is replaced by the new polymorphic STR locus possessing the PE of above 0.6, the PE of this battery of tests will reach approximately 0.999999 thus gaining more statistical power and hence a more powerful tool will be obtained. Better statistical indices render the test more powerful.

Nowadays, the commercial multiplex reaction kits such as CSF1PO-TPOX-TH01 and F13A01-FESFPS-vWA for silver stain detection have been marketed. They are rapid, convenient, effective and have a strong statistical index power when used in combination.

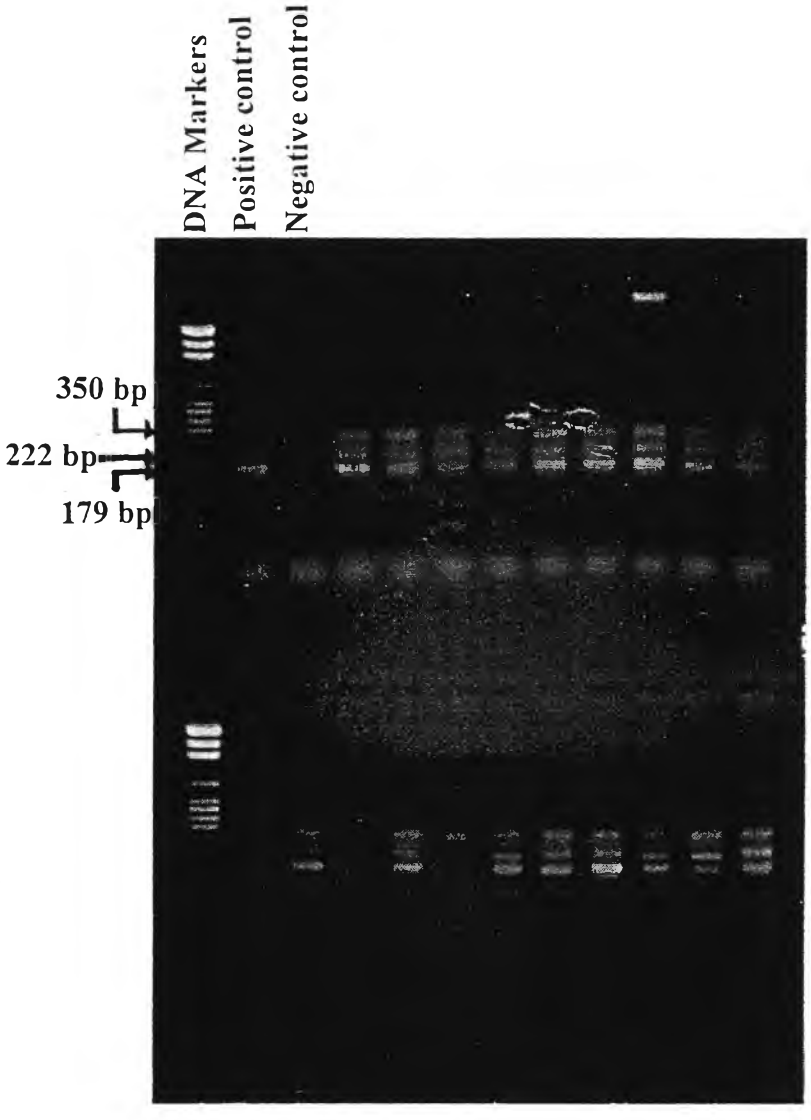


Fig. 7. Agarose gel electrophoresis for checking the PCR products after the amplification of the CSF1PO, TPOX, TH01 multiplex system.

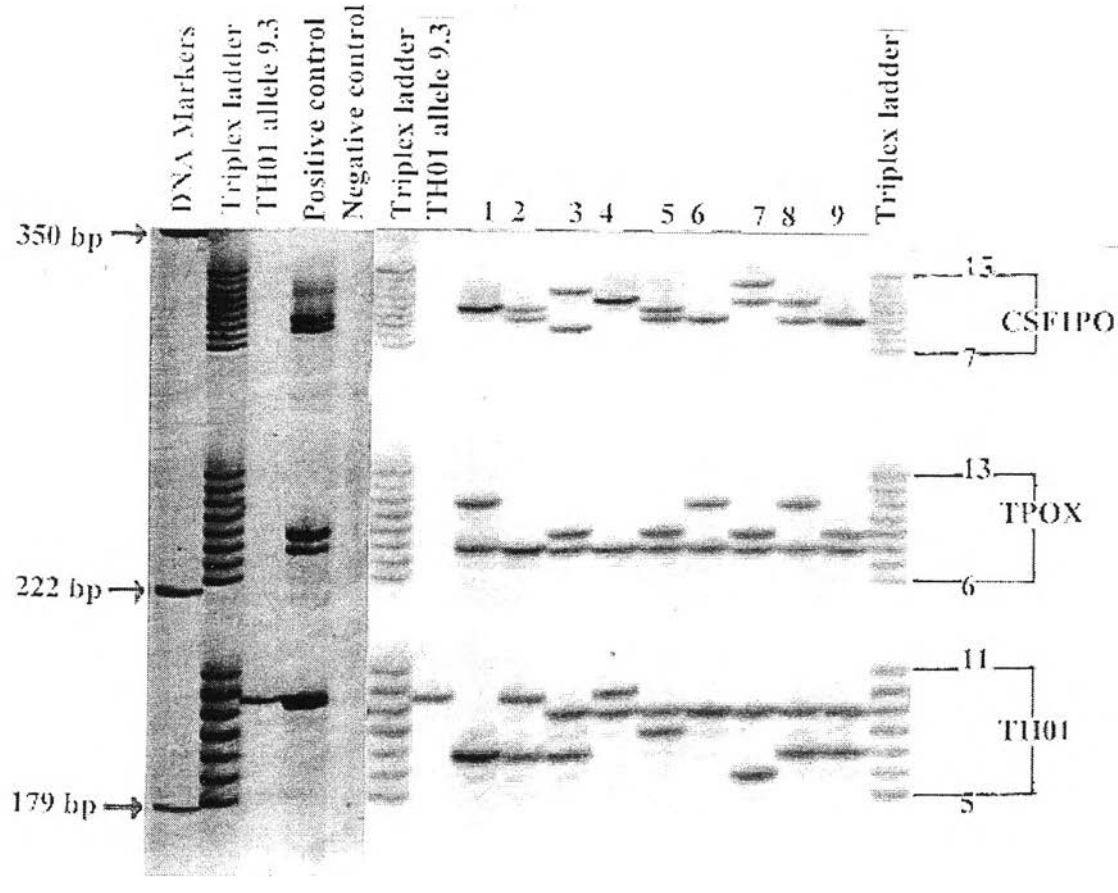
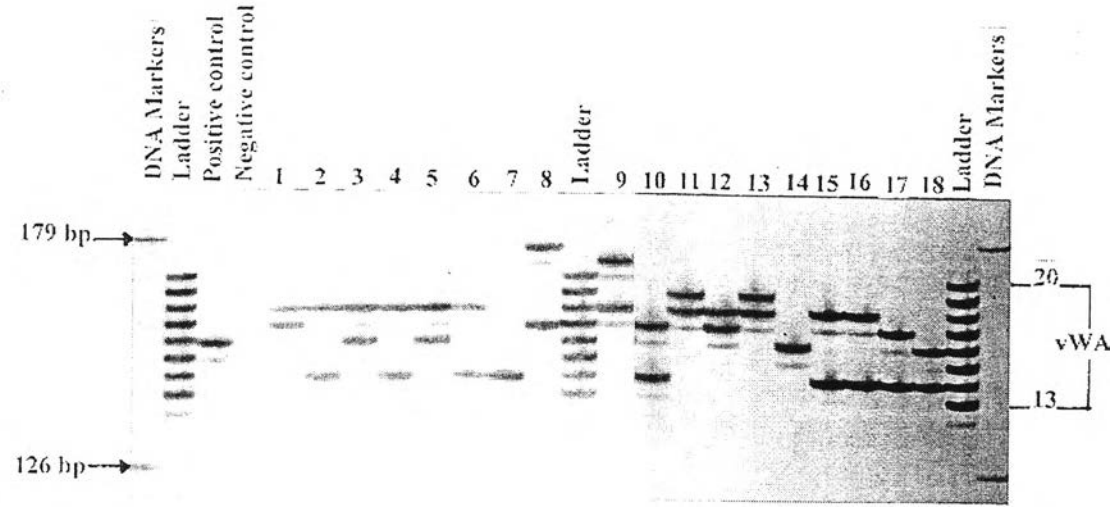


Fig. 8. Analysis of different STR loci obtained from a multiplex PCR reaction at CSF1PO, TPOX, TH01 loci after loading on denaturing polyacrylamide gel and subsequent detection by silver stain. Lane 2 represents genotype 7,9.3 for TH01 locus.

A.



B.

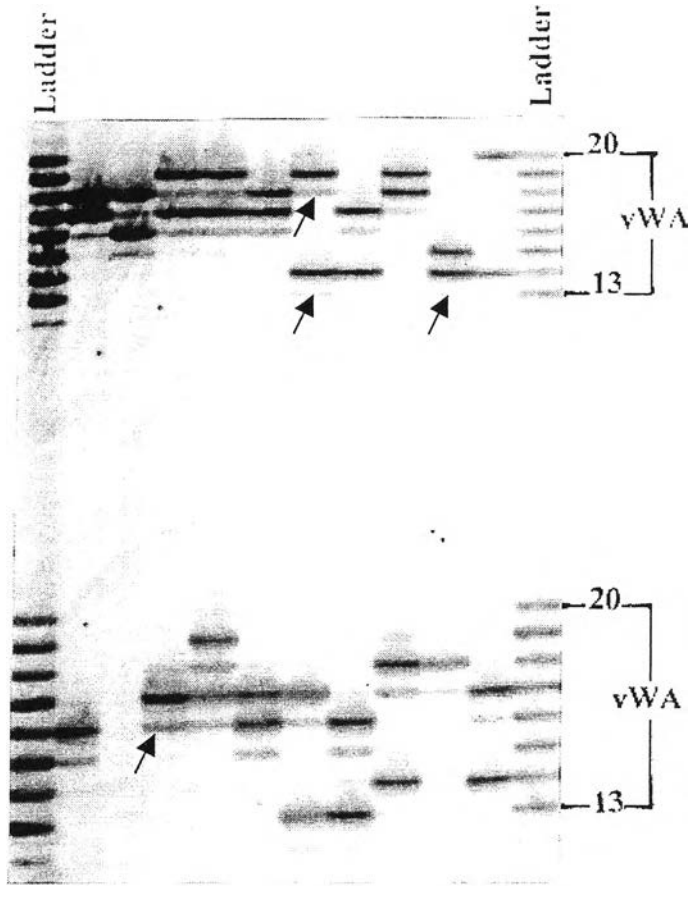


Fig. 9. A. The STR profile of vWA locus after a single-locus PCR reaction. Lane 8 presents genotype 17,22. Lane 9 presents genotype 18,21. B. Sequencial loading of the PCR products on denaturing polyacrylamide gel and subsequent detection by silver staining. The arrows indicate stutter bands.

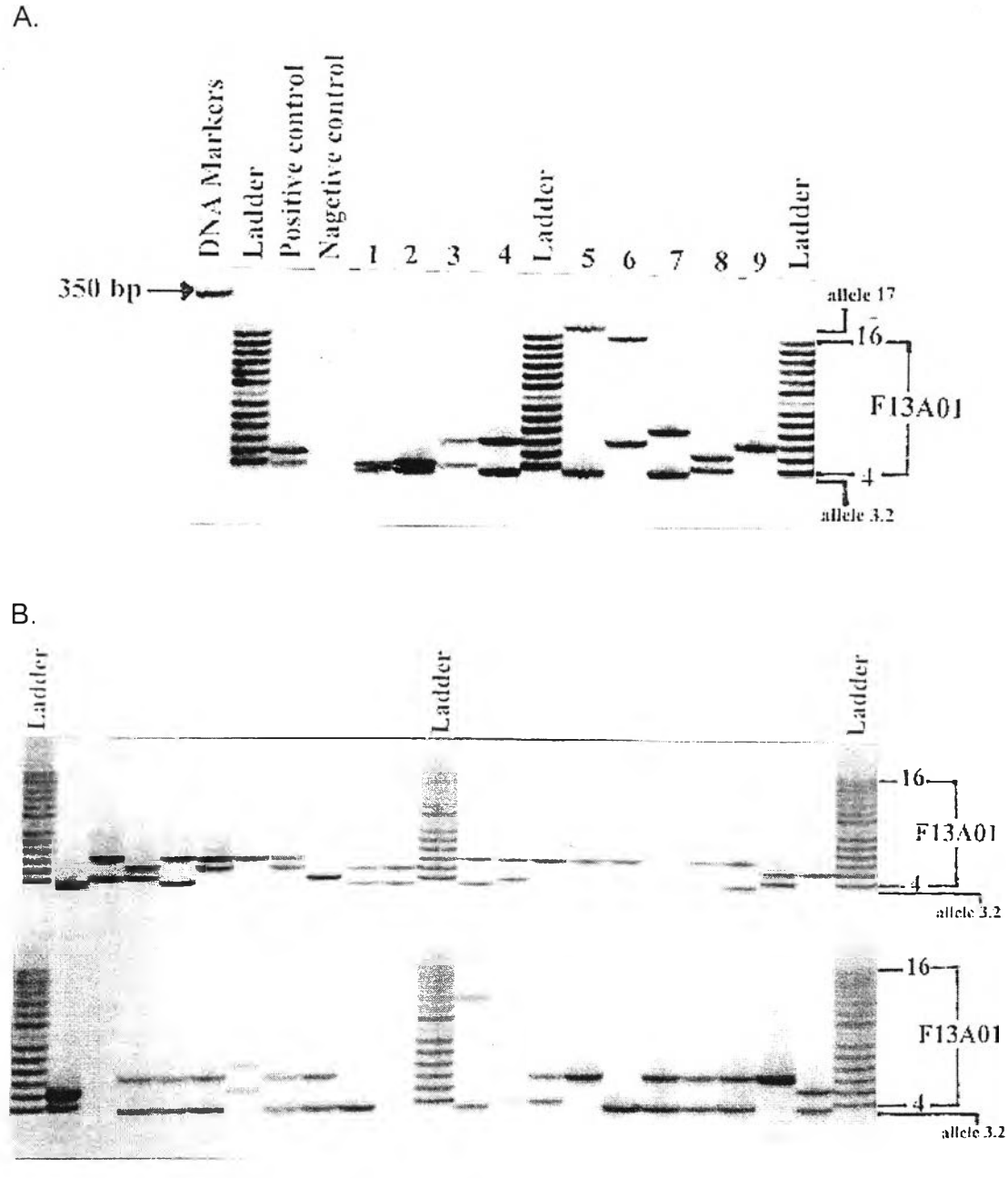


Fig. 10. A. F13A01 genotypes obtained from polyacrylamide gel electrophoresis. Lane 5 presents genotype 3.2,17. B. Sequencial loading of F13A01.

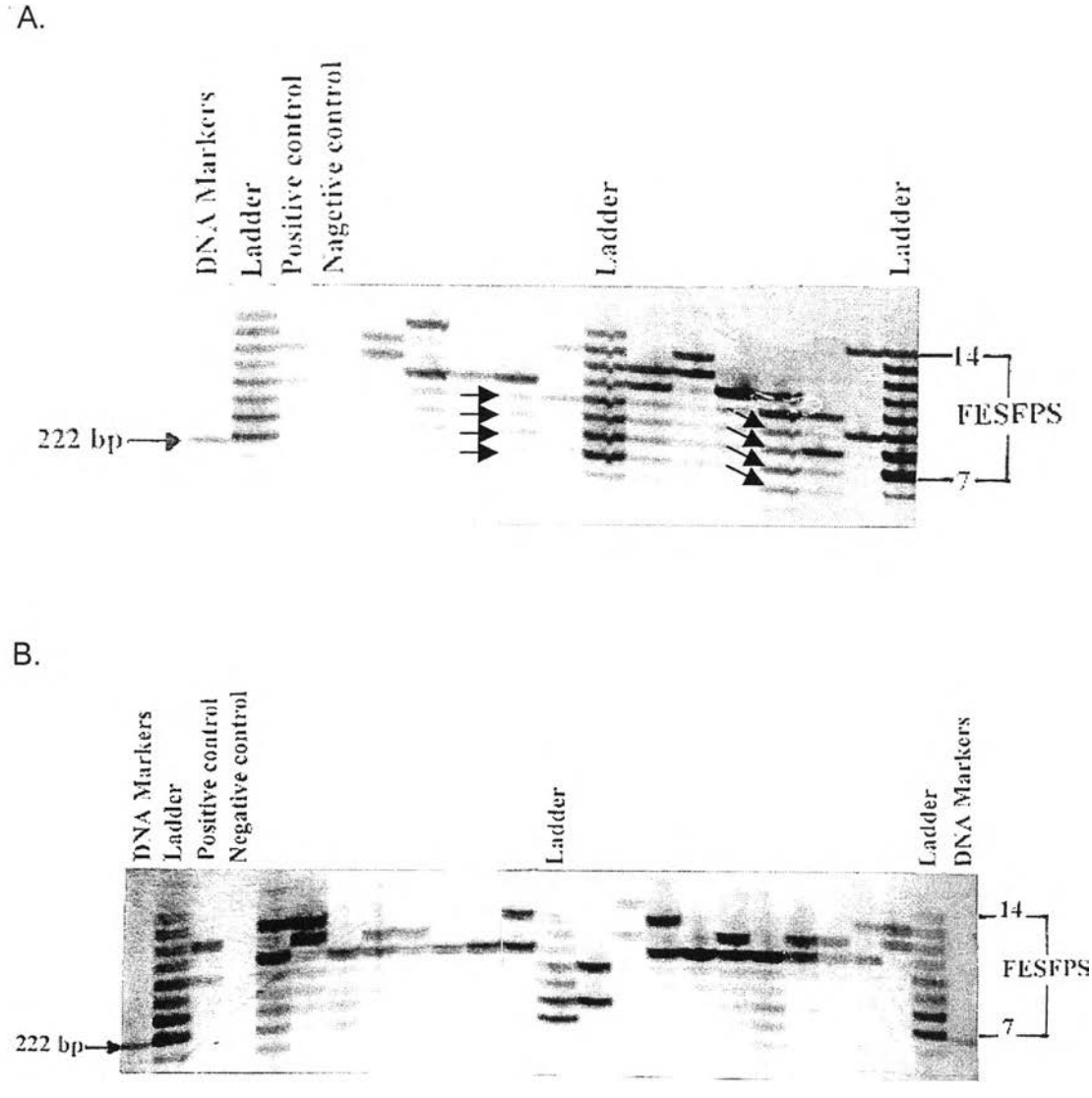


Fig. 11. FESFPS genotypes obtained with polyacrylamide gel electrophoresis. A. The arrows indicate examples of ladderring artifacts. B. Represents closely spaced duplet bands that can be visible as two bands per allele.

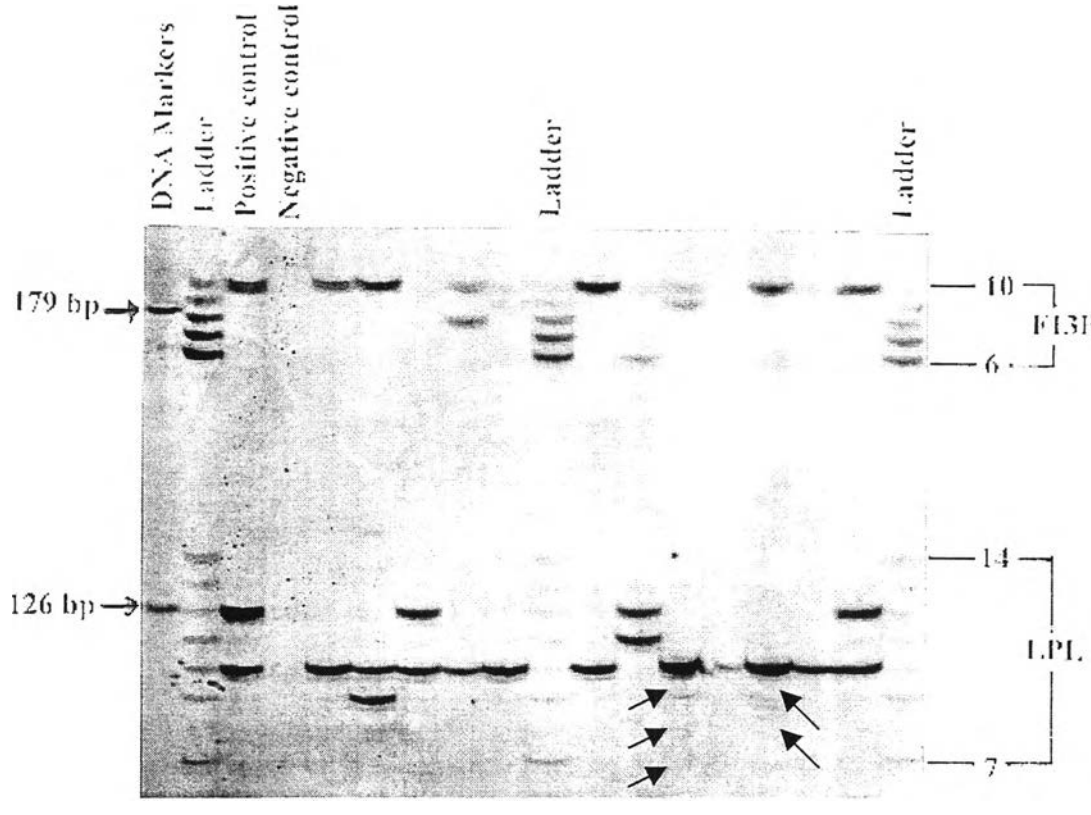


Fig. 12. Analysis of F13B and LPL amplicons obtained from separate single PCR and simultaneous loading on polyacrylamide sequencing gel. The arrows indicate ladderring artifacts.

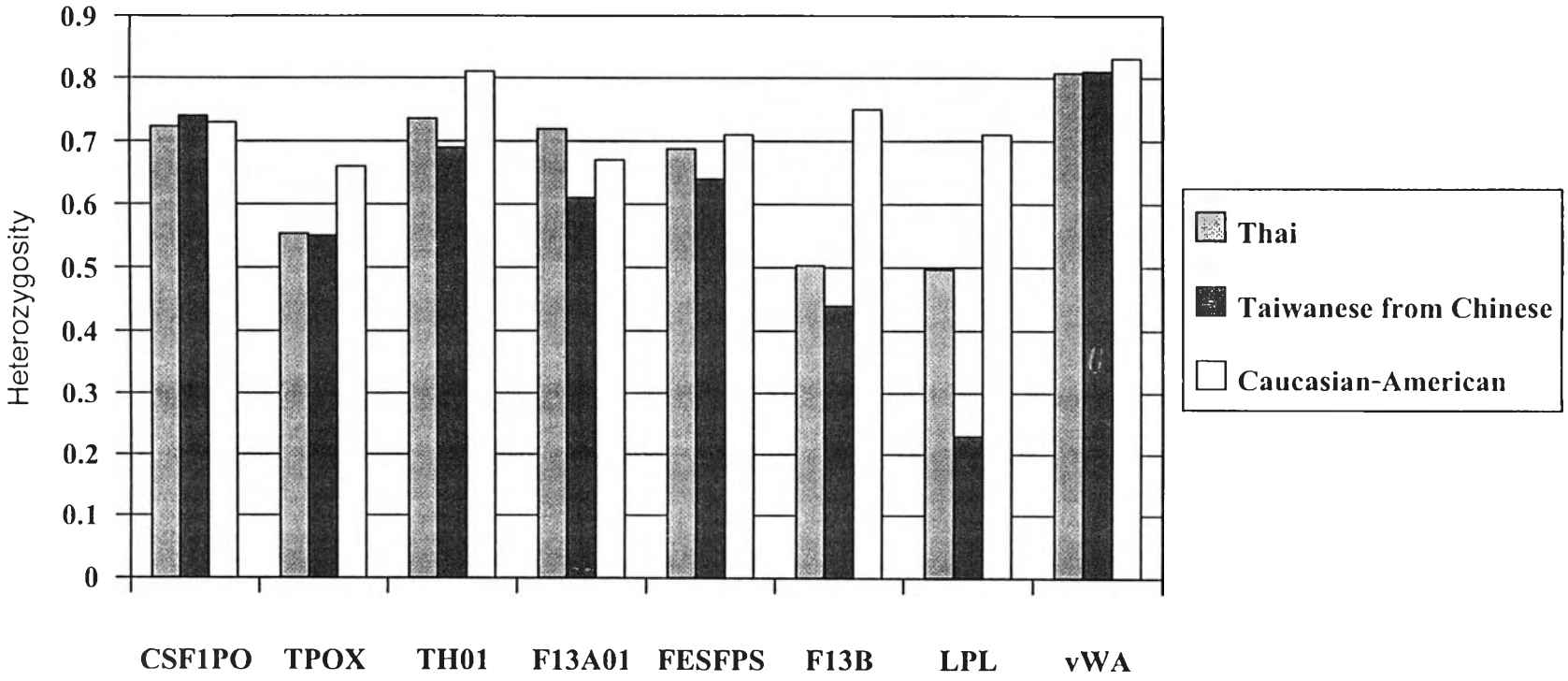


Fig. 13. The expected heterozygosity of each STR locus of the Thai population compared with a Taiwanese-Chinese[67] and a Caucasian-American population[38].

Table 4. CSF1PO genotypes and allele frequencies.

Genotypes	Observed number
10,10	10
10,11	21
10,12	37
10,13	3
10,14	1
11,11	15
11,12	43
11,13	9
11,14	2
11,15	1
12,12	29
12,13	8
12,14	6
13,13	1
8,10	1
8,11	1
9,10	4
9,11	2
9,12	4
9,13	2
others	0
total	200

$\sum\chi^2$= 6.6062; df 9 (class pooled), p> 0.05 (not significant)

Allele frequencies; *8* 0.0050; *9* 0.0030; *10* 0.2175; *11* 0.2725; *12* 0.3900; *13* 0.0060; *14* 0.0225; *15* 0.0025.

Table 5. TPOX genotypes and allele frequencies.

Genotypes	Observed number
10,11	2
11,11	22
11,12	1
12,12	1
8,10	7
8,11	53
8,8	80
8,9	21
9,11	10
9,9	3
others	0
total	200

$\sum\chi^2$= 7.2757; df 5 (class pooled), p> 0.05 (not significant)

Allele frequencies; *8* 0.6025; *9* 0.0925; *10* 0.0225; *11* 0.2750; *12* 0.0075.

Table 6. TH01 genotypes and allele frequencies.

Genotypes	Observed number
10,10	3
6,10	1
6,6	5
6,7	8
6,8	2
6,9	14
6,9.3	6
7,10	5
7,7	11
7,8	12
7,9	51
7,9.3	15
8,10	3
8,8	1
8,9	9
8,9.3	1
9,10	7
9,9	34
9,9.3	11
9.3,9.3	1
others	0
total	200

$\sum\chi^2$= 11.5336; df 10 (class pooled), p> 0.05 (not significant)

Allele frequencies; *6* 0.1025; *7* 0.2825; *8* 0.0725; *9* 0.4000; *9.3* 0.0875; *10* 0.0550.

Table 7. F13A01 genotypes and allele frequencies.

Genotypes	Observed number
3.2,13	1
3.2,15	1
3.2,17	1
3.2,3.2	18
3.2,4	15
3.2,5	16
3.2,6	50
3.2,7	5
3.2,8	1
4,4	3
4,5	10
4,6	17
4,7	2
4,8	1
5,5	4
5,6	15
5,7	2
6,15	1
6,16	1
6,6	35
6,7	1
others	0
total	200

$\sum\chi^2$= 2.6954; df 7 (class pooled), p> 0.05 (not significant)

Allele frequencies; ***3.2**** 0.3150; *4* 0.1275; *5* 0.1275; *6* 0.3875; *7* 0.0250; *8* 0.0050; *13* 0.0025; *15* 0.0050; *16* 0.0025; *17** 0.0025.

* The allele ladders do not have allele 3.2 and 17.

Table 8. FESFPS genotypes and allele frequencies.

Genotypes	Observed number
10,10	2
10,11	6
10,12	9
10,13	5
11,11	39
11,12	54
11,13	29
11,14	5
12,12	16
12,13	25
12,14	2
13,13	6
8,10	1
9,14	1
others	0
total	200

$\sum\chi^2$= 2.7789; df 7 (class pooled), p> 0.05 (not significant)

Allele frequencies; *8* 0.0025; *9* 0.0025; *10* 0.0625; *11* 0.4300; *12* 0.3050; *13* 0.1755; *14* 0.0200.

Table 9. F13B genotypes and allele frequencies.

Genotypes	Observed number
10,10	92
11,11	1
6,10	3
8,10	18
8,8	3
8,9	6
9,10	56
9,9	20
6,9	1
others	0
total	200

$\sum\chi^2$= 3.5895; df 4 (class pooled), p> 0.05 (not significant)

Allele frequencies; *6* 0.1000; *8* 0.0750; *9* 0.2575; *10* 0.6525; *11* 0.0050.

Table 10. LPL genotypes and allele frequencies.

Genotypes	Observed number
10,10	87
10,11	25
10,12	62
10,13	4
11,12	7
12,12	7
12,13	4
7,10	1
9,10	2
9,12	1
others	0
total	200

$\sum\chi^2$= 1.3282; df 4 (class pooled), p> 0.05 (not significant)

Allele frequencies; *7* 0.0025; *9* 0.0075; *10* 0.6700; *11* 0.0800; *12* 0.2200; *13* 0.0200.

Table 11. vWA genotypes and allele frequencies.

Genotypes	Observed number
13,16	1
13,17	2
14,14	9
14,15	5
14,16	10
14,17	20
14,18	25
14,19	6
14,20	4
15,17	3
15,18	7
16,16	3
16,17	18
16,18	10
16,19	7
16,20	3
17,17	11
17,18	24
17,19	16
17,20	1
17,22	1
18,18	5
18,19	8
18,21	1
others	0
total	200

$\sum\chi^2$= 18.2305; df 12 (class pooled), p> 0.05 (not significant) Allele frequencies; *13* 0.0075; *14* 0.2200; *15* 0.0375; *16* 0.1375; *17* 0.2675; *18* 0.2125; *19* 0.0925; *20* 0.0200; *21** 0.0025; *22** 0.0025.

* The allele ladders do not have allele 21 and 22.

Table 12. Statistics for Forensic identification and Parentage studies.

	Obs. h	*Exp. h*	PIC	DP	Pm	PE_{trio}	PE_{duo}	$PI_{typical}$
CSF1PO	0.7250	0.7231	0.6913	0.8756	0.1244	0.5005	0.2025	1.8057
TPOX	0.4700	0.5536	0.4960	0.7366	0.2634	0.3054	0.1335	1.1201
TH01	0.7250	0.7356	0.7063	0.8904	0.1096	0.5285	0.1922	1.8911
F13B	0.4200	0.5035	0.4450	0.6901	0.3099	0.2627	0.1155	1.0070
LPL	0.5300	0.4971	0.4518	0.6955	0.3044	0.2737	0.1097	0.9942
FESFPS	0.6850	0.6880	0.6457	0.8470	0.1530	0.4475	0.1926	1.6026
F13A01	0.7000	0.7192	0.6838	0.8732	0.1267	0.4950	0.1978	1.7806
VWA	0.8600	0.8076	0.7922	0.9343	0.0656	0.6361	0.2264	2.5988
CTT[a]	-	-	-	0.9964	3.5913×10^{-3}	0.8364	0.4418	3.8249
FFv[b]	-	-	-	0.9987	1.2717×10^{-3}	0.8985	0.4989	7.4159
CTTFFv[c]	-	-	-	0.9999 95433	4.5669×10^{-6}	0.9834	0.7203	28.3649
8 loci[d]	-	-	-	0.9999 99569	4.3081×10^{-7}	0.9911	0.7797	28.3979

a. combined CSF1PO-TPOX-TH01.

b. combined F13A01-FESFPS-vWA.

c. combined CSF1PO-TPOX-TH01- F13A01-FESFPS-vWA.

d. all 8 loci combined.

Obs. h = Observed homozygosity, *Exp. h* = Unbiased estimate of the expected heterozygosity, PIC = Polymorphism information content

DP = Discrimination power, Pm = Probability of matching

PE = Power of exclusion, PI = Paternity Index.

Table 13. The different of allele distributions between populations are shown by p value.

	Thai-Chinese[a, *]	Thai-Spanish[b, *]
CSF1PO	0.561	0.171
TPOX	0.511	0.042
TH01	0.004	0.000
F13A01	0.959	0.000
FESFPS	0.032	0.000
F13B	0.294	0.000
LPL	0.000	-
VWA	0.725	0.000

a Allele frequency distributions of the Chinese population from Taiwan.[67]

b Allele frequency distributions of the Spanish population.[37]

* There was significant difference when p value < 0.05.