CHAPTER 5

RESULTS

5.1 Self-report of demographic characteristic data

In this study, the polymorphisms of *CYP2A13* gene was restricted to the subjects who were Thai nationality and born in the Northern region of Thailand. A total of 336 unrelated healthy blood donor volunteers (175 males and 161 females) were participated in this study. Demographic characteristics of all subjects are shown in Table 5. The mean ages at interview of male and female were not significantly different (years 34.07 $\pm$ 10.30 versus 33.10 $\pm$ 8.55 years). No subjects who smoke were included in this study.

5.2 Identification of SNPs in the *CYP2A13* Gene and Determination of Allele Frequency.

In this study, a two-step PCR method was developed to detect the presence of the six variant alleles of the CYP2A13 gene. First, a full-length of the *CYP2A13* gene was amplified from each of 336 genomic DNA samples (175 males and 161 females). The amplified fragment spans 9,456 bp (Figure 13) starting at -1,117 bp before the ATG start codon and ending at 869 bp after the TGA stop codon as described by Zhang (2003). The sequence regions cover 5'-untranslated, all nine exon, eight intron, exon-intron junction and 3'-untranslated region of *CYP2A13*.

Table 5 Subject Characteristics

Characteristics	**No. of subject**
Total	336
Gender	
Male (%)	175 (52.1)
Female (%)	161 (47.9)
Age (years)	
20 – 29 (%)	128 (38.1)
30 – 39 (%)	118 (35.1)
40 – 49 (%)	68 (20.2)
≥ 50 (%)	22 (6.5)
Mean age (years)	33.57±9.52
Education	
Primary school (%)	110 (32.7)
Secondary school (%)	90 (26.8)
Diploma (%)	49 (14.6)
University (%)	87 (25.9)
Occupation	
Employee (%)	112 (33.3)
Agriculturist (%)	50 (14.9)
Commerce (%)	32 (9.5)
Government officer (%)	52 (15.5)
Student (%)	42 (12.5)
Officer (%)	39 (11.6)
Others* (%)	9 (2.7)

*Others represented housewife and priests.

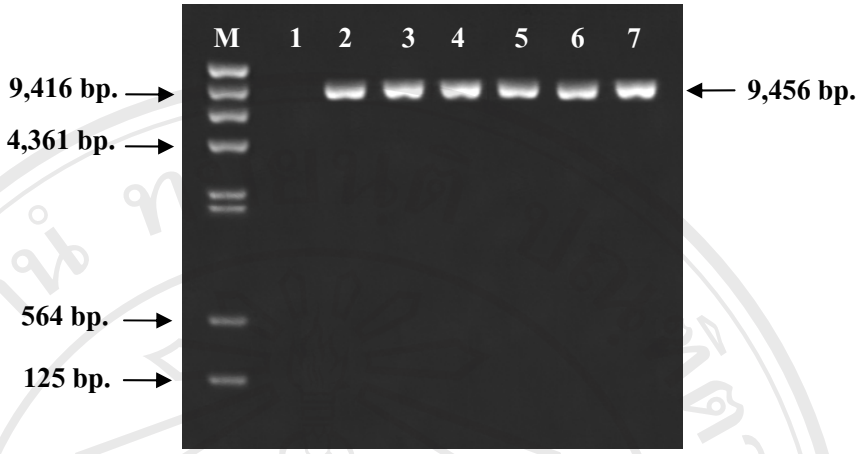


Figure 13 Representative agarose gel electrophoresis of amplified *CYP2A13* gene by long-distance PCR. Lane M: Lambda DNA/*Hind*III Marker; Lane 1: a negative control in which on genomic DNA template was added, Lanes 2-7: PCR products from genomic DNA samples showed a band at 9,456 bp. The position and size of the expected PCR products are indicated (←).

PCR-RFLP assay was employed for genotyping of the 1662G>C, 3375C>T, and 7520C>G except for the 74G>A variant allele which used the mismatch PCR-RFLP assay. The 578C>T and 7571G>C variant allele were analyzed using Mismatch allele-specific PCR and Tetra-primer PCR assay, respectively.

5.2.1 Genotyping of the 74G>A in exon 1 of *CYP2A13* gene

Genotyping of the 74G>A were performed by mismatch PCR amplification coupled with RFLP analysis (mismatch PCR-RFLP) with the forward primer Ex1F and reverse primer Ex1RFLP. PCR amplification produced the expected size of the

247 bp PCR product (Figure 14A). Then, the product was digested with *Hha*I restriction enzymes. A *Hha*I restriction site was created by introducing a C-G change at the 3'penultimate position of the reverse primers Ex1RFLP. The 74G wild-type allele that carries a *Hha*I restriction site GCGC produced two fragments of 24 and 223 bp, while the 74A mutant-type allele amplicon was uncut (Figure 14B).

5.2.2 Genotyping of the 578C>T in exon 2 of *CYP2A13* gene

Genotyping of the 578C>T were performed by mismatch allele-specific amplification (mASA) PCR assay by using allele-specific primer (Ex2WtF and Ex2MtR) and outer primer (Ex1F and Ex2R). The expected size of ASA products from both wild-type and mutant type were 186 and 898 bp, where 898 bp fragment was used as an internal control. The small fragment 186 bp used to identify 578C wild-type allele or 578T variant allele as shown in Figure 15.

5.2.3. Genotyping of the 1662G>C in exon 3 of *CYP2A13* gene

Genotyping of the 1662G>C was performed by PCR-RFLP based analysis. After amplification with forward primer Ex3F and reverse primer Ex3R, the fragment of 406 bp PCR products (Figure16A) were digested with restriction enzyme *Hae*II. The 1662G wild-type allele that carries a *Hae*II restriction site at one position produced two fragments of 177 and 229 bp, while the 1662C mutant-type allele amplicon that created one pseudo-restriction site produced three fragments of 42, 177, and 187 bp (Figure 16B).

5.2.4 Genotyping of the 3375C>T in exon 5 of *CYP2A13* gene

Genotyping of the 3375C>T was performed by PCR-RFLP based analysis. After amplification with forward primer Ex5F and reverse primer Ex5R, the fragment of 400 bp PCR product (Figure 17A) was then digested with *Hha*I and *Apal*I restriction enzymes. The *Hha*I digest resulted in two fragments for 3375C wild-type allele (99 bp and 301 bp), but does not cut the 3375T mutant-type allele, thus producing only one band (400 bp). In addition, all samples were also digested with *Apal*I, which does not cut the 3375C wild-type allele but cuts the 3375T mutant-type allele to give two fragments (100 and 300 bp) (Figure 17B).

5.2.5 Genotyping of the 7520C>G in 3'Untranslated region of *CYP2A13* gene

Genotyping of the 7520C>G was performed by PCR-RFLP based analysis. After amplification with forward primer UtrF and reverse primer UtrR, the fragment of 576 bp PCR product (Figure 18A) was then digested with *Xma*III restriction enzymes. The 7520C wild-type allele that carries a *Xma*III restriction site at one position produce two fragments (259 and 317 bp), but does not cut the 7520G mutant-type allele, thus producing only one band (576 bp) (Figure 18B).

5.2.6 Genotyping of the 7571G>C in the 3'Untranslated region of *CYP2A13* gene

A tetra-primer PCR assay (Ye et al., 1992) was developed to genotype the 7571G>C variant allele in the 3'Untranslated region of *CYP2A13* gene. Two (outer) primers that are complementary to unique sequences of intron 8 and 3'UTR of *CYP2A13* (UtrF, UtrR) and two (inner) primers that are designed for the allele

specific amplification (ASA) of the wild-type 7571G allele (primer UtrWtF) and the mutant-type 7571C (primer UtrMtR) variant allele were combined in a single tube PCR assay. PCR Amplification of the 3'UTR with primers UtrF and UtrR produced a large 576 bp fragment that served as internal control for the quality of PCR amplification and as template for ASA. The ASA reaction produced the 389 bp PCR product specific for the mutant-allele (7571C) allele and the 223 bp PCR product specific for the wild-type (7571G) allele. Since the 7571G allele and the 7571C allele were analyzed on the sense and anti-sense DNA strands, respectively both amplifications can occur simultaneously in one tube. Based on this technique, two and three PCR bands were detected for homozygote and heterozygote, respectively.

5.3 DNA sequencing

A set of 24 samples representative of each genotype were sequenced to confirm the mutations detected by PCR-based assay and the results were identical to those obtained by PCR-based assay as shown in figure 19-22.

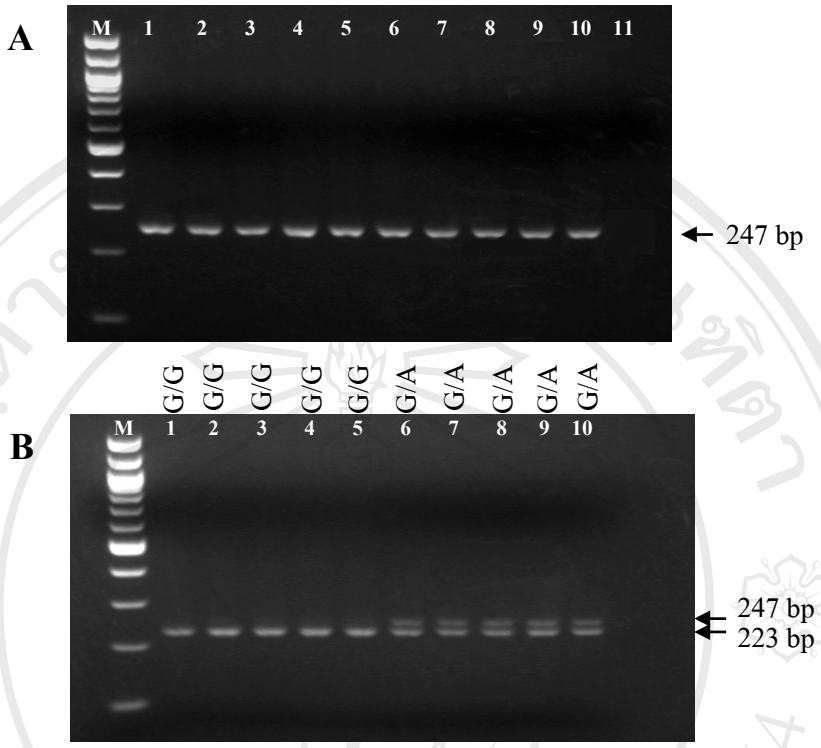


Figure 14 Typical electrophoresis pattern in detection of the 74G>A in exon 1 of CYP2A13 gene. (A) The first PCR amplification analysis. Lane M: 100 bp DNA ladder marker; lanes 1-10: first PCR products (247 bp) amplified from ten DNA samples; Lane 11: a negative control in which on DNA template was added. (B) The mismatch RFLP fragments obtained using *Hha*I restriction enzyme cleavage. Lane M: 100 bp DNA ladder marker; lanes 1-5: homozygous wild-type (G/G) with *Hha*I restriction site results in two bands of 24 (not be seen) and 223 bp; lanes 6-10: heterozygous (G/A) variant show three bands of 24 (not be seen), 223 and 247 bp.

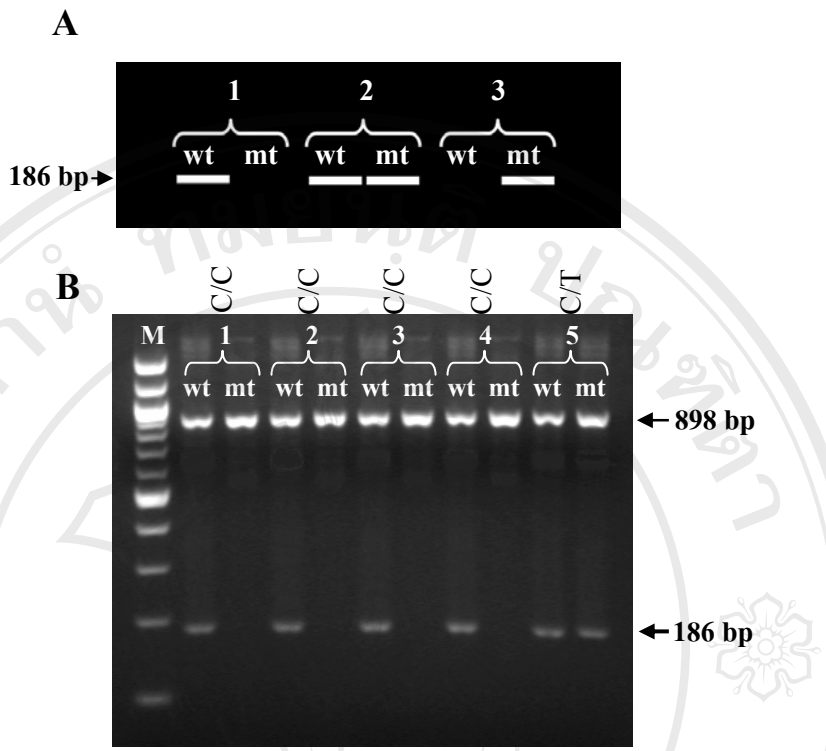


Figure 15 Typical electrophoresis pattern in detection of the 578C>T in exon 2 of *CYP2A13* gene. (A) Schematic mismatched allele-specific PCR amplification (ASA) for detection of the CYP2A13 578C>T polymorphism; wt: PCR product obtained by wild-type specific primer (Ex2WtF); mt: PCR product obtained by mutant-type specific primer (Ex2MtF). Expect sizes of PCR products are 898 bp and 186 bp, but the small fragment 186 bp was used to identify 578C wild-type allele or 578T variant allele. (B) Mismatch ASA analysis for 578C>T. Diagnostic genotypes of 578C>T were considered from appearance of size-specific and number of DNA fragments such as homozygous wild-type (C/C) has one fragment of wild-type, Heterozygous (C/T) has two fragment of both wild-type and mutant-type, and homozygous mutant-type has only one fragment of mutant-type. Lane M: 100 bp DNA ladder; lanes 1-4: genotype with homozygous wide-type (C/C); lane 5: genotype with heterozygous (C/T).

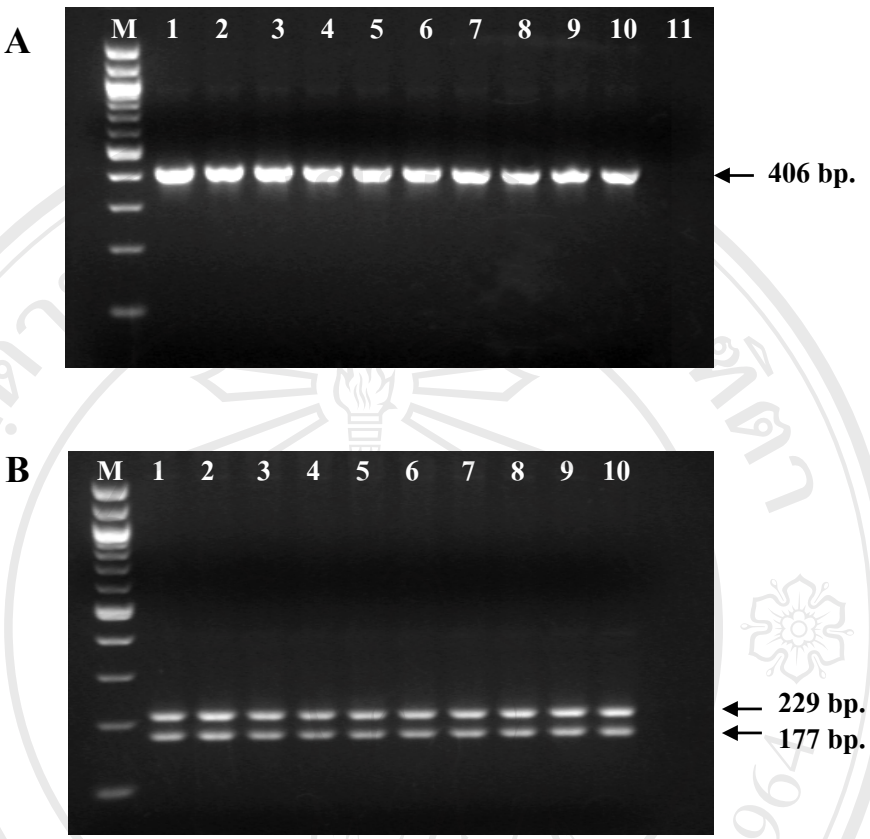


Figure 16 Typical electrophoresis pattern and DNA sequencing in detection of the 1662G>C in exon 3 of *CYP2A13* gene. (A) The first PCR amplification analysis. Lane M: 100 bp DNA ladder marker; lanes 1-10: first PCR products (406 bp) amplified from ten DNA samples; lane 11: a native control in which on DNA template was added. (B) The RFLP fragments obtained using *Hae*II. restriction enzyme cleavage. Lane M: 100 bp DNA ladder marker; lanes 1-10: homozygous wild-type (G/G) with *Hha*I restriction site results in two bands of 177 and 229 bp.

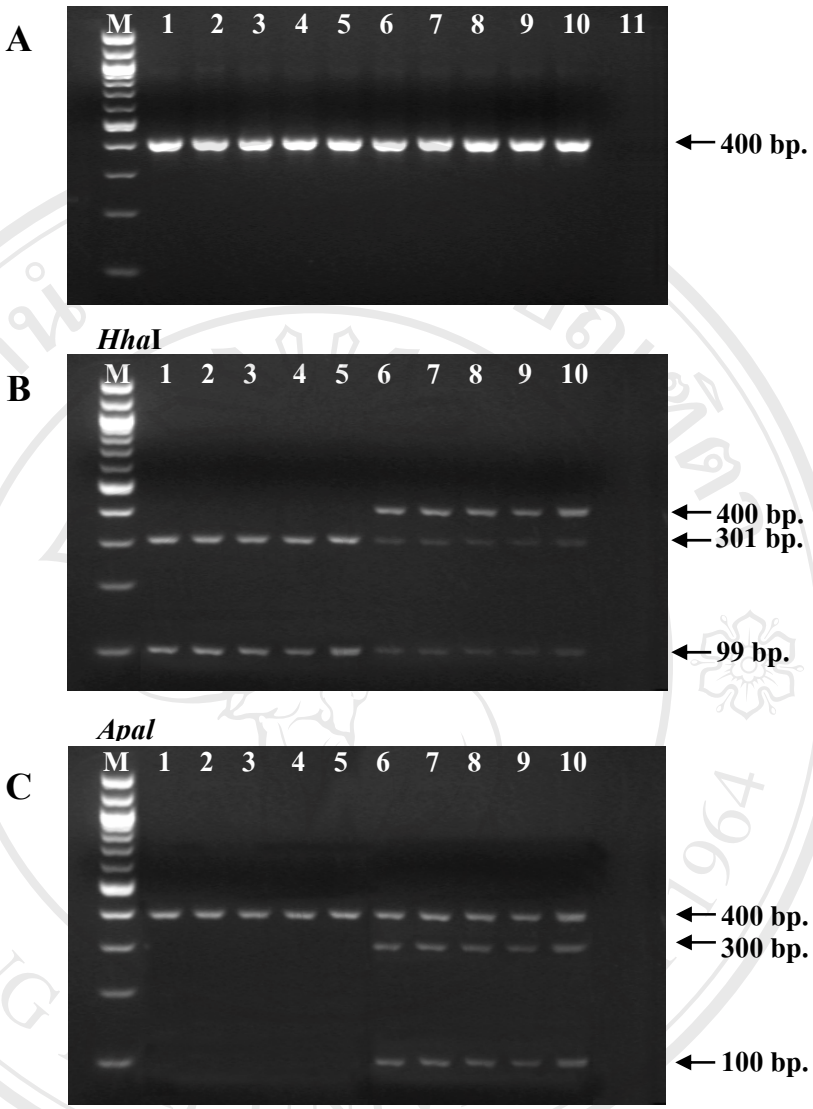


Figure 17 Typical electrophoresis pattern and DNA sequencing in detection of the 3375C>T in exon 5 of *CYP2A13* gene. (A) The first PCR amplification analysis. lanes 1-10: first PCR products (400 bp) amplified from ten DNA samples; lane 11: a native control in which on DNA template was added. (B and C) The RFLP fragments obtained using *Hha*I and *Apal*I restriction enzyme cleavage, respectively. Lanes M: 100 bp DNA ladder marker; lane 1-5: homozygous wild-type (C/C) with *Hha*I restriction site results in two bands of 99 and 301 bp but without *Apal*I restriction show a 400 bp band ; lane 6-10: genotype with heterozygous (C/T).

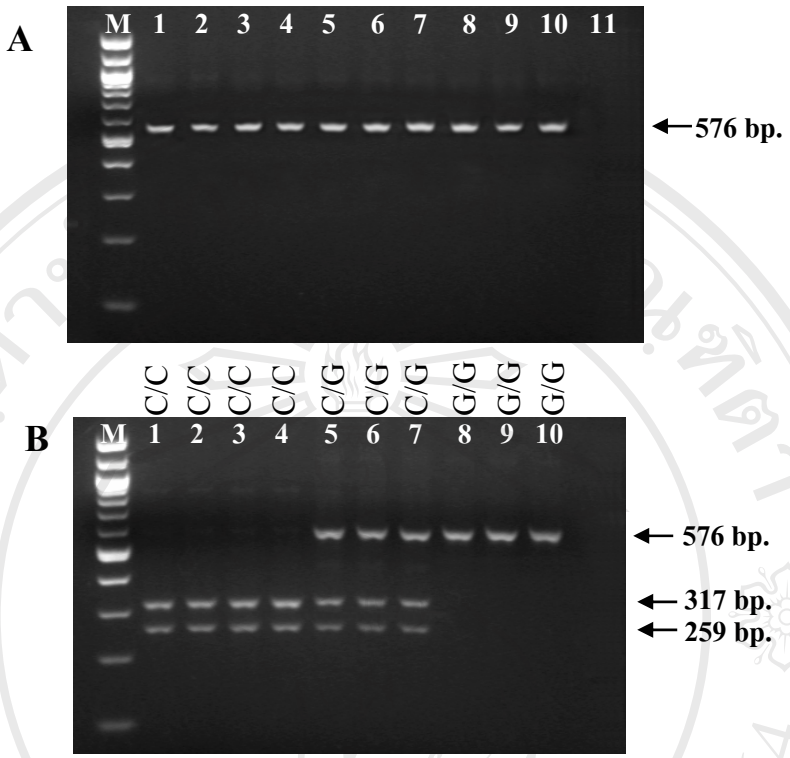


Figure 18 Typical electrophoresis pattern in detection of the 7520C>G 3'UTR of *CYP2A13* gene. (A) The first PCR amplification analysis. Lane M: 100 bp DNA ladder marker; lanes 1-10: first PCR products (576 bp) amplified from ten DNA samples; lane 11: a native control in which on DNA template was added. (B) The RFLP fragments obtained using *Xma*III restriction enzyme cleavage. Lane M: 100 bp DNA ladder marker; lanes 1-4: homozygous wild-type () with *Xma*III restriction site results in two bands of 259 and 317 bp.; lanes 5-7: heterozygous (C/G) with al three bands of 259, 317and 576 bp; lanes 8-10 homozygous mutant-type (G/G) without *Xma*III restriction show a 576 bp band.

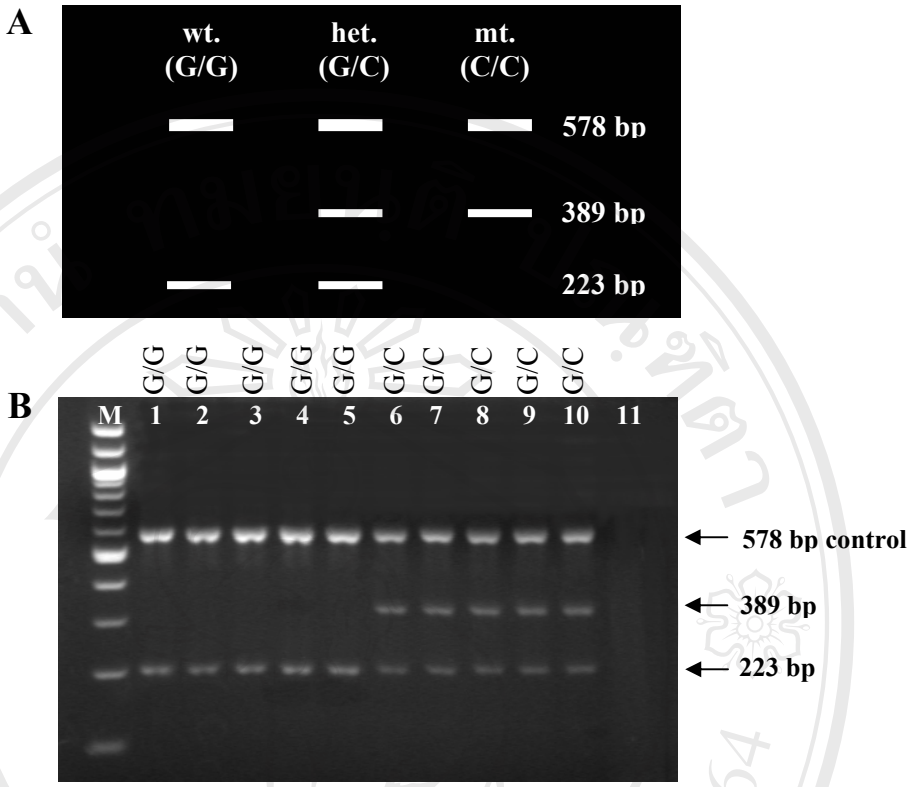


Figure 19 Typical electrophoresis pattern in detection of the 7571G>C 3'UTR of *CYP2A13* gene. (A) Representative agarose gel electrophoresis pattern of tetra-primer PCR for detection of the 7571G>C polymorphism. One is homozygous for 7571G allele (G/G), one is heterozygous for 7571G>C (G/C) and one homozygous for the 7571C allele (C/C). (B) The Tetra-primer analysis of the 7571G>C polymorphism. Lane M: 100 bp DNA ladder; lanes 1-5: genotype with homozygous wide-type (); lanes 6-10: genotype with heterozygous (G/C); lane 11: a negative control in which on DNA template was added.

Figure 20 DNA sequencing confirmation of the 74G>A and 1662G>C variants. (A) Mismatch PCR-RFLP analysis for the 74G>A SNP in exon 1 with *Hha*I restriction enzyme cleavage. The digest of a heterozygous sample (G/C) resulted in three fragments (24, 223, and 247 bp.) and the wild-type sample (G/G) was completely digested in two fragments (24 and 223 bp.). (B) PCR-RFLP analysis for the 1662G>C in exon 3 with *Hae*II restriction enzyme cleavage. The wild-type sample (G/G) that carries a *Hae*II restriction site at one position produced two fragments of 177 and 229 bp. (C) DNA sequencing confirmed heterozygosity for the 74G>A allele in exon 1 and homozygosity for the 1662G allele in exon 3.

Figure 21 DNA sequencing confirmation of the 3375C>T SNP in exon 5. (A) Digest with *Hha*I restriction enzyme. The wild-type sample (C/C) was completely digested in two fragments (99 and 301 bp) and the digest of a heterozygous sample (C/T) resulted in three fragments (99, 301 and 400 bp). (B) Digest with *Apal* I restriction enzyme, which does not cut the 3375C wild-type allele but cuts the 3375T mutant-type allele to give two fragments (100 and 300 bp). Thus the wild-type sample (C/C) was a fragment of 400 bp. The digest of a heterozygous sample (C/T) resulted in three fragments (100, 300 and 400 bp). (C) DNA sequencing confirmed heterozygosity and homozygosity for the 3375C>T allele in exon 5.

Figure 22 DNA sequencing confirmation of the 7520C>G SNP in 3'UTR. (A) PCR-RFLP analysis for the 7520C>G SNP in 3'UTR with *Xma*III restriction enzyme cleavage. The digest of a heterozygous sample (C/G) resulted in three fragments (259, 317 and 576 bp), the wild-type sample (C/C) was completely digested in two fragments (24 and 223 bp) and the homozygous mutant sample (G/G) was not digested, thus resulted in only one fragment (576 bp). (B) DNA sequencing confirmed heterozygosity for the 7520C>G allele and homozygosity for the 7520C and 7520G allele in 3'UTR.

Figure 23 DNA sequencing confirmation of the 578C>T SNP in exon 2 and 7571G>C SNP in 3'UTR. (A) Mismatch allele-specific PCR amplification analysis for the 578C>T SNP in exon 2. The fragment of 186 bp was used to identify 578C wild-type allele or 578T variant allele. The wild-type sample (C/C) has one fragment of 186 bp. Heterozygous sample (C/T) has two fragment of both wild-type and mutant-type (186 bp). (B) Tetra-primer PCR analysis for the 7571G>C SNP in 3'UTR. PCR Amplification of the 3'UTR with the outer primer pairs produced a large 576 bp fragment that served as internal control. The ASA reaction produced the 389 bp PCR product specific for the mutant-allele (7571C) allele and the 223 bp PCR product specific for the wild-type (7571G) allele. Thus, the wild-type sample generated two fragments of 223 and 576 bp, and the heterozygous sample producted three fragments of 223, 389 and 576 bp. (C) DNA sequencing confirmed heterozygosity for the 578C>T allele in exon 2 and 7571G>C allele in 3'UTR.

5.4 The frequencies of the *CYP2A13* alleles and genotypes

The allele frequencies and genotype distributions of the CYP2A13 in this study population were shown in table 6 and 7. In this study, the allele frequencies of the variant alleles were investigated in a total 336 Northern Thai subjects (175 males and 161 females). The allelic frequencies of 74G>A, 578C>T, 3375C>T, 7520C>G and 7571G>C were 4.9%, 0.1%, 4.9%, 4.2%, and 4.6%, respectively. The most of variant alleles were detected as heterozygotes except for the 7520G>C, there were both heterozygotes and homozygotes. Surprisingly, the 1662G>C variant allele was not found in this population examined.

Of the six SNPs, four are located in the coding region, whereas the other two are located in 3' untranslated region. The SNP in exon 1 was a 74G>A missense mutation, leading to a predicted amino acid change from Arg to Gln at position 25 (Arg25Gln). Of the 336 Thai subjects, only 33 subjects (9.8%) were heterozygous for the 74G>A SNP and no homozygous mutant was observed. Among the thirty-three heterozygotes, 13 were males and 20 were females, with an apparent frequency of 74A allele being 4.9% (95%CI 3.5-6.9).

The SNP detected in exon 2 was a 578C>T (responsible for CYP2A13*7) missense mutation, resulting in an amino acid change from an Arg to a stop codon at position 101 (Arg101Stop). This SNP was detected as heterozygote in a male subject (0.3%). The 578C>T polymorphism was relatively rare in this population, with an apparent 578T allele frequency of 0.1% (95%CI 0.0-1.0).

A rare mutation in exon 3, previously detected in a Hispanic newborn (Zhang et al., 2002) was a missense mutation (1662G>C), which is expected to cause a

conserved Gly144Arg substitution. It was the only one SNP that was not found variant allele in this Northern Thai population examined.

The SNP in exon 5 (3375C>T) was a missense mutation, resulting in an amino acid change from an Arg to a Cys codon at position 257 (Arg257Cys). Of the 336 Thai subjects, this SNP was detected as hetorzygotes in 33 (9.8%) subjects and no homozygous mutant was observed. Among the thirty-three heterozygotes, 17 were males and 16 were females, with an apparent allele frequency of 4.9% (95%CI 3.5-6.9).

The SNP detected in 3'Untranslated region (7520C>G), which have no amino acid change. This SNP was detected as hetorzygotes in 22 (6.5%) subjects (9 males and 13 females) and homozyguos mutant in 3 (0.9%) subjects (2 males and 1 female), with an apparent allele frequency of 4.2% (95%CI 2.8-6.1).

The last SNP as detected in 3'Untranslated region (7571G>C) and also have no amino acid change. This SNP was detected as only heterozygote in 31 (9.2%) subjects and no homozygote, with an apparent allele frequency of 4.6% (95%CI 3.2-6.6). Among the thirty-one heterozygotes, 18 were males and 13 were females.

The distribution of genotypes for the 74G>A, 578C>T, 3375C>T, 7520C>G, and 7571G>C SNPs conformed well to the Hardy-Weinberg equilibrium (*P*= or 0.502 higher) (appendix E)

No significant difference was observed in the *CYP2A13* allele frequencies among male and female subjects (*P*= 0.053 or higher) (appendix F).

Genotype data of 336 Thai subjects were used for haplotype configuration analysis of the *CYP2A13* gene. Ten haplotype configurations were observed and designed as haplotype I-X. As shown in Table 8, the most frequent haplotype was

haplotype I, the wild type (frequency 79.2%) followed by VII (5.7%), IV (5.4%), II (3.9%) and IX (3.3%) which were corresponded to *CYP2A13*1A*, **1H* (or **3*), **2B*, and **1G*, respectively. The frequencies the other haplotypes were relatively low.

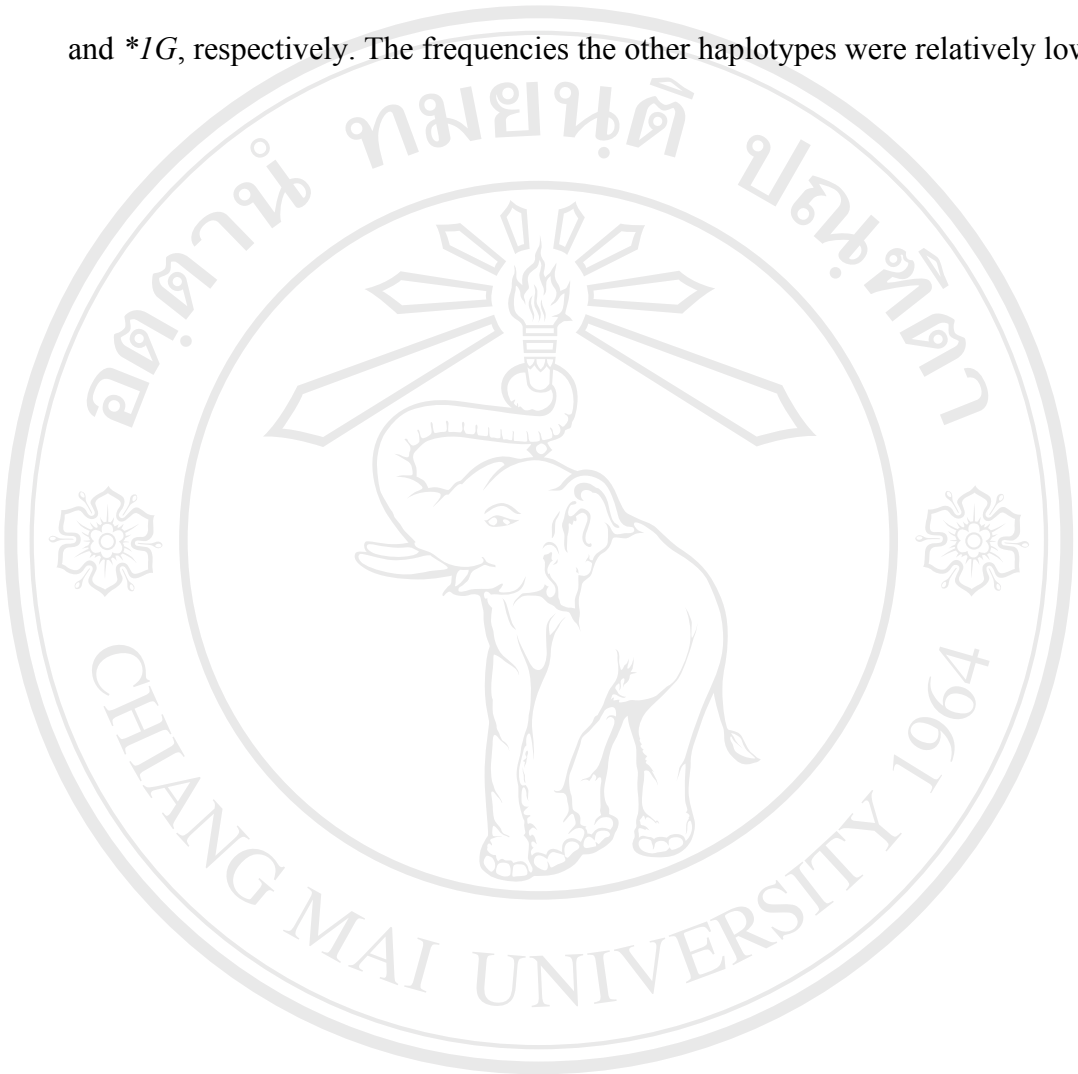


Table 6 Distribution of *CYP2A13* allele frequencies in a total 336 Northern Thai subjects.

SNPs		No. of alleles	% Frequency (95%CI)
Exon 1: 74G>A (Arg25Gln)			
G allele		639	95.1 (93.1-96.6)
A allele		33	4.9 (3.5-6.9)
	total	672	100.0
Exon 2: 578C>T (Arg101Stop)			
C allele		671	99.9 (99.0-100.0)
T allele		1	0.1 (0.0-1.0)
	total	672	100.0
Exon 3: 1662G>C (Gly144Arg)			
G allele		672	100.0 (99.3-100.0)
C allele		0	0.0 (0.0-0.7)
	total	672	100.0
Exon 5: **3375C>T** (Arg257Cys)			
C allele		639	95.1 (93.1-96.6)
T allele		33	4.9 (3.5-6.9)
	total	672	100.0
3'Untranslated region			
7520C>G			
C allele		644	95.8 (94.0-97.2)
G allele		28	4.2 (2.8-6.1)
	total	672	100.0
7571G>C			
G allele		641	95.4 (93.4-96.8)
C allele		31	4.6 (3.2-6.6)
	total	672	100.0

Notes: The sequence of the complete human CYP2A13 gene described in genBank (accession no AC008962) was used as a reference sequence, with the A of ATG start codon designated as +1 (nucleotide 66807).

Table 7 Distribution of *CYP2A13* genotyping in total 336 Northern Thai subjects.

SNPs		*genotype frequency*		
		Observed	**%**	**Expected**[a]
Exon 1 (74G>A)				
Wild-type	**G/G**	303	90.2	304
Heterozygous	**G/A**	33	9.8	31
Mutant-type	**A/A**	0	0	1
	Total	336	100	
Exon 2 (578C>T)				
Wild-type	**C/C**	335	99.7	335
Heterozygous	**C/T**	1	0.3	1
Mutant-type	**T/T**	0	0	0
	Total	336	100	
Exon 3 (1662G>C)				
Wild-type	**G/G**	336	100	336
Heterozygous	**G/C**	0	0	0
Mutant-type	**C/C**	0	0	0
	Total	336	100	
Exon 5 (3375C>T)				
Wild-type	**C/C**	303	90.2	304
Heterozygous	**C/T**	33	9.8	31
Mutant-type	**T/T**	0	0	1
	Total	336	100	
3'Untranslated region 7520C>G				
Wild-type	**C/C**	311	92.6	308
Heterozygous	**C/G**	22	6.5	27
Mutant-type	**G/G**	3	0.9	1
	Total	336	100	
7571G>C				
Wild-type	**G/G**	305	90.8	306
Heterozygous	**G/C**	31	9.2	29
Mutant-type	**C/C**	0	0	1
	Total	336	100	

[a] The expected values were deduced by the Hardy-Weinberg equation from allele frequencies in Table 6

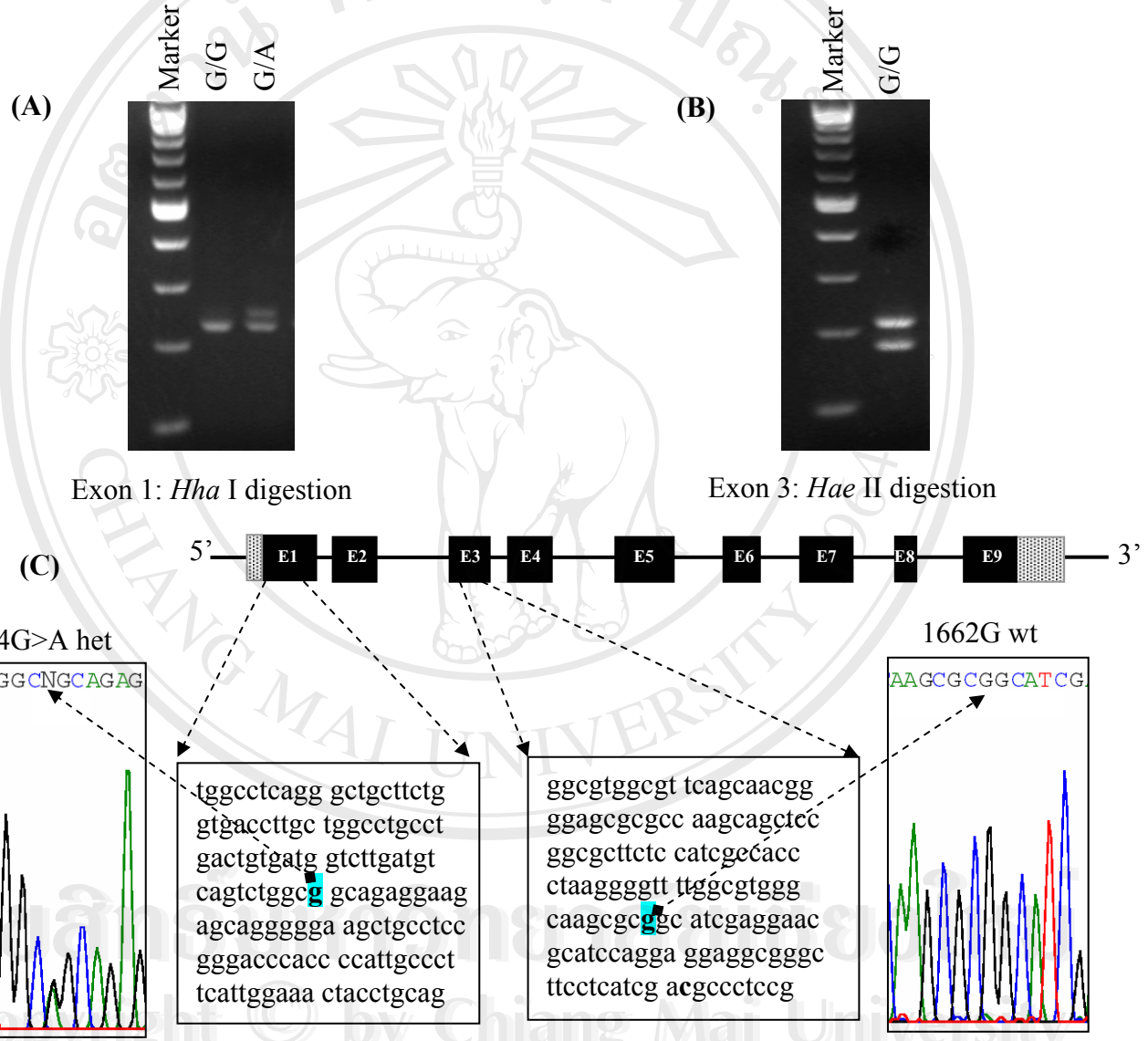
Notes: The Chi-square test (χ^2 test) was used to compare the expected and observed genotype. The observed and expected counts are not significantly different form one another (appendix E).

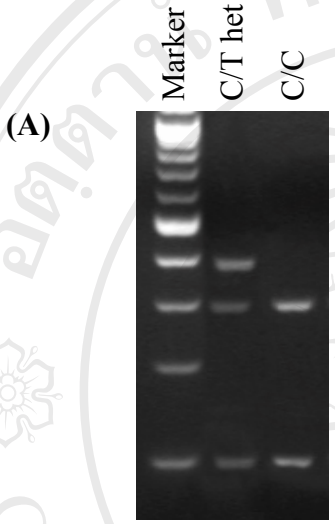
Table 8 Haplotype configurations of the *CYP2A13* gene estimated in 336 Thai subjects.

Haplotype	Exon1 74G>A	Exon 2 578C>T	Exon 3 1662G>C	Exon 5 3375C>T	3'UTR 7520C>G	3'UTR 7571G>C	*n* (%)	Corresponded allele[a]
I	G/G	C/C	G/G	C/C	C/C	G/G	266 (79.2)	**1A*
II	**G/A**	C/C	G/G	**C/T**	C/C	G/G	13 (3.9)	**2A*
III	**G/A**	C/C	G/G	**C/T**	**C/G**	G/G	1 (0.3)	
IV	**G/A**	C/C	G/G	**C/T**	C/C	**G/C**	18 (5.4)	**2B*
V	**G/A**	C/C	G/G	**C/T**	**C/G**	**G/C**	1 (0.3)	
VI	G/G	**C/T**	G/G	C/C	C/C	G/G	1 (0.3)	**7*
VII	G/G	C/C	G/G	C/C	**C/G**	G/G	19 (5.7)	**1H* or **3*
VIII	G/G	C/C	G/G	C/C	**C/G**	**G/C**	1 (0.3)	
IX	G/G	C/C	G/G	C/C	C/C	**G/C**	11 (3.3)	**1G*
X	G/G	C/C	G/G	C/C	**G/G**	G/G	3 (0.9)	
				total			336 (100)	

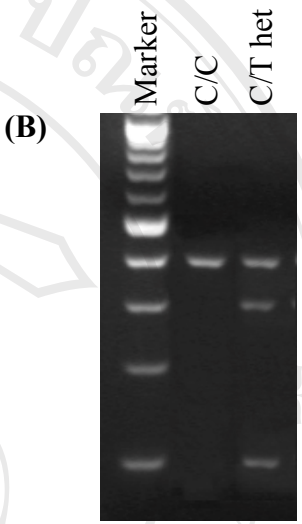
n = number of subjects evaluated.

[a] haplotype nomenclature is as recommended by the Human Cytochrome P450 Allele Nomenclature (www. imm.ki.se/CYPalleles).





Exon 5: *Hha* I digestion



Exon 5: *Apal* I digestion