# CHAPTER 3

# RESULTS

## Identification of microsatellite and SNP alleles

The alleles of each polymorphic marker were genotyped by gel analysis (Figures 4.1-4.8). The nomenclature of the microsatellite alleles was based on the number of AC dinucleotide repeats as determined by sequencing. The discrimination of each SNP allele corresponded to their PCR product fragment size following careful consideration of the different product sizes obtained from the primer designs of the Biallelic-ARMS PCR technique.

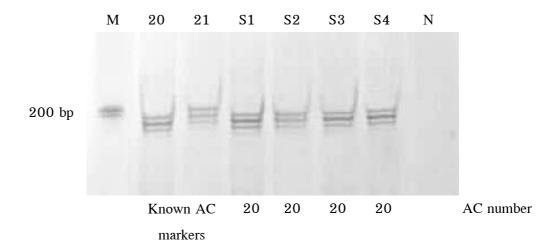


Figure 4.1. The representative polyacrylamide gel for DXS548. Marker (M) was the low molecular weight DNA ladder (200 ng). Lanes S1–S4 show individual males presenting a DXS548 microsatellite allele; the genotyping results are shown below each sample lane. There was no DNA template in Lane N (negative control). All lanes showed DXS548 allele 20 when compared with known alleles of 20 and 21 AC repeats (2<sup>nd</sup> and 3<sup>rd</sup> lanes from the left, respectively).

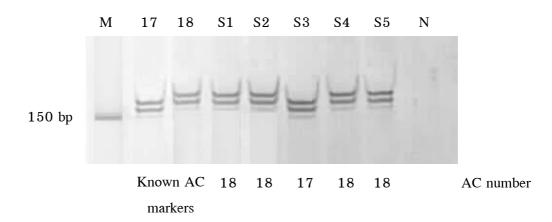


Figure 4.2. The representative polyacrylamide gel for FRAXAC1. Marker (M) was the low molecular weight DNA ladder (200 ng). Lanes S1–S5 were individual males presenting a FRAXAC1 microsatellite allele; the genotyping results are shown below each sample lane. There was no DNA template in Lane N (negative control). Lane S3 had FRAXAC1 allele 17 while lanes S1, S2, S4 and S5 showed FRAXAC1 allele 18 when compared with known alleles of 17 and 18 AC repeats (2<sup>nd</sup> and 3<sup>rd</sup> lanes from the left, respectively).

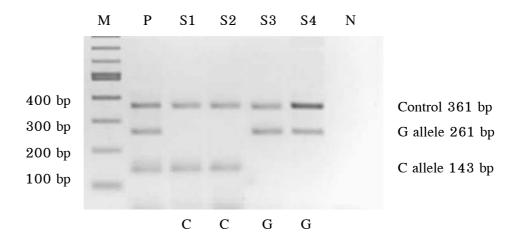


Figure 4.3. The representative agarose gel for WEX5. Marker (M) was the 100 bp DNA ladder (400 ng). Lane P (positive control) contained the DNA control of the heterozygous female indicating two different SNP alleles and there was no DNA template in Lane N (negative control). Lanes S1–S4 show samples of individual males presenting a particular SNP allele along with the genotyping results shown below each sample lane. The control product of 361 bp used as the internal control to prove the successful PCR amplification appears in every lane. Lanes S1 and S2 exhibit the fragment size of 143 bp resulting from the WEX5 genotype C as well as lanes S3 and S4 indicate the PCR product of 261 bp scoring of WEX5 genotype G.

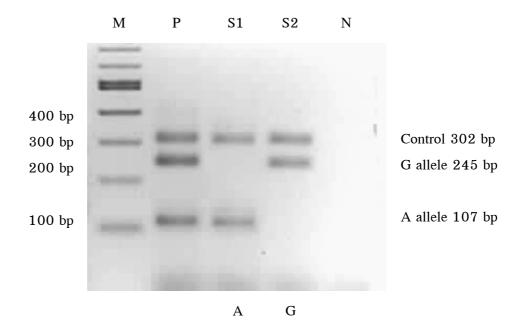


Figure 4.4. The representative agarose gel for ATL1. Marker (M) was the 100 bp DNA Ladder (400 ng). Lane P (positive control) contains the DNA control of the heterozygous female indicating two different SNP alleles and there is no DNA template in Lane N (negative control). Lanes S1 and S2 are samples of individual males presenting a particular SNP allele along with the genotyping results shown below each sample lane. The control product of 302 bp was used as an internal control to prove the successful PCR amplification which appears in every lane. Lane S1 exhibits the fragment size of 107 bp resulting in the ATL1 genotype A and lane S2 shows the PCR product of 245 bp scoring of ATL1 genotype G.

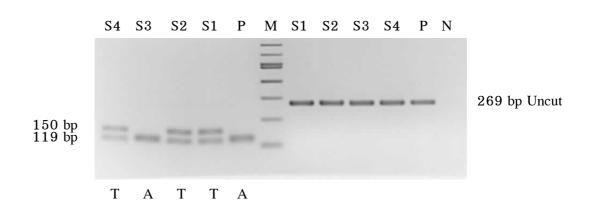


Figure 4.5. The representative agarose gel for rs25731. Marker (M) was the 100 bp DNA ladder (400 ng). Lane P (positive control) contains the male DNA control and there is no DNA template in Lane N (negative control). Lanes S1–S4 show samples of individual males presenting a particular SNP allele along with the genotyping results shown below each sample lane. Unlike the multiplex PCR of other SNP loci, the singleplex PCR with only control forward and reverse primer of SNP rs25731 amplified a PCR product of 269 bp (lanes S1–S4 and P on the right part of the gel). After these products were digested with restriction endonuclease, *Dra* I, the A allele of rs25731 displayed three bands of 125, 119 and 25 bp and the pattern looked like one band on 2.5% gel (lanes P and S3 on the left side of gel) because the 2.5 percent of the gel could not resolve the 6-bp difference fragments and 25-bp fragment was run out of gel. The distinguishing pattern of the T allele presented two obvious distinct bands of 150 and 119 bp (lanes S1, S2 and S4 on the left side of the gel).

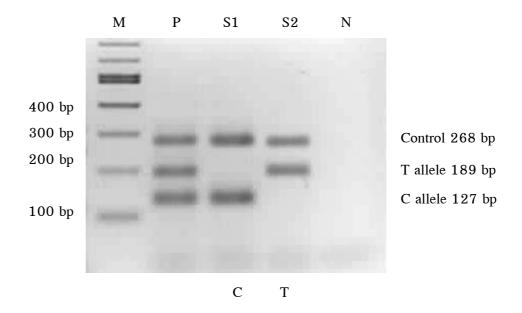


Figure 4.6. The representative agarose gel for IVS10. Marker (M) was the 100 bp DNA Ladder (400 ng). Lane P (positive control) contains the DNA control of the heterozygous female indicating two different SNP alleles and there is no DNA template in Lane N (negative control). Lanes S1 and S2 are samples of individual males presenting a particular SNP allele along with the genotyping results shown below each sample lane. The control product of 268 bp used as the internal control to prove the successful PCR amplification appears in every lane. Lane S1 exhibits the fragment size of 127 bp resulting in IVS10 genotype C and lane S2 shows the PCR product of 189 bp scoring of IVS10 genotype T.

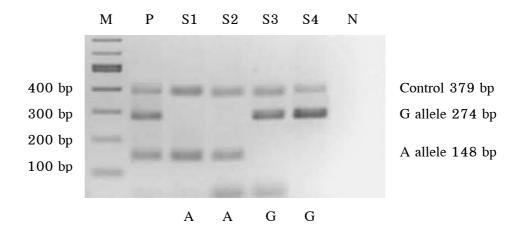


Figure 4.7. The representative agarose gel for rs25702. Marker (M) was the 100 bp DNA ladder (400 ng). Lane P (positive control) contains the DNA control of the heterozygous female indicating two different SNP alleles and there is no DNA template in Lane N (negative control). Lanes S1-S4 are samples of individual males presenting a particular SNP allele along with the genotyping results shown below each sample lane. The control product of 379 bp used as the internal control to prove the successful PCR amplification appears in every lane. Lanes S1 and S2 exhibit the fragment size of 148 bp resulting in rs25702 genotype A as well as lanes S3 and S4 show the PCR product of 274 bp scoring of rs25702 genotype G.

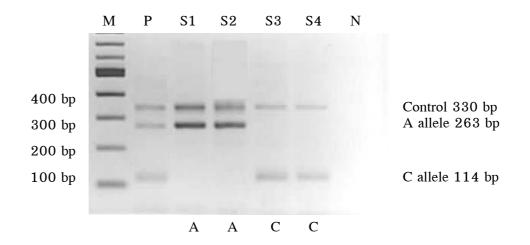


Figure 4.8. The representative agarose gel for rs25723. Marker (M) was the 100 bp DNA ladder (400 ng). Lane P (positive control) contains the DNA control of the heterozygous female indicating two different SNP alleles and there is no DNA template in Lane N (negative control). Lanes S1-S4 are samples of individual males presenting a particular SNP allele along with the genotyping results shown below each sample lane. The control product of 330 bp used as the internal control to prove the successful PCR amplification appears in every lane. Lanes S1 and S2 exhibit the fragment size of 263 bp resulting in rs25723 genotype A as well as lanes S3 and S4 which show the PCR product of 114 bp scoring of rs25723 genotype C.

#### Allele distributions and frequencies of the microsatellite and SNP markers

We evaluated two types of markers, microsatellites and single nucleotide polymorphisms (SNPs). The microsatellite marker is more polymorphic than SNP which has only two alleles. As can be seen in Table 6, five DXS548 alleles of 18, 20, 21, 22 and 25 AC repeats presented in normal subjects, while only two alleles of 20 and 21 repeats observed in FXS subjects. The AC 20 repeat allele was the most common allele accounting for 90.2% (120/133) of all alleles in the normal groups and 96% (48/50) in the FXS groups. The AC alleles of FRAXAC1 in both cohorts had a similar allele distribution pattern of 17, 18 and 19 alleles (there was no allele 19 in the FXS samples). The 17 AC allele followed by the 18 AC allele were the most prevalent genotypes of both normal and FXS groups with similar allele frequencies, 61.7% as compared with 62% for 17 AC repeats and 36.8% as compared with 38% for 18 AC repeats, respectively. We found no statistically significant differences in allele distributions and frequencies of all polymorphic markers between the normal and FXS groups (P > 0.05, Table 6). In both sample groups, DXS548 was less polymorphic than other markers. The heterozygosity (Het) of DXS548 in normal chromosomes was 18.1%, while the remaining markers presented similar heterozygosities ranging from ~46-48%. Additionally on FXS chromosomes, the smallest heterozygosity was 7.7% in the DXS548 locus, which greatly differed from other loci (~46-47%). We observed that the heterozygosity of FRAXAC1 and all SNPs in normal samples were similar to the FXS groups.

Locus	A 11 a 1 a	Normal		FXS		
Locus	Allele	NO.	%	NO.	%	
DXS548	18	1	0.8	0	0.0	
	20	120	90.2	48	96.0	
	21	9	6.8	2	4.0	
	22	1	0.8	0	0.0	
	25	2	1.5	0	0.0	
	total	133	100.0	50	100.0	
	Het		18.1		7.7	

 Table 6. Allele distributions and frequencies of the microsatellite and SNP markers in normal controls and FXS patients.

Normal vs FXS: Chi-square = 2.10, df = 4, P = 0.72 (not significant)

Larma	Allele	No	rmal	FXS		
Locus	Allele	NO.	%	NO.	%	
FRAXAC1	17	82	61.7	31	62.0	
	18	49	36.8	19	38.0	
	19	2	1.5	0	0.0	
	total	133	100.0	50	100.0	
	Het		48.4		47.1	

Normal vs FXS: Chi-square = 0.77, df = 2, P = 0.68 (not significant)

Locus	Allele	Normal		F2	XS
Locus	Allele	NO.	%	NO.	%
WEX5	С	84	63.2	31	62.0
	G	49	36.8	19	38.0
	total	133 100.0 50		50	100.0
	Het		46.5		47.1

Normal vs FXS: Chi-square = 0.02, df = 1, P = 0.89 (not significant)

## Table 6. (continued)

Loous	Locus Allele	Normal		FXS		
Locus	Allele	NO.	%	NO.	%	
ATL1	А	48	36.1	18	36.0	
	G	85	63.9	32	64.0	
	total	133	100.0	50	100.0	
	Het		46.1		46.1	

Normal vs FXS: Chi-square = 0.00, df = 1, P = 0.99 (not significant)

Leoya	Allele	Normal		F	XS
Locus	Allele	NO.	%	NO.	%
rs25731	А	48	36.1	19	38.0
	Т	85	63.9	31	62.0
	total	133	100.0	50	100.0
	Het		46.1		47.1

Normal vs FXS: Chi-square = 0.06, df = 1, P = 0.81 (not significant)

Locus Allele	Allele	Normal		FXS		
Locus	Allele	NO.	%	NO.	%	
IVS10	С	52	39.1	19	38.0	
	Т	81	60.9	31	62.0	
	total	133	100.0	50	100.0	
	Het		47.6		47.1	

Normal vs FXS: Chi-square = 0.02, df = 1, P = 0.89 (not significant)

Lagua	Allele	Normal		FXS		
Locus	Allele	NO.	%	NO.	%	
rs25702	А	81	60.9	31	62.0	
	G	52	39.1	19	38.0	
	total	133	100.0	50	100.0	
	Het		47.6		47.1	

Normal vs FXS: Chi-square = 0.02, df = 1, P = 0.89 (not significant)

#### Table 6. (continued)

Leona	Allele	Norr		FXS		
Locus	Allele	NO.	%	NO.	%	
rs25723	А	81	60.9	31	62.0	
	С	52	39.1	19	38.0	
	total	133	100.0	50	100.0	
	Het		47.6		47.1	

Normal vs FXS: Chi-square = 0.02, df = 1, P = 0.89 (not significant)

#### Haplotype analysis

We established several series of combined haplotype and analyzed all of them (Table 7). This analysis revealed that all sets of combined haplotypes showed no statistically significant differences of haplotype associations between the normal and FXS groups (P > 0.05), similar to the study of Thai subjects reported by Limprasert (2001), even though this study included more polymorphic markers and sample pools than the previous study. Nevertheless, we clearly observed a common haplotype background (bold characters) accounting for the majority of both groups in every haplotype set. Thus, we divided the combined haplotype of seven markers (FRAXAC1-WEX5-ATL1-rs25731-IVS10-rs25702-rs25723) into 4 groups including 17-G-G-A-T-A-A (designated as Hap A), 18-C-A-T-C-G-C (designated as Hap B), 17-C-G-T-T-A-A (designated as Hap C) and other haplotypes (Table 8). Reanalysis of such major and other haplotypes still detected no specific haplotype association (P = 0.88). In addition, Hap A was similar to Hap C with minor differences of only 2 SNP loci (WEX5 and rs25731), whereas Hap B was different from Hap A in all polymorphic markers and from Hap C in 5 markers (FRAXAC1, ATL1, IVS10, rs25702 and rs25723). Hap B was slightly different from the haplotypes of the OTHER group at 1 or 2 markers (FRAXAC1, WEX5 and ATL1). If it was assumed that these common haplotypes occurred on a related background, we could group them into 2 major haplotype backgrounds. The majority of samples in normal and FXS individuals occurred on haplotype backgrounds Hap A and Hap C (60.9% and 62%, respectively) while the minority of samples occurred on haplotype backgrounds Hap B and OTHER (39.1% and 38%, respectively).

Table 7. Frequencies of each haplotype set in normal and FXS groups. (D = DXS548, AC = FRAXAC1, W = WEX5, AT = ATL1, 1 = rs25731, IV = IVS10, 2 = rs25702, 3 = rs25723)

Haplotype set	Combined haplotypes	Nor	rmal	FXS	
napiotype set	Combined napiotypes	NO.	%	NO.	%
D-AC	18-17	1	0.8	0	0.0
	20-17	74	55.6	31	62.0
	20-18	44	33.1	17	34.0
	20-19	2	1.5	0	0.0
	21-17	5	3.8	0	0.0
	21-18	4	3.0	2	4.0
	22-18	1	0.8	0	0.0
	25-17	2	1.5	0	0.0
	total	133	100.0	50	100.0

Normal vs. FXS: Chi-square = 4.51, df = 7, P = 0.72 (not significant)

Haplatupa sat	Combined honletunes	Noi	rmal	F	XS
Haplotype set	Combined haplotypes	1     0.3       1     0.3       30     22       43     32       41     30       2     1.3       2     1.3       2     1.3       2     1.3       2     1.3       3     2.3	%	NO.	%
D-AC-W-AT-1-IV-2-3	18-17-G-G-A-T-A-A	1	0.8	0	0.0
	20-17-C-G-T-C-G-C	1	0.8	0	0.0
	20-17-С-G-Т-Т-А-А	30	22.6	12	24.0
	20-17-G-G-A-T-A-A	43	32.3	19	38.0
	20-18-C-A-T-C-G-C	41	30.8	16	32.0
	20-18-C-G-T-C-G-C	2	1.5	1	2.0
	20-18-G-A-T-C-G-C	1	0.8	0	0.0
	20-19-C-A-T-C-G-C	1	0.8	0	0.0
	20-19-C-G-T-C-G-C	1	0.8	0	0.0
	21-17-C-G-T-T-A-A	3	2.3	0	0.0
	21-17-G-G-A-T-A-A	2	1.5	0	0.0
	21-18-C-A-T-C-G-C	4	3.0	2	4.0
	22-18-C-A-T-C-G-C	1	0.8	0	0.0
	25-17-G-G-A-T-A-A	2	1.5	0	0.0
	total	133	100.0	50	100.0

Normal vs. FXS: Chi-square = 5.44, df = 13, P = 0.96 (not significant)

Table 7. (continued)

Haplotype set	Combined haplotypes	Nor	Normal		XS
Haplotype set	Combined naplotypes	NO.	%	NO.	%
AC-W-AT-1-IV-2-3	17-G-G-A-T-A-A	48	36.1	19	38.0
	17-C-G-T-C-G-C	1	0.8	0	0.0
	17-С-G-Т-Т-А-А	33	24.8	12	24.0
	18-C-A-T-C-G-C	46	34.6	18	36.0
	18-C-G-T-C-G-C	2	1.5	1	2.0
	18-G-A-T-C-G-C	1	0.8	0	0.0
	19-C-A-T-C-G-C	1	0.8	0	0.0
	19-C-G-T-C-G-C	1	0.8	0	0.0
	total	133	100.0	50	100.0

Normal vs. FXS: Chi-square = 1.63, df = 7, P = 0.98 (not significant)

Table 8. Common haplotypes among normal and FXS subjects using seven polymorphicmarkers (FRAXAC1-WEX5-ATL1-rs25731-IVS10-rs25702-rs25723).

Haplotype		Normal		FXS	
AC-W-AT-1-IV-2-3	Designated name	NO.	%	NO.	%
17-G-G-A-T-A-A	Hap A	48	36.1	19	38.0
17-С-G-Т-Т-А-А	Hap C	33	24.8	12	24.0
18-C-A-T-C-G-C	Hap B	46	34.6	18	36.0
OTHER	_	6	4.5	1	2.0
total		133	100.0	50	100.0

Normal vs. FXS: Chi-square = 0.67, df = 3, P = 0.88 (not significant)

#### Haplotype association

When we divided the normal cohort into 6 subgroups according to the number of CGG repeats (19-28, 29, 30, 31-35, 36 and 37-56), we found a striking and significant association between the common CGG repeats (29, 30 and 36) and haplotype (FRAXAC1-WEX5-ATL1-rs25731-IVS10-rs25702-rs25723) (P = 0.00, Table 9). The 29 CGG alleles were associated with the Hap A (27/32 or 84.4%), the 30 CGG alleles were associated with the Hap B (29/31 or 93.5%) and the 36 CGG alleles were associated with the Hap C (23/28 or 82.1%). We also found, however, these common haplotypes in the FXS chromosomes (38% of Hap A, 36% of Hap B and 24% of Hap C). We observed that there was no Hap B in the 29 and 36 CGG repeats subgroups. Also, Hap C was not present in the 30 CGG repeats subgroup. Only one chromosome with Hap A was observed in the 30 CGG repeats subgroup. Interestingly, the majorities of large CGG repeats (36-56) were related to Hap A and Hap C (32/35 or 91.4%), while the other subgroups of the smaller CGG alleles (19-28 and 31-35) had only 43.8% (7/16) and 54.5% (12/22) of Hap A and Hap C, respectively. Nevertheless, no association of Hap A and Hap C with the FXS groups (31/50 or 62%) was found.

Table 9. Haplotype associations in subgroups of normal CGG repeats. (AC = FRAXAC1, W = WEX5, AT = ATL1, 1 = rs25731, IV = IVS10,

2 = rs25702,	3 = rs25723)
--------------	--------------

Haplotype (Hap)	19	-28	2	9	3	0	31	-35	3	6	37-	-56	FZ	XS
AC-W-AT-1-IV-2-3	NO.	%												
17-G-G-A-T-A-A (A)	5	31.3	27	84.4	1	3.2	5	38.5	5	17.9	5	38.5	19	38.0
17-C-G-T-T-A-A (C)	2	12.5	2	6.3	0	0.0	1	7.7	23	82.1	5	38.5	12	24.0
18-C-A-T-C-G-C (B)	8	50.0	0	0.0	29	93.5	6	46.2	0	0.0	3	23.1	18	36.0
OTHER	1	6.3	3	9.4	1	3.2	1	7.7	0	0.0	0	0.0	1	2.0
total	16	100.0	32	100.0	31	100.0	13	100.0	28	100.0	13	100.0	50	100.0

29 vs 30: Chi-square = 56.14, df = 3, P = 0.00 (significant)

29 vs 36: Chi-square = 36.03, df = 2, P = 0.00 (significant)

- 29 vs FXS: Chi-square = 24.78, df = 3, P = 0.00 (significant)
- 30 vs 36: Chi-square = 55.66, df = 3, P = 0.00 (significant)
- 30 vs FXS: Chi-square = 27.85, df = 3, P = 0.00 (significant)
- 36 vs FXS: Chi-square = 26.53, df = 3, P = 0.00 (significant)

#### **AGG Interruption**

The results from the sequencing analysis in each subgroup of normal chromosomes revealed some variety in both the number of AGGs and the AGG substructures within the FMR1 CGG alleles (Figure 5 and Table 10). Most normal alleles had two interspersed AGG triplets (49/95 or 51.6%). Alleles with a single or more (3-4) AGG interspersions had similar frequencies of 20% (19/95) and 24.2% (23/95), respectively. Alleles devoid of AGG were rare in this study, accounting for only 4.2% (4/95). The position of the first AGG interruption varied, but in 91.2% (83/91) of normal alleles having one or more AGGs it punctuated at the tenth (9A) or eleventh (10A)position within the repeat. The most common CGG lengths between the first and second AGG for alleles with two or more AGG interruptions were nine or ten CGG repeat units (63/72 or 87.5%). Thus, the general pattern for the most normal alleles was two AGGs intervening every 9 or 10 CGG repeats with repeat length variability at the 3' end. Strikingly, the majority of the common CGG alleles 29, 30 and 36, had the specific AGG organization of 9A9A9 (15/17 or 88.2%), 10A9A9 (16/18 or 88.9%) and 9A9A6A9 (15/18 or 83.3%), respectively, whereas the uncommon CGG alleles (19-28, 31-35 and 37-56) presented various AGG interspersion patterns corresponding to their CGG lengths. We also observed an allele possessing a 5' tract with 20 CGG repeats (20A9; 1/95 or 1.1%) and alleles with a 3' pure CGG tract (9/10An, where  $n \ge 20$ ; 9/95 or 9.5%).

	10	20	30	40	50	60	7.0
					1		T
ref CGG repeat	ACGGAGGCGC	CGCTGCCAGG	GGGCGTGCGG	CAGCGCGGCG	GCGGCGGCGG	CGGCGGCGGC	GGCGGAGGCG
EP57-571R	AcGgAGGcGC	CGCTGCCAgg	gGGCGTGcgG	CAGCgCgGCG	GCGGCGGCGG	CGGCGGCGGC	GGCGGCGGCG
P137FXS-FXS571R	AcGgAGGcGC	CGCTGCCAGg	gGGCGTGCgG	CAGCgCgGCG	GCGGCGGCGG	CGGCGGCGGC	GGAGGCGGCG
4 P125FXS-FXS571R	AcGgAGGcGC	CGCTGCCAgg	gGGCGTGCgG	CAGCGCGGCG	GCGGCGGCGG	CGGCGGCGGC	GG <u>CGGAGG</u> CG
215FXS-FXS571R	AcGgAGGcGc	CGCTGCcAgg	gGGcGTgcgG	CAGCgcggCG	GCGGCGGCGG	CGGCGGCGGC	GGAGGCGGCG
5 <b>E</b> RM43-3-571R				CAGCgCgGCG			
7 RM144-3FXS-FXS-seqF	AcGGAgGCGC	CGCTgCCAGG	GGGCGTGCGG	CAGCGCGGCG	GCGGCGGCGG	CGGCGGCGGC	GGAGGCGGCG
	80	90	100	110	120	130	140
l ref CGG repeat	60000000000	ceeceeceec	66066				<b>I</b>
2 P57-571R		CGGCGGCGGC					
P137FXS-FXS571R				GCGGCGGCGG	CGGCGGCGGC	66	
4 P125FXS-FXS571R				GCGGCGGCGG			
5 215 FXS - FXS 571R				GCGGCGGCGG			
5 RM43-3-571R				GCGGCGGCGG			
7 RM144-3FXS-FXS-seqF				GCGGCGGCGG			
MIIII-SIND-IND-Bedi	0000000000	6996996996	GGAGGCGGCG	0000000000	COGNOGCOGC	GGCGGCGGCG	GCGGAGGCGG
	150	160	170	) 180	190	200	210
l ref CGG repeat		I		CTCGAGCGCC	CCCACCCAC	CTCTCGGGGGG	CGGGCTCCCC
2 P57-571R				CTCGAGCGCC			
E P137FXS-FXS571R				CTCGAGCGCC			
							CAGACICCCA
							CCCCCCCCCC
4 P125FXS-FXS571R			CTGGGC	CTCGAGCGCC	CGCAGCCCAC	CTCTCGGGGG	
4 <b>E</b> p125FXS-FXS571R 5 <b>E</b> 215FXS-FXS571R			CTGGGC CTGGGC	CTCGAGCGCC CTCGAGCGCC	CGCAGCCCAC CGCAGCCCAC	CTCTCGGGGGG CTCTCGGGGGG	CGGGCTCCCG
4 P125FXS-FXS571R 215FXS-FXS571R 8 RM43-3-571R	 CGG		CTGGGC CTGGGC CTGGGC	CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC	CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC	CTCTCGGGGG CTCTCGGGGG CTCTCGGGGG	CGGGCTCCCG CGGGCTCCCG
4 <b>E</b> p125FXS-FXS571R 5 <b>E</b> 215FXS-FXS571R		GGCGGCGGCG	CTGGGC CTGGGC CTGGGC	CTCGAGCGCC CTCGAGCGCC	CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC	CTCTCGGGGG CTCTCGGGGG CTCTCGGGGG	CGGGCTCCCG CGGGCTCCCG
4 P125FXS-FXS571R 215FXS-FXS571R 5 RM43-3-571R			CTGGGC CTGGGC GCgGCTGGgC	CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC	CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC	CTCTCGGGGG CTCTCGGGGG CTCTCGGGGG cTCTCGGGGGG	CGGGCTCCCG CGGGCTCCCG CgggCTCCCG
4 P125FXS-FXS571R 5 215FXS-FXS571R 6 RM43-3-571R 7 RM144-3FXS-FXS-seqF		230 I	CTGGGC CTGGGC GCgGCTGGGC 240	CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC	CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC	CTCTCGGGGG CTCTCGGGGG CTCTCGGGGG cTCTCGGGGGG	CGGGCTCCCG CGGGCTCCCG CgggCTCCCG
E P125FXS-FXS571R 215FXS-FXS571R E RM43-3-571R 7 RM144-3FXS-FXS-seqF 1 ref CGG repeat	CGGCGGCGGC 220 I GCGCTAGCAG	230   GGCTGAAGAG	CTGGGC CTGGGC GCgGCTGGgC AAGATGGAGG	CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC 250 AGCTGGTG	CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC	CTCTCGGGGG CTCTCGGGGG CTCTCGGGGG cTCTCGGGGGG	CGGGCTCCCG CGGGCTCCCG CgggCTCCCG
2 P125FXS-FXS571R 215FXS-FXS571R 6 RM43-3-571R 7 RM144-3FXS-FXS-seqF 1 ref CGG repeat 2 P57-571R	CGGCGGCGGC 220 GCGCTAGCAG GCGCTAGCAG	230   GGCTGAAGAG GGCTGAAGAG	CTGGGC CTGGGC GCgGCTGGGC 240 AAGATGGAGG AAGATGGAGG	CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC 250 AGCTGGTG AGCTGGTG	CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC	CTCTCGGGGG CTCTCGGGGG CTCTCGGGGG cTCTCGGGGGG	CGGGCTCCCG CGGGCTCCCG CgggCTCCCG
4       P125FXS-FXS571R         5       215FXS-FXS571R         5       RM43-3-571R         7       RM144-3FXS-FXS-seqF         1       ref       CGG       repeat         2       P57-571R       S       P137FXS-FXS571R	CGGCGGCGGC 220 GCGCTAGCAG GCGCTAGCAG GCGCTAGCAG	230   GGCTGAAGAG GGCTGAAGAG GGCTGAAGAG	CTGGGC CTGGGC GCgGCTGGGC 240 AAGATGGAGG AAGATGGAGG AAGATGGAGG	CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC I AGCTGGTG AGCTGGTG AGCTGGTG	CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC	CTCTCGGGGG CTCTCGGGGG CTCTCGGGGG cTCTCGGGGGG	CGGGCTCCCG CGGGCTCCCG CgggCTCCCG
4       P125FXS-FXS571R         5       215FXS-FXS571R         5       RM43-3-571R         7       RM144-3FXS-FXS-seqF         1       ref       CGG       repeat         2       P57-571R       S       P137FXS-FXS571R         3       P137FXS-FXS571R       S       F75571R	CGGCGGCGGC 220 GCGCTAGCAG GCGCTAGCAG GCGCTAGCAG GCGCTAGCAG	230 J GGCTGAAGAG GGCTGAAGAG GGCTGAAGAG GGCTGAAGAG	CTGGGC CTGGGC GCgGCTGGGC 240 AAGATGGAGG AAGATGGAGG AAGATGGAGG AAGATGGAGG	CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC I AGCTGGTG AGCTGGTG AGCTGGTG AGCTGGTG	CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC	CTCTCGGGGG CTCTCGGGGG CTCTCGGGGG cTCTCGGGGGG	CGGGCTCCCG CGGGCTCCCG CgggCTCCCG
4       P125FXS-FXS571R         5       215FXS-FXS571R         5       RM43-3-571R         7       RM144-3FXS-FXS-seqF         1       ref       CGG       repeat         2       P57-571R       3       P137FXS-FXS571R         3       P137FXS-FXS571R       4       P125FXS-FXS571R         5       215FXS-FXS571R       215FXS-FXS571R	CGGCGGCGGC 220 GCGCTAGCAG GCGCTAGCAG GCGCTAGCAG GCGCTAGCAG GCGCTAGCAG	230 J GGCTGAAGAG GGCTGAAGAG GGCTGAAGAG GGCTGAAGAG GGCTGAAGAG	CTGGGC CTGGGC GCgGCTGGGC AAGATGGAGG AAGATGGAGG AAGATGGAGG AAGATGGAGG AAGATGGAGG AAGATGGAGG	CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC I AGCTGGTG AGCTGGTG AGCTGGTG AGCTGGTG AGCTGGTG	CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC	CTCTCGGGGG CTCTCGGGGG CTCTCGGGGG cTCTCGGGGGG	CGGGCTCCCG CGGGCTCCCG CgggCTCCCG
4       P125FXS-FXS571R         5       215FXS-FXS571R         5       RM43-3-571R         7       RM144-3FXS-FXS-seqF         1       ref       CGG       repeat         2       P57-571R       S       P137FXS-FXS571R         3       P137FXS-FXS571R       S       F75571R	CGGCGGCGGC 220 I GCGCTAGCAG GCGCTAGCAG GCGCTAGCAG GCGCTAGCAG GCGCTAGCAG	230 J GGCTGAAGAG GGCTGAAGAG GGCTGAAGAG GGCTGAAGAG	CTGGGC CTGGGC GCgGCTGGGC AAGATGGAGG AAGATGGAGG AAGATGGAGG AAGATGGAGG AAGATGGAGG AAGATGGAGG AAGATGGAGG	CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC CTCGAGCGCC I AGCTGGTG AGCTGGTG AGCTGGTG AGCTGGTG AGCTGGTG AGCTGGTG	CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC CGCAGCCCAC	CTCTCGGGGG CTCTCGGGGG CTCTCGGGGG cTCTCGGGGGG	CGGGCTCCCG CGGGCTCCCG CgggCTCCCG

Figure 5. The alignment of CGG repeat regions resulting from the Sequence Navigator Software. The sequences of samples (P57, P137, P125, 215, RM43-3 and RM144-3) which shown different CGG repeat numbers (21, 29, 30, 33, 36 and 43 CGG repeats, respectively) are aligned with the reference sequences (ref CGG repeat) which were downloaded from GenBank (Accession number: L29074). The boundary of the CGG repeat regions are indicated with fishhook vertical arrows. The AGG triplet is squared.

AGG Pattern	19-28 (%)	29 (%)	30 (%)	31-35 (%)	36 (%)	37-56 (%)
9A9	1 (6.3)					
9A13	2(12.5)					
9A20			1(5.6)			
9A21				1 (7.7)		
9A23				1 (7.7)		
9A25				1 (7.7)		
9A29						1 (7.7)
10A9	2(12.5)					
10A23				2 (15.4)		
10A27						1 (7.7)
10A39						1 (7.7)
11A12	1 (6.3)					
12A9	3 (18.8)					
20A9			1(5.6)			
9A9A9		15 (88.2)				
9A9A12				2 (15.4)		
9A9A15				1 (7.7)		
9A9A16					1(5.6)	
10A6A9	2(12.5)					
10A7A9	1 (6.3)					
10A9A3	1 (6.3)					
10A9A5	1 (6.3)					
10A9A9			16 (88.9)			
10A9A11				1 (7.7)		
10A12A9				1 (7.7)		
11A9A9				3 (23.1)		
12A6A9		2 (11.8)				
19A6A9					2(11.1)	
9A9A6A9					15 (83.3)	
9A9A7A9						1 (7.7)
9A9A9A9						1 (7.7)
9A11A9A9						1 (7.7)
9A9A6A6A9						4 (30.8)
9A9A6A8A9						1 (7.7)

Table 10. AGG Interruption patterns in each subgroup of control samples.

Table 10. (continued)

AGG Pattern	19-28 (%)	29 (%)	30 (%)	31-35 (%)	36 (%)	37-56 (%)
21	2(12.5)					
43						1(7.7)
56						1(7.7)
Total (95)	16 (100.0)	17 (100.0)	18 (100.0)	13 (100.0)	18 (100.0)	13 (100.0)

Note: The position of an AGG is designated by A and the number refers to the triplet length of uninterrupted CGG repeats.

# Relationship among CGG repeat numbers, haplotypes and AGG interruption patterns

When we analyzed all factors known to be responsible for repeat instability (CGG numbers, haplotypes and AGG interspersion patterns), we surprisingly found a strong association among the common CGG lengths (29, 30 and 36), the associated haplotypes (Hap A, Hap B and Hap C) and the specific AGG interspersion patterns (Figure 6 and Appendix A, Table 18). The 29 CGG repeats with Hap A showed the AGG pattern of 9A9A9 (12/17 or 70.6%), the 30 CGG repeats with Hap B showed the AGG pattern of 10A9A9 (16/18 or 88.9%) and the 36 CGG repeats with Hap C showed the AGG pattern of 9A9A6A9 (13/18 or 72.2%).

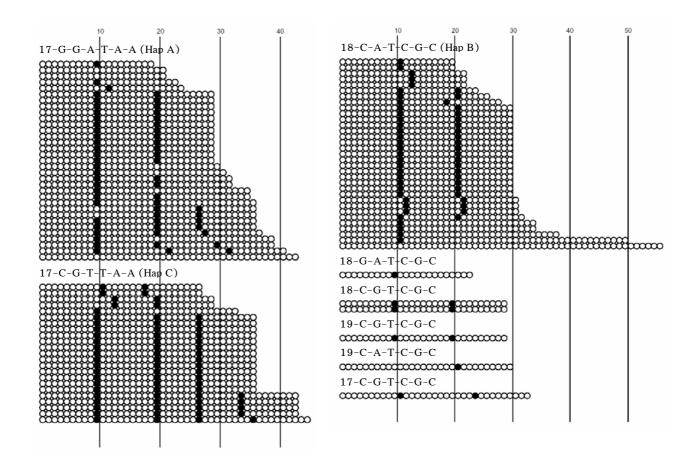


Figure 6. AGG interspersion patterns of 95 CGG alleles. The alleles are classified according to the associated FRAXAC1-WEX5-ATL1-rs25731-IVS10-rs25702-rs25723 haplotype. The AGG interspersion patterns are shown from the 5' to the 3' end. White circles represent CGG triplets and black circles represent AGG triplets. The numbers of triplets are indicated by the vertical lines.

#### Preliminary testing of DNA samples in affected families

We established the haplotype of seven markers (FRAXAC1-WEX5-ATL1-rs25731-IVS10-rs25702-rs25723) from each member of the affected families (Figure 7.1-7.2). Haplotype analysis of these members presented unsurprising results because common haplotype backgrounds (Hap A and Hap B) were found in each sample. The occurrence of recombination transmitted by maternal lineage was not observed in these affected families over the studied generations.

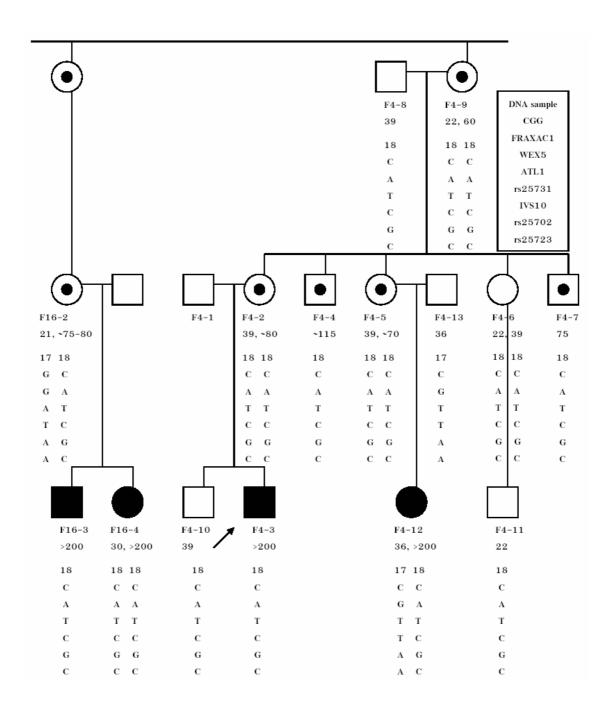


Figure 7.1. The pedigree of affected family 4 and 16. The number of CGG repeats and the results of haplotype analysis are indicated below each sample. This family had 3 generations and the inheritance of unstable chromosomes was through premutation carrier females.

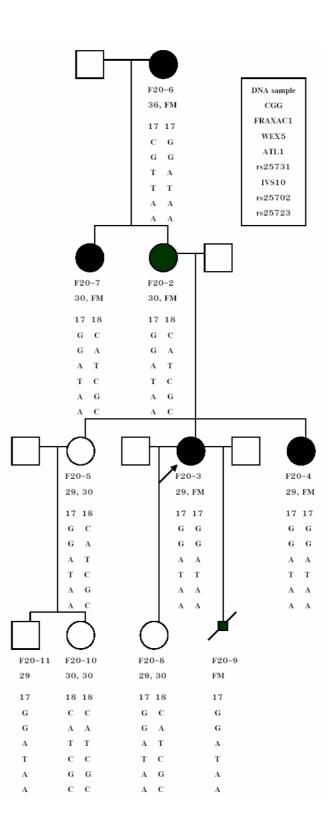


Figure 7.2. The pedigree of affected family 20. The number of CGG repeats and the results of haplotype analysis are indicated below each sample. There were 4 generations and the affected females passed the FXS mutation through the next successions.

#### Verification of microsatellite and SNP alleles

We confirmed the genotyping results of all polymorphic markers by sequencing in selected male samples. These sequencing results are shown in the right panel (square), and can be seen to be consistent with the genotyping alleles on the left panel obtained by PCR techniques as depicted in Figures 8.1–8.7 (DXS548 not analyzed).

FRAXAC1 9/8/06 10:09 am

	10	20	30	40	50	60	70
1 ref FRAXAC1	TGATCTAATC	AACATCTATA	GACTTTATTG	TGTGTGTGTĠ	TGTGTGTGTG	TGTGTATGTĠ	TGTGTCAGTC
2[ AC1-17	TGATCTAATC	AACAtCTATA	GACTTTATTG	TGTGTGTGTG	TGTGTGTGTG	TGTGTATGTG	TGTCAGTC
3[ AC1-18	TGaTCTAATC	AACAtCTATA	GACTTTATTG	TGTGTGTGTG	TGTGTGTGTG	TGTGTATGTG	TGTGTCAGTC
	8.0	90	100	11	0 120	) 130	140
1 ref FRAXAC1	TCACTCTGTC	ACTCAGGCTT	GGAGTGCAGT	GGGCAAT	1	1	1
2 <b>E</b> AC1-17	TCACTCTGTC	ACTCAGGCTT	GGAGTGCaGT	GGGCAAT			
BF AC1-18	TCACTCTGTC	ACTCAGGCTT	GGAGTGCAGT	GGGCaAT			

Figure 8.1. The alignment of AC repeat regions for FRAXAC1 using Sequence Navigator Software. The sequences of samples (AC1-17 and AC1-18) which shown different AC repeat numbers (17 and 18 AC repeats, respectively) are aligned with the reference sequences (ref FRAXAC1) which were downloaded from GenBank (Accession number: L29074). The boundary of the AC repeat regions is squared (seen as GT in the opposite strand).

W	EX5 4/22/06	3:11 am			R	-		
123	ref rs1805420 WEX5C WEX5G	GCTGG	1	GCAGCAACAT GCAGCAACAT	40 I CCTCTCATTC CCTCTCATTC CCTCTCATTC	50 TGGGGCACCT TGGGGCACCT TGGGGCACCT	60   GCCTGGGGGCA GCCTGGGGCA GCCTGGGGCA	GGTCATCCTG
4 1234		CCTCTGCCAA	CTCAGTGCTA	 TTAGTTAACT TTAGTTAACT	1	ATATTCCAGC ATATTCCAGC	TGGAATCATC TGGAATCATC	ТССССТТСТС ТССССТТСТС
1234		CACCCCAGAC	TAGGTCATGT TAGGTCATGT	TCCGCCATCA TCCGCCATCA	TGGAAGCGCC	 TATTCTTCAT TATTCTTCAT	-1	ACAGCTGCAA ACAGCTGCAA
1 2 3 4	ref rs1805420 WEX5C WEX5G	CTACTCATTT		CAATTTGATT CAATTTGATT	1	 TACTTTGCTA TACTTTGCTA	GGTACTAAGT	TCAATGCTGG TCAATGCTGG
1 2 3 4		290 I CAGTCGTTTC CAGTCGTTTC CAGTCGTTTC	300 I TTCTTTTTTT TTCTTTTTTT TTCTTTTTTT TTCTTTTTT	TTC TTC TTC TTC	320 	330 	9 340 I	) 35   

Figure 8.2. The alignment of WEX5 using Sequence Navigator Software. The sequences of samples (WEX5C and WEX5G) which shown different genotype (C and G) are aligned with the reference sequences (ref rs1805420) which were downloaded from the dbSNP of the NCBI. The asterisk (\*) indicates two distinct alleles at the SNP site (the IUB code on the reference sequences). The dash (-) indicates the consensus of all nucleotides at that position.

ATL1 9/8/06	9:16 am
1 ref ATL1 2 ATL1A 3 <mark>[</mark> ATL1G 4	10203040506070IIIIIIIITTTCTCAAATTCCAAGATTCTCATCACATTTTTTTTTCTCCCAAACTCTAAATAACCTTTTAATATTAAGTTTCTCAAATTCCAAGATTCTCATCACATTTTTTTTCTCCCCAAACTCTAAATAACCTTTTAATATTAAGTTTCTCAAATTCCAAGATTCTCATCACATTTTTTTTCTCCCCAAACTCTAAATAACCTTTTAATATTAAGTTTCTCAAATTCCAAGATTCTCATCACATTTTTTTTCTCCCCAAACTCTAAATAACCTTTTAATATTAAG
1 ref ATL1 2 ATL1A 3 <mark>E</mark> ATL1G 4	8090100110120130140IIIIIIIITATCTTTGTGGAAACATTGTTTTCTTTTCTATCCCAATTTTTAAAGCTTTTTTAAAAAAAAGAGTGCTTTATCTTGTGGAAACATTGTTTTCTTTTCTATCCCAATTTTTAAAGCTTTTTTAAAAAAAAGAGTGCTTTATCTTGTGGAAACATTGTTTTCTTTTCTATCCCAATTTTTAAAGCTTTTTTAAAAAAAAGAGTGCTTTATCTTGTGGAAACATTGTTTTCTTTTCTATCCCAATTTTTAAAGCTTTTTTAAAAAAAAGAGTGCTT
1 ref ATL1 2 ATL1A 3 <mark>[</mark> ATL1G 4	150 160 170 180 190 200 210 TTGTTGGGAT GTACATTTC CAAATGCAAA ARCATTTATG ATTCTGTGTC TCTTATAAAA TATGACACTC TTGTTGGGAT GTACATTTC CAAATGCAAA AACATTTATG ATTCTGTGTC TCTTATAAAA TATGACACTC TTGTTGGGAT GTACATTTC CAAATGCAAA AGCATTTATG ATTCTGTGTC TCTTATAAAA TATGACACTC 

Figure 8.3. The alignment of ATL1 using Sequence Navigator Software. The sequences of samples (ATL1A and ATL1G) which shown different genotype (A and G) are aligned with the reference sequences (ref ATL1) which were downloaded from the dbSNP of the NCBI. The asterisk (\*) indicates two distinct alleles at the SNP site (the IUB code on the reference sequences). The dash (-) indicates the consensus of all nucleotides at that position.

	10	20	30	40	50	60	7
			I		1	1	
	AGATTCCCAC	CTCCTGTAGG		GATATAAATG		AGTTGAGGTG	AGTTTTCCCT
/	AGATTCCCAC	CTCCTGTAGG	TTATAATAAA		AAAGTGATGA	AGTTGAGGTG	AGTTTTCCCT
731T	AGATTCCCAC	CTCCTGTAGG	TTATAATAAA	GATATAAATG	AAAGTGATGA	AGTTGAGGTG	AGTTTTCCCT
1	80	90 I	100	110	120	130	14
ref rs25731	GCCATAAAGT	CATTTAGCAC	TGAAAGAGTG	GGGTTAATT	ATCTGTGTTT	TTTTTWAATA	CTTTGTCTTT
	GCCATAAAGT	CATTTAGCAC	TGAAAGAGTG	GGGTTAATTT	ATCTGTGTTT	ттттааата	CTTTGTCTT
		CATTTAGCAC			ATCTGTGTTT	TTTTTTTAATA	CTTTGTCTT
						*	
	150	160	) 170 I	180	190	200	2
ref rs25731	AACACTGTTT	AAATTACTTT	GAGAATTACA	GCTGGAATGG	ACACGTGCTT	TTGACTAACT	CATCTTATT
731A	AACACTGTTT	AAATTACTTT	GAGAATTACA	GCTGGAATGG	ACACGTGCTT	TTGACTAACT	CATCTTATT
	AACACTGTTT	AAATTACTTT	GAGAATTACA	GCTGGAATGG	ACACGTGCTT	TTGACTAACT	CATCTTATTA
	220	230	240	250	260	270	2
	2.10			230	200	2.10	2
ref rs25731	ATAATTCAAÅ	ATGATACATĠ	ATGCTTACAT	TTGGCTATTŤ	GAGCAGTACT	CAGAGCAT	
731A	ATAATTCAAA	ATGATACATG	ATGCTTACAT	TTGGCTATTT	GAGCAGTACT	CAGAGCAT	
731т	ΔͲΔΔͲͲϹΔΔΔ	ATGATACATG	ATGCTTACAT	TTGGCTATTT	GAGCAGTACT	CAGAGCAT	

Figure 8.4. The alignment of rs25731 using Sequence Navigator Software. The sequences of samples (731A and 731T) which shown different genotype (A and T) are aligned with the reference sequences (ref rs25731) which were downloaded from the dbSNP of the NCBI. The asterisk (\*) indicates two distinct alleles at the SNP site (the IUB code on the reference sequences). The dash (-) indicates the consensus of all nucleotides at that position.

<u> 10 9/8/06 10 10 10 10 10 10 10 10 10 10 10 10 10 </u>	9:44 am						
	10	20	30	40	50	60	70
ref IVS10	GAGGCTGAAA	ATGAGAAAAA	TGTTCCACAA	GAAGAGGTAT	GTTACAGTO	GAATATTTG	TGGCACATAT
IVS10C				GAAGAGGTAT		GAATATTTTG	
IVS10T	GAGGCTGAAA	ATGAGAAAAA	TGTTCCACAA	GAAGAGGTAT	GTTACAGTGT	GAATATTTTG	
	80	90 I	100	110 I	120 	130	140
ref IVS10	AATAAAAGTA	AAAGTTTTTT	ATGTGATATĠ	TTGAGGACCT	CTAATATGTĠ	CATAAAGTGA	ATGCAAATAT
IVS10C	AATAAAAGTA	AAAGTTTTTT	ATGTGATATG	TTGAGGACCT	CTAATATGTG	CATAAAGTGA	ATGCAAATAT
IVS10T	AATAAAAGTA	AAAGTTTTTT 	ATGTGATATG	TTGAGGACCT	CTAATATGTG	CATAAAGTGA	ATGCAAATAT
	150	160	) 170 	180 I	190 	200	21
ref IVS10	TCTGATTATĊ	AAGCATGCCT	GCTGTAATTA	ATG	•	•	•
IVS10C	TCTGATTATC	AAGCATGCCT	GCTGTAATTA	ATG			
IVS10T	TCTGATTATC	AAGCATGCCT	GCTGtAATtA	ATG			

Figure 8.5. The alignment of IVS10 using Sequence Navigator Software. The sequences of samples (IVS10C and IVS10T) which shown different genotype (C and T) are aligned with the reference sequences (ref IVS10) which were downloaded from the dbSNP of the NCBI. The asterisk (\*) indicates two distinct alleles at the SNP site (the IUB code on the reference sequences). The dash (-) indicates the consensus of all nucleotides at that position.

	10	20	30	40	50	60	
	Ĩ	20	JU I	Ĩ	Ĭ	ĩ	
ref rs25702	СТСТ	TAATCCATTT	GATCCTTTCT	AGCATTTTGĠ	TTTTTTCCAG	АСАААААТСТ	GTGTCTATA
702A	CTGT	TAATCCATTT	GATCCTTTCT	AGCATTTTGG	TTTTTTCCAG	ACAAAAATCT	GTGTCTATA
702G	CTGT	TAATCCATTT	GATCCTTTCT	AGCATTTTGG	TTTTTTCCAG	ACAAAAATCT	GTGTCTATA
	80	. 90	100	110	120	130	) 1
		I	1		1	1	
		TATTGTTCCT		TTATCAAGGA			
702A	CTCTTGCCTT	TATTGTTCCT	TTTATGTCAT	TTATCAAGGA			TTCTTTGTI
702G	CTCTTGCCTT	TATTGTTCCT	TTTATGTCAT	TTATCAAGGA	AATCTAAGCA	TTTCATGAAG	TTCTTTGTT
	150	160	170	180	190	200	)
ref rs25702		CCTATTTTTC		 TATTCAAAAT			~~~~~~~~
	AAATGTTTAC		TCTTTACTGT		ACAAACTATT	TGTGTCTTCT	CATAAATGI
	AAATGTTTAC		TCTTTACTGT	TATTCAAAAT		TGTGTCTTCT	CATAAATGI
1020							
	220	230	240	250	260	270	)
	Ĩ	Ĩ	— Ĩ	1	Ĩ	21	
	CAGTTTAGTŤ	AGTGTGATGČ	RGTTATGCCT	CCTTAAAATT	TCAAACTGGA	AGATAGGAAĊ	GAGGAGGCT
				CCTTAAAATT			
702G	CAGTTTAGTT	AGTGTGATGC	GGTTATGCCT	CCTTAAAATT	TCAAACTGGA	AGATAGGAAC	GAGGAGGCI
			*				
	290	300	310	320	330	340	) :
ref rs25702		magaaaaaa		0.3.3.00.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.			
702A	CAAGGTGTAG			CAATGCAGCA CAATGCAGCA			
702G		+	+	CAATGCAGCA			
702G	CAAGGTGTAG	TGGAGCAAGC	ACTGGATGGG	CAATGCAGCA	ATTTACTTAT	AAGA	

Figure 8.6. The alignment of rs25702 using Sequence Navigator Software. The sequences of samples (702A and 702G) which shown different genotype (A and G) are aligned with the reference sequences (ref rs25702) which were downloaded from the dbSNP of the NCBI. The asterisk (\*) indicates two distinct alleles at the SNP site (the IUB code on the reference sequences). The dash (-) indicates the consensus of all nucleotides at that position.

30 40 50 60 70	0	30	20	10	
ACAAA TCAAAGTACT GAATCCTTGG TAACGAGACA TTTAAAACAC	 А.Т(	AAGATACAAA	aaaaaaaaaa	gtctgggaaa	ref rs25723
			AAAAAAAAAA		<b>7</b> 23A
CACAAA TCAAAGTACT GAATCCTTGG TAACGAGACA TTTAAAACAC	A T(	AAGATACAAA	ААААААААААА	GtcTGGGaAA	723C
	00	100	90 I	80 	
PACTTT GAAATTACAA CCATTTGGGG ATGTTTTTAG CATTTGTGCT	r Gi	AACATACTTT	MCACTACTTA	ATGCACATAĊ	ref rs25723
ACTTT GAAATTACAA CCATTTGGGG ATGTTTTTAG CATTTGTGCI	r Gž	AACATACTTT	ACACTACTTA	ATGCACATAC	723A
TACTTT GAAATTACAA CCATTTGGGGG ATGTTTTTAG CATTTGTGC	Г G2	AACATACTTT	CCACTACTTA	ATGCACATAC	723C
170 180 190 200 21	70	17(	) 160	150	
TTTTA GTTCCATTTG TCACTGTTAA CTTTCATTTG TACCTCTGGA	A G	AGTTGTTTTA	CAATATTTGT	TGAAGTAGAT	ref rs25723
STTTTA GTTCCATTTG TCACTGTTAA CTTTCATTTG TACCTCTGG	A G	AGTTGTTTTA	CAATATTTGT	TGAAGTAGAT	723A
GTTTTA GTTCCATTTG TCACTGTTAA CTTTCATTTG TACCTCTGGA	A G'	AGTTGTTTTA	CAATATTTGT	TGAAGTAGAT	723C
	40	24(	) 230	220	
GCACT TAAAATATTT TATAGCTCTT AGAACACTA	г т	CATTGGCACT	CTGTATTCAG	ATTAGCAGTG	ref rs25723
GCACT TAAAATATTT TATAGCTCTT AGAACACTA	г т	CATTGGCACT	CTGTATTCAG	ATTAGCAGTG	723A
GCACT TAAAATATTT TATAGCTCTt AGAACACTA	г т	CATTGGCACT	CTGTATTCAG	ATTAGCAGTG	
					E 723C

Figure 8.7. The alignment of rs25723 using Sequence Navigator Software. The sequences of samples (723A and 723C) which shown different genotype (A and C) are aligned with the reference sequences (ref rs25723) which were downloaded from the dbSNP of the NCBI. The asterisk (\*) indicates two distinct alleles at the SNP site (the IUB code on the reference sequences). The dash (-) indicates the consensus of all nucleotides at that position.