CHAPTER 4: RESULTS

4.1 Primers Designed and Synthesized for Microsatellite Markers (SSRs)

4.1.1 Chloroplast Genome of *Gracilaria tenuistipitata*

Eight perfect SSRs with designed primer-pairs (Table 4.1) were obtained from 183,883 base pair (bp) *Gracilaria tenuistipitata* var. *liui* chloroplast genome (AY673996). Of the eight primer-pairs, five (62.5%) were mononucleotide repeats, two (25%) were dinucleotide repeats, and only one (12.5%) was a trinucleotide repeat. All cpSSRs obtained were A, T or AT repeats. No CG microsatellites were identified.

4.1.2 Expressed Sequenced Tags (ESTs)

4.1.2.1 *Gracilaria* species

A total of 8,525 *Gracilaria* species (*G. changii*, *Gracilaria gracilis* (Stackhouse) Steentoft, Irvine, and Farnham, and *Gracilariopsis lemaneiformis* (Bory de Saint-Vincent) Dawson, Acleto & Foldvik) ESTs (Table 4.2) were screened for redundancy and resulting in 5,369 of non-redundant *Gracilaria* species ESTs. From these, 1,273 SSRs were identified in 1,027 ESTs. Only perfect SSRs were used in this study, thus the 1,273 EST-SSRs contained 880 perfect EST-SSRs with 584 (46%) mononucleotide, 170 (14%) dinucleotide, 120 (9.5%) trinucleotide, 3 (0.3%) tetranucleotide and 3 (0.3%) haxanucleotide. However, not all SSRs were suitable for primer design, only 45 (3.5%) SSRs were obtained with designed primers. Out of 45 designed primers, 33 were synthesized with 25 (75%) trinucleotide, 7 (21%) dinucleotide and 1 (3%) tetranucleotide (Table 4.3). Among the trimeric repeats that have been synthesized, the motifs GGC are the most common, followed by CGC, GAC and GCA. However, there are no common motifs for the dimeric repeats in our products.

Table 4.1: Primer-pairs derived from chloroplast genome of *G. tenuistipitata* var. *liui*.
Tm Melting temperature

Name	Forward primer (5'-3')	Reverse primer (5'-3')	Repeat motif	Predicted product size (bp)	Genomic coordinates of the SSR loci	Tm (℃)
GT1	TTTATCAACGATCCCTGTAG	AATGGACTGTAATTCACCAA	$(T)_{10}$	336	4314 – 4323	52
GT2	TTTTTGAGCGATATTTTGAC	AGAATAAGACCACCTGAACC	$(T)_{10}$	310	78590 – 78599	53
GT3	CCATAATGGAGATCTGTTTG	CTGGCAACATAGTTAGCATT	$(T)_{10}$	306	84904 – 84913	53
GT4	AGCAATCCTAAATTGACAAC	AAGGTAGACCAGGAGAAAAA	$(ATT)_5$	295	99538 – 99552	53
GT5	AGAAATTGATCAAGCTGTT	TTTTCAGCAATTGGAGTATC	$(AT)_8$	329	105129 – 105144	52
GT6	CCTACAATCAGAATGGAATG	AGCTTCCAAGAAAAATGAGT	$(A)_{10}$	303	143312 – 143321	53
GT7	ATCCTTCTTTTAAGCCGTAG	TCTTCCATGAAGTCTTCTTTT	$(T)_{10}$	309	161988 – 161997	53
GT8	CTCCTGACATGATAAACACC	CTTCGATTTGTTTAATGAGC	$(AT)_7$	288	179473 - 179486	52

Table 4.2: Number of ESTs for three *Gracilaria* species (*G. changii*, *G. gracilis* and *Gp. lemaneiformis*) downloaded from GenBank.

***Gracilaria* species**	**Number of ESTs**	**Total numbers of ESTs**
Gracilaria changii	8147	8525
Gracilaria gracilis	200	
Gracilariopsis lemaneiformis	178	

4.1.2.2 *Gracilaria tenuistipitata*

A total of 4,006 *Gracilaria tenuistipitata* var. *liui* ESTs downloaded from GenBank were screened for redundancy and searched for the SSR markers. Only perfect SSRs were used in this study and not all SSRs were suitable for primer design, thus five (50%) di-nucleotide and five (50%) tri-nucleotide were obtained with designed primers. All the SSR markers that we developed were unique with no common motifs observed for the dimeric or trimeric repeats (Table 4.4).

Table 4.3: Primer-pairs derived from expressed sequence tags (ESTs) of *Gracilaria* species (*G. changii*, *G. gracilis* and *Gp. lemaneiformis*).
Tm Melting temperature

No.	Primer Name	Microsatellite Repeats	Primers(5'-> 3')	Tm(℃)
P3	GC_Contig1289	$(CCA)_8$	F:CATATAATTGGGCACAAGAG R:CATCAGACGCTTGAATAGAT	53
P5	GC_Contig752	$(GCG)_{11}$	F:ATTGAATGCATTACCAATGT R:TGAATGACTGGTTTCAGATT	52
P6	GC_Contig969	$(GTGA)_7$	F:GATTGATCGGAGCATAGC R:TTTCTTTTGTAAAGCATTGG	54
P7	GC_Contig1209	$(CG)_6$	F:CCACCAGAACGGAGTTAT R:ACAACAGTAGCAATCAGCTT	53
P8	GC_Contig1477	$(TG)_{12}$	F:CTGCACGAGAGTGAGAAG R:GCTTGCTCATCAATGATAC	52
P9	GC_Contig1364	$(AGA)_5$	F:GAGCGATGATATCAAGAAAA R:GAGCGATGATATCAAGAAAA	53
P10	GC_Contig546	$(GGC)_5$	F:ACTGGTATGGGTTTCGAG R:ATCTTTACCTCTCCTGTCGT	53
P11	GC_Contig673	$(GA)_6$	F:AGACAGAAGGGAGAAGAGG R:AGAGGAGAACTGACAAGGAT	53
P12	GC_Contig761	$(GAC)_5$	F:GAACATCTTCACTCTCAACG R:TTCTCCTTCAATAGTGGTGA	53
P13	GC_Contig1171	$(GT)_9$	F:TTTCTGTTTTCCTTGTGTGT R:ACAGCGCACCTACTTAATAC	53
P14	GC_Contig1110	$(GC)_6$	F:GTCAGTTCATATGCCGAGT R:GGCTCTTTATATGAGAGGGT	53
P15	GC_Contig1241	$(AGC)_5$	F:ACAAGGACCCTAACTACGAC R:ATCAAATCTGCAGAAAAGC	53
P16	GC_Contig1232	$(CCG)_8$	F:CATATCATAGCCGTTTTACG R:TGTTAGAGCTGGTGGTAAAG	53
P17	GC_Contig792	$(AT)_6$	F:GGTGATGTGTACACCTTTGT R:GTATGTGCAATGAGGTTTTC	53
P18	GC_Contig814	$(CGG)_9$	F:TGTGAATCAGACGGTGTC R:GTAAAGCGTAGCGTTCAGT	53
P19	GC_Contig645	$(GAC)_{12}$	F:GTGAACTAGGTAGTTTGGCA R:CAAGAAATGACAAGCAAGAT	52
P20	GC_Contig569'	$(GGC)_5$	F:GTTTTTCCACCATATCCAG R:ACCACCAGGCAGATACTC	52
P21	GC_Contig607	$(GCA)_5$	F:CCTCTTCTCAATATGAGCG R:AGATTGGTAGAGGAGGTACG	53

Table 4.3, continued.

No.	Primer Name	Microsatellite Repeats	Primers(5'-> 3')	Tm(℃)
P22	GC_Contig607'	$(CCG)_{10}$	F:CCTCTTCTCAATATGAGCG R:AGATTGGTAGAGGAGGTACG	53
P23	GC_Contig703	$(CT)_8$	F:AAACACTCTCCGAACAACTA R:AATATTGCGAAAATGATAGC	52
P24	GC_Contig574	$(AAC)_9$	F:GTACATGCTCCAAACTTCC R:GTTCATGAGCAGCTTGAG	53
P25	GC_Contig1353	$(GTG)_5$	F:TGAAGATTTGTATTTGGTGG R:GCGTGTACTAATGAGGCTT	53
P26	GC_Contig1386	$(CAA)_7$	F:CGACTACTCCAAGACGTTC R:GTTTTCAACTACCTTTGCTG	53
P27	GC_Contig1220	$(CGC)_5$	F:AAATACTGGTCACGGAAGA R:CTGTTAGCGAATGATTTAGG	53
P28	GC_Contig1254	$(CCT)_7$	F:ACAAGTGGAGAAGCAGAAG R:GAAGGCAGAAAATAGAACCT	52
P29	GC_Contig1129	$(CGC)_5$	F:ACATGTCTTCGCTTGCTC R:CTGTGGTTGAAGACGAAAC	54
P30	GC_Contig698	$(GCA)_6$	F:CTTTCTGTTTTGAGCTCTGT R:ATGCTGTTACTGCGAATG	52
P31	GC_Contig1211	$(CGC)_5$	F:ACAGTGTGTTGTACTCCTGC R:AGTGTTGCAAGTGAAGACTG	53
P32	GC_Contig800	$(CGC)_8$	F:CATACGCTCCGTCAACTA R:GAAGTCGAAGTCTGGCAA	53
P33	GC_Contig972	$(GGC)_6$	F:ATAACACCAGTTCAGTGCC R:TTTTTAAAATTAAGGTGCCA	53
P34	GC_Contig583	$(GGC)_5$	F:GTACATGGGATCACGTCTT R:AAGTGGAGTGGACAAATGTA	53
P35	GC_Contig569	$(GGC)_5$	F:CTTATCCTTGATCATACCGA R:CTGGATATGGTGGAAAAAC	52
P36	GC_Contig785	$(ACA)_5$	F:CCTTCAACCACTACAAACAG R:GTCGTCTTCATCATCTTCAC	53

Table 4.4: Primer-pairs derived from ESTs of *G. tenuistipitata* var. *liui*.
Tm Melting temperature

Name	Forward primer (5'-3')	Reverse primer (5'-3')	Repeat motif	Predicted product size (bp)	Tm (℃)
GT9	ACTGATATACCCTTCGGATT	AAGATACCCGGACTGTTG	$(GGC)_5$	329	53
GT10	TCTAGAGGCTGAATGAAAGA	ATGAATCTAATCTGGTGTGC	$(AG)_8$	319	53
GT11	ACTTTAAGATGGTGACATGG	TGGAGAATAATTGAGCAAAC	$(GT)_9$	340	53
GT12	AAACATGTCATTCAAGTGGT	GGAAAAATAGTTGAGGAAGG	$(CAT)_5$	308	53
GT13	CATACGTTCACAGGGCTAC	ATTACGTACTCAGCTTCGTG	$(AT)_6$	348	53
GT14	CTCCACTAACACTTCAGTCG	TTGAGAAGAGAGTTCTGTCG	$(TGC)_5$	297	53
GT15	CACGGACTTATGGTAATTTG	CCATATGGTCCAGTACCTTT	$(CT)_6$	341	54
GT16	CAGTCTTCTGAGTATCACGG	AACTGTTGAAATGCTACCTG	$(TA)_8$	321	53
GT17	GTTCCCAGCTAGACTTCAG	GAAGCGGAGAGGAATTAG	$(CCG)_5$	303	52
GT18	AGGAATTTAACTTTTCCGAC	AACAAGTCTATCACAGCGAG	$(CAA)_9$	282	53

4.2 Fragment Analysis (FA) Results

4.2.1 Data Analysis for SSRs from Chloroplast Genome of *Gracilaria tenuistipitata*

4.2.1.1 Populations of *Gracilaria tenuistipitata* from Various Localities

All the primer-pairs were tested with eighty specimens of *G. tenuistipitata* from six different localities. The eight primer-pairs showed good amplification in 3% MetaPhor agarose gel electrophoresis (Figure 4.1 and 4.2).

Mononucleotide primer-pairs (i.e. GT1, GT2, GT3, GT6 and GT7) and trinucleotide primer-pair (i.e. GT4) showed good amplification but was monomorphic on the samples tested (Figure 4.3). Only one defined amplified fragment size, 336 bp for primer-pair GT1, 312 bp for primer-pair GT2, 308 bp for primer-pair GT3, 301 bp for primer-pair GT6 and 307 bp for primer-pair GT7 was derived, and the fragment results indicated that all samples fell within the same defined amplified fragment size (Table 4.5).

For the primer-pair GT5, three defined amplified fragment sizes, 327 bp, 329 bp and 333 bp were derived and samples with similar base size peaks fell into the same amplified fragment size. The dendrogram of primer-pair GT5 indicated that the *G. tenuistipitata* specimens were grouped into three main clades: (a) *G. tenuistipitata* from Kelantan, Malaysia; (b) *G. tenuistipitata* from Quy Kim, Hai Phong, Vietnam and Pattani, Thailand; and (c) *G. tenuistipitata* from Kuah, Batu Laut and Middle Banks (Malaysia). All three clades were supported with similarity coefficients of 0.5. The support value of 0.5 is the 50% similarity between two clades (Figure 4.4).

For the primer-pair GT8, two defined amplified fragment sizes, 284 bp and 295 bp, were derived, and samples with similar amplified fragment sizes exhibited a cell value in that column. The dendrogram of primer-pair GT8 indicated that the *G. tenuistipitata* specimens were grouped into two main clades: (a) *G. tenuistipitata* from Batu Laut, Middle Banks and Kuah (Malaysia); and (b) *G. tenuistipitata* from Pattani (Thailand), Quy Kim, Hai Phong (Vietnam), and Kelantan (Malaysia). Both clades were also supported with similarity coefficients of 0.5 (Figure 4.5).

The combined dataset from the two primer-pairs (i.e. GT5 and GT8) produced a dendrogram with similarity coefficients ranging from 0.17 to 0.5. From the generated dendrogram, the eighty specimens of *G. tenuistipitata* were grouped into two major clades: Clade A, consisting of *G. tenuistipitata* from Batu Laut, MiddleBanks and Kuah (Malaysia); and Clade B, consisting of *G. tenuistipitata* from Pattani (Thailand), Quy Kim, Hai Phong (Vietnam), and Kelantan (Malaysia). Clade B is further divided into two sub-clades: Clade B1, consisting of *G. tenuistipitata* from Pattani (Thailand) and Quy Kim, Hai Phong (Vietnam); and Clade B2, consisting of *G. tenuistipitata* from Kelantan (Malaysia). Both clades B1 and B2 were supported with a similarity coefficient of 0.33 (Figure 4.6).

The sequencing results indicated that the amplification products of these primer-pairs (i.e. GT1, GT2, GT3, GT4, GT5, GT6, GT7 and GT8) were generated because of polymorphisms and not indels in the regions flanking the cpSSRs (Appendix A1 - A7).

Figure 4.1: Gel picture shows the differences in size using different markers (primer-pairs GT4, GT5 and GT8) on some of the *Gracilaria tenuistipitata* from Batu Laut, Selangor, Malaysia.
LRL: Low Range Ladder

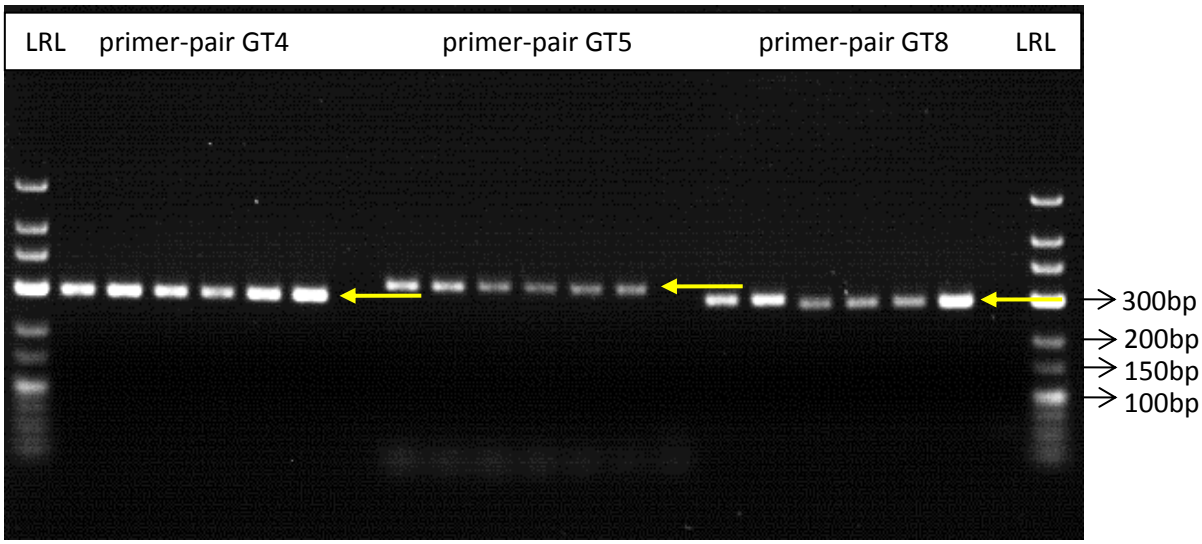


Figure 4.2: Gel picture shows the differences in size using primer-pair GT8 on *Gracilaria tenuistipitata* from different populations (PKB: Middle Banks, Penang, Malaysia; Thailand: Pattani, Thailand).
LRL: Low Range Ladder; -ve: Negative control

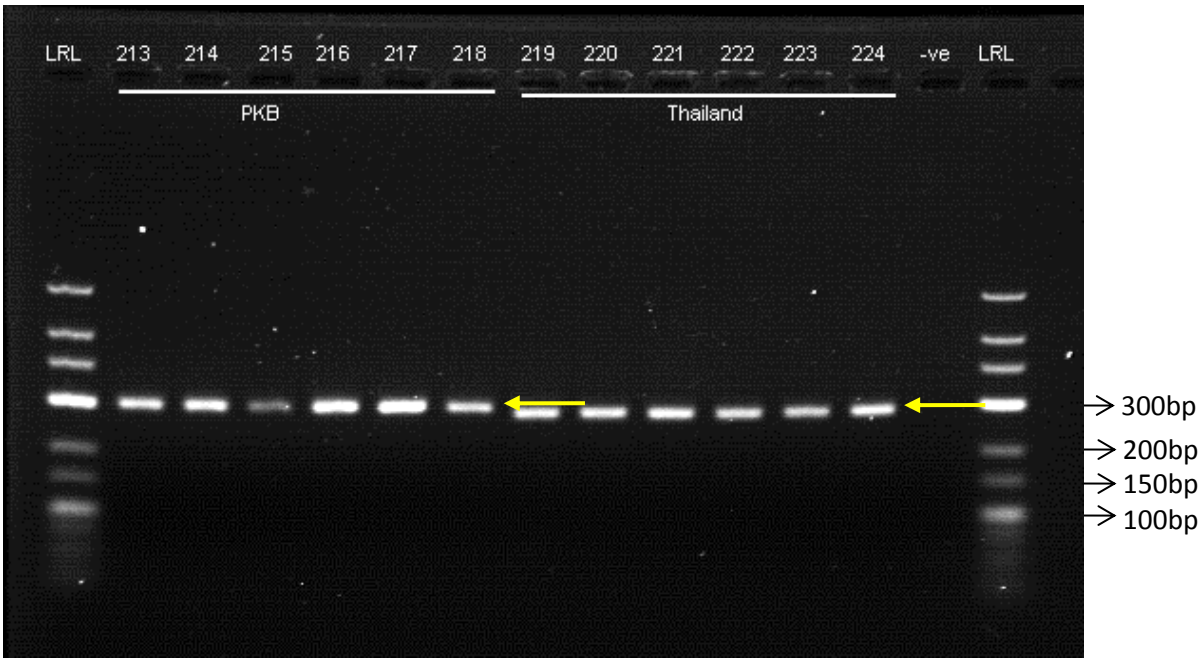


Figure 4.3: Gel picture shows good amplification but monomorphic using primer-pair GT1 on *Gracilaria tenuistipitata* from different localities (92-93, 269-277, TW: Batu Laut, Selangor, Malaysia; 216-218, 237-238, 247-248, 253-254, 278-280: Middle Banks, Penang, Malaysia).
LRL: Low Range Ladder

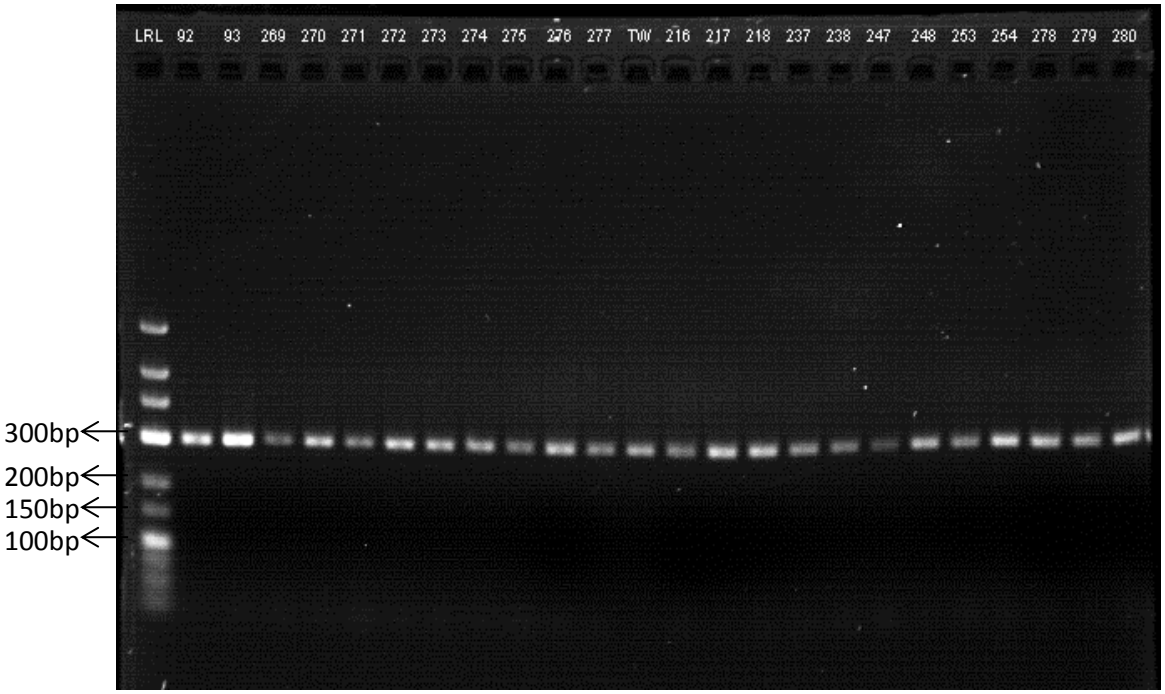


Table 4.5: Amplified allele size of each *G. tenuistipitata* from six different localities for primer-pairs GT4, GT5 and GT8.

1 = Presence of SSR product
0 = Absence of SSR product

Location	Number of Samples Tested	Primer-pair GT4	Primer-pair GT5			Primer-pair GT8	
		Allele (295 bp)	Allele (327 bp)	Allele (329 bp)	Allele (333 bp)	Allele (284 bp)	Allele (295 bp)
Batu Laut, Selangor, Malaysia	15	1	1	0	0	0	1
Middle Banks, Penang, Malaysia	15	1	1	0	0	0	1
Quy Kim, Hai Phong, Vietnam	15	1	0	1	0	1	0
Pattani, Thailand	15	1	0	1	0	1	0
Kelantan, Malaysia	10	1	0	0	1	1	0
Kuah, Pulau Langkawi, Malaysia	10	1	1	0	0	0	1

Figure 4.4:

Figure 4.5:

Figure 4.6:

4.2.1.2 Species Differentiation in *Gracilaria*

Further study was carried out using one of the primer-pair (i.e. GT5) generated from the chloroplast genome of *G. tenuistipitata* to test on other *Gracilaria* species, such as *G. changii*, *G. salicornia*, *G. cacalia*, *G. crassa* and *G. canaliculata* from Malaysia. Results indicated that the specimens were grouped into five main clades: (a) *G. tenuistipitata* from Kelantan, Malaysia; (b) *G. tenuistipitata* from Vietnam and Thailand; (c) *G. tenuistipitata* from Batu Laut, Middle Banks and Kuah, Malaysia; (d) *G. changii* from Morib, Malaysia; and (e) *G. canaliculata*, *G. crassa*, *G. salicornia* and *G. cacalia* from Malaysia. All five clades were supported with similarity coefficients of 0.5. The support value of 0.5 is the 50% similarity between two clades (Figure 4.7).

4.2.2 Data Analysis for SSRs from ESTs of *Gracilaria* Species (*G. changii*, *G. gracilis* and *Gp. lemaneiformis*)

4.2.2.1 *Gracilaria changii* from Different Populations

Gracilaria changii collected at various localities in Malaysia (e.g. Selangor, Malacca, Penang, Kedah, Negeri Sembilan, Sabah and Sarawak) were analysed with 33 EST-SSR primer-pairs. Based on the variations in repeated nucleotides of microsatellite marker, one primer-pair, that is P3 showed significant results that can distinguish *Gracilaria changii* from various localities. The other 32 primer-pairs were either produced null alleles or did not show polymorphism among the *Gracilaria* specimens tested and hence they were excluded from data analysis. There is no difference in the SSR results for female gametophytes and tetrasporophytes. The UPGMA dendrogram analysis grouped *Gracilaria changii* into three main clades: (a) *G. changii* from Batu Besar (Malacca), Sandakan (Sabah), Bintulu (Sarawak), Batu Tengah (Malacca), Gua Tanah (Malacca), Middle Banks (Penang), Sg. Merbok (Kedah), Teluk Pelandok

(Negeri Sembilan), Pantai Dickson (Negeri Sembilan), Sg. Kong-Kong (Johor), and Sg. Pulai (Johor); (b) *G. changii* from Morib (Selangor); and (c) *Gracilaria changii* from Pantai Dickson (Negeri Sembilan), Gua Tanah (Malacca), Sg. Merbok (Kedah), Sg. Kong-Kong (Johor) and Sg. Pulai (Johor). Three different genotypes of *Gracilaria changii* were obtained and *G. changii* from Morib, (Selangor) has its unique allele that can be distinguished from other populations. All three clades were supported with similarity coefficients of 0.5. The support value of 0.5 is the 50% similarity between two clades (Figure 4.8).

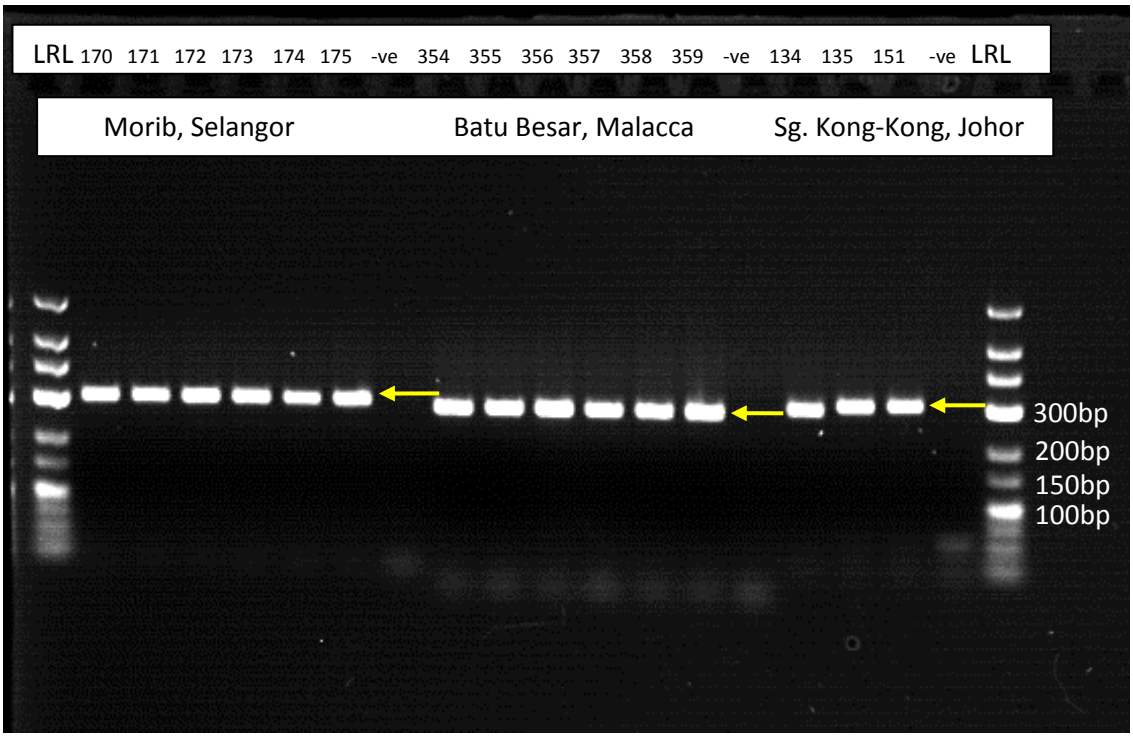
4.2.2.2 *Gracilaria* Species

Other than the analysis of *Gracilaria changii* from Malaysia, the 33 primer-pairs were also used to distinguish between different *Gracilaria* species: *Gracilaria manilaensis* from Cebu, Philippines, *Gracilaria fisheri* from Pattani, Thailand, *G. edulis* and *G. salicornia* from Malaysia. The same primer-pair P3 showed good amplification and polymorphism on *Gracilaria manilaensis* and *Gracilaria fisheri* using 3% Metaphor agarose gel electrophoresis (Figure 4.9). However, *G. edulis* and *G. salicornia* were either produced null allele or faint band even repeated PCR amplification was conducted. Amplification with primer-pair P3 generated products of five amplified fragment sizes: 326 bp, 329 bp, 331 bp, 334 bp and 337 bp and the samples with similar base size peaks fell into the respective column (Table 4.6).

Figure 4.7:

Figure 4.8:

Figure 4.9: Gel picture shows the differences in size using primer-pair P3 on *Gracilaria changii* from different localities (170-175: Morib, Selangor; 354-359: Batu Besar, Malacca; 134, 135, 151: Sg. Kong-Kong, Johor, Malaysia).
LRL: Low Range Ladder; -ve: Negative control



The dendrogram of primer-pair P3 showed that *Gracilaria* species were grouped into five main clades: (i) Clade A: *Gracilaria changii* from Batu Besar, Batu Tengah & Gua Tanah (Malacca), Sandakan (Sabah), Bintulu (Sarawak), Middle Banks (Penang), Sg. Merbok (Kedah), Teluk Pelandok & Pantai Dickson (Negeri Sembilan), Sg. Kong-Kong & Sg. Pulai (Johor); (ii) Clade B: *G. manilaensis* from Cebu, Philippines; (iii) Clade C: *G. changii* from Morib, Selangor; (iv) Clade D: *G. fisheri* from Pattani, Thailand; and (v) Clade E: *G. changii* from Pantai Dickson (Negeri Sembilan), Gua Tanah (Malacca), Sg. Merbok (Kedah), Sg. Kong-Kong and Sg. Pulai (Johor). All five clades were supported with similarity coefficients of 0.5 (Figure 4.10).

Table 4.6: Amplified allele sizes obtained from Malaysian *Gracilaria changii*, *Gracilaria manilaensis* and *Gracilaria fisheri* using primer-pair P3, listed by species and location.
The numeric number in bracket after the allele size is the number of samples.

Species Name	Locations	Amplified Allele Size (bp)	
Gracilaria changii	Morib, Selangor, Malaysia	337 (10)	
	Middle Banks, Penang, Malaysia	331 (10)	
	Batu Besar, Malacca, Malaysia	331 (10)	
	Batu Tengah, Malacca, Malaysia	331 (10)	
	Gua Tanah, Malacca, Malaysia	331 (4)	334 (6)
	Sg. Merbok, Kedah, Malaysia	331 (4)	334 (4)
	Teluk Pelandok, Negeri Sembilan, Malaysia	331 (10)	
	Pantai Dickson, Negeri Sembilan, Malaysia	331 (8)	334 (2)
	Sg. Kong-Kong, Johor, Malaysia	331 (1)	334 (8)
	Sg. Pulai, Johor, Malaysia	331 (9)	334 (1)
	Sandakan, Sabah, Malaysia	331 (10)	
	Bintulu, Sarawak, Malaysia	331 (10)	
Gracilaria manilaensis	Cebu, Philippines	329 (10)	
Gracilaria fisheri	Pattani, Thailand	326 (10)	

Figure 4.10:

4.2.2.3 Comparison with *cox*1 Gene Marker

The same primer-pair P3 was also used to test with *Gracilaria changii* from the west coast of Peninsular Malaysia in order to compare the variability of this marker and the commonly used marker, mitochondrial cytochrome c oxidase subunit I (*cox*1). In the previous study, six haplotypes were obtained with haplotype C1 showed to be the basal haplotype as it was the most widespread haplotype in all the different localities. These results were published in 2011 by Yow et al. and the DNAs used in the former study were used in this study with the courtesy of Yow et al., 2011. The microsatellite marker that we developed was able to separate the *Gracilaria changii* into three main clades: (a) *G. changii* from Gua Tanah, Middle Banks, Batu Besar, Batu Tengah, Sg. Pulai, Sg. Merbok, and Teluk Pelandok; (b) *G. changii* from Gua Tanah and Sg. Merbok; and (c) *G. changii* from Morib (Figure 4.11). This showed that three genotypes of *G. changii* were obtained using the primer-pair P3 and the common genotype was M1 as specimens collected from seven out of eight localities shared the same genotype. However, two different genotypes were obtained, namely M1 and M2, for *Gracilaria changii* from Gua Tanah, Malacca and Sg. Merbok, Kedah (Table 4.7).

Figure 4.11:

Table 4.7: Three different genotypes of *G. changii* were obtained based on the results from primer-pair P3. The numeric number in bracket after the allele size is the number of samples.

Species Name	Locations	Amplified Allele Size (bp)		
Gracilaria changii	Morib, Selangor			337 (10)
	Middle Banks, Penang	331 (11)		
	Batu Besar, Malacca	331 (8)		
	Batu Tengah, Malacca	331 (11)		
	Gua Tanah, Malacca	331 (4)	334 (5)	
	Sg. Merbok, Kedah	331 (3)	334 (4)	
	Teluk Pelandok, Negeri Sembilan	331 (4)		
	Sungai Pulai, Johor	331 (6)		

4.2.3 Data Analysis for SSRs from ESTs of *Gracilaria tenuistipitata*

All the 10 primer-pairs derived from ESTs of *G. tenuistipitata* were tested on eighty specimens of *G. tenuistipitata* from six different localities. Only two primer-pairs (i.e. GT12 and GT18) showed good amplification in 3% MetaPhor agarose gel electrophoresis. The other eight primer-pairs either produced null alleles or did not show polymorphism among the *G. tenuistipitata* specimens tested and hence they were excluded from data analysis. The dendrogram for primer-pair GT12 indicated that the *Gracilaria tenuistipitata* were grouped into two main clades: (a) *G. tenuistipitata* from Batu Laut, Middle Banks, Kelantan (Malaysia), Pattani (Thailand) and Quy Kim, Hai Phong (Vietnam); and (b) *G. tenuistipitata* from Kuah, Malaysia (Figure 4.12). For the primer-pair GT18, the dendrogram indicated that *G. tenuistipitata* were grouped into three main clades: (i) Clade A: *G. tenuistipitata* from Middle Banks, Batu Laut and Kuah (Malaysia); (ii) Clade B: *G. tenuistipitata* from Pattani (Thailand); and (iii) Clade C: *G. tenuistipitata* from Quy Kim, Hai Phong (Vietnam) and Kelantan (Malaysia). All three clades were supported with similarity coefficients of 0.5 (Figure 4.13).

The combined dataset from the two primer-pairs (Table 4.8) generated a dendrogram with similarity coefficients ranging from 0.1 to 0.43. From the generated dendrogram, the eighty specimens of *G. tenuistipitata* were grouped into two major clades: Clade A was further divided into three sub-clades: Clade A1 consisting of *G. tenuistipitata* from Batu Laut and MiddleBanks (Malaysia); Clade A2 consisting of *G. tenuistipitata* from Pattani (Thailand); and Clade A3 consisting of *G. tenuistipitata* from Quy Kim, Hai Phong (Vietnam) and Kelantan (Malaysia). All three clades were supported with similarity coefficients of 0.33. Clade B consisting of *G. tenuistipitata* from Kuah (Malaysia) and was supported with similarity coefficient of 0.43 (Figure 4.14).

Table 4.8: Amplified allele size of each *G. tenuistipitata* from six different localities for primer-pairs GT12 and GT18.

1 = Presence of SSR product
0 = Absence of SSR product

Location	**Number of Samples Tested**	**Primer-pair GT12**		**Primer-pair GT18**		
		Allele (307 bp)	**Allele (310 bp)**	**Allele (269 bp)**	**Allele (275 bp)**	**Allele (281 bp)**
Batu Laut, Selangor, Malaysia	15	1	0	1	0	0
Middle Banks, Penang, Malaysia	15	1	0	1	0	0
Quy Kim, Hai Phong, Vietnam	15	1	0	0	1	0
Pattani, Thailand	15	1	0	0	0	1
Kelantan, Malaysia	10	1	0	0	1	0
Kuah, Pulau Langkawi, Malaysia	10	0	1	1	0	0

Figure 4.12:

Figure 4.13:

Figure 4.14:

CHAPTER 5: DISCUSSION

5.1 Development of Microsatellite Markers (SSRs) from GenBank

Eight SSR primer-pairs including mononucleotide repeats developed from the chloroplast genome of *G. tenuistipitata* var. *liui* showed a relatively lower number of cpSSRs compared to plants. In plants, the highest frequency of cpSSRs was 160 found in *Arabidopsis thaliana*, and the lower numbers of cpSSRs were found in *Nymphaea alba*, the water lilies with 35 and in *Nuphar advena* with 39 (Ebert & Peakall, 2009). Studies have suggested that the cpSSRs are abundant in most plants (Powell et al., 1995a,b; Provan et al., 2001) but less abundant in algae. This finding is consistent with our analysis on SSRs identified from the 23 algae chloroplast genomes deposited at the National Center for Biotechnology Information (NCBI) database using the same SSRs mining method as in 3.2 (not shown in results) (Table 5.1). The analysis indicated that approximately 50% of the chloroplast genome contains low numbers of repetitive sequences, particularly on SSRs. *Ectocarpus siliculosus, Parachlorella kessleri, Monomastix* sp., *Scenedesmus obliquus, Chlorella vulgaris* were found to have one SSR; *Bryopsis hypnoides, Micromonas* sp., *Cyanidioschyzon merolae* and *Cyanidium caldarium* were found to have two SSRs; *Gracilaria tenuistipitata* var. *liui* and *Porphyra yezoensis* have three SSRs; and only *Leptosira terrestris* has more than 60 SSRs. These results show that *G. tenuistipitata* var. *liui* used in this study is in the category of low numbers of repetitive sequences.

Of the eight primer-pairs we designed, more than 50% of the total cpSSRs obtained was single nucleotide repeat. This is congruent with the initial finding of cpSSRs by Powell et al. that single nucleotide repeat predominate, with numerous A or T mononucleotide microsatellites were found in all six plant species (Powell et al., 1995a,b). Our study showed that microsatellites with $(AT)_n$ sequences were commonly

found in genomic libraries of algae which also supports the findings that the most abundant SSR motifs in plants are A/T rich (Morgante & Olivieri, 1993; Wang et al., 1994; Cardlea et al., 2000). This result may be observed because of the excess of AT repeats and deficiencies of AC/TG and CG repeats. C or G mononucleotide repeats were rare or absent (Powell et al., 1995; Ebert & Peakall, 2009).

For the microsatellite markers developed from the ESTs of *Gracilaria* species, a total of 3,166 (37%) redundant ESTs were screened and excluded from the SSR analysis. Non-redundant set of EST were used to ensure better quality and length of reads for the identification and development of EST-SSR markers by obtaining a consensus sequence after EST sequence assembly (Kantety et al., 2002). In some studies, the redundant EST sequences were scanned for the presence of ESTs containing SSRs (SSR-ESTs) before the smaller dataset of redundant SSR-ESTs used to identify non-redundant SSR-ESTs or EST-SSRs after clustering and defining the unigene SSR-ESTs. For example, the study of SSR markers in barley showed that out of 1,453 microsatellites containing ESTs with designed primers, 945 ESTs belonging to 244 clusters and 508 ESTs being singletons. Thus, 752 unique sequences were potential candidates for SSR marker development (Thiel et al., 2003). Clustering of ESTs reduces the redundancy in dataset and more accurately reflects the density of SSRs in the transcribed portion of the genome (Kantety et al., 2002; Varshney et al., 2005). Using the SSRs pipeline that we developed, 5,369 of non-redundant ESTs were obtained with 1,027 ESTs being contigs and 4,342 ESTs being singletons. These unique sequences were then used to screen for the microsatellite markers. About 1.24% of the *Gracilaria* species (*Gracilaria changii, Gracilaria gracilis* and *Gracilariopsis lemaneiformis*) ESTs contained SSRs. In *Porphyra* species, approximately 6.02% of *Porphyra haitanensis* ESTs contained SSRs (Xie et al., 2009) and 2.1% for *P. yezoensis* (Sun et al., 2006). Our study showed a lower percentage compared to *Porphyra* species, barley

(3.4%), rice (4.7%) and wheat (3.2%) but approximately similar with the maize (1.5%) (Kantety et al., 2002).

Not all SSRs were suitable for primer design, some SSRs were located too close to the end of the flanking region for primer design, or the base composition of the flanking sequence was unsuitable (Sun et al., 2006), or the EST sequences containing microsatellites are too short (Varshney et al., 2002). To obtain useful primer-pairs for PCR amplification, a few parameters (e.g. oligonucleotide melting temperature, size of the primer-pairs, GC content and PCR product size) were defined during the designing of the primer-pairs. The size of the primer-pairs flanking the region of SSRs should be at least 20 bp to avoid non-specific priming (Schlötterer & Pemberton, 1994) and the primer-pairs were recommended to have 50 – 60% of GC content with similar melting temperature. The oligonucleotide melting temperature is influenced by the base composition, thus primer-pairs with high GC content can be annealed at higher temperature (melting temperature increases with GC content) with greater specificity than the primer-pairs with low GC content. The annealing temperature should be high enough to prevent primers from annealing to non-target sequences but low enough to ensure the primers are able to anneal appropriately. The amplification product size shouldn't be too long as longer amplification products required longer gel running times to resolve unambiguously and they are difficult to score (Scribner & Pearce, 2000). Hence, our parameters for designing the primer-pairs were: 48℃ to 55℃ for primer melting temperature with primer length within 20 – 24 bp; GC content between 50% - 70% and the expected product size between 280 bp and 350 bp. Consequently, only 45 (3.5%) SSRs were obtained with designed primer. Of the 45 EST-SSRs, 33 primer-pairs with longer repeats were chosen for analyses as they were shown in the literature review by Varshney et al. that higher numbers of tandem repeats were found to be more polymorphic (Varshney et al., 2002).

Our results showed that trinucleotides were the most common type of SSR which are similar to the EST-SSRs study on *Porphyra haitanensis* with 64.29% of trinucleotides was obtained (Xie et al., 2009). The second most common type of SSRs was dinucleotide followed by tetranucleotide. In agreement with Varshney et al. (2002) and Xie et al. (2009), the sequence GGC was the most frequent trinucleotide repeat. However, the study on microsatellite loci in *Drosophila melanogaster* showed that dinucleotide repeats were more abundant (66%) than trinucleotide (30%) or tetranucleotide (4%) (Schug et al., 1998). Although in the study of eggplant by Stàgel et al. showed that trinucleotides were the most frequent repeats (51.4%), the percentage of dinucleotides (15.7%) were lower than the mononucleotide repeats (28.6%) (Stàgel et al., 2008).

Table 5.1: Number of SSRs generated for 23 algae chloroplast genome.
P2 – P6: Primer with nucleotide repeats from two to six.

No.	Species Name	Accession Number	Number of SSRs (P2-P6)
1.	*Ectocarpus siliculosus*	FP102296	1
2.	*Floydiella terrestris*	GU196268	13
3.	*Bryopsis hypnoides*	GQ892829	2
4.	*Parachlorella kessleri*	FJ968741	1
5.	*Micromonas sp.*	FJ858267	2
6.	*Micromonas pusilla*	FJ858269	0
7.	*Monomastix sp.*	FJ493497	1
8.	*Pycnococcus provasolii*	FJ493498	0
9.	*Pyramimonas parkeae*	FJ493499	2
10.	*Oedogonium cardiacum*	EU677193	11
11.	*Leptosira terrestris*	EF506945	62
12.	*Stigeoclonium helveticum*	DQ630521	4
13.	*Scenedesmus obliquus*	DQ396875	1
14.	*Oltmannsiellopsis viridis*	DQ291132	0
15.	*Pseudendoclonium akinetum*	AY835431	0
16.	*Chlamydomonas reinhardtii*	BK000554	7
17.	*Chlorella vulgaris*	AB001684	1
18.	*Porphyra yezoensis*	AP006715	3
19.	*Gracilaria tenuistipitata var. liui*	AY673996	3
20.	*Cyanidioschyzon merolae*	AB002583	2
21.	*Cyanidium caldarium*	AF022186	2
22.	*Porphyra purpurea*	U38804	6
23.	*Vaucheria litorea*	EU912438	5

5.2 Microsatellite Markers from Chloroplast Genome for Genetic Analysis of *Gracilaria tenuistipitata*

Of the eight primer-pairs we designed, only two primer-pairs revealed genetic polymorphisms on the eighty specimens tested. Primer-pair GT5 differentiated the *G. tenuistipitata* into three main groups at the 0.5 similarity level by cluster analysis. Primer-pair GT8 was able to differentiate the populations of *G. tenuistipitata* from Vietnam and Thailand from the populations of *G. tenuistipitata* from Malaysia (Batu Laut, Middle Banks and Kuah), but not the population from Kelantan, Malaysia, as was achieved with the use of primer-pair GT5. Primer-pair GT5 showed that there are three possible genotypes of *G. tenuistipitata* from the samples collected, whereas primer-pair GT8 only showed two genotypes of *G. tenuistipitata* (Figure 4.4 and 4.5).

Using the combined dataset of the two primer-pairs, *G. tenuistipitata* from Kelantan, Malaysia was more closely related to the populations from Quy Kim, Hai Phong (Vietnam) and Pattani (Thailand) and distantly separated from *G. tenuistipitata* from Batu Laut, Middle Banks and Kuah (Malaysia) (Figure 4.6). This result also showed that the populations of *G. tenuistipitata* from Batu Laut, Middle Banks and Kuah that grow in the Straits of Malacca, on the west coast of Peninsular Malaysia, were separated from the populations of Thailand, Vietnam and Kelantan, Malaysia, which face the South China Sea (Figure 5.1).

Figure 5.1:

SSR markers that were generated from the chloroplast genome of *G. tenuistipitata* were able to distinguish the populations of *G. tenuistipitata* from different geographical origins (west and east coast of Peninsular Malaysia) and the population from Kelantan, Malaysia from other localities. This showed that the variation in the repeated nucleotides can be used not only in genetic mapping and marker-trait association studies (Ellis & Burke, 2007; Wang et al., 2009) but also in genetic diversity and population studies. The combined primer-pairs (i.e. GT5 and GT8) showed even more clearly that *G. tenuistipitata* from Kelantan, Malaysia were distantly separated from *G. tenuistipitata* from Batu Laut, Middle Banks and Kuah (Malaysia). This information regarding genetic diversity and differentiation of strains of the same species that obtained is important particularly in cultivation where red algae have contributed to 24.5% of the total world seaweed production (FAO, 2001) for food, fodder, ingredients in cosmetics and fertilizers, and hydrocolloid production (e.g. agar and alginate) (Chan et al., 2006).

Although all the cpSSRs that we developed have repeated nucleotide of 10 or less than 10, our study showed that primer-pairs GT5 and GT8 with longer repeats, $(AT)_8$ and $(AT)_7$, were more polymorphic than primer-pair GT4, that has shorter repeats, $(ATT)_5$. Primer-pair GT5 was able to give three genotypes and primer-pair GT8 gave two possible genotypes whereas primer-pair GT4 showed monomorphic alleles. Our study supports the experimental analysis of microsatellite in rice (*Oryza sativa* L.) that the class I microsatellites which contain perfect SSRs with longer repeats ($\geq$ 20) are highly polymorphic compared to class II microsatellite markers that contain SSRs > 12 nucleotides and < 20 nucleotides in length (Temnykh et al., 2001). This is also in agreement with data reported for human that according to Weber, the lack of polymorphism is always associated with a low number of repeats (Weber, 1990).

Chloroplast SSRs (cpSSRs) have been used in the study of red pine (*Pinus resinosa*) in eastern North America. Nine chloroplast microsatellite loci were tested on 159 individuals from seven populations and 23 haplotypes were discovered (Echt et al., 1998). In a previous study of this similar species, no allozyme diversity was detected and only limited RAPD diversity was observed (Mosseler et al., 1991; Mosseler et al., 1992; De verno & Mosseler, 1997). This shows that cpSSRs have more discriminative power than allozyme and RAPD method. In our study, only three genotypes were obtained for the analysis on eighty *Gracilaria tenuistipitata* from six different localities. There is limited gene flow and a weak geographic structuring pattern because the numbers of genotypes were rather low. However, this can only be verified with wider samples from more populations. For example, the analysis of the distribution of genetic diversity of the seaweed *Chondrus crispus* by Provan and Maggs (2012) included 19 populations from both sides of the North Atlantic using mitochondrial single nucleotide polymorphisms (SNPs), sequence data from two single copy nuclear regions and microsatellite loci. These researchers' results revealed unique genetic variation for all marker classes in the rear-edge populations in Iberia, but not in the rear-edge populations in North America. Thus, more samples from different populations and also different techniques of molecular assays will probably help us to understand the migration and dispersal ability of the species. There is also an extreme case that no variation was detected in the chloroplast genome of Torrey pine (*Pinus torreyana*) (Provan, 1999).

Based on Figure 4.7, the cpSSR (primer-pair GT5) was able to differentiate not only within the same species of *G. tenuistipitata* according to geographical distribution but also between different species of *Gracilaria*. Based on morphology, *G. changii*, *G. tenuistipitata* and *G. salicornia* are quite distinct from each other. However, the classification of *G. salicornia* and its closely related species (*G. canaliculata* = *G.*

crassa and *G. cacalia*) has been debated over the last two decades (Xia, 1986; Meneses & Abbott, 1987; Withell et al., 1994; Wynne 1995) due to the variable morphology. *G. salicornia* and *G. canaliculata* = *G. crassa* (Xia, 1986) were proven to be two different species using RAPD approach (Lim et al., 2001) and RuBisCO spacer sequence data (Iyer et al., 2005). Their morphological differences are genetically based and not due to plasticity. Although the microsatellite marker that has been developed was not able to differentiate them as distinctive species, our study has proved that they can be used for amplification among different species such as *G. changii*, *G. salicornia*, *G. canaliculata* and *G. cacalia*. One of the studies by Ginwal *et al.* showed that cpSSRs developed from *Acorus calamus* were able to use for cross-species amplification in 20 different individuals of *Asparagus racemosus*. Out of 18 microsatellite primer-pairs screened, five cpSSRs showed good amplification and polymorphism in *Asparagus racemosus* (Ginwal et al., 2011). Their study has proven the possibility of cross-species amplification in other individuals which can also be applied in our study. Cross-species amplification of microsatellite markers not only can save time but also cost-effective as we don't need to develop specific marker for new species.

Our study on *Gracilaria tenuistipitata* supports that the variation in SSR length and stability varies among loci within species (Cho et al., 2000; Temnykh et al., 2000), and microsatellites obtained from genomic libraries have been found to be more polymorphic (Cho et al., 2000; Eujayl et al., 2001). However, whether microsatellites are more polymorphic than ESTs of *G. tenuistipitata* must be proven by comparing SSRs developed from the two sources. The comparison between EST and genomic SSR markers of *G. tenuistipitata* is discussed in section 5.5.

5.3 Microsatellite Markers from Expressed Sequence Tags (ESTs) of Seaweeds in Differentiating Various *Gracilaria* Species

The taxonomic and systematic problems within the members of Gracilariaceae have not been resolved due to the lack of distinct morphological diagnostic characters and unreliable developmental characters of the female reproductive system (Wattier et al., 1997; Iyer et al., 2004). For example, the distinction between *G. changii*, *G. manilaensis* and *G. fisheri* using morphological and anatomical characters is very difficult as the only difference between these three species is the proliferous branching in *G. manilaensis* and *G. fisheri* while less proliferous in *G. changii*. The difference between *G. manilaensis* and *G. fisheri* is mainly the male reproductive structure in which *G. manilaensis* has *verrucosa* type of spermatangia that are crowded and separated by modified vegetative cell walls while there is no modified vegetative cell walls in *G. fisheri* (Chang & Xia, 1988; Xia & Abbott, 1987; Abbott et al., 1991; Yamamoto & Trono, 1994; Phang & Lewmanomont, 2001). Whether the branching mode is proliferative or not is very subjective and the occurrence of male spermatangia is rare, therefore, it is very hard to differentiate them morphologically. Thus, molecular markers are important to distinguish them as distinctive species. Our study has shown that primer-pair P3 was able to distinguish these three different *Gracilaria* species into five different clades (Clade B, *G. manilaensis* and Clade D, *G. fisheri* are clearly differentiated from the three clades of *G. changii* (Clades A, C and E) and hence can be used as molecular marker to distinguish the three species (Figure 4.10).

Our study also showed that the population of *Gracilaria changii* from Morib, Selangor, Malaysia (Clade C) has their unique allele that can distinguish *Gracilaria changii* from the other populations (Clade A and E). Although the populations of *Gracilaria changii* from Sabah and Sarawak are located in east Malaysia, they shared a similar genotype with some of the populations in west coast of Peninsular Malaysia

(Figure 5.2). In 2009, microsatellite markers were developed from *Gracilariopsis lemaneiformis* to study about the genetic diversity of wild populations of this species. This is an agarophyte that is important for biological research and has significant economic value. Pang et al. developed 16 SSR primer-pairs using the conventional method but the PCR products obtained were monomorphic among 40 *Gp. lemaneiformis* tested. This indicated that genetic homogeneity was observed between different *Gp. lemaneiformis* specimens. Our study has included the ESTs from *G. changii*, *G. gracilis* as well as *Gp. lemaneiformis* and the microsatellite markers developed from these ESTs can be used to distinguish between *Gracilaria changii* from different population and also different species of *Gracilaria*. This shows that the EST-SSR that we developed may be used to test on *Gp. lemaneiformis* from wild populations for genetic diversity study. Pang et al. also tested the same primer-pairs with closely related *Gracilaria* species but they only found one SSR generated PCR products from *Gracilaria verrucosa* and *Gracilaria textorii* (Pang et al., 2010). Although they did not mention about the polymorphism of SSR marker on the specimens examined, both of our studies showed the possibility of amplification on different species.

Study on microsatellite markers was also conducted in animals other than seaweeds. For example, the species identification of *Drosophila virilis* (fruit fly) using microsatellite markers were carried out by Routtu *et al.* to identify possible misclassifications based on morphological characters. They found that microsatellite locus Vir72ms were able to identify the North-European species of the group (*D. montana*, *D. littoralis*, *D. ezoana* and *D. lummei*) with specific allele or alleles at this locus. This locus will be valuable for species like *D. lummei* as they are rare species and has not been collected and identified in Fennoscandia, Finland during the last 20 years (Routtu et al., 2007). This showed that microsatellite markers can be used across different application such as the study of germplasm identification, genetic diversity

analysis, parentage analysis, gene mapping, marker assistant selection and population study for seaweeds, plants and animals.

5.4 Comparison between SSR Marker and *cox*1 Gene for Investigating the Genetic Diversity of *G. changii* from West Coast, Peninsular Malaysia

The mitochondrial cytochrome c oxidase subunit I gene (*cox*1) are highly variable markers for use between the genus and species level. However, the *cox*1 gene marker was found to be inappropriate for plants due to the slower rate of evolution of *cox*1 gene in higher plants than in animals (Kress et al., 2005; Chase et al., 2005). Saunder (2005) showed that this gene was suitable to be used in red algae in their study on three species: (i) *Mazzaella* species in the Northeast Pacific; (ii) species of the genera *Dilsea* and *Neodilsea* in the Northeast Pacific; and (iii) *Asteromenia peltata* from three oceans and the marker has successfully assigned collections to the correct species. Robba et al. confirmed the suitability of this marker as an ideal DNA barcoding marker in red algae (Robba et al., 2006). The first study on intra- and inter- specific phylogenetic relationships of *Gracilaria vermiculophylla* using mitochondrial gene *cox*1 has shown that this marker was highly variable and is suitable for species identification (Yang et al., 2007). In 2011, Yow et al. studied the genetic diversity of *Gracilaria changii* using mitochondrial *cox*1 gene and showed that this marker was suitable for resolving intraspecific relationships and is a reliable marker for the study of genetic variation in *Gracilaria*. Comparison between SSR marker and *cox*1 gene showed that the microsatellite marker that we developed, was able to resolve only three genotypes while *cox*1 gene was able to give six different haplotypes (Figure 5.3). This suggested that *cox*1 gene is more variable compared to the microsatellite marker developed from the ESTs sequences that situated in more conserved regions of the genome.

Comparison of two different molecular markers has also been carried out in the genetic diversity study on *Gracilariopsis lemaneiformis* using SSR and AFLP analysis (Pang et al., 2010). The results showed that eight AFLP primer-pairs were able to give 347 reproducible bands while no polymorphism was detected using SSR markers. The UPGMA dendrogram revealed two main clusters with one clade included all the Qingdao isolates. This shows that other molecular marker method, such as *cox*1 gene and AFLP were more reliable and suitable than SSR marker in the genetic diversity study of seaweeds. Although *cox*1 gene was not suitable for the study on plants, the combination of the coding locus *rbcL* (*rbcL*-a) with the non-coding *trnH-psbA* spacer is recommended for the species identification in land plants (Kress & Erickson, 2007). This indicated that there is no ideal marker for a particular study and the suitability of a molecular marker can be tested using different approach.

Figure 5.2:

Figure 5.3:

5.5 Comparison between EST and genomic SSR markers of *Gracilaria tenuistipitata*

Microsatellite markers from chloroplast genome of *G. tenuistipitata* grouped *G. tenuistipitata* into two main clades: (a) *G. tenuistipitata* from Batu Laut, Middle Banks and Kuah (Malaysia); and (b) *G. tenuistipitata* from Pattani (Thailand), Quy Kim, Hai Phong (Vietnam), and Kelantan (Malaysia). This showed that the specimens from Batu Laut, Middle Banks and Kuah (Malaysia) grow in the Straits of Malacca, west coast of Peninsular Malaysia shared unique allele and were separated from the populations of Thailand, Vietnam and Kelantan, Malaysia facing the South China Sea. The population from Kelantan, Malaysia was further grouped in another sub-clade, which was separated from two other localities (Quy Kim, Hai Phong, Vietnam and Pattani, Thailand). On the other hand, the EST-SSR markers showed four different genotypes on the specimens tested and the populations from Kuah (Malaysia) and Pattani (Thailand) were grouped into two distinct clades. The population from Kuah (Malaysia) was shown to be most distant from the other populations while the populations from Batu Laut and Middle Banks (Malaysia) as well as the populations from Quy Kim, Hai Phong, Vietnam and Kelantan, Malaysia shared distinct genotypes. This showed that the EST-SSR markers were not able to differentiate them according to geographical regions as was achieved by using the chloroplast SSR markers but they gave higher number of genotypes.

Most of the studies showed that the microsatellite markers derived from EST database were less polymorphic than those derived from the genomic DNA (Cho et al., 2000; Eujayl et al., 2001; Gupta et al., 2003). However, in our study, the SSR markers obtained from the ESTs were shown to have higher polymorphism in comparison to genomic SSR markers. The number of SSR markers obtained were also higher (primer-pairs GT9, GT10, GT11, GT12, GT13, GT14, GT15, GT16, GT17 and GT18) than the genomic SSR markers (primer-pairs GT1, GT2, GT3, GT4, GT5, GT6, GT7 and GT8).

Several studies have also reported that the level of polymorphism was higher in genic SSRs than genomic SSRs. For example, the microsatellite markers developed from loblolly pine ESTs were showed to be more conserved than traditional SSRs in which the sequence analyses revealed that the frequencies of insertions/deletions and base substitutions were lowed in EST-SSRs (Liewlaksaneeyanawin et al., 2004). Besides, in the study of identification and genome mapping of EST-SSRs in kiwifruit, 93.5% of the markers were showed to be polymorphic in two genotypes of *Actinidia chinensis* (Fraser et al., 2004). This shows that not all genomic SSRs are more polymorphic than EST-SSR marker and this has been proven by our study and also study in different organisms.

In addition, the pattern of SSR markers obtained from the genomic DNA of *G. tenuistipitata* were five mononucleotides, two dinucleotides and one trinucleotide while the SSR markers from the ESTs of *G. tenuistipitata* were five dinucleotides and five trinucleotides with no mononucleotide obtained. This may due to the longer repeats of the loci obtained from EST database as longer repeat length are generally more polymorphic than loci composed of short motifs. This could also be attributed by the size of chloroplast genome downloaded from GenBank which is usually smaller than mitochondrial genome (Zhang et al., 2012) and nuclear genome as chloroplast is only an organelle found in plant cells and some other eukaryotic genomes where photosynthesis take place.

The availability of genomic sequence data is limited in comparison to EST database in most of the plant species. For example, SSRs have been developed from ESTs in *Arabidopsis* (Morgante et al., 2002), cotton (Saha et al., 2003; Han et al., 2004), soybean (Gao et al., 2003), sugarcane (Cordeiro et al., 2001), rice (Temnykh et al., 2000; Kantety et al., 2002), rye (Khlestkina et al., 2004; Hackauf & Wehling, 2002) and wheat (Gao et al., 2003; Nicot et al., 2004). In wheat (Stephenson et al., 1998) and barley

(Ramsay et al., 2000), faint or stuttering bands were observed using genomic SSRs but SSRs derived from ESTs produced high-quality markers with strong bands in most reports, for example in rice (Cho et al., 2000), wheat (Yu et al., 2004) and barley (Kota et al., 2001). However, in our study, all the genomic SSRs derived from *G. tenuistipitata* were able to give amplification with strong bands but some of the EST-SSRs produced null alleles or very faint bands. Besides, null alleles were also observed in the studies of kiwifruit, rice, and wheat using EST-SSR markers. The occurrence of null alleles can be caused by the deletion of microsatellite at a specific locus (Callen et al., 1993) or mutations (insertion, deletion of substitutions) has occurred in the primer binding site (Lehman et al., 1996).

In summary, our study showed that the SSR markers derived from the ESTs were more polymorphic than the genomic SSR markers and they also gave amplification with strong bands. In addition, the numbers of EST-SSR markers obtained were higher than genomic SSRs while the patterns of repeat motif for EST-SSR markers were lower than the genomic SSR markers.

5.6 DNA Fragments Separation Approaches

In our data analysis, the MetaPhor agarose gel electrophoresis (MAGE) was chosen to detect the length polymorphism as they give high resolution separation of 20 – 800 bp DNA fragments that differ in size by 2%. Other electrophoresis methods such as polyacrylamide gel electrophoresis (PAGE) give excellent separation but the preparation of this technique is tedious, requires expensive and hazardous radioactive chemicals. PAGE was the first electrophoresis separation method used in the analysis of the DNA fragments (Cipriani et al., 1999). Capillary electrophoresis was also used to determine length polymorphisms of microsatellite markers (Marino et al., 1995) but this

technique requires sophisticated instruments which are very expensive (Asif et al., 2008). Research has been carried out to evaluate the efficiency of MAGE in the separation of microsatellite markers products in cotton and result has shown that MAGE can clearly separate alleles of SSRs with a size difference of five bp (Asif et al., 2008). Sánchez-Pérez et al. compared the efficiency of using automated capillary sequencers, PAGE and MAGE for the separation and analysis of the DNA fragments in almond and they found that the automated SSR detection and polyacrylamide gel electrophoresis were the most efficient methods. These methods were able to resolve allelic variation at a finer scale than MAGE and the SSR detection using PAGE showed similar results with automated sequencing. However, polyacrylamide gel electrophoresis method is time consuming and toxic as they used radioactive and silver staining in the detection of PCR products. Thus, the analysis of amplified PCR products by MAGE has to be completed by the automated sequencer capillary electrophoresis or PAGE analysis in order to get a higher level of accuracy for the detection of SSR polymorphism (Sánchez-Pérez et al., 2006). We have employed MetaPhor agarose gel electrophoresis in our study and results showed that MetaPhor agarose electrophoresis gave better resolution than the normal gel electrophoresis. Further analysis was carried out using automated sequencing that gave a higher resolution than MAGE which aid in our species identification study by using different softwares. Other researcher has also performed automated sequencer capillary electrophoresis for SSR analysis, for example in the study of EST-SSR markers for *Fucus*, Coyer et al. has analysed the amplification products with an ABI 3730 autosequencer (Applied Biosystems) before the determination of genotypes using software (Coyer et al., 2008). This showed that the SSR analysis by automated sequencing is important to better resolve the PCR product amplified by the microsatellite markers.