Animal Cells and Systems 13: 305-314, 2009

Animal Cells
and
Systems

http://www.AnimalCells.or.kr

Phylogenetic Analysis of Mitochondrial DNA Control Region in the Swimming Crab, *Portunus trituberculatus*

Eun Min Cho, Gi-Sik Min[1], Sumaira Kanwal, Young Se Hyun, Sun Wha Park, and Ki Wha Chung*

Department of Biological Sciences, Kongju National University, Gongju 314-701, Korea; [1]Department of Biological Sciences, Inha University, Incheon 402-751, Korea

Abstract: The control region of mitochondrial DNA (13516-14619) is located between srRNA and $tRNA^{Ile}$ gene in swimming crab, *Portunus trituberculatus*. The present study was investigated the genetic polymorphisms of the control region in samples of *P. trituberculatus* collected at coastal waters of the Yellow Sea in Korea. A total of 300 substitution and indel polymorphic sites were identified. In addition to SNPs and indel variation, a hypervariable microsatellite motif was also identified at position from 14358 to 14391, which exhibited 10 alleles including 53 different suballeles. When the hypervariable microsatellite motif was removed from the alignment, 95 haplotypes were identified (93 unique haplotypes). The nucleotide and haplotype diversities were ranged from 0.024 to 0.028 and from 0.952 to 1.000, respectively. The statistically significant evidence for geographical structure was not detected from the analyses of neighbor-joining tree and minimum-spanning network, neither. This result suggest that population of *P. trituberculatus* are capable of extensive gene flow among populations. We believed that the polymorphisms of the control region will be used for informative markers to study phylogenetic relationships of *P. trituberculatus*.

Key words: control region, microsatellite, mitochondrial DNA, polymorphism, *Portunus trituberculatus*, swimming crab

INTRODUCTION

The crustacean brachyuran species, which are usually called crabs, are found worldwide in both seawater and freshwater, comprising about 6000 species (Bowman and Abele, 1982). As one of the most well-known crustacean species, the swimming crab, *Portunus trituberculatus* (Crustacea: Decapoda: Brachyura), inhabits the seafloor habitats with sand or pebbles, being widely distributed in the coastal waters. *P. trituberculatus* is one of the most common edible crab species in East Asia, and has been artificially propagated and stocked in several countries, *e.g.* Korea and Japan.

The mitochondrial genome (mtDNA) has been extensively studied for population genetics, phylogeographics, and phylogenetics studies in a variety of metazoan species, because of the unique features, *e.g.* haploidy, maternal transmission, high copy number per cell, little or no intermolecular recombination, and rapid mutation rate (Boore, 1999; Ballard and Rand, 2005). Increasing number of the complete mtDNA sequences have been reported in brachyuran species including the Portunidae: *P. trituberculatus* (Yamauchi et al., 2003), *Callinectes sapidus* (Place et al., 2005), *Scylla tranquebarica*, *S. serrata*, *S. olivacea*, and *S. paramamosain*. In recent, the information for the metazoan mtDNA is accessible at the Metazoan Mitochondrial Genomes Accessible database (METAMiGA) website (http://amiga.cbmeg.unicamp.br/) (Feijao et al., 2006).

The control region of the mtDNA which plays a role for replication and expression of the mtDNA does not include a gene (Shadel and Clayton, 1997). The size and structure of the control region are considerably variable among animal groups. Even in the brachyuran species, the sizes range from 0.51 kb in *Geothelphusa dehaani* (Segawa and Aotsuka, 2005) to 1.44 kb in *Callinectes sapidus* (Place et al., 2005). Because the control region is far more polymorphic than other region of mtDNA, it is widely used for tracing maternal origin, haplogroup association analysis, and forensic applications in human (Lee et al., 2008; van Oven et al., 2009). However, studies on the structure and genetic polymorphism of the control region in crustaceans are still very limited. Diniz et al. (2005) reported the potential of using the control region as a genetic marker for phylogeographical studies in the spiny lobster, *Panulirus argus*. Kilpert and Podsiadlowski (2006) revealed that the

*To whom correspondence should be addressed.
Tel: +82-41-850-8506; Fax: +82-41-854-8505
E-mail: kwchung@kongju.ac.kr

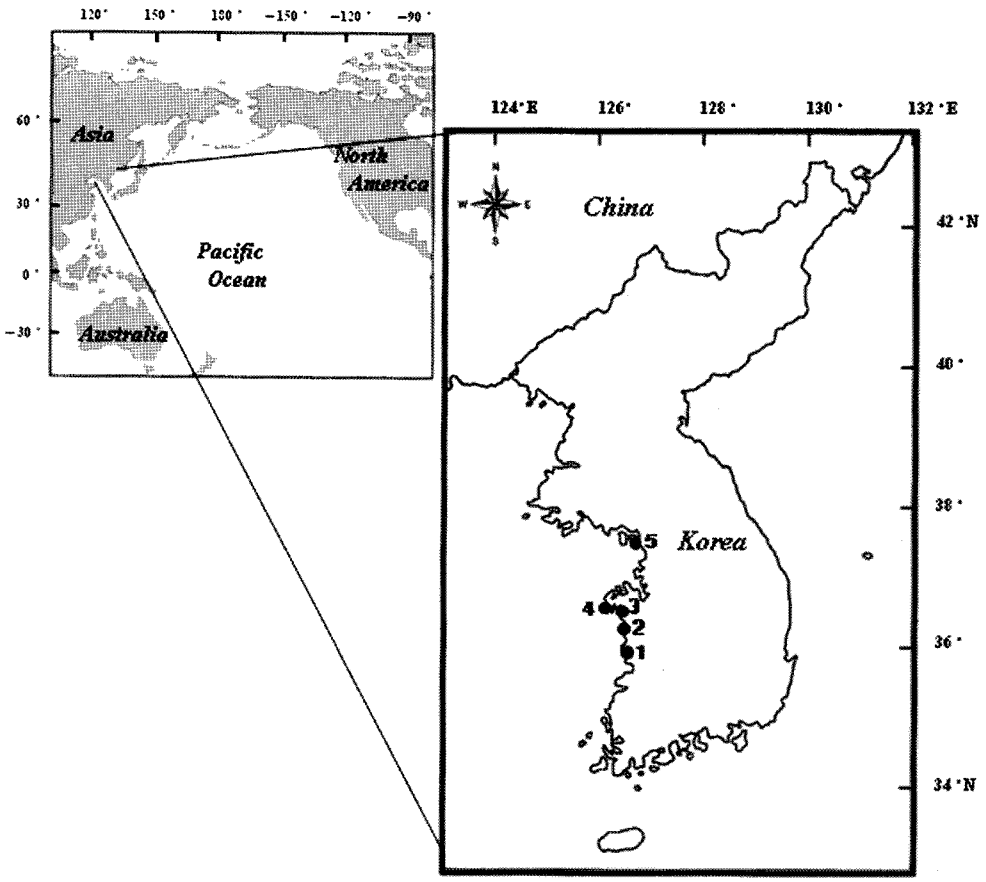

Fig. 1. Sampling locations of swimming crab, *P. trituberculatus*. Swimming crabs were collected from five sites along the coast of the Yellow Sea in Korea (1: Gunsan, 2: Boryeong, 3: Hongseong, 4: Anmyeon-do island, and 5: Incheon).

control region in the common sea slater, *Ligia oceanica* contains two tandem repeats and a GC-rich region containing the putative hairpin structure. Pie et al. (2008) recently revealed that two crab species, *Ucides cordatus* and *Cardisoma guanhumi* shared no similarity at the nucleotide level but showed similar secondary structures and positions of the regulatory elements within the control region.

The complete mtDNA sequence of *P. trituberculatus* was reported by Yamauchi et al. (2003). The mtDNA of *P. trituberculatus* consists of 16026 bp, and the genome includes the 13 protein-coding, 22 tRNA, and two rRNA genes as found in other metazoan animals. In addition, it includes an 1104-bp non-coding putative control region (13516-14619) between the 12S rRNA and $tRNA^{Ile}$ genes.

Because of their worldwide phylogeographic distribution and economical importance, the investigation of genetic diversity of the mtDNA in *P. trituberculatus* would be very important. Imai et al. (1999) investigated the restriction fragment length polymorphisms of mtDNA in *P. trituberculatus*. Liu et al. (2009) and Xu et al. (2009) studied on genetic differentiation among geographical populations of *P. trituberculatus* based on the analyses of mitochondrial 16S rDNA and cytochrome c oxidase subunit 1 (CO1) genes. The control region sequences have been used previously as informative markers to detect population structure in marine decapods, *Panulirus argus*, *Ucides cordatus*, and *Cardisoma guanhumi* (Diniz et al., 2005; Pie et al., 2008). As far as we are recognizing, however, no population genetic study using the mtDNA control region of *P. trituberculatus* was performed yet.

In the present study, we investigated the genetic variation and population genetic structure of *P. trituberculatus* collected from coastal waters of the Yellow Sea in Korea, based on the analyses from the putative mtDNA control region. The aim of the present study is to characterize the variability of the control region and to explore the potential of this polymorphism using fishery stock identification as a genetic marker in *P. trituberculatus*.

MATERIALS AND METHODS

Samples and DNA extraction

Samples of swimming crab, *P. trituberculatus* (n=98) were collected from five sites along the coast of the Yellow Sea in Korea, Gunsan, Boryeong, Hongseong, Anmyon-do island and Incheon (Fig. 1). Total DNA was extracted from the leg tissue using a QIAamp DNA Micro kit (Qiagen,

Table 1. Primer sequences for amplification and sequencing of control region

Name	Position	Sequence (5'→3')	Size (bp)
CR-F1	13441-13466	TCGAGCTACTACACGCAACAACTCT	491
CR-R1	13908-13932	GAGGGAATGCTTCATCTCTCTATTT	
CR-F2	13785-13810	ACTCCTACATTAAGACAGCAATTAGA	415
CR-R2	14176-14200	ATGATAGGATGGAGTTCCTATTTC	
CR-F3	13931-13950	TCAGGGAATGCTTTGGTTCT	553
CR-R3	14465-14484	AAAAGTAAGATTAGGATAGG	
CR-F4	14151-14170	CTCTTCACTCAACGGCAGAA	560
CR-R4	14691-14711	GGGGACGGTAATCCGTTAC	

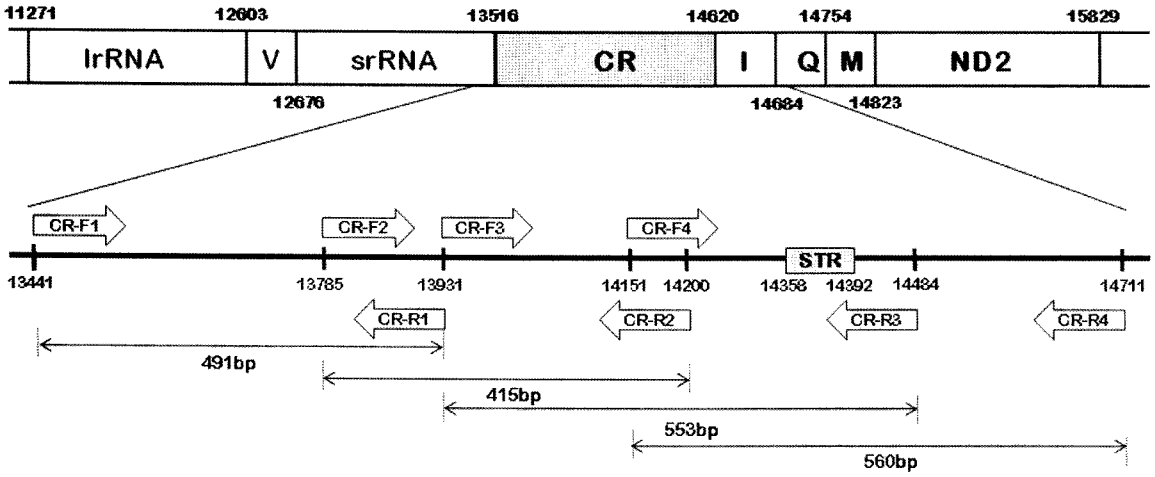

Fig. 2. Partial mtDNA map around the control region and strategy for PCR amplification in *P. trituberculatus*. The control region (CR) and microsatellite (short tandem repeat: STR) motif were indicated by the grey boxes. Primers and PCR sizes were shown on the bottom of the figure (lrRNA: large ribosomal RNA, V: $tRNA^{Val}$, srRNA: small ribosomal RNA, I: $tRNA^{Ile}$, Q: $tRNA^{Gln}$, M: $tRNA^{Met}$, and ND2: NADH dehydrogenase subunit 2).

Hilden, Germany). The extracted DNA was quantified by a spectrophotometer (Smart Spec 3000, BIO-RAD, USA) and stored at −20°C until use.

PCR amplification and sequencing analysis

The control region and its adjacent sequence of mitochondrial genome were amplified with four overlapping fragments by the PCR method using the primers shown in Table 1 (Fig. 2). PCR was performed in 20 μL of reaction mixture containing 10-20 ng of genomic DNA, 10 pmole of each primer, 200 μM of each dNTP, 2 mM of $MgCl_2$, 0.5 unit of *Taq* DNA polymerase (Promega, Madison, WI), and 1X supplied buffer using a GeneAmp PCR system 9700 (Applied Biosystems, Foster city, CA). The PCR amplification condition consisted of initial denaturation at 95°C for 5 min, followed by 32 cycles at 95°C for 30 sec, 60°C for 40 sec, 72°C for 1 min, and a final extension at 72°C for 10 min. PCR products were purified by an Exo-SAP-IT kit (Fermentas, Burlington, Canada), and subsequently used for the sequencing reaction. The sequence was determined from both directions by an automatic genetic analyzer (ABI 3100, Applied Biosystems) using the big dye terminator cycle sequencing ready reaction kit (Applied Biosystems). The variations different from the sequence of Yamauchi et al. (2003) were identified using the SeqScape program (Applied Biosystems).

Characterization of microsatellite motif

For the microsatellite motif (14358-14391) within the control region, the exact allele sizes and repeat structures were determined by direct sequencing of each DNA sample. Alleles were basically designated by the number of their repeats, however suballeles were defined as same repeat number but different repeat structure.

Statistical analysis

The obtained sequences along with that of Yamauchi et al. (2003) retrieved from the GenBank were aligned with ClustalX (Thompson et al., 1997). Nucleotide diversity, haplotype diversity (h) and the average number of nucleotide difference (π) were calculated using DnaSp (Ver. 4.00) (Rozas et al., 2003). A phylogenetic tree was constructed by employing the neighbor-joining (NJ) tree method using the MEGA 4.0 computer program with a

Kimura 2-parameter (Tamura et al., 2007). Gaps were treated as a fifth character. The gene diversity (GD) of microsatellite motif was determined using the formula $N(1-\Sigma P_i^2)/(N-1)$, where N represents the number of individuals, and P_i is the allele frequencies.

RESULTS

Substitution and insertion/deletion variation

A total of 300 nucleotide substitution and insertion/deletion (indel) polymorphic sites were identified within the control region (except for a microsatellite motif 14358-14391) in the 98 samples of *P. trituberculatus* collected from five locations along the coast of the Yellow Sea in Korea (Fig. 1). Of them, 149 sites showed that the most frequent allele frequencies were calculated to be below the 0.97 (observed sample No.: 95 or less). Table 2 summarized the nucleotide positions and substitutions, mutation types and allele frequencies of these 149 polymorphic sites. In the table, the numbering of nucleotides and reference sequences were followed by Yamauchi et al. (2003). Allelic frequencies represent those of alleles different from the reference sequences. At many sites, the original sequences reported by Yamauchi et al. (2003) were not the most frequent (13579, 13580, 13687, 13712, 13738, 13742, 13749, 13755, 13813, 13889, 14115, 14138, 14140, 14158, 14209, 14228, 14232, 14339, 14399, and 14470).

Of the 149 polymorphic sites, most substitutions exhibited

Table 2. Polymorphic nucleotides of the mtDNA control region in *P. trituberculatus*

Position	Ref	A	T	G	C	Indel	Freq	Position	Ref	A	T	G	C	Indel	Freq
13579[a,b]	G	91		6	1		A:0.929	13736[a]	C		7		91		0.071
							C:0.010	13737[a]	A	94		4			0.041
13580[a,b]	C		1	81	16		T:0.010	13738[a]	T		11		87		0.888
							G:0.827	13742[a]	A	12		86			0.878
13587[a]	G	10		88			0.102	13744[a,c]	G	12		85		Del:1	A:0.122
13614[a]	A	95		3			0.031								D:0.010
13615[a]	C		5		93		0.051	13747[a]	A	76		22			0.224
13621[a]	T		87		11		0.112	13748[a]	A	95		3			0.031
13624[a]	C		3		95		0.031	13749[a]	A	4		94			0.959
13654[a]	G	7		91			0.071	13752[a,c]	T		93		3	Del:2	C:0.031
13656[a]	G	4		94			0.041								D:0.020
13667[a]	A	93		5			0.051	13754[a]	A	90		8			0.082
13669[a]	G	3		95			0.031	13754-1[c]	In					A:7	0.071
13678[a]	C		7		91		0.071	13755[a]	C		93		5		0.949
13684[a]	G	6		92			0.061	13764[a,b]	A	94	1	3			T:0.010
13687[a,b]	G	81	12	5			A:0.827								G:0.031
							T:0.122	13765-1[c]	In					A:4	0.041
13699[a]	T		86		12		0.122	13767[a]	C		6		92		0.061
13702[a,b]	A	90	1	7			T:0.010	13768[a]	A	95		3			0.031
							G:0.071	13770[a]	T		89		9		0.092
13709[a]	A	92		6			0.061	13771[a]	A	91		7			0.071
13711[a]	T		94		4		0.041	13776[a]	A	94		4			0.041
13712[a]	G	95		3			0.969	13778[a,b]	T	2	93	3			A:0.020
13717[a]	G	16		82			0.163								G:0.031
13721[a]	T		90		8		0.082	13803[a,b]	C	5	3		90		A:0.051
13727[a]	C		12		86		0.122								T:0.031
13729[b]	C	10			88		0.102	13804[a]	A	76		22			0.224
13733[a,b]	C	1	2		95		A:0.010	13807[a]	T		91		7		0.071
							T:0.020	13809[a]	G	16		82			0.163
13734[a]	T		95		3		0.031	13813[a]	A	27		71			0.724
13735[a]	T		93		5	Ht:1	0.051	13815[a]	T		95		3		0.031

[a]transition; [b]transversion; [c]indel mutations. References were followed by Yamauchi *et al* (2003).

Table 2. Continued

Position	Ref	A	T	G	C	Indel	Freq	Position	Ref	A	T	G	C	Indel	Freq
13830[a]	T		85		13		0.133	14138[a]	C		52		46		0.531
13864[a]	T		79		19		0.194	14140[a,b]	C	13	45	1	39		A:0.133
13888[a,b]	C	3	10		85		A:0.031								T:0.459
							T:0.102								G:0.010
13889[a]	T		14		84		0.857	14145[a,b]	T	1	69		28		A:0.010
13931[a]	T		80		18		0.184								C:0.286
13932[a,b]	C	10	14	3	71		A:0.102	14146[a]	G	34		64			0.347
							T:0.143	14148[a,b]	A	84		12	2		G:0.122
							G:0.031								C:0.020
13933[a,c]	A	86		11		Del:1	G:0.112	14149[a]	C		23		75		0.235
							D:0.010	14154[a,b]	T	1	88		9		A:0.010
13934[a]	G	24		74			0.245								C:0.092
13956[a]	A	92		6			0.061	14156[b]	C			17	81		0.173
13957[a]	G	5		93			0.051	14158[a,b]	C	5	73		20		A:0.051
13958[a]	T		88		10		0.102								T:0.745
13983[a]	C		9		89		0.092	14163[a]	C		4		94		0.041
13988[a,c]	A	91		1		Del:6	G:0.010	14165[a]	G	19		79			0.194
							D:0.061	14166[a,b]	C		5	19	74		T:0.051
13989[a,b,c]	G	1	1	95		Del:1	A:0.010								G:0.194
							T:0.010	14167[a]	A	90		8			0.082
							D:0.010	14181[a]	T		78		20		0.204
13993[b]	A	95	3				0.031	14187[a]	C		11		87		0.112
14114[a]	A	94		4			0.041	14189[a,b]	C		3	1	94		T:0.031
14115[a]	C		88		10		0.898								G:0.010
14117[a]	A	92		6			0.061	14191[a]	A	76		22			0.224
14118[a]	A	91		7			0.071	14195[a]	T		93		5		0.051
14125[a]	T		92		6		0.061	14206[a]	T		94		4		0.041
14126[a]	T		95		3		0.031	14206-1[c]	In					A:3	0.031
14128[a,b]	A	91		2	5		G:0.020	14209[c]	T		19			Del:79	0.806
							C:0.051	14209-1[c]	In					T:3	0.031
14132[b]	C	6			92		0.061	14209-3[c]	In					C:86	0.878
14137[a]	A	90		8			0.082	14211[a]	A	91		7			0.071

transition mutations (95 sites), however, transversions were rarely present (5 sites). The indel polymorphisms were found at 15 sites. Additional 34 sites exhibited complex types which consist of two or more types of mutations. Particularly, position 13932, 14140, 14212, and 14228 showed four kinds of nucleotides. The high rate of transition mutation compared to transversion is consistent with the general phenomenon of mutation events.

Polymorphism of microsatellite motif

In addition to substitution and indel variations, a hyper-polymorphic 2-bp repeat microsatellite motif was identified at position from 14358 to 14391 (Fig. 2). The repeat structures were very complex, but the major repeat unit was 'TA' 2-bp. From the analysis of 98 samples, the motif showed 10 alleles with different repeat numbers from 12 to 24, however, it could be divided into 53 different suballeles with additional repeat sequence variations (Table 3). Particularly, the commonest allele 17 (allele frequency: 0.418) displayed 18 different suballeles. As the commonest suballele, 14f and 17c were observed in 8 samples, respectively, and followed by 17h in 7 samples. Of the 53 suballeles, 33 types were uniquely identified only in a sample. It is interesting that most alleles had a 'CGTATGCA' conserved sequence at the end of repeats (37 among 53 suballeles). When this microsatellite motif was considered as a single genetic marker, the gene diversity was estimated to be 0.977. Other 'TA(G)' 2-bp repeat' motif

Table 2. Continued

Position	Ref	A	T	G	C	Indel	Freq	Position	Ref	A	T	G	C	Indel	Freq
14212[a,b,]	A	83	3	10	1	Del:1	T:0.031	14301[a]	A	82		16			0.163
							G:0.102	14310[a]	T		80		18		0.184
							C:0.010	14319[c]	T		93			Del:5	0.051
							D:0.010	14320[a]	A	94		4			0.041
14212-1[c]	In					A:68	0.694	14327[a]	A	84		14			0.143
14215[a,b]	C	4	3		91		A:0.041	14328[a]	T		92		6		0.061
							T:0.031	14329[a]	A	94		4			0.041
14217-1[c]	In					A:3	0.031	14334[a]	A	95		3			0.031
14226[a,b]	T	1	93		4		A:0.010	14335[a,b]	A	71		26	1		G:0.265
							C:0.041								C:0.010
14228[a,b]	A	34	1	62	1		T:0.010	14337[a]	G	5		93			0.051
							G:0.633	14339[a]	A	42		56			0.571
							C:0.010	14343[a]	A	95		3			0.031
14229[a]	A	87		11			0.112	14344[a]	C		7		91		0.071
14232[a]	T		16		82		0.837	14352[a,b]	T		93	3	2		G:0.031
14233[a]	C		34		64		0.347								C:0.020
14234[a]	A	93		5			0.051	14399[a]	G	56		42			0.571
14238[a]	A	88		10			0.102	14426[c]	A	89				Del:9	0.092
14253[a]	A	75		23			0.235	14441[a]	C		7		91		0.071
14270[b]	T		94	4			0.041	14470[a]	C		87		11		0.888
14278[a]	T		93		5		0.051	14476[a]	C		8		90		0.082
14284[a]	C		14		84		0.143	14496-1[c]	In					A:2	0.020
14287[a,b]	C	1	2		95		A:0.010							T:1	0.010
							T:0.020	14528[c]	T		78			Del:20	0.204
14288[a]	G	23		75			0.235	14543[a,c]	C		1		95	Del:2	T:0.010
14289[a]	G	9		89			0.092								D:0.020
14293[a,b]	A	52	10	36			T:0.102	14550[a,c]	T	1	95			Del:2	A:0.010
							G:0.367								D:0.020
14294[a]	T		87		11		0.112	14553[c]	A	95				Del:3	0.031
14296[a]	T		88		10		0.102	14554[c]	T		95			Del:3	0.031
14298[a,b]	T		50	5	43		G:0.051	14555[c]	T		95			Del:3	0.031
							C:0.439	15592[a]	A	94		4			0.041

was found at position 13601-13614 (TATATGTATATATA), however, no repeat number variation was found from the examined samples.

Polymorphic sites with rare allele

About half of the polymorphic sites showed the commonest allele frequencies with 0.97 or below. The information for these less polymorphic sites not shown in Table 2 will be provided upon request to the corresponding author. Major parts of them are transition mutations, but transversion and indel mutations were also occasionally observed. As the particular cases, 12-bp and 13-bp long deletions were identified at the partially overlapped 'A/T'-rich sequences. Deletions of 13-bp nucleotides (position 14043-14055: CTCTTAATTTATT) and of 12-bp nucleotides (position 14053-14064: ATTTTCTTAAAT) were observed in two and one samples, respectively. A 'GTT' 3-bp insertion between position 14227 and 14228 was identified only in a sample. The position 13735 exhibited a heteroplasmic variation with co-existence of T and C in a sample. The heteroplasmy was confirmed by the PCR cloning of the corresponding fragment into the pGEM-T easy vector (Promega) and subsequent sequencing of plasmids from 20 random selected *E. coli* colonies.

Table 3. Repeat structures of microsatellite motif (position 14358-14391) and allele frequencies

Allele	Suballele	Repeat structure	Obs. No.	Freq.
12	12a	$(TA)_8(CG)(TA)(TG)(CA)$	1	0.010
	12b	$(TA)_2(CA)(TA)_7(TG)(TA)$	1	0.010
13	13a	$(TA)_9(CG)(TA)_2(CA)$	1	0.010
	13b	$(TA)_9(CG)(TA)(TG)(CA)$	2	0.020
	13c	$(TA)_7(TG)(TA)(CG)(TA)_2CA)$	1	0.010
	13d	$(TA)_5(CA)_2(TA)_2(CG)(TA)(TG)(CA)$	1	0.010
	13e	$(TA)_2(CA)(TA)_6(CG)(TA)(TG)(CA)$	1	0.010
14	14a	$(TA)_{10}(CG)(TA)_2(CA)$	1	0.010
	14b	$(TA)_{10}(CG)(TA)(TG)(CA)$	4	0.041
	14c	$(TA)_8(CA)(TA)(CG)(TA)(TG)(CA)$	1	0.010
	14d	$(TA)_8(TG)(TA)(CG)(TA)_2(CA)$	1	0.010
	14e	$(TA)_7(CG)(TA)_2(CG)(TA)(TG)(CA)$	1	0.010
	14f	$(TA)_5(CA)_3(TA)_2(CG)(TA)(TG)(CA)$	8	0.082
	14g	$(TA)_2(CA)(TA)_7(CG)(TA)(TG)(CA)$	3	0.031
	14h	$(TA)_2(CA)(TA)_7(TG)(TA)(TG)(CA)$	1	0.010
15	15a	$(TA)_8(CA)(TA)_2(CG)(TA)(TG)(CA)$	1	0.010
16	16a	$(TA)_9(CA)(TA)_2(CG)(TA)(TG)(CA)$	4	0.041
	16b	$(TA)_5(CA)(TA)_4(TG)(TA)(CG)(TA)(TG)(CA)$	2	0.020
	16c	$(TA)_5(TG)(TA)_3(CA)(TA)_2(CG)(TA)(TG)(CA)$	1	0.010
	16d	$(TA)_2(CA)(TA)_6(CA)(TA)(CA)(CG)(TA)(TG)(CA)$	3	0.031
17	17a	$(TA)_{10}(CA)(TA)_2(CG)(TA)_3$	1	0.010
	17b	$(TA)_{10}(CA)(TA)_2(CG)(TA)_2(CA)$	1	0.010
	17c	$(TA)_{10}(CA)(TA)_2(CG)(TA)(TG)(CA)$	8	0.082
	17d	$(TA)_9(CA)_2(TA)_2(CG)(TA)(TG)(CA)$	3	0.031
	17e	$(TA)_9(CA)(TA)_3(TG)(TA)(CG)(TA)$	1	0.010
	17f	$(TA)_8(CA)_3(TA)_2(CG)(TA)(TG)(CA)$	2	0.020
	17g	$(TA)_8(CA)(CG)(TA)_3(CG)(TA)_2(CA)$	2	0.020
	17h	$(TA)_8(CA)(TA)(CA)(TA)_2(CG)(TA)(TG)(CA)$	7	0.071
	17i	$(TA)_7(CA)_4(TA)_2(CG)(TA)(TG)(CA)$	1	0.010
	17j	$(TA)_5(CA)(TA)_5(TG)(CA)(CG)(TG)_2(CA)$	1	0.010
	17k	$(TA)_5(CA)(TA)_3(CA)(TA)(TG)(TA)(CG)(TA)(TG)(CA)$	3	0.031
	17l	$(TA)_5(TG)(TA)_3(CA)(TA)_3(CG)(TA)(TG)(CA)$	2	0.020
	17m	$(TA)_2(CA)(TA)_7(CA)(TA)_2(CG)(TA)_3$	2	0.020
	17n	$(TA)_2(CA)(TA)_7(CA)(TA)(CA)(CG)(TA)(TG)(CA)$	2	0.020
	17o	$(TA)_2(CA)(TA)_7(CA)(TA)(CA)(CG)(TG)_2(CA)$	1	0.010
	17p	$(TA)_2(CA)(TA)_6(CA)_2(TA)_2(CG)(TA)(TG)(CA)$	1	0.010
	17q	$(TA)_2(CA)(TA)_6(CA)_2(TA)(CA)(CG)(TA)(TG)(CA)$	1	0.010
	17r	$(TA)_2(CA)(TA)_2(CA)(TA)_3(CA)(TA)(TG)(TA)(CG)(TA)(TG)(CA)$	2	0.020
18	18a	$(TA)_9(CA)_3(TA)_2(CG)(TA)(TG)(CA)$	1	0.010
	18b	$(TA)_9(CA)_2(TA)_3(CG)(TA)(TG)(CA)$	2	0.020
	18c	$(TA)_8(CA)_4(TA)_2(CG)(TA)(TG)(CA)$	2	0.020
	18d	$(TA)_7(CA)_2(TA)_3(TG)(TA)(CG)(TA)(TG)(CA)$	1	0.010
	18e	$(TA)_5(CA)(TA)_6(TG)(CA)(CG)(TG)_2(CA)$	1	0.010
	18f	$(TA)_2(CA)(TA)_7(CG)(CA)(TA)_2(CG)(TA)(TG)(CA)$	1	0.010
19	19a	$(TA)_9(CA)_3(TA)_3(CG)(TA)(TG)(CA)$	1	0.010
	19b	$(TA)_9(CA)(TA)_3(TG)(TA)(CG)(TA)(TG)(CA)$	1	0.010
	19c	$(TA)_8(CA)(TA)_4(TG)(TA)(CG)(TA)(TG)(CA)$	1	0.010
	19d	$(TA)_7(CA)_5(TA)_3(CG)(TA)(TG)(CA)$	1	0.010
	19e	$(TA)_2(CA)(TA)_8(CA)_2(TA)_2(CG)(TA)(TG)(CA)$	1	0.010
	19f	$(TA)_2(CA)(TA)_7(CA)_2(TA)_3(CG)(CA)(TG)(CA)$	1	0.010
20	20a	$(TA)_9(CA)_2(TA)_2(CA)(TA)_2(CG)(TA)(TG)(CA)$	1	0.010
	20b	$(TA)_7(TG)(TA)_2(CA)_4(TA)_2(CG)(TA)(TG)(CA)$	2	0.020
24	24a	$(TA)_9(CA)_4(TA)_2(CA)(TA)_4(CG)(TA)_2(CA)$	1	0.010

Table 4. Collection locations, number of samples and statistical summary of genetic variability for *P. trituberculatus* along the coast of the Yellow Sea, Korea[a]

Population	Sampling site	No. sample	No. haplotype	Haplotype diversity (*h*)	Nucleotide diversity (ð)	No. private haplotype
1	Gunsan	19	18	0.9942	0.0274	17 (89%)
2	Boryeong	7	6	0.9524	0.0241	5 (83%)
3	Hongseong	6	6	1.0000	0.0284	6 (100%)
4	Anmyeon-do Is.	47	46	0.9991	0.0275	46 (98%)
5	Incheon	19	19	1.0000	0.0257	19 (100%)
Total		98	95	0.9987	0.0272	93

[a] Hypervariable microsatellite motifs were excluded in the analysis.

Population structure

There was considerable genetic variation in the sampling populations of *P. trituberculatus*. From the haplotype analysis, 96 different types were identified, and 95 were defined as unique haplotype only in single individual. Even the hypervariable microsatellite motif was removed, 95 haplotypes were still remained (93 haplotypes were unique), of which only one shared between two locations: population 2 (SC2 and SC5) and population 4 (SC50) in Table 3. The diversities of nucleotide (π) and haplotype (h) within population ranged from 0.024 to 0.028 and from 0.952 to 1.000, respectively (Table 4). The genetic diversity level of control region was compared with other crab species, the genetic diversity of *P. trituberculatus* was slightly low. For example, the nucleotide diversity of Brazilian mangrove land crab, *Ucides cordatus*, ranged from 0.032 to 0.041 (Oliveira-Neto et al., 2007). While those of other gene such as cytochrome c oxidase subunit 1 (CO1) in *P. trituberculatus*, the genetic diversities of the control region were very high: the nucleotide diversity and the haplotype diversity of CO1 for China populations of *P. trituberculatus* ranged from 0.002 to 0.003 and 0.582 to 0.847, respectively (Xu et al., 2009).

High rate of unique haplotypes and a few numbers of shared haplotypes among populations may reflect that extensive gene flow among the Korean populations of *P. trituberculatus* was occurred and there are no clear geographical boundaries among Korean populations. This idea was also strongly supported from the phylogenetic analysis. We could not find any statistically significant geographical structure from a neighbor-joining tree (Fig. 3), in which there is no apparent relationship between the location where a given sample was located and its genetic relationship with the other samples.

DISCUSSION

From the investigation of the mtDNA control region in *P. trituberculatus*, we could observe about 300 polymorphic nucleotides including a microsatellite motif. Particularly, the hyperpolymorphic microsatellite motif exhibited 10 alleles including 53 different suballeles. When the microsatellite motif was not considered, 95 haplotypes were identified, of which 93 were defined as unique haplotype. The diversities of nucleotide and haplotype from the total samples were determined to be 0.0272 and 0.9987, respectively.

No statistically significant geographical structure was detected from the phylogenetic analysis (Fig. 3). It has been known well that species in the genus *Portunus* have high dispersal potential and possibly migrate more than several hundred kilometers (Klinbunga et al., 2007; Liu et al., 2009). In this aspect, our sampling locations might be too close to prevent the mixture of crabs from different areas. Therefore to observe the intraspecific population structuring from the analysis of the mtDNA control region sequences, studies using larger sample size and more widespread sampling throughout the distribution of *P. trituberculatus* are required in the future.

Sequence similarities were weak (*Callinectes sapidus*: 49.0%, *Scylla paramamosain*: 48.9%), when the control region of *P. trituberculatus* was compared with other species in the class Portunidae using the EMBOSS Pairwise Alignment Algorithms (http://www.ebi.ac.uk/Tools/emboss/align/). Similar microsatellite motif regions were found in the control regions of *Callinectes sapidus* and *Scylla paramamosain*, however, they were not strongly matched with the sequence of *P. trituberculatus*. The A+T ratio of the control region in *P. trituberculatus* (76.3%) was similar to those of closely related other crab species, *Callinectes sapidus* (78.3%), *Ucides cordatus* (77.1%) and *Cardisoma guanhumi* (77.2%), but, below than that of *Scylla paramamosain* (86.2%) (Place et al., 2005; Pie et al., 2008).

In conclusion, our present study provides the evidences that the mtDNA control region is highly polymorphic in *P. trituberculatus*. This population genetic study using the mtDNA control region of *P. trituberculatus* would be first. We believed that the polymorphisms of the control region will be used as highly informative markers to study phylogenetic relationships of *P. trituberculatus*. These results will be also helpful to prepare a management

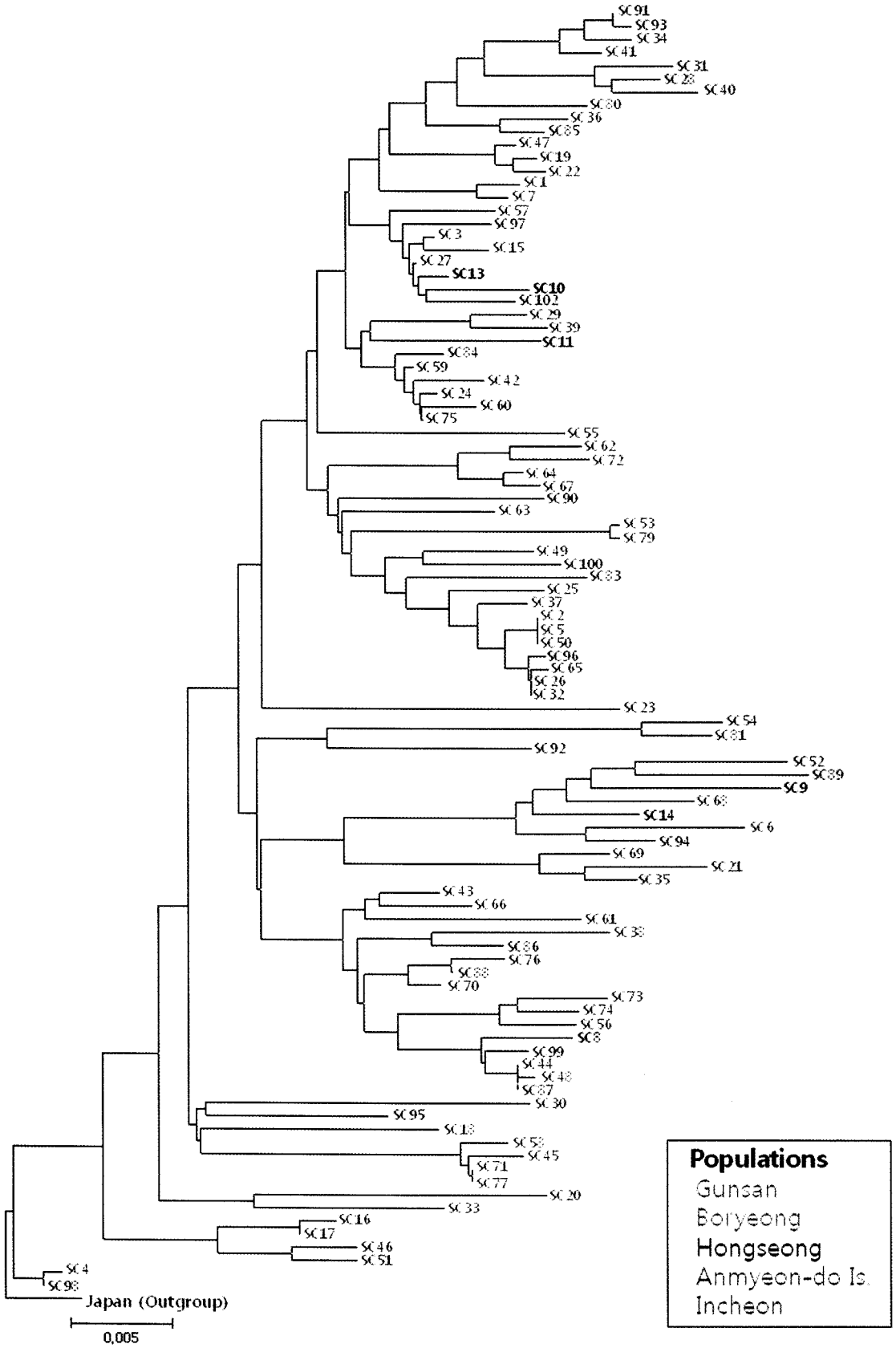

Fig. 3. Phylogenetic relationship of swimming crab, *P. trituberculatus* based on the mtDNA control region sequences. The tree was constructed by neighbor-joining distance analyses. Outgroup (Japan) sequence was retrieved from NCBI (No. AB093006). Sample names were marked with five different colors according to the collection localities as shown in Fig. 1.

strategy for the stock identification and conservation due to its commercial importance.

ACKNOWLEDGMENTS

This study was supported the National Research Foundation of Korea (NRF) grant in Ministry of Education, Science and technology (KOSEF grant No: R01-2008-000-20604-0), and by the Agricultural R&D Promotion Center in Ministry for Food, Agriculture, Forestry and Fisheries (ARPC grant No: 109092-03) funded by the Korean Government.

REFERENCES

Ballard JWO and Rand DM (2005) The population biology of mitochondrial DNA and its phylogenetic implications. *Annu Rev Ecol Evol Syst* 36: 621-642.

Boore JL (1999) Animal mitochondrial genomes. *Nucleic Acids Res* 27: 1767-1780.

Bowman TE and Abele LG (1982) Classification of the recent Crustacea. In: Abele LG (ed), The Biology of Crustacea. Vol. 1: Systematics, the Fossil Record, and Biogeography. Academic Press, New York, pp 1-92.

Diniz FM, Maclean N, Ogawa M, Cintra IH, and Bentzen P (2005) The hypervariable domain of the mitochondrial control region in Atlantic spiny lobsters and its potential as a marker for investigating phylogeographic structuring. *Mar Biotechnol (NY)* 7: 462-473.

Feijao PC, Neiva LS, Azeredo-Espin AML, and Lessinger AC (2006) AMiGA: The Arthropodan Mitochondrial Genomes Accessible database. *Bioinformatics* 22: 902-903.

Imai H, Fujii Y, and Karakawa J (1999) Analysis of the population structure of the swimming crab, *Portunus trituberculatus* in the coastal waters of Okayama Prefecture, by RFLPs in the whole region of mitochondrial DNA. *Fish Sci* 65: 655-656.

Kilpert F and Podsiadlowski L (2006) The complete mitochondrial genome of the common sea slater, *Ligia oceanica* (Crustacea, Isopoda) bears a novel gene order and unusual control region features. *BMC Genomics* 7: 241.

Klinbunga S, Khetpu K, Khamnamtong B, and Menasveta P (2007) Genetic heterogeneity of the blue swimming crab (*Portunus pelagicus*) in Thailand determined by AFLP analysis. *Biochem Genet* 45: 725-736.

Lee HY, Yoo JE, Park MJ, Kim CY, and Shin KJ (2008) Genetic characterization and assessment of authenticity of ancient Korean skeletal remains. *Hum Biol* 80: 239-250.

Liu Y, Liu R, Ye L, Liang J, Xuan F, and Xu Q (2009) Genetic differentiation between populations of swimming crab *Portunus trituberculatus* along the coastal waters of the East China Sea. *Hydrobiologia* 618: 125-137.

Oliveira-Neto JF, Boeger WA, Pie MR, Ostrensky A, and Hungria DB (2007) Genetic structure of populations of the mangrove crab *Ucides cordatus* (Decapoda: Ocypodidae) at local and regional scales. *Hydrobiologia* 583: 69-76.

Pie MR, Oliveira-Neto JF, Boeger WA, Ostrensky A, and Baggio RA (2008) The organization of the mitochondrial control region in 2 Brachyuran Crustaceans: *Ucides cordatus* (Ocypodidae) and *Cardisoma guanhumi* (Gecarcinidae). *J Hered* 99: 432-437.

Place AR, Feng X, Steven CR, Fourcade HM, and Boore JL (2005) Genetic markers in blue crabs (*Callinectes sapidus*) II: Complete mitochondrial genome sequence and characterization of variation. *J Exp Mar Biol Ecol* 319: 15-27.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, and Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496-2497.

Segawa RD and Aotsuka T (2005) The mitochondrial genome of the Japanese freshwater crab, *Geothelphusa dehaani* (Crustacea: Brachyura): evidence for its evolution via gene duplication. *Gene* 355: 28-39.

Shadel GS and Clayton DA (1997) Mitochondrial DNA maintenance in vertebrates. *Annu Rev Biochem* 66: 409-435.

Tamura K, Dudley J, Nei M, and Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596-1599.

Thompson, JD, Gibson TJ, Plewniak F, Jeanmougin F, and Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 24: 4876-4882.

van Oven M and Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat* 30: E386-394.

Xu Q, Liu R, and Liu Y (2009) Genetic population structure of the swimming crab, *Portunus trituberculatus* in the East China Sea based on mtDNA 16S rRNA sequences. *J Experimental Mar Biol Ecol* 371: 121-129.

Yamauchi MM, Miya MU, and Nishida M (2003) Complete mitochondrial DNA sequence of the swimming crab, *Portunus trituberculatus* (Crustacea: Decapoda: Brachyura). *Gene* 311: 129-135.

[Received July 10, 2009; accepted September 11, 2009]