J. Ecol. Field Biol. 30 (4): 331～340, 2007

Biogeography and Distribution Pattern of a Korean Wood-eating Cockroach Species, *Cryptocercus kyebangensis*, Based on Genetic Network Analysis and DNA Sequence Information

Park, Yung Chul[1] and Jae Chun Choe[2]*

[1]*Research Institute of EcoScience, Ewha Womans University, Seoul 120-750, Korea*

[2]*Division of EcoScience, Ewha Womans University, Seoul 120-750, Korea*

ABSTRACT: We examined the evolutionary and ecological processes shaping current geographical distributions of a Korean wood-eating cockroach species, *Cryptocercus kyebangensis*. Our research aims were to understand evolutionary pattern of DNA sequences, to construct genetic network of *Cryptocercus kyebangensis* local populations and to understand evolutionary and ecological processes shaping their current geographical distribution patterns via DNA sequence information and genetic networks, using sequence data of two genes (ITS-2 and AT region) from local populations of *C. kyebangensis*. The results suggest that the ITS-2 and AT region are appropriate molecular markers for elucidating *C. kyebangensis* geographic patterns at the population level. The MSN-A based on the ITS-2 showed two possible routes, the Hwaak-san and Myeongji-san route and the Seorak-san and Gyebang-san route, for migration of ancestral *C. kyebangensis* into South Korea. The MSNs (MSN-A and -B) elucidate migration routes well within South Korea, especially the route of Group Ⅰ and Group Ⅱ.

Key words: AT rich region, Biogeography, *Cryptocercus kyebangensis*, Genetic network, ITS-2, Migration

INTRODUCTION

Woodroaches of the genus *Cryptocercus* are subsocial and xylophagous insects which are closely related to primitive termites (Cleveland et al. 1934, Seelinger and Seelinger 1983, Nalepa 1984, Park et al. 2002, Park and Choe 2003). The distribution of the genus *Cryptocercus* is limited in high forest mountains in West China, Northeast Asia and western and eastern North America (Cleveland et al. 1934, Bey-Bienko 1950, Park et al. 2004, Grandcolas et al. 2005). In Korea, they are distributed over wide mountainous regions at high altitudes (Park et al. 2002, Park and Choe 2003, Park et al. 2004). Since they have no wings and live in rotten logs, their distribution in local areas is patchy and discontinuous. Park et al. (2004) suggested that *C. kyebangensis* may have migrated into Korea during the Miocene (16.5～11.2 Myr), and that their current distribution within Korea resulted from migration across the Taebaek Mountains during the Pleistocene (0.8～0.4 Myr ago). Because of their evolutionary history and patchy geographical distribution, *Cryptocecrcus* species have been models for understanding evolutionary and ecological processes acting on speciation and geographical distribution of local populations in ecosystems (Park et al. 2004, Maekawa et al. 2005).

Our research aims were to understand evolutionary pattern of DNA sequences, to construct genetic network of *Cryptocercus* local populations and to understand evolutionary and ecological processes shaping their current geographical distribution patterns via DNA sequence information and genetic networks, using sequence data of two genes from local populations of *C. kyebangensis*.

MATERIALS AND METHODS

Sample Collection

There are 13 sites of *C. kyebangensis* distribution recognized in South Korea (Park et al. 2004, Maekawa et al. 2005). We collected *C. kyebangensis* populations from 11 sites in South Korea and *C. parvus* (the outgroup species; MANCHURA) from Manchuria (Fig. 5). Samples were preserved in 100% ethanol. Details of sample information are listed in Table 1.

DNA Extraction

Two genes, one from a nuclear internal transcribed spacer (ITS-2) region and one from a part of the mitochondrial DNA control region (AT region; hereafter referred to as AT region), were sequenced from each local population. For extraction of genomic DNA, we removed leg tissues from the ethanol-preserved specimens and dried them on heating blocks to remove the ethanol. We then extracted whole genomic DNA using the AccuPrep Genomic DNA

* Corresponding author; Phone: +82-2-3277-4513, Email: jaechoe@ewha.ac.kr

Table 1. Sampling locality and gene information of *C. kyebangensis* local populations

Species	Sampling locality	Abbreviation of local populations	Genbank accession number	
			ITS	AT
C. kyebangensis	Juwang-san (=Juwangsan Natl. Park, 720 m), Gyeogsang Province, South Korea	JUWANG	EU267232	–
C. kyebangensis	Yongmun-san (1,157 m), Gyonggi Province, South Korea	YONGMUN	EU267233	EU267223
C. kyebangensis	Cheongok-san (1,256 m), Gangwon Province, South Korea	CHEONGOK	EU267234	EU267224
C. kyebangensis	Chiak-san (=Chiak-san Natl. Park, 1,288m), Gangwon Province, South Korea	CHIAK	EU267235	EU267225
C. kyebangensis	Worak-san (=Wolak-san Natl. Park, 1,093 m), Chungcheong Province, South Korea	WEORAK	EU267236	–
C. kyebangensis	Yumyeong-san (864 m), Gyonggi Province, South Korea	YUMYEONG	EU267237	EU267226
C. kyebangensis	Hwaak-san (1,468 m), Gyonggi Province, South Korea	HWAAK	EU267238	EU267227
C. kyebangensis	Myeongji-san (1,267 m), Gyonggi Province, South Korea	MYEONGJI	EU267239	EU267228
C. kyebangensis	Gyebang-san (1,577 m), Gangwon Province, South Korea	GYEBANG	EU267240	EU267229
C. kyebangensis	Seorak-san (=Seorak-san Natl. Park, 1,708 m), Gangwon Province, South Korea	SEORAK	EU267241	EU267230
C. kyebangensis	Songri-san (=Songni-san Natl. Park, 1,058 m), Chungcheong Province, South Korea	SONGRI	EU267242	EU267231
C. parvus†	Manchuria, Northeast China	MANCHURIA	EU267243	–

† *Cryptocercus parvus* was used as an outgroup species for Korean local populations.

Extraction Kit (Bioneer) and manufacturer-supplied protocols.

DNA Amplification, Purification and Sequencing

We amplified the target DNA sequences using the polymerase chain reaction (PCR) using 2.5 μl of 1 to 10 dilution of the genomic DNA as a template. We used a universal primer set for the amplification of the ITS-2 sequences as the follows: (forward) 5'-ACCCTGGACGGTGGATCA-3' and (reverse) 5'-GTTGGTTTC-TTTTCCTCC-3'. For the amplification of AT region we used the following primer set: (forward) 5'-TAGGGTATCTAATCCTAGTT-3' and (reverse) 5'-TGGGGTATGAACCCAGTAGC-3' (Taylor et al. 1993). We used Top DNA polymerase Kit (Bioneer) for PCR mixture, and the amplification of PCR were conducted using a PTC 100 (BIO-RAD). The temperature profiles for amplifying nuclear ITS-2 and AT region were as the follows: denaturation at 94℃ for 10 min, and 35 cycles of 94℃ for 0.5 min, 47～56℃ (47～56℃ for AT region and 48～52℃ for ITS-2) for 1.3 min and 72℃ for 2 min, followed by a final extension of 72℃ for 10 min. PCR products were purified using LaboPass^{TM} PCR kit (COSMO Co, Ltd.). The purified products were then used as templates for sequencing. Sequencing reactions were performed using an ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer) according to the manufacturer's instructions, and sequences were determined by automatic sequencing on a 3730 DNA Sequencer (ABI).

Data Analysis

Nucleotide diversity and other information about the DNA sequences were obtained using the ARLEQUIN 3.0 package (Excoffier et al. 2005). Nucleotide diversity is computed as the probability that two randomly chosen homologous (nucleotide) sites are different (Tajima 1983). It is equivalent to gene diversity at the nucleotide level for DNA data. The pairwise difference, or mean number of differences between all pairs of haplotypes in the sample, was calculated according to Tajima (1983). Pairwise genetic distances (*p*-distances) between haplotypes were calculated using the HKY85 model (Hasegawa et al. 1985) using PAUP 4.0b10 (Swofford 2002). We tested for significance of differences in the average pairwise genetic distances using an independent *t*-test.

The DNA sequences were aligned using the Clustal-X program package (Thompson et al. 1997) and manually edited using SeAl 1.0a (Rambaut 1996) and BioEdit version 5.0.6. Sequence datasets of ITS-2 and AT region were constructed separately, and then combined. Each dataset and the combined dataset for the two genes were used in all analyses. For analysis of genetic structure and networks, we used neighbor joining (NJ) with PAUP4.0b10 (Swofford 2002) and minimum spanning network (MSN) using the ARLEQUIN 3.0 package (Excoffier et al. 2000). The minimum spanning network was computed from the matrix of pairwise differences calculated between all pairs of haplotypes using a modification of the algorithm described in Rohlf (1973).

RESULTS

DNA Sequence Information

Nuclear ITS-2 sequences (406 bp) were collected from populations at 12 localities including one representing an outgroup species, *C. parvus*. Mitochondrial AT region sequences (471 bp) were also collected from populations in 9 of 11 localities. The combined dataset for ITS-2 and AT region sequences from these nine populations therefore included 877 bp (Fig. 1).

Analysis of the aligned sequences detected 5 haplotypes in the ITS-2 of 11 local populations and 9 haplotypes in the AT regions of 9 local populations, resulting in a total of 9 haplotypes in the combined sequences of 9 local populations from which mitochondrial AT region sequences were available (Fig. 1).

In the ITS-2, we detected nucleotide substitutions in four polymorphic sites (Fig. 1, Table 2). The substitutions in the ITS-2 regions included only nucleotide transitions. We also detected an insertion/deletion mutation in the ITS-2 region differentiating the *C. kyebangensis* samples from MANCHURIA (Manchurian outgroup sample) (Fig. 1). YUMYEONG, WEORAK and HWAAK showed higher sequence similarity to MANCHURIA than *Cryptocercus* populations from the other localities, whereas dissimilarity with ITS-2 sequences of MANCHURIA was highest in the GYEBANG (Fig. 1).

There were nucleotide transitions in 9 sites and transversions in 3 sites of 471 nucleotide sites in AT region (Fig. 1, Table 2). Nucleotide diversity was about 2.5 times higher in the AT region (0.00932 ± 0.00576) than the ITS-2 (0.00367 ± 0.00271) (Table 2), and the average pairwise genetic distances were 0.003697 ± 0.003023 (n = 55, range = 0.00000～0.00996) in the ITS-2 and 0.007995 ± 0.003077 (n = 36, range = 0.00213～0.01511) in the AT region (Tables 3 and 4). The mean genetic divergence in the AT region sequences was significantly higher than that in the ITS-2 sequences (t-test, p<0.005).

Genetic Network of Local Populations

Neighbor Joining Analysis

Neighbor joining tree of ITS-2 sequences showed two major clades (Clades I and Ⅱ) in the Korean local populations (Fig. 2). Clade I included two subclades *a* and *b*. There was no sequence variation among populations within each subclade in Clade I (Fig. 1). A nucleotide substitution (G↔A) was detected in only one site (site 386 in Fig. 1) between the subclades. In pairwise distance comparisons with the outgroup species (Table 3), subclade *b* had the lowest genetic distance (n = 4, 0.07277 ± 0.000), whereas populations of Clade Ⅱ (n = 3, 0.07563 ± 0.00287) and subclade *a* (n = 4, 0.07562 ± 0.000) had similar genetic distance values. The genetic distance of the GYEBANG was the highest (0.0785) among all pairwise comparisons with the outgroup species. NJ analysis showed that the populations of subclade *b* (Clade I) were more closely related to MANCHURIA than were the populations of Clade Ⅱ.

The unrooted NJ tree based on AT region sequences (Fig. 3), clustered the CHIAK and CHEONGOK together and the SONGRI and MYEONGJI together, while the GYEBANG was highly divergent from the other populations. The genetic distances within the two clusters were both 0.00213. Each genetic distance within these clusters was much lower than the mean genetic distance for all of the populations (n = 36, 0.007995 ± 0.003077, range = 0.00213～0.01511). The GYEBANG had the highest distance value (n = 8, 0.01209 ± 0.001997, range = 0.00856～0.01511) in pairwise comparisons of genetic distance among populations (Table 4).

Minimum Spanning Network (MSN)

A minimum spanning network (MSN-A) based on the ITS-2 sequences showed that the SEORAK and Group I populations were separated from MANCHURIA by substitutions in 60 of 406 nucleotides (Fig. 4). Within South Korea, there might be two dispersal routes of *C. kyebangensis* populations. One dispersal starts from SEORAK, followed by a spread to GYEBANG, which is distinguished from SEORAK by two nucleotide substitutions. The other dispersal may have started with the dispersal of Group I across the western side of South Korea, and then the subsequent emergence of Group Ⅱ, which is distinguished from Group I by a single nucleotide difference. Connection of the two dispersal routes occurred via the SONGRI (Fig. 4, MSN-A).

A second minimum spanning network was constructed based on the combined dataset including both ITS-2 and AT region sequences (Fig. 4, MSN-B). Sequences from the AT region were not obtained from JUWANG and WEORAK. Since high divergence caused alignment problem between AT sequences of *C. kyebangensis* and MANCHURIA (*C. parvus*), the sequence from MANCHURIA was also excluded from our dataset. Thus the corresponding ITS-2 sequences were excluded from the combined dataset, resulting in a genetic network in 9 of 12 local populations (MSN-B). This network identified two additional clusters of local populations, indicated in shaded circles in Fig. 4 (MSN-B), including a group consisting of the MYEONGJI and SONGRI, which are distinguished by two substitutions, and a group consisting of the CHEONGOK and CHIAK, which are distinguished by a single nucleotide substitution.

Group I of MSN-A was linked to the SEORAK and GYEBANG via the SONGRI in MSN-B. In contrast to MSN-A, MSN-B found no linkage between the SEORAK and GYEBANG. Groups Ⅰ and Ⅱ of MSN-A were distinguished by substitutions in 3 nucleotides

	ITS region																																												
	146	212	213	214	217	218	219	220	222	223	243	244	245	246	247	281	290	310	311	312	313	314	346	348	349	350	352	353	355	361	362	363	365	372	374	375	377	379	381	382	383	385	386	388	389	391
YUMYEONG	A	C	C	T	G	T	G	G	C	T	T	A	C	A	T	C	G	C	A	C	T	T	T	T	C	G	C	G	A	C	C	C	C	G	A	G	T	T	T	C	T	T	G	A	A	C
WEORAK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HWAAK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MYEONGJI	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
JUWANG	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.
YONGMUN	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.
CHEONGOK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.
CHIAK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.
SONGRI	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
GYEBANG	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	G	.
SEORAK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
MANCHURIA	G	G	G	G	−	−	−	−	T	C	−	−	−	−	−	−	C	G	−	−	−	T	C	C	T	C	T	A	T	−	−	−	G	T	C	C	G	G	C	G	A	A	G	C	A	T

	AT region												Haplotype No.		
	139	150	197	318	321	326	373	381	412	414	458	468	ITS	AT	ITS and AT
YUMYEONG	C	T	G	A	C	T	T	G	T	A	T	G	1	1	1
WEORAK	?	?	?	?	?	?	?	?	?	?	?	?	1	−	−
HWAAK	T	.	.	G	T	.	.	.	.	.	C	.	1	2	2
MYEONGJI	.	.	.	.	T	.	C	.	.	.	C	.	1	3	3
JUWANG	?	?	?	?	?	?	?	?	?	?	?	?	2	−	−
YONGMUN	T	.	.	.	.	.	.	.	.	.	C	A	2	4	4
CHEONGOK	.	.	.	.	T	.	.	.	.	G	C	.	2	5	5
CHIAK	.	.	.	.	.	.	.	.	.	G	C	.	2	6	6
SONGRI	.	.	.	.	.	.	C	.	.	.	C	.	3	7	7
GYEBANG	T	A	T	.	.	C	C	.	.	.	C	.	4	8	8
SEORAK	T	.	.	.	.	.	.	T	C	.	C	.	5	9	9
MANCHURIA	?	?	?	?	?	?	?	?	?	?	?	?			

-Dots indicate nucleotide identity in each site with the sequece sites of *C. kyebangensis* (YUMYEONG). The site numbers on the sequence indicate nucleotide positions in 406 sites of ITS sequence and 471 sites of AT sequence, respectively. *Cryptocercus parvus* was used as an outgroup species for *C. kyebangensis* populations.

Fig. 1. DNA sequence information and haplotype patterns from the nuclear ITS-2 region and the mitochondrial AT region from *C. kyebangensis* populations.

Table 2. Information of haplotype sequences from ITS-2, AT region and the combined dataset of the two genes

Gene	Sample size	No. of haplotypes	No. of nucleotide sites	No. of polymorphic site	Transition/ Transiversion	Mean no. of pairwise difference (Mean ± SD)	Nucleotide diversity (Mean ± SD)
ITS-2	11	5	406	4	4/0	1.491 ± 0.972	0.00367 ± 0.00271
AT	9	9	471	12	9/3	4.389 ± 2.391	0.00932 ± 0.00576
ITS-2 + AT	9	9	877	16	13/3	6.056 ± 3.184	0.00691 ± 0.00412

- Nucleotide diversity and other information of DNA sequences were obtained using the ARLEQUIN 3.0 package (Excoffier et al. 2005). Nucleotide diversity is computed as the probability that two randomly chosen homologous (nucleotide) sites are different (Tajima 1983). It is equivalent to the gene diversity at the nucleotide level for DNA data.
- Mean number of pairwise difference between all pairs of haplotypes in the sample was calculated according to Tajima (1983).

Table 3. Pairwise genetic distance between nuclear ITS-2 sequences of *C. kyebangensis* local populations

	JU-WANG	YONG-MUN	CHEONG-OK	CHIAK	YUM-YEONG	WEO-RAK	HWA-AK	MYEONG-JI	SONG-RI	GYE-BANG	SEO-RAK	MAN-CHURIA
JUWANG	−											
YONGMUN	0	−										
CHEONGOK	0	0	−									
CHIAK	0	0	0	−								
YUMYEONG	0.00247	0.00247	0.00247	0.00247	−							
WEORAK	0.00247	0.00247	0.00247	0.00247	0	−						
HWAAK	0.00247	0.00247	0.00247	0.00247	0	0	−					
MYEONGJI	0.00247	0.00247	0.00247	0.00247	0	0	0	−				
SONGRI	0.00495	0.00495	0.00495	0.00495	0.00247	0.00247	0.00247	0.00247	−			
GYEBANG	0.00745	0.00745	0.00745	0.00745	0.00996	0.00996	0.00996	0.00996	0.00745	−		
SEORAK	0.00745	0.00745	0.00745	0.00745	0.00495	0.00495	0.00495	0.00495	0.00247	0.00495	−	
MANCHURIA	0.07562	0.07562	0.07562	0.07562	0.07277	0.07277	0.07277	0.07277	0.07563	0.0785	0.07277	−

- Pairwise genetic distances (*p*-distances) between haplotypes were calculated using HKY85 model (Hasegawa et al. 1985) implemented using PAUP 4.0b10 (Swofford 2002).
- The sequence length for pairwise genetic distance was 406 nucleotide size.

in MSN-B. In MSN-B, Group Ⅱ was linked to the SONGRI either by the MYEONGJI or the YUMYEONG.

DISCUSSION

DNA Sequence Information of the Nuclear ITS-2 and Mitochondrial AT region

Evolutionary rates of DNA sequences vary among genes and regions within each gene. Thus it is very important to select appropriate molecular markers for evolutionary studies of ecosystems at the molecular level. The choice of molecular markers needs to be considered according to the taxonomic level (e.g. genus, species and population) of target samples as well as the study aims.

Park et al. (2004) used sequences from two mitochondrial genes, COⅡ (cytochrome oxydase Ⅱ) and 16S (16S ribosomal DNA), for analyses of molecular evolution and evolutionary biogeography of Korean (*C. kyebangensis*) and Manchurian (*C. parvus*) wood-eating cockroach species. However, the molecular markers chosen provided little information about geographic distributions and molecular evolution at the population level, although geographic patterns and biogeography could be relatively well explained at the species level.

The results of this study suggest that the ITS-2 and AT region are appropriate molecular markers for elucidating *C. kyebangensis*

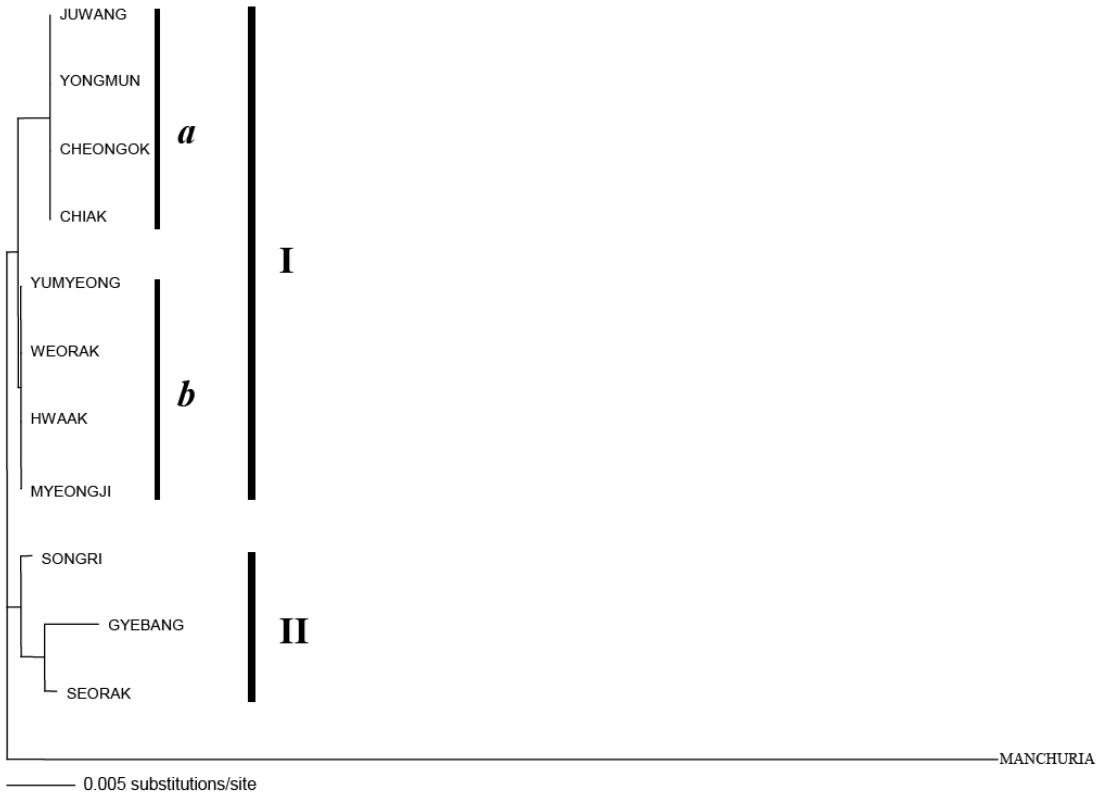

Fig. 2. Neighbor joining tree of *C. kyebangensis* haplotypes inferred from ITS-2 sequences. *Cryptocercus parvus* from Manchuria (MANCHURIA) was used as the outgroup. Substitution rates for Tv/Ti were calibrated using the HKY85 model (Hasegawa et al. 1985).

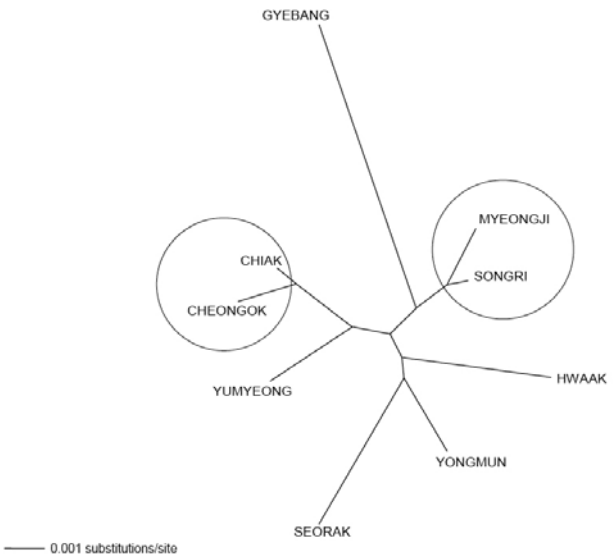

Fig. 3. Unrooted neighbor joining tree of *C. kyebangensis* haplotypes inferred from AT sequences. Substitutions rates for Tv/Ti were calibrated using the HKY85 model (Hasegawa et al. 1985). The CHIAK was closely related to the CHEONGOK and the MYEONGJI was related to the SONGRI. Genetic distances were 0.00213 for both sets of populations (Table 4).

geographic patterns at the population level. In comparison with protein coding genes like mitochondrial COⅡ, the evolution of AT region and ITS-2 sequences should be controlled by different selective constraints (Sbisa et al. 1997, Randi and Lucchini 1998, Yang 1998). In our analysis of the ITS-2 and AT regions, nucleotide substitutions seem to have occurred rapidly at various positions of the two gene sequences relative to the protein-coding gene (COⅡ) (refer to Park et al. 2004 for COⅡ gene sequence information). In particular, the AT region of mtDNA seems to accumulate transversion mutations (Tv) more rapidly than the protein coding gene, thus causing relatively deep sequence divergence in phylogenetic analyses which weigh Tv more heavily, like HKY85 and Kimura-2 parameter (Kimura 1980, Hasegawa et al. 1985, Randi and Lucchini 1988, Lockard et al. 1995, Kimball et al. 1999). Thus the AT region is a better molecular marker than the COⅡ gene for biogeographic analyses of local populations that diverged relatively recently.

A comparison of the two datasets shows that the AT region of mitochondrial DNA has evolved more rapidly than the nuclear ITS-2 in the study populations. The number of polymorphic sites was about 3 times higher in the AT region than ITS-2, and nucleotide

Table. 4. Pairwise genetic distance between mitochondrial AT gene sequences of *C. kyebangensis* local populations

	YONGMUN	CHEONGOK	CHIAK	YUMYEONG	HWAAK	MYEONGJI	SONGRI	GYEBANG	SEORAK
YONGMUN	-								
CHEONGOK	0.00861	–							
CHIAK	0.00644	0.00213	–						
YUMYEONG	0.00644	0.00644	0.00428	–					
HWAAK	0.00644	0.00644	0.00861	0.00861	–				
MYEONGJI	0.00861	0.00428	0.00644	0.00644	0.00644	–			
SONGRI	0.00644	0.00644	0.00428	0.00428	0.00861	0.00213	–		
GYEBANG	0.01073	0.01511	0.01291	0.01291	0.01291	0.01073	0.00856	–	
SEORAK	0.00641	0.01076	0.00858	0.00858	0.00858	0.01076	0.00858	0.01288	–

- Pairwise genetic distances (*p*-distances) between haplotypes were calculated using HKY85 model (Hasegawa et al. 1985) implemented using PAUP 4.0b10 (Swofford 2002).
- The sequence length for pairwise genetic distance was 471 nucleotide size.
- Genetic distances were lowest between CHIAK and CHEONGOK and between SONGRI and MYEONGJI.

diversity was about 2.5 times higher in the former than the latter. The average pairwise genetic divergence was significantly higher for AT region sequences than for ITS-2 sequences. In addition, our results showed that the AT region of mtDNA accumulated more Tv than ITS did (Ti/Tv = 4/0 in ITS-2; Ti/Tv = 9/3 in AT region). Since analyses of genetic divergence generally put greater weight on Tv, the AT region permits detection of relatively deeper sequence divergence than ITS-2.

Genetic Networks of ITS-2 and AT Region Sequences

One of most interesting results of the genetic network analysis was that two possible dispersal routes would be possible for dispersal of Korean *Cryptocercus* populations into South Korea from Manchuria (Fig. 4, MSN-A). According to the NJ tree based on the ITS-2 sequences (Fig. 2), two main lineages (Clade Ⅰ and Ⅱ) are found within South Korea as well. The NJ analysis of AT region sequences showed that the GYEBANG was most genetically distinct from the other populations (Fig. 3, Table 4), whereas pairwise genetic distances based on AT region showed close genetic relationships between the MYEONGJI and SONGRI and between the CHIAK and CHEONGOK (Table 4). The MSN-A (MSN based on the ITS-2 dataset) suggested the same basic pattern of relationships among local populations as the MSN-B (based on the combined dataset of ITS-2 and AT region). However, the two networks display interesting differences in the linkage patterns for the GYEBANG. In the MSN-A the GYEBANG is linked to the SEORAK, but the MSN-B links the GYEBANG to the SONGRI, but not to the SEORAK. As expected from the unrooted NJ tree of AT region (Fig. 3), the MSN-B indicated few substitutions between the MYEONGJI and SONGRI and between the CHIAK and CHEONGOK.

In the comparison of the NJ trees, Clade *a* and *b* were separated relatively well in the NJ of ITS-2, but not that of AT region. This could be caused by both small sample size and rapid evolution of AT region. The average pairwise genetic divergence and nucleotide diversity were significantly higher for AT region sequences than for ITS-2 sequences. Thus the AT region of mtDNA accumulated more complex haplotypes and deeper divergence than ITS-2. In this case, higher resolution in NJ of AT region could be obtained by addition of more samples. It indicates that ITS-2 should be more appropriate molecular marker for aims of this study based on small sample size than AT region.

In a previous study (Park et al. 2004) mitochondrial COⅡ and 16S were not appropriate molecular markers for migration or dispersal at local population level because of slow evolution of their nucleotides, though they could explain migration routes of large scale well. This study showed that ITS-2 slowly evolved than AT region, but evolved rapidly than COⅡ and 16S. Thanks to more rapid evolution of ITS-2 than the latter markers, the NJ tree of ITS-2 could get higher resolution than those COⅡ and 16S (Park et al. 2004).

We used small sample size in this study. Since AT region evolves more rapidly than ITS-2, it might have deeper divergence and complex haplotypes. In this case more samples and sites need to be added for high-resolved relationship among populations. Thanks to slower evolution of ITS-2 than AT region, ITS-2 could show more clear relationships among local populations, though we used small sample size.

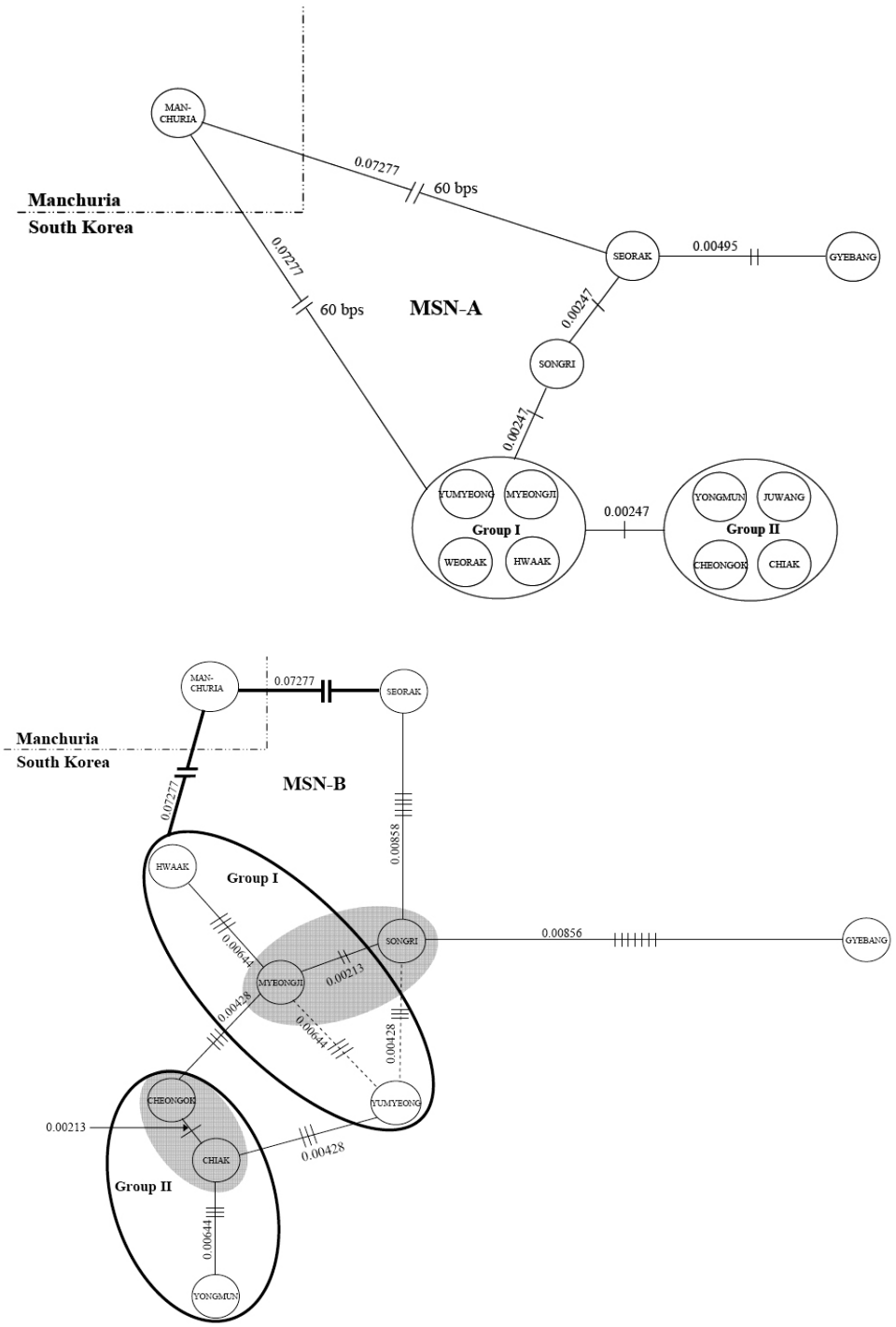

Fig. 4. Minimum spanning networks based on ITS-2 (MSN-A) and the combined dataset of ITS-2 and AT region (MSN-B). MSN-A suggests two possible routes, Group Ⅰ~Ⅱ and SEORAK-GYEBANG, for Korean *Cryptocercus* dispersal. In MSN-A, numbers on the lines indicate values of the pairwise genetic distances between the connected ITS-2 haplotypes (refer to Table 3). In MSN-B, bold lines indicate groupings based on only ITS-2 (refer to MSN-A). Shaded circles indicate groupings based on only AT sequences. MSN-B, like the unrooted NJ tree (Fig. 3), groups the CHIAK and CHEONGOK and the MYEONGJI and SONGRI, respectively. In contrast to MSN-A, however, the MSN-B links the GYEBANG to SONGRI rather than SEORAK. In MSN-B, numbers on the lines indicate values of the pairwise genetic distances of each connected groups based on the combined dataset of ITS-2 and AT region. Dot lines indicate alternative routes of the minimum spanning networks.

We inferred possible dispersal routes of Korean local populations from the MSNs and NJ trees (Fig. 5). We inferred two migration routes into South Korea from Manchuria via the network relationship between *C. kyebangensis* and MANCHURIA (*C. parvus*). Since higher divergence between AT regions of *C. kyebangensis* and *C. parvus* caused alignment problem and might also make errors like long branch attraction, we used only ITS-2 for these routes from Manchurian to South Korea. For a scenario for dispersal routes within South Korea, we depended on MSN-A and -B (Fig. 4).

In our scenario, some ancestral groups of *C. kyebangensis* (the ancestors of Group Ⅰ and Ⅱ) in the northern forests of Manchuria migrated southward to the mid-northern forests of South Korea (Hwaak- san and Myeongji-san), and then some of these populations dispersed further to the southern forests (Yumyeong-san and Weorak- san). Other populations dispersed to the eastern and southern forests (Yongmun-san, Chiak-san, Cheongok-san and Juwang-san) (Fig. 5). Different ancestral groups (the ancestors of SEORAK and GYEBANG) dispersed via the eastern mountainous forests of the Korean Peninsula. In our scenario, they migrated southward to Seorak-san, and then to Gyebang-san. Connection of the two clades (populations in Group I and in Seorak-san and Gyebang-san) was then possible in Songri-san. However, in contrast to MSN-A, MSN-B detected no linkage between the populations in Gyebang-san and those in Seorak-san, though Gyebang-san is located closer to the southern part of Seorak-san than the other populations.

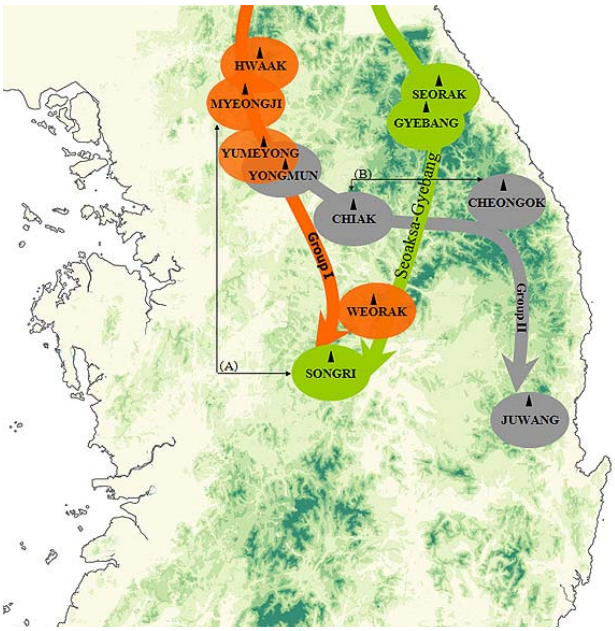

Fig. 5. Scenarios for possible migration routes of ancestral *Cryptocercus* into South Korea. Ancestral groups of *C. kyebangensis* (the ancestors of Group I) in northern forests migrated southward to the mid-northern forests of South Korea (Hwaak-san), and then some of the populations dispersed more to the southern forests (Myeongji-san, Yumyeong-san, Weorak-san) while others migrated to the southeastern forests (Group Ⅱ). Other ancestral groups (the ancestors of *C. kyebangensis* in Seorak-san and Gyebang-san) used the eastern mountainous forests of the Korean Peninsula (the SEORAK-GYEBANG line). They migrated southward to Seorak-san, and then to Gyebang-san. Connection of the two clades (the populations in Group I and in the Seorak-san and Gyebang-san) was possible in Songri-san. (A) and (B) on the map indicate the groupings of the MYEONGJI and SONGRI and the CHIAK and CHEONGOK described in MSN-B, respectively.

Park et al. (2004) suggested the eastern high mountainous forests of the Korean Peninsula as a possible migration route for ancestral *C. kyebangensis* into South Korea. In the pairwise genetic distances of the ITS-2 region, however, the average genetic distance from the outgroup (MANCHURIA) was lower in populations from the western forests (0.07277, Subclade *b*) than those from the eastern forests (0.07563, Clade Ⅱ) of South Korea.

This suggests that the western populations are more closely related to the outgroup species. In addition, the MSN-A based on the ITS-2 showed two possible routes, the Hwaak-san and Myeongji-san route and the Seorak-san and Gyebang-san route, for migration of ancestral *C. kyebangensis* into South Korea. In this study, we included only a small sample of populations from the eastern and southern mountains. Broader sampling of populations from these regions should be conducted to develop a more complete genetic network for analyses of the biogeography and distribution patterns of *Cryptocercus* in Korea.

ACKNOWLEDGMENTS

This study was supported by Ewha Womans University Research Fund 2007. We'd like to express our special thanks for technical support and help to Dr. Kiyoto Maekawa, Toyama University, Japan.

LITERATURE CITED

Bey-Bienko GI. 1950. Fauna of the USSR. Insects. Blattodea. Institute of Zoology, Academy of Sciences, Moscow, 342 P. (in Russian).

Cleveland LR, Hall SR, Sanders EP, Collier J. 1934. The wood-feeding roach *Cryptocercus*, its protozoa, and the symbiosis between protozoa and roach. Mem Am Acad Arts Sci 17: 85-342.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinformatics Online 1:47-50.

Grandcolas P, Legendre F, Park YC, Belles X, Murienne J, Pellens R. 2005. The genus *Cryptocercus* in East Asia: distribution and new

species (Insecta, Dictyoptedra, Blattaria, Polyphagidae). Zoosystema 27: 725-732.

Hasegawa M, Kishino H, Yano T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22: 160-174.

Kimball RT, Braun EL, Zwartjes PW, Crowe TM, Ligon JD. 1999. A molecular phylogeny of the pheasants and partridges suggests these lineages are polyphyletic. Mol Phylogenet Evol 11: 38-54.

Kimura M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. J Mol Evol 16: 111-120.

Lockhart PJ, Penny D, Meyer A. 1995. Testing the phylogeny of swordtail fishes using split decomposition and spectral analysis. J Mol Evol 41: 666-674.

Maekawa K, Park YC, Lo N. 2005. Phylogeny of endosymbiont bacteria harbored by the woodroach *Cryptocercus* spp. (Cryptocercidae: Blattaria): Molecular clock evidence for a late Cretaceous-early Tertiary split of Asian and American lineages. Mol Phyl Evol 36: 728-733.

Nalepa CA. 1984. Colony composition, protozoan transfer and some life history characteristics of the woodroach *Cryptocercus punctulatus* Scudder (Dictyoptera: Cryptocercidae). Behav Ecol Sociobiol 14: 273-279.

Park YC, Grandcolas P, Choe JC. 2002. Colony composition, social behavior and some ecological characteristics of the Korean wood-feeding cockroach (*Cryptocercus kyebangensis*). Zoo Sci 19: 1133-1139.

Park YC, Choe JC. 2003. Life history and population dynamics of Korean woodroach (*Cryptocercus kyebangensis*) populations. Korean J Biol Sci 7: 111-117.

Park YC, Maekawa K, Matsumoto T, Santoni R, Choe JC. 2004. Molecular phylogeny and biogeography of the Korean woodroaches. *Cryptocercus* spp. Mol Phylogenet Evol 30: 450-464.

Rambaut A. 1996. Se-Al: Sequence Alignment Editor. Available at http://evolve.zoo.ox.ac.uk/

Rohlf FJ. 1973. Algorithm 76. Hierarchical clustering using the minimum spanning tree. The Computer Journal 16: 93-95.

Randi, Lucchini, 1998. Organization and evolution of the mitochondrial DNA control-region in the avian genus *Alectoris*. J Mol Evol 47: 449-462.

Seelinger G, Seelinger U. 1983. On the social organization, alarm and fighting in the primitive cockroach *Cryptocercus punctulatus* Scudder. Z Tierpsychol 61: 315-333.

Sbisa E, Tanzariello F, Reyes A, Pesole G, Saccone C. 1997. Mammalian mitochondrial D-loop region structural analysis: Identification of new conserved sequences and their functional and evolutionary implications. Gene 205: 125-140.

Swofford DL. 2002. PAUP*, Phylogenetic Analysis Using Parsimony (*and other Methods), Version 4.10b. Sinauer Associates, Sunderland, MA.

Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics 105: 437-460.

Taylor MFJ, McKechnie SW, Pierce N, Kreitman M. 1993. The Lepidopteran mitochondrial control region: Structure and evolution. Mol Bid Ed 10: 1259- 1212.

Tompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4879-4882.

Yang Z. 1998. On the best evolutionary rate for phylogenetic analysis. Syst Biol 47: 125-133.

(Received October 31, 2007; Accepted November 17, 2007)