Genetic structure of *Lycorma delicatula* (Hemiptera: Fulgoridae) populations in Korea: implication for invasion processes in heterogeneous landscapes

Marana Park[1], Kyung-Seok Kim[2] and Joon-Ho Lee[1,3]*

[1]Entomology Program, Department of Agricultural Biotechnology, Seoul National University, Seoul, 151-921, Republic of Korea: [2]College of Veterinary Medicine, Seoul National University, Seoul, Republic of Korea: [3]Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul, 151-921, Republic of Korea

Abstract

Lycorma delicatula (White) was identified in 2004 as an invasive pest in South Korea, where it causes serious damage to vineyard crops. To investigate the population structure and dispersal pattern of *L. delicatula* in South Korea, we estimated the population genetic structure and gene flow among nine locations across the country using seven microsatellite markers. Although *L. delicatula* spread throughout most of its geographical range in South Korea within 5–7 years following invasion, its populations show evidence of genetic structuring across the range with a low but significant global F_{ST} (genetic differentiation across all populations) of 0.0474. Bayesian-based clustering analysis indicates the presence of at least three genetically unique populations in South Korea, including populations in northeastern South Korea, which show a distinct genetic background. However, isolation by distance suggests that populations in South Korea have not yet reached genetic equilibrium. Estimates of the historical rate of gene flow ($N_e\,m$) indicate that relatively high rates of flow have been maintained among populations within the western region, which may indicate recent range expansion. A population assignment test using the first-generation migrant detection method suggested that long-distance dispersal of *L. delicatula* may have occurred over large areas of South Korea. More complex dispersal patterns may have occurred during *L. delicatula* invasion of heterogeneous landscapes in South Korea.

Keywords: *Lycorma delicatula*, invasive pest, microsatellite, genetic structure

Introduction

The original distribution of *Lycorma delicatula* White (Hemiptera: Fulgoridae) was in northern China, in the

*Author for correspondence
Phone: +82 2 880 4705
Fax: +82 2 873 2319
E-mail: jh7lee@snu.ac.kr

Shanxi, Shandong, and Hubei provinces (Liu, 1939). *L. delicatula* was first detected as a non-indigenous species in the mid-western region of South Korea in 2004 (Kim & Kim, 2005). *L. delicatula* has spread from this region and become abundant across South Korea (Han *et al.*, 2008), where it causes crop damage, particularly in vineyards. The geographical expansion of *L. delicatula* has continued to the eastern region of South Korea, but it remains unclear where the initial introduction occurred and what is the source location of spread to other regions. Range expansion by *L. delicatula* is speculated to

have been asymmetric from west to east, which may be the result of geographical barriers including the mountain range dividing the western and eastern regions of South Korea.

The first instar nymphs of *L. delicatula* appear in May; these molt four times, becoming adults in late July. Mating, ovi-positing, and the death of adults occur prior to winter, and the eggs overwinter (Park *et al.*, 2009). Successful establishment of this species is thought to be associated with its overwintering abilities and a recent increase in Korean winter temperatures (Lee *et al.*, 2011). As the range of *L. delicatula* increased rapidly within South Korea, it is thought that movement occurs by both short-range expansion into adjacent areas and also by long-distance dispersal among distant sites, reflecting a stratified dispersal pattern. Short-distance dispersal of *L. delicatula* may be related to different host plant preferences between nymphs and adults. The host plant preferences of *L. delicatula* change during its growth cycle, with a broad range of host plants being fed upon during the nymph stages, but only a few plant species, including *Ailanthus altissima*, acting as food sources in the adult stage (Kim *et al.*, 2011). Therefore, the short-range dispersal behavior of *L. delicatula* may be influenced by the spatial distribution of available host plants. However, its long-distance dispersal ability and the pattern of *L. delicatula* range expansion remain unknown. Knowledge of the rate of range expansion and the mode of dispersal is required to enable mitigation and pest control strategies to be devised.

Molecular genetic markers enable estimation of the genetic diversity, movement of individuals (Kim *et al.*, 2008), inbreeding, and historical patterns of dispersal (Miller *et al.*, 2005). Investigations of population demography using molecular genetic marker data have been facilitated by the development of cost-effective methods of data acquisition, as well as the development of statistical approaches that have improved the capacity to estimate the proportion of a population that has moved various distances (Hastings *et al.*, 2005). Multilocus genotyping techniques using microsatellite markers have proven to be useful tools for understanding the biology of invasive species. Using microsatellite markers to estimate gene flow is a powerful alternative population assignment technique that has the potential to complement direct methods for measuring contemporary migration (Kim & Sappington, 2006). Using analogous methods, we conducted a population genetics study of *L. delicatula* in South Korea, analyzing microsatellite marker data to estimate gene flow and genetic structuring.

Material and methods

Study insect and sample collection

L. delicatula was collected in 2011 from nine locations throughout their current distributional range in Korea, including the initial occurrence locations (table 1, fig. 1). The sedentary behavior of *L. delicatula* nymphs on host plants enabled collection of one individual *L. delicatula* specimen from each host plant, *Vitis vinifera* L. (Vitaceae) or *A. altissima* Swingle. Sampled plants were at least 5 m apart to avoid collection of full siblings. The collected specimens were placed in 95% ethanol and stored at –20°C until DNA extraction was performed.

Microsatellite genotyping

DNA was extracted from the stored *L. delicatula* specimens using the Qiagen DNeasy Blood and Tissue Kit (Qiagen,

Table 1. Sampling information for *L. delicatula* specimens collected in South Korea during 2011.

Sample site	Sample name	First detection	Sampling year	Coordinates
Seoul	SE	2006	2011	N37°27′ E126°56′
Suwon	SW	2008	2011	N37°15′ E126°59′
Samcheok	SC	2010	2011	N37°26′ E129°09′
Chuncheon	CC	2008	2011	N37°52′ E127°44′
Cheonan	CA	2004	2011	N36°52′ E127°10′
Okcheon	OC	2006	2011	N36°18′ E127°34′
Gunsan	GS	2010	2011	N35°57′ E128°30′
Gwang-Ju	GJ	2010	2011	N35°09′ E126°55′
Daegu	DG	2010	2011	N35°52′ E128°30′

Valencia, CA, USA), and the eluted DNA template was diluted tenfold with deionized water. Seven microsatellite loci previously developed for *L. delicatula* by Park *et al.* (2012) were used for genotyping. Multiplex polymerase chain reaction (PCR) was conducted in two separate reactions: (i) for markers LD-D4, LD-D5, LD-T1, and LD-T3; and (ii) for markers LD-D1, LD-D2, and LD-T2. For these reactions, we used the i-star Taq PCR Kit (Qiagen) in a total volume of 10 μl, which contained 5.55 μl distilled water, 1.0 μl 10 × buffer, 0.8 μl dNTPs, 0.2 μl of each primer, 0.05 Taq, and 1 μl template DNA. The PCR profiles followed a 'touchdown' protocol (Don *et al.*, 1991), whereby an initial denaturation of 15 min at 95°C was followed by seven cycles of PCR, each consisting of 30 s denaturation at 94°C, 90 s annealing at 67°C, 60 s extension at 72°C, and a 0.5°C decrease per cycle. A total of 25 cycles were then run with 1 min denaturation at 60°C.

Statistical analysis

Genetic variation and genetic structure

Micro-Checker (Van Oosterhout *et al.*, 2004) was used to evaluate potential scoring errors resulting from stuttering, large allele drop-out, and null alleles in the *L. delicatula* microsatellite genotypes. The mean number of alleles per locus and the observed (H_O) and expected (H_E) heterozygosities were calculated using the Microsatellite Toolkit (Park, 2001). Multiple comparisons were made after applying the sequential Bonferroni correction (Rice, 1989). The Genepop program (Raymond & Rousset, 1995) was used to test deviations from Hardy–Weinberg equilibrium (HWE) conditions.

The program Structure v. 2.3.1 (Pritchard *et al.*, 2000) was used to estimate the most likely number of clusters for the datasets by determining the change in the marginal likelihood of the data Pr (X/K), where K was fixed at different values. The range of possible clusters (K) tested was set from 1 to 10, with five iterations. The lengths of the Markov Chain Monte Carlo (MCMC) iteration and burn-in were set at 100,000 and 200,000, respectively. We used an ancestry model allowing for admixture and correlated allele frequency among populations. The K-value was estimated using the maximal value of the log-likelihood [ln Pr (X/K)] of the posterior probability of the data for a given K (Pritchard *et al.*, 2000). The second-order rate of change in the log probability of the data between successive values of ΔK (the 'true' number of K within the *L. delicatula* sample dataset) was also calculated using $\Delta K = m \ |L''(K)| / s[L(K)]$ (Evanno *et al.*, 2005).
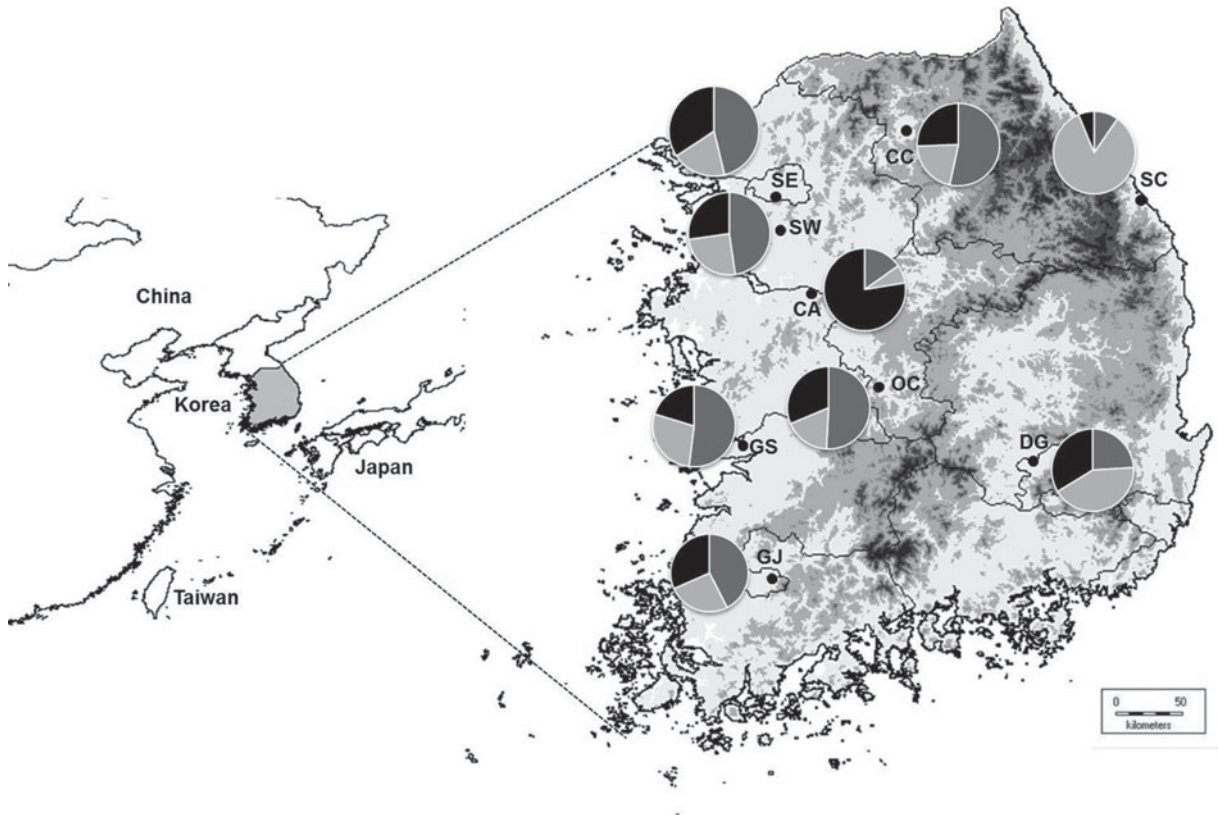
Fig. 1. Sampling locations for *L. delicatula* in South Korea, 2011. The pie graphs show the results of a Bayesian cluster analysis of multilocus microsatellite genotypes. Each location is partitioned into *K* = 3 components.

Table 2. Genetic variability estimates for each *L. delicatula* population, inferred from seven microsatellite loci. Number of alleles, expected heterozygosity (H_E) at HWE, observed heterozygosity (H_O), inbreeding coefficient (F_{IS}), probability (*P*-value) of being in HWE, and loci showing potential null alleles.

Population	Sample size	No. of alleles	Allelic richness	H_O	H_E	F_{IS}	*P*-value[1]	Loci with null alleles
SW	30	6.71	6.443	0.637	0.709	0.119	0.0024	LD-D2, LD-T3
DG	27	5.29	5.206	0.603	0.711	0.155	0.0008	LD-D4
GS	28	7	6.709	0.630	0.725	0.133	0.0016	LD-T2
OC	31	6.57	6.021	0.695	0.707	0.017	0.3183	LD-T3
SE	30	6.86	6.370	0.692	0.731	0.053	0.1103	LD-D2
GJ	29	5.86	5.661	0.589	0.668	0.12	0.004	LD-D4, LD-T3
SC	30	5	4.904	0.532	0.654	0.189	0.0016	LD-D2
CC	26	6.71	6.599	0.599	0.757	0.212	0.0008	LD-D2, LD-T2
CA	29	5.86	5.590	0.596	0.662	0.101	0.0135	LD-D2

[1] Hardy–Weinberg exact test (Raymond & Rousset, 1995) with Bonferroni correction (*P* = 0.00079).

We carried out a principal coordinate analysis (PCoA) using the Genalex program (Peakall & Smouse, 2006). A scatter diagram was plotted based on factor scores along the two PCo axes accounting for most variation.

Gene flow measures

Indirect estimates of the historical rates of gene flow between populations (N_em) were calculated according to the relationship $N_em = (1 - F_{ST})/4F_{ST}$ (Wright, 1931), where N_em is the effective number of migrants per generation, N_e is the effective population size, and *m* is the migrant rate. Pairwise estimates of the genetic differentiation (F_{ST}) between populations were made using FSTAT v. 2.9.3 (Goudet, 2001). As Micro-Checker revealed the potential occurrence of null alleles on at least one locus for each population (see table 2), the FreeNA program (Chapuis & Estoup, 2007) was used to estimate F_{ST}, which was adjusted for null alleles (excluding null alleles), and the result was compared with that of F_{ST} assuming no null alleles.

Isolation by distance (IBD) was tested by regressing pairwise population estimates of linearized $F_{ST}/(1 - F_{ST})$

Table 3. Pairwise estimates of genetic differentiation (F_{ST}) (below the diagonal) between *L. delicatula* populations, and gene flow ($N_e m = (1 - F_{ST})/4F_{ST}$) inferred from each estimate (above diagonal).

	SW	DG	GS	OC	SE	GJ	SC	CC	CA
SW	–	7.0599	39.4325	Infinity	29.8705	7.6364	3.0137	34.9613	4.1283
DG	0.0342*[1]	–	5.564	8.6786	7.4187	9.75	2.0186	10.1234	4.681
GS	0.0063[NS]	0.043*	–	17.9982	10.1234	3.9922	2.1219	13.5622	2.2252
OC	− 0.0008[NS]	0.028*	0.0137[NS]	–	Infinity	11.3241	3.0439	41.4167	4.2383
SE	0.0083[NS]	0.0326*	0.0241*	− 0.0007[NS]	–	6.7332	2.801	34.9613	4.8416
GJ	0.0317*	0.025*	0.0605*	0.0216*	0.0358*	–	2.03519	11.2179	3.835
SC	0.0766*	0.1102*	0.1054*	0.0759*	0.0817*	0.1094*	–	2.4011	1.682
CC	0.0071[NS]	0.0241*	0.0181[NS]	0.006[NS]	0.0071[NS]	0.0218[NS]	0.0943*	–	3.1983
CA	0.0571*	0.0507*	0.101*	0.0557*	0.0491*	0.0612*	0.1294*	0.0725*	–

[1] Probability of being different from zero following correction for multiple comparisons. *$P < 0.05$; NS, not significant. The adjusted nominal level (5%) for multiple comparisons was 0.001389.

(Rousset, 2000) on the natural log of the geographical distance between all pairs of sample locations, using the Mantel test implemented with Genalex software. Hierarchical partitioning of genetic variation was assessed using analysis of molecular variance (AMOVA) for populations and individuals. AMOVA provides an estimate of the proportion of genetic variation within and between populations.

Population assignment/exclusion tests were conducted by direct and simulation methods using the Geneclass2 program (Piry *et al.*, 2004) to detect genetic signatures of dispersal and immigration (Rannala & Mountain, 1997). The direct assignment test allocates an individual to one of the reference populations without probability computation. The test calculates the proportion of individuals correctly assigned to the most likely population of origin, even though the true population of origin is not among the reference populations. In contrast, the exclusion method uses a simulation approach in which the likelihood of a genotype occurring in the population is computed by simulating multilocus genotypes based on the allele frequencies of each reference population. In this method, the likelihood of the genotype of an individual is compared with the distribution of likelihoods of simulated genotypes for each reference population. If the genotype likelihood (a) of an individual is below a predetermined threshold (e.g. $a = 0.01$), the population is excluded as a possible origin of the individual (Cornuet *et al.*, 1999). Unlike the direct assignment method, the exclusion method does not assume that the true population of origin has been sampled because each population is treated independently (Cornuet *et al.*, 1999). Frequency probabilities of multilocus genotypes in each reference population were determined in the exclusion test using Monte Carlo simulations of 10,000 independent individuals for the population (Paetkau *et al.*, 2004). We followed a Bayesian statistical approach (Rannala & Mountain, 1997) using the Monte Carlo resampling method (Paetkau *et al.*, 2004).

To infer contemporary migration of individuals between populations, we employed the *detection of first-generation migrants* criterion implemented in Geneclass2 (Piry *et al.*, 2004), which assigns each potential individual that traveled from Site A to Site B in year X, or individuals born in year X to a gravid female that moved from Site A to Site B in year X − 1. As we do not know whether all source populations for immigrants were sampled in the current study, two test statistics (L_{home} and the ratio L_{home}/L_{max}) were used to compute the likelihood of migrant detection (L) (Paetkau *et al.*, 2004). The analysis was conducted using a simulation of 10,000 independent individuals at thresholds of $a = 0.05$ and $a = 0.01$. As the method of Paetkau *et al.* (2004) is intended to measure real-time migration between populations, it assumes that all populations are sampled in the same year.

Bottleneck tests

The Bottleneck (Cornuet & Luikart, 1996) program was used to assess the evidence for past bottlenecks. We used both a strict stepwise mutation model (SMM) (Ohta & Kimura, 1973) and a two-phase model (TPM) (Di Rienzo *et al.*, 1994). The bottleneck test analyzed the heterozygosity excess by comparing the observed number of alleles at each locus, assuming mutation–drift equilibrium. Estimated values were determined using a TPM (Piry *et al.*, 1999) with an 80% single-step mutation proportion, a variance among multiple steps of 12 and 5000 iterations. The probability of significant heterozygosity excess was determined using the Wilcoxon signed-rank test. We also used a model shift in allele frequency distribution as a qualitative indicator of population bottlenecks (Luikart *et al.*, 1998). The *M*-ratio of Garza & Williamson (2001), which is the mean ratio of the number of alleles to the range of allele size, was calculated using the AGARST program (Harley, 2001). The *M*-ratio has a long recovery time following a decline in population size (e.g. >100 generations), and so enables recent population reductions to be distinguished from those occurring a long time ago. Garza & Williamson (2001) suggested that the *M*-value and its variance across loci could be used as an alternative test for detecting reductions in population size over a much longer time frame.

Results

Genetic variability

A total of 86 alleles were detected across seven microsatellite loci for 260 *L. delicatula* individuals from among the nine locations in South Korea. Scoring errors resulting from large allele dropout were not detected in any *L. delicatula* population or locus, but Micro-Checker identified possible stuttering for the marker LD-D4 in population DG and marker LD-T2 in population GS. The presence of potential null alleles was indicated by a general excess of homozygotes for most allele size classes for one or two loci within at least one population (table 2). The genetic variability estimates for each *L. delicatula* population deduced from the seven microsatellite loci included allelic diversity, the observed (H_o) and expected

Table 4. *M*-ratio test results using the SMM (Ohta & Kimura, 1973) and the TPM (Di Rienzo *et al.*, 1994) to detect a recent population bottleneck event within each *L. delicatula* population. Significance tested using the Wilcoxon sum-rank test (α=0.05).

Population	Wilcoxon tests[1]		Mode shift	M^2
	TPM	SMM		
SW	0.7109	0.9609	Normal	0.499 (0.050)
DG	0.0078	0.1875	Shifted	0.497 (0.097)
GS	0.8556	0.9609	Normal	0.578 (0.050)
OC	0.9726	0.9961	Normal	0.559 (0.067)
SE	0.7656	0.9804	Normal	0.539 (0.053)
GJ	0.7656	0.9726	Normal	0.540 (0.055)
SC	0.2891	0.8516	Normal	0.537 (0.055)
CC	0.2891	0.8516	Normal	0.503 (0.030)
CA	0.9563	0.7656	Normal	0.537 (0.047)

[1] One-tail probability for an excess or deficit of observed heterozygosity relative to the expected equilibrium heterozygosity, computed from the observed number of alleles under mutation–drift equilibrium.
[2] M = mean ratio of the number of alleles to the range of allele size (Garza & Williamson, 2001); variance in parentheses.
SMM, stepwise mutation model; TPM, two-phase model of mutation.

Table 5. AMOVA among *L. delicatula* samples from nine locations in South Korea.

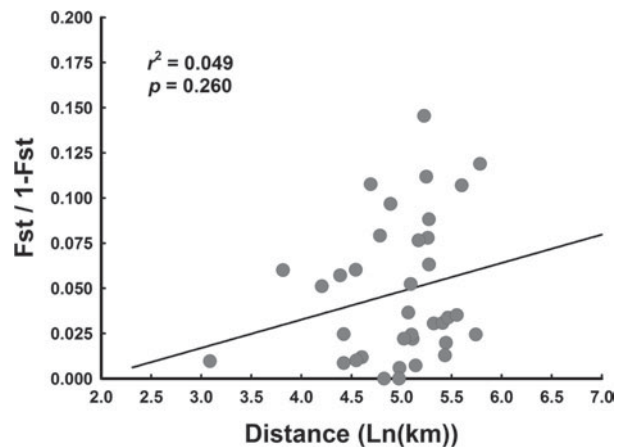| Source of variation	df	Sum of squares	Mean sums of squares	Estimated variance	%
Among populations	8	157.106	19.638	0.476	7
Individuals within populations	251	1479.002	5.892	5.892	93
Total		259	1636.108		6.368	100

Fig. 2. Geographical distance versus genetic distance ($F_{ST}/1 - F_{ST}$) for populations of *L. delicatula*, using pairwise F_{ST}. Correlations and probabilities were estimated from a Mantel test with 10,000 bootstrap repeats.

(H_E) heterozygosity, the inbreeding coefficient (F_{IS}), and the *P*-values for deviations from the HWE. Allelic richness varied from 4.904 to 6.709, and H_E ranged from 0.654 to 0.757. Six of nine populations exhibited a significant deviation from HWE following sequential Bonferroni correction for multiple testing. These six populations had positive F_{IS} values across loci with an excess of observed homozygotes (table 2).

Genetic structure within and among populations

The genetic differentiation between each pair of populations (uncorrected and corrected pairwise F_{ST}) and the effective number of migrants exchanged per generation ($N_e m$) are shown in table 3. Uncorrected estimates of pairwise F_{ST} values ranged from −0.0008 for the SW and OC populations (ENA corrected F_{ST}=0.0009; SE and OC populations) to 0.1294 for the SC and CA populations (ENA corrected F_{ST}=0.1365; CA and SC populations). Both estimates of F_{ST} (F_{ST} adjusted for null alleles and F_{ST} assuming no null allele results) were similar. Global estimates of F_{ST} across all loci and all populations were low but significant (uncorrected F_{ST}=0.0474, 95% CI=0.0327–0.0634; ENA corrected F_{ST}=0.0477, 95% CI=0.0338–0.0631). The $N_e m$ calculated from the uncorrected F_{ST} ranged from 1.68 (CA and SC) to infinity (OC and SW, SE

and OC), implying an intermediate to very high level of gene flow.

The *M*-ratio values, which are used to detect long-term bottleneck events, were generally low in all populations, ranging from 0.497 to 0.578 (table 4). The results indicated a significant bottleneck event for all populations of *L. delicatula* as recently invading species. However, deviation from mutation–drift equilibrium (under the TPM model) and mode shift revealed a signature of recent population reduction only for the DG population (P=0.0078).

AMOVA analysis among the *L. delicatula* samples revealed that most of the genetic variation was partitioned to among populations and individuals within populations. More than 93% of the total genetic variation was accounted for by individuals within a population and, correspondingly, 7% of the total genetic variation was among populations (table 5). No significant correlation was found between genetic and geographical distances among the populations, as evidenced by the Mantel tests of IBD over all samples (R^2=0.049, P=0.260), indicating recent range expansion and/or frequent gene flow among populations in South Korea (fig. 2).

In the PCoA, the mean factor scores for the nine populations were plotted along the first two principal component axes, which together accounted for 72% of the total variance (40.2% for axis 1 and 31.97% for axis 2; fig. 3). This analysis showed conspicuous divergence of the SC and CA populations from the other populations in South Korea.

Bayesian clustering revealed three clusters. The value of ΔK calculated from ln *P*(D) of the structure output revealed a maximum value of 23.41 for K=3 among the genotypes (fig. 4). The average distances among individuals in the same cluster were 0.38 for cluster 1, 0.30 for cluster 2, and 0.32 for cluster 3; the variance in the mean individual membership within each cluster was accounted for by between-sample site differences (fig. 1, table 6). This showed that co-ancestry of genotypes provides evidence for three distinct populations (fig. 4): (i) one encompassing most of Korea (populations SE, SW, GJ, GS, DG, and OC); (ii) population CA; and (iii) population SC. Population CA has a genetic structure that differs from that of the other western population.
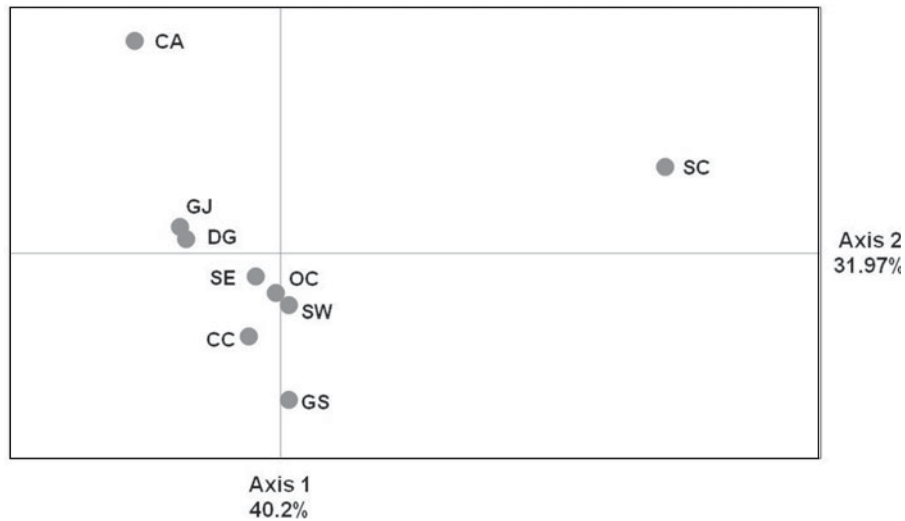
Fig. 3. Scatter diagram of factor scores from a PCoA of genotype data for seven microsatellite loci in samples of *L. delicatula* collected from nine locations in South Korea (see fig. 1). The percentage of total variation attributed to each axis is indicated.
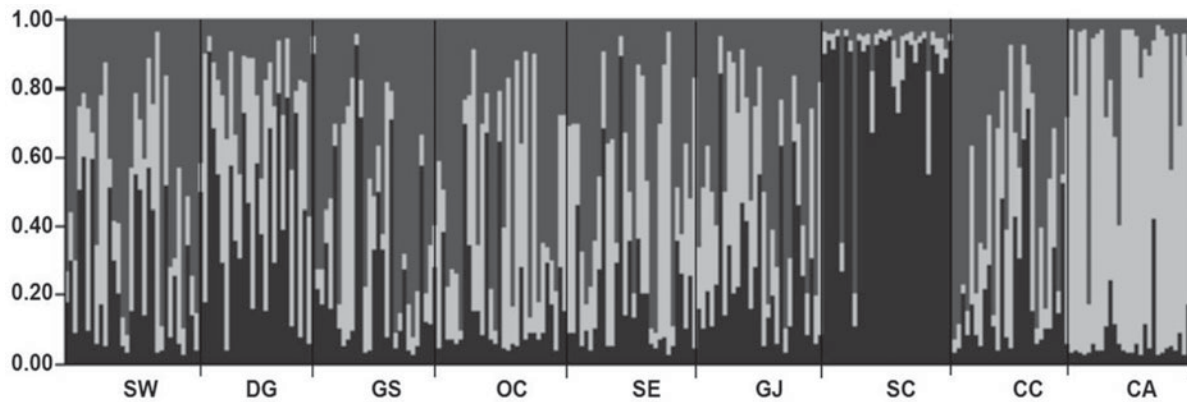

Fig. 4. Bar plot of population structure estimates for 260 *L. delicatula* specimens collected from nine locations in Korea, generated by structure. The maximum value among genotypes was 23.41 at $\Delta K = 3$, using $\Delta K = m|L''(K)|/s[L(K)]$ (Evanno *et al.*, 2005).

Table 6. Average coefficient of ancestry obtained from a structure analysis with $K=3$ for 260 *L. delicatula* specimens collected from nine sampling locations in South Korea.

Sample	Cluster		
	1	2	3
SW	0.477[1]	0.25	0.272
GS	0.52	0.278	0.203
OC	0.509	0.179	0.313
SE	0.462	0.195	0.343
GJ	0.425	0.259	0.316
CC	0.534	0.21	0.256
CA	0.152	0.072	0.776
SC	0.096	0.843	0.062
DG	0.24	0.423	0.337

[1] The highest value of co-ancestry for each population in a cluster is shown in bold.

Assignment/exclusion test and detection of first-generation migrants

We calculated the percentage of *L. delicatula* individuals sampled from each population and excluded as potential immigrants at a threshold of $a = 0.01$, and the mean assignment log-likelihood for each possible donor population (table 7). Populations from most locations contained members whose potential origins in other populations could be excluded with ≥99% certainty. Individuals from the SC population could be assigned to their own population with >90% certainty, and individuals from DG and CA populations were assigned to their own population with 74.1% and 65.5% certainty, respectively. In the exclusion test, the SC population could be excluded with 34.5–75.9% certainty (0.01 threshold) as a putative origin of all populations. The assignment and exclusion values were evenly distributed among the other populations. The mean estimated individual assignment likelihood

Table 7. Percentage of *L. delicatula* individuals assigned to and excluded from (i.e. determined to not be a potential immigrant from) each reference population and the mean assignment log-likelihood for individuals from each geographic population to possible source populations.

Sample location	Method[1]	Potential source (reference) population								
		SW	DG	GS	OC	SE	GJ	SC	CC	CA
SW	Assignment[2]	6.67 (2)	10 (3)	16.7 (5)	33.3 (10)	13.3 (4)	3.3 (1)	0 (0)	13.3 (4)	3.3 (1)
	Exclusion[3]	3.33 (1)	23.3(7)	3.3 (1)	3.3 (1)	0 (0)	16.7 (5)	40 (12)	3.3 (1)	33.3 (10)
	−LOG(L)[4]	8.22	10.28	8.48	**8.20**	8.60	8.81	11.23	8.42	9.76
DG	Assignment	0 (0)	74.1 (20)	7.4 (2)	0 (0)	3.7 (1)	11.1 (3)	3.7 (1)	0 (0)	0 (0)
	Exclusion	11.1 (3)	3.7 (1)	7.4 (2)	7.4 (2)	18.5 (5)	29.6 (8)	55.6 (15)	14.8 (4)	51.9 (14)
	−LOG(L)	9.60	7.45	9.55	9.68	10.06	**9.41**	12.40	9.66	11.33
GS	Assignment	14.3 (4)	7.1 (2)	42.9 (12)	21.4 (6)	0 (0)	0 (0)	3.6 (1)	10.7 (3)	0 (0)
	Exclusion	3.6 (1)	21.4 (6)	0 (0)	3.6 (1)	10.7 (3)	25 (7)	57.1 (16)	3.6 (1)	50 (14)
	−LOG(L)	8.81	10.23	8.35	**8.66**	9.41	9.82	12.24	9.07	11.03
OC	Assignment	25.8 (8)	3.2 (1)	13.0 (4)	22.6 (7)	13.0 (4)	6.5 (2)	0 (0)	6.5 (2)	9.7 (3)
	Exclusion	0 (0)	6.5 (2)	0 (0)	0 (0)	6.5 (2)	6.5 (2)	45.2 (14)	0 (0)	22.6 (7)
	−LOG(L)	**7.68**	9.28	7.94	7.76	8.06	8.53	11.41	8.18	9.17
SE	Assignment	10 (3)	3.3 (1)	6.7 (2)	16.7 (5)	36.7 (11)	6.7 (2)	3.3 (1)	13.3 (4)	3.3 (1)
	Exclusion	6.7 (2)	20 (6)	6.7 (2)	9.7 (3)	3.3 (1)	26.7 (8)	53.3 (16)	6.7 (2)	30 (9)
	−LOG(L)	8.94	10.64	9.50	**8.78**	8.43	9.71	12.49	8.95	10.28
GJ	Assignment	6.9 (2)	6.9 (2)	0 (0)	6.9 (2)	6.9 (2)	51.7 (15)	0 (0)	20.7 (6)	0 (0)
	Exclusion	0 (0)	6.8 (2)	0 (0)	0 (0)	0 (0)	3.4 (1)	34.5 (10)	0 (0)	17.2 (5)
	−LOG(L)	8.72	9.47	9.06	8.58	8.89	7.46	11.45	**8.27**	9.66
SC	Assignment	0 (0)	0 (0)	6.7 (2)	3.3 (1)	0 (0)	0 (0)	90 (27)	0 (0)	0 (0)
	Exclusion	10 (3)	36.7 (11)	6.7 (2)	23.3 (7)	16.7 (5)	46.7 (14)	0 (0)	16.7 (5)	63.3 (19)
	−LOG(L)	**10.15**	11.78	10.34	10.61	10.17	10.40	6.76	10.70	12.43
CC	Assignment	11.5 (3)	3.8 (1)	15.4(4)	11.5(3)	15.3 (4)	19.2 (5)	0 (0)	19.2 (5)	3.8 (1)
	Exclusion	3.8 (1)	38.5 (10)	0 (0)	3.9 (1)	0 (0)	30.8 (8)	73.1 (19)	0 (0)	65.4 (17)
	−LOG(L)	**9.10**	10.89	9.32	9.26	9.16	9.23	13.01	8.80	11.31
CA	Assignment	10.3 (3)	3.4 (1)	0 (0)	10.3(3)	0 (0)	6.9 (2)	0 (0)	3.4 (1)	65.5 (19)
	Exclusion	0 (0)	27.6 (8)	24.1 (7)	13.8 (4)	6.9 (2)	44.8 (13)	75.9 (22)	6.9 (2)	3.4 (1)
	−LOG(L)	**9.29**	11.42	11.62	9.96	9.61	10.98	13.87	10.77	7.70

[1] The assignment test was carried out using the direct approach without probability computation and the exclusion test was carried out using the simulation method (Cornuet *et al.*, 1999). Both tests used the Bayesian statistical approach described by Rannala & Mountain (1997). The simulation method developed by Paetkau *et al.* (2004) was used in the exclusion test.
[2] The number of individuals assigned to the most likely population is shown in parentheses.
[3] The number of individuals excluded from the reference population for $a = 0.01$ is shown in parentheses.
[4] Mean assignment −log likelihood (L) value for individuals from a given sample population. Bold value indicates the value most similar to that of the sample population, and therefore represents the population from which it most likely originated, under the assumptions of the test.

Table 8. Number of probable first-generation migrants identified in each population of *L. delicatula* and its putative source population at thresholds of $a = 0.05$ and $a = 0.01$ (in parentheses).

Population	Test statistic[1]	Putative source population								
		SW	DG	GS	OC	SE	GJ	SC	CC	CA
SW	L_{home}	–								
	L_{home}/L_{max}	–	1(1)	1(0)						1(1)
DG	L_{home}		–							
	L_{home}/L_{max}		–	1(0)						
GS	L_{home}			–						
	L_{home}/L_{max}	2(0)	1(1)	–					1(1)	
OC	L_{home}				–					
	L_{home}/L_{max}	2(0)	1(0)		–	2(0)			1(1)	1(1)
SE	L_{home}					–				
	L_{home}/L_{max}	1(0)			2(0)	–	1(1)		1(0)	1(1)
GJ	L_{home}						–			
	L_{home}/L_{max}					2(1)	–		1(0)	
SC	L_{home}							–		
	L_{home}/L_{max}			2(2)				–		
CC	L_{home}								–	
	L_{home}/L_{max}	1(1)		2(0)			1(1)		–	
CA	L_{home}									–
	L_{home}/L_{max}	2(0)	1(0)		1(1)		2(0)			–

The analysis used the assignment criterion described by Rannala & Mountain (1997) and the Monte Carlo resampling method reported by Paetkau *et al.* (2004).
[1] L = likelihood of migrant detection (Paetkau *et al.*, 2004).

indicated that the highest assignment likelihood of individuals of the OC population (apart from itself) come from the SW population (mean assignment log-likelihood = –7.68). Similarly, the highest assignment likelihood of SW individuals was from the OC population (–8.20) (table 7).

The number of immigrant individuals estimated for the current generation is summarized in table 8. When the L_{home}/L_{max} ratio was considered for the detection of first-generation migrants between sample locations, a total of 36 and 14 individuals were detected as probable first-generation migrants at thresholds of $a = 0.05$ and $a = 0.01$, respectively. This result suggests that the OC population received individuals from each of the SE, CC, and CA populations. The proposed dispersal pathway among populations is shown in fig. 5, based on the result of 'detection of first-generation migrants'. The results indicate that there was movement of *L. delicatula* to adjacent locations and, unexpectedly, that long-distance dispersal beyond the geographical barrier also occurred.

Discussion

Evidence from this study suggests that there is no significant difference in genetic diversity, allelic richness, or H_E among the *L. delicatula* populations in South Korea. In contrast, a low but significant level of genetic differentiation, based upon F_{ST} estimates, was observed among the populations. A general trend of low level but significant population differentiation has previously been reported among populations of the Chinese mitten crab that have recently established in Europe (Herborg *et al.*, 2007). Inbreeding coefficients (F_{IS}) indicate a significant difference between observed and expected heterozygosity as a consequence of non-random mating, the presence of a null allele, and/or the Wahlund effect (Wahlund, 1928). Analogously, a non-significant IBD

relationship is expected among populations recently introduced to South Korea (such as *L. delicatula*), which show little evidence of genetic divergence. The absence of IBD is indicative of a recent range expansion/high gene flow, or a lack of gene flow coinciding with extensive genetic differentiation among all populations. The latter can be excluded as a possible explanation in relation to the South Korean *L. delicatula* populations, as low levels of genetic differentiation were estimated among the sample populations (table 3). These analyses further suggest that the *L. delicatula* populations in South Korea have not reached equilibrium since colonization.

Although the IBD test has been widely used to investigate the spatial patterns of gene flow and genetic relatedness between populations (Wright, 1943), the presence of geographical barriers to dispersal can also limit gene flow. McRae (2006) suggested that the resistance distance of the two-dimensional geographical distances might be a more realistic reflection of the underlying barriers to dispersal, and that geographical distances that integrate heterogeneities in dispersal pathways might be more relevant. For heterogeneous landscapes, use of the resistance distance may help to reveal patterns of IBD that are absent from Euclidean distance estimates. For the spread stage of invading populations, the major outstanding study aspects include linking the traits of species to their movement ability and exploring the impacts of landscape heterogeneity on dispersal success (Puth & Post, 2005). Consequently, to gain a greater understanding of gene flow among populations of *L. delicatula*, our future studies (both laboratory and field) will focus on the dispersal abilities of the species.

The genetic structure analysis, based on PCoA and structure, indicated that three genetically divergent *L. delicatula* populations are present in South Korea. The results suggest that simultaneous establishment, rather than point or
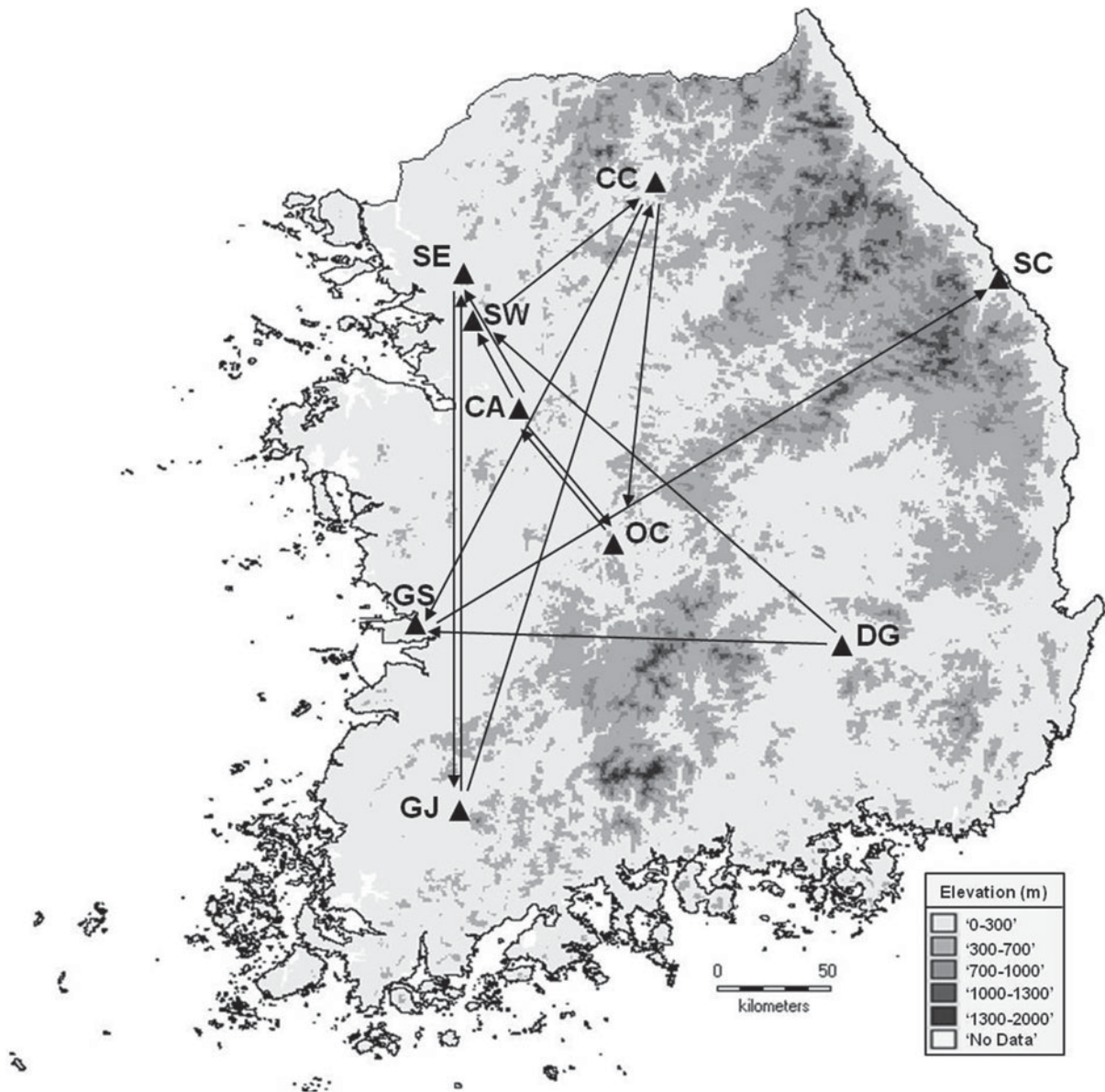
Fig. 5. Dispersal pathway of *L. delicatula* populations in South Korea. The arrows indicate the probable source and recipient populations of first generation migrants detected using the L_{home}/L_{max} statistic.

independent introduction, probably occurred in western South Korea, and that the subsequent outbreak can in part be attributed to colonization of the eastern region (populations SC and DG). It is possible that two temporally separated incursions of *L. delicatula* into the eastern regions occurred. While the CA population (which has an apparently different genetic structure) may have been introduced independently, additional studies are needed to clarify the mode of introduction; such studies should include samples from eastern China.

Assessing bottlenecks is important in determining founder effects underpinning demographical events, especially recent species invasions. Microsatellites are particularly informative in the study of recent population phenomena. A combination of approaches was applied to *L. delicatula*, based on a range of

microsatellite characteristics. The *M*-values for the studied populations were less than those expected from historically stable populations (0.82), and also below the *M*-ratio range (0.599–0.693) for populations that have undergone a historical reduction in population size or recently founding population (Garza & Williamson, 2001). Although the *M*-ratio for a stable population of *L. delicatula* needs to be established, the results indicate that in Korea this species was subject to a bottleneck event at some time in the past, as is expected for invasive species. Our mutation–drift equilibrium (under the TPM model) and mode-shift analyses indicated that a bottleneck occurred in the DG population. More reliable demographic evidence will require investigations of longer time scales of biological invasion, because short-term population structures

may provide only weak and/or incorrect indications of demographic events associated with invasion processes (Fitzpatrick *et al.*, 2012).

Although gene flow can serve as a surrogate for dispersal ability (Bohonak, 1999), it has limitations when movement cannot be verified because of insufficient genetic information. Moreover, high gene flow yields within open landscapes can produce populations that are genetically homogenous over great geographical distances (fig. 2). Our assignment/exclusion tests showed similarity of the −log (*L*) estimates (excluding the highest values) in the potential source population (table 7), suggesting that the majority of *L. delicatula* sampled in this study were not representative of the initial outbreak source population in South Korea. In a previous study, the possible source population of the introduced insect species *Ostrinia nubilalis* in North America was investigated by sampling each geographic population and assessing the qualitative information on gene flow (Kim *et al.*, 2009). However, there were limitations in the method used, as the source population could not be identified because of the high rates of dispersal among populations in the range expansion phase.

First-generation migrant detection can enable inferences to be made regarding the point origin of individuals, and provides the potential to estimate real-time dispersal through detection of immigrant individuals (Rannala & Mountain, 1997; Paetkau *et al.*, 2004). The probable source and recipient populations of first-generation *L. delicatula* migrants were investigated using the L_{home}/L_{max} statistic, which revealed that dispersal probably occurred over long distances as well as among adjacent locations (fig. 5). It is notable that more frequent dispersal may occur in the western region compared with the eastern region, and that long-distance dispersal may be occurring beyond the mountain range.

Since its first appearance in Cheonan (western Korea) in 2004, the distribution of *L. delicatula* has expanded across South Korea in only 5–7 years. In this study, we applied a combination of population genetics analyses to study the invasion dynamics. The presence of a mountain range appears to have caused only a minor delay in the dispersal of *L. delicatula* from west to east across the country. The dispersal patterns may have influenced the invasion dynamics with naturogenic deceleration (heterogeneous topography) and anthropogenic acceleration (human activity and ground transportation) of gene flow during the spread phase of *L. delicatula* in South Korea.

Understanding of the invasion process and dispersal pathways is important in the development of effective quarantine measures and for improving biosecurity in relation to invasive insect pests. Genetics should play a greater role in the development of policy to manage and control invasive species because it is important in understanding the invasion process and developing defense mechanisms against invading pests (Allendorf & Lundquist, 2003). Our study indicates that the Korean Peninsula is susceptible to invasive by insects from adjacent countries. Following species invasions, subsequent establishment and spread within South Korea can occur if conducive climatic and ecological conditions are present. For example, the rice water weevil *Lissorhoptrus oryzophilus* Kuschel spread throughout the paddy fields of Korea in a relatively short time period. Rapid dispersal was possible because the host plant is widely distributed, and long-distance dispersal was facilitated by ground transport and human activity in Korea (Lee & Uhm, 1993). This example underscores the importance of a specific framework for preventing the range expansion of invasive pests, based on knowledge of the invasion dynamics in the heterogeneous landscape of Korea.

Invasion is often cryptic in terrestrial habitats during the initial phase of introduction, and detection of pests typically occurs during the dispersal stages, when the species begins to damage native biota. Generalizations about the dynamics of establishment and spread of invasive species are difficult, and consequently assessment of population structure and movement are fundamental to understanding the potential effects of pest introductions within specific regions. Temporal analysis of the structure of *L. delicatula* populations is needed to clarify the expansion dynamics of this invasive species.

References

Allendorf, F.W. & Lundquist, L.L. (2003) Introduction: population biology, evolution, and control of invasive species. *Conservation Biology* **17**, 24–30.

Bohonak, A.J. (1999) Dispersal, gene flow, and population structure. *Quarterly Review of Biology* **74**, 21–45.

Chapuis, M.P. & Estoup, A. (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* **24**, 621–631.

Cornuet, J.M. & Luikart, G. (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**, 2001–2014.

Cornuet, J.M., Piry, S., Luikart, G., Estoup, A. & Solignac, M. (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**, 1989–2000.

Di Rienzo, A., Peterson, A., Garza, J.C., Valdes, A., Slatkin, M. & Freimer, N.B. (1994) Mutational processes of simple-sequence repeat loci in human populations. *Genetics* **91**, 3166–3170.

Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K. & Mattick, J.S. (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research* **19**, 4008.

Evanno, G., Regnaut, S. & Goudet, J. (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14**, 2611–2620.

Fitzpatrick, B.M., Fordyce, J.A., Neimiller, L. & Reynolds, G. (2012) What can DNA tell us about biological invasions? *Biological Invasions* **14**, 245–253.

Garza, J.C. & Williamson, E.G. (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* **10**, 305–318.

Goudet, J. (2001) Fstat, a program to estimate and test gene diversities and fixation indices (Version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html

Han, J.M., Kim, H.J., Lim, E.J., Lee, S.H., Kwon, Y.J. & Cho, S.W. (2008) *Lycorma delicatula* (Hemiptera: Auchenorrhyncha: Fulgoridae: Aphaeninae) finally, but suddenly arrived in Korea. *Entomological Research* **38**, 281–286.

Harley, E.H. (2001) *AGARst. A Program for Calculating Allele Frequencies, GST, and RST from Microsatellite Data, version 2*. South Africa, University of Cape Town.

Hastings, A., Cuddington, K., Davies, K.F., Dugaw, C.J., Elmendorf, S., Freestone, A., Harrison, S., Holland, M., Lambrinos, J., Malvadkar, U., Melbourne, B.A., Moore, K., Taylor, C. & Thomson, D. (2005) The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters* **8**, 91–101.

Herborg, L.M., Weetman, D., Van Oosterhout, C. & Hänfling, B. (2007) Genetic population structure and contemporary dispersal patterns of a recent European invader, the Chinese mitten crab, *Eriocheir sinensis*. *Molecular Ecology* **16**, 231–242.

Kim, J., Lee, E.H., Seo, Y.M. & Kim, N.Y. (2011) Cyclic behavior of *Lycorma delicatula* (Insecta: Hemiptera: Fulgoridae) on host plants. *Journal of Insect Behavior* **24**, 423–435.

Kim, K.S. & Sappington, T.W. (2006) Molecular genetic variation of boll weevil populations in North America estimated with microsatellites: implications for patterns of dispersal. *Genetica* **127**, 143–161.

Kim, K.S., Ratcliffe, S.T., French, B.W. & Sappington, T.W. (2008) Utility of EST-derived SSRs as population genetics markers in a beetle. *Journal of Heredity* **99**, 112–124.

Kim, K.S., Bagley, M.J., Coates, B.S., Hellmich, R.L. & Sappington, T.W. (2009) Spatial and temporal genetic analyses show high gene flow among European corn borer (Lepidoptera: Crambidae) populations across the central U.S. corn belt. *Environmental Entomology* **38**, 1312–1323.

Kim, S.S. & Kim, T.W. (2005) *Lycorma delicatula* (White) (Hemiptera : Fulgoridae) in Korea. *Lucanus* **5**, 9–10.

Lee, I.Y. & Uhm, K.B. (1993) Landing, settling and spreading of the rice water weevil in Korea. pp. 42–57 *in* Hirai, K. (*Ed.*) *Establishment, Spread, and Management of the Rice Water Weevil and Migratory Rice Insect Pests in East Asia*. NARC, Tsukuba.

Lee, J.S., Kim, I.K., Koh, S.H., Cho, S.J., Jang, S.J., Pyo, S.H. & Choi, W.I. (2011) Impact of minimum winter temperature on *Lycorma delicatula* (Hemiptera: Fulgoridae) egg mortality. *Journal of Asia-Pacific Entomology* **14**, 123–125.

Liu, G. (1939) Some extracts from the history of entomology in China. *Psyche* **46**, 23–28.

Luikart, G., Allendorf, F., Cornuet, J.M. & Sherwin, W. (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* **89**, 238–247.

McRae, B.H. (2006) Isolation by resistance. *Evolution* **60**, 1551–1561.

Miller, N., Estoup, A., Toepfer, S., Bourguet, D., Lapchin, L., Derridj, S., Kim, K.S., Reynaud, P., Furlan, L. & Guillemaud, T. (2005) Multiple transatlantic introductions of the western corn rootworm. *Science* **310**, 992.

Ohta, T. & Kimura, M. (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetical Research* **22**, 201–204.

Paetkau, D., Slade, R., Burden, M. & Estoup, A. (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13**, 55–65.

Park, J.D., Kim, M., Lee, S.G., Shin, S.C., Kim, J. & Park, I.K. (2009) Biological characteristics of *Lycorma delicatula* and the control effects of some insecticides. *Korean Journal of Applied Entomology* **48**, 53–57.

Park, M., Kim, K.S. & Lee, J.H. (2012) Isolation and characterization of eight microsatellite loci from *Lycorma delicatula* (White) (Hemiptera: Fulgoridae) for population genetic analysis in Korea. *Molecular Biology Reports* **39**, 5637–5641.

Park, S.D.E. (2001) *Trypanotolerance in West African cattle and population genetic effects of selection*. PhD thesis, University of Dublin, Dublin, Ireland.

Peakall, R. & Smouse, P.E. (2006) Genalex6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288–295.

Piry, S., Luikart, G. & Cornuet, J.M. (1999) Bottleneck: a program for detecting recent effective population size reductions from allele data frequencies. *Journal of Heredity* **90**, 502–503.

Piry, S., Alapetite, A., Cornuet, J.M., Paetkau, D., Baudouin, L. & Esop, A. (2004) Geneclass2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity* **95**, 536–539.

Pritchard, J.K., Stephens, M. & Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Puth, L.M. & Post, D.M. (2005) Studying invasion: have we missed the boat? *Ecology Letters* **8**, 715–721.

Rannala, B. & Mountain, J.L. (1997) Detecting immigration by using multilocus genotypes. *Genetics* **94**, 9197–9201.

Raymond, M. & Rousset, F. (1995) An exact test for population differentiation. *Evolution* **49**, 1280–1283.

Rice, W.R. (1989) Analyzing tables of statistical tests. *Evolution* **43**, 223–225.

Rousset, F. (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology* **13**, 58–62.

Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M. & Shipley, P. (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**, 535–538.

Wahlund, S. (1928) Zusammensetzung von populationen und korrelationserscheinungen vom standpunkt der vererbungslehre aus betrachtet. *Hereditas* **11**, 65–106.

Wright, S. (1931) Evolution in Mendelian populations. *Genetics* **16**, 290.

Wright, S. (1943) Isolation by distance. *Genetic* **28**, 114–138.