Genetic structure and demographic history of *Lymantria dispar* (Linnaeus, 1758) (Lepidoptera: Erebidae) in its area of origin and adjacent areas

Tae Hwa Kang1 | Sang Hoon Han2 | Heung Sik Lee3

1Bio Control Research Center, Jeonnam Bioindustry Foundation, Gokseong-gun, Korea
2Department of Life Science, College of Natural Science, Kyonggi University, Suwon, Korea
3Plant Quarantine Technology Center, Animal and Plant Quarantine Agency, Gimcheon-si, Korea

Correspondence
Heung Sik Lee, Plant Quarantine Technology Center, Animal and Plant Quarantine Agency, Gimcheon-si, Korea.
Email: lhsgo@korea.kr

Funding information
Animal and Plant Quarantine Agency, Grant/Award Number: B-1541785-2013-15-02

Abstract
We analyzed the population genetic structure and demographic history of 20 *Lymantria dispar* populations from Far East Asia using microsatellite loci and mitochondrial genes. In the microsatellite analysis, the genetic distances based on pairwise $F_{ST}$ values ranged from 0.0087 to 0.1171. A NeighborNet network based on pairwise $F_{ST}$ genetic distances showed that the 20 regional populations were divided into five groups. Bayesian clustering analysis ($K = 3$) demonstrated the same groupings. The populations in the Korean Peninsula and adjacent regions, in particular, showed a mixed genetic pattern. In the mitochondrial genetic analysis based on 98 haplotypes, the median-joining network exhibited a star shape that was focused on three high-frequency haplotypes (Haplotype 1: central Korea and adjacent regions, Group 1; Haplotype 37: southern Korea, Group 2; and Haplotype 90: Hokkaido area, Group 3) connected by low-frequency haplotypes. The mismatch distribution dividing the three groups was unimodal. In the neutral test, Tajima's D and Fu's FS tests were negative. We can thus infer that the Far East Asian populations of *L. dispar* underwent a sudden population expansion. Based on the age expansion parameter, the expansion time was inferred to be approximately 53,652 years before present (ybp) for Group 1, approximately 65,043 ybp for Group 2, and approximately 76,086 ybp for Group 3. We propose that the mixed genetic pattern of the inland populations of Far East Asia is due to these expansions and that the inland populations of the region should be treated as valid subspecies that are distinguishable from other subspecies by genetic traits.

KEYWORDS
demographic history, Far East Asia, *Lymantria dispar*, population genetic structure, species origin region

1 INTRODUCTION

The gypsy moth, *Lymantria dispar* (Linnaeus, 1758), originating from Hokkaido, Japan (Bogdanowicz, Mastro, Prasher, & Harrison, 1997; Bogdanowicz, Schaefer, & Harrison, 2000; Goldschmidt, 1934, 1940), is widely distributed in the Palearctic region (Pogue & Schaefer, 2007; Schintlmeister, 2004). There are three subspecies: *L. dispar dispar*, *L. dispar asiatica* Vnukovskij, 1926, and *L. dispar japonica* Motschulsky,
1860 (Pogue & Schaefer, 2007). Lymnatria dispar dispar is mainly distributed in Europe, L. dispar asiatica occurs from Central Asia to East Asia, and L. dispar japonica is present only in Japan (Pogue & Schaefer, 2007). Among these subspecies, the validity of the scientific name L. dispar asiatica (Figure 1) has been debated by many authors (Lee, Kang, Jeong, Ryu, & Lee, 2015). Schintlmeister treated L. dispar asiatica as a synonym of L. dispar dispar on the basis of their type locality; however, Pogue and Schaefer treated the subspecies as valid based on the morphological characteristics of the females, which have larger wings than the females of L. dispar dispar (Lee et al., 2015; Pogue & Schaefer, 2007; Schintlmeister, 2004). The dispersal ability of the two subspecies may differ because of these differences in wing size. Based on research of male deaths after interbreeding, Higashiura et al. (2011) accepted the five subspecies of Inoue (1982). Thus, the subspecies of L. dispar are clearly in a state of confusion.

Lymnatria dispar dispar was intentionally brought to North America for hybridization experiments; however, some individuals escaped in either 1868 or 1869 (Liebhold, Mastro, & Schaefer, 1989). Since then, the subspecies has become an invasive forest pest, causing injury to approximately 400 species of plants (Lowe, Browne, Boudjelas, & De Poorter, 2000; Pogue & Schaefer, 2007). Approximately US$11 million is spent on European gypsy moth control every year (Pimentel, Zuniga, & Morrison, 2005; Pogue & Schaefer, 2007). For these reasons, L. dispar asiatica, which has higher flight capability than L. dispar dispar, has been treated as a quarantine pest in North America (Pogue & Schaefer, 2007).

Due to the quarantine and danger this invasive species represents, studies of the differences among local populations are actively conducted. In previous decades, population genetic analyses of L. dispar were performed using various methods, such as allozyme detection, amplified fragment length polymorphism, restriction fragment length polymorphism, sequence-based analysis, and microsatellites (Bogdanowicz et al., 1997, 2000; deWaard et al., 2010; George, 1984; Kang, Lee, & Lee, 2015; Keena, Côté, Grinberg, & Wallner, 2008; Koshio, Tomishima, Shimizu, Kim, & Takenaka, 2002; Qian et al., 2014; Wu et al., 2015). Area of origin studies, in particular, using microsatellite loci were mainly conducted by North American researchers. In the first attempt by Bogdanowicz et al. (1997), four markers were developed and used to assay allelic variation in four gypsy moth populations (Japan, Far East Russia, China, and North America). Subsequently, Keena et al. (2008) evaluated flight capability and related traits using four microsatellite loci (from Bogdanowicz et al., 1997) and mitochondrial DNA analyses of samples obtained from 46 geographic strains. In Far East Asia, Koshio et al. (2002) compared the allele types of regional populations using three microsatellite loci of Japanese samples from three local populations; however, they did not consider population structure because of small sample sizes. Recently, Wu et al. (2015) thoroughly analyzed the population structure of the Holarctic gypsy moth and performed an origin test for each regional population using nine microsatellite loci, including three from Bogdanowicz et al. (1997).

These studies were conducted from the perspective of quarantine inspection (or invasive species control), and the number of sampled individuals was large; however, the number of sampled areas in each region was small, leading to taxonomic confusion with respect to the subspecies of L. dispar. For example, it was reported that two Asian subspecies, L. dispar asiatica and L. dispar japonica, were difficult to distinguish using morphological characters, with individuals of L. dispar asiatica collected from the southern coastal area of Korea having characteristics similar to L. dispar japonica (Lee et al., 2015; Pogue & Schaefer, 2007). To resolve this taxonomic confusion at the subspecific level, a demographic history of the Far East Asian populations of L. dispar based on intensive sampling is required. Therefore, the goal of this study was to reveal the population genetic structure and demographic history of L. dispar in Far East Asia, including in the region of species origin: Hokkaido, Japan. For this purpose, we analyzed the genetic diversity and demographic history of L. dispar from Far East Asia using eight microsatellite loci and three mitochondrial genes (cytochrome c oxidase I [COI], ATP6, and ATP8 genes).

Genetic diversity analyses using microsatellite loci have been conducted for various eukaryotes (Balloux & Lugon-Moulin, 2002; Sakai et al., 2001; Sunnucks, 2000). Recently, they have been used to track the influx of invasive species (Hess, Swalla, & Moran, 2008; Hunter & Hart, 2013; Keena et al., 2008; Kim et al., 2011; King, Eackles, & Chapman, 2011; Tóth, Gáspári, & Jurka, 2000). For the use of microsatellite loci, however, a primer set for each polymorphic locus is required. The general method employed is an enrichment strategy (López-Uribe, Santiago, Bogdanowicz, & Danforth, 2012; Richardson, Stanley, & Sherman, 2012), which is expensive and time-consuming, as it is based on traditional cloning strategies (Perry & Rowe, 2011; Santana et al., 2009; Zane, Bargelloni, & Patarnello, 2002). However, the next-generation sequencing (NGS) technique is very useful for the construction of microsatellite loci libraries at a lower cost and far more quickly than traditional cloning-based approaches (Hess et al., 2008;

**FIGURE 1** Adult habitus of Lymnatria dispar asiatica (a, male; b, female)
Kang, Han, & Park, 2016; Kang, Han, & Park, 2015; Perry & Rowe, 2011; Yu, Won, Jun, Lim, & Kwak, 2011). Because of the problems associated with the traditional cloning strategies, we used Illumina sequencing, one of the NGS techniques, for reading the genomic DNA of L. dispar and then developed microsatellite markers from the results.

2 | MATERIALS AND METHODS

2.1 | Sampling and genomic DNA extraction for NGS and pyrosequencing

For NGS, we extracted genomic DNA from an egg mass of L. dispar. The egg mass was collected from Suwon, Korea (37°14.092′N, 127°02.840′E; Figure 2b: Site A). In the egg mass, we selected 50 eggs and extracted genomic DNA using a NucleoSpin® Tissue Kit (Macherey-Nagel GmbH, Düren, Germany) following the manufacturer’s instructions. The sequencing was performed with a MiSeq Sequencer (Illumina, San Diego, CA, USA) by the DNA sequencing company DisGene (Daejeon, Korea). The resulting contigs were assembled in CLC workbench (CLC Bio, Aarhus, Denmark).

2.2 | Sampling and genomic DNA extraction for genetic structure analysis

For polymerase chain reaction (PCR) analysis of polymorphisms with the developed microsatellite markers and for genetic structure analysis, 552 samples were collected from 20 sites in Mongolia (1), Russia (1), China (1), Korea (12), and Japan (5) using pheromone attraction traps (Figure 2, Table 1). The thoracic muscle of each individual was removed for the extraction of genomic DNA. For morphological examination, fore and hind wings were prepared as specimens on a glue board. Abdomens were maintained at −20°C for examination of genitalia. Genomic DNA was extracted using a DNeasy® Blood & Tissue Kit (Qiagen, Leipzig, Germany) according to the manufacturer’s instructions.

2.3 | Microsatellite locus identification and marker development

Microsatellite loci were identified using Phobos ver. 3.3.12 (Leese, Mayer, & Held, 2008; Mayer, Leese, & Tollrian, 2010) with the following conditions: repeated sequence length, 2–4 base pairs (bp) and repeat count, greater than four. AT-rich loci were excluded from the investigated microsatellite loci, and for loci that were repeated more than six times, primer sets were chosen using the primer design software PRIMER 3 (Koressaar & Remm, 2007; Untergrasser et al., 2012) with the following criteria: melting temperature, 55.5–56.5°C; GC content, over 30%; and primer length, 18–22 bp. A hundred and fifty primer sets were designed, and PCR tested for specificity and the presence of polymorphic amplification using one sample from each of the twelve regional populations from Korea. PCRs for the primer qualification test were conducted with AccuPower PCR
PreMix (Bioneer, Daejeon, Korea) in a final volume of 20 μl containing 30 ng of template DNA and 5 pmol of each primer. Extra MgCl₂ was not added. The amplification profile was 5 min at 94°C; 30 cycles of 10 s at 94°C, 10 s at 56°C, and 20 s at 72°C; and a final 5 min extension at 72°C. The specificity and presence of polymorphic amplification for each primer set were checked using a QIAxcel DNA high-resolution cartridge (Qiagen, Leipzig, Germany). For the markers showing polymorphism in the electrophoresis, each forward (sense) primer for genotyping was labeled with 6-carboxyfluorescein at the 5’ end (Schuelke, 2000). Of the labeled markers, eight were selected for microsatellite marker assessment by a PCR amplification test. For microsatellite marker assessment, 432 samples from the 20 regional populations were genotyped (Table 1). These PCRs were performed by the DNA sequencing company Bionics (Seoul, Korea).

2.4 Mitochondrial DNA sequencing

For the analysis of L. dispar genealogy in Far East Asia, we selected three mitochondrial genes: COI, ATP6, and ATP8. The COI gene may not be suitable for population analysis because its intraspecific variation is relatively low and its interspecific variation is relatively high (Cameron & Whiting, 2008; Wu et al., 2015); however, when combined with other genes, it may be useful (Hajibabaei, Singer, Hebert, & Hickey, 2007). The ATP6 and ATP8 genes show relatively higher intraspecific variation and are known to be suitable for population genetic analysis (Cameron & Whiting, 2008; Wu et al., 2015). The former region of the COI gene was amplified using the LCO1490 (5’-GGTCAACAAATCAAATAAGATATTG-3’) and HCO2198 (5’TAAACTTCAGGGTGACCAAAATCA-3’) primer set (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994) and a GeneMax Tc-s-B PCR cycler (BIOER, Hangzhou, China). PCR conditions were set as in Hebert, Cywinska, Ball, and deWaard (2003). The ATP6 and ATP8 genes were amplified using the primer set from Wu et al. (2015) and an ABI Veriti 96-well Thermal Cycler (Applied Biosystems); Thermofisher Scientific Inc., MA, USA). PCR products were checked using 1% agarose gel electrophoresis. The PCR products were purified and sequenced using the sequencing services of Macrogen (Seoul, Korea) and Bionics (Seoul, Korea). The obtained sequences were submitted to NCBI GenBank (Table 1).

2.5 Microsatellite loci data analysis

Genotyping errors (such as null alleles and scoring errors) on selected markers were checked with MICRO-CHECKER ver. 2.2.3 (Oosterhout, Hutchinson, Wills, & Shipley, 2004). The pairwise linkage disequilibrium values for pairs of loci were then examined using Arlequin ver. 3.1 (Excoffier, Laval, & Schneider, 2005). Genetic diversity parameters such as allele frequency, genotype number, allele type, gene diversity, heterozygosity, and polymorphism information content (PIC) were calculated with PowerMarker ver. 3.5 (Liu & Muse, 2005). Hardy–Weinberg equilibrium (HWE) across loci was estimated after sequential Bonferroni correction (Rice, 1989). To test the isolation by distance (IBD) model, the correlation between genetic distance and geographic distance was calculated using Mantel’s test with 30,000 randomizations in IBD ver. 3.23 (Jensen, Bohonak, & Kelley, 2005). To estimate genetic differentiation among regional populations, analysis of molecular variance (AMOVA) was used. AMOVA was calculated using the Kimura two-parameter model in Arlequin ver. 3.1 (Excoffier et al., 2005). We ascertained the allele type frequencies based on microsatellite loci for each population and estimated the pairwise genetic distances between the populations based on allele type frequencies with PowerMarker ver. 3.5 (Liu & Muse, 2005). Based on the pairwise genetic distances, a network estimating the genealogical relations among the 20 regional populations was calculated with SplitTree4 (Huson & Bryant, 2006). We tested the genetic differentiation among the populations using a model-based Bayesian analysis with STRUCTURE ver. 2.3.4 (Falush, Stephens, & Pritchard, 2003; Pritchard, Stephens, & Donnelly, 2000) under the following conditions: a correlated-allele model with a 500,000 burn-in period, 750,000 MCMC reps after burn-in, K from 2 to 8, and 20 iterations. The value of the ad hoc statistics Δ(K) was then estimated with Harvester (Earl & von Holdt, 2012) using the average value of LnP(D) to estimate the number of genetic groups (Evanno, Regnaut, & Goudet, 2005).

2.6 Mitochondrial sequence data analysis

The obtained sequences were manipulated as a raw data set using MEGA 6 (Tamura, Stecher, Peterson, Filipski, & Kumar, 2013), and sequence divergence was estimated. The standard diversity indices (the number of haplotypes and polymorphic sites) were estimated using DnaSP ver. 5.10.01 (Librado & Rozas, 2009), and the raw data set was converted for analysis in Arlequin and NETWORK. The molecular diversity indices (haplotype diversity and nucleotide diversity) were estimated using Arlequin ver. 3.1 (Excoffier et al., 2005). To estimate the genealogical relations among the haplotypes, a median-joining network was calculated using NETWORK ver. 4.6.1.3 (http://www.fluxus-engineering.com). FST distances among all pairs in the population were used to assess the genetic structure of L. dispar asiatica. Population pairwise FST values were calculated using the Kimura two-parameter model in Arlequin ver. 3.1 (Excoffier et al., 2005) (significance test = 0.05; significance level = 1,000 permutations). To estimate genetic differentiation among regional populations, AMOVA was used following the Kimura two-parameter model in Arlequin ver. 3.1 (Excoffier et al., 2005). To test the IBD model, the correlation between genetic distance and geographic distance was calculated using Mantel’s test with 30,000 randomizations in IBD ver. 3.23 (Jensen et al., 2005).

To estimate the demographic history of the gypsy moth populations, a mismatch distribution analysis was conducted using Arlequin ver. 3.1 (Excoffier et al., 2005). Sudden expansion of the population was first estimated in a mismatch distribution graph: unimodal or not unimodal. Deviation from the demographic expansion model was estimated using the sum of the squared deviation and Harpending’s raggedness index (Harpending, 1994). To estimate population equilibrium, Fu’s FS test (Fu, 1997) and Tajima’s D test (Tajima, 1989) were conducted. If population expansion was detected by the mismatch
| Sn | CL | GPS | Cln | Sn | SSn | COI GAn | ATP6/ATP8 GAn | GSn |
|----|----|-----|-----|----|-----|---------|---------------|-----|
| A  | Korea, GG, Suwon-si, Yeongtong-gu, Mangpo-dong | 37°14.092′N 127°02.840′E | Egg mass | For NGS | – | – | – | – |
| 1  | Incheon, Gyeyang-gu, Gyesan-dong | 37°32′57.9″N 126°15′40.1″E | 30 | 192–221 | 30 | KT245570–KT245599 | XX945522–XX945551 | 20 |
| 6  | GW, Inje-gun, Buk-myeon, Hange-ri | 38°08′09.5″N 127°42′22.7″E | 30 | 312–341 | 28 | KT245288–KT245317 | XX945552–XX945579 | 20 |
| 10 | CN, Seosan-si, Haemi-myeon, Daegok-ri | 36°41′46.2″N 127°36′27.3″E | 30 | 492–521 | 23 | KT245496–KT245540 | XX945560–XX945562 | 20 |
| 12 | CB, Cheongweon-gun, Misweon-myeon, Daesin-ri | 36°50′23.4″N 129°16′22.5″E | 30 | 612–641 | 20 | KT245496–KT245540 | XX945560–XX945562 | 20 |
| 16 | GB, Yeongyang-gun, Subi-myeon, Suha-ri | 36°50′23.4″N 129°16′22.5″E | 30 | 612–641 | 17 | KT245558–KT245584 | XX945629–XX945645 | 20 |
| 18 | JB, Jinan-gun, Jinan-eup, Danyang-ri | 35°45′55.8″N 127°25′00.6″E | 30 | 672–701 | 28 | KT245609–KT245636 | XX945646–XX945673 | 20 |
| 22 | JN, Gangjin-gun, Jakcheon-myeon, Galdong-ri | 34°43′00.3″N 126°43′49.5″E | 30 | 792–821 | 29 | KT245722–KT245750 | XX945674–XX945679 | 20 |
| 26 | GB, Gyeongju-si, Yonggang-dong | 35°51′45.4″N 129°14′14.7″E | 30 | 912–941 | 20 | KT245840–KT245869 | XX945703–XX945732 | 20 |
| 27 | GN, Hapcheon-gun, Daebyeong-myeon, Hageum-ri | 35°31′27.9″N 127°59′12.1″E | 30 | 942–971 | 22 | KT245870–KT245899 | XX945733–XX945754 | 20 |
| 28 | GN, Milyang-si, Bubuk-myeon, Jeonsapo-ri | 35°27′30.5″N 128°44′11.6″E | 30 | 972–1,001 | 28 | KT245900–KT245929 | XX945755–XX945782 | 20 |
| 30 | GN, Gimhae-si, Saman-dong | 35°15′16.0″N 128°54′51.3″E | 30 | 1,032–1,061 | 30 | KT245960–KT245989 | XX945783–XX945812 | 20 |
| 31 | JJ, Jeju-si, Bonggae-dong | 33°26′15.0″N 126°37′43.8″E | 30 | 1,062–1,091 | 28 | KT245990–KT245990 | XX945813–XX945842 | 20 |
| 33 | Russia, Vladivostok | 43°23′44.6″N 132°09′56.6″E | 30 | 1,703–1,732 | 28 | KT246046–KT246075 | XX945843–XX945870 | 20 |
| 34 | Mongolia Selenge Province Shaganuur | 50°15′N 105°30′E | 30 | 1,733–1,749 | 13 | XX945391–XX945403 | XX945871–XX945883 | 17 |
| 35 | Japan Hokkaido Otaru Asarigawa-onsen, 1 Chome | 43°8.056′N 141°2.395′E | 30 | 1,870–1,899 | 23 | XX945404–XX945426 | XX945884–XX945906 | 30 |
| 36 | Japan Hokkaido Sapporo Minami-ku Jozankei-jojo-cho Fushimi | 42°48.098′N 140°58.172′E | 30 | 1,900–1,929 | 23 | XX945427–XX945449 | XX945907–XX945929 | 30 |
| 37 | Japan Hokkaido Abuta Kimshitsu-cho Fushimi | 42°57.666′N 141°9.423′E | 30 | 1,930–1,959 | 23 | XX945450–XX945472 | XX945930–XX945952 | 30 |
| 38 | Japan Hokkaido Sapporo Minami-ku Jozankei-onsen Higashi 4 Chome | 42°57.666′N 141°9.431′E | 30 | 1,960–1,989 | 26 | XX945473–XX945498 | XX945953–XX945978 | 30 |
| 39 | China Jilin Helong Qingshanli | 42°26′22.6″N 128°51′50.3″E | 18 | 1,990–2,007 | 17 | XX945499–XX945515 | XX945979–XX945995 | 18 |
| 41 | Japan Kyushu Fukuoka Miyawaka Mt. Inunaki | 33°40′54.0″N 130°33′15.4″E | 7 | 2,013–2,019 | 6 | XX945516–XX945521 | XX945960–XX946001 | 7 |
| Total | 20 sites | – | 565 | – | 480 | – | – | 432 |

Sn, site number; CL, collecting location; Cln, number of collected individuals; Sn, sample number; SSn, number of sequenced samples; GAn, GenBank accession number; GSn, number of genotyped samples.
distribution, the expansion time of the population was calculated using the formula \( \tau = 2ut \) (\( \tau \) = age expansion parameter; \( u = \) the aggregate mutation rate over the region of DNA under study; and \( t = \) generation time) (Roger & Harpending, 1992), with the assumption that the mutation rate of insect mitochondrial DNA is 2.3% per million years (Brower, 1994).

### RESULTS

#### 3.1 NGS sequencing and microsatellite marker development

Illumina sequencing of the genomic DNA from *L. dispar* eggs obtained 3,974,358,483 bp from 15,988,036 reads, with an average of 248 bp per read, which assembled into 718,940 contigs, with an average of 511 bp per contig (Table S1). The contigs contained 1,867 microsatellite loci (excluding AT repeats; the length of the repeated base, 2–4 bp; and repeated more than four times). Of these, 430 loci showing more than six repeated motifs were tested for the probability of marker design with PRIMER3 (Koressaar & Remm, 2007; Untergrasser et al., 2012). We were able to design primer sets for 207 loci, from which we randomly selected 150 loci for PCR tests to examine polymorphism.

Capillary electrophoresis revealed that 29 of 150 loci showed clear polymorphism (Table 2). From PCR amplification tests on the labeled markers of selected 29 loci, we selected eight microsatellite loci (39,767, 58,587, 124,259, 134,079, 230,995, 297,455, 344,041, and 346,977) (Table 3). The sequences of the eight selected microsatellite loci were submitted to NCBI GenBank (Table 3).

### 3.2 Microsatellite marker assessment

Microsatellite markers were initially assessed using 20 regional populations (one population from Mongolia, one from China, one from

---

**TABLE 2** Selected primer sets (29 of 150) showing clear polymorphism

| No | MSL no | SPS | RM |
|----|--------|-----|----|
| 1  | 19,028 | GCGTACAAACTACGCAAGTC | (CT)\(_{14}\) |
| 2  | 20,500 | CCCCTAGTCTCCGTGTTAACC | (ATG)\(_{9}\) |
| 3  | 22,651 | GTCGGACCCGGAGCATAAC | (ATC)\(_{10}\) |
| 4  | 39,767 | AGCGGTCTCTAATTGGTTAT | (GT)\(_{15}\) |
| 5  | 44,678 | GAGTGAAGTGTGATGGGTGAG | (ATG)\(_{9}\) |
| 6  | 58,587 | TGAAACGAGCTACTAACACGAG | (ATG)\(_{8}\) |
| 7  | 109,715 | GGGTTTCCCTGACTTTTTGACAAATA | (AC)\(_{13}\) |
| 8  | 119,274 | GCGACCCTCTCTAAGGCGGTCAG | (ATG)\(_{9}\) |
| 9  | 124,259 | TTGAACTCTGCGTACATTAGGCAAC | (AG)\(_{13}\) |
| 10 | 134,079 | TGGCACTACGAAAGGACATGAAACCGGC | (ATG)\(_{9}\) |
| 11 | 167,938 | GAAAATCGCAGCTTTGGAATGGTTTTCAATC | (AG)\(_{15}\) |
| 12 | 178,435 | CTTGCCCGCTGATATCCGAA | (GT)\(_{13}\) |
| 13 | 178,855 | AATGCAGAGGCGAAGGGAATGG | (GT)\(_{16}\) |
| 14 | 203,511 | GACTTTAGCGGTGCTAGGTCAGT | (ACAT)\(_{7}\) |
| 15 | 205,435 | GGTGGGTTGTTTAGTACGATA | (CT)\(_{13}\) |
| 16 | 205,922 | CCGCTGATGGGACAGAGGCC | (CT)\(_{13}\) |
| 17 | 230,995 | CCACTGTGACCATGTGCTATG | (ATC)\(_{10}\) |
| 18 | 233,404 | GCCCGACGCTTAGTTTGAAGAT | (AC)\(_{16}\) |
| 19 | 239,543 | TTTTGGGGCAACATGAGTAA | (ATG)\(_{9}\) |
| 20 | 243,906 | CGGACCAAGAATATCGGAAACGGTTAC | (ATG)\(_{9}\) |
| 21 | 253,129 | GAGTCACGCTACAGGCAAGTTCGA | (ATC)\(_{10}\) |
| 22 | 297,455 | GTGGCGCTTCCTAATTGGTTAT | (CT)\(_{23}\) |

MSL, microsatellite loci; SPS, sequence of primer set; RM, repeat motif.

---

(Continues)
Two populations, Sites 12 and 26, had observed heterozygosity values lower than the gene diversity values in all loci except locus 344,041. Lower observed heterozygosity values than the gene diversity values suggest significant homozygosity, and this implies the presence of null alleles or allelic dropout, linkage of alleles, or inbreeding (Damm, Armstrong, Arjo, & Piaggio, 2015). However, we tested for the presence of null alleles or allelic dropout and linkage of alleles through the previous analyses. Lastly, if the violation were a consequence of inbreeding, we would have expected to observe such a phenotype at many or all loci, not just at a single locus (Damm et al., 2015; Selkoe & Toonen, 2006). The samples from Sites 12 and 26 might indicate inbreeding or sib sampling. In this study, however, we retained the samples from the two sites in our analysis because the deviation from HWE was not detected at all analyzed loci (Table S3). Therefore, we suggest that the developed eight novel microsatellite markers may be useful for a population genetic analysis of *L. dispar*.

### 3.3 Population structure using microsatellite loci

#### 3.3.1 Pairwise $F_{ST}$ genetic distances

The population genetic structure of *L. dispar* in Far East Asia was calculated with $F_{ST}$ values. Pairwise $F_{ST}$ distances among regional populations ranged from −0.0087 to 0.1171 (Table 5: Lower side). Considering the genetic distances in each geographical region, the regional populations in Hokkaido (Sites 35–38), the species origin region, showed relatively low genetic distances (from −0.0055 to −0.0010); however, compared to other regional populations, their genetic distance was relatively high (0.0472 to 0.1171). A Mongolian regional population (Site 34), which was further from Hokkaido than other regional populations, showed relatively high genetic distances.

**TABLE 3** Ten selected markers for microsatellite loci analysis of *Lymantria dispar*

| MSL no | Marker name | Sequence | RM | Size | GAn  |
|--------|-------------|----------|----|------|------|
| 39,767 | 39767-FAM   | AGCGCTTCTAATTGGTTAT | (GT)$_{15}$ | 129–179 | KT633401 |
|        | 39767R      | ACGGCTGTTATAATTTCCA |      |      |      |
| 58,587 | 58587-FAM   | TCAGTCGAATTTAGGCAAA | (ATG)$_{8}$ | 214–299 | KT633402 |
|        | 58587R      | TCGAAACAAATCTCGGAT |      |      |      |
| 124,259| 124259-FAM  | TTGACACTGACCCGAAAATT | (AG)$_{13}$ | 184–218 | KT633403 |
|        | 124259R     | ATTTGCACATATGACCCAC |      |      |      |
| 134,079| 134079-FAM  | TGAAGAGCAGCTAAACGCG | (ATC)$_{9}$ | 159–270 | KT633404 |
|        | 134079R     | GACTTGTACAGGAATTGGTT |      |      |      |
| 230,995| 230995-FAM  | CCAAGCAATTTGACCTAT | (ATC)$_{10}$ | 148–196 | KT633405 |
|        | 230995R     | TAAAGGCACTTGCTTGTAT |      |      |      |
| 297,455| 297455-FAM  | GTTGGCTTGCTTGATGTTATG | (CT)$_{23}$ | 170–254 | KT633407 |
|        | 297455R     | GTGACTTCGCTGTAACACTC |      |      |      |
| 344,041| 344041-FAM  | GTGCCACAGTGAAATAATAC | (ATC)$_{9}$ | 131–239 | KT633409 |
|        | 344041R     | CTTGCTTGCTTGGTGCTCAT |      |      |      |
| 346,977| 346977-FAM  | CTTGCTTGACTTATCTGTGG | (AGT)$_{8}$ | 165–201 | KT633410 |
|        | 346977R     | AGCTTTTTCTAGGGTGATG |      |      |      |

MSL, microsatellite loci; RM, repeat motif; GAn, GenBank accession number.
**TABLE 4** Results of the test of null alleles and the PCR error present in eight filtered markers

| MSL     | Site 1 | Site 6 | Site 10 | Site 12 |
|---------|--------|--------|---------|---------|
|         | NP     | Freq   | NP      | Freq    | NP      | Freq    | NP      | Freq    |
| 39,767  | No     | 0.0164 | No      | −0.0053 | No      | −0.1204 | No      | 0.0735  |
| 58,587  | No     | −0.1305| No      | −0.5138 | No      | −0.4287 | No      | 0.0323  |
| 124,259 | No     | −0.0127| No      | 0.0763  | No      | −0.0307 | No      | 0.0725  |
| 134,079 | No     | −0.0313| No      | −0.1476 | No      | −0.0360 | No      | 0.0456  |
| 230,995 | No     | −0.0139| No      | 0.0417  | No      | −0.0508 | No      | −0.0014 |
| 297,455 | No     | −0.0548| No      | −0.0010 | No      | −0.0810 | No      | 0.0379  |
| 344,041 | No     | −0.1074| No      | −0.0994 | No      | −0.1225 | No      | −0.1475 |
| 346,977 | No     | −0.0233| No      | 0.0644  | No      | −0.0468 | No      | 0.0002  |

| MSL     | Site 16 | Site 18 | Site 22 | Site 26 |
|---------|---------|---------|---------|---------|
|         | NP     | Freq   | NP      | Freq    | NP      | Freq    | NP      | Freq    |
| 39,767  | No     | 0.0704 | No      | 0.0236  | Yes     | 0.1386  | No      | 0.0001  |
| 58,587  | No     | −0.2454| No      | −0.4557 | No      | 0.0147  | No      | 0.0096  |
| 124,259 | No     | 0.0164 | No      | −0.0084 | No      | −0.0284 | No      | 0.1223  |
| 134,079 | No     | −0.0530| No      | 0.0206  | No      | 0.0575  | Yes     | 0.1540  |
| 230,995 | No     | 0.0303 | No      | −0.0086 | No      | −0.0073 | No      | 0.0995  |
| 297,455 | No     | 0.0070 | No      | −0.0123 | No      | 0.0362  | No      | 0.0708  |
| 344,041 | No     | −0.1725| No      | −0.1539 | No      | −0.1185 | No      | −0.1949 |
| 346,977 | No     | −0.0357| No      | 0.0835  | Yes     | 0.1947  | No      | 0.1397  |

| MSL     | Site 27 | Site 28 | Site 30 | Site 31 |
|---------|---------|---------|---------|---------|
|         | NP     | Freq   | NP      | Freq    | NP      | Freq    | NP      | Freq    |
| 39,767  | No     | 0.0022 | No      | −0.1312 | No      | 0.0307  | No      | 0.1159  |
| 58,587  | No     | 0.0443 | No      | 0.1366  | No      | 0.1281  | No      | 0.0343  |
| 124,259 | No     | −0.0552| No      | 0.0166  | No      | −0.0625 | No      | −0.1003 |
| 134,079 | No     | −0.1057| No      | −0.0089 | No      | −0.0659 | No      | −0.0150 |
| 230,995 | No     | −0.0280| Yes     | 0.1603  | No      | 0.0402  | No      | −0.0294 |
| 297,455 | No     | 0.0071 | No      | 0.0192  | No      | 0.0271  | No      | −0.0036 |
| 344,041 | No     | −0.5174| No      | −0.2371 | No      | −0.1892 | No      | −0.1916 |
| 346,977 | No     | 0.0421 | No      | 0.0802  | No      | 0.0023  | No      | 0.1520  |

| MSL     | Site 33 | Site 34 | Site 35 | Site 36 |
|---------|---------|---------|---------|---------|
|         | NP     | Freq   | NP      | Freq    | NP      | Freq    | NP      | Freq    |
| 39,767  | No     | 0.0554 | No      | −0.1725 | No      | −0.0752 | No      | −0.0275 |
| 58,587  | No     | −0.0017| No      | −0.0801 | No      | 0.0340  | No      | 0.0901  |
| 124,259 | No     | 0.0812 | No      | 0.0861  | No      | 0.0094  | No      | −0.1114 |
| 134,079 | No     | −0.0264| No      | −0.1606 | No      | 0.0087  | No      | 0.0502  |
| 230,995 | Yes    | 0.0980 | No      | −0.6078 | No      | −0.0350 | No      | −0.0157 |
| 297,455 | No     | 0.0196 | No      | −0.0727 | No      | −0.2100 | No      | −0.0614 |
| 344,041 | No     | −0.0685| No      | −0.1697 | No      | −0.2168 | No      | −0.3067 |
| 346,977 | No     | −0.0182| No      | 0.0000  | No      | 0.0964  | No      | 0.0824  |

| MSL     | Site 37 | Site 38 | Site 39 | Site 41 |
|---------|---------|---------|---------|---------|
|         | NP     | Freq   | NP      | Freq    | NP      | Freq    | NP      | Freq    |
| 39,767  | No     | 0.0402 | No      | −0.0001 | No      | 0.0300  | No      | 0.0332  |
| 58,587  | No     | −0.0135| No      | −0.0135 | No      | −0.1435 | No      | −0.0089 |

(Continues)
of 0.0464 to 0.1171 from other regional populations. A Vladivostok population (Site 33) was similar to a Korean inland population (Site 6) in population genetic structure \( (F_{ST} = -0.00006) \), yet showed a high genetic distance of 0.1030 when compared to a Jeju regional population (Site 31). A Chinese regional population (Site 39) was more similar to a Korean inland population (Site 28) with a relatively large geographical distance (genetic distance: 0.0050 and geographic distance: 776.35 km) than to a Russian population (Site 33) with a relatively short distance (genetic distance: 0.0199 and geographic distance: 289.10 km). A Kyushu regional population (Site 41) with a small sample size had a genetic distance of 0.0307 to 0.0964 from other populations. Lastly, the Korean inland populations (Sites 1–30) ranged from −0.0087 to 0.0358 and were similar in genetic structure. In the analysis of isolation by distance, genetic distance increased with increasing geographic distance (Figure 3a, \( r = 0.7909, p = .0000 \)).

### 3.3.2 NeighborNet network

In a NeighborNet network based on pairwise \( F_{ST} \) genetic distances, the 20 regional populations could be divided into five groups: Group 1, Hokkaido (Sites 35, 36, 37, and 38); Group 2, Kyushu (Site 41); Group 3, Jeju Island (Site 31); Group 4, Korean Peninsula and adjacent areas (Sites 1, 6, 10, 12, 16, 18, 22, 26, 27, 28, 30, 33, and 39); and Group 5, Mongolia Selenge (Site 34) (Figure 4). Among these sites, Sites 28 and 30 from inland Korea were closest to each other in geographic

### Table 5 Pairwise \( F_{ST} \) distances among regional populations of *Lymantria dispar* in Far East Asia

| RP | Site 1 | Site 6 | Site 10 | Site 12 | Site 16 | Site 18 | Site 22 | Site 26 | Site 27 |
|----|-------|-------|--------|--------|--------|--------|--------|--------|--------|
| Site 1 | −      | −0.00406 | −0.01463 | 0.00437 | −0.01334 | 0.30988 | 0.26863 | 0.34350 | 0.36249 |
| Site 2 | 0.0082 | −      | 0.02061 | 0.04141 | 0.01087 | 0.31302 | 0.27210 | 0.34671 | 0.36975 |
| Site 10 | 0.0041 | 0.0046 | −      | −0.00547 | −0.00048 | 0.27709 | 0.23826 | 0.31041 | 0.32849 |
| Site 12 | 0.0095 | −0.0087 | 0.0047 | −      | −0.01072 | 0.19614 | 0.16332 | 0.23044 | 0.24587 |
| Site 16 | 0.0215 | 0.0085 | 0.0249 | 0.0023 | −      | 0.24157 | 0.20618 | 0.27773 | 0.29740 |
| Site 18 | 0.0154 | 0.0044 | 0.0074 | 0.0048 | 0.0179 | −      | −0.02895 | −0.02244 | 0.00001 |
| Site 22 | 0.0211 | 0.0223 | 0.0096 | 0.0350 | 0.0427 | 0.0082 | −      | −0.01622 | 0.00406 |
| Site 26 | 0.0067 | 0.0132 | 0.0076 | 0.0172 | 0.0302 | 0.0055 | 0.0002 | −      | −0.01407 |
| Site 27 | 0.0195 | 0.0167 | 0.0135 | 0.0278 | 0.0358 | 0.0116 | 0.0082 | 0.00041 | −      |
| Site 28 | 0.0024 | −0.0061 | 0.0083 | −0.0011 | 0.0167 | 0.0079 | 0.0233 | 0.0031 | 0.0136 |
| Site 30 | 0.0153 | −0.0029 | 0.0146 | −0.0006 | −0.0006 | 0.0004 | 0.0158 | 0.0130 | 0.0180 |
| Site 31 | 0.0418 | 0.0463 | 0.0502 | 0.0470 | 0.0582 | 0.0280 | 0.0347 | 0.0359 | 0.0451 |
| Site 33 | 0.0050 | −0.0003 | 0.0069 | 0.0043 | 0.0248 | 0.0215 | 0.0383 | 0.0225 | 0.0346 |
| Site 34 | 0.0561 | 0.0607 | 0.0590 | 0.0794 | 0.0960 | 0.0709 | 0.0569 | 0.0651 | 0.0664 |
| Site 35 | 0.0585 | 0.0738 | 0.0698 | 0.0616 | 0.0619 | 0.0554 | 0.0639 | 0.0582 | 0.0626 |
| Site 36 | 0.0558 | 0.0659 | 0.0642 | 0.0594 | 0.0572 | 0.0472 | 0.0588 | 0.0568 | 0.0572 |
| Site 37 | 0.0510 | 0.0636 | 0.0643 | 0.0557 | 0.0560 | 0.0526 | 0.0617 | 0.0509 | 0.0568 |
| Site 38 | 0.0655 | 0.0741 | 0.0779 | 0.0739 | 0.0662 | 0.0618 | 0.0654 | 0.0610 | 0.0627 |
| Site 39 | 0.0137 | 0.0178 | 0.0167 | 0.0237 | 0.0380 | 0.0302 | 0.0353 | 0.0345 | 0.0282 |
| Site 41 | 0.0516 | 0.0448 | 0.0453 | 0.0380 | 0.0467 | 0.0398 | 0.0307 | 0.0359 | 0.0442 |

RP, regional population; lower side, microsatellite loci; upper side, mitochondrial genes.
distance; however, their genetic distance was similar to Site 1 (geographically close to Incheon Harbor) and Site 16 (geographically close to Uljin Harbor). These two regions are geographically close to Busan Harbor, which is a frequent entry port for vessels (Choi, 2014). We therefore suspect that these two regional populations may frequently interbreed with the regional populations near Incheon Harbor and Uljin Harbor.

### 3.3.3 Bayesian clustering

For the model-based Bayesian analysis, K was estimated by varying it from two to eight, and the ad hoc statistics ΔK (Evanno et al., 2005) indicate the maximum level of structure in three genetic groups (Figure 5). *Lymantria dispar* has been divided into two subspecies in Asia, *L. dispar asiatica* (or *L. dispar dispar*) and *L. dispar japonica*, based on mitochondrial DNA and microsatellite analysis (Bogdanowicz et al., 2000; Wu et al., 2015). Our study showed similar results; however, the Far East Asian gypsy moth populations were distinguishable as three types according to sampling region (Figure 6). Comparing the individual colored bar plots among the regional populations revealed that the frequency of the green-colored genetic content was high in Hokkaido regional populations (the species origin region) (Figures 6a, r, s, and t), the frequency of the red genetic content was high in Jeju regional populations (Figure 6o), and the frequency of the blue genetic content was high in Mongolian regional populations (Figure 6a). The regional populations from the Korean Peninsula and adjacent areas showed a mixed pattern in comparison with the Jeju regional populations and Mongolian regional populations.

Comparing the individual colored bar plots of each regional population, Sites 35, 36, 37, and 38 from Hokkaido were clearly distinct in genetic makeup from the regional populations of the Korean Peninsula and adjacent areas (Figure 6). Several individuals (Figure 6: individual 346; Figure 6o: individual 375; and Figure 6t: individuals 406, 407, and 410) showed a genetic makeup similar to that of other regional populations; however, in the majority of individuals, the main genetic makeup was the green-colored one. A genetic content frequency similar to that of the Hokkaido regional populations could be seen in Site 41 (Kyushu population), Site 27 (Hapcheon population), Site 12 (Cheongwon population), and Site 39 (Jilin population). Among them, the Kyushu regional population, with only seven individuals analyzed, was divided into two types: three individuals showed features similar to the Hokkaido regional populations, and four individuals showed features similar to the Jeju regional populations (Figure 6p).

The individual colored bar plots of the Korean inland populations show high frequencies of the blue or red genetic content in each individual. These two genetic content types showed similar frequencies in several individuals. This result may be caused by the higher genetic diversity in these populations than in other regions, and the gene flow among the Korean inland regions may be relatively higher than with other regions (Table 6, $F_{ST} = 0.04192$). In the Chinese and Russian regional populations, however, the blue genetic content was higher than other genetic content types. Several individuals (Figure 6i: 0.01105; 0.03056; 0.002079; 0.02576; 0.034123; 0.00213; 0.001105; 0.002450; 0.000950; 0.00004; 0.000542; 0.00004; 0.000676; 0.000676; 0.000591; 0.000575; 0.000697; 0.000505; 0.000555) had higher blue genetic content frequencies than red genetic content frequencies.
3.4 Population structure using mitochondrial DNA

3.4.1 Mitochondrial DNA sequence variation

DNA barcodes of the COI, ATP6, and ATP8 genes were sequenced from 480 of 552 *L. dispar asiatica* samples collected from individual 173; Figure 6m: individuals 248 and 262; and Figure 6c: individuals 53 and 56) had features similar to those of the Hokkaido regional populations.

FIGURE 3 Isolation by distance for matrix correlation between genetic distance and geographic distance (a, microsatellite loci, \( r = 0.7909, p = .0000 \); b, mitochondrial genes, \( r = 0.5312, p = .0006 \))

FIGURE 4 NeighborNet network using pairwise \( F_{ST} \) distances from 20 regional populations of *Lymantria dispar* from Far East Asia

FIGURE 5 The ad hoc statistics \( \Delta(K) \) on the basis of LnP(D) estimated from 20 iterations for each \( K \). The ad hoc statistics exhibited a signal of at best \( K = 3 \)
the 20 study sites (n = 6–30 per site). Mitochondrial DNA sequence divergences obtained from the 480 samples ranged from null to 0.5%, with 98 haplotypes distinguished by 85 polymorphic sites (Table S4). The mean gene diversity was 0.6529 ± 0.0929 (lowest value 0.1538 ± 0.1261 from Site 34 and highest value 0.9407 ± 0.0432 from Site 12), and the mean nucleotide diversity was 0.013798 ± 0.010223 (lowest value 0.001810 ± 0.003024 from Site 34 and highest value 0.027156 ± 0.017203 from Site 12) (Table S5).

3.4.2 | Mitochondrial genealogy

In the median-joining network, three high-frequency haplotypes (H1, 151ex; H37, 75ex; and H90, 73ex) were connected to each other by low-frequency haplotypes (Figure 7). This pattern was revealed in the pairwise $F_{ST}$ distances (Table 5: Upper side). We found that the 20 studied populations of *L. dispar* were divided into three groups according to genetic distance: Group 1, Korean inland region and adjacent areas (Sites 01, 06, 10, 12, 16, 31, 33, 34, 39, and 41);
Group 2, Korean southern region (Sites 18, 22, 26, 27, 28, and 30); and Group 3, Hokkaido region (Sites 35, 36, 37, and 38) (Table S5, $F_{CT} = 0.48280$, $F_{ST} = 0.48753$). The results of the analysis of IBD were similar to the microsatellite results (Figure 3b, $r = 0.5312$, $p = .0006$).

In particular, haplotype H90 appeared only in Hokkaido regional populations and was connected with haplotype H1 by haplotype H95, which is another Hokkaido haplotype (Figure 7). The Kyushu regional population (Site 41) contained five haplotypes, of which three haplotypes showed in the inland (H1, 2ex; H27, 1ex; and H82, 1ex), one in Hokkaido (H93, 1ex), and one in only Kyushu (H98, 1ex). One of the high-frequency haplotypes, H1, was distributed in all the inland collecting regions and was detected in an individual from Site 38 in Hokkaido (Figure 7). Haplotype H1 was connected with Haplotype H37, a high-frequency haplotype in the southern area of the Korean Peninsula, by low-frequency haplotypes. Overall, the Far East Asian gypsy moth populations showed a star-shaped network in which three high-frequency haplotypes (H1, 151ex; H37, 75ex; and H90, 73ex) were connected with each other through low-frequency haplotypes (Figure 7). Therefore, the Far East Asian gypsy moth populations may have undergone sudden population expansion.

### 3.4.3 Mitochondrial DNA haplotype mismatch distribution

The median-joining network revealed a star-shaped mtDNA genealogy, so we analyzed the mismatch distribution, applying a sudden population expansion model. We conducted the analysis using the three groups recognized above, and we found that the mismatch
gaphs of the groups were unimodal and the mismatch parameters were insignificant (Figure 8). In neutral equilibrium, Tajima’s D and Fu’s FS tests also had negative values in all three groups (Figure 8). We therefore consider that the mismatch analysis supports a sudden population expansion.

The expansion time of each group was inferred using the observed age expansion parameter (tau), the equation $t = \frac{\tau}{2\mu}$ (Roger & Harpending, 1992), and an insect mtDNA mutation rate of 2.3% per MY per lineage for silent sites (Brower, 1994). The tau of each group was 1.234 in Group 1, 1.496 in Group 2, and 1.750 in Group 3, and the expansion times were estimated to be 53,652 generations ago in Group 1, 65,043 in Group 2, and 76,086 in Group 3. Considering that $L. \ dispar$ produces one generation per year (Pogue & Schaefer, 2007), the population expansion time of each group in Far East Asia was inferred to be approximately 53,652 years before present (ybp) in Group 1, 65,043 ybp in Group 2, and 76,086 ybp in Group 3.

4 | DISCUSSION

The taxonomic status of the two subspecies of $L. \ dispers$ in Far East Asia has been debated (Arimoto & Iwaizumi, 2014; Pogue & Schaefer, 2007; Schintlmeyer, 2004). In a recent study using molecular data (Wu et al., 2015), $L. \ dispers$ (European subspecies) was clearly distinct from the Asian two subspecies; however, the Asian subspecies were difficult to distinguish from each other. The Japanese subspecies, $L. \ dispers$ japonica, was genetically similar to the populations from the southern end of the Korean Peninsula, and the Korean populations had mixed genetic content (Wu et al., 2015). We examined the previous study’s collecting sites and found they were mainly located near seaports. We therefore included inland populations in the present study (Figure 2).

In mitochondrial genealogy, we found that three lineages of $L. \ dispers$ were distributed in Far East Asia: two in the Korean Peninsula and adjacent inland areas, and one in Hokkaido, Japan. Inferring the demographic history of each lineage through mismatch analysis, Group 1 expanded suddenly approximately 53,652 ybp, Group 2 approximately 65,043 ybp, and Group 3 approximately 76,086 ybp, all within the Würm glacial period (110,000–12,000 ybp) (Gao, Hou, & Guo, 2016).

The Würm glacial period can be divided into three glacial stages and two subinterglacial stages (Gao et al., 2016; Han & Meng, 1996; Ma, Yu, Wang, & Yao, 2006). The mean temperature during the period was approximately 5°C lower than at present, based on the snow-line elevation on Mt. Fuji in Japan (Kim, 2011). In Europe, three ice sheets (Scandinavian, British, and Alpine) developed to cover a large part of the continent (Trojan, 1997). The advancing glacier forced the flora and fauna of the warm and temperate zones southward, and refugia were formed in the Mediterranean region (Trojan, 1997). In a great amount of Siberia, large ice masses eliminated all plants and animals; however, eastern regions (including Ussuri Land, Korea, Manchuria, and Japan) remained ice-free as fauna- and flora-preserving areas during the glaciation period (Trojan, 1997). During these periods, the flora of the southern part of Korea showed the features of a cool, temperate climate (Chung, Lee, Lim, & Kim, 2005). For example, Polygodiaceae, $Alnus$ spp., $Carpinus$ spp., and deciduous $Quercus$ spp. were distributed in the area (Chung et al., 2005; Kim, 2011). $Carpinus$ spp. and $Quercus$ spp. are the food plants of $L. \ dispers$ asiatica in Korea (Lee et al., 2002). During the last glacial maximum (approximately 20,000–18,000 ybp), however, $Picea$ spp., $Abies$ spp., $Pinus$ spp., and $Larix$ spp. were distributed in Far East Asia as it changed to a subarctic climate (Kim, 2011; Yoon & Hwang, 2009). The coastline during this period was quite different from the present. The west sea of the Korean Peninsula was a low hilly area because the sea level was approximately 30–130 m below present levels (Kim, 2011; Park & Cho, 1998; Park, Yoo, Lee, & Lee, 2000). The Japanese islands were connected with Sakhalin and the southeast part of the Korean Peninsula by a land bridge (Park et al., 2000; Trojan, 1997).

The sudden expansion of the Japanese Hokkaido lineage (Group 3) may have taken place in the middle of the Würm glacial stage I (approximately 76,086 ybp), a period with a cold and dry climate that might have led them to move to more southern regions. The southeastern lineage of the Korean Peninsula (Group 2) might have expanded...
During the late Würm glacial stage I (approximately 65,043 ybp). During this period, the gypsy moth populations might have dispersed into the southern part of the Korean Peninsula because of the cold climate. Lastly, Group 1 might have dispersed into the Vladivostok area, the middle region of Korea, and even Mongolia because the estimated expansion period is approximately 53,652 ybp, which is known as the subinterglacial stage I (60,000–50,000 ybp), a period with a hot and wet climate. We can thus infer that the gypsy moth populations dispersed from Far East Asia into middle Asia. In the Korean Peninsula, however, they may not have dispersed southward because the Noryeong and Charyeong Mountains were formed in the Miocene (Park & Son, 2008), and therefore, genetic interaction between Group 1 and Group 2 would not have been possible. In the model-based Bayesian analysis using microsatellite loci, K (assumed as the number of populations) was calculated to be three, the same as the number of groups examined in the mitochondrial genealogy. The genetic diversity of the regional populations was higher in the Korean Peninsula than in other regions, with the Korean Peninsula populations showing the same mixed pattern reported previously (Wu et al., 2015). We suggest that this genetic pattern might have been caused by multiple sudden population expansions, and the demographic patterns caused by the Würm glacial period may have resulted in the present genetic diversity. In genetic makeup, however, the regional populations near the Busan seaport (Sites 27 and 28) were similar to the middle area of the Korean Peninsula. This might have been caused by vessels arriving in Korea and anchoring at the ports of Incheon or Busan (Choi, 2014; Kim, Kim, Kim, & Lee, 2008). We also looked for this genetic pattern in several samples from Russian and Japanese populations (Figures 6b, p, q, r, s, and t). Thus, we suggest that several individuals might have been introduced into each region via vessels arriving at seaports.

We can suggest that L. dispar in Far East Asia are divided into two types (the inland type and the Hokkaido type), although the analyzed samples did not cover the full distributional region of the species in Far East Asia. Taxonomically, 15 nomino-subspecies have been assigned to L. dispar: L. dispar linnaeus, 1758 (type locality [TL]: Europe); L. dispar erebos Thierry Mieg, 1886 (TL: England, Us proviennt de Darlington); L. dispar asiatica Vnukovskij, 1926 (TL: Russia, Siberia meridionales, Altaj et Sajan occidentales, Prov. Semipalatinsk); L. dispar praeterea Kardakoff, 1928 (TL: Russia, Ussuri-Gebiet, “Russ. Insel und in Narwa”); L. dispar hokkaidoensis Goldschmidt, 1940 (TL: Japan, Hokkaido); L. dispar koreibia Bryk, 1948 (TL: Korea, Motojondo); L. dispar koltchoff Bryk, 1948 (TL: China, Jiangsu [=Jiangsu]); L. dispar andalusica Reinig, 1938 (TL: Spain, Sierra de Alfacar); L. dispar mediterraneae Goldschmidt, 1940 (TL: Southern Europe); L. dispar bocharae Goldschmidt, 1940 (TL: Turkestan); L. dispar chosensis Goldschmidt, 1940 (TL: Korea); L. dispar japonica Motschulsky, [1861] (TL: Japan); L. dispar umbrosa Butler, 1881 (TL: Japan, Tokei, Yokohama, Hakodate); L. dispar hadina Butler, 1881 (TL: Honshu, Yokohama); L. dispar obscura Goldschmidt, 1940 (TL: Japan, Honshu); and L. dispar nesiobia Bryk, 1942 (TL: Japan, Kuril Island). Most of these nomino-subspecies were synonymized and merged into two subspecies, L. dispar dispar and L. dispar japonica, by Schintlmeister (2004) from morphological analysis and consideration of the type locality of each subspecies. Recently, Pogue and Schaefer (2007) reinstated L. dispar asiatica and suggested a three subspecies system (L. dispar dispar, L. dispar asiatica, and L. dispar japonica). We partly agree with Schintlmeister’s (2004) view that L. dispar asiatica is a synonym of L. dispar dispar because the type locality of L. dispar asiatica is close to Europe, which is the type locality of L. dispar dispar. However, we consider that thorough genetic analyses on regional populations have to be undertaken in other regions of Eurasia to characterize the lineages of gypsy moth across its native range. A taxonomic system of the subspecies of L. dispar could therefore be re-established if each regional lineage revealed by genetic analysis is analyzed and compared with the toptotypes collected from the type locality of each subspecies.

ACKNOWLEDGMENTS

We thank Dr. B. Bayartogtokh (Ecology and Environmental Science, National University of Mongolia) for help with the collection of the Mongolian samples, Dr. J. Tchistyakov (Institute of Biology and Soil Science, Russia) for providing the Russian samples, and Dr. S. W. Park (Research Institute of Forest Insects Diversity, Suwon, Korea) for collecting and providing the Japanese samples. The present study is supported by R&D project “Development of determination method of the species and the origin of Asian Gypsy Moth intercepted in the port area (project number: B-1541785-2013-15-02”).

DATA ACCESSIBILITY

DNA sequences of 10 selected microsatellite loci are available in GenBank (KT633401–KT633410); DNA sequences of COI gene are available in GenBank (KT245170–KT246075; KX945391–KX945521); DNA sequences of ATP6/ATP8 gene are available in GenBank (KX945522–KX946001).

AUTHOR CONTRIBUTIONS

Tae Hwa Kang participated in the correction of the main idea of the study, coordinated the experiment, collected samples, participated in the analysis on the genetic diversity, and drafted the manuscript; Sang Hoon Han participated in the correction of the main idea of the study, coordinated the experiment, and participated in the analysis on the genetic diversity; Heung Sik Lee participated in the design of the main idea the study, collected samples, and managed funding.

ORCID

Tae Hwa Kang http://orcid.org/0000-0001-7834-4143

REFERENCES

Arimoto, M., & Iwaizumi, R. (2014). Identification of Japanese Lymantria species (Lepidoptera: Lymantriidae) based on morphological characteristics of adults. Research Bulletin of the Plant Protection Service Japan, 50, 89–110.
Koshio, C., Tomishima, M., Shimizu, K., Kim, H., & Takenaka, O. (2002). Microsatellites in the gypsy moth, Lymantria dispar L. (Lepidoptera: Lymantriidae). Applied Entomology and Zoology, 37, 309–312.

Lee, K. S., Kang, T. H., Jeong, J. W., Ryu, D. P., & Lee, H. S. (2015). Taxonomic review of the genus Lymantria (Lepidoptera: Erebidae: Lymantriinae) in Korea. Entomological Research, 45, 225–234.

Lee, J. H., Lee, H. P., Schafer, P. W., Fuester, R. W., Park, J. D., Lee, B. Y., & Shin, C. H. (2002). Gypsy moth parasitoid complex at Mt. Halla National Park, Cheju Island, Korea. Entomological News, 113, 103–112.

Leese, F., Mayer, C., & Held, C. (2008). Isolation of microsatellites from unknown genomes using known genomes as enrichment templates. Linnology and Oceanography: Methods, 7, 412–426.

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25, 1451–1452.

Liebold, A., Mastro, V., & Schafer, P. W. (1989). Learning from the legacy of Léopold Trouvelot. Bulletin of the Entomological Society of America, 35, 20–22.

Linnaeus, C. (1758). Systema Naturae per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species, cum Characteribus, Differentiis, Synonymis, Locis. Tomis I (10th edn). Homiae: Laurentii Slavii.

Liu, K., & Muse, S. V. (2005). PowerMarker: Integrated analysis environment for genetic marker data. Bioinformatics, 21, 2128–2129.

López-Uribre, M. M., Santiago, C. K., Bogdanowicz, S. M., & Danforth, B. N. (2012). Discovery and characterization of microsatellites for the solitary bee Colletes inaequalis using Sanger and 454 pyrosequencing. Apidologie, 44, 163–172.

Lowe, S., Browne, M., Boudjellal, S., & De Poorter, M. (2000). 100 of the world’s worst invasive alien species: a selection from the Global Invasive Species Database. Invasive Species Specialist Group (ISSG), Species Survival Commission (SSC), World Conservation Union (IUCN), Gland, Switzerland.

Ma, L., Yu, H. J., Wang, S. K., & Yao, J. (2006). Late quaternary environmental evolution in the Bohai Sea and formation of quaternary subsurface brine. Coastal Engineering, 25, 1–6. [in Chinese].

Mayer, C., Leese, F., & Tollrian, R. (2010). Genome-wide analysis of tandem repeats in Daphnia pulex – a comparative approach. BMC Genomics, 11, 277.

Miegh, T. H. (1886). Lépidopéres Européens. Le Naturaliste. Paris, 8, 236–237.

Oosterhout, C. V., Hutchinson, W. F., Willis, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes, 4, 535–538.

Park, Y. A., & Cho, S. K. (1998). Marine geology. In The Geological Society of Korea (Ed.), Geology of Korea (pp. 621–696). Seoul: Sigma Press [in Korean].

Park, S. J., & Son, I. (2008). Discussions on distribution and genesis of mountain ranges in the Korean Peninsula (III): Proposing a new mountain range map. Journal of the Geological Society of Korea, 43, 276–295. [in Korean].

Park, S. C., Yoo, D. G., Lee, C. W., & Lee, E. I. (2000). Last glacial sea-level changes and paleogeography of the Korea (Tsushima) Strait. Geo-Marine Letters, 20, 64–71.

Perry, J. C., & Rowe, L. (2011). Rapid microsatellite development for water striders by next-generation sequencing. Journal of Heredity, 102, 125–129.

Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update of the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics, 52, 273–288.

Pogue, M. G., & Schafer, P. W. (2007). A review of selected species of Lymantria Hübner [1819] including three new species (Lepidoptera: Noctuidae: Lymantriinae). Washington, DC: US Department of Agriculture Forest Health Technology Enterprise Team.

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

Qian, L., An, Y., Song, J., Xu, J., Ye, J., Wu, C., ... Hao, D. (2014). COI gene geographic variation of gypsy moth (Lepidoptera: Lymantridae) and a TaqMan PCR diagnostic assay. DNA Barcodes, 2, 10–16.

Reinig, W. F. (1938). Elimination und selektion. Jena: Gustav Fischer.

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

Richardson, M. F., Stanley, A. M., & Sherman, C. D. H. (2012). Development of novel microsatellite markers for the invasive Northern Pacific seastar, Asterias amurensis. Conservation Genetic Resources, 4, 327–330.

Roger, A. R., & Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. Molecular Biology and Evolution, 9, 552–569.

Sakai, A. K., Allendorf, F. W., Holt, J. S., Lodge, D. M., Molofsky, J., With, K. A., ... Weller, S. G. (2001). The population biology of invasive species. Annual Review of Ecology and Systematics, 32, 305–332.

Santana, Q. C., Coetzee, M. P. A., Steenkamp, E. T., Monyeni, O. X., Hammond, G. N. A., Wingfield, M. J., & Wingfield, B. D. (2009). Microsatellite discovery by deep sequencing of enriched genomic libraries. BioTechniques, 46, 217–223.

Schintzius, A. (2004). The taxonomy of the genus Lymantra Hübner, (1819) (Lepidoptera: Lymantriidae). Quadridina, 7, 1–248.

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology, 18, 233–234.

Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. Ecology Letters, 9, 615–629.

Sunnucks, P. (2000). Efficient genetic markers for population biology. Trends Ecology and Evolution, 15, 199–203.

Tajima, F. (1989). Statistical methods for testing the neutral mutation hypothesis by DNA polymorphism. Genetics, 123, 585–595.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution, 30, 2725–2729.

Tóth, G., Gáspári, Z., & Jurka, J. (2000). Microsatellites in different eukaryotic genomes: Survey and analysis. Genome Research, 10, 967–981.

Troyan, P. (1997). The floristic and faunistic Korean refugium during the last glacial period and its significance in postglacial biota formation. Fragmenta Faunistica, 40, 215–221.

Untergrasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3–new capabilities and interfaces. Nucleic Acids Research, 40, e115.

Vnukovskij, V. (1926). Nouvelles formes de Lymantria dispar L. de la Sibérie et de semirestshje. Revue Russe d’Entomologie, 20, 78–81.

Wu, Y., Molongoski, J. J., Winograd, D. F., Bogdanowicz, S. M., Louyakis, A. S., Lance, D. R., ... Harrison, R. G. (2015). Genetic structure, admixture and invasion success in a Holarctic defoliator, the gypsy moth (Lymantria dispar, Lepidoptera: Erebidae). Molecular Ecology, 24, 1275–1291.

Yoon, S. O., & Hwang, S. (2009). The reconstruction of natural environment for the last glacial maximum around Korea and adjacent areas. Journal of the Korean Geomorphological Association, 16, 101–112. [in Korean].

Yoo, J. N., Won, C., Jun, J., Lim, Y. W., & Kwak, M. (2011). Fast and cost-effective mining of microsatellite markers using NGS Technology: An example of a Korean water deer Hydroptes inermis argyropus. PLoS One, 6, e26933.

Zane, L., Bargelloni, L., & Patarnello, T. (2002). Strategies for microsatellite isolation: A review. Molecular Ecology, 11, 1–16.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.