Development of 17 polymorphic microsatellite loci from Jeju striped field mouse, *Apodemus agrarius chejuensis* (Rodentia: Muridae), by 454 pyrosequencing

Han-Na Kim¹,², Han-Ul Kim¹, Yeong-Seok Jo²* and Jongwoo Jung¹,³*

Abstract

**Background:** The striped field mouse, *Apodemus agrarius*, is the most common mammal in Korea. Although microsatellite loci for the species have been identified from populations in southwestern China, amplification of those markers for Korean populations have been unsuccessful. The complicated taxonomy of Korean striped field mouse including populations on Jeju Island (*A. a. chejuensis*) necessitates identification of additional molecular markers.

**Findings:** We applied 454 pyrosequencing systems to develop a suite of microsatellite markers. Muscle tissue was harvested and sequenced from 30 Jeju striped field mouse specimens which yielded 12,165 reads with a mean length per read of 287 bp. From these reads, we identified 17 microsatellite loci for *A. a. chejuensis* and tested these new markers against samples of both *A. a. chejuensis* and *A. a. coreae*, the mainland taxon. All 17 loci were amplified successfully for both taxa. Of the total 17 loci, one locus failed to amplify for a population on Heuksan Island. The cross-species transferability was also tested with the allied taxon, *A. peninsulae* and confirmed successful for 12 loci.

**Conclusions:** These newly developed markers will benefit studies of genetic structure, evolution, and resolving taxonomic problems of striped field mice and allied taxa in Korea.

**Keywords:** *Apodemus agrarius chejuensis*, Microsatellites, South Korea, Jeju Island, Pyrosequencing

Introduction

With 24 subspecies, the striped field mouse, *Apodemus agrarius* (Pallas, 1771) is widely distributed throughout Eurasia [1]. Striped field mice are the most abundant mammal in Korea with an important place in the ecosystems as a primary prey item for predators and a vector of seed dispersal [2]. In Korea, four subspecies (*A. a. manchuricus, A. a. chejuensis, A. a. pallenscens*, and *A. a. coreae*) have been reported, but only *A. a. chejuensis* on Jeju Island has a definitive geographic barrier from the other subspecies [3]. Based on differences in morphology [4–7], physiology [8], and mitochondrial DNA [9–11], the taxonomic status of *A. a. chejuensis* has changed frequently in rank between species and subspecies. The striped field mouse is the first known vector of Hanta virus, which killed roughly 3000 United Nations soldiers per year during the Korean War, which has attracted great attention by epidemiologists [12]. Because hemorrhagic fever never occurred in Jeju Island, the island subspecies was used as a control for the isolation of Hanta virus [12].

Simple Sequence Repeats (SSRs, microsatellites) are one of the most common and efficient nuclear DNA markers for population genetics and phylogenetic studies of recently diverged taxa because they are multiallelic, highly polymorphic, abundant and show a codominant mode of inheritance [13, 14]. However, development of microsatellite markers has been costly, ineffective, and time consuming [15]. The application of next generation
sequencing technology has greatly improved these challenges over the past few years [14].

To address ongoing taxonomic debates, the status of *A. a. chejuensis* and its relationship to mainland species should be examined using both haplotypic (mitochondrial DNA markers) and genotypic molecular markers (nuclear DNA markers). Wu et al. [16] developed 14 loci of microsatellite markers (nuclear DNA markers) and genotypic molecular markers (nuclear DNA markers) for future investigation of genetic properties of the striped field mouse and then tested cross-species transferability run because Jo et al. [3] suggested that the *A. agrarius* population in Heuksan Island might be a newly diverged subspecies from the mainland populations (*A. a. coreae*) in Korea.

The following genetic diversity parameters were calculated using ARLEQUIN 3.5 [20]: the number of alleles at each locus (*A*), observed heterozygosity (*H*), expected heterozygosity (*H*) and population divergence (*F*), between the two populations (BH and YD). Polymorphic information contents (PIC) were also estimated in CENES software (http://www.ufv.br/dbg/genes/genes.htm; Cruz 2007). We tested whether significant deviation from Hardy-Weinberg equilibrium (HWE) within each population across all amplified loci in ARLEQUIN 3.5. Linkage disequilibrium (LD) between pairs of loci within each

**Sample preparation and genotyping**

We used muscle tissues of *A. a. chejuensis* collected on Jeju Island to develop microsatellites loci. We collected 15 individuals each at Yongdu-dong (YD), Jeju-si (N33°31′08.72″ E126°29′47.77″) and 15 individuals at Beophwan-dong (BH), Seogwipo-si, Korea (N33°24′42.97″ E126°51′52.74″). We captured mice in Sherman traps following guidelines of the American Society of Mammalogists [17]. Specimens were deposited in the mammal collection at the National Institute of Biological Resources, Korea (NIBRMM0000105863).

Genomic DNA was extracted using DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA, USA). The isolated DNA was visualized on 1% agarose gel and quantified in NanoDrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA). We used a total amount of over 2 μg with concentration of 50 ng/μL or higher for 454 pyrosequencing library preparation. Library preparation, amplification, and sequencing were conducted at Macrogen Inc. (Seoul, Korea). Briefly, after DNA quality control steps, isolated genomic DNAs were fragmented using restriction enzymes and tagged them with two multiplex identifiers. After amplification of DNA fragments, the prepared libraries were pooled and sequenced in 454 GS FLX platform (Roche).

**Primer design and testing makers**

Candidate loci for microsatellites were selected using QDD3 software [18]. A total of 335 primer pairs were obtained for candidate loci using PRIMER3 [19]. Each locus was then evaluated with PCR amplification performed with a total volume of 25 μl containing 1.0 μl of genomic DNA, 2.5 μl of 10× buffer (Takara), 0.7 μl of dNTP (2.5 mM each), 1.5 μl of MgCl₂, 0.5 μl of primers each, and 0.3 U of Taq polymerase (Takara r-Taq, Takara). PCR conditions were follows: initial denaturation at 95 °C for 5 min followed by 35 cycles of denaturation at 94 °C for 60 s, annealing at 45–60 °C for 60 s, elongation at 72 °C for 90 s, and a final extension at 72 °C for 10 min. PCR products were selected and sequenced to confirm whether they contained microsatellite sequences.

A M13-tail (FAM: 5′-TTTCCCCAGTCACGACGTGG-3′, VIC: 5′-GGAAACACGTGATGACCA-3′, PET: 5′-GGCGATAAACCATTCACACAGG-3′, NED: 5′-TAAACGACGGCCAGTGC-3′) was added to the 5′ end of each forward primers while a pig tail (GTTCCTTT) was added to the 5′ end of each reverse primers. For multiplex PCR, we performed PCR amplifications in a total reaction volume of 16 μl containing 5 μl of 2× QIAGEN Multiplex PCR master mix (QIAGEN). 0.08 μl of M13-tailed forward primer, 0.8 μl of pig tailed reverse primer, 0.3 μl of template DNA and 0.16 μl of each of fluorescent primer (FAM/VIC/PET/NED). Multiplex PCR cycling conditions were as follows: initial denaturation at 95 °C for 15 min followed by denaturation at 95 °C for 30 s, annealing (14 cycles at 63 °C, 7 cycles at 58 °C, 20 cycles at 55 °C) for 90 s and 72 °C for 30 s, with a final elongation at 72 °C for 20 min.)

**Genotyping and population analysis**

We genotyped 30 Jeju striped field mice (*A. a. chejuensis*) from two populations (YD and BH) with the newly designed microsatellites. The amplified fragments were separated in ABI 3730x1 DNA analyzer (Applied Biosysstem, Inc, Forster City, CA, USA) and GeneMapper software v3.7 (Applied Biosystems) was used for manual scoring.

To test for cross-species transferability, we amplified one congener (*Agrarius peniculae*) and one subspecies (*A. a. coreae*) using the 17 developed markers with the same PCR conditions. Additionally, we added an *A. agrarius* sample collected from Heuksan Island in the cross-species transferability run because Jo et al. [3] suggested that the *A. agrarius* population in Heuksan Island might be a newly diverged subspecies from the mainland populations (*A. a. coreae*) in Korea.
population was tested using GENEPOP 4.2.1 [21, 22]. All statistical significances of multiple comparisons were adjusted with sequential Bonferroni corrections [23]. We performed Analysis of Molecular Variance (AMOVA) to hierarchically partition the total genetic variance within and between the two populations in GENALEX v. 6.5 [24]. We also carried out principal coordinate analysis (PCoA) based on pairwise $F_{ST}$ for all 30 individuals to examine clustering pattern among the individual genotypes in GENALEX v. 6.5 [24].

### Results & discussion

Four hundred and fifty-four pyrosequencing produced 12,165 reads with a mean length of 287 bp. We selected 335 candidate microsatellite loci from the produced reads and confirmed that 17 loci were polymorphic (Table 1).

The number of alleles per locus in the two populations of *A. a. chejuensis* varied from 2 to 11 in YD and 2–13 in BH. Ranges of $H_E$ and $H_O$ in YD were 0.067–0.905 and 0.067–0.867, respectively. In BH, $H_E$ and $H_O$ were

| Locus | Repeat motif | Repeat number range | Primer sequence (5’-3’) | Dye | Size range (bp) |
|-------|--------------|---------------------|-------------------------|-----|-----------------|
| AC2 L | TGTA         | 5–10                | F: TTTCCAGTCACGACTGGAGTCTCGAGGAAGAGTGGCATCCTAAAGAAGG           | 6FAM | 213–233 |
|       |              |                     | R: GTTTCTTGGGATCATGACGATGGCAAGG                          |      |                 |
| ACH7 | AACC         | 5–9                 | F: GGCCTGTACTACCACTCCTCC | 6FAM | 139–155 |
|       |              |                     | R: GGTAGGTCAATGAGTGGGTTC                                  |      |                 |
| ACW7 | CA           | 17–27               | F: TTTCCAGTCACGACTGGGATGACTGTAAGATGGCATCCTAAAGAAGG          | 6FAM | 211–231 |
|       |              |                     | R: GTTTCTTGGCCAGCTCGACTGACTACATGG                        |      |                 |
| AC9S | CA           | 18–25               | F: TAAAACGACTGGCAACGACTGGGATGACTGTAAGATGGCATCCTAAAGAAGG | 6FAM | 260–274 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| ACAQ | TG           | 15–31               | F: TTTCCAGTCACGACTGGGATGACTGTAAGATGGCATCCTAAAGAAGG         | 6FAM | 162–194 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| ACWV | CCA          | 6–11                | F: TTTCCAGTCACGACTGGGATGACTGTAAGATGGCATCCTAAAGAAGG         | 6FAM | 316–331 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| ACAV | AG           | 5–9                 | F: TTTCCAGTCACGACTGGGATGACTGTAAGATGGCATCCTAAAGAAGG         | 6FAM | 192–200 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| ACU8 | AACC         | 6–18                | F: TTGGTACATTACGATGTAAGG                                  | HEX  | 291–339 |
|       |              |                     | R: TGCAAGAAATACAAACGACTGATCATAACGACGTCACGTACG             |      |                 |
| AC9 M | CA         | 4–5                 | F: GGAAACAGCTATGACCCACGACGCAAGGAGGAGGAGGAGGACTGAGG         | VIC  | 165–167 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| AC124 | TG         | 6–8                 | F: GGAAACAGCTATGACCCACGACGCAAGGAGGAGGAGGACTGAGG         | VIC  | 278–282 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| ACEC | ATGA         | 5–10                | F: GGAAACAGCTATGACCCACGACGCAAGGAGGAGGAGGACTGAGG         | VIC  | 148–168 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| ACQT | AG           | 7–12                | F: GGAAACAGCTATGACCCACGACGCAAGGAGGAGGAGGACTGAGG         | VIC  | 160–170 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| ACHZ | TTTTG        | 7–16                | F: GCGGATAAACACATTACACAGGGCCCCTGCTGCTGCTGCTGCTGCTGCT    | PET  | 312–357 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| ACS3 | AAAC         | 7–15                | F: GCGGATAAACACATTACACAGGGCCCCTGCTGCTGCTGCTGCTGCTGCT    | PET  | 313–348 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| ACLI | TG           | 5–6                 | F: TAAAACGACCCGACGACGCAAGGAGGAGGAGGACTGAGG         | VIC  | 231–233 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| ACS5 | AC           | 20–49               | F: TAAAACGACCCGACGACGCAAGGAGGAGGACTGAGG         | NED  | 153–211 |
|       |              |                     | R: GTTTCTTACCAATGAGTGGCATCCTAAAGAAGG                      |      |                 |
| ACPF | AC           | 10–36               | F: TAAAACGACCCGACGACGCAAGGAGGAGGACTGAGG         | NED  | 132–184 |

F denote forward sequences and R denote reverse sequences
The value of polymorphic information content ranged from 0.067 to 0.856 with a mean of 0.537 in YD. It ranged from 0.062 to 0.872 with a mean of 0.672 in BH. Three loci in YD and four loci in BH were significantly deviated from HWE after Bonferroni correction ($p < 0.0001$). No linkage disequilibrium was detected at any pair of loci.

$F_{ST}$ (0.05) estimated between the two Jeju striped field mouse populations was not significantly different from zero ($P > 0.99$). The AMOVA results revealed that in the two populations most of genetic variation (88%) was harbored within each population, whereas among population genetic variation was proportionally minor (12%; Fig. 1). Coupled with the small $F_{ST}$ value, the AMOVA results suggest that YD and BH are not greatly diverged in part due to high gene flow between the two populations. However, the PCoA analysis showed spatial structure among the two populations. Overall, the genotypes from each population are affiliated with the population from which the genotypes were drawn (Fig. 2). The clustering pattern indicates that the 17 markers developed from our study have sufficient power to separate out the two local populations.

Cross-species amplification examined for a subspecies, *A. a. coreae*, and its congeneric species, *A. peninsulae*, in South Korea showed high cross-species
transferability. All 17 loci were successfully amplified for *A. a. coreae* and 12 of 17 loci were successfully amplified for *A. peninsulae* (Table 3). Except for the AC53 locus, all microsatellite loci were successfully amplified for individuals collected on Heuksan Island. The cross-species amplification could be an indicator of genetic distance among species. The Jeju island subspecies, *A. a. chejuensis* has often been regarded a distinct species in Korea [3]. In contrast, the classification of the population on Heuksan Island has been *A. a. coreae* (the same as the mainland populations) in most of the literature despite its high genetic divergence from mainland populations as revealed in a microsatellite analysis [3]. Our cross-species amplification tests suggested that the subspecies *A. a. chejuensis* might be more closely related to the mainland populations of *A. a. coreae* than the populations on Heuksan Island. Intensive population level studies with these developed markers might elucidate the true phylogenetic relationships among these three taxa.

Although this species is the main vector of Hanta virus and the most common mammal in the Korean Peninsula, few studies have addressed populations of the striped field mouse in Korea. One of the limitations in the study of *A. agrarius* is the deficiency of genetic markers [4]. The microsatellite loci developed in this study have proven to be useful for investigating genetic structure, introgression, hybrid, and taxonomic status of *A. agrarius* in not only the Korean Peninsula but the entire distribution for the species. The genetic markers might also be used to establish a management strategy for *A. a. chejuensis*, the most common endemic rodent on Jeju Island.

### Table 3 Results of cross-species amplifications

| Loci | *Apodemus agrarius coreae* | *Apodemus peninsulae* |
|------|-----------------------------|----------------------|
| AC2L | +                           | +                    |
| ACH7 | +                           | +                    |
| ACW7 | +                           | +                    |
| AC9S | +                           | +                    |
| ACAQ | +                           | +                    |
| ACW | +                           | –                    |
| ACAV | +                           | –                    |
| ACU8 | +                           | +                    |
| AC9M | +                           | +                    |
| AC124 | +                           | +                    |
| ACEC | +                           | –                    |
| ACQT | +                           | –                    |
| ACHZ | +                           | +                    |
| AC3S | +                           | +                    |
| ACL1 | +                           | –                    |
| AC5S | +                           | +                    |
| ACPJ | +                           | +                    |

Seven individual of each species were tested.

Note: + successful PCR amplification; − PCR failed.

![Fig. 2 PCoA plot for 30 individual genotypes of *A. agrarius chejuensis* from two populations. The first two variance components PC1 and PC2 were plotted. BHs are samples from Beophwan-dong and YDs are from Yongdu-dong.](image-url)
Acknowledgments
This study was supported by a grant (NIBR201403201) from the National Institute of Biological Resources (NIBR), funded by the Ministry of Environment (MOE), Republic of Korea. Also, NIBR supported field collections of this study as part of Wildlife Survey (NIBR20182401). We also thank Drs. J. Baccus, P. Schlichting, M. Forstner, anonymous reviewers with the editors Y. Shi and S. Baumgarter for valuable suggestions and comments.

Funding
Genetic works were supported by a grant (NIBR201403201) and field collection was supported by Wildlife Survey (NIBR20182401).

Availability of data and materials
Dataset is provided by an additional file. All specimens are stored in National Institute of Biological Resources, Korea. It’s available after obtaining permission from the agency.

Authors’ contributions
HNK and JJ conceived the idea of this study. HNK collected samples. HNK and HUK performed lab works. YSJ and JJ wrote the manuscript. All authors read and approved the final manuscript.

Ethics approval
Sample collections were followed by guidelines of American Society of Mammalogists. This study was approved by Research Management Committee of National Institute of Biological Resources, Ministry of Environment, Korea.

Consent for publication
Mammalogists. This study was approved by Research Management Committee of National Institute of Biological Resources, Ministry of Environment, Korea.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1The Division of EcoCreative, Ewha Womans University, Seoul 03760, Korea.
2Animal Resources, National Institute of Biological Resources, Incheon 22689, Korea.
3Department of Science Education, Ewha Woman University, Seoul 03760, Korea.

Received: 24 April 2018 Accepted: 20 September 2018
Published online: 26 September 2018

References
1. Musser GG, Carleton MD. Superfamily Muroidea. In: Winson DE, Reeder DM, editors. Mammal species world a Taxon Geogr Ref. 3rd ed. Baltimore: Johns Hopkins University Press; 2005. p. 894–1532.
2. Won PH. Illustrated encyclopedia of fauna and flora of Korea volume 7, mammals. Seoul: Ministry of Education; 1967.
3. Jo Y, Kim H, Baccus JT, Jung J. Genetic differentiation of the Korean striped field mouse, Apodemus agrarius (Muridae, Rodentia), based on microsatellite polymorphism. Mammalia. 2017;81:297–307.
4. Koh HS. Geographic variation of morphometric characters among three subspecies of striped field mouse, Apodemus agrarius Pallas (Rodentiae, Mammalia), from Korea. Korean J Biol Zool. 1986;29:272–82.
5. Koh HS. Systematic studies of Korean rodents. III. Morphometric and chromosomal analyses of striped field mouse, Apodemus agrarius chejuensis Jones and Johnson, from Jeju-Do. Korean J Syst Zool. 1987;3:24–40.
6. Koh HS. Systematic studies of Korean rodents. V. Morphometric and chromosomal analyses on island populations of striped field mouse (Apodemus agrarius coreae) in southwestern coasts of the Korean peninsula. Korean J Syst Zool. 1989;5:1–12.
7. Koh HS. Morphometric analyses with eight subspecies of striped field mice, Apodemus agrarius Pallas (Rodentia, Mammalia), in Asia: the taxonomic status of subspecies chejuensis at Cheju island in Korea. Korean J Syst Zool. 1991;7:179–88.
8. Oh HS, Mori T. Taxonomic re-examination of the striped field mouse, Apodemus agrarius coreae and A. a. chejuensis: evidence from crossbreeding experiments. J Fac Agric Kyushu Univ. 1998;43:143–51.
9. Koh HS, Yoo BS. Variation of mitochondrial DNA in two subspecies of striped field mouse, Apodemus agrarius coreae and Apodemus agrarius chejuensis, from Korea. Korean J Zool. 1992;35:332–8.
10. Koh HS, Lee WJ, Kocher TD. The genetic relationships of two subspecies of striped field mice, Apodemus agrarius coreae and Apodemus agrarius chejuensis. Heredity (Edinb). 2000;85:30–6.
11. Oh D-J, Kim T-W, Chang MH, Han S-H, Oh H-S, Kim S-J. Migration route estimation of the Jeju striped field mouse Apodemus agrarius chejuensis (Rodentia, Muridae). Mitochondrial DNA. 2013;24:137–44.
12. Johnson KM, Hanov. Rabies. history and overview. Curr Top Microbiol Immunol. 2001;256:1–14.
13. Elleheen H. Microsatellites: simple sequences with complex evolution. Nat Rev Genet. 2004;5:435–45.
14. Lopez L, Barreiro R, Fischer M, Koch MA. Mining microsatellite markers from public expressed sequence tags databases for the study of threatened plants. BMC Genomics [Internet]. BMC Genomics, 2015;16. Available from: doi:https://doi.org/10.1186/s12864-015-2031-1
15. Squirell J, Hollingsworth PM, Woodhead M, Russell J, Lowe AJ, Gibby M, et al. How much effort is required to isolate nuclear microsatellites from plants? Mol Ecol. 2003;12:139–48.
16. Wu H, Zhan XJ, Yan L, Liu SY, Li M, Hu JC, et al. Isolation and characterization of fourteen microsatellite loci for striped field mouse (Apodemus agrarius). Conserv Genet. 2008;9:1691–3.
17. Sikes RS. 2016 guidelines of the American Society of Mammalogists for the use of wild mammals in research and education. J Mammal. 2016;97:643–88.
18. Meglécz E, Pech N, Gilles A, Dubut V, Hingamp P, Trilles A, et al. QDD version 3.1: a user-friendly computer program for microsatellite selection and primer design revisited: experimental validation of variables determining genotyping success rate. Mol Ecol Resour. 2014;14:1302–13.
19. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3-new capabilities and interfaces. Nucleic Acids Res. 2012;40:1–12.
20. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010;10:564–7.
21. Raymond M, Rouset F. GENEPOP (version 1.2): population genetics software for exact tests and euclimicms. J Hered. 1995;86:248–9.
22. Rouset F. GENEPOP’007: a complete re-implemention of the GENEPOP software for windows and Linux. Mol Ecol Resour. 2008;8:103–6.
23. Rice WR. Analyzing tables of statistical tests. Evolution (N Y). 1989;43:223–5.
24. Peakall R, Smouse PE. GenALEx 6.5: genetic analysis in excel. Population genetic software for teaching and research-an update. Bioinformatics. 2012;28:2537–9.

Ready to submit your research? Choose BMC and benefit from:
- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions