The Journal of Veterinary Medical Science

NOTE

Wildlife Science

Development of microsatellite markers and the genetic diversity of *Myocastor coypus* introduced to South Korea

Dong-Young KIM[1)#], Ho Young SUK[1)#], Seul-Ki PARK[1)], Sun-Young KANG[2)], Seong-Hoon SEOK[2)], Seong-Kyu LEE[3)] and Seong-Chan YEON[4)*]

[1)]Department of Life Sciences, Yeungnam University, 280 Daehak-ro, Gyeongsan, Gyeongsangbuk-do 38541, South Korea
[2)]The Gyeongnam WildLife Center, Gyeongsang National University, 501 Jinju-daero, Jinju, Gyeongnam 52828, South Korea
[3)]Nakdong River Basin Environment Officer, 5 Jungang-daero, Gyeongnam 51439, South Korea
[4)]Laboratory of Wildlife Medicine and the Research Institute for Veterinary Science, College of Veterinary Medicine, Seoul National University, Seoul 08826, South Korea

J. Vet. Med. Sci.
81(3): 499–503, 2019
doi: 10.1292/jvms.18-0564

Received: 19 September 2018
Accepted: 22 January 2019
Published online in J-STAGE: 1 February 2019

ABSTRACT. The nutria (*Myocastor coypus*) was introduced to South Korea in 1987 for breeding of individuals for fur and meat industry, and was accidentally released into the wild. Here, we report the development of microsatellites for the nutria collected from South Korea using Illumina MiSeq genome sequencing to identify the genetic variability and demographic history of these introduced populations. A total of 626,282 microsatellite sequences were identified, and nine polymorphic loci were characterized. We used four novel loci developed and three previously known loci to investigate the genetic diversity of twelve South Korean populations. A low level of diversity was found, and no signature of genetic structuring was revealed among populations, indicating that Korean nutria individuals originated from a single population or a highly inbred reared herd.

KEY WORDS: genetic structure, Illumina MiSeq, invasive species, microsatellites, *Myocastor coypus*

The nutria (*Myocastor coypus*) is a large herbivorous rodent species belonging to Echimyidae [11, 19]. Originally native to subtropical and temperate South America [2, 3], this species was introduced via Bulgaria to South Korea in 1987 primarily for breeding of individuals that could be used for fur and meat industry [3, 9]. This species was accidentally released into the wild, and has become a mammalian species that is frequently found in nature. The Ministry of Environment in South Korea designated this species as an ecological disturbance species in 2009, and with the local governments, started to control the population in nature [9]. This species is still considered as pests in many areas of South Korea, especially on the southern part, destroying aquatic vegetation, marshes and irrigation systems [9].

There are no basic ecological and historical records in South Korea on how many individuals are currently living in the wild and how many or diverse broodstocks have been introduced to South Korea. The absence of information on introduced wildlife also means that it is difficult to control the disease or the risks that this organism may cause. Insights can be obtained into contemporary demographic history by the practice of hypervariable and codominant neutral molecular loci such as microsatellites for the estimation of genetic diversity and structure among populations [1, 17]. Primers for microsatellites of this species are already known [2]. Given that introduced species are likely to be genetically very monomorphic [15], however, novel polymorphic loci are required to be characterized from these introduced individuals. Here we report the development of microsatellite markers for the nutria collected from South Korea using Illumina MiSeq genome sequencing technology. The obtained polymorphic loci were used to identify the genetic variability and demographic history of these introduced populations. Many individuals of this species are found mainly in the southeastern part of South Korea, and the sampling was accomplished throughout most of the known areas (Fig. 1).

For the development of microsatellite markers, total genomic DNA of 48 individuals collected from two localities (Goryeong and Haman) in South Korea (Table 1; Fig. 1) was extracted with a DNeasy Blood and Tissue Kit (Qiagen, Dusseldorf, Germany). A high-quality DNA sample was selected and sent to Macrogen Inc. (Seoul, South Korea), where library construction, quality check and sequencing were performed. One microgram was processed in a hydrodynamic shearing system (Covaris, Woburn, CA,

*Correspondence to: Yeon, S.-C.: scyeon1@snu.ac.kr #These authors contributed equally to this work.

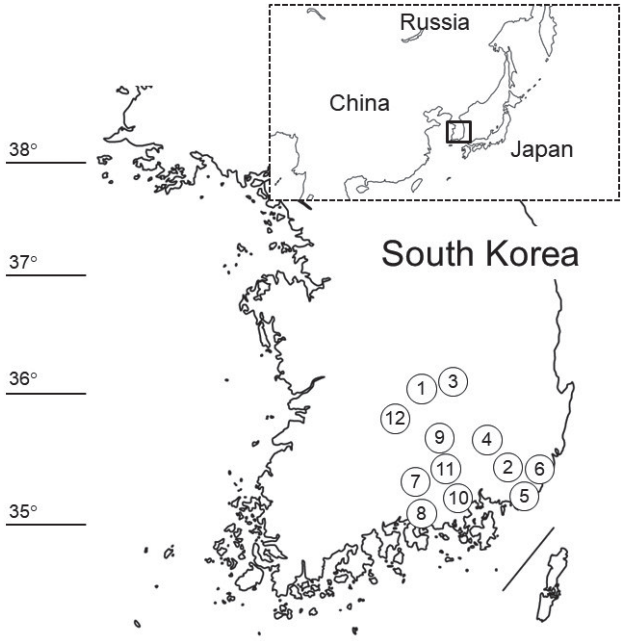

Fig. 1. Sampling location of the 12 *Myocastor coypus* populations on the Korean Peninsula. See Table 3 for locality numbers.

Table 1. The nine novel microsatellite loci developed in this study and the diversity estimates obtained from two *Myocastor coypus* populations, Goryeong and Haman (N=48)

Name	A	A_R	H_O	H_E	F_{IS}	F_{ST}	R_{ST}	NCBI ACCN #
Mck044	5	4.929	0.704	0.661	−0.091	0.034	−0.018	MH346456
Mck055	3	2.998	0.576	0.445	−0.253	0.088	0.103	MH346457
Mck087	4	3.925	0.774	0.661	−0.152	0.009	−0.003	MH346458
Mck101	3	2.975	0.591	0.558	−0.087	−0.020	−0.020	MH346459
Mck130	5	4.993	0.691	0.743	0.068	0.076	0.017	MH346460
Mck141	3	2.970	0.313	0.301	−0.017	0.017	−0.022	MH346461
Mck226	3	2.986	0.626	0.571	−0.100	0.008	0.006	MH346467
Mck250	4	3.637	0.695	0.657	−0.066	0.007	−0.017	MH346469
Mck291	3	3.000	0.636	0.583	−0.068	0.076	0.009	MH346471

Data include total number of alleles (A), allelic richness (A_R), observed (H_O) and expected (H_E) heterozygosities, fixation indices (F_{IS}), and NCBI GenBank accession numbers (NCBI ACCN #).

U.S.A.) to produce fragments of less than 800 bp. The fragments were blunt-ended and phosphorylated, and the appropriate library sizes (around 450 bp) were selected by purification beads. A paired-end library was constructed using the TruSeq DNA Sample Prep kit (Illumina, San Diego, CA, U.S.A.). The library constructed was electrophoresed on 1% agarose gel to extract fragments of approximately 500 bp, and the purity of eluted fragments was checked using a Quant-iT^{TM} dsDNA High Sensitivity Assay Kit (Invitrogen, Carlsbad, CA, U.S.A.) on an Agilent 2100 Bioanalyzer (Agilent Technologies, San Diego, CA, U.S.A.). The final library preparation was sequenced using the MiSeq sequencing (Illumina) platform to generate 300-bp paired-end reads (600-cycle Miseq Reagent Kit v3). Raw image files obtained in the platform were processed to generate base calling through RTA (Real Time Analysis v1.18, Illumina), which was converted into FASTQ utilizing MSR (Miseq Reporter, Illumina).

A total of 31,805,658 high-quality reads, consisting of 9,545,204,206 bp (Q20=72.90%), were generated, and the reads with Phred quality scores of less than 20 and lengths of less than 20 bp were removed. De novo genome assembly was performed with the high quality filtered reads using the de Brujin graph algorithm and the SOAPdenovo2 software package (v. 2.04) under the default settings [10]. Consequently, 626,282 microsatellite sequences were identified using the MicroSAtellite Identification Tool (MISA) [18]; the composition of di-, tri- and tetranucleotide repeats were 32.5% (203,662 loci), 40.2% (251,719 loci) and 24.3% (152,147 loci), respectively. All of the microsatellite information are stored in Yeungnam University database (http://suk.yu.ac.kr/english.html) and can be released freely to anyone who needs this. Of these, 169 loci were selected among 15–50 repeat types for design of primer pairs using Primer3 software [16]. Primer pairs were synthesized for those loci and were examined on whether successfully generating PCR products in eight randomly selected individuals. All PCR amplifications were performed in a 10 μl volume containing 1 μl DNA template, 1X Taq Buffer, 0.2 mM dNTPs, 2.5 μM of each primer and 2.5 units of Taq DNA

Table 2. The seven microsatellite loci developed in this study (*Mck044, Mck087, Mck130* and *Mck250*) and previously reported (*McoD212, McoD215* and *McoD69*) and the diversity estimates obtained from twelve *Myocastor coypus* populations (N=269)

Name	A	A_R	H_O	H_E	F_{IS}	F_{ST}	R_{ST}
McoD212	3	2.066	0.528	0.456	−0.171	0.053	0.057
McoD215	4	3.941	0.233	0.662	0.592	0.111	0.251
McoD69	3	2.782	0.549	0.536	−0.018	0.049	0.048
Mck044	5	3.977	0.647	0.655	0.004	0.043	0.056
Mck087	6	4.292	0.721	0.668	−0.061	0.078	0.062
Mck130	6	5.544	0.817	0.763	−0.095	0.060	0.045
Mck250	5	3.606	0.626	0.629	0.001	0.037	0.024

Data include total number of alleles (A), allelic richness (A_R), observed (H_O) and expected (H_E) heterozygosities, fixation indices (F_{IS}), F_{ST} and R_{ST}.

Table 3. List of twelve *Myocastor coypus* populations on the Korean Peninsula and the diversity estimates obtained from the genetic structure analysis with seven microsatellites (see Table 2)

Name	Locality	N	A	A_R	H_O	H_E	F_{IS}	F_{ST}	R_{ST}
1. Goryeong	35°42'N 128°21'E	32	3.7	3.545	0.692	0.656	−0.055	0.040	0.031
2. Gimhae	35°11'N 128°51'E	23	3.9	3.460	0.590	0.631	0.066	0.052	0.039
3. Daegu	35°51'N 128°44'E	27	3.7	3.446	0.677	0.629	−0.078	0.068	0.074
4. Milyang	35°23'N 128°50'E	24	3.9	3.555	0.625	0.664	0.060	0.058	0.038
5. Busan	35°06'N 128°54'E	34	3.7	3.288	0.542	0.586	0.075	0.075	0.052
6. Yangsan	35°19'N 129°01'E	22	3.9	3.495	0.552	0.583	0.054	0.080	0.064
7. Uiryeong	35°21'N 128°22'E	10	3.6	3.526	0.543	0.637	0.155	0.083	0.042
8. Jinju	35°12'N 128°09'E	25	3.1	3.032	0.560	0.614	0.090	0.068	0.039
9. Changnyeong	35°30'N 128°28'E	22	3.3	3.152	0.610	0.627	0.028	0.079	0.106
10. Changwon	35°18'N 128°40'E	20	3.6	3.348	0.507	0.596	0.153	0.057	0.030
11. Haman	35°18'N 128°23'E	21	3.4	3.268	0.612	0.623	0.017	0.068	0.061
12. Hapcheon	35°30'N 128°10'E	9	3.7	3.714	0.556	0.631	0.126	0.086	0.107

Data include average number of alleles (A), allelic richness (A_R), observed (H_O) and expected (H_E) heterozygosities, fixation indices (F_{IS}), F_{ST} and R_{ST}.

polymerase (Genetbio, Daejeon, South Korea) under the following thermal cycling conditions (GenePro, Bioer, PR China): an initial denaturation at 94°C for 5 min followed by 30 cycles each of denaturation at 94°C for 30 sec, annealing at 58°C for 30 sec, and an extension at 72°C for 45 sec, with a final extension at 72°C for 10 min.

Only nine primer pairs showed polymorphism in the eight individuals, which was understandable considering that this species is a recently introduced species and its genetic variation would be extremely low. The information of the remaining 160 loci is stored and freely available from http://suk.yu.ac.kr/english.html. The forward primers of those loci were labeled with a fluorescent dye of 6-FAM or Hex (Applied Biosystems, San Diego, CA, U.S.A.). The fluorescence-labeled PCR products were fragmented on an ABI3730XL genetic analyzer following the standard protocol (Applied Biosystems) in Macrogen Inc. Allele peaks were scored using GeneMapper 3.7 (Applied Biosystems) and Peak Scanner 1.0 (Applied Biosystems). FSTAT 2.9.3.2 [7] and Arlequin 3.5 [6] were applied to quantify various diversity indices, including average number of alleles per locus (A), allelic richness (A_R), observed (H_O) and expected heterozygosities (H_E), and inbreeding coefficient (F_{IS}). The deviations from genotype frequencies expected under Hardy-Weinberg equilibrium (HWE) were examined with the Fisher exact test following Markov chain parameters, with 1,000 batches and 10,000 iterations per batch [8] for each locus, implemented in GENEPOP 4.2 [13]. The linkage disequilibrium between each pair of loci was detected by Fisher exact tests under the Markov chain algorithm in GENEPOP. The level of statistical significance for simultaneous and multiple tests was corrected using sequential Bonferroni adjustments [14].

The overall allelic diversity of the novel 9 loci was as low as 3 to 5 (Table 1; N=48). H_O ranged from 0.313 (*Mck141*) to 0.774 (*Mck087*) and H_E ranged from 0.301 (*Mck141*) to 0.743 (*Mck130*; Table 1). No locus deviated significantly from the expectation of HWE (Table 1). No signature of linkage disequilibrium was detected between any possible pair of the nine loci analyzed following the Bonferroni adjustment (all nominal P values >0.05; data not provided).

The 5 loci, except for *Mck044, Mck087, Mck130* and *Mck250*, were very monomorphic in the other populations except Goryeong and Haman. For the investigation of genetic structure, the 15 alleles of these 5 loci were not used and replaced by 3 loci (*McoD212, McoD215* and *McoD69*; Table 2) previously reported [2]. All 12 populations (N=269) had very similar levels and very low genetic variation (Table 3). No signature of genetic structure was found among populations, considering the low F_{ST} and R_{ST} values (mostly less than 0.1) in all loci examined (Table 2), probably indicating that Korean nutria individuals originated from a single population or a highly inbred reared herd, though it is premature to conclude only with our results yet. To identify the fine-

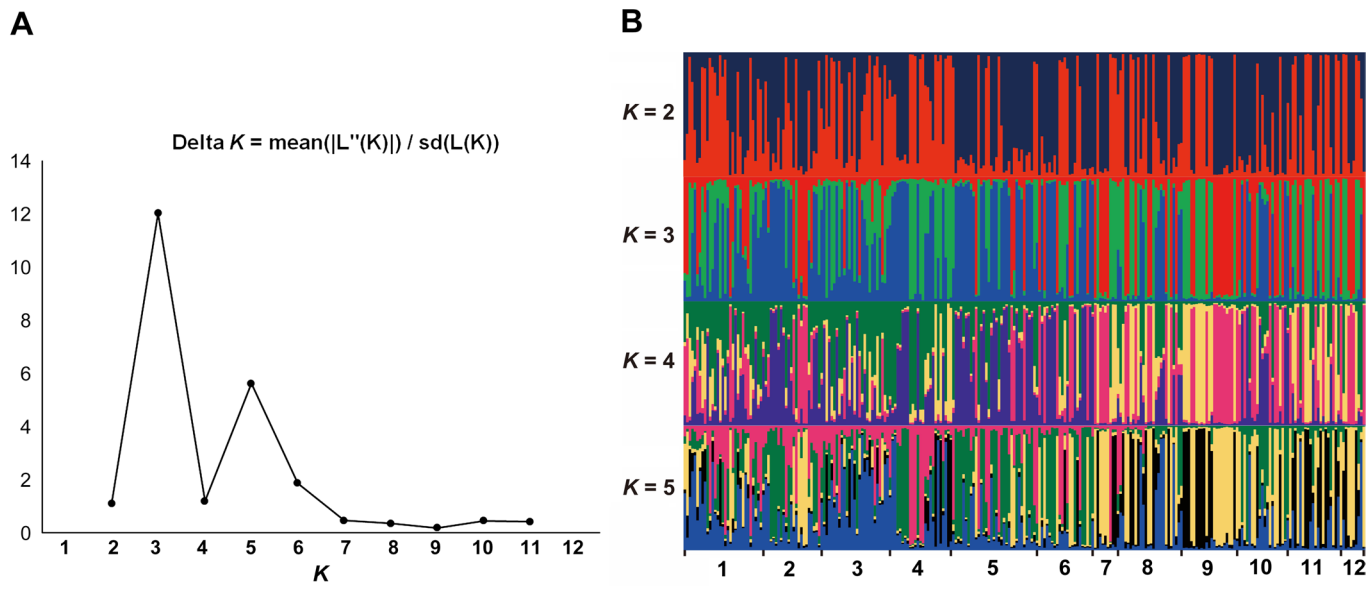

Fig. 2. A. The delta *K* method indicated that the most reliable number of cluster was three. B. Bayesian clustering obtained using Structure v2.3.4 for the result of seven microsatellite loci genotyping of twelve South Korean *Myocastor coypus* populations (x-axis; 1–12; see Table 2). The analysis was performed for all *K* values (putative number of clusters; y-axis) from 2 to 5. The number of clusters was indicated by a different color. Each population was represented by a mixture of all colors, not separated by different ones, indicating that there was signature of genetic structure among populations.

scale structuring among populations, Bayesian clustering analysis was performed using Structure v2.3.4 [12]. The putative numbers of genetic clusters (*K*) were inferred by the delta *K* method [5] implemented in Structure Harvester v0.6.94 [4]. Although three was estimated to be the most predictable number of genetic clusters in delta *K* method (Fig. 2A), the analysis was performed for all *K* values from 2 to 6. Ten independent MCMC runs were performed for each number of genetic clusters with 4×10^5 iterations after a burn-in of 10^5 iterations. Evident signature of genetic structure was not found in Structure analysis, though there was a slight bias in the frequency of genetic characteristics among populations (Fig. 2B), which can be predicted from genetic drift or sampling error in populations with extremely low genetic variability. Our genetic data can be used to monitor the demographic change of Korean populations or for the population study in other areas.

REFERENCES

1. Abdelkrim, J., Robertson, B., Stanton, J. A. L. and Gemmell, N. 2009. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *Biotechniques* **46**: 185–192. [Medline] [CrossRef]
2. Callahan, C. R., Henderson, A. P., Eackles, M. S. and King, T. L. 2005. Microsatellite DNA markers for the study of population structure and dynamics in nutria (*Myocastor coypus*). *Mol. Ecol. Notes* **5**: 124–126. [CrossRef]
3. Carter, J. and Leonard, B. P. 2002. A review of the literature on the worldwide distribution, spread of, and efforts to eradicate the coypu (*Myocastor coypus*). *Wildl. Soc. Bull.* **30**: 162–175.
4. Earl, D. A. and vonHoldt, B. M. 2012. Structure harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**: 359–361. [CrossRef]
5. Evanno, G., Regnaut, S. and Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**: 2611–2620. [Medline] [CrossRef]
6. Excoffier, L. and Lischer, H. E. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**: 564–567. [Medline] [CrossRef]
7. Goudet, J. 2001. FSTAT, a program to estimate and test gene diversity and fixation indices (version 2.9.3). http://www2.unil.ch/popgen/softwares/fstat.htm [accessed on August 7, 2018].
8. Guo, S. W. and Thompson, E. A. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**: 361–372. [Medline] [CrossRef]
9. Hong, I. H., Kang, S. Y., Kim, J. H., Seok, S. H., Lee, S. K., Hong, S. J., Lee, S. Y., Park, S. J., Kong, J. Y. and Yeon, S. C. 2017. Histopathological findings in wild Nutrias (*Myocastor coypus*) with *Capillaria hepatica* infection. *J. Vet. Med. Sci.* **78**: 1887–1891. [Medline] [CrossRef]
10. Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., Cheung, D. W., Yiu, S. M., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li, Y., Yang, H., Wang, J., Lam, T. W. and Wang, J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* **1**: 18. [Medline] [CrossRef]
11. Murphy, W. J., Eizirik, E., Johnson, W. E., Zhang, Y. P., Ryder, O. A. and O'Brien, S. J. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature* **409**: 614–618. [Medline] [CrossRef]
12. Pritchard, J. K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959. [Medline]

13. Raymond, M. and Rousset, F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* **86**: 248–249. [CrossRef]
14. Rice, W. R. 1989. Analyzing tables of statistical tests. *Evolution* **43**: 223–225. [Medline] [CrossRef]
15. Rollins, L. A., Moles, A. T., Lam, S., Buitenwerf, R., Buswell, J. M., Brandenburger, C. R., Flores-Moreno, H., Nielsen, K. B., Couchman, E., Brown, G. S., Thomson, F. J., Hemmings, F., Frankham, R. and Sherwin, W. B. 2013. High genetic diversity is not essential for successful introduction. *Ecol. Evol.* **3**: 4501–4517. [Medline] [CrossRef]
16. Rozen, S. and Skaletsky, H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* **132**: 365–386. [Medline]
17. Selkoe, K. A. and Toonen, R. J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecol. Lett.* **9**: 615–629. [Medline] [CrossRef]
18. Thiel, T., Michalek, W., Varshney, R. K. and Graner, A. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **106**: 411–422. [Medline] [CrossRef]
19. Woods, C. A. and Howland, E. B. 1979. Adaptive radiation of capromyid rodents: anatomy of the masticatory apparatus. *J. Mammal.* **60**: 95–116. [CrossRef]