Genetic Diversity and Population Genetic Structure of Black-spotted Pond Frog (*Pelophylax nigromaculatus*) Distributed in South Korean River Basins

Jun-Kyu Park[^1^], Nakyung Yoo[^2^], Yuno Do[^*^][^1^]

[^1^]Department of Biological Science, Kongju National University, Gongju, Korea

**ABSTRACT**

The objective of this study was to analyze the genotype of black-spotted pond frog (*Pelophylax nigromaculatus*) using seven microsatellite loci to quantify its genetic diversity and population structure throughout the spatial scale of basins of Han, Geum, Yeongsan, and Nakdong Rivers in South Korea. Genetic diversities in these four areas were compared using diversity index and inbreeding coefficient obtained from the number and frequency of alleles as well as heterozygosity. Additionally, the population structure was confirmed with population differentiation, Nei’s genetic distance, multivariate analysis, and Bayesian clustering analysis. Interestingly, a negative genetic diversity pattern was observed in the Han River basin, indicating possible recent habitat disturbances or population declines. In contrast, a positive genetic diversity pattern was found for the population in the Nakdong River basin that had remained the most stable. Results of population structure suggested that populations of black-spotted pond frogs distributed in these four river basins were genetically independent. In particular, the population of the Nakdong River basin had the greatest genetic distance, indicating that it might have originated from an independent population. These results support the use of genetics in addition to designations strictly based on geographic stream areas to define the spatial scale of populations for management and conservation practices.

**Keywords:** Bayesian clustering analysis, Genetic diversity, Inbreeding probability, Microsatellite, Multivariate clustering analysis, Population differentiation

**Introduction**

The black-spotted pond frog (*Pelophylax nigromaculatus*) is widely distributed in East Asia (Korea, China, Japan, and Russia). It is a common species that is relatively easily observed in various freshwater habitats (e.g., ponds around rivers, lakes, swamps, rice paddies, and so on) in South Korea. Owing to several characteristics such as strong philopatry to breeding sites and habitats, low mobility, relatively easy sampling, and interbreeding in the laboratory, the black-spotted pond frog is a suitable model for studying patterns of population genetic structure (Beebee, 1996; 2005; Gong *et al.*, 2013). In previous study, black-spotted pond frog was used to understand variation patterns of genetic diversity and/or population structure of amphibians according to life history traits of species, the landscape of freshwater habitats, and the isolation on a time scale (Garcia *et al.*, 2017; Ma *et al.*, 2015; Wang *et al.*, 2014). Likewise, frogs can be used to study genetic variation patterns of amphibians distributed in South Korea.

In recent years, with the development of genetic analysis techniques, statistical processing methods, and computer information processing capabilities, access to population genetics analysis has become easier (Sunnucks, 2000).
For this reason, various molecular markers such as partial sequences of mitochondrial or chloroplast DNA, random amplified polymorphic DNA (RAPD), amplified fragment length polymorphic DNA (AFLP), and microsatellite markers have been used to determine the structure and genetic diversity of a population (Sunnucks, 2000). Among them, the microsatellite technique can quantify genotypes by identifying high levels of length polymorphisms in a serial repetition section of the nucleotide sequence (Zane et al., 2002). With this method, it is possible to obtain information on demographic trends through gene genealogies representing current and historical processes (Sunnucks, 2000). The identified genotype can be used to determine genetic diversity, population structure, origin of the population, and/or migration rate (Garris et al., 2005; Hatmaker et al., 2018; Luikart & England, 1999; Waraniak et al., 2019), which can be used to quantify the range of the population and to identify environmental factors that influence compositions of genes in the population.

This study was conducted to identify differences of genetic composition in the black-spotted pond frog in river basins of South Korea using the microsatellite technique. In freshwater ecosystems, rivers serve as corridors to connect groups of organisms in different habitats. They also act as barriers to separate groups depending on their types, sizes, and surrounding habitats (Finn et al., 2006; 2016; Lam et al., 2018; Mullen et al., 2010). The largest river basins in South Korea include Han River, Geum River, Yeongsan River, and Nakdong River basins due to various mountain ranges and catchment areas of rivers. The gene flow and/or population genetic structure of the black-spotted pond frog, which is a semi-aquatic frog, can be affected by the distribution of water resources. Therefore, we analyzed differences in genetic diversity and the population genetic structure of black-spotted pond frogs in four river basins.

Materials and Methods

Field investigation

A total of 95 black-spotted pond frogs were collected from paddy fields at three sites of the Han River basin (Yeoju-si, 27 individuals), three sites of the Geum River basin (Gongju-si, 23 individuals), three sites of the Yeongsan River basin (Naju-si, 20 individuals), and three sites of the Nakdong River basin (Chilgok-gun, 25 individuals) in September and October 2020 during the non-breeding season (Fig. 1). Frogs ranged from 1-year-old individuals that had not yet reached sexual maturity to individuals that were fully sexual mature. These collected individuals were euthanized by pithing and stored in 70% ethanol for later analysis.

Fig. 1. Sampling sites for black-spotted pond frogs (P. nigromaculatus) in four river basins of South Korea: (A) Han River, (B) Geum River, (C) Youngsan River, and (D) Nakdong River. Frogs were collected from three sites in each basin.
DNA extraction and genotype analysis

Muscle samples were collected from femurs of 95 black-spotted pond frogs and homogenized using liquid nitrogen. Genomic DNA was extracted using a DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany). The concentration and quality of genomic DNA were measured with a NanoDrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA). Using a total of seven primers (Rnh-1, Rnh-2, Rnh-3, Rnh-4, Rnh-6, Rnh-10, Rnh-12), the locus of each microsatellite was amplified according to previously reported procedures (Gong et al., 2013) (Table 1). A Seq-Studio Genetic Analyzer (Thermo Fisher-Applied Biosystems, Foster City, CA, USA) was used to visualize amplified genotypes. GeneMapper version 6.1 (Thermo Fisher-Applied Biosystems) was employed to evaluate genotype errors and the presence of a null allele. GENEPOP version 4.7 (Rousset, 2008) was used to determine deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium (LD) at the loci of seven microsatellites. There were no null alleles, significant deviations from Hardy-Weinberg equilibrium, or evidence of LD at the seven loci tested. In addition, by identifying the number of multiple genotypes, it was confirmed that these seven loci were sufficient to distinguish population differentiation from 95 frogs (Fig. 2).

Table 1. Information of seven microsatellite loci in black-spotted pond frogs

<table>
<thead>
<tr>
<th>Locus ID</th>
<th>Size (bp)</th>
<th>Repeat motif</th>
<th>Primer sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rnh-1</td>
<td>269–281</td>
<td>(TGC)_7</td>
<td>F: TGAAGTATTCTAGGTACAACAGGT R: GGCGGCAAAAGAGAGGGTT</td>
</tr>
<tr>
<td>Rnh-2</td>
<td>235–250</td>
<td>(TGC)_7TGTG(GTT)_7(TGC)_6</td>
<td>F: GCCTCCGGCTATAAATCAAACAACAGGT R: GCCGGCAACTACAG</td>
</tr>
<tr>
<td>Rnh-3</td>
<td>221–254</td>
<td>(AAC)_7</td>
<td>F: GCCTGABCTGGTGGAGGACA R: ATGGACATGCCTGGTGGTAGG</td>
</tr>
<tr>
<td>Rnh-4</td>
<td>154–166</td>
<td>(GA)_5</td>
<td>F: CGCTTACTAGGGGGGATATA R: GCCGAGAAGGGCTGGTAGCT</td>
</tr>
<tr>
<td>Rnh-6</td>
<td>208–214</td>
<td>(AAC)_5</td>
<td>F: TCTCAGGGAGGAGAATAGG R: AAGGGATCTGGAGATCTAAAGAC</td>
</tr>
<tr>
<td>Rnh-9</td>
<td>198–228</td>
<td>(GCA)_7</td>
<td>F: GCACAGTATAGGGGATAGG R: GCTACATGGCTGGTGATAT</td>
</tr>
<tr>
<td>Rnh-10</td>
<td>119–200</td>
<td>(GCT)_6</td>
<td>F: AGTGCAACATCAAATTTGGGTAAG G: GCAGAGTCCTGTCTGGGA</td>
</tr>
</tbody>
</table>

Fig. 2. Number of multi-locus genotypes by number of loci from 95 individuals of black-spotted pond frogs. Complete separation of 92 frogs was achieved with six loci.

Confirmation of genetic diversity and distance

Genetic diversity and population differentiation were quantified using GenAlEx version 6.5 (Peakall & Smouse, 2012) and Arlequin version 3.5 (Excoffier & Lischer, 2010). Shannon’s information index (I), molecular diversity (h), and inbreeding coefficient relative to the subpopulation (F_I) were calculated using mean number of alleles (N_A), effective number of alleles (N_E), observed heterozygosity (H_O), and expected heterozygosity (H_E). Three types of diversity indices (I, h, F_I) were used to compare genetic diversities of frog populations from four watershed areas. Population differentiation (F_ST) and p-value of F_ST were obtained using Arlequin version 3.5. They were used to express the genetic distance among populations of frogs collected from those four watersheds. In addition, Nei’s genetic distance was obtained using GenAlEx version 6.5 and expressed using an unweighted pair group method with an arithmetic mean (UPGMA) hierarchical tree form using Past 3 software (Hammer et al., 2001).

Analysis of population genetic structure

A discriminant analysis of principal components (DAPC)
was performed using the “adegenet” package (Jombart et al., 2010) in R to analyze clusters of population structure through multivariate clustering analysis. DAPC uses principal component analysis (PCA) to reduce the dimension of genetic variation. It then uses principal components generated to produce a linear combination of correlated alleles in linear discriminant analysis (Jombart et al., 2010; Waraniak et al., 2019). Through this analysis, population structures in those four watershed areas were identified using the discriminant function (DF). STRUCTURE version 2.3.4 (Pritchard et al., 2000) was used to conduct Bayesian clustering analysis to infer whether or not each individual received a portion of the allele from a potential population K. STRUCTURE analysis was performed with 100,000 burn-ins and 100,000 simulations. The ΔK method (Evanno et al., 2005) of STRUCTURE Harvester (Earl & Vonholdt, 2012) was used to determine the most suitable K value. Results from each of three independent STRUCTURE analyses revealed the most suitable cluster range of populations for separating frog populations collected from the four watersheds.

**Result**

**Comparison of genetic diversity**
For the seven microsatellite loci of 95 frogs, overall $N_A$, $N_E$, $H_O$, and $H_E$ were 3.893 (3.571-4.286, Han-Nakdong), 2.187 (1.965-2.489, Han-Yeongsan), 0.323 (0.259-0.398, Han-Nakdong), and 0.389 (0.336-0.422, Han-Geum), respectively (Table 2). The population of Han River had a lower genetic diversity ($I = 0.768$, $h = 0.397$) than the overall population ($I = 0.768$, $h = 0.397$), although it had a similar inbreeding coefficient ($F_{IS} = 0.191$) to the overall population ($F_{IS} = 0.179$). The population of Geum River had similar genetic diversity ($I = 0.785$, $h = 0.405$) and inbreeding coefficient ($F_{IS} = 0.181$) to the overall population. The population of Yeongsan River had higher genetic diversity ($I = 0.819$, $h = 0.412$) and inbreeding coefficient ($F_{IS} = 0.369$) than the overall population, while the population of Nakdong had higher genetic diversity ($I = 0.800$, $h = 0.430$) and lower inbreeding coefficient ($F_{IS} = -0.024$) than the overall population.

**Identification of genetic distance**
Paired $F_{ST}$ ranged from 0.031 (Han-Yeongsan) to 0.181 (Han-Nakdong) for populations of four watersheds (Fig. 3). All p-values in paired $F_{ST}$ from the population of four watersheds were significant ($p < 0.05$). The population of Han River was genetically similar to the population from the Geum River ($F_{ST} = 0.041$) and Yeongsan River ($F_{ST} = 0.031$). Likewise, the population of Geum River was genetically similar to the population from Yeongsan River ($F_{ST} = 0.048$). On the other hand, paired $F_{ST}$ values of the population in Nakdong River with the other three populations were high (Han River, 0.181; Geum River, 0.087; and Yeongsan, 0.143). Therefore, the population of Nakdong River is likely to be genetically distinct from populations of other rivers.

**Analysis of the population genetic structure**
In DAPC, the discriminant function 1 (DF1) explained 60.79% of the genetic variation in 95 black-spotted pond frogs sampled and DF2 explained 25.80% of the genetic variation. The DF1 value showed that populations of frogs collected from the Han River, Geum River, and Yeongsan River were of the same group while populations collected from the Nakdong River were distinct from other populations (Fig. 4A). Using 100% of the explanatory power of all

### Table 2. Genetic diversity of seven nuclear microsatellite loci estimated for four populations of 95 black-spotted pond frogs

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>$N_A$</th>
<th>$N_E$</th>
<th>$H_O$</th>
<th>$H_E$</th>
<th>$I$</th>
<th>$F_{IS}$</th>
<th>$h$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Han</td>
<td>27</td>
<td>3.571</td>
<td>1.965</td>
<td>0.259</td>
<td>0.336</td>
<td>0.670</td>
<td>0.191</td>
<td>0.342</td>
</tr>
<tr>
<td>Geum</td>
<td>23</td>
<td>3.714</td>
<td>2.059</td>
<td>0.349</td>
<td>0.422</td>
<td>0.785</td>
<td>0.181</td>
<td>0.405</td>
</tr>
<tr>
<td>Youngsan</td>
<td>20</td>
<td>4.000</td>
<td>2.489</td>
<td>0.286</td>
<td>0.402</td>
<td>0.819</td>
<td>0.369</td>
<td>0.412</td>
</tr>
<tr>
<td>Nakdong</td>
<td>25</td>
<td>4.286</td>
<td>2.236</td>
<td>0.398</td>
<td>0.397</td>
<td>0.800</td>
<td>-0.024</td>
<td>0.430</td>
</tr>
<tr>
<td>Overall</td>
<td>3.893</td>
<td>2.187</td>
<td>0.323</td>
<td>0.389</td>
<td>0.768</td>
<td>0.179</td>
<td>0.397</td>
<td></td>
</tr>
</tbody>
</table>

$N$, samples size; $N_A$, mean number of alleles; $N_E$, effective number of alleles; $H_O$, observed heterozygosity; $H_E$, expected heterozygosity; $I$, Shannon’s information index; $F_{IS}$, inbreeding coefficient relative to the subpopulation; $h$, molecular diversity.
discriminant functions of DAPC, the fine-scale genetic structure indicated that frogs from four watershed areas could be separated into independent populations according to their genetic variance (Fig. 4B).

In the STRUCTURE harvester analysis, the most suitable K value for distinguishing these frog populations was 3 (Fig. 5). The UPGMA graph of Nei’s genetic distance showed that populations from Han River and Geum River were genetically similar while the population from Nakdong River was genetically distinct from the other three populations (Fig. 6A). STRUCTURE analysis also showed that the population of Nakdong River was a genetically distinct K population from populations of Han, Geum, and Yeongsan Rivers. On the other hand, populations of Han, Geum, and Yeongsan Rivers appeared to have the same origin in the STRUCTURE clustering analysis (Fig. 6B).

Discussion

In this study, the genetic diversity and population structure of black-spotted pond frogs collected in four watershed areas in Korea were compared using seven microsatellite loci. The population with the lowest genetic diversity (I and h) and the highest inbreeding coefficient ($F_{IS}$) was the population collected from the Han River. The population of the Nakdong River showed opposite results. In the population of Geum River, genetic diversity and inbreeding coefficient were similar to those of the entire population. The population of Yeongsan River had both high genetic diversity and inbreeding coefficient. Population differentiation ($F_{ST}$), DAPC, and Nei’s genetic distance showed that these four populations were distinct. In particular, it implied that the population of the Nakdong River was genetically separated from population from the other three river basins (Han River, Geum River, and Yeongsan River). Bayesian cluster structure analysis (STRUCTURE) also showed that the population of Nakdong River was likely to have originated from other populations than the three populations that appeared to have a common genetic origin.

Like biodiversity, the high genetic diversity of species can increase resistance and resilience of organisms to environmental changes or disturbances, thereby serving as a biological insurance that guarantees the stability of the species over time and the ecosystem in turn (Hughes & Stachowicz, 2004; Spielman et al., 2004). Genetic diversity can change due to habitat disturbance, population size, and/or population connectivity. In general, small populations living in disturbed or isolated habitats often tend to show decreased genetic diversity with increased inbreeding and less gene flow (Banks et al., 2013; Young et al., 1996). The negative pattern of genetic diversity in the population from Han River suggests that this population might have experienced a habitat disturbance or population decline more recently compared to other populations. On the other hand, the population of Geum River had an average pattern of genetic diversity, indicating that this population was more stable than the population of Han River. Inbreeding is generally known to reduce genetic diversity by causing loss of alleles (Willoughby et al., 2015). However, the population of Yeongsan River showed both high genetic diversity and inbreeding coefficient. This result might be because the Yeongsan River population is relatively large, but partially isolated. The population of Nakdong River showed a positive pattern of genetic diversity through high genetic diversity and low inbreeding probability, indicating that its genetic exchange might be active and that the population maintained its stability.

Analyses for $F_{ST}$, DAPC, and Nei’s genetic distance showed that populations of frogs collected from four watersheds were genetically separated. Among them, population of the Nakdong River was found to be more clearly separated from populations of Han River, Geum River, and Yeongsan River. Furthermore, STRUCTURE analysis showed that the population of the Nakdong River was more likely to have a different genetic origin than populations of Han, Geum, and Yeongsan rivers. This pattern might be because Sobaek Mountains, including Jirisan (1,915 m), have restricted population distribution (Fig. 1). In fact, when the domestic territory is geographically divided,

Fig. 3. Heatmap of paired $F_{ST}$ values among black-spotted pond frogs from four populations. The number on the right represents paired $F_{ST}$ value. All paired $F_{ST}$ values were significant ($p < 0.05$).
The most suitable $K$ of black-spotted pond frogs obtained by the $\Delta K$ method with STRUCTURE Harvester. The value of $\Delta K$ was the highest at three ($\Delta K: 25.252$).

Fig. 4. Population genetic structure by multivariate clustering methods (discriminant analysis of principal components, DAPC): (A) Discriminant function 1 (DF1) explained 60.79% of the genetic variation in black-spotted pond frogs from four river basins (H: Han river, G: Geum river, Y: Youngsan river, N: Nakdong river), and (B) DF2 explained 25.80% of the genetic variation in these frogs. Fine-scale structures represent the probability that frogs are assigned to each group.

Fig. 5. The most suitable $K$ of black-spotted pond frogs obtained by the $\Delta K$ (delta $K$) method with STRUCTURE Harvester. The value of delta $K$ was the highest at three ($\Delta K: 25.252$).
Fig. 6. Results for population clustering by Nei’s genetic distance and Bayesian clustering algorithms: (A) UPGMA (Euclidean) tree of genetic distance among black-spotted pond frogs in four river basins, (B) Population genetic structure obtained by Bayesian clustering algorithms (STRUCTURE analysis) among black-spotted pond frogs in four river basins.

Species. Population genetic analysis of a single locus of relatively well-preserved mitochondrial DNA is commonly used for many taxa. However, it has lower variability than microsatellite analysis which analyzes the composition of a gene in a population using multiple loci (Sunnucks, 2000). Ecological characteristics of the black-spotted pond frog, a semi-aquatic frog that has high philopatry with gene flow depending on water source are different from those of the Japanese tree frog, a terrestrial amphibian that can alternately move between its habitat and the breeding ground. Thus, the difference observed might be due to different population genetic structures. In the future, it may be necessary to compare population genetic patterns among various species with different ecological characteristics using genetic markers of the same or similar variability to identify differences and changes in the population structure according to natural and artificial factors. Identifying the origin and extent of a population in an artificially designated watersheds might help us define the spatial scale to protect and manage the population. In addition, this may be applicable to conservation genetic studies that identify factors of population reduction by comparing composition patterns of genes in common species with phylogenetically adjacent endangered amphibians.

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2018R1C1B3008645).

References


Conflict of Interest

The authors declare that they have no competing interests.
the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology, 14, 2611–2620. doi:10.1111/j.1365-294X.2005.02553.x


Waraniak, J.M., Fisher, J.D., Purcell, K., Mushet, D.M., and Stockwell, C.A. (2019). Landscape genetics reveal broad and fine-scale population structure due to landsc-
ape features and climate history in the northern leopard frog (*Rana pipiens*) in North Dakota. Ecology and Evolution, 9, 1041-1060. doi:10.1002/ece3.4745


