Thesis for the Degree of Doctor of Philosophy

Population genetic structure of genus *Semisulcospira* using genetic markers

by

Yeon Jung Park

Department of Microbiology

The Graduate School

Pukyong National University

February 2020

Population genetic structure of genus *Semisulcospira* using genetic markers

유전자 마커를 이용한 다슬기류의 집단 유전학적 구조 분석

Advisor: Prof. Tae-Jin Choi

by

Yeon Jung Park

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the Department of Microbiology, Graduate School,

Pukyong National University

February 2020

Population genetic structure of genus *Semisulcospira* using genetic markers

A dissertation

by

Yeon Jung Park

Approved by

Chairman : Myung-Suk Lee

Member : Gun-Do Kim

Member : Jung-Ha Kang

Member : Eun-Mi Kim

Member : Tae-Jin Choi

February 2020

CONTENTS

List of Tables

Chapter II.

Chapter III.

Chapter IV.

Chapter V.

List of Figures

Chapter II.

Chapter III.

Chapter IV.

Chapter V.

유전자 마커를 이용한 다슬기류의 집단 유전학적 구조 분석

박 연 정

부경대학교 대학원 미생물학과

요 약

생물다양성협약은 자국내 서식하는 생물 자원에 대한 주권적 권리를 인정하며, 자국 생물종의 자세한 목록 제시와 감시 체계 구축을 의무화 하고 있다. 이에 따라 세계 각국은 국가적 생물 다양성 목록 작성과 생물 다양성 변화 추이의 감시, 생물 자원에 대한 접근과 유출 관리, 유전자원에 대한 접근과 이익의 공유에 대한 시스템을 강화하고 있다. 우리나라에서 현재까지 보고된 생물 종 정보에 의하면 유사 환경을 지닌 일본 및 영국에 비해 생물 종 다양성 정도가 낮게 파악된다고 보고하고 있다. 이를 개선하기 위해 환경부에서 한반도 고유 생물을 지정하고 이들의 국외 반출 승인 제도를 운영하고 있으며, 해양수산부에서도 고유종을 보호대상해양생물로 지정하여 보호하고 있다. 이들의 과학적, 전문적 관리는 국내 생물 자원의 지속적 보호를 목적으로 한다. 고유종 (endemic species)이란 지리적으로 한정된 지역에만 자연적으로 서식하는 생물분류군을 의미한다. 이들은 어느 특정 지역 내에서만 자연발생적으로 서식하기 때문에, 개체군의 크기와 분포 범위가 작아 환경 변화에 취약하며, 외래종과의 경쟁에서 이기지 못하는 경우가 허다하고, 유전적 교란이 쉽게 일어나기 때문에 위기 종으로 분류 및 보호하고 있으며, 고유종의 보전을 위해 범국가적인 보호 및 관리를 수행하는 국제기구인 ‘세계자연보전연맹 (International Union of Conservation of Nature (IUCN)’ 에서 고유종 밀집지역에 대한 조사 및 연구를 진행하고 있다. 또한, 자생생물에 대한 연구가 체계적으로 수행되고 있는 국가에서는 다양한 생물자원 요소 중 특히 고유종의 관리와 보존을 우선순위에 두고 있다. 이에 대해 우리나라에서도 국내에 서식하는 고유종의 분포 범위 및 서식 환경 등에 관해 정보를 정리한 ‘한반도 고유종 총람’을 발간하였다. 그러나 고유종의 보전을 위해서는 국가 단위로 특정 지역의 고유종 목록을 단순 제시하기 보다는 종의 위협 상황에

대한 모니터링과 희귀성에 대한 유형 분석, 멸종 위기와 관련된 유전학적 집단 분석과 같은 역동적인 실태 조사가 필요하다. 하지만 고유종에 대한 중요성의 인식과 관리 방안의 구축이 주요한 국가적 관심사로 부각됨에도 불구하고, 아직 우리나라는 고유종에 대한 연구가 매우 미흡하다.

한반도 고유 생물 중 담수산 연체 동물에는 다슬기과가 3속 5종이 있으며 (*Semisulcospira coreana, S. forticosta, S. tegulata, Koreanomelania nodifila, Koreoleptoxis globus ovalis*), 이들 중 다슬기속 (*genus Semisulcospira*) 에는 참다슬기 (*S. coreana*), 좀주름다슬기 (*S. tegulata*), 주름다슬기 (*S. forticosta*) 가 한반도 고유종으로 지정되어있고, 한반도와 중국에 서식하는 곳체다슬기 (*S. gottschei*) 는 동아시아 고유종으로 지정 되어있다. 다슬기는 건강 식품으로 인식되어 경제적 가치가 매우 높은 담수 종이다. 하지만 환경 변화 및 오염, 서식지 감소 및 남획 등으로 인해 다슬기 자원이 감소하고 있는 상황에서 다슬기의 국내 어획량은 수요량에 비하여 제한적이므로 이를 보완하기 위해 종묘 방류 및 수입 의존도가 증가하고있는 실정이다. 국내 서식하는 다슬기 중 주요 경제성 종으로 참다슬기와 곳체다슬기가 있으며, 특히 국내 수요를 충족하기 위하여 중국에서 곳체다슬기가 다량 수입되고 있다. 이처럼 다슬기류는 생물 자원 및 생물지리학상 연구 가치가 매우 높은 분류군이지만, 다슬기류에 대한 연구는 번식 및 생태학적 연구, 식품 영양학적 연구가 주로 수행되었다. 또한 형태 및 유전학적 계통 분류에 대한 연구가 수행되었으나 여전히 분류학상 혼란이 있으며, 고유종의 자원 관리 및 보전을 위한 유전적 다양성 분석, 개체군 간의 유전적 분화도 및 유연 관계 등을 포함하는 집단 유전학적 연구는 미비하다.

본 연구에서는 집단 유전학 연구에 널리 사용되고 있는 분자 마커인 미토콘드리아 DNA (mtDNA)와 초위성체 (microsatellite, MS) 마커를 활용하여 국내 서식하는 다슬기류의 유전적 다양성 및 집단 구조와 집단 사이의 유전적 분화 등의 집단 유전학 연구를 수행하고, 우리나라와 중국에 서식하는 동아시아 고유종인 곳체

다슬기의 집단 유전학적 분석을 통해 국내산과 중국산을 구별할 수 있는 MS 마커의 활용 방안을 제시하였다. 본 연구에서 수행된 결과를 바탕으로 한반도 고유종인 다슬기류의 보전 및 관리 방안을 제시하고자 한다.

국내 서식하는 다슬기의 계통 분류를 위해 universal primer와 species specific primer를 활용하여 COI 과 ND4 유전자의 염기서열을 각각 확보하였다. 이를 바탕으로 COI 유전자에서의 종간 유전적 변이 정도를 확인한 결과, 참다슬기와 좀주름다슬기가 유전적으로 가장 가깝게 나타났으며(1.5%), 주름다슬기와 다슬기가 가장 멀게 나타났다(8.9%). 이들의 종내 유전적 변이는 0.5-0.8%로 종간 유전적 변이 보다 낮게 나타났다. 또한 종 분류가 명확하지 않은 *Semisulcospira sp.* 2그룹이 관찰되었으며, 이들의 유전적 거리는 종내 유전적 변이보다 높으므로 잠재종 (cryptic species)으로 사료된다. 또한 ND4 유전자를 활용한 종간 유전적 변이율은 COI 유전자보다 약간 높게 관찰되었다. 이들 유전자 영역 (COI, ND4, COI+ND4)을 활용하여 *Semisulcospira sp.* 2그룹을 제외한 형태와 유전학적 분류가 일치하는 종을 대상으로 한반도 및 아시아 고유종인 다슬기류 4종의 유전적 다양성을 분석하였다. 4종 모두에서 높은 haplotype diversity와 낮은 nucleotide diversity가 나타났으며, 중립성 검사 (neutrality test)에서 모두 음의 값을 나타내어, 최근에 인구 팽창이 이루어 지고 있는 것으로 확인 되었다. 또한 좀주름다슬기, 참다슬기와 곳체다슬기는 분석된 모든 집단에서 동일한 main haplotype이 나타나고, 일부 집단 별 rare haplotype이 보였다. 반대로 주름다슬기는 집단 간 rare haplotype을 가지는 형태를 보였으며, 이는 집단간 분화 정도를 나타내는 F_{ST} 결과와 동일하게 나타났다. 따라서 주름다슬기는 국내 지역별로 유전적 분화가 나타났으며, 반면에 좀주름다슬기, 참다슬기, 곳체다슬기는 지역간 유전적 차이가 일부 지역 사이에만 나타남을 확인 할 수 있었다.

다슬기류 4종 중 경제적 가치가 높은 참다슬기와 곳체다슬기 2종의 집단 내 유전적 다양성, 집단간 유전적 구조 및 분화도를 알아보기 위하여 참다슬기 유래 MS 마커 18개를 개발하였다. 이들의 평균 대립유전자 수는 6.9개이며, 평균 다형성

지수(PIC)는 0.626으로 높게 나타났다. 또한 hetrozygosity의 기대치 (H_e)와 관측치 (H_o)는 각각 0.679와 0.434를 나타내었다. 또한 다른 4종의 다슬기류에 교차 시험을 한 결과 80-85%의 높은 transferability를 나타내었다. 본 연구에서 개발된 참다슬기 유래 MS 마커 중 9개를 선별하여 국내 서식하는 참다슬기와 곳체다슬기의 집단 유전학적 분석을 수행하였다. 두 종 모두 평균 H_e 가 H_o 보다 모두 높게 나타났으며, 근친교배계수를 나타내는 F_{IS} 역시 양의 값을 나타내었다. 참다슬기 10개 집단은 Pairwise F_{ST}, NJ and PCA 분석, STRUCTURE 분석에서 2개의 그룹으로 나뉘는 것을 확인 하였으며, 이 결과는 mtDNA 결과와 동일하였다. 곳체다슬기의 경우 Pairwise F_{ST} 와 STRUCTURE 분석 결과 국내 집단의 유전적 흐름 (gene flow)이 활발하여 지역간에 유전적 균일화가 진행된것으로 사료된다. 또한 MS 마커를 이용한 곳체다슬기의 집단간 유전적 분화도가 mtDNA를 이용한 결과보다 낮게 나타났는데 이는 mtDNA만 단독으로 분석하였을 땐 집단간 유의적인 차이가 있을 것으로 생각했으나 MS를 분석 시 지역간 gene flow가 활발하게 일어나 매우 단순한 유전자형을 구성하여 지역에 따른 유전적 분화가 없는 것을 확인 할 수 있었다.

국내산 다슬기를 보호하고 불법 행위를 근절시키며 안정성을 보장하기 위하여 국내산과 중국산 곳체다슬기의 원산지 판별에 mtDNA와 MS마커를 응용하였다. COI 유전자를 이용하여 총 43개의 haplotype을 얻었으나, 이들에서 국내산과 중국산을 구별할 수 있는 특이 SNP은 발견 되지 않았다. 9개의 MS 마커의 대립유전자형을 이용한 결과 곳체다슬기 11개 집단은 96.9%의 정확도로 국내산과 중국산으로 구별되었다. 이는 국내산과 중국산 다슬기의 구별이 필요한 여러 분야에 활용이 가능할 것이라 사료된다.

이와 같은 결과는 생물의 진화 및 역사를 이해하기위해 mtDNA는 활용 가능하지만 집단 간 유전적 구조 분석의 경우 모계 유전 및 유효 집단의 크기에 따른 한계가 있으므로, 이를 보완 하기 위해 핵 DNA인 MS마커를 활용 함으로써 집단간 유전적 분화도 및 원산지 판별에 활용 가능할 것이라 생각된다. 기존의 연구들을 보면, 많은

멸종 위기종 및 고유종은 수계별로 분화가 진행되어 있으므로, 독립 수계별로 관리 방안이 필요하다고 보고되었다. 이러한 양상은 주름다슬기와 좀주름다슬기, 참다슬기의 일부 집단에서도 확인되었다. 따라서 한반도 고유종인 다슬기류의 보전 및 관리를 위해서는 유전적으로 차이가 나타나는 그룹 간의 무분별한 방류는 반드시 지양되어야 하며 지역 집단 별로 보존 대책이 필요하다고 판단된다.

본 연구 결과는 유전자 마커를 활용하여 한반도 고유종인 다슬기류의 유전적 특성을 규명한 것으로써 국내 고유종에 대한 자원 관리 및 보전을 위한 구체적인 기초 자료 및 나고야의정서에 대응하기 위한 우리나라 고유종의 생물 주권 확보를 위한 과학적 근거로도 활용 가능할 것이라 생각된다.

Chapter I.

General Introduction

General Introduction

1. Background of this study

Based on the Convention on Biological Diversity, countries are creating national biodiversity lists, monitoring biodiversity, increasing access to biological resources, managing leakage, and sharing genetic resources. The biodiversity in Korea is lower than in countries with similar environments, such as Japan and the United Kingdom. To address this, the Ministry of the Environment of Korea has identified endemic Korean Peninsula organisms and implemented an approval system that must be satisfied before they can leave the country. The purpose of this project is to promote the protection of domestic biological resources through scientific management (www.me.go.kr).

Endemism is an ecological phenomenon in which species occur only in specific geological locations, such as confined areas, or on specific islands or in certain countries. Because endemic species only inhabit certain regions, they tend to be protected species because of their typically small population sizes and distribution ranges, and are vulnerable to environmental changes. They are often less competitive than invasive species, and genetic confusion is common (Bouzat 2010). Therefore, countries often prioritize the management and preservation of endemic species (Kruckerberg and Rabinowitz 1985). The loss of endemic species does not simply represent the loss of a national biological resource, but rather the total extinction of that species. The preservation of endemic species requires targeted actions to protect and manage them, both at the country and international level. The International Union of Nations of Nature (IUCN) established the Specialties Survival Commission and created 250 Regional

Endemic Centers to investigate and manage endemic species worldwide. We collected information on species endemic to the Korean Peninsula and published 'Endemic Species of Korea' (https://www.nibr.go.kr/). In Korea, 2,177 endemic species were identified, including 73 mollusks. Except for *Corbicula papyracea*, the genetic diversity and population structure of these mollusks has not been analyzed (Huh 1998). Although the management of endemic species has emerged as a major national concern, research on endemic species in Korea is still lacking.

As one of the Korean endemic species, the family Cerithiodiea includes three genera and five species of freshwater Korean mollusk: *Semisulcospira coreana*, *S. forticosta*, *S. tegulata*, *Koreanomelania nodifila*, and *Koreoleptoxis globus ovalis*. While *S. gottschei* is not endemic to the Korean Peninsula, it inhabits only East Asia (www.me.go.kr). This snail family is the second-most important for inland fisheries, behind only the family Viviparidae (http://kostat.go.). Freshwater snails are economically important species that are valued in Korea as a source of food and medicine. *Semisulcospira* has also become a candidate for inland aquaculture in Korea, because of the decline in wild production due to overexploitation (Pham et al, 2010). The current domestic consumption is about 2,000 tons per year, and the wild stock catch is 700–1,000 tons. To meet the demand, over 1,000 tons of freshwater snails are imported from China and North Korea in processed form (Moon et al, 2015). Freshwater snails also play important roles as biological markers of the status of freshwater environments (Lee and Lim 2005). However, despite their economic and ecological importance, there have been few phylogenetic, phenotypic, breeding, or nutritional analyses of *Semisulcospira* (Urabae 2000; Zeng et al, 2015; Kim et al, 2010; Moon et al, 2015).

The phylogenetic position of Cerithiodiea in Gastropoda is unclear (Colgan et

al, 2007). Miyanaga (1942) reported that the genus *Semisulcospira* consisted of six nominal species and three subspecies, while Kwon and Habe (1979) and Kwon (1990) reported two genera: *Semisulcospira* and *Koreanomelania*. Burch et al (1987) proposed a third genus, *Koreoptoxis*, based on shell and reproductive characters. More recently, Ko et al (2001) identified seven species in the three genera by analyzing external radulae. However, the snails undergo marked changes in shape and color depending on environmental factors, such as sediment and water flow rate. Recently, Köhler et al (2017) noted considerable discrepancies in terms of the delineation of species based on the branching patterns of the mitochondrial DNA (mtDNA) tree versus morphological features. In addition, Kim et al (2010) reported that the clade *Semisulcospira* was clearly separated from *Koreoptoxis* and *Koreanomelania* based on the mtDNA COI gene sequence; they suggested the existence of cryptic species. Therefore, it is necessary to reconstruct phylogenetic relationships to identify species that are genetically identical and check the degree of concordance with analyses based on morphological characters.

Population genetics is important for managing endangered and endemic species, because it provides information on genetic diversity and population structure (Jang et al, 2017). Molecular markers are widely used to assess the evolution and patterns of genetic diversity of fish and shellfish (Liu and Cordes 2004). Because the tests are rapid, highly specific, and precise (Yu et al, 2012). Among the various molecular markers, microsatellites (MS) and mtDNA can be used to analyze the genetic diversity and population structure of Korean melania snails.

mtDNA represents a tiny part of an organism's genome and has been used as a marker of molecular diversity for 30 years (Galtier et al, 2009). Its advantages as

a molecular marker include the fact that it is strongly conserved across organisms, with very few duplications, no introns, and very short intergenic regions; moreover, it is easily amplified because there are multiple copies in a single cell (Gissi et al, 2008). The cytochrome oxidase subunit 1 (COI), a mitochondrial fragment, recently became The Consortium for the Barcode of Life (CBOL) standard for molecular taxonomy and identification (Ratnashingham and Hebert 2007). In addition, among the mitochondrial genes, NADH dehydrogenase subunit 4 (ND4) is an excellent marker for analyzing the genetic population structure of fish and shellfish (Doosey et al, 2010). This gene outperforms other commonly used mtDNA genes, such as COI, cytochrome b, and 12S/16S rRNA, in phylogenetic analysis because of its relatively long sequence and phylogenetically informative variation at the first and second codon positions (Miya et al, 2006). To resolve taxonomical confusion, some studies have used more than two mitochondrial genes (Zheng et al, 2016).

MS are single sequence repeats (SSRs) or short tandem repeats that have been used widely as indicators of genetic variation between fishery populations (Yu et al, 2012; Kang et al, 2014). MS are codominant, ubiquitous and abundant, and highly polymorphic in prokaryotic and eukaryotic genomes (Li et al, 2002). These characteristics have led to their wide application in many areas of biology, including genome mapping, detection of pedigrees, population genetics, conservation, and forensic studies (Jarne et al, 1996; Schuler et al, 1996; Luikart et al, 2003). The traditional cloning methods used to develop MS involve considerable trial and error, including DNA sequencing of flanking regions, and are time-consuming, cost-prohibitive, and have low throughput (Queller et al, 1993). These disadvantages were partly resolved with the advent of next-

generation sequencing (NGS), a cost- and time-effective method that produces millions of base pairs of short fragment reads (Cai et al, 2009). The highly polymorphic markers developed with NGS have been used to examine population genetic structure, including in terms of intraspecific population differentiation and gene flow (Ball and Chapman 2003). However, no population-level study using MS markers has been performed to clarify the genetic diversity and population structure of the genus *Semisulcospira*. To obtain an in-depth understanding of the genetic diversity of *Semisulcospira*, its population structure should be analyzed using mtDNA and MS.

2. Purpose of this study

This study sought to clarify the phylogenetic relationships of *Semisulcospira*, to facilitate management of the endemic Korean species and investigate the genetic variation within and among populations inhabiting Korea, using mtDNA and MS markers to elucidate their biological characteristics.

(1) In chapter 2, the mitochondrial DNA and MS markers developed for phylogenetic and genetic diversity analyses of the genus *Semisulcospira*

(2) In chapter 3, gene flow and population structure of Korean Peninsula endemic species in the genus *Semisulcospira* were analyzed using mitochondrial DNA

(3) In chapter 4, the genetic diversity and population structure of the most economically valuable molluscs, *S. coreana* and *S. gottschei*, were studied using microsatellite markers.

(4) In chapter 5, The population genetic structure of *S. gottschei* was studied in Korean and Chinese populations and used to determine the country of origin.

The result of this study was identified the genetic characteristics of melania snail, an endemic species to the Korean Peninsula, using genetic markers, and was thought to be available as a concrete basic data for resource management and conservation for endemic species.

3. Reference

Colgan D.J., Ponder W.F., Beacham E., Macaranas J. (2007) Molecular phylogenetics of Caenogastropoda (Gastropoda: Mollusca). Mol. Phylogenet. Evol, 42(3), 717-737.

Miyanara M. (1942) Kwwanina snails of the genus Semisulcospira from Korea. Chosen Hakubutsu Zasshi. 9(36), 114-130.

Kwon O.K. and Habe T. (1979) A list of non-marine molluscan fauna of Korea. Korean Journal of Malacology. 12, 24-31.

Kwon O.K. (1990) Illustrated encyclopedia of fauna and flora of Korea Vol. 32. Mollusca (I). Ministry of Education. Seoul, Korea.

Burch J.B., Chung P.R., Jung Y. (1987) A guide to the fresh water snails of Korea. Walkerana. 2, 195-232.

Ko J.H., Lee J.S., Kwon O.K. (2001) Study on radulae of seven species of the Family Pleuroceridae in Korea. The Korean Journal of Malacology. 17, 105-115.

Kim W.J., Kim D.H., Lee J.S., Bang I.C., Lee W.O., Jung H.T. (2010) Systematic Relationships of Korean Freshwater Snails of *Semisulcospira*, *Koreanomelania*, and *Koreoleptoxis* (Cerithiodiea; Pleuroceridae) revealed by Mitochondrial Cytochrome Oxidase I Sequences. Korean. J. Malacol. 26(4), 275-283.

Davis G.M. (1969) A taxonomic study of some species of Semisulcospira in Japan (Mesogastropoda: Pleuroceridae). Malacologia. 7, 211-294.

Hsu K.C., Bor H., Lin H.D., Kuo P.H., Tan M.S., Chiu Y.W. (2014) Mitochondrial DNA phylogeography of *Semisulcospira libertina* (Gastropoda: Cerithioidea: Pleuroceridae): implications the history of landform changes in Taiwan. Mol. Biol. Rep. 41(6), 3733-3743.

Urabe M. (1992) A discrimination and morphological comparison of two snail species of the

genus *Semisulcospira* in a single river. Venus. 50, 270-286.

Pham M.A., Hwang G.D., Kim Y.O., Seo J.Y., Lee S.M. (2010) Soybean meal and wheat flour, proper dietary protein sources for optimal growth of snail (*Semisulcospira coreana*). Aquacult Int. 18(5), 883-895.

Moon S.K., Kim I.S., Lim C.W., Yoon N.Y., Jeong B.Y. (2015) Proximate and Fatty Acid Compositions of Commercial Domestic and Imported Melania Snails *Semisulscospira* sp. Korean. J. Fish. Aquat. Sci., 48, 977-981.

Rasmussen R.S., Morrissey M.T. (2008) DNA-based methods for the identification of commercial fish and seafood species. Comprehensive reviews in food science and food safety, 7, 280-295.

Lee S.M., Lim T.J. (2005) Effects of dietary protein and energy levels on growth and lipid composition of juvenile snail (*Semisulcospira gottschei*). Journal of Shellfish Research. 24, 99-102.

URABE M. (2000) Phenotypic modulation by the substratum of shell sculpture in *Semisulcospira reiniana* (Prosobranchia: Pleuroceridae). J. Mol. Stud. 66, 53-60.

Zeng T., Yin W., Xia R., Fu C., Jin B. (2015) Complete mitochondrial genome of a freshwater snail, *Semisulcospira libertina* (Cerithioidea: Semisulcospiridae). Mitochondrial DNA. 26, 897-898.

Liu Z.J. and Cordes J.F. (2004) DNA marker technologies and their applications in aquaculture genetics. Aquaculture, 238, 1-37.

Yu J.N., Azuma N., Abe S. (2012). Genetic differentiation between collections of hatchery and wild masu salmon (*Oncorhynchus masou*) inferred from mitochondrial and microsatellite DNA analyses. Environ. Biol. of Fish. 94, 259-271.

Galtier N., Nabholz B., Glémin S., Hurst G.D.D. (2009) Mitochondrial DNA as a marker of molecular diversity: a reappraisal. Molecular ecology. 18(22), 4541-4550.

Gissi C., Iannelli F., Pesole G. (2008) Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity. 101(4), 301-320.

Ratnasingham S., Hebert P.D. (2007) bold: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes. 7, 355-364.

Doosey M.H., Bart Jr H.L., Saitoh K., Miya M. (2010) Phylogenetic relationships of catostomid fishes (Actinopterygii: Cypriniformes) based on mitochondrial ND4/ND5 gene sequences. Mol. Phylogenet. Evol. 54(3), 1028-1034.

Miya M., Saitoh K., Wood R., Nishida, M., Mayden R.L. (2006) New primers for amplifying and sequencing the mitochondrial ND4/ND5 gene region of the Cypriniformes (Actinopterygii: Ostariophysi). Ichthyol. Res. 53(1), 75-81.

Zheng L.P., Yang J.X., Chen X.Y. (2016) Molecular phylogeny and systematics of the Barbinae (Teleostei: Cyprinidae) in China inferred from mitochondrial DNA sequences. Biochem. Syst. Ecol. 68, 250-259.

Kang J.H., Kang H.S., Noh E.S., Park J.Y., An C.M. (2014) Isolation and characterization of novel microsatellite markers for the northern mauxia shrimp, *Acetes chinensis*, using pyrosequencing. Marine genomics, 18, 67-69.

Li Y.C., Korol A.B., Fahima T., Beiles A., Nevo E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol. Ecol. 11(12), 2453-2465.

Jarne P., and Lagoda P.J. (1996) Microsatellites, from molecules to populations and back. Trends. Ecol. Evol. 11(10), 424-429.

Schuler G.D., Boguski M.S., Stewart E.A., Stein L.D., Gyapay G., Rice K., et al. (1996) A gene map of the human genome. Science, 274(5287), 540-546.

Luikart G., England P.R., Tallmon D., Jordan S., Taberlet P. (2003) The power and promise of population genomics: from genotyping to genome typing. Nat. Rev. Genet., 4(12), 981-994. Schlötterer C. (2000) Evolutionary dynamics of microsatellite DNA. Chromosoma,

109(6), 365-371.

Queller D.C., Strassmann J.E., Hughes C.R. (1993) Microsatellites and kinship. Trends. Ecol. Evol. 8(8), 285-288.

Cai Z., Li G., Lin C., Shi T., Zhai L., Chen Y., Huang G. (2013) Identifying pathogenicity genes in the rubber tree anthracnose fungus *Colletotrichum gloeosporioides* through random insertional mutagenesis. Microbiol. Res. 168(6), 340-350.

Geist J., Rottmann O., Schröder W., Kühn R. (2003) Development of microsatellite markers for the endangered freshwater pearl mussel *Margaritifera margaritifera* L. (Bivalvia: Unionoidea). Mol. Ecol.Notes. 3(3), 444-446.

Köhler F. (2017) Against the odds of unusual mtDNA inheritance, introgressive hybridisation and phenotypic plasticity: systematic revision of Korean freshwater gastropods (Semisulcospiridae, Cerithioidea). Invertebr. Syst. 31(3), 249-268.

Lee T., Hong H.C., Kim J.J., Foighil D.Ó. (2007) Phylogenetic and taxonomic incongruence involving nuclear and mitochondrial markers in Korean populations of the freshwater snail genus *Semisulcospira* (Cerithioidea: Pleuroceridae). Mol. Phylogenet. Evol. 43(2), 386-397.

Ball A.O. and Chapman R.W. (2003) Population genetic analysis of white shrimp, *Litopenaeus setiferus*, using microsatellite genetic markers. Molecular ecology. 12(9), 2319-2330.

Chapter II.

Marker development for

mitochondrial DNA and microsatellite

Marker development for mitochondrial DNA and microsatellite

1. Abstract

Korean freshwater snails of the genus *Semisulcospira* are widely distributed across East Asia. It has been a very popular nutritional food in Korea, and is ecologically important a water quality indicator because it lives only in clean water. However, population genetic diversity study of this genus using molecular markers was extremely limited. Various molecular markers are used to overcome the limitations of morphological analysis methods. Common genetic markers in fish and shellfish include allozymes, mitochondrial DNA, RFLP, RAPD, AFLP, microsatellite, SNP, and EST markers. Of the various genetic marker, this study chose mtDNA and microsatellite marker to analyze for resource management and conservation based on results of reconstruct taxonomy, genetic diversity and population structure. The COI and ND4 sequence of mtDNA were amplified using universal primer and species specific primer, respectively. The microsatellites were isolated using 454 GS-FLX Titanium sequencing and 18 markers were used for genotyping in *S. coreana*. In addition, we also tested the cross-species transferability of the microsatellite markers in four additional *Semisulcospira* spp. We identified 18 polymorphic loci and the number of alleles per loci, and their polymorphism information content values ranged from 2 to 17 and 0.203 to 0.902, respectively. The observed and expected heterozygosities of the loci ranged from 0.063 to 0.924 and 0.226 to 0.924, respectively. According to analyze the cross-species transferability of these markers, the four species, *S. forticosta*, *S. gottschei*, *S. tegulata*, *S. libertina*, showed a very high transferability (80-85%). These results show that this set of nuclear markers could be useful for population genetics studies of this species and closely related species.

Keyword: genetic marker; microsatellite; mtDNA; genus *Semisulcospira*

2. Introduction

Genetic resources have evolved with the history of the Earth, and once they are lost, they are products that will never be regenerated and must be passed on to our descendants. However, the diversity of genetic resources is rapidly decreasing as urban and industrialization progresses in modern times and the habitats of living things are damaged along with climate change. Accordingly, the Convention on Biological Diversity (CBD) was put into effect in 1993 for the purpose of preserving biodiversity internationally to cope with the reduction of biodiversity, and it strives globally to preserve marine organism. Before the adoption of the CBD, bio-genetic sources were recognized as a common asset of mankind and could benefit from free access to and utilization of genetic resources. Developed countries have benefited greatly from the development of new species and new materials, mainly by discovering and utilizing the world's useful genetic resources. However, resource holding countries that have many genetic resource has required recognition of genetic resources as national property and insists that they couldn't no longer using genetic resource for free of charge. Therefore, the CBD to equal share of access and benefit sharing to genetic resources (ABS) became effective. Accordingly, research and development needs to be expanded to secure diversity in the genetic resources that can claim the nation's sovereignty and to establish a foundation for utilizing them both domestically and abroad. To do this, Korean traditional, wild and domestic species should be studied. In addition, genetic analysis and discriminating technology should be developed to assert the sovereignty of domestic genetic resources. This is important for resolving disputes regarding the sovereignty of genetic resources that may arise in the

future (Lee at al, 2012).

Among the Korea domestic species, the freshwater snails, genus *Semisulcospira*, is one of economically important a speciose family of Cerithiodiea freshwater gastropods. In 2017, this snail is a major shellfish of domestic inland fishery second only to family Viviparidae (http://kostat.go.). Also, the freshwater snail has been recognized as a healthy food in Korea, with domestic consumption continuing to rise. However, despite their important role, this snail has caused taxonomic confusion in morphological classification due to variety of shell and form caused by flow rate. In previous studies, this snail was researched about morphology analysis (Ko et al, 2001), reproductive anatomy (Kim et al, 2012), karyotypic (Kim et al, 1987), electrophoretic (Jung et al, 1999) and mitochondrial variation (Lee et al, 2007; Kim et al. 2010; Köhler 2016). Various studies have been carried out, but still genetic diversity studies of the freshwater snail, endemic species of Korea, were insufficient. Genetic diversity research provides information on evolutionary factors, such as the adaptability of living organisms, and gene flow (Hartl and Clark 1997). Especially, because the endemic species naturally inhabit only certain regions, they are generally classified and protected as crisis species because of their small size and distribution range, which are vulnerable to environmental changes, often do not win competition from foreign species, and easily occurs genetic confusion. Therefore, in order to manage and preserve the resources of an endemic species, population genetics studies should be conducted, including genetic diversity, genetic differentiation between populations, and phylogenetic relationships (Kim et al, 2011).

To manage and preserve of Korean Peninsula endemic freshwater snails, it is

necessary to analyze genetic resources. Molecular genetic markers are an important tool for assessing the evolutionary level and patterns of genetic diversity of fish and shellfish (Liu and Cordes 2004). Common genetic markers in fish and shellfish include allozymes, mitochondrial DNA, RFLP, RAPD, AFLP, microsatellite, SNP, and EST markers. The principle of these molecular markers is to find genetic variants that have occurred at the DNA level that have accumulated over a long period of time and have led to the next generation, which have contributed to determining the diversity of species. This genetic variation is explored and secured using DNA marker technology (Jang et al, 2009). The mtDNA sequence can be easily obtained from a variety of taxa without cloning, because of characteristic of haploid. Also, the advantage of high evolutionary rate and effective population size allows to recover the pattern and tempo of recent historical events. Of various mtDNA gene, COI, 16S rRNA, Cytochrome b was usually used to analyze genetic differentiation between intra- and inter-species. But, 16S rRNA and cytochrome b have lower genetic variances than COI, limiting the analysis of genetic population structures (Soliman et al, 2017). The mtDNA D-loop is very common to analyze the phylogeny and structure of populations within species due to the fastest evolving region, greatest level of variation (Stoneking et al, 1991). Gene arrangement of mtDNA of Gasptropoda are different from the typical vertebrate (Yang et al, 2014). The absence of D-loop is consistent with the Gastropoda, and so is genus *Semisulcospira* (Zeng et al, 2014). Among the mitochondrial regions, the NADH dehydrogenase subunit 4 (*ND4*) has shown to be an excellent marker for analyzing the genetic population structure in fish and shellfish (Doosey et al, 2010). These genes have been shown to outperform commonly utilized mtDNA

genes such as COI, cytochorome b, 12S/16S rRNA genes in phylogenetic analysis because it is relatively long sequence and contains more phylogenetically informative variation at first and second codon positions (Miya et al, 2006). To reveal the confused taxonomy, over two mitochondrial genes were used in other studies (Zheng et al, 2016).

Allozyme and mtDNA have been traditionally used in fish and shellfish, but their differentiating power is limited compared to more recently developed markers such as AFLP, SNP and microsatellite, there are highly polymorphism (Liu and Codes 2004). The strengths of the AFLP include high reproducibility and polymorphisms, but the method of experimentation is tricky. Microsatellite DNA markers, also named simple sequence repeats (SSRs), are useful tools to assess the genetic variance and structure of populations because of their several desirable features such as variability, co-dominance, and high mutation rates (Feral 2002). They have become the most widely useful DNA technology in many fields of biology, including genome mapping, determining pedigree, population genetics, biological resource conservation, and forensic studies (Knapik et al, 1998; Luikart et al, 2003). Because MS markers could especially provide information on relatively recent evolutionary processes, phylogeographic and population genetic analysis was common used (Avise 2012). The traditional cloning methods used to develop microsatellites involved significant trial and error, required knowledge of the flanking region DNA sequence, and were time-consuming, cost prohibitive, and low throughput (Queller et al, 1993). At present, these problems have been partly resolved with the advent of next-generation sequencing (NGS) technologies that are both cost- and time-effective, as they can manufacture millions of base pairs of short fragment reads in a single run (Moges

et al, 2016). Furthermore, when primers newly developed for one species can be used for broad taxonomic groups is increased the cost- and time-effectiveness of microsatellites. The cross-species amplification of microsatellites has successfully been applied to several marine species (Greenley et al, 2012). Single nucleotide polymorphism (SNP) caused by point mutations that given rise to different alleles containing alternative bases at a given nucleotide position within a locus. SNP marker are inherited as co-dominant, high PIC (polymorphic information content). The advantage of SNPs is that they are richer in available positions than microsatellite and are more freely scattered than any specific area of the genome without bias. Although microsatellite tends to be scattered throughout the genome, this sequence is unlikely to exist in region directly related to the expression of protein, SNPs are less constrained. Although microsatellite should be satisfied by analyzing the genetic structure of two populations, SNPs provide an opportunity to explain genetic variation as well as the genetic structure of the population (Flanagan et al, 2019). However, many SNP markers were needed to provide the same effectiveness as MS for genetic identification and structure analysis (Fernández et al, 2013). More studies have inferred genetic relationships among populations based on the combined analysis of different classes of markers (Allendorf and Seeb 2000; Lu et al, 2001).

In the present study, we developed species-specific ND4 marker of mtDNA of freshwater snail individuals that were identified as 5 species of genus *Semisulcospira* based on sequence analysis of mtDNA COI gene. Also, the first novel polymorphic microsatellite markers for *S. coreana* and tested cross-species amplification in four additional *Semisulcospira* spp. These loci will assist in obtaining genetic information for resource management of the genus

Semisulcospira throughout the Korean drainage area.

3. Material and Method

3.1 Development of mtDNA marker

3.1.1 Sample collection and DNA extraction

The 416 total individual of freshwater snail collected from 19 localities were collected. Information of sample using in the study was shown Table 1. Foot muscles from the samples were maintained in 100% ethanol before being transported to the laboratory. Genomic DNA was done to extract using the QIAGEN DNeasy Blood and Tissue® kit (Qiagen, Valencia, CA) following the manufacturer's instruction. The concentration and quality of extracted genomic DNA was estimated using a spectrophotometer NanoDrop ND-1000 (Thermo Fisher Scientific, Barrington, IL, USA).

3.1.2 Nucleotide sequence analysis

Amplification and sequencing of the partial region of the mtDNA COI were conducted using universal primers: HCO2198 (5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3') and LCO1490 (5'- GGT CAA CAA ATC ATA AAG ATA TTG G-3').

In order to identify within and between species mutations and preserved region in the ND4 gene of mitochondria DNA, the mitochondrial DNA complete gene information of *S. coreana* (NC037771.1 and LC333861.1), *S. libertina* (NC023364.1 and KF736848.1) belonging to genus *Semisulcospira* registered in NCBI GeneBank was aligned using the Bioedit program (http://mbio.ncsu.edu/BioEdit/bioedit. html). Based on the conserved region of

gene, specific forward and reverse primer were developed to amplify the ND4 gene of these snails (Fig. 1).

The PCR profile for the mtDNA COI and ND4 consisted of initial denaturation at 94°C for 4 min, followed by 34 cycles of denaturation at 94°C for 30s, annealing at 56°C and 60°C for 30s, extension at 72°C for 30s, and a final extension at 72°C for 8 min, respectively. The PCR products were purified using a Davinch™ PCR Purification Kit (Davinch-K Co., Ltd., Seoul, Korea). Sequencing of PCR products were done with an Applied Biosystems ABI 3730XL sequencer (Applied Biosystems, Foster City, CA) using an ABI PRISM BigDyeTM Terminator Cycle Sequencing Ready Reaction Kit v3.1 (Applied Biosystems).

Table 1. Freshwater snail species collected in this study

Sample locations	Code	Sample number
Yanggu-gun, Gangwon-do	YG	7
Gosung-gun, Gangwon-do	GS	10
Yeoungdong-gun, Chungcheongbuk-do	YD	7
Hamyang-gun, Gyeongsangnam-do	HY	14
Gaypyoung-gun, Gyeonggi-do	GP	7
Jaecheon, Gyeonggi-do	JC	39
Mt. Jiri, Gyeongsangnam-do	JRS	39
Hadong, Gyeongsangnam-do	HD	40
Namhan river, Gyeonggi-do	NHG	27
Hongcheon, Gangwon-do	HC	42
Inje, Gangwon-do	IJ	9
Yangpyoung, Gyeonggi-do	YP	15
Gapyoung, Gyeonggi-do	GP	26
Jeju-do	JJ	9
Danyang, Jeollanam-do	DY	4
Chungcheongbuk-do	CHU	13
Samcheok, Gangwon-do	SC	11
Yangsan, Gyeongsangnam-do	YS	76
Ulgin, Gyeongsangbuk-do	UG	21

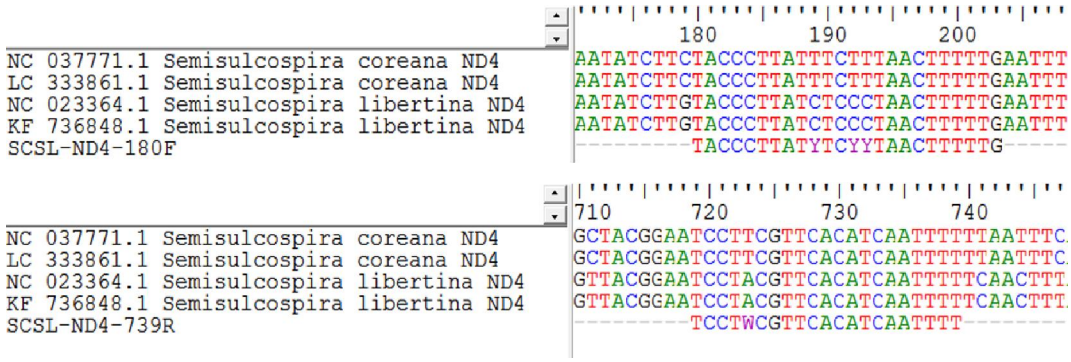

Figure 1. Primer information for ND4 region designed from complete mitochondrial sequences of *S. coreana* (Accession No. NC037771.1 and LC 333861.1), *S. libertine* (Accession No. NC023364.1 and KF736848.1).

3.2 Development of microsatellite marker

3.2.1 Sample collection and DNA extraction

Thirty-seven individuals of *S. coreana* were obtained from the National Institute of Fisheries Science (NIFS) in Korea. Foot muscles from the samples were maintained in 100% ethanol before being transported to the laboratory. Genomic DNA was done to extract using the QIAGEN DNeasy Blood and Tissue® kit (Qiagen, Valencia, CA) following the manufacturer's instruction. The concentration and quality of extracted genomic DNA was estimated using a spectrophotometer NanoDrop ND-1000 (Thermo Fisher Scientific, Barrington, IL, USA). For the cross-species transferability test to other *Semisulcospira* species including *S. tegulata*, *S. libertina*, *S. forticosta*, and *S. gottschei*, DNA extraction was performed by the same method from ethanol-fixed tissue from four related *Semisulcospira* spp. obtained from NIFS, Korea.

3.2.2 454 GS-FLX Titanium sequencing and SSR identification

The NGS library was generated from ~10 μg genomic DNA and sequenced on a GS-FLX-454 pyrosequencing system (454 Life Sciences, Branford, CT, USA) at the NICEM (National Instrumentation Center for Environmental Management of Seoul National University). The obtained sequence reads were trimmed to 96% minimum overlap identity using Newbler 2.6 (Roche Diagnostics, Mannheim, Germany). To search for SSRs in the genomic sequence, the dinucleotide and trinucleotide repeats of more than seven iterations were screened using the Perl program SSR_finder.pl (Tóth et al., 2000). Primer pairs complementary to the

sequence flanking the repeat element were designed using Primer 3 software (Untergasser et al., 2012). The optimal size for primer was set to a range as 18–26 bases and the optimal annealing temperature was set at 58°C. The optimal product size was set to 130–400 bp and the remaining parameters were kept at default settings.

3.2.3. DNA amplification and genotyping

The performances of the newly developed primer sets were tested for optimum concentration for PCR amplification using DNA from eight individuals of *S. coreana*. The electrophoresis of PCR products were operated on a 1.5% agarose gel, and 40 primer sets produced PCR products of 100–300 bp in length. Thirty of these primer sets were labeled and used to amplify DNA from the eight individuals. The forward primer was labeled at 5' end with the fluorescent dyes 6-FAM, TAMRA, and HEX (Applied Biosystems, Foster City, CA, USA). PCR amplifications were performed in 10-μl total volumes that included 0.4 μM of the forward and reverse primer, 0.2 mM dNTPs, 1× PCR buffer, 0.25 U Ex Taq DNA polymerase (Takara Biomedical, Inc., Shiga, Japan), and approximately 100 ng template DNA under the following conditions: pre-denaturation at 95°C for 10 min, followed by 35 cycles of 45 s at 94°C, 45 s of 58°C, and 45 s at 72°C with a finishing of 5 min at 72°C using an ABI 2720 Thermocycler (Applied Biosystems, USA). The PCR products were detected on a 1.5% agarose gel. Fragment analysis was run on a 3730xl DNA Analyzer (Applied Biosystems, USA) using GeneScan 400HD ROX dye as the internal size standard (Life Technologies, Carlsbad, CA, USA) and analyzed using GeneMapper version 5.0 (Applied Biosystems). Primer

sets that generated numerous signal peaks or amplified monomorphic microsatellite loci were excluded.

3.2.4 Transferability analysis

To test the cross-species transferability of the SSR loci, we used all 18 polymorphic sites of *S. coreana* on samples of four related taxa (Table 3). Transferability tests were performed using the abovementioned PCR conditions. Amplification products were monitored on 1.5% agarose gels, and qualitatively graded on the presence or absence of the band.

3.2.5 Data analysis

Numbers of alleles, effective numbers of alleles, expected and observed heterozygosities were assessed using Arlequin 3.0 (Excoffier and Lischer, 2010). The significance level (*P*-value) for Hardy-Weinberg equilibrium (HWE) was evaluated using Genepop v4.0 (Rousset, 2008). Polymorphic information content (PIC) was assessed statistically using the Cervus 3.0 program (Kalinowski et al., 2007). The presence of null alleles and potential genotyping errors was calculated statistically using MicroChecker version 2.2.3 (Van Oosterhout et al., 2004).

4. Result and Discussion

4.1 Development of mtDNA marker

Species identification of genus *Semisulcospira* have been studied using morphological characters such as shape, sculpture, suture, eroding of apex, aperture lip (Lee et al, 2017). However, with the discovery of variation in external characteristics by flow velocity and drainage environment, species identification by morphological analysis has its limits. To complement this problem, several taxonomic studies have been conducted using the mtDNA COI, 16S rRNA gene at genus *Semisulcospira* (Köhler 2016; Kim et al, 2010). Köhler (2016) suggest that *S. coreana*, *S. gottschei* and *S. forticosta* has been clear separation in systematic and phlygenetic studies, but *S. tegulata* were transferred to *Koreoleptoxis tegulata*. Otherwise, Kim et al (2010) reported that *S. tegulata* appeared to have little genetic difference with *S. coreana*, making it the closest to a genetic relationship in phylogeny tree. Therefore, it is necessary to reconstruct phylogenetic relationships to target species that are genetically identical to morphological characters. The 456 bp sequence was acquired and aligned using universal primer of COI, resulting in seven clade showing on the phylogenetic tree (Chapter 3, Fig 2). Among them, species-specific ND4 primer designed from complete mtDNA sequence of *S. coreana* and *S. libertina* in NCBI was used to amplify the species with matching form and molecular classification. The information of spices-specific ND4 primers was and SCSL-ND4-180F (5'- TAC CCT TAT YTC YYT AAC TTT TTG-3') and SCSL-ND4-739R (5'-AAA ATT GAT GTG AAC GWA GGA-3'), be shown Fig 1. The acquired COI and ND4 of sequence information could be used for genetic diversity and population structure

analysis. Furthermore, since two markers with higher genetic diversity were used than studies using conventional COI and 16S rRNA, it is expected that the resolving power of genetic research will increase further.

4.2 Development of microsatellite marker

NGS generated 17,719,000 reads of a total of 8,894,938,000 bp for *S. coreana*. The total number of contigs was 132,872 and the number of singletons was 1,754,529. The total number of reads after trimming low-quality sequences was 9,967,563. A total of 3,565,825,918 regions were found to contain from nucleotide repeats for more than four iterations. The amplicon sizes varied from 90 to 320 bp. Of these, 41%, 27%, and 32% were di-, tri-, and tetra-nucleotides, respectively. CA repeats were the most common dinucleotide repeats, TTG repeats were the most common trinucleotide repeats, and TGTG repeats were the most common tetra-nucleotide repeats.

All of 78 SSRs with high copy numbers of all repeat motifs for the amplification and assessment of polymorphisms were synthesized to measure amplification efficiency and extent of polymorphism. Of these, a total of 18 primer pairs (23%) showing clear amplification of polymorphic loci were applied for PCR analysis and genotyping. Primer details and fluorescent labeling information are shown in Table 2.

In the present, microsatellite developed in most marine resource, was mostly consisted as di-nucleotide repeats. However, tri- and tetra-nucleotides motif have the advantages like highly polymorphic and stable than di-nucleotide repeats (Lindqvist et al, 1996). We developed 18 novel polymorphic microsatellite markers consisted of all of di-, tri-, and tetranucleotide motifs. The allele number per locus resulted from 2 to 17, with a mean of 6.9. The PIC for each locus ranged from 0.203 to 0.902, with a mean of 0.626. Botstein et al (1980) described that a PIC value above 0.5 has sufficient discrimination within a population. Also,

comparisons of PIC values can present an estimate of the power of the marker. Fifteen of the 18 loci were supposed to be highly informative (PIC > 0.5). Another two (Sc-27 and Sc-60) were indicated reasonably informative (0.25 < PIC < 0.5), and one (Sc-76) was only slightly informative (PIC < 0.25). This suggests that the set of microsatellites generated in the present study shows considerable potential for analyzing genetic polymorphisms.

The observed and expected heterozygosities of the loci resulted from 0.063 to 0.906and 0.226 to 0.924, respectively. The inbreeding coefficient (F_{IS}) ranged from -0.750 to 0.844. Null alleles were detected at most of the loci, except for Sc06, Sc53, and Sc60, which showed deviation from HWE after Bonferroni correction (P < 0.0027) with expected heterozygosity (H_e) values greater than observed heterozygosity (H_o) values. The significant deficits of heterozygosity are associated with the existence of null alleles (Selkoe and Toonen 2006), which are common in freshwater mollusc (Gu at al, 2012). This is also demonstrated by the significant positive F_{IS} values at most of the loci, except for Sc-60, indicating that the excess of homozygotes was likely due to the existence of null alleles or stuttering at these loci, as suggested by the Micro-Checker results. This is also caused by inbreeding within a population in invertebrates due to the poor mobility of freshwater snails (Gu et al, 2015).

Cross-species amplification of the 18 polymorphic loci was conducted within four species of *Semisulcospira*, *S. forticosta*, *S. gottschei*, *S. libertina*, and *S. tegulata*. As shown in Table 3, three polymorphic loci (Sc-50, Sc-60, and Sc-76) were not amplified by any species tested. These three loci could be used as identification of *S. coreana*. Loci Sc-24 showed polymorphic amplification in all species tested except *S. libertina*. The transferability test of these microsatellite

markers for the four species showed a very high transferability (80-85%). The transferability of SSR markers in different snail species is affected by the nucleotide sequence similarity in the SSR marker primer sites among related species (Barbara et al, 2007). The cross-species transferability of SSR primers in the snails was high, indicating that there is high conservation of markers between *Semisulcospira* spp.

We developed 18 new microsatellite markers for *S. coreana* using 454 GS-FLX NGS. This is the first set of polymorphic microsatellite markers designed for *S. coreana* and related species. These microsatellite markers will potentially be useful for genetic diversity and population genetic structure analyses of snails throughout the Korean drainage area.

Table 2. Characteristics of *Semisulcospira coreana* microsatellite loci

Locus	Dye	Primer sequence (5'→3')	Motif	N_A	PIC	H_O	H_E	F_{IS}
Sc-01	FAM	F: CTAAAACTCCCACAATTGAAGC	(TG)7	3	0.524	0.290	0.607	0.521*
		R: CTGAGATTCGCGTGTCTACATA						
Sc-02	HEX	F: TTGAGAGATACTCTGACACCCC	(CA)7	4	0.514	0.344	0.580	0.407
		R: TAGGACCCACAGGACTAAACTG						
Sc-06	HEX	F: GGAAAGTGATATGTGGAGGTGT	(TG)8	7	0.782	0.719	0.821	0.124*
		R: AACTTAACGTTCTGCTCTTTCG						
Sc-11	TAMRA	F: CTGTTTGACTCTCAAGGCTTCT	(CA)9	5	0.583	0.281	0.647	0.566*
		R: AAATGAATCTGTTTTCAGCGAT						
Sc-13	FAM	F: TGGTTTTTCTGTGGTAGGTTTC	(TG)10	5	0.729	0.400	0.784	0.490*
		R: TAGACACCAAACAATTATGGCA						
Sc-16	FAM	F: GTTGATGATGGTGAAGATTGTG	(TG)11	8	0.726	0.600	0.767	0.218*
		R: ACAAACCAACAGCATTCTTTCT						
Sc-19	TAMRA	F: AGCACCGAGGTAAAATGAGTTA	(CT)7	10	0.81	0.469	0.844	0.444*
		R: GGGTCTGGTCATGAGGTAAATA						
Sc-24	FAM	F: GTCTATGTTCCCACACCACTCT	(CT)9	12	0.769	0.548	0.804	0.318*
		R: AATTCGGTAGACAAAACCACAC						
Sc-26	TAMRA	F: TTGCAGAAAGAAATGTCAACAG	(CT)9	11	0.771	0.419	0.809	0.481*
		R: TAAAACAAGATGTGATGACCCA						
Sc-27	FAM	F: CTAGCATTTGCTAATCTGACCC	(CAA)7	3	0.345	0.063	0.400	0.844*
		R: AAGGTTGTTGCCTAACGTAGAA						
Sc-28	HEX	F:CATCAAAGAACACTCAAACATCA	(CAA)7	13	0.828	0.406	0.859	0.527*
		R: ATACATGTGCGTAGTTGAATCG						
Sc-29	HEX	F: GTGGTTGTTATGGTGGTCTTTT	(TTG)7	7	0.563	0.429	0.603	0.290*
		R: TCTATACCTCTTGCGGTTTGAT						
Sc-30	TAMRA	F: CTTGATCCACTACCCATGATCT	(TTG)7	6	0.725	0.375	0.768	0.512*
		R: ACGAAATGAGTTCGAGGTTTTA						
Sc-36	HEX	F: CTGCTGGTGTTGTTTTCTGTTA	(TTG)9	4	0.546	0.387	0.620	0.376
		R: CCTGAAAGAGGCACATATTCTC						
Sc-50	TAMRA	F: GCCCTTCGTCTATCTTGTACTG	(TAT)7	4	0.583	0.207	0.664	0.688*
		R: ACAAAGACTAACCATTTGGGTG						
Sc-53	FAM	F:TGAGTCTTATTTTCAAACGGAAA	(TGTG)8	17	0.902	0.906	0.924	0.019
		R: TAGCAAGCCCCTTATGTGTAGT						
Sc-60	TAMRA	F: AAAGAAACAAAGGATTCTGCAA	(CACA)9	2	0.371	0.875	0.500	0.750*
		R: TTCCGCTAATCTACGTTCAAGT						
Sc-76	TAMRA	F: TGTGTATGTCTAAGCCGCTATG	(GTGT)8	3	0.203	0.094	0.226	0.585*
		R: ATACGACAACAAAAGTCCCAAC						

N_A, number of alleles; PIC, polymorphic information content; H_O and H_E, observed and expected heterozygosities; F_{IS}, inbreeding coefficient; *F_{IS}, significant deviation from Hardy-Weinberg equilibrium after Bonferroni correction ($P < 0.0027$).

Table 3. Cross-species amplification in related *Semisulcospira* spp. using *S. coreana* microsatellite loci

Locus	Genus *Semisulcospira*				
	S. coreana (N = 37)	*S. forticosta* (N =41)	*S. gottschei* (N = 24)	*S. libertina* (N = 9)	*S. tegulata* (N = 8)
Sc-01	147-151	147-151	147-153	147-155	147-151
Sc-02	193-203	193-203	159-203	195	193-203
Sc-06	206-250	206-218	204-248	240-242	206-210
Sc-11	223-261	223-315	233-259	235	233-261
Sc-13	154-168	162-172	162-168	178	154-168
Sc-16	150-172	146-182	158-168	160	156-168
Sc-19	227-263	243-307	241-263	301-307	231-263
Sc-24	154-198	150-186	148-202	-	162-198
Sc-26	261-291	257-291	261-291	263	261-289
Sc-27	146-152	146-158	146-155	143	146-155
Sc-28	191-239	191-260	194-212	197	191-224
Sc-29	243-261	243-258	243-249	234	246-249
Sc-30	294-330	279-309	291-336	303	291-330
Sc-36	208-217	193-226	208-214	211	211-214
Sc-50	249-261	-	-	-	-
Sc-53	126-190	126-194	134-190	182-194	134-178
Sc-60	249-261	-	-	-	-
Sc-76	276-284	-	-	-	-

5. Reference

Lee H.S., Lee S.Y., Baek H.J., Jeong E.G., Kim C.K., Kim Y.G. (2012) International Trends and Implication on Access to Plant Genetic Resources and Benefit Sharing (ABS). Korean. J. Intl. Agri., 24, 106-115.

Ko J.H., Lee J.S., Kwon O.K. (2001) Study on radulae of seven species of the family Pleuroceridae in Korea. Korean. J. Malacol., 17, 105-115.

Kim D.H., Bang I.C., Lee W.O., Baek J.M. (2012) Reproductive Ecology of the Freshwater Melania Snail, Semisulcospira coreana (v. Martens) in Bukhan River. Korean. J. Malacol., 28, 175-185.

Kim S.S., Kim D.C., Chung P.R., Chung Y.J. (1987) A cytological study of two species of genus Semisulcospira (Gastropoda: Pleuroceridae) in Korea. Korean. J. Malacol., 3, 24-34.

Jung Y., Park G.M., Park J., Kim J.J., Min, D.Y., Chung P.R. (1999) Allozyme variability in three species of genus Semisulcospira (Prosobranchia: Pleuroceridae). Korean. J. Malacol., 15, 13-20.

Lee T., Hong H.C., Kim J.J., Foighil D.Ó. (2007) Phylogenetic and taxonomic incongruence involving nuclear and mitochondrial markers in Korean populations of the freshwater snail genus Semisulcospira (Cerithioidea: Pleuroceridae). Mol. Phylogenet. Evol., 43, 386-397.

Kim W.J., Kim D.H., Lee J.S., Bang I.C., Lee W.O., Jung H.T. (2010) Systematic Relationships of Korean Freshwater Snails of Semisulcospira, Koreanomelania, and Koreoleptoxis (Cerithiodiea; Pleuroceridae) revealed by Mitochondrial Cytochrome Oxidase I Sequences. Korean. J. Malacol., 26, 275-283.

Köhler F. (2016) Rampant taxonomic incongruence in a mitochondrial phylogeny of Semisulcospira freshwater snails from Japan (Cerithioidea: Semisulcospiridae). J. Molluscan. Stud., 82, 268-281.

Hartl D.L. and Clark A.G. (1997) Principles of Population Genetics. Sinauer Associates Inc., Sunderland, Massachusetts, 542pp.

Kim C.K., Kim H.J., Choi H.K. (2011) Assessment of the minimum population size for ex situ conservation of genetic diversity in Aster altaicus var. uchiyamae populations inferred from AFLP markers. Kor. J. Env. Eco., 25, 470-478.

Liu Z.J., and Cordes J.F. (2004) DNA marker technologies and their applications in aquaculture genetics. Aquaculture, 238(1-4), 1-37.

Jang Y.S., Yi S.K., Noh C.H., Oh S.Y. (2009) Genetic Diversity of the Mud Crab Scylla serrata in Micronesia based on Microsatellite Marker Analysis. Ocean and Polar Research, 31, 319-326.

Soliman T., Aly W., Fahim R.M., Berumen M.L., Jenke-Kodama H., Bernardi G. (2017) Comparative population genetic structure of redbelly tilapia (Coptodon zillii (Gervais, 1848)) from three different aquatic habitats in Egypt. Ecol. Evol., 7, 11092-11099.

Stoneking M., Hedgecock D., Higuchi R.G., Vigilant L., Erlich H.A. (1991) Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. Am. J. Hum. Genet., 48, 370-382.

Yang H., Xie Z., Li S., Wu X., Peng C., Zhang Y., Lin H. (2016) The complete mitochondrial genome of the orange-spotted grouper Epinephelus coioides (Perciformes, Serranidae). Mitochondrial DNA Part A, 27, 1674-1676.

Zeng T., Yin W., Xia R., Fu C., Jin B. (2015) Complete mitochondrial genome of a freshwater snail, Semisulcospira libertina (Cerithioidea: Semisulcospiridae). Mitochondrial DNA, 26, 897-898.

Doosey M.H., Bart Jr H.L., Saitoh K., Miya M. (2010) Phylogenetic relationships of catostomid fishes (Actinopterygii: Cypriniformes) based on mitochondrial ND4/ND5 gene sequences. Mol. Phylogenet. Evol. 54(3), 1028-1034.

Miya M., Saitoh K., Wood R., Nishida, M., Mayden R.L. (2006) New primers for amplifying and sequencing the mitochondrial ND4/ND5 gene region of the Cypriniformes (Actinopterygii: Ostariophysi). Ichthyol. Res. 53(1), 75-81.

Zheng L.P., Yang J.X., Chen X.Y. (2016) Molecular phylogeny and systematics of the Barbinae (Teleostei: Cyprinidae) in China inferred from mitochondrial DNA sequences. Biochem. Syst. Ecol. 68, 250-259.

Feral J.P. (2002) How useful are the genetic markers in attempts to understand and manage marine biodiversity?. J. Exp. Mar. Biol. Ecol., 268, 121-145.

Knapik E.W., Goodman A., Ekker M., Chevrette M., Delgado J., Neuhauss S., Jacob H.J. (1998) A microsatellite genetic linkage map for zebrafish (Danio rerio). Nat. genet., 18, 338.

Luikart G., England P.R., Tallmon D., Jordan S., Taberlet, P. (2003) The power and promise of population genomics: from genotyping to genome typing. Nat. Rev. Genet., 4, 981.

Avise J.C. (2012) Molecular markers, natural history and evolution. 2nd ed. Sunderland, MA: Sinuate.

Queller D.C., Strassmann J.E., Hughes C.R. (1993) Microsatellites and kinship. Trends. Ecol. Evol., 8, 285-288.

Moges A.D., Admassu B., Belew D., Yesuf M., Njuguna J., Kyalo M., Ghimire S.R. (2016) Development of microsatellite markers and analysis of genetic diversity and population structure of Colletotrichum gloeosporioides from Ethiopia. PloS one., 11, e0151257.

Greenley A.P., Muguia-Vega A., Saenz-Arroyo A., Micheli F. (2012) New tetranucleotide microsatellite loci in pink abalone (Haliotis corrugata) isolated via 454 pyrosequencing. Conservation. Genet. Resour., 4, 265-268.

Flanagan, S. P., & Jones, A. G. (2019). The future of parentage analysis: From microsatellites to SNPs and beyond. Molecular ecology, 28(3), 544-567.

Fernández M.E., Goszczynski D.E., Lirón J.P., Villegas-Castagnasso E.E., Carino M.H., Ripoli

M.V. et al. (2013) Comparison of the effectiveness of microsatellites and SNP panels for genetic identification, traceability and assessment of parentage in an inbred Angus herd. Genetics and molecular biology, 36, 185-191.

Allendorf F.W. and Seeb L.W. (2000) Concordance of genetic divergence among sockeye salmon populations at allozyme, nuclear DNA, and mitochondrial DNA markers. Evolution, 54, 640-651.

Lu G., Basley D.J., Bernatchez L. (2001) Contrasting patterns of mitochondrial DNA and microsatellite introgressive hybridization between lineages of lake whitefish (Coregonus clupeaformis); relevance for speciation. Mol. Ecol., 10, 965-985.

Tóth G., Gáspári Z., Jurka J. (2000) Microsatellites in different eukaryotic genomes: survey and analysis. Genome research., 10, 967-981.

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

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.

Rousset F. (2008) GENEPOP'007: A complete re-implementation of the genepop software for Windows and Linux. Mol. Ecol. Resour., 8, 103-106.

Kalinowski S.T., Taper M.L., Marshall T.C. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol. Ecol., 16, 1099-1106.

Van Oosterhout C., Hutchinson W.F., Wills D.P., Shipley P. (2004) MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Mol. Ecol. Notes., 4, 535–538.

Lee J.S. and Ko J.H. (2017) The study of external morphological characters of the family Pleuroceridae in Korea. Journal of the Evironment, 1, 27-39.

Lindqvist A.K., Magnusson P.K., Blciuniene J., Wadelius C., Lindholm E., Alacon-Riquelme M.E., Gyllensten U.B. (1996) Chromosome specific panels of tri- and tetranucleotide microsatellite marekrs for multiplex fluorescent detection and automated genotyping: evaluation of their utility in pathology and forensics. Genome. Res., 6, 1170-1176.

Botstein D., White R.L., Skolnick M., Davis R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet., 32, 314-331.

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

Gu Q., Xiong B., Zhu Y., Wang Q., Shi P. (2012) Isolation and characterization of polymorphic microsatellites loci in the freshwater snail Bellamya aeruginosa (Mollusca, Gastropoda). Conservation. Genet. Resour., 4, 387-390.

Gu Q.H., Cheng Q.Q., Li X.J., Zhou C.J. (2015) Novel polymorphic microsatellite markers for Bellamya and their application in population genetics of three species. Genet. Mol. Res., 14, 15201-15212.

Barbara T., PALMA-SILVA C., Paggi G.M., Bered F., Fay M.F., Lexer C. (2007) Cross-species transfer of nuclear microsatellite markers: potential and limitations. Mol. Ecol., 16, 3759-3767.

Chapter III.

Analysis of genetic differentiation and population structure of the Korean-peninsula-endemic genus, *Semisulcospira*, using mitochondrial markers

Analysis of genetic differentiation and population structure of the Korean-peninsula-endemic genus, *Semisulcospira*, using mitochondrial markers

1. Abstract

The genus *Semisulcospira*, which is endemic to the Korean Peninsula and consists of *S. coreana*, *S. forticosta*, *S. gottschei*, and *S. tegulata*, is an economically and ecologically valuable freshwater resource. Therefore, maintenance and conservation of wild populations of melania snail are important. We investigated the genetic diversity and population structure of *Semisulcospira* based on the mitochondrial COI, ND4, and combined mtDNA (COI+ND4) sequences. All four species and various genetic makers showed a high level of haplotype diversity and a low level of nucleotide diversity. In addition, Fu's *F*s and Tajima's *D* neutrality tests were performed to assess the variation in size among populations. Neutrality tests of the four species yielded negative Fu's *F*s and Tajima's *D* values, except for populations with one haplotype. The minimum spanning network indicated a common haplotype for populations of *S. coreana*, *S. tegulata* and *S. gottschei*, whereas *S. forticosta* had a rare haplotype. Also, genetic differences and gene flows between populations were assessed by analysis of molecular variance and using the pairwise fixation index. Our findings provided insight into the degree of preservation of the species' genetic diversity and could be utilized to enhance the management of endemic species.

Key words: *Semisulcospira*, endemic species, population genetic diversity, mtDNA

2. Introduction

The decline in biodiversity in freshwater ecosystems is greater than that in many other ecosystems because of rising human demand for certain of the organisms therein over the past century (Dudgeon et al, 2006). This "biodiversity crisis" requires priority setting at global, regional, and local scales to concentrate limited resources to address the most pressing conservation issues (Darwall and Vie 2005). Endemic species, which are conservation priorities, are typically seen in freshwater habitats due to cause evolution of many species with small geographic range (Strayer and Dudgeon 2010). Therefore, endemic species are primary candidates for extinction (Anderso 1994). Especially, molluscs have higher extinction rates than any other taxa, making them the most threatened group of animals on Earth (Regnier et al, 2009). Efficient resource management requires management of population structure based on genetic research into species and populations (Kang et al, 2010). To this end, the Ministry of Environment of the Republic of Korea (ROK) has identified Korean Peninsula-endemic organisms and operates an approval system that must be satisfied before they can leave the country. Among the taxa endemic to the Korean peninsula, the genus *Semisulcospira* is a speciose family of Cerithiodiea freshwater gastropods distributed throughout eastern Asia, including in the ROK, Japan, China, and Taiwan (Chiu et al, 2016). Although five species of the genus *Semisulcospira* inhabit the ROK, only four (*S. coreana*, *S. forticosta*, *S. gottschei*, and *S. tegulata*) have been designated as endemic species by the Ministry of Environment. *Semisulcospira* is a major shellfish for domestic inland fisheries, as evidenced by an annual output of 574 tons and annual production value of 6.4 billion Korean

Won, second only to the family Viviparidae (http://kostat.go). Moreover, the snails are intermediate hosts of human pathogens and are used as a biological marker of the environmental status of freshwater systems (Lou et al, 2011; Pan et al, 2011). Despite their important role, there were still the insufficient taxonomic system and genetic population structure. Until now, the morphology, reproductive anatomy, and karyotypic, electrophoretic, and mitochondrial variation of this snail species have been analyzed but its taxonomy remains unclear (Ko et al, 2001; Kim et al, 2012; Kim et al, 1987; Jung et al, 1999; Lee et al, 2007; Kim et al, 2010; Miura et al, 2013; Köhler 2016). Therefore, it is necessary to reconstruct phylogenetic relationships to determine species that are genetically identical based on their morphological characters, and to analyze the genetic variation of populations throughout the ROK. Also, investigation of within-species genetic diversity is needed to obtain information on population migration and history.

Studies of genetic population structure have sought to improve the management and preservation of freshwater resources. To date, most genetic diversity studies have used isotopic enzymes as markers, but developments in molecular-biological technology have enabled large amounts of genetic information to be processed. (Kang et al, 1996). This genetic information facilitates the development of new markers for analysis of populations of marine organisms. In analyses of genetic diversity, the sensitivity of the marker gene determines its performance, and mitochondrial DNA (mtDNA) is typically used due to its high sequence variation (Dunham 2004). This variation enables the exploration of parameters such as subdivision and gene flow (Hillis et al, 1996; Zhu et al, 1994). The mitochondrial cytochrome c oxidase subunit I (COI) gene (1,500 bp) has been used as a DNA barcode for population studies of fish and

shellfish because of its marked nucleotide variation (Ward et al, 2005). In addition, the NADH dehydrogenase subunit 4 (ND4) gene has been used as a marker for analyzing the genetic population structure of fish and shellfish (Doosey et al, 2010). The population genetic structure of shellfish using the mtDNA COI gene has been variously studied (An et al, 2000; Arnaud et al, 2000; Kim et al, 2000; Lee and Kim 2003; Matsumoto 2003). However, our understanding of intraspecies genetic variation by geographic area for this snail is limited.

The aims of this study were to (1) reconstruct the relationships among *Semisulcospira* species, (2) analyze the pattern of genetic diversity among four *Semisulcospira* species endemic to Korea, (3) infer the history of the populations,

3. Materials and Methods

3.1 Sampling collection and DNA extraction

Korean melania snails of the genus *Semisulcospira* were obtained from 19 localities in the ROK from December 2015 to August 2019. Information on the samples is shown in Fig. 1. The samples were preserved in 99% ethanol until use. Genomic DNA was extracted using the DNeasy Blood and Tissue® Kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions. Briefly, ~ 20 mg of tissue was lysed in Buffer ATL containing proteinase K (Qiagen) at 56°C for 3 h. The lysates were mixed with Buffer TL and ethanol, and DNA was purified using a resin column. The concentration of genomic DNA was estimated using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

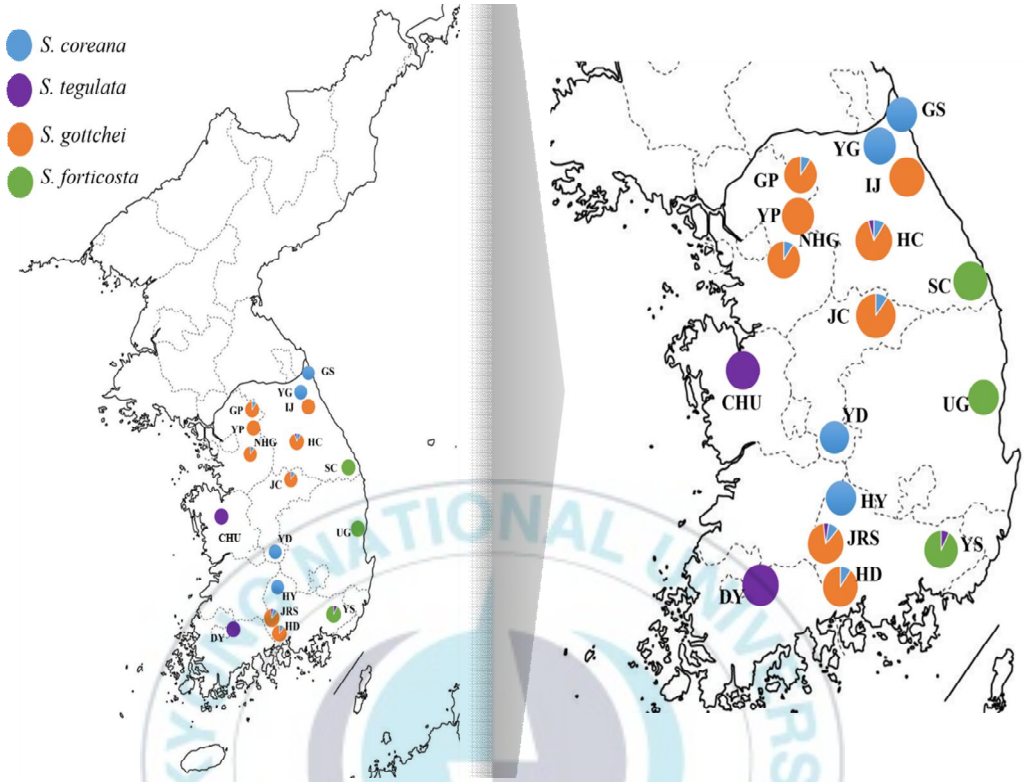

Figure 1. *Semisulcospira* sampling sites in the Republic of Korea (ROK). Each color indicates a different species. GS, Goseong; YG, Yanggu; IJ, Inje; GP, Gapyoung; YP, Yangpyoung; HC, Hongcheon; SC, Samcheok; JC, Jaecheon; YD, Yeoungdong; HY, Hamyang; UG, Ulgin; DY, Damyang; HD, Hadong; YS, Yangsan; JRS, Mt. Jiri; NHG, Namhan river; CHU, Chungcheongbuk-do. Each of color circle indicated species and the ratio of specimen.

3.2 PCR amplification and Nucleotide sequence

Amplification and sequencing of the partial region of the mtDNA COI and ND4 gene were conducted using universal and species-specific designed primers: HCO2198 and LCO1490 and SCSL-ND4-180F (5'- TAC CCT TAT YTC YYT AAC TTT TTG-3') and SCSL-ND4-739R (5'-AAA ATT GAT GTG AAC GWA GGA-3'). The PCR profile and sequencing condition for the mtDNA COI and ND4 were formed of mention of chapter 2.

3.3 Statistic analyses

The acquired sequences were assembled using ClustalW in SeqMan software (DNASTAR, Madison, WI, USA). Phylogenetic analysis was performed by the neighbor-joining (NJ) method with bootstrap analysis based on the Kimura two-parameter distance in MEGA software (www.megasoftware.net). Haplotype diversity (h) and nucleotide diversity (π) were estimated using ARLEQUIN v. 3.5.1.2 (http://lgb.unige.ch/arlequin/ software/). Demographic history was estimated based on the Tajima's D and Fu's Fs values, obtained by neutrality testing in ARLEQUIN v. 3.5.5. Analysis of molecular variance (AMOVA) was performed to assess differentiation between populations using ARLEQUIN v. 3.5.1.2. The significance of pairwise FST values was calculated at the 0.05 level based on 10,000 permutations. A minimum spanning network was generated from the 456 bp alignment of sequences using PopArt v. 1.7 software (http://popart.otago.ac.nz), with epsilon set to zero.

4. Results

4.1 Systematic analysis using genetic markers

The sequence of the 456 bp fragment of the mtDNA COI of melania snail has been submitted to GenBank (accession number NC 037771, HM 991879, KY 675029, KY675052, HM991871). NJ trees were constructed based on the COI haplotypes distributed among seven well-supported clades. Each clade comprised the same species, with two clades consisting of *Semisulcospira* sp. (Fig 2). Regarding interspecific genetic distances, *S. coreana* and *S. tegulata* were closest (1.5%), while *S. forticosta* and *S. libertina* were the most distant (8.9%). Also, the genetic distance of clade C from *S. coreana* was 2.3%, which was greater than that between *S. tegulata* and *S. coreana*. Clade F had a genetic distance of 17.4–18.7% from clades A–E. The intergenic distance value for clade D (*S. coreana*) was 0.8% and that of clade A (*S. gottschei*) was 0.7%. Clades B and E (*S. forticosta* and *S. tegulata*, respectively) showed a 0.5% interspecific difference, while clade F (*S. libertina*) comprised only one haplotype. We analyzed the genetic relationships based on the ND4 gene of melania snail species identified based on the COI gene. In NJ trees based on the ND4 gene (Fig 3), the interspecific genetic distance ranged from 1.8% to 10.7%, which was greater than the values determined based on the COI gene. This was higher than the intraspecific variation based on the COI gene (0.4–0.9%); *S. coreana* showed the greatest genetic variation, followed by *S. forticosta*, *S. gottschei*, and *S. tegulata*.

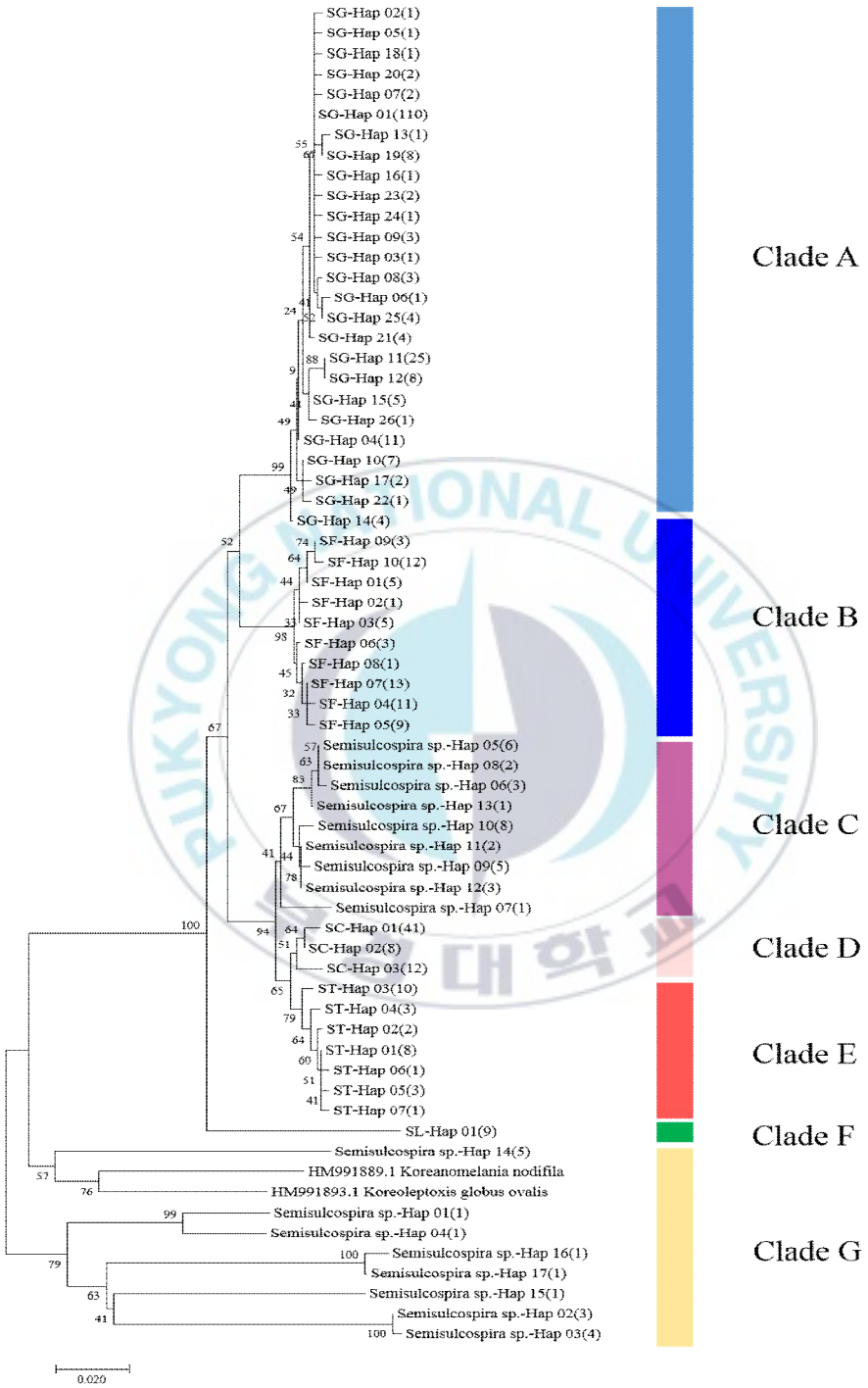

Figure 2. Neighbor-joining (NJ) phylogenetic tree based on the mitochondrial DNA (mtDNA) cytochrome c oxidase subunit I (COI) haplotypes of each clade.

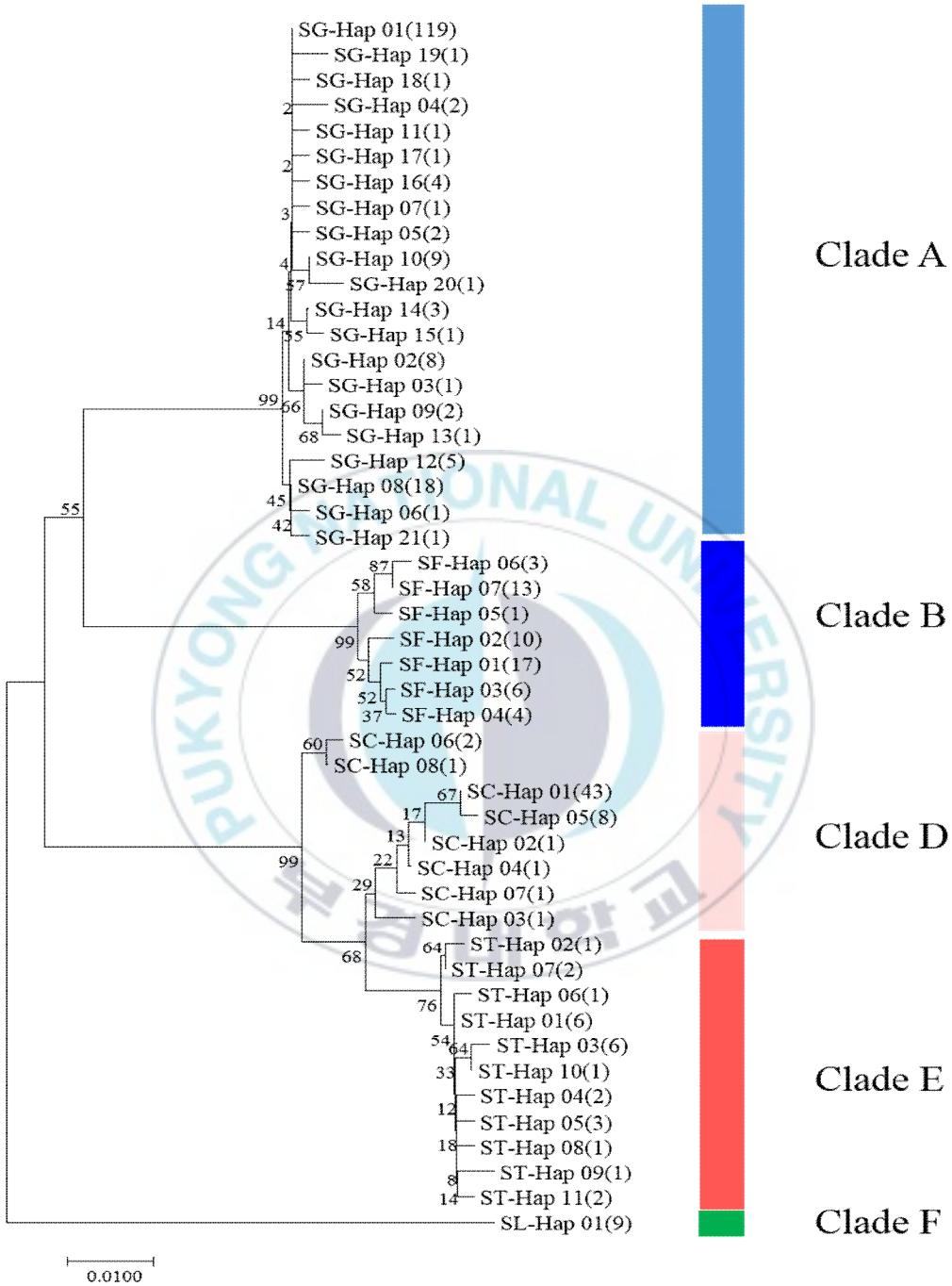

Figure 3. Neighbor-joining (NJ) phylogenetic tree based on the mitochondrial DNA (mtDNA) NADH dehydrogenase subunit 4 (ND4) haplotypes of each clade.

4.2 Genetic diversity and population structure of *S. coreana*

A total of 61 *S. coreana* individuals of 10 populations was obtained from the 456 bp mtDNA COI gene fragment. Three haplotypes were detected only among COI genes; haplotype 1 (Hap_1) predominated (67.2%). Eight haplotypes were detected in the 486 bp ND4 gene from 58 samples of *S. coreana*, with Hap_1 being predominant (70.4%). Analysis of the 942 bp COI+ND4 identified 11 haplotypes among 58 individuals from 10 locations. COI+ND4 yielded a higher haplotype diversity than COI and ND4 in each population (Table 1). The GP, JRS, and HC populations displayed only one haplotype for all of the genetic markers. The YD population had a higher level of haplotype diversity than the other populations for all genetic markers (0.476–0.952). The nucleotide diversity of the YD population was higher than that of the other populations.

The results of neutrality tests are shown in Table 1. Fu's *F*s value was negative for all populations (excluding those with only one haplotype) and those of the YD and NHG populations were not significant ($P < 0.05$). Also, Tajima's *D* test yielded a negative result for the YG, GS, and HD populations ($P < 0.05$).

Table 1. Genetic diversity based on mtDNA markers of *S. coreana* populations from each sampling location in the ROK

Genetic marker	Locality	n	N	*h*	π	Fu's *F*s	p-value	Tajima's *D*	p-value
mtDNA COI	YG	7	2	0.2857	0.00125	-11.147	0	-1.237	0.122
	GS	10	3	0.6	0.00409	-10.986	0	-0.496	0.304
	YD	7	2	0.476	0.00627	-4.557	0.002	0.847	0.812
	HY	10	3	0.6	0.00673	-8.194	0	1.845	0.988
	GP	7	1	0	0	0	1	0	0
	JC	3	2	0.667	0.00292	-1.216	0.063	0	0
	JRS	4	1	0	0	0	1	0	1
	HD	6	2	0.333	0.00146	-8.093	0	-1.131	0.151
	NHG	3	2	0.667	0.00292	-1.216	0.062	0	0.942
	HC	4	1	0	0	0	1	0	1
Genetic marker	Locality	n	N	*h*	π	Fu's *F*s	p-value	Tajima's *D*	p-value
mtDNA ND4	YG	7	3	0.524	0.00274	-7.303	0	-0.877	0.252
	GS	10	3	0.378	0.00165	-16.717	0	-1.667	0.031
	YD	7	5	0.905	0.0098	-3.13	0.024	1.575	0.946
	HY	8	2	0.429	0.00088	-15.453	0	0.3335	0.809
	GP	7	1	0	0	0	1	0	1
	JC	2	1	0	0	0	1	0	1
	JRS	4	1	0	0	0	1	0	1
	HD	6	1	0	0	0	0	0	1
	NHG	3	1	0	0	0	1	0	1
	HC	4	1	0	0	0	1	0	1

Table 1. (*Continued*)

Genetic marker	Locality	n	N	*h*	π	Fu's *F*s	p-value	Tajima's *D*	p-value
Combinded mtDNA (COI+ND4)	YG	7	4	0.714	0.00202	-5.926	0	-1.129	0.161
	GS	10	4	0.645	0.00283	-8.933	0	-1.089	0.145
	YD	7	7	0.952	0.00809	-2.078	0.074	1.351	0.918
	HY	8	3	0.643	0.00364	-5.135	0.028	1.285	0.907
	GP	7	1	0	0	0	1	0	1
	JC	2	1	0	0	0	1	0	1
	JRS	4	1	0	0	0	1	0	1
	HD	6	2	0.333	0.00071	-8.093	0	-1.132	0.141
	NHG	3	2	0.667	0.00142	-1.216	0.064	0	0.936
	HC	4	1	0	0	0	1	0	1

n, number of specimens; N, number of haplotypes; h, haplotype diversity; π, nucleotide diversity

The phylogeny of the 11 haplotypes of the COI+ND4 sequence with no visible geographic population structure was assessed using the minimum spanning algorithm (Fig 4). The number of haplotypes per population ranged from one to seven (Table 1). The predominant haplotype was present in all 10 penguin colonies (62% of all individuals). The second most common haplotype was present in 12% of all individuals, but was shared only by the HY, YD, and GS populations. The YD, YG, HY, and GS populations had 4, 2, 1, and 1 unique haplotype, respectively.

AMOVA showed that the variation among populations (36.1%) was lower than that within populations (63.9 The overall F_{ST} value (0.36098) indicated weak genetic differentiation among 10 populations (Table 2). The pairwise F_{ST} values of the COI+ND4 sequence ranged from −0.007 to 0.620 (Table 3). The HY and YD populations exhibited significantly greater pairwise genetic differentiation than the other populations. The YG and GS populations had the lowest genetic divergence, and the HY and GP populations the highest.

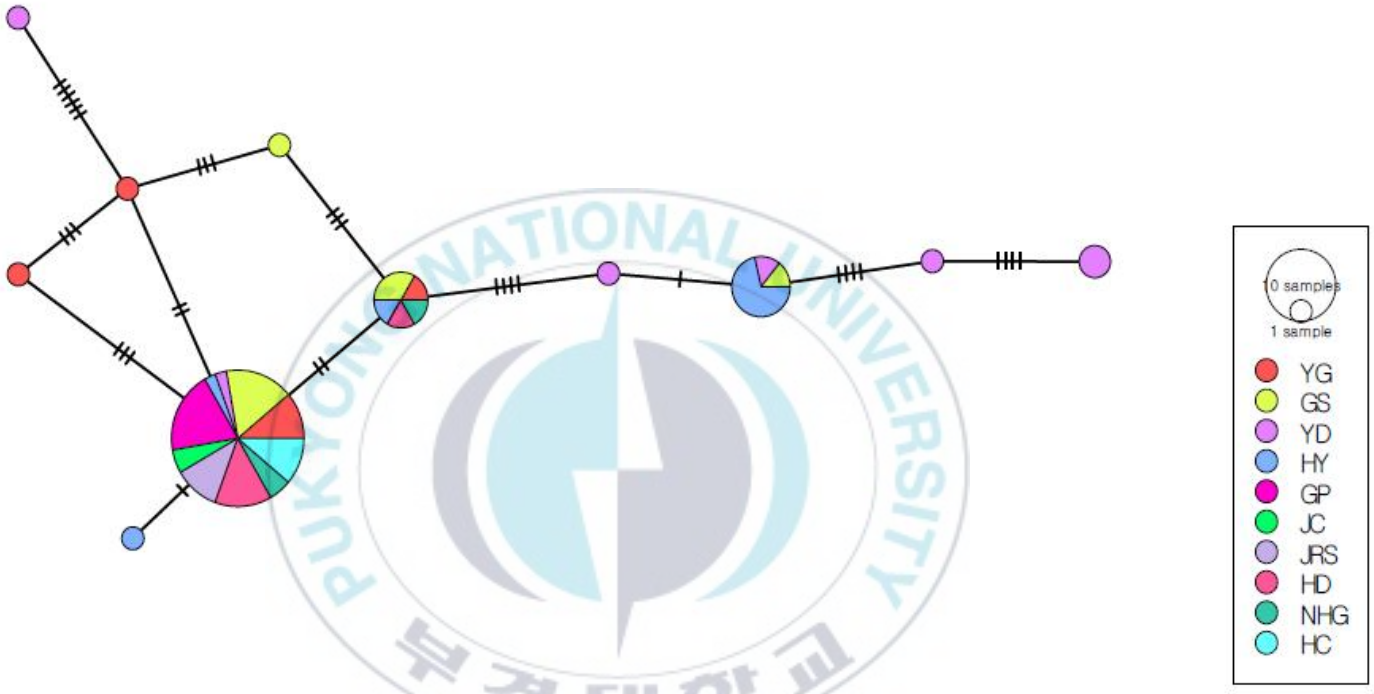

Figure 4. Minimum spanning network analysis of 11 haplotypes inferred from mtDNA markers of *S. coreana* using PopART v. 1.7. Each circle represents a different haplotype and each color represents a different locality.

Table 2. AMOVA within and between populations of *S. coreana* from 10 localities

Source of variation	*d.f.*	Sum of squares	Variance components	Percentage of variation (%)
Among populations	9	44.282	0.65922	36.1
Within populations	48	56.051	1.16698	63.9
Fixation	Index	FST	:	0.36098

Table 3. Pairwise F_{ST} values for mtDNA markers among 10 populations of *S. coreana*

	YG	GS	YD	HY	GP	JC	JRS	HD	NHG	HC
YG	0.000									
GS	-0.007	0.000								
YD	0.431*	0.372*	0.000							
HY	0.493*	0.328*	0.209	0.000						
GP	0.047	0.114	0.541*	0.620*	0.000					
JC	-0.265	-0.154	0.310	0.436	0.000	0.000				
JRS	-0.055	0.024	0.443*	0.539*	0.000	0.000	0.000			
HD	-0.050	-0.038	0.468*	0.521*	0.028	-0.304	-0.081	0.000		
NHG	-0.131	-0.220	0.321	0.377	0.300	-0.200	0.111	-0.227	0.000	
HC	-0.055	0.024	0.443*	0.539*	0.000	0.000	0.000	-0.081	0.111	0.000

Significant *P* values are indicated * $P < 0.05$

4.3 Genetic diversity and population structure of *S. forticosta*

The total number of specimens analyzed was 63 for COI, 53 for ND4, and 53 for COI+ND4. In total, 12, 7, and 25 mtDNA haplotypes were obtained from the COI, ND4, and COI+ND4 sequences for the three *S. forticosta* populations, respectively (Table 4). Haplotype diversity was typically high, ranging from 0.529 to 0.949. The YS population had higher haplotype diversity and nucleotide diversity than the other populations based on the COI+ND4 sequence (h = 0.949; π = 0.00301). The YS population had the second-highest haplotype diversity based on all genetic markers (0.636–0.927).

Neutrality testing was performed to determine the variation in population sizes. Tajima's D test based on the ND4 gene yielded a positive value for all populations except UG ($P < 0.05$). However, Fu's Fs values differed significantly among the markers ($P < 0.05$). Additionally, Fu's Fs values were negative for all markers in all populations, indicating expansion.

.

Table 4. Genetic diversity based on mtDNA markers of *S. forticosta* populations from each sampling location in the ROK

Genetic marker	Locality	n	N	*h*	π	Fu's *F*s	p-value	Tajima's *D*	p-value
mtDNA COI	SC	11	3	0.636	0.00159	-20.1139	0	0.19903	0.678
	YS	34	5	0.752	0.00236	-29.5068	0	0.24986	0.646
	UG	18	3	0.529	0.00297	-28.2902	0	0.49002	0.735
Genetic marker	Locality	n	N	*h*	π	Fu's *F*s	p-value	Tajima's *D*	p-value
mtDNA ND4	SC	11	4	0.709	0.00359	-13.3052	0	1.01828	0.859
	YS	24	5	0.75	0.00357	-27.3754	0	0.2424	0.646
	UG	18	4	0.542	0.00265	-28.5045	0	-0.36166	0.382
Genetic marker	Locality	n	N	*h*	π	Fu's *F*s	p-value	Tajima's *D*	p-value
Combined mtDNA	SC	11	7	0.927	0.00262	-10.9659	0	0.81983	0.789
(COI+ND4)	YS	24	14	0.949	0.00301	-26.4359	0	0.19367	0.634
	UG	18	4	0.542	0.0028	-23.3523	0	0.03237	0.543

n, number of specimens; N, number of haplotypes; h, haplotype diversity; π, nucleotide diversity

The minimum spanning parsimony network based on combined mtDNA A minimum-spanning parsimony analysis was performed based on the mtDNA haplotypes of three *S. forticosta* populations (Fig 5). There was no dominant haplotype in which the network contained all populations, and haplotypes with only two populations were observed. Also, the networks were clustered into three distinct groups: SC, YS, and UG. Most of the UG groups were branching out based on the haplotype of SC.

AMOVA showed that the variation among populations (54.82%) was higher than that within populations (45.18%). The overall F_{ST} value (0.54815) indicated genetic differentiation among SC, YS, and UG (Table 5). Pairwise comparison revealed a high degree of genetic structure among the populations (Table 6). The three pairwise comparisons yielded significant F_{ST} values (0.314–0.624). The YS and UG populations had the highest genetic differentiation, and the YS and SC populations the lowest.

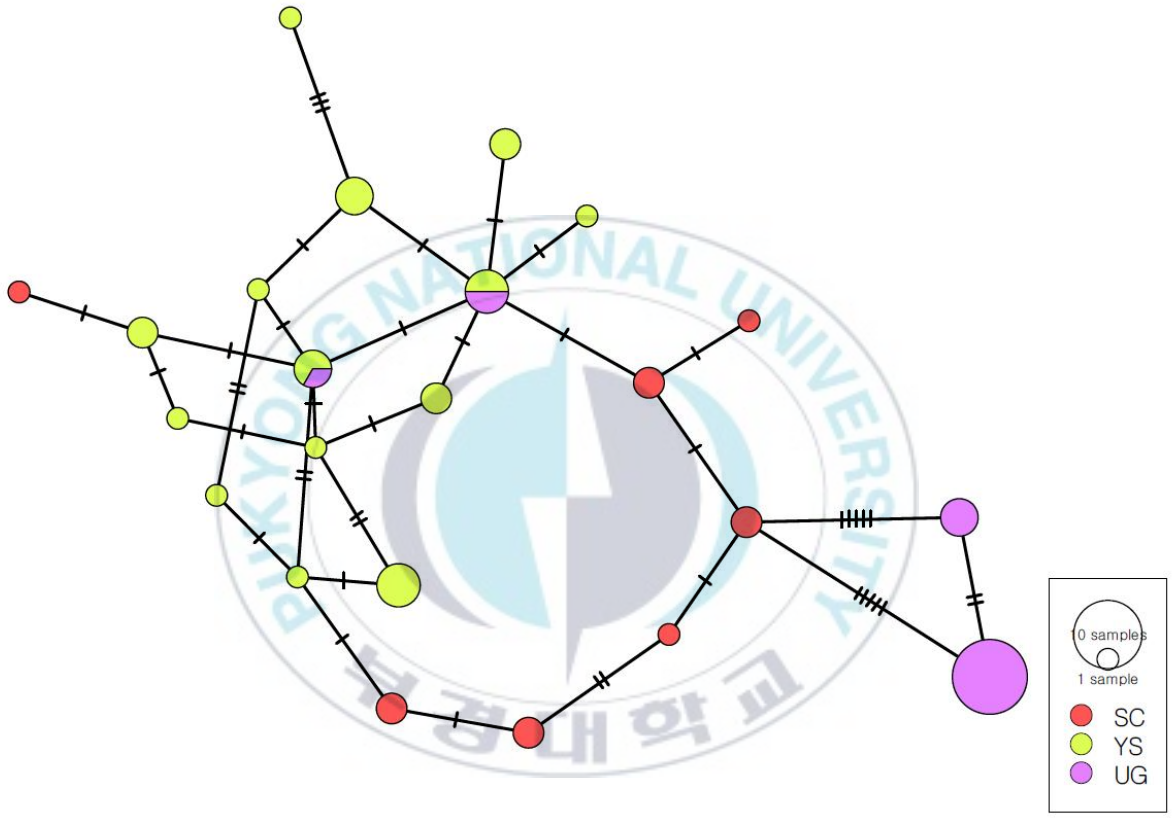

Figure 5. Minimum spanning network analysis of 14 haplotypes inferred from mtDNA markers of *S. forticosta* using PopART v. 1.7. Each circle represents a different haplotype and each color represents a different locality.

Table 5. AMOVA within and between populations of *S. forticosta* from three locality

Source of variation	*d.f.*	Sum of squares	Variance components	Percentage of variation (%)
Among populations	2	58.168	1.6439	54.82
Within populations	50	67.754	1.35508	45.18
Fixation	Index	FST	:	0.54815

Table 6. Pairwise F_{ST} values for mtDNA markers among among three populations of *S. forticosta*

	SC	YS	UG
SC	0.000		
YS	0.314*	0.000	
UG	0.550*	0.624*	0.000

Significant *P* values are indicated * $P < 0.05$

4.4 Genetic diversity and population structure of *S. gottschei*

The *S. gottschei* specimens consisted of eight populations (IJ, YP, GP, JC, JRS, HD, NHG, and HC) were 208 of COI, 183 of ND4, and 183 of COI+ND4. In total, 26, 21, and 40 mtDNA haplotypes were obtained from the COI, ND4, and COI+ND4 sequences, respectively, for the eight populations of *S. gottschei*. The number of haplotypes in each group was highest for COI, with 11 in the GP population, followed by 10 in JRS and HC, 7 in HD, 5 in NHG and YP, followed by in JC and IJ. There were nine ND4 haplotypes in the HD population, eight in JRS and HC, five in YP, four in GP, and in JC and NHG; IJ contained only one haplotype. Finally, the number of COI+ND4 haplotypes was in the HC population, 12 in JRS and HD, 9 in YP, 7 in GP, 6 in NHG, 4 in JC, and 3 in IJ (Table 7). The haplotype diversity and nucleotide diversity of COI were higher in the GP population than in the other populations (h = 0.868; π = 0.0054). For ND4 and COI+ND4, the HC population had the highest haplotype diversity and nucleotide diversity (h = 0.775; π = 0.00275 and h = 0.876; π = 0.00374, respectively).

Fu's *F*s and Tajima's *D* test were performed to assess demographic history (Table 7). Tajima's *D* values were negative for all populations based on all genetic markers, except for the HC population based on COI. However, based on all genetic markers, most populations differed from the extreme expansion model ($P < 0.05$), with the exceptions of IJ based on COI; JRS, HD, and NHG based on ND4; and JC, JRS, and NHG based on COI+ND4. Fu's *F*s test yielded negative or zero values.

Table 7. Genetic diversity based on mtDNA markers of *S. gottschei* populations from each sampling location in the ROK

Genetic marker	Locality	n	N	*h*	π	Fu's *F*s	p-value	Tajima's *D*	p-value
mtDNA COI	IJ	9	3	0.417	0.001	-18.1762	0	-1.3624	0.09
	YP	15	5	0.628	0.0034	-22.872	0	-0.53759	0.316
	GP	26	11	0.868	0.0054	-26.6507	0	-0.94202	0.18
	JC	31	4	0.243	0.0015	0	0	-1.54988	0.037
	JRS	33	10	0.695	0.0034	-27.6268	0	-1.14212	0.123
	HD	35	7	0.639	0.0037	-27.4415	0	-0.70772	0.282
	NHG	25	5	0.523	0.0013	0	0	-1.18867	0.109
	HC	36	10	0.837	0.0048	-26.8799	0	0.02113	0.562
Genetic marker	Locality	n	N	*h*	π	Fu's *F*s	p-value	Tajima's *D*	p-value
mtDNA ND4	IJ	9	1	0	0	0	1	0	1
	YP	15	5	0.733	0.0027	-24.9154	0	-0.53194	0.321
	GP	12	4	0.455	0.0013	-24.1621	0	-1.17901	0.09
	JC	24	3	0.236	0.0008	0	0	-1.25609	0.059
	JRS	32	8	0.528	0.0021	-30.9699	0	-1.67863	0.035
	HD	30	9	0.687	0.002	0	0	-1.78687	0.01
	NHG	25	2	0.08	0.0003	0	0	-1.51406	0.035
	HC	36	8	0.775	0.0028	-28.1736	0	-1.1452	0.13

Table 7. (*Continued*)

Genetic marker	Locality	n	N	*h*	π	Fu's *F*s	p-value	Tajima's *D*	p-value
Combined mtDNA (COI+ND4)	IJ	9	3	0.417	0.0005	-18.1762	0	-1.3624	0.1
	YP	15	9	0.8	0.003	-16.784	0	-0.59911	0.289
	GP	12	7	0.772	0.0024	-13.4749	0	-1.36573	0.086
	JC	24	4	0.308	0.0014	-28.4015	0	-1.55305	0.043
	JRS	32	12	0.774	0.0027	-26.5758	0	-1.58116	0.045
	HD	30	12	0.784	0.0028	-26.5274	0	-1.46812	0.063
	NHG	25	6	0.533	0.0008	0	0	-1.57299	0.033
	HC	36	13	0.876	0.0037	-26.0477	0	-0.62626	0.283

n, number of specimens; N, number of haplotypes; h, haplotype diversity; π, nucleotide diversity

The minimum spanning network was star-like, with a central haplotype and sporadic haplotypes connected by several mutational steps; the 40 haplotypes were separated into two major clusters (Fig 6). The first cluster consisted of specimens from all eight populations, including a central haplotype and unique haplotypes derived from the main haplotype. The predominant haplotype was present in all eight penguin colonies (48.6% of all individuals). The central haplotype was present at the highest frequency in the JC population, followed by NHG, JRS, HD, HC, and YP (1.12%). The second cluster consists of individuals from six (HC, HD, JRS, GP, YP, JC) of all population. The second cluster consisted of individuals from the HC, HD, JRS, GP, YP, and JC populations. The second most common haplotype was present in 7% of all individuals, and was shared among the HC, HD, JRS, and GP populations.

In addition, the network showed many unique haplotypes in each AMOVA showed that the variation among populations (11.73%) was lower than that within populations (88.27%) (Table 8). The overall F_{ST} value (0.117) showed that genetic differentiation among the eight populations was low. The pairwise F_{ST} values ranged from −0.024 to 0.397 (Table 9). The YP population showed significant differences from the other populations, while the HC population showed significant differences from the other populations with the exception of the HD population.

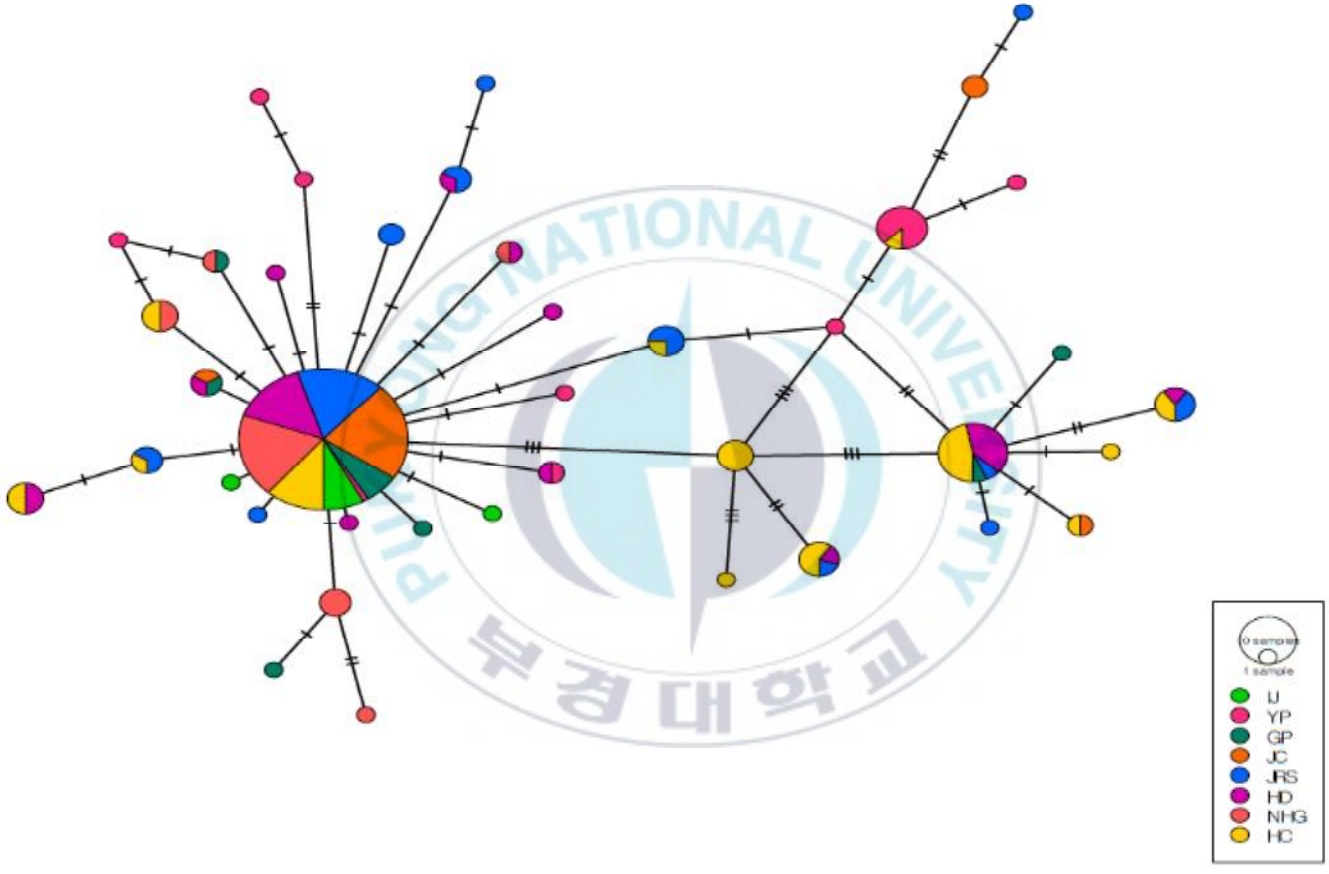

Figure 6. Minimum spanning network analysis of 40 haplotypes inferred from mtDNA markers of *S. gottschei* using PopART1.7. Each circle represents a different haplotype and each color represents a different locality.

Table 8. AMOVA within and between populations of *S. gottschei* from eight locality

Source of variation	*d.f.*	Sum of squares	Variance components	Percentage of variation (%)
Among populations	7	31.319	0.1498	11.73
Within populations	175	197.348	1.1277	88.27
Fixation	Index	F_{ST}	:	0.117

Table 9. Pairwise F_{ST} value for mtDNA markers among eight populations of *S. gottschei*

	IJ	YP	GP	JC	JRS	HD	NHG	HC
IJ	0.000							
YP	0.316*	0.000						
GP	0.023	0.193*	0.000					
JC	0.003	0.253*	-0.003	0.000				
JRS	0.024	0.174*	-0.021	0.007	0.000			
HD	0.063	0.183*	-0.024	0.041	-0.009	0.000		
NHG	0.018	0.397*	0.064	0.053*	0.080*	0.122	0.000	
HC	0.196*	0.168*	0.083*	0.176*	0.089*	0.042	0.261*	0.000

Significant P values are indicated * $P < 0.05$

4.5 Genetic diversity and population structure of *S. tegulata*

The 23 *S. tegulata* individuals in the DY, CHU, JRS, HC, and YS populations were analyzed, and the CHU population was found to have the largest number of specimens (Table 10). In total, 7, 11, and 15 mtDNA haplotypes were obtained from the COI, ND4, and COI+ND4 sequences, respectively, in the eight *S. forticosta* populations. The DY population had three haplotypes based on three genetic markers and the CHU population encompassed 6, 7, and 10 haplotypes based on the COI, ND4, and COI+ND4 sequences, respectively. The JRS population exhibited one, two and one haplotype based on the COI, ND4, and COI+ND4 sequences, respectively. In addition, the HC and YS populations displayed only one haplotype based on the COI, ND4, and COI+ND4 sequences. Haplotype diversity was high, except in populations with only one haplotype. The haplotype diversity of all genetic markers was higher in the CHU population than in the other populations (h = 0.846, 0.872 and 0.961, respectively), but nucleotide diversity was greatest in the DY population (π = 0.00548, 0.0048 and 0.00513, respectively).

Fu's *F*s and Tajima's *D* test were conducted to evaluate differences in population size (Table 10). The Tajima's *D* of the CHU population based on all genetic markers were non-significantly negative. The Fu's *F*s values of the DY and CHU populations based on all genetic markers were negative, but the value of the DY population based on the COI+ND4 sequence was not significant.

.

Table 10. Genetic diversity based on mtDNA markers of *S. tegulata* populations from each sampling location in the ROK

Genetic marker	Locality	n	N	*h*	π	Fu's *Fs*	p-value	Tajima's *D*	p-value
mtDNA COI	DY	4	3	0.833	0.00548	-1.51381	0.049	1.36522	0.862
	CHU	13	6	0.846	0.00345	-18.0438	0	-0.60752	0.289
	JRS	2	1	0	0	0	1	0	1
	HC	2	1	0	0	0	1	0	1
	YS	2	1	0	0	0	1	0	1
Genetic marker	Locality	n	N	*h*	π	Fu's *Fs*	p-value	Tajima's *D*	p-value
mtDNA ND4	DY	4	3	0.833	0.0048	-1.62218	0.044	0.6501	0.781
	CHU	13	7	0.872	0.00295	-19.1115	0	-1.3784	0.097
	JRS	2	2	1	0.00206	0	0.258	0	1
	HC	2	1	0	0	0	1	0	1
	YS	2	1	0	0	0	1	0	1
Genetic marker	Locality	n	N	*h*	π	Fu's *Fs*	p-value	Tajima's *D*	p-value
Combined mtDNA (COI+ND4)	DY	4	3	0.833	0.00513	-0.56837	0.196	1.06443	0.847
	CHU	13	10	0.961	0.00324	-12.7966	0	-1.11866	0.134
	JRS	2	2	1	0.00106	0	0.228	0	1
	HC	2	1	0	0	0	1	0	1
	YS	2	1	0	0	0	1	0	1

n, number of specimens; N, number of haplotypes; h, haplotype diversity; π, nucleotide diversity

The minimum spanning network of 15 haplotypes based on the COI+ND4 sequence consisted of many unique haplotypes (Fig 7). The predominant haplotype includes the DY, JRS, HC populations, and appears to be genetically distant from the other haplotypes. The CHU population had 10 unique haplotypes, and there were also haplotype structures that were unique to the DY group. The YS population had one haplotype that was not shared with any other population. The haplotype network suggested that the JRS and HC populations were genetically distant from the other local populations.

AMOVA showed that the variation among populations (43.58%) was lower than that within populations (56.42%). The overall F_{ST} value (0.436) indicated marked genetic differentiation among five populations (Table 11). The pairwise F_{ST} values ranged from −0.074 to 0.913. Importantly, the pairwise F_{ST} values between geographic localities were not significant for all populations, with the exceptions of the CHU and DY, CHU and JRS, and CHU and HC populations (Table 12).

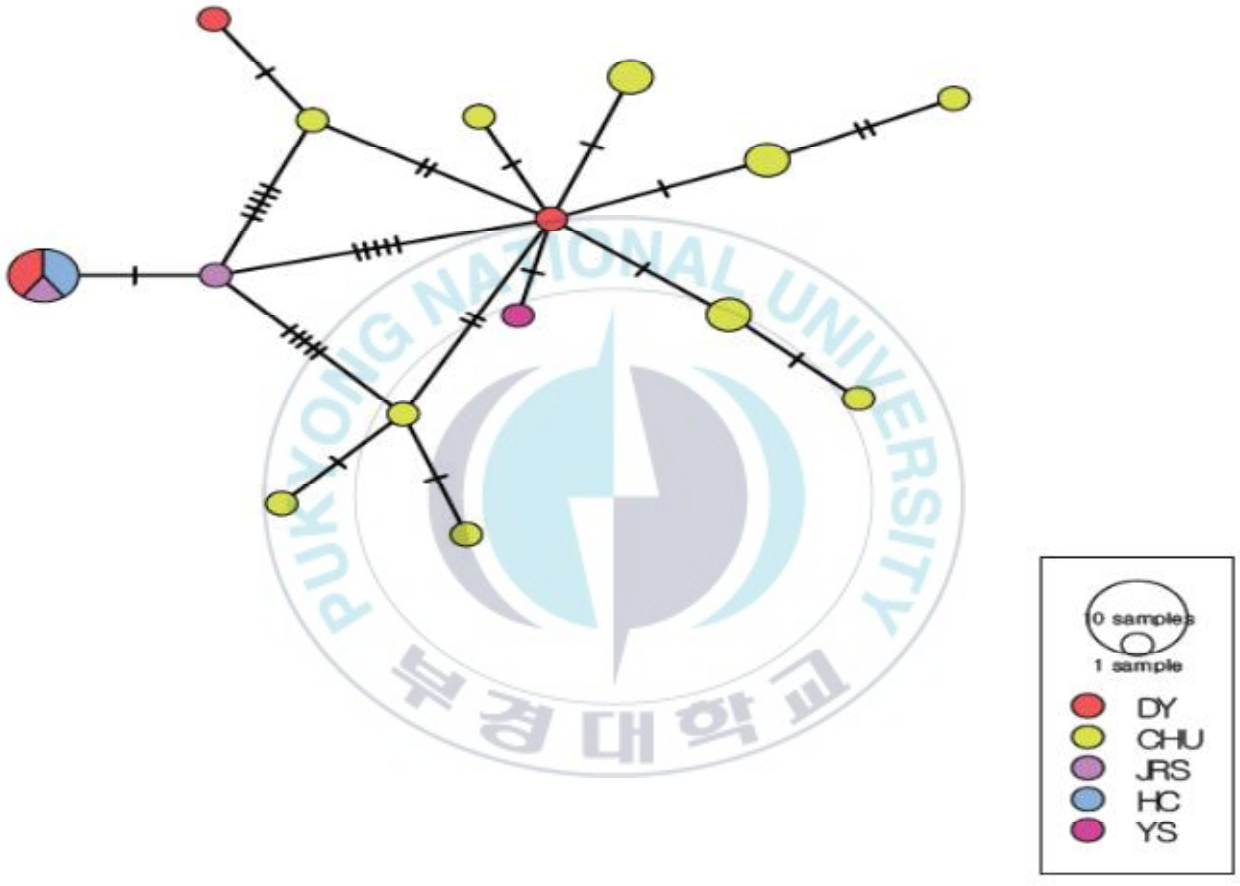

Figure 7. Minimum spanning network analysis of 15 haplotypes inferred from mtDNA markers of *S. tegulata* using PopART1.7. Each circle represents a different haplotype and each color represents a different locality.

Table 11. AMOVA within and between populations of *S. tegulata* from five locality

Source of variation	*d.f.*	Sum of squares	Variance components	Percentage of variation
Among populations	4	22.038	1.124	43.58
Within populations	18	26.188	1.455	56.42
Fixation	Index	FST	:	0.436

Table 12. Pairwise F_{ST} values for mtDNA markers among three populations of *S. tegulata.*

	DY	CHU	JRS	HC	YS
DY	0.000				
CHU	0.274*	0.000			
JRS	-0.074	0.587*	0.000		
HC	0.021	0.634*	0.000	0.000	
YS	0.303	0.143	0.824	0.913	0.000

Significant *P* values are indicated * $P < 0.05$

5. Discussion

5.1 Species composition of *Semisulcospira*

Several taxonomic studies of *Semisulcospira* have used mtDNA genes (Köhler 2016; Lee et al, 2007). We used the COI and ND4 genes to assess the genetic relationships of melania snails in the ROK (Fig 2 and 3). NJ analysis based on the COI sequence showed the presence of multiple species in the ROK (clades A–G). All individuals of clades A–F formed a monophyly in the natural population. In NJ analysis, the *S. tegulata* clade was sister to the *S. coreana* clade, in agreement with the isozyme relationships (Kim et al, 1995). Based on the mtDNA 16S rRNA and nDNA 28S rRNA, *S. coreana* and *S. tegulata* were not classified as sister clades (Lee et al, 2007). The nuclear gene exhibited varying degrees of incomplete sorting among the species, as was the case for *Alviniconcha* snails (Johnson et al, 2015). Previous studies did not classify *S. gottschei* as *Semisulcospira* due to insufficient bootstrap support (Kim et al, 2010). In this study, *S. gottschei* and *S. forticosta* formed sister clades, as also reported by Köhler (2016). In addition, the remaining *S. libertina* were separated with high statistical support as a sister clade to *Semisulcospira* (Fig 2). Kim et al (2010) reported cryptic and new species separated from the monophyly of *Semisulcospira*, and we identified two clades of presumed cryptic species. Clade C was a sister lineage to clades D and E (genetic distance 2.3–3.9%). To distinguish among species, intraspecific variation must be lower than interspecies variability (Hebert et al, 2003). Clade C was assumed to be a cryptic or new species, as its interspecific variability was markedly higher than its intraspecific variability. In addition, clade G was separated from the monophyly of

Semisulcospira (genetic distance 17.4–18.7%). Identification of melania snail species based on morphology is hampered by their marked variation in external characteristics and the existence of cryptic species (Lee et al, 2017). To overcome this, species identification can be performed using the COI, 16S rRNA, and cytochrome b sequences. However, 16S rRNA and cytochrome b have low genetic variability, indicating similar intraspecies and interspecies genetic variability. Therefore, species identification should be carried out using novel genetic markers, such as ND4. We used the ND4 gene to identify five *Semisulcospira* species, and the interspecies genetic distance was similar to that based on the COI gene (Fig 3). The genetic information obtained from a single mtDNA marker is limited and biased; multiple genetic markers are needed to increase the resolving power (Gruenthal et al, 2007). Use of the COI and ND4 genes will improve population genetics research, including on the genetic diversity in melania snails, genetic differentiation between populations, and phylogenetic relationships.

5.2 Population genetic structure of *Semisulcospira*

We estimated the level of genetic diversity and historical gene flow among populations of four *Semisulcospira* species. The COI, ND4, and COI+ND4 sequences showed that the genetic diversity of the four species varies throughout the ROK. Nucleotide and haplotype diversity provide demographic information on the history of populations (Grant and Bowen 1998). A high level of haplotype diversity and low level of nucleotide diversity were observed in 18 populations, with the exception of the *S. coreana* YD population (Tables 1, 4, 7, and 10). Similar patterns have been reported for other snail species, such as *Lunella granulate* (Chiu et al, 2013) and *Succinea caduca* (Holland and Cowie 2007). A high level of haplotype diversity is associated with high gene flow among populations (Katsares et al, 2008). In contrast, low nucleotide diversity values indicate small differences between haplotypes (Khosravi et al, 2019), possibly due to expansion after a period characterized by a low effective population size (De Jong et al, 2011). The intraspecific variation of the four snail species in this study was very low (0.05–0.08%), suggesting small differences among haplotypes. In addition, negative Fu's *F*s and Tajima's *D* values imply excess low-frequency polymorphisms, possibly due to demographic expansion. Fu's *F*s and Tajima's *D* tests were performed to assess the demographic history of the populations. Negative values indicate excess rare haplotypes and nucleotide site variants under a neutral model of evolution, whereas positive values indicate a recent population bottleneck (Puvanasundram et al, 2017). The four species showed negative Fu's *F*s and Tajima's *D* values, except for populations with one haplotype, especially result of Fu's *F*s with a significant (Tables 1, 4, 7, and 10).

Fu's *F*s value is regarded as more powerful than Tajima's *D* value for detecting the signatures of past population expansions (Fu 1997). Population expansion is supported by a unimodal pattern of negative Fu's *F*s and Tajima's *D* values, together with a high level of haplotype diversity and low level of nucleotide diversity (Chen at al, 2004). Despite the concern over resource reduction in the ROK due to overharvesting, habitat destruction, and human consumption of the four *Semisulcospira* species, all four species showed population expansion.

In this study, minimum spanning networks based on the COI+ND4 sequence were constructed to investigate the intraspecific relationships of the four *Semisulcospira* species. The size of the circle is proportional to the number of specimens with that haplotype, and a single line connecting circles represents a one-base difference. The *S. coreana* and *S. gottschei* network consisted of the central haplotype included all penguin populations and several unique haplotypes (Fig 4 and 6). In the *S. coreana* network, each haplotype was located around the central haplotype, indicating a greater genetic distance between haplotypes compared to the other three species. The YD population had the lowest frequency (2.7%) of the main haplotype, while the rare haplotype had the most with four, compared to the other *S. coreana* populations. This is in agreement with the corresponding pairwise F_{ST} values (Table 3). The *S. gottschei* network was separated into 40 haplotypes, a larger number compared to the other three species. Also, the *S. gottschei* network formed two clusters, but it also showed many unique haplotypes. The frequency of the main haplotype in the YP population was 1.12% (Fig 6). The *S. forticosta* network showed that this species was the only one with clear regional genetic differences, and no obvious central haplotype. In addition, multiple unique haplotypes were detected at each location, and the UG

population was more genetically distinct than the SC and YS populations (Fig 5). For *S. tegulata*, no haplotype was common to all populations, and there were only three penguin colonies, which were genetically distant from the CHU population (Fig 7). Therefore, all populations of the four *Semisulcospira* species had a large number of unique haplotypes, indicating an excess of rare haplotypes and suggesting rapid genetic differentiation and population expansion. The minimum spanning network did not reveal any haplotypes restricted to specific geographic regions, except for *S. forticosta*, suggesting a weak phylogeographic structure. This is in accordance with the results of studies on other snail taxa (Nehemia et al, 2019).

The AMOVA, pairwise and overall FST values, based on the COI+ND4 sequences, indicated genetic differentiation of the *Semisulcospira* populations. AMOVA based on COI+ND4 indicated that 36.1, 11.73, and 43.58% of the overall variation was among populations, and 63.9, 88.27, and 56.42% was within populations, of *S. coreana*, *S. gottschei*, and *S. tegulata*, respectively. For *S. forticosta*, 54.82 and 45.18% of the variation was among and within populations, respectively (Tables 2, 5, 8 and 11), in agreement with the minimum spanning network results. This indicates low genetic variation among geographic locations, except for *S. forticosta* populations, and suggests weak phylogeographic structures. In addition, the small inter-population variation values indicate restricted gene flow among regional populations. Inter-population genetic differentiation is affected by the blocking of gene flow by physical and physiological barriers, and by genetic drift and reduced species abundance (Jacquemyn et al, 2014). Wright et al (1965) reported that pairwise F_{ST} values of 0.05–0.15, 0.15–0.25, and > 0.25 indicate slight, considerable, and large genetic

differentiation among populations, respectively. In this study, the average F_{ST} value was 0.139 for *S. coreana*, 0.106 for *S. gottschei*, 0.496 for *S. forticosta*, and 0.381 for *S. tegulata*, indicating high genetic variation and low genetic communication for *S. forticosta* and *S. tegulata*. The populations of melania species from different river systems showed significant genetic differentiation due to limited gene flow, which is in accordance with the findings of other studies (Hou et al, 2013; McGlashan and Hughes 2000). However, only *S. forticosta* showed limited gene flow among populations and high pairwise F_{ST} values (0.314–0.624). The level of gene flow among local populations of the three other species was high, although distinct and rare haplotypes existed in different regions. The pairwise F_{ST} values of *S. coreana* ranged from 0.209 to 0.541 and 0.328 to 0.620 between the YD population and the other populations, and between the HY population and the other populations, respectively, suggesting a longer evolutionary history than the other regional populations. The pairwise F_{ST} values between the YP population and the other populations of *S. gottschei* were 0.168–0.253. In addition, the pairwise F_{ST} values between the CHU population and the other populations of *S. tegulata*, and between the YS population and the other populations of *S. tegulata*, were 0.143–0.634 and 0.303–0.913, respectively. These high genetic differentiation and low gene flow values are likely attributable to ecological factors, such as the sedentary lifestyle of the snail (Gu et al, 2015). However, the genetic distance and geographic distance of most populations of the two species were not correlated, indicating active gene flow. This was likely caused by transport by waterfowl, headwater capture, tectonic-driven changes in drainage patterns (Pfenninger et al, 2011) or marked larval dispersal (An et al, 2012). Similar results have been reported for other snail species, such as *Bulinus*

truncatus (Zein-Eddine et al, 2017) and *Juga* (Strong and Whelan 2019). Domestic drainage typically involves relatively short streams—with the obvious exceptions of the Geum River, Yeongsan/Seomjin River, Nakdong River, and Han River—and the populations in those streams showed various degrees of geographical genetic differentiation (Jang et al, 2016; Bae and Suk 2015). The Korean inland endemic fish, *Iksookimia pacifica*, was collected in streams draining into the East Sea and clustered into a northern group and a southern group; however, gene flow was active in only one population in each of the northern and southern groups. This population was presented wahlund effect as the result of the change in biological distribution due to artificial disturbance by release the seed as a result of the wall effect (Jang et al, 2017). Melania snails must be released into the same water system from which their mothers were collected to prevent artificial disturbance of the ecosystem (Cho et al, 2013). However, *S. coreana* and *S. gottschei* collected nationwide had the lowest geographic genetic differentiation; this finding warrants further investigation.

We analyzed the genetic diversity and population structure of four *Semisulcospira* species using mtDNA markers. In summary, a high level of haplotype diversity and low level of nucleotide diversity were detected among the 18 populations of *Semisulcospira*. Although the previous reports mentioned drastic decline in over-harvesting in Korea drainage due to over-harvesting has not resulted in low genetic diversity. Although the populations are expanding, this species requires management and conservation because of its restricted distribution, endemism, and consumption by humans. In addition, although *S. forticosta* clearly showed geographical genetic differentiation, *S. tegulata*, *S. coreana* and *S. gottschei* indicated only a small amount of geographical genetic

differentiation, we attempt to present conservation and management measures by verifying the genetic structural analysis of population through the nuclear gene, the MS marker, in the Chapter 4.

6. Reference

Dudgeon D., Arthington A.H., Gessner M.O., Kawabata Z.I., Knowler D.J., Lévêque C., et al. (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. Biological reviews, 81, 163-182.

Darwall W.R. T. and Vié J.C. (2005) Identifying important sites for conservation of freshwater biodiversity: extending the species-based approach. Fisheries. Manag. Ecol., 12, 287-293.

Strayer D.L. and Dudgeon D. (2010) Freshwater biodiversity conservation: recent progress and future challenges. J. North. Am. Benthol. Soc., 29, 344-358.

Anderson S. (1994) Area and endemism. The Quarterly Review of Biology, 69, 451-471.

Régnier C., Fontaine B., Bouchet P. (2009) Not knowing, not recording, not listing: numerous unnoticed mollusk extinctions. Conservation Biology, 23, 1214-1221.

Kang J.H., Yu K.H., Kim S.K., Park J.Y., Kim B.S., An C.M. (2010) Species identification and genetic structure of *Octopus minor* from Korea and China on the basis of partial sequences of mitochondrial cytochrome oxidase I. Korean. J. Malacol., 26, 285-290.

Chiu Y.W., Bor H., Kuo P.H., Hsu K.C., Tan M.S., Wang W.K., Lin H.D. (2017) Origins of *Semisulcospira libertina* (gastropoda: semisulcospiridae) in Taiwan. Mitochondrial DNA Part A, 28, 518-525.

Li W.H., Gojobori T., Nei M. (1981) Pseudogenes as a paradigm of neutral evolution. Nature, 292, 237-239.

Pan B., Wang Z., He X. (2011) Studies on assemblage characteristics of macrozoobenthos in the West River. Acta Hydrobiologica Sinica, 35, 851-856.

Ko J.H., Lee J.S., Kwon O.K. (2001) Study on radulae of seven species of the family Pleuroceridae in Korea. Korean. J. Malacol., 17, 105-115.

Kim D.H., Bang I.C., Lee W.O., Baek J.M. (2012) Reproductive Ecology of the Freshwater

Melania Snail, *Semisulcospira coreana* (v. Martens) in Bukhan River. Korean. J. Malacol., 28, 175-185.

Jung Y., Park G.M., Park J., Kim J.J., Min, D.Y., Chung P.R. (1999) Allozyme variability in three species of genus *Semisulcospira* (Prosobranchia: Pleuroceridae). Korean. J. Malacol., 15, 13-20.

Lee T., Hong H.C., Kim J.J., Foighil D.Ó. (2007) Phylogenetic and taxonomic incongruence involving nuclear and mitochondrial markers in Korean populations of the freshwater snail genus *Semisulcospira* (Cerithioidea: Pleuroceridae). Mol. Phylogenet. Evol., 43, 386-397.

Kim W.J., Kim D.H., Lee J.S., Bang I.C., Lee W.O., Jung H.T. (2010) Systematic Relationships of Korean Freshwater Snails of *Semisulcospira*, *Koreanomelania*, and *Koreoleptoxis* (Cerithiodiea; Pleuroceridae) revealed by Mitochondrial Cytochrome Oxidase I Sequences. Korean. J. Malacol., 26, 275-283.

Köhler F. (2016) Rampant taxonomic incongruence in a mitochondrial phylogeny of *Semisulcospira* freshwater snails from Japan (Cerithioidea: Semisulcospiridae). J. Molluscan. Stud., 82, 268-281.

Kang Y.J., Kim Y.H., Hong Y.K., Park J.Y., Park K.Y. (1996) A population genetic analysis of the common squid, *Todarodes pacificus* Steenstrup in the Korean waters. J. Kor. Fish. Soc., 29, 320-331.

Dunham R.A. (2004) Aquaculture and fisheries biotechnology. CABI Publishing press, Wallingford, pp. 104-109.

Hillis D.M., Mabel B.K., Moritz C. (1996) Applications of molecular systematics: The state of the field and a look to the future. In: Molecular Systematics. Sinauer. Sunderland. 515-543.

Zhu D., Jamieson B.G., Hugall A., Moritz C. (1994) Sequence evolution and phylogenetic signal in control-region and cytochrome b sequences of rainbow fishes (Melanotaeniidae). Mol.

Biol. Evol., 11, 672-683.

Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., & Hebert, P. D. (2005). DNA barcoding Australia's fish species. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 360, 1847-1857.

Doosey M.H., Bart Jr H.L., Saitoh K., Miya M. (2010) Phylogenetic relationships of catostomid fishes (Actinopterygii: Cypriniformes) based on mitochondrial ND4/ND5 gene sequences. Mol. Phylogenet. Evol. 54(3), 1028-1034.

An H.S., Jee Y.U., Park D.W., Ryu H.Y., Min K.S. (2000) Genetic variation in populations of the Pacific oyster (*Crassostrea gigas*), based on the mitochondrial COI gene sequence. Korean. J. Genet., 22, 249-255.

Arnaud S., Monteforte M., Galtier N., Bonhomme F., Blanc F. (2000) Population structure and genetic variability of pearl oyster *Pinctada mazatlanica* along Pacific coasts from Mexico to Panama. Cons. Genet., 1, 299-308.

Kim I.S., Yoon M.H., Min B.Y., Kim D.H. (2000) Gene flow of the blue mussel (*Mytilus edulis*) species complex in Korea detected by mitochondrial COIII gene sequences. Korean. J. Genet., 22, 305-318.

Lee S.Y., Kim S.H. (2003) Genetic variation and discrimination of Korean arkshell *Scapharca* species (Bivalvia, Arcoida) based on mitochondrial COI gene sequences and PCR-RFLP. Korean. J. Genet., 25, 309-315.

Matsumoto M. (2003) Phylogenetic analysis of the subclass *Pteriomorphia* (Bivalvia) from mtDNA COI sequences. Mol. Phylogenet. Evol., 27, 429-440.

Fu Y.X. (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics, 147, 915-925.

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

Kim J.J. (1995) Isozyme variations of the genus *Semisulcospira* (Pleuroceridae: Gastropoda) in Korea. Korean. J. Malacol., 11, 171-179.

Lee T., Hong H.C., Kim J.J., Foighil, D.Ó. (2007) Phylogenetic and taxonomic incongruence involving nuclear and mitochondrial markers in Korean populations of the freshwater snail genus *Semisulcospira* (Cerithioidea: Pleuroceridae). Mol. Phylogenet. Evol., 43, 386-397.

Johnson S.B., Warén A., Tunnicliffe V., Dover C.V., Wheat C.G., Schultz T.F., Vrijenhoek R.C. (2015) Molecular taxonomy and naming of five cryptic species of *Alviniconcha snails* (Gastropoda: Abyssochrysoidea) from hydrothermal vents. Systematics and Biodiversity, 13, 278-295.

Hebert P.D., Ratnasingham S., De Waard J. R. (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc.R. Soc. Lond. B. Supl., 270, 96-99.

Lee J.S. and Ko J.H. (2017) The study of external morphological characters of the family Pleuroceridae in Korea. Journal of the Environment, 12, 27-39.

Gruenthal K.M., Acheson L.K., Burton R.S. (2007) Genetic structure of natural populations of California red abalone (*Haliotis rufescens*) using multiple genetic markers. Mar. Biol., 152, 1237-1248.

Grant W.S., Bowen B.W. (1998) Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. J. Hered., 89, 415-426.

Chiu Y.W., Bor H., Tan M.S., Lin H.D., Jean C.T. (2013) Phylogeography and genetic differentiation among populations of the moon turban snail *Lunella granulata* Gmelin, 1791 (Gastropoda: Turbinidae). Int. J. Mol. Sci., 14, 9062-9079.

Holland B.S., Cowie R.H. (2007) A geographic mosaic of passive dispersal: population structure

in the endemic Hawaiian amber snail *Succinea caduca* (Mighels, 1845). Mol. Ecol., 16, 2422-2435.

Katsares V., Tsiora A., Galinou-Mitsoudi S., Imsiridou A. (2008) Genetic structure of the endangered species *Pinna nobilis* (Mollusca: Bivalvia) inferred from mtDNA sequences. Biologia, 63, 412-417.

Khosravi R., Malekian M., Hemami M.R., Silva T.L., Brito J.C. (2019) Low genetic diversity in the vulnerable Goitred Gazelle, *Gazella subgutturosa* (Cetartiodactyla: Bovidae), in Iran: potential genetic consequence of recent population declines. Zool. Middle. East., 65, 104-115.

De Jong M.A., Wahlberg N., Van Eijk M., Brakefield P.M., Zwaan B.J. (2011) Mitochondrial DNA signature for range-wide populations of *Bicyclus anynana* suggests a rapid expansion from recent refugia. PloS one, 6, e21385.

Puvanasundram P., Esa Y.B., Rahim K.A.A., Amin S.N. (2018) Phylogeography and population structure of *Tenualosa toli* inferred from Cytochrome b mitochondrial DNA fragment. Journal of Environmental Biology, 39, 895-906.

Chen C.A., Ablan M.C.A., McManus J.W., Bell J.D., Tuan V.S., Cabanban A.S., Shao K.T. (2004) Population structure and genetic variability of six bar wrasse (*Thallasoma hardwicki*) in northern South China Sea revealed by mitochondrial control region sequences. Mar. Biotechnol., 6, 312-326.

Nehemia A., Ngendu Y., Kochzius M. (2019) Genetic population structure of the mangrove snails *Littoraria subvittata* and *L. pallescens* in the Western Indian Ocean. J. Exp. Mar. Bio. Ecol., 514, 27-33.

Jacquemyn H., Honnay O., Galbusera P., Roldán-Ruiz I. (2004) Genetic structure of the forest herb *Primula elatior* in a changing landscape. Mol. Ecol., 13, 211-219.

Wright S. (1965) The interpretation of population structure by F-statistics with special regard to

systems of mating. Evolution, 19, 395-420.

Hou X., Zhu F., Yin S., Zhang L., Hu Y., Wang Y., et al. (2014) Genetic diversity of *Odontobutis potamophila* from different geographic populations inferred from mtDNA control region. Mitochondrial DNA, 25, 400-406.

McGlashan D.J., Hughes J.M. (2000) Reconciling patterns of genetic variation with stream structure, earth history and biology in the Australian freshwater fish *Craterocephalus stercusmuscarum* (Atherinidae). Mol. Ecol., 9, 1737-1751.

Gu Q. H., Cheng Q. Q., Li X. J., Zhou C. J. (2015) Novel polymorphic microsatellite markers for Bellamya and their application in population genetics of three species. Genet. Mol. Res., 14, 15201-15212.

Pfenninger M., Salinger M., Haun T., Feldmeyer B. (2011) Factors and processes shaping the population structure and distribution of genetic variation across the species range of the freshwater snail *Radix balthica* (Pulmonata, Basommatophora). BMC Evol. Biol., 11, 135.

An H.S., Lee J.W., Dong C.M. (2012) Population genetic structure of Korean pen shell (*Atrina pectinata*) in Korea inferred from microsatellite marker analysis. Genes & Genomics, 34(6), 681-688.

Zein-Eddine R., Djuikwo-Teukeng F.F., Dar Y., Dreyfuss G., Van den Broeck F. (2017) Population genetics of the Schistosoma snail host *Bulinus truncatus* in Egypt. Acta. Trop., 172, 36-43.

Strong E.E., Whelan N.V. (2019) Assessing the diversity of Western North American *Juga* (Semisulcospiridae, Gastropoda). Mol. Phylogenet. Evol., 136, 87-103.

Chapter IV.

Genetic diversity of *S. coreana* and *S. gottschei* using the microsatellite markers

Genetic diversity of *S. coreana* and *S. gottschei* using the microsatellite markers

1. Abstract

Information on genetic diversity is essential for assessing long-term population survival and adaptive evolution. Genetic markers such as microsatellites (MS) are used to study genetic diversity and population structure. A MS is a nuclear DNA marker that undergoes biparental inheritance and can be used to assess genetic variation within a species due to its highly polymorphic nature. While *Semisulcospira coreana* and *S. gottschei* are economically important mollusks, little is known about their within-population genetic diversity and between-population gene flow. In both species, the average expected heterozygosity (H_e) exceeded the observed heterozygosity (H_o) in all populations, indicating an excess of homozygotes; this often results from inbreeding. Most populations also had a positive inbreeding coefficient (F_{IS}). In *S. coreana*, analysis of molecular variance (AMOVA) indicated that 66.52% of the genetic variance occurred within sites, versus 8.78% among sites. Pairwise FST values showed that populations HY and YD had significant pairwise genetic differentiation compared with most other populations. Pairwise FST values, NJ and principal component analysis (PCA) analysis, and STRUCTURE identified two genetically distinct regions. In *S. gottschei*, AMOVA indicated that 82.77% of the genetic variance occurred within sites, compared with 3.46% among sites. The between-population genetic differentiation index was very low, as it was in the STRUCTURE analysis. This report provides critical information to support conservation strategies for melania snails based on MS analysis.

Key word: genetic population structure, microsatellite, genus *Semisulcospira*

2. Introduction

Knowledge of genetic diversity is essential for assessing the long-term survival of populations and potential for adaptive evolution (Pablos et al, 2015). Genetic diversity reflects the degree of adaptation and fitness, and is directly related to species survival (Desalle and Amato 2004). Genetic variation is reduced by genetic drift, limited gene flow, and inbreeding, and when the population size is small (Madsen 1999). Studies of genetic structure and diversity are vital for planning conservation strategies and identifying evolutionary significant units for biodiversity preservation (Harley et al, 2005). Mitochondrial genetic analysis can lead to erroneous conclusions about population structure and phylogeographic characteristics, where the genealogy may differ from the species history (Flanderset et al, 2009). To overcome this, nuclear markers, such as MS, can be used for analyzing population genetic structure because they have high allelic diversity and multiple independent loci (Bryja et al, 2010). Moreover, unlike mtDNA, MS can identify the genetic structure according to biparental inheritance, and can thus be used to assess genetic diversity.

The Ministry of the Environment has designated most melania snails in the genus *Semisulcospira* as Korean Peninsula or Asian endemic species Population genetic analysis is essential for management of *Semisulcospira*, as the species in this genus are considered to be endangered due to their being restricted to specific regions. Melania snails are considered a health food in Korea and play important roles in the monitoring freshwater systems (Lee and Lim 2005). Currently, the snail population is declining rapidly in Korea due to increasing human consumption, overharvesting, habitat degradation, and pollution caused by insecticides and the restructuring of riverbeds (Kim et al, 2010). Artificial seed

production has been promoted to facilitate recovery of decreasing resources. In the genus *Semisulcospira*, *S. coreana* is a major economic resource in inland waters, while *S. gottschei* is an Asian endemic species found in China and Korea; it is imported from China to meet the demand in Korea.

Therefore, this study used MS to examine the genetic diversity and population structure of *S. coreana* and *S. gottschei* in Korea, to obtain basic data to facilitate resource management.

3. Material and method

3.1 sample collection and DNA extraction

Of genus *Semisulcospira*, *S. coreana* and *S. gottschei* was collected from 10 and 8 localities, respectively. Total individuals of *S. coreana* and *S. gottschei* was 61 and 208, respectively, which coincide with the samples carried out in Chapter 3. Genomic DNA was extracted using the DNeasy Blood and Tissue® Kit (Qiagen, Valencia, CA, USA) following the manufacturer's instructions. Briefly, ~ 20 mg of tissue was lysed in Buffer ATL containing proteinase K (Qiagen) at 56℃ for 3 h. The lysates were mixed with Buffer TL and ethanol, and DNA was purified using a resin column. The concentration of genomic DNA was estimated using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

3.2 PCR amplification and genotyping

Eighteen MS loci for *S. coreana* were developed previously and assessed in terms of cross-species transferability (Chapter 2). After screening, nine SSR loci were chosen to analyze the population genetic structure of *S. coreana* and *S. gottschei* (Table 2). A fluorescent dye (FAM, TAMRA, or HEX) was attached to each forward primer at the 5'-end. Polymerase chain reaction (PCR) was performed on a total solution of 10 μL containing 1 μL of DNA, 0.25 U of Ex *Taq* DNA polymerase, 10× PCR buffer, 0.2 mM dNTP, and 10 pmol of each primer. PCR amplification was performed using an ABI 2720 Thermocycler (10 min at 95℃, 35 45-s cycles at 94℃, 45 s at 58℃, 45 s at 72℃, and a final 5-min

extension at 72℃). The size of the PCR products was estimated by reference to molecular size markers using an ABI PRISM 3730 sequencer.

Table 1. Characteristics of the microsatellite (MS) markers used for population analysis of *S. coreana* and *S. gottschei*

Locus	Forward primer sequence (5'→3')	Reverse primer sequence (5'→3')	Motif
Sc-02	TTGAGAGATACTCTGACACCCC	TAGGACCCACAGGACTAAACTG	(CA)7
Sc-06	GGAAAGTGATATGTGGAGGTGT	AACTTAACGTTCTGCTCTTTCG	(TG)8
Sc-19	AGCACCGAGGTAAAATGAGTTA	GGGTCTGGTCATGAGGTAAATA	(CT)7
Sc-26	TTGCAGAAAGAAATGTCAACAG	TAAAACAAGATGTGATGACCCA	(CT)9
Sc-27	CTAGCATTTGCTAATCTGACCC	AAGGTTGTTGCCTAACGTAGAA	(CAA)7
Sc-28	CATCAAAGAACACTCAAACATCA	ATACATGTGCGTAGTTGAATCG	(CAA)7
Sc-30	CTTGATCCACTACCCATGATCT	ACGAAATGAGTTCGAGGTTTTA	(TTG)7
Sc-36	CTGCTGGTGTTGTTTTCTGTTA	CCTGAAAGAGGCACATATTCTC	(TTG)9
Sc-53	TGAGTCTTATTTTCAAACGGAAA	TAGCAAGCCCCTTATGTGTAGT	(TGTG)8

3.3 Statistical analyses

The MS genotypes were determined using GeneMapper software (Applied Biosystems, UK). The number of alleles per locus, expected heterozygosity (*He*) and observed heterozygosity (*Ho*), inbreeding coefficient (F_{IS}), and departure from Hardy–Weinberg equilibrium (HWE) were estimated for each locus using GENEPOP software (Raymond and Rousset 1995) and allelic richness (AR) was calculated using FSTAT v 2.9.3.2 (Goudet 2001) to correct for the number of alleles. Micro-Checker v 2.2.3 was used to test for the presence of genotyping errors, including stuttering, allelic dropouts, and null alleles.

To assess the relationships between populations, principal component analysis (PCA) was performed using FSTAT (ver. 2.9.3.2), and ARLEQUIN 3.5 software (Excoffier & Lischer 2010) was used to calculate F_{ST} values. Values of F_{ST} were tested for significant deviation from zero using 10,000 permutations per run. Analysis of molecular variance (AMOVA) based on allele frequency was performed using ARLEQUIN 3.5.

Genetic differences and relationships among populations were examined with POPULATION 1.2.28 (http://www.cnrs-gif.fr/pge). Nei's genetic distances between all pairs of populations were calculated and a dendrogram was constructed using the unweighted pair group method with arithmetic mean (UPGMA) (1,000 bootstrap resampling across loci) (Langella 2007); the results were visualized using TreeView (http://rana.lbl.gov/EisenSoftware.htm).

Population genetic structure was assessed using STRUCTURE ver. 2.3.4 via a Bayesian model-based cluster analysis based on admixture and correlated allele frequencies models (Pritchard et al, 2000). For each value of K from 1 to 5, a 5,000-iteration burn-in period followed by 50,000 Markov Chain Monte Carlo

(MCMC) was used in each run. The most appropriate number of genetic clusters was derived using delta K (Evanno et al, 2005).

4. Results

4.1 Within-population genetic variation of *S. coreana*

All of the SSR loci were slightly polymorphic, and 91 alleles were detected at the nine SSR loci. Locus Sc-53 had the highest variation, and Sc-27 and Sc-36 the lowest. Table 2 summarizes the genetic diversity of the 10 *S. coreana* populations. No population samples were monomorphic across all loci, except population YD at Sc-27, with the number of alleles per locus in each population ranging from 2 to 10. The allelic richness varied from 1.00 to 5.46 for each sample. The lowest mean numbers of alleles were in populations JRS and NHG, whereas YG had the most. Population JC had the highest average allelic richness, while NHG had the lowest. The average Ho was highest in NHG (0.630), followed by JC and JRS (0.593), HD (0.522), GP (0.519), GS (0.506), HC (0.500), YD (0.488), HY (0.444), and YG (0.413). The genetic diversity was highest in population YD (H_e = 0.776) and lowest in YG (H_e = 0.413). After Bonferroni correction for multiple tests, only 5 of the 90 population–locus cases (10 populations × 9 loci) were in HWE (YG at Sc-27, GS at Sc-19, GS at Sc-19, YD at Sc-28, and HY at Sc-26; P<0.005), indicating heterozygote deficiencies. The mean F_{IS} values were positive, except for population NHG, which had a significant heterozygote deficiency at these loci.

Table 2. Variability in the *S. coreana* alleles at nine MS markers

Population		Microsatellite loci									Mean all loci
		Sc-02	Sc-06	Sc-19	Sc-26	Sc-27	Sc-28	Sc-30	Sc-36	Sc-53	
YG (N=7)	N	7	7	7	7	7	7	7	7	7	7
	Na	3	5	3	6	3	3	4	4	6	4
	AR	2.42	3.13	2.28	4.06	2.62	2.66	3.20	2.79	4.17	3.04
	R	195-203	206-248	231-263	261-289	146-152	191-200	294-309	208-217	134-170	
	H_o	0.429	0.429	0.143	0.857	0.000	0.143	0.286	0.571	0.857	0.413
	H_e	0.604	0.659	0.473	0.835	0.615	0.648	0.736	0.626	0.846	0.672
	F_{IS}	0.308	0.368	0.714	-0.029	1.000*	0.793	0.631	0.094	-0.014	0.429
	HW	0.440	0.183	0.021	0.272	0.002	0.011	0.023	1.000	0.959	0.323
GS (N=10)	N	10	10	10	9	10	10	10	10	10	10
	Na	3	7	9	7	2	8	6	4	8	6
	AR	1.98	4.15	4.17	3.78	1.99	4.03	3.58	2.50	4.26	3.38
	R	193-197	206-250	227-263	261-291	146-149	197-230	294-330	208-217	154-182	
	H_o	0.400	0.800	0.400	0.556	0.100	0.500	0.400	0.400	1.000	0.506
	H_e	0.353	0.853	0.837	0.778	0.521	0.826	0.758	0.500	0.847	0.697
	F_{IS}	-0.143	0.065	0.535*	0.298*	0.816	0.408	0.486	0.209	-0.192	0.276
	HW	1	0.17107	0.00026	0	0.01489	0.00773	0.01937	0.443	0.32788	0.220
YD (N=7)	N	7	7	7	7	7	7	7	6	7	7
	Na	3	6	6	2	1	5	4	3	9	4
	AR	2.77	4.10	4.28	1.99	1.00	3.94	3.55	2.87	5.18	3.30
	R	195-203	206-250	227-263	263-273	146	194-230	294-309	211-217	126-178	
	H_o	0.143	1.000	0.714	0.000	NA	0.286	0.429	0.333	1.000	0.488
	H_e	0.670	0.846	0.868	0.527	NA	0.835	0.802	0.712	0.945	0.776

Table 2. (*Continued*)

	F_{IS}	0.800	-0.200	0.189	1.000	NA	0.676*	0.486	0.556	-0.063	0.431
	HW	0.007	0.208	0.041	0.012	NA	0.000	0.025	0.204	0.091	0.074
	N	10	10	10	10	10	10	10	10	10	10
	Na	3	3	5	7	2	8	6	3	10	5
	AR	2.00	2.57	3.32	3.99	1.38	4.58	3.62	2.60	4.60	3.18
HY	R	195-203	208-212	229-249	263-291	146-152	194-239	294-330	211-217	130-190	
(N=10)	H_o	0.375	0.625	0.625	0.250	0.125	0.625	0.375	0.250	0.750	0.444
	H_e	0.342	0.608	0.742	0.800	0.125	0.892	0.767	0.633	0.867	0.642
	F_{IS}	-0.105	-0.029	0.167	0.702*	0.000	0.314	0.528	0.622	0.143	0.260
	HW	1.000	1.000	0.261	0.000	1.000	0.101	0.008	0.012	0.233	0.402
	N	7	7	7	7	7	7	7	*7	7	7
	Na	4	5	5	7	3	4	4	3	6	5
	AR	2.91	3.54	3.71	4.73	2.55	2.77	3.18	2.00	3.90	3.25
GP	R	195-205	206-244	239-253	261-291	146-152	191-206	294-306	208-214	134-178	
(N=7)	H_o	0.333	0.500	0.500	0.833	0.333	0.667	0.333	0.333	0.833	0.519
	H_e	0.636	0.742	0.803	0.909	0.545	0.561	0.712	0.318	0.803	0.670
	F_{IS}	0.500	0.348	0.400	0.091	0.412	-0.212	0.556	-0.053	-0.042	0.222
	HW	0.091	0.169	0.071	0.595	0.192	1.000	0.021	1.000	1.000	0.460
	N	3	3	3	3	3	3	3	3	3	3
	Na	3	5	2	4	2	5	4	3	5	4
JC	AR	3.00	5.00	2.00	4.00	2.00	5.00	4.00	3.00	5.00	3.67
(N=3)	R	195-199	206-258	233-249	263-299	146-149	191-224	300-330	208-214	150-202	
	H_o	0.667	1.000	0.000	1.000	0.000	0.667	0.667	0.667	0.667	0.593
	H_e	0.733	0.933	0.533	0.800	0.533	0.933	0.800	0.600	0.933	0.756

Table 2. (*Continued*)

	F_{IS}	0.111	-0.091	1.000	-0.333	1.000	0.333	0.200	-0.143	0.333	0.268
	HW	1.000	1.000	0.201	1.000	0.200	0.196	0.602	1.000	0.205	0.600
	N	4	4	4	4	3	4	4	4	4	4
	Na	2	3	4	5	3	3	4	2	5	3
	AR	1.75	2.93	3.68	4.00	3.00	2.71	3.46	2.00	4.21	3.08
JRS	R	195-203	206-210	241-249	263-291	146-155	194-200	294-309	211-214	134-174	
(N=4)	H_o	0.250	0.750	1.000	0.750	0.333	0.500	0.500	0.250	1.000	0.593
	H_e	0.250	0.714	0.821	0.786	0.600	0.607	0.750	0.536	0.857	0.658
	F_{IS}	0.000	-0.059	-0.263	0.053	0.500	0.200	0.368	0.571	-0.200	0.130
	HW	1.000	1.000	1.000	0.774	0.201	0.428	0.313	0.429	0.662	0.645
	N	6	6	6	6	6	6	6	5	6	6
	Na	3	4	3	7	3	4	3	2	5	4
	AR	2.41	3.18	2.76	4.38	2.55	2.77	2.27	1.97	3.92	2.91
HD	R	195-203	206-234	247-253	261-279	146-155	191-200	294-303	211-214	134-154	
(N=6)	H_o	0.667	0.667	0.667	1.000	0.000	0.667	0.167	0.200	0.667	0.522
	H_e	0.530	0.712	0.682	0.864	0.545	0.561	0.439	0.467	0.833	0.626
	F_{IS}	-0.290	0.070	0.024	-0.176	1.000	-0.212	0.643	0.600	0.216	0.208
	HW	1.000	0.343	1.000	0.867	0.010	1.000	0.091	0.333	0.555	0.577
	N	3	3	3	3	3	3	3	3	3	3
	Na	2	2	3	4	3	3	3	2	4	3
NHG	AR	2.00	2.00	3.00	4.00	3.00	3.00	3.00	2.00	4.00	2.89
(N=3)	R	195-203	208-210	247-251	263-273	146-152	191-200	294-303	208-211	134-186	
	H_o	0.667	0.667	0.667	1.000	0.333	1.000	0.333	0.667	0.333	0.630
	H_e	0.533	0.533	0.600	0.867	0.600	0.733	0.733	0.533	0.867	0.667

Table 2. (*Continued*)

	F_{IS}	-0.333	-0.333	-0.143	-0.200	0.500	-0.500	0.600	-0.333	0.667	-0.008
	HW	1	1	1	1	0.19754	1	0.19897	1	0.06725	0.718
	N	4	4	4	4	3	4	4	4	4	4
	Na	2	4	2	5	3	5	5	3	7	4
	AR	1.96	3.25	1.75	4.21	3.00	4.21	4.21	2.50	5.46	3.40
HC	R	195-197	208-246	231-249	263-289	146-152	191-206	294-309	208-214	134-190	
(N=4)	H_o	0.000	0.500	0.250	0.750	0.000	0.500	1.000	0.500	1.000	0.500
	H_e	0.429	0.643	0.250	0.857	0.800	0.857	0.857	0.464	0.964	0.680
	F_{IS}	1.000	0.250	0.000	0.143	1.000	0.455	-0.200	-0.091	-0.043	0.279
	HW	0.143	0.428	1.000	0.661	0.066	0.086	1.000	1.000	1.000	0.598
	N	6	6	6	6	6	6	6	6	6	6
	Na	3	4	4	5	3	5	4	3	7	4
	AR	2.32	3.39	3.09	3.91	2.31	3.57	3.41	2.42	4.47	3.21
Mean all	R	193-205	206-258	227-263	261-299	146-155	191-239	294-330	208-217	126-202	
pop.	H_o	0.393	0.694	0.497	0.700	0.136	0.555	0.449	0.417	0.811	0.521
	H_e	0.508	0.724	0.661	0.802	0.543	0.745	0.736	0.539	0.876	0.684
	F_{IS}	0.185	0.039	0.262	0.155	0.692	0.226	0.430	0.203	0.081	0.250
	HW	0.668	0.550	0.459	0.518	0.209	0.383	0.230	0.642	0.510	0.462

N_A, number of alleles; AR, allelic richness; H_O and H_E, observed and expected heterozygosities; F_{IS}, inbreeding coefficient; HWE, significant deviation from Hardy-Weinberg equilibrium after Bonferroni correction ($P < 0.005$).

4.2 Among population genetic variation of *S. coreana*

For all nine loci for all 10 populations, F_{IS} = 0.271 and F_{ST} = 0.0878, which indicate low levels of differentiation over all populations. The pairwise F_{ST} is used to determine the correlation between geographic distance and population differentiation. The estimated pairwise F_{ST} among *S. coreana* populations showed the lowest genetic distance (0.030) between populations JRS and YG, and the highest (0.183) between GS and JRS (Table 3). Population HY showed significant genetic differences from the other populations, except for YD. AMOVA showed that 66.52% of the genetic variance in *S. coreana* occurred within sites, versus 8.78% among sites and 24.7% among sites within populations (Table 4).

Table 3. Pairwise F_{ST} (below diagonal) and genetic distance (above the diagonal) for the 10 sample sites of *S. coreana*

F_{ST} D_{IS}	YG	GS	YD	HY	GP	JC	JRS	HD	NHG	HC
YG	-	0.398	0.453	0.516	0.281	0.401	0.248	0.262	0.377	0.433
GS	0.099*	-	0.478	0.387	0.516	0.368	0.822	0.655	0.462	0.398
YD	0.113*	0.118*	-	0.235	0.458	0.811	0.377	0.421	0.827	0.766
HY	0.146*	0.107*	0.051	-	0.608	0.761	0.594	0.572	0.603	0.494
GP	0.058	0.129*	0.113*	0.167*	-	0.657	0.182	0.120	0.340	0.565
JC	0.055	0.043	0.147	0.161*	0.120	-	0.769	0.668	0.509	0.362
JRS	0.030	0.183*	0.088*	0.167*	-0.004	0.134	-	0.127	0.522	0.767
HD	0.060	0.174*	0.115*	0.172*	-0.011	0.145*	-0.029	-	0.464	0.654
NHG	0.070	0.096	0.180*	0.153*	0.052	0.060	0.103	0.114	-	0.267
HC	0.097	0.082	0.174*	0.129*	0.133	0.023	0.179*	0.174*	0.005	-

Table 4. Analysis of molecular variance (AMOVA) within and between *S. coreana* populations from 10 localities

Source of variation	*d.f.*	Sum of squares	Variance components	Percentage of variation (%)
Among populations	9	65.05	0.29359	8.78
Among individuals within population	48	186.028	0.82572	24.7
Within individuals	58	129	2.2414	66.52

The genetic relationships among the populations are shown in the UPGMA tree based on Nei's genetic distance. The UPGMA tree separated the 10 populations into two clusters (Fig 1). Cluster 1 included six populations (YG, GP, HD, JRS, GS, JC, NHG, and HC) and Cluster 2 included the remaining four populations. Similarly, PCA based on D_{ce} values produced two clusters (Fig 2). The PCA of the 10 populations showed that the top three principal coordinates explained 94.84% of the genetic variation, with the first, second and third principal coordinates explaining 48.86, 35.01, and 10.97% respectively. The STRUCTURE analysis predicted that K = 3 would give a similar cluster pattern (Fig 3).

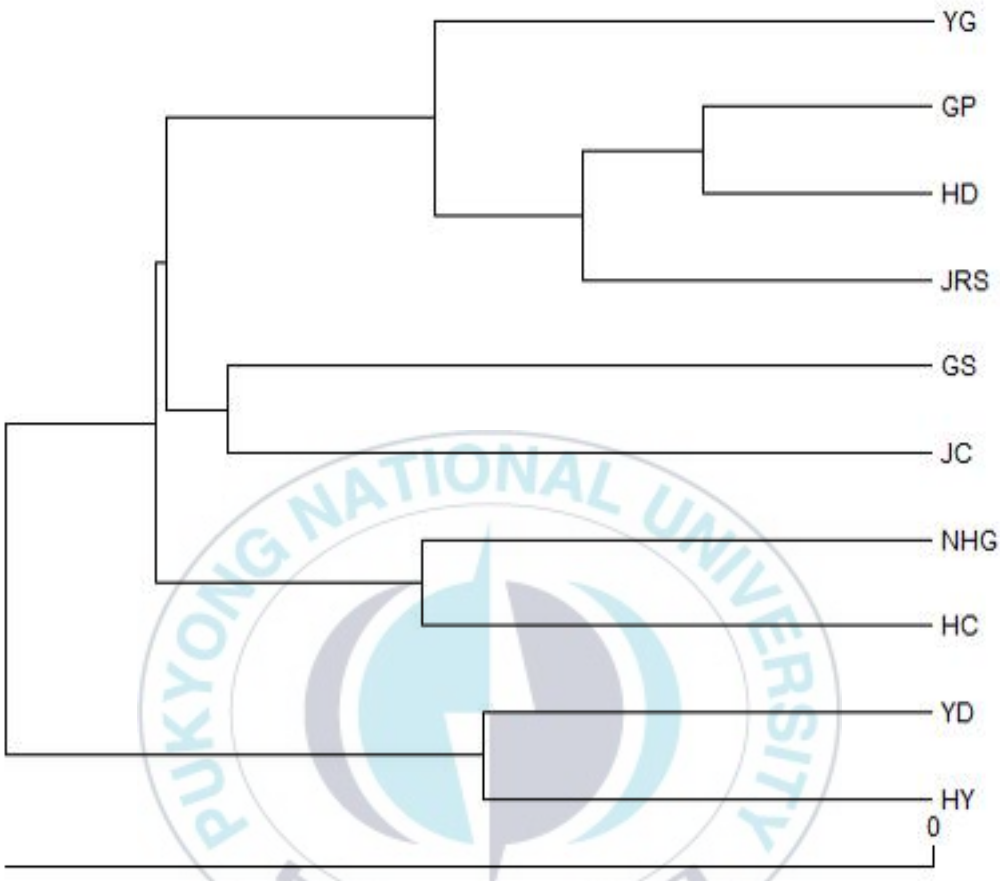

Figure 1. Unweighted pair group method with arithmetic mean (UPGMA) tree based on matrices of Nei's distances for nine MS markers.

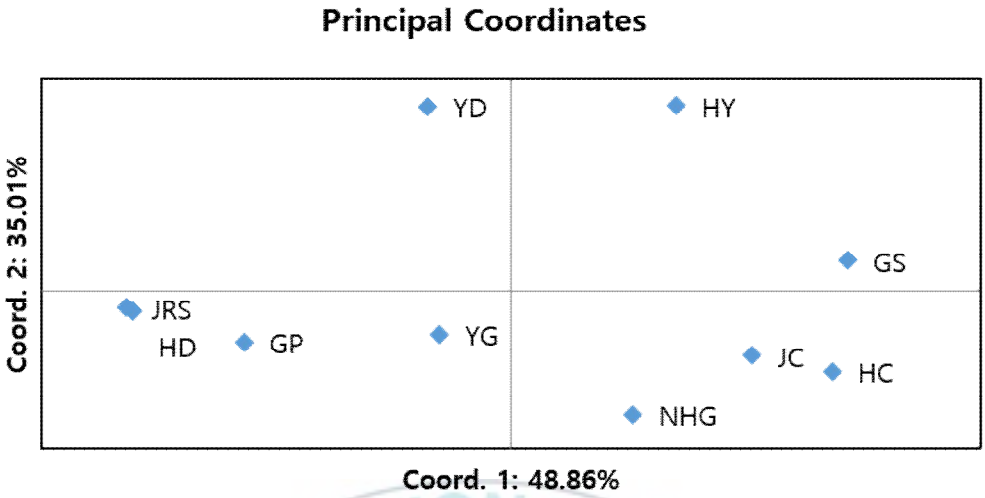

Figure 2. Principal coordinate analysis based on Nei's distances using nine MS loci from 10 populations.

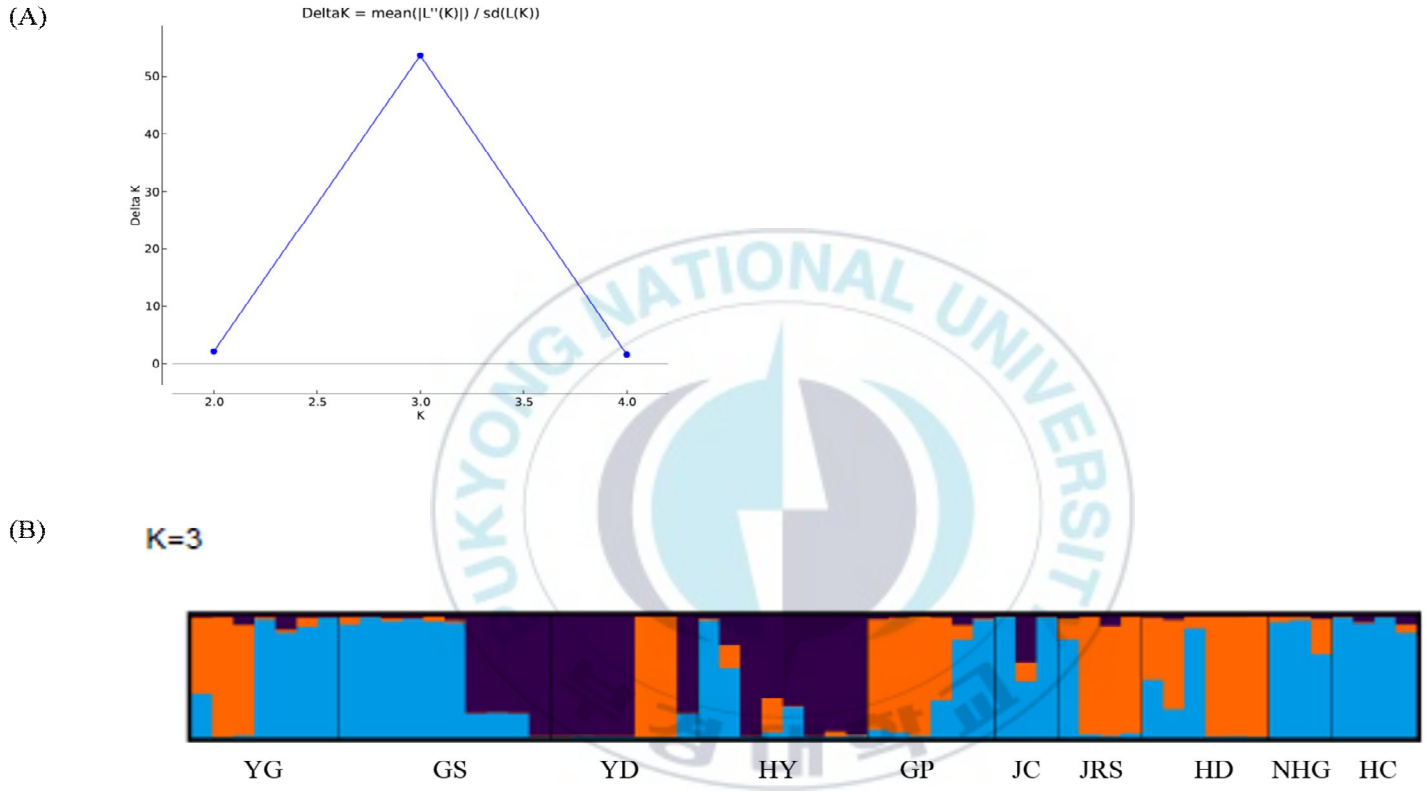

Figure 3. STRUCTURE results for nine polymorphic MS loci in *S. coreana*. (A) Population structure inferred by Bayesian clustering at K = 3. (B) The three colors in the column represent differences among clusters.

4.3 Within-population genetic variation of *S. gottschei*

All SSR loci were polymorphic, with 88 alleles detected at the nine SSR loci. Table 5 summarizes the genetic diversity of the eight *S. gottschei* populations. No population samples were monomorphic across all loci and the number of alleles ranged from 2 to 14. The allelic richness varied from 2.00 to 7.33 for each sample. Populations IJ and YP had the lowest mean number of alleles, whereas JC and HD had the most. Population JC had the highest average allelic richness, and YP the lowest. The average Ho was highest in NHG (0.610), followed by HD (0.559), HC (0.494), JC (0.487), JRS (0.465), GP (0.448), IJ (0.438), and YP (0.341). The genetic diversity was highest in population IJ (H_e = 0.619) and lowest in HD (H_e = 0.558). No loci deviated from HWE after applying the Bonferroni correction (P=0.05/9). However, He was greater than Ho at all SSR loci except for populations HD and NHG, indicating excessive homozygosity at these loci. Null alleles were present at most of MS loci, although populations HD and NHG had fewer null alleles than the other populations. The F_{IS} value was high in all populations except NHG.

Table 5. Variability in the *S. gottschei* alleles at nine MS markers

Population		Microsatellite loci									Mean all loci
		Sc-02	Sc-06	Sc-19	Sc-26	Sc-27	Sc-28	Sc-30	Sc-36	Sc-53	
IJ (N=9)	N	9	9	9	9	9	9	9	8	9	9
	Na	2	5	4	5	3	4	5	3	7	4
	AR	2.00	4.77	3.78	4.88	2.99	3.88	4.88	3.00	6.44	4.07
	R	199-203	204-216	245-253	261-271	146-155	194-212	291-309	208-214	134-190	
	Ho	0.222	0.444	0.667	0.556	0.222	0.222	0.333	0.500	0.778	0.438
	He	0.366	0.556	0.608	0.752	0.621	0.542	0.771	0.592	0.765	0.619
	FIS	0.407	0.210	-0.103	0.273	0.656	0.605	0.583	0.164	-0.018	0.309
	HW	0.3660	0.5556	0.6078	0.7516	0.6209	0.5425	0.7712	0.5917	0.7647	0.619
YP (N=15)	N	15	15	15	15	14	15	15	15	15	15
	Na	4	3	5	8	3	3	5	2	6	4
	AR	3.32	2.53	4.78	6.00	2.50	2.95	4.11	2.00	5.02	3.69
	R	195-203	206-248	241-263	261-291	146-152	197-206	294-336	211-214	134-186	
	Ho	0.333	0.267	0.333	0.467	0.071	0.333	0.067	0.533	0.667	0.341
	He	0.582	0.384	0.802	0.766	0.262	0.480	0.568	0.515	0.772	0.570
	FIS	0.435	0.313	0.593	0.399	0.735	0.314	0.886	-0.037	0.141	0.420
	HW	0.582	0.384	0.802	0.766	0.262	0.480	0.568	0.515	0.772	0.570
GP (N=26)	N	26	26	25	26	26	26	26	26	26	26
	Na	3	5	8	8	4	6	5	3	6	6
	AR	2.72	3.62	4.66	6.00	3.08	3.98	3.08	2.31	4.52	4.12
	R	195-203	206-214	241-263	261-279	140-155	191-224	291-309	211-214	134-162	
	Ho	0.346	0.577	0.560	0.577	0.115	0.346	0.308	0.423	0.654	0.448
	He	0.361	0.615	0.603	0.784	0.479	0.480	0.365	0.461	0.689	0.574

Table 5. (*Continued*)

	FIS	0.043	0.063	0.073	0.268	0.763	0.282	0.160	0.083	0.052	0.229
	HW	0.361	0.615	0.603	0.784	0.479	0.480	0.365	0.461	0.689	0.574
	N	31	31	31	31	31	31	31	31	31	31
	Na	3	7	9	12	6	5	6	4	9	7
	AR	2.68	4.83	5.66	6.62	4.28	3.83	3.51	2.90	6.30	4.51
JC	R	195-203	204-218	239-263	261-289	146-161	191-224	291-309	208-217	126-174	
(N=31)	Ho	0.323	0.645	0.581	0.742	0.194	0.581	0.129	0.484	0.710	0.487
	He	0.336	0.744	0.701	0.774	0.646	0.643	0.347	0.493	0.827	0.612
	FIS	0.040	0.134	0.174	0.042	0.704*	0.098	0.632*	0.020	0.144	0.221
	HW	0.496	0.224	0.226	0.431	0.000	0.104	0.000	0.728	0.014	0.247
	N	32	32	32	32	32	32	32	32	32	32
	Na	3	4	9	10	6	4	5	4	10	6
	AR	2.23	3.50	5.32	5.63	4.18	2.88	3.10	2.69	5.35	3.88
JRS	R	195-203	206-214	237-263	261-291	146-161	191-203	291-309	202-214	134-174	
(N=32)	Ho	0.344	0.813	0.719	0.594	0.125	0.406	0.156	0.469	0.563	0.465
	He	0.335	0.649	0.705	0.732	0.527	0.372	0.310	0.499	0.749	0.542
	FIS	-0.027	-0.256	-0.019	0.191	0.766*	-0.095	0.500*	0.061	0.252	0.153
	HW	1.000	0.481	0.455	0.076	0.000	0.341	0.001	0.048	0.021	0.269
	N	34	34	34	34	34	34	33	34	34	34
	Na	4	5	9	14	7	3	4	5	11	7
HD	AR	2.52	4.23	5.44	7.33	3.77	2.71	2.49	2.85	6.89	4.25
(N=34)	R	193-203	206-214	239-263	261-297	140-161	191-200	291-309	202-214	126-174	
	Ho	0.294	0.794	0.706	0.794	0.206	0.294	0.909	0.235	0.794	0.559
	He	0.265	0.690	0.705	0.814	0.445	0.389	0.537	0.371	0.809	0.558

Table 5. (*Continued*)

	FIS	-0.113	-0.154	-0.001	0.025	0.541*	0.247	-0.711*	0.368	0.019	0.025
	HW	1.000	0.675	0.771	0.223	0.000	0.103	0.000	0.005	0.737	0.390
NHG (N=25)	N	25	25	25	24	24	25	25	25	25	25
	Na	3	4	6	9	5	4	2	4	8	5
	AR	2.31	3.02	4.25	6.38	4.22	2.99	2.00	3.02	5.65	3.76
	R	195-203	204-210	239-263	259-277	140-161	191-200	294-297	211-220	134-170	
	Ho	0.320	0.720	0.720	0.833	0.333	0.360	1.000	0.480	0.720	0.610
	He	0.381	0.581	0.643	0.794	0.663	0.424	0.510	0.577	0.788	0.596
	FIS	0.163	-0.245	-0.122	-0.050	0.503*	0.155	-1.000*	0.171	0.088	-0.037
	HW	0.657	0.460	0.394	0.246	0.001	0.215	0.000	0.132	0.533	0.293
HC (N=36)	N	36	36	36	34	36	36	36	36	36	36
	Na	4	5	8	10	6	5	3	2	9	6
	AR	3.03	3.71	5.33	5.16	4.83	3.61	2.21	2.00	5.65	3.947
	R	193-203	206-214	241-263	261-281	140-161	191-224	291-297	211-214	126-174	
	Ho	0.500	0.444	0.583	0.559	0.361	0.472	0.306	0.389	0.833	0.494
	He	0.471	0.610	0.736	0.718	0.737	0.519	0.372	0.380	0.781	0.592
	FIS	-0.062	0.274	0.210	0.224	0.514*	0.092	0.181	-0.023	-0.069	0.149
	HW	0.040	0.029	0.053	0.109	0.000	0.144	0.480	1.000	0.033	0.210
Mean all pop.	N	26	26	26	26	26	26	26	26	26	
	Na	3	5	7	10	5	4	4	3	8	
	AR	2.60	3.77	4.90	6.00	3.73	3.35	3.17	2.59	5.73	
	R	193-203	204-248	245-263	259-297	140-161	191-224	291-336	202-220	126-190	
	Ho	0.335	0.588	0.609	0.640	0.203	0.377	0.401	0.439	0.715	
	He	0.387	0.603	0.688	0.767	0.548	0.481	0.473	0.486	0.772	

Table 5. (*Continued*)

FIS	0.111	0.042	0.101	0.172	0.648	0.212	0.154	0.101	0.076
HW	0.563	0.428	0.489	0.423	0.170	0.301	0.273	0.435	0.446

N_A, number of alleles; AR, allelic richness; H_O and H_E, observed and expected heterozygosities; F_{IS}, inbreeding coefficient; HWE, significant deviation from Hardy-Weinberg equilibrium after Bonferroni correction ($P < 0.005$).

4.4 Among population genetic variation of *S. gottschei*

For all nine loci for all eight populations, F_{IS} = 0.143 and F_{ST} =0.0346, indicating low differentiation over all populations. The estimated pairwise F_{ST} among populations of *S. gottschei* showed the lowest genetic distance (–0.001) between populations GP and JRS, and the highest (0.115) between YP and IJ, which also showed significant genetic differences from the other populations (Table 6). AMOVA showed that 82.77% of the genetic variance in *S. gottschei* occurred within sites, compared with 3.46% among sites and 13.78% among sites within populations (Table 7).

Table 6. Pairwise F_{ST} (below diagonal) and genetic distance (above the diagonal) for the 8 sample site of *S. gottschei*

F_{ST} D_{IS}	IJ	YP	GP	JC	JRS	HD	NHG	HC
IJ	-	0.306	0.166	0.182	0.138	0.188	0.122	0.190
YP	0.115*	-	0.134	0.147	0.118	0.181	0.169	0.160
GP	0.071*	0.066*	-	0.041	0.022	0.056	0.074	0.035
JC	0.061*	0.062*	0.012	-	0.034	0.086	0.102	0.036
JRS	0.057*	0.058*	-0.001	0.009	-	0.061	0.076	0.040
HD	0.085*	0.095*	0.027*	0.042*	0.032*	-	0.048	0.065
NHG	0.037*	0.077*	0.035*	0.043*	0.037*	0.018	-	0.068
HC	0.073*	0.074*	0.008	0.006	0.014	0.031	0.027*	-

Significant P values are indicated * $P < 0.05$

Table 7. Analysis of molecular variance (AMOVA) within and between *S. gottschei* populations from eight localities

Source of variation	*d.f.*	Sum of squares	Variance components	Percentage of variation (%)
Among populations	7	53.52	0.09208	3.46
Among individuals within population	200	587.646	0.36695	13.78
Within individuals	208	458.5	2.20433	82.77

The genetic relationships among populations are shown in the UPGMA tree based on Nei's genetic distance. The UPGMA tree separated the eight populations into three genetic clusters (Fig 4). Cluster 1 was population IJ, Cluster 2 was population YP, and Cluster 3 consisted of the remaining six populations. The PCA based on D_{ce} values also gave three clusters (Fig 5). The PCA of the eight populations showed that the top three principal coordinates explained 90.15% of the genetic variation, with the first to third explaining 42.89, 27.80, and 19.47%, respectively. However, STRUCTURE analysis predicted that K = 2 showed no genetic differentiation among the eight populations (Fig 6).

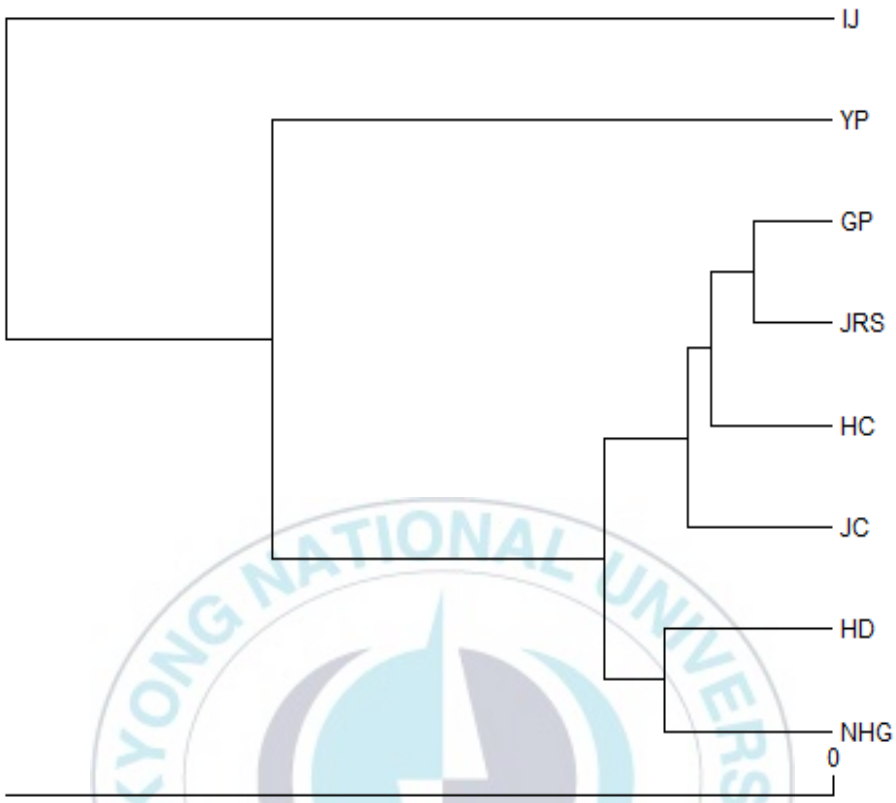

Figure 4. Unweighted pair group method with arithmetic mean (UPGMA) tree based on matrices of Nei's distances for nine MS markers.

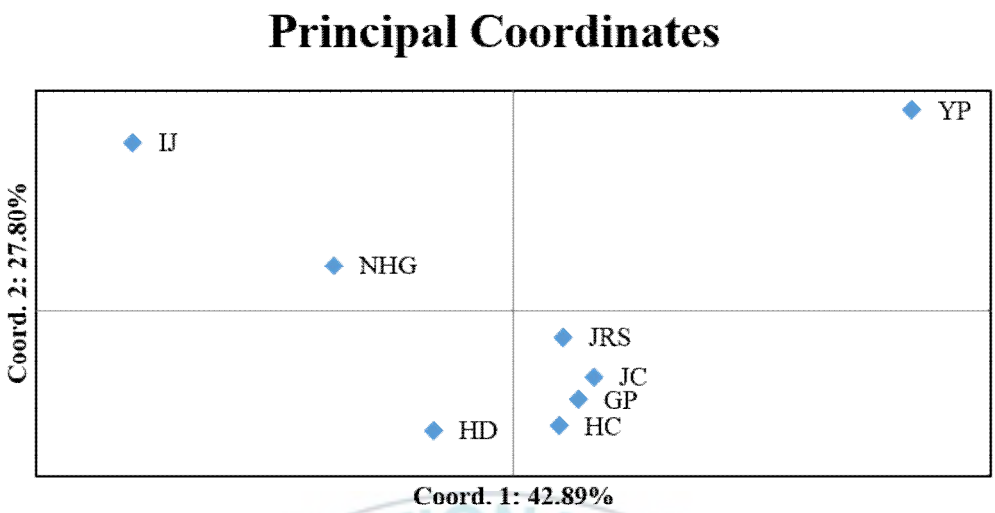

Figure 5. Principal coordinate analysis based on Nei's distances using nine MS loci from 8 populations.

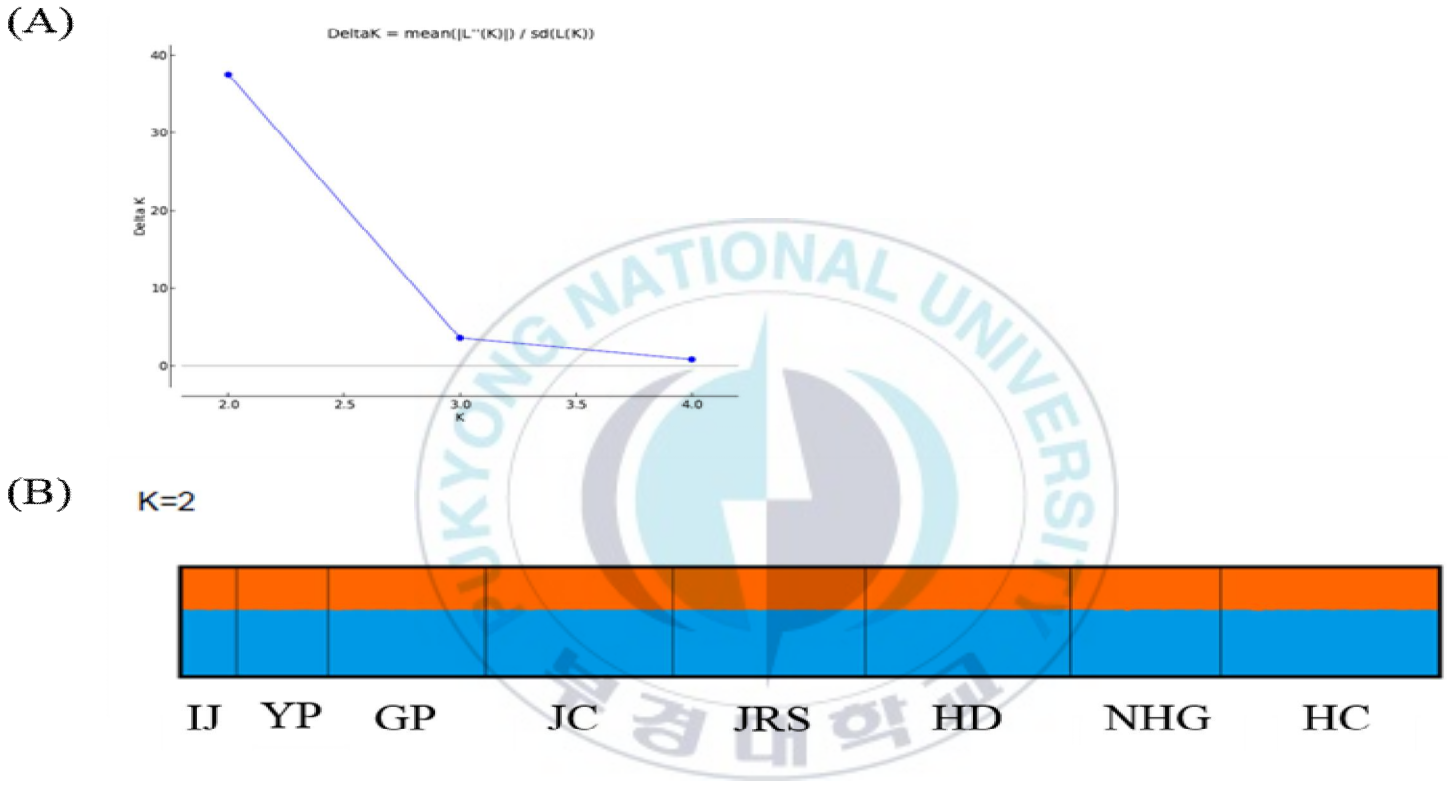

Figure 6. STRUCTURE results for nine polymorphic MS loci in *S. gottschei*. (A) Population structure inferred by Bayesian clustering at K = 2. (B) The three colors in the column represent differences among clusters

5. Discussion

Genetic markers, including mtDNA and nuclear DNA, show different patterns between males and females based on population structure and genetic variation, as revealed by analyses using two genetic markers, such as allozymes and MS (Prugnolle and De Meeˆus 2002). The technique of using markers with different genetic patterns to estimate population structures has been criticized, because there can be differences in mutations and effective population sizes between marker types even when the gene flow is similar between the sexes (Buonaccorsi et al, 2001). Estimates based on biparentally inherited markers are potentially more informative for evaluating population genetic structure (Goudet et al, 2002). This study estimated the level of genetic diversity and population structure, using MS markers based on biparental heredity, in two economically important species of the genus *Semisulcospira*. Our results suggest that the genetic diversity in *S. coreana* and *S. gottschei* is slightly higher (average allele number: 4.20 and 5.63; average H_e = 0.684 and 0.583, respectively) than that in the melania snails *Physa acuta* and *P. cubensis* (Sourrouille et al, 2003). Especially, Sc-26 and Sc-53 showed greater allelic diversity and higher heterozygosity in both *S. coreana* and *S. gottschei*. *In S.coreana*, YD population showed highest genetic diversity value on the results using both mtDNA and MS (Table 2; Chapter 3 Table 1). The average H_e was higher than H_o in all 10 populations, indicating a greater excess of homozygotes, which is typically the result of inbreeding. The F_{IS} can be used to quantify inbreeding within populations. The estimated F_{IS} for the 10 populations in this study ranged from –0.008 (NHG) to 0.431 (YD), with an average of 0.250. The high FIS values indicated that some of the loci in this species were homozygous, presumably resulting from genetic drift (Dixit et al, 2011).

The overall genetic differentiation among the 10 populations was low (F_{ST} = 0.0878, P <0.001), but significantly higher than zero. AMOVA indicated that 66.52% of the genetic variance occurred within sites, versus 8.78% among sites, which highly similar in their genetic compositions of *S. coreana* based on mtDNA analysis (Chapter 3, Table 2). The average the pairwise F_{ST} value was 0.104, showing slight genetic differentiation among populations. Nevertheless, HY and YD showed significant pairwise genetic differentiation compared with most other populations (Table 3). The pairwise F_{ST} values, UPGMA, PCA, and STRUCTURE all identified two genetically distinct regions. Although YD and HY were close geographically, they were not separated geographically from the other regions. However, *S. coreana* seemed to have relatively high genetic differentiation compared with the other molluscs.

In *S. gottschei*, population IJ had the highest genetic diversity among populations (average H_e = 0.619). On the other hand, the results from mtDNA showed high genetic diversity in HC population (Chapater 3, Table 7). Discord of results of the two types of marker could be interpreted as a result of sex-specific differences in dispersal (Teske et al, 2012). Indeed, the imbalance in sex ratio of male and female of melania snail appeared very severe (Kim et al, 2012), and this difference has also been reported in the study of another Gastropoda, *Glossaulax didyma* (Kim et al, 2007). No significant deviation from HWE was observed in any of the *S. gottschei* populations studied. However, the average of H_e exceeded H_o in six populations, except for NHG and HD, indicating a greater excess of homozygotes, which is typically a result of inbreeding. The estimated F_{IS} for the eight populations ranged from -0.037 (NHG) to 0.420 (YP), with an average value of 0.183. The positive F_{IS} values suggest significant heterozygote

deficiency in all populations, which is common in freshwater molluscs (Gu at al, 2012).

The overall genetic differentiation among the 10 populations was low (F_{ST} = 0.035, P <0.001), but significantly higher than zero. AMOVA indicated that 82.77% of the genetic variance occurred within sites, versus 3.46% among sites, which highly similar in their genetic compositions of *S. gottschei* based on mtDNA analysis (Chapter 3, Table 8). The average of the pairwise F_{ST} value was 0.045, showing slight genetic differentiation among populations. Significant genetic differentiation was found between populations IJ and YP versus the other populations, but the mean pairwise F_{ST} values were only 0.071 and 0.078, respectively. The results of mtDNA showed that the mean values of pairwise F_{ST} for IJ and YP were 0.123 and 0.235 respectively, higher than the MS results. This difference could be due to the fact that the mtDNA gene, whose effective population size is four times smaller, has a much faster and more sensitive effect on intergroup isolation than the nuclear gene (Birky et al, 1983). UPGMA and PCA identified three genetically distinct regions (IJ, YP, and all other populations). However, STRUCTURE analysis showed no genetic differentiation among the eight populations (Fig 6). This accords with the pairwise F_{ST} results, which showed active genetic flow among the eight *S. gottschei* populations.

A comparative analysis based on mtDNA and MS of economic mollusc *S. coreana* and *S. gottschei* found that the population expansion in both species was confirmed in mtDNA (Chapter 3), but the MS results showed that a lack of heterozygosity was suspected, and that a high degree of inbreeding. In addition, the analysis of MS showed no genetic differentiation between populations, indicating a low level of inter population genetic differentiation based on mtDNA.

Studies have reported that many endangered and endemic species have diverged in different water systems, necessitating independent management of each water system (ME, 2011; MLTM, 2010). These studies also showed regional genetic differences in *S. forticosta* and in part of *S. tegulata* and *S. coreana* (Chapter 3). Therefore, to manage melania snails, indiscriminate seed release in genetically distinct populations must be prohibited and preservation methods are needed for each population. The results of our analysis of the genetic flow, drift, diversity and population structure of the genus *Semisulcospira* will serve as basic data for resource management and conservation.

6. Reference

Pablos M.T.P., Saragusty J., Santiago-Moreno J., Stagegaard J., Göritz F., Hildebrandt T.B., Hermes R. (2015) Cryopreservation of onager (*Equus hemionus onager*) epididymal spermatozoa. J. Zoo. Wildl. Med., 46, 517-525.

DeSalle R., Amato G. (2004) The expansion of conservation genetics. Nat. Rev. Genet., 5, 702-712.

Madsen T., Shine R., Olsson M., Wittzell H. (1999) Conservation biology: restoration of an inbred adder population. Nature, 402, 34-35.

Harley E.H., Baumgarten I., Cunningham J., O'RYAN C. (2005) Genetic variation and population structure in remnant populations of black rhinoceros, *Diceros bicornis*, in Africa. Mol. Ecol., 14, 2981-2990.

Flanders J., Jones G., Benda P., Dietz C., Zhang S., Li G. et al. (2009) Phylogeography of the greater horseshoe bat, *Rhinolophus ferrumequinum*: contrasting results from mitochondrial and microsatellite data. Mol. Ecol., 18, 306-318.

Bryja J., Smith C., Konečný A., Reichard M. (2010) Range-wide population genetic structure of the European bitterling (*Rhodeus amarus*) based on microsatellite and mitochondrial DNA analysis. Mol. Ecol., 19, 4708-4722.

Lee S.M. and Lim T.J. (2005) Effects of dietary protein and energy levels on growth and lipid composition of juvenile snail (*Semisulcospira gottschei*). J. Shellfish Res., 24, 99-102.

Kim W.J., Kim D.H., Lee J.S., Bang I.C., Lee W.O., Jung H.T. (2010) Systematic Relationships of Korean Freshwater Snails of *Semisulcospira*, *Koreanomelania*, and *Koreoleptoxis* (Cerithiodiea; Pleuroceridae) revealed by Mitochondrial Cytochrome Oxidase I Sequences. Korean. J. Malacol., 26, 275-283.

Prugnolle F., and De Meeˆus T. (2002). Inferring sex-biased dispersal from population genetic

tools: a review. Heredity, 88(3), 161-165.

Buonaccorsi V.P., McDowell J.R., Graves J.E. (2001) Reconciling patterns of inter-ocean molecular variance from four classes of molecular markers in blue marlin (*Makaira nigricans*). Mol. Ecol., 10, 1179-1196.

Goudet J., Perrin N., Waser P. (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. Mol. Ecol., 11, 1103-1114.

Sourrouille P., Debain C., Jarne P. (2003) Microsatellite variation in the freshwater snail *Physa acuta*. Mol. Ecol. Notes., 3, 21-23.

Dixit S.P., Aggarwal R.A.K., Verma N.K., Vyas M.K., Rana J., Sharma A., Chander R. (2011) Genetic variability and bottleneck analyses of Kanniadu goat breed based on microsatellite markers. Indian. J. Anim. Sci., 81, 40-43.

Teske P.R., Papadopoulos I., Barker N.P., McQuaid C.D. (2012) Mitochondrial DNA paradox: sex-specific genetic structure in a marine mussel–despite maternal inheritance and passive dispersal. BMC Genet., 13, 45.

Kim D.H., Bang I.C., Lee W.O., Baek J.M. (2012) Reproductive Ecology of the Freshwater Melania Snail, *Semisulcospira coreana* (v. Martens) in Bukhan River. Korean. J. Malacol., 28, 175-185.

Kim D.G., Chung E.Y., Shin M.S., Hwang K. (2007) Reproductive ecology of the bladder moon, *Glossaulax didyma* (Gastropoda: Naticidae) in Western Korea. Korean. J. Malacol., 23, 189-198.

Gu Q., Xiong B., Zhu Y., Wang Q., Shi P. (2012) Isolation and characterization of polymorphic microsatellites loci in the freshwater snail *Bellamya aeruginosa* (Mollusca, Gastropoda). Conservation. Genet. Resour., 4, 387-390.

Burg T.M. and Croxall J.P. (2001) Global relationships amongst black-browed and grey-headed albatrosses: analysis of population structure using mitochondrial DNA and

microsatellites. Mol. Ecol., 10, 2647-2660.

Chapter V.

Population genetic structure of *Semisulcospira gottschei*: simultaneous examination of mtDNA and microsatellite markers

Population genetic structure of *Semisulcospira gottschei*: simultaneous examination of mtDNA and microsatellite markers

1. Abstract

Semisulcospira gottschei is an Asian domestic species inhabiting Korea and China. However, genetic structure analysis of the resource management of this species has not been performed. To investigate the genetic diversity among populations, microsatellites can be used to determine the geographic origins of marine and freshwater species. This study investigated the genetic structures of the Korean and Chinese populations of *S. gottschei* based on mitochondrial DNA (mtDNA) Cytochrome oxidase subunit I (COI) and polymorphic microsatellite loci developed from *Semisulcospira coreana*. Analysis of the mtDNA COI sequence revealed 43 haplotypes, which indicated no gene flow between the Korean and Chinese populations. To further elucidate the genetic structures of the Korean and Chinese populations, the population genetics of *S. gottschei* were analyzed using nine microsatellite markers. The genetic diversity analysis showed an average of 5.25 alleles per locus, with an average allelic richness of 4.02. Excessive homozygosity was found at all loci, which was expected to be due to the presence of null alleles at all loci. Populations of *S. gottschei* formed two separate clusters according to pairwise F_{ST} and AMOVA. Also, the UPGMA tree, PCA, STRUCTURE, and GeneClass indicated separation of the 11 populations into two clusters: Korea and China. These results have potential use in the management, restoration, and distinction of the origin country of populations.

Key words: *Semisulcospira*, Microsatellite, Populations, Country of origins

2. Introduction

The melania snail, genus *Semisulcospira*, is one of the most popular commercial gastropod resources in South Korea. In Korea, the melania snail is designated as an endemic species, and its population distribution is limited. As a result, the melania snail is recognized as an endangered species (www.me.go.kr). It is critical to investigate the genetic structure of endangered species populations to choose the relevant strategy for conservation management. (Sydes and Peakall 1998). Of the genus *Semisulcospira*, *Semisulcospira gottschei* is an endemic species of East Asia, but not the Korean Peninsula, found only in East Asia and is an important biological resource for biogeographical research.

The melania snail has become accepted as a salubrious food in Korea, and its domestic consumption continues to rise. Domestic consumption is ~2,000 tons per year, with the wild-stock catch reaching 700–1,000 tons. To meet the demands, over 1,000 tons of processed melania snails are imported from China (Moon et al, 2015). In case of different values of same or similar species in exporting and importing countries could cause serious economic problems. However, the differences in morphological characteristics among snails of different origins have not been identified, which can result in economic fraud when one seafood species is illegally substituted for another (Rasmussen et al, 2008). Therefore, a rapid, reliable and reproducible method for identification of the country of origin of the melania snail is necessary.

Genetic diversity estimation, a method of assessing the effectiveness of fishery resource management, is a very important indicator of the adaptation of biological populations to the environment and has become a crucial part of the management

of marine resources. Molecular markers have been widely used to characterize the population genetic structures of fish and shellfish. These markers include allozyme (Planes and Fauvelot 2002), mitochondrial DNA (mtDNA) (Roe et al, 2001), random amplified polymorphic DNA (RAPD) (Toro et al, 2004), amplified fragment length polymorphisms (AFLP) (Kong and Li 2007), and microsatellite (Gruenthal et al, 2007). Microsatellites (MS), single sequence repeats (SSRs) or short tandem repeats (STR), are very effective indicators that have been widely used to investigate genetic variation among fishery populations (Kang et al, 2014; Dong et al, 2106). Because of the high level of polymorphism, these techniques have been used in studies on population genetics in shared fishery stocks, which allow for gene flow among populations (Hauser et al, 2002). In addition, MS can be used to identify the genetic variation of each population, providing useful data for resource conservation (Allendorf et al, 2008). The geographic origins of marine and freshwater populations, such as east Asian giant scallop (An et al, 2009), wreckfish (Ball et al, 2000), noble crayfish (Gross et al, 2013), can potentially be differentiated by population genetic studies using MS. A multivalent trace metal analysis has been performed to determine the countries exporting fish such as tuna (Rooker et al, 2003) and eels (Yamashita et al, 2006). However, this method is only applicable to a limited number of target species.

Recently, genetic methods based on allele-specific PCR of mtDNA 16S rRNA genes have been developed to identify the origin of northern mauxia shrimp (Kang et al, 2015) and brackish water bivalves (Park et al, 2018). Also, a PCR-RFLP method has been developed to detect the origin of Japanese horse mackerel (Takashima et al, 2006). However, these methods are based on single nucleotide polymorphisms (SNPs), which require sequence information, including

differences between alleles, and the use of a primer with a 3' end to locate the SNP (Rahman et al, 2013).

Microsatellites were only recently developed from *S. coreana* to evaluate the genetic diversity and to assess the population genetic structure of *S. gottschei* in Korea and China. We also aimed to detect the country of origin, comparing the molecular genetic characteristics of Korean *S. gottschei* with those of a non-native population.

3. Materials and Methods

3.1 Sample collection and genomic DNA preparation

Genetic analysis was performed using a total of 326 individuals collected from 11 localities in Korea and China. Collection details are shown in Table 1. *S. gottschei* specimens were collected from across Korea (n=208). Packages of snails whose shells had been removed during food processing were purchased from a retail market in China in 2016 and 2017, and high-quality samples were used in the analysis. Genomic DNA was extracted using the QIAGEN DNeasy Blood and Tissue kit and kept at −20°C until used for analysis.

Table 1. Collection details for the *S. gottschei* samples

Taxa	Code	Country of origin	Sample size	Locality
S. gottschei	IJ	Republic of Korea	9	Inje-gun, Gangwon-do
	YP	Republic of Korea	15	Yangpyeong-gun, Gyeonggi-do
	GP	Republic of Korea	26	Gapyoung-gun, Gyeonggi-do
	JC	Republic of Korea	31	Jecheon-si, Chungcheongbuk-do
	JRS	Republic of Korea	32	Jiri mountain, Hamyang-gun, Gyeongsangnam-do
	HD	Republic of Korea	34	Hadong-gun, Gyeongsangnam-do
	NHG	Republic of Korea	25	Namhangang, Chungju-si, Chungcheongbuk-do
	HC	Republic of Korea	36	Hongcheon-gun, Gangwon-do
	CH1	China	48	
	CH2	China	24	
	CH3	China	26	

3.2 Species identification

For species identification, the COI gene of mtDNA was PCR amplified using universal primers: HCO2198 (5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3') and LCO1490 (5'- GGT CAA CAA ATC ATA AAG ATA TTG G-3'). The PCR mixture consisted of 1× PCR buffer, 250 μM each dNTP, 2.5 U Ex-Taq DNA polymerase (Takara), 20 ng DNA sample, and 10 pmol universal primers in a total volume of 20 μL. PCR amplification was performed using the ABI 2720 Thermocycler under the following conditions: 10 min denaturation at 95°C, 35 cycles of amplification (45 s at 95°C, 45 s at 56°C, and 1 min at 72°C), and a 5 min final extension at 72°C. The resulting PCR products were visualized by gel electrophoresis and purified using Expin PCR SV. The PCR products were sequenced using the Big Dye Terminator v3.1 Cycle Sequencing kit and the ABI 3730XL DNA analyzer.

3.3 PCR and genotyping

Eighteen microsatellite loci in *S. coreana* were previously developed and tested for cross-species transferability (Chapter II). After screening these loci, nine SSR loci were chosen for population genetic structure analysis in *S. gottschei* (Table 2). The fluorescent dye (FAM, TAMRA, HEX) was attached to each of the forward primers at the 5' end. PCR was performed in a 10 μL mixture containing 1 μL DNA, 0.25 U Ex-Taq DNA polymerase (Takara, Shiga, Japan), 10 x PCR buffer, 0.2 mM dNTP, 10pmol of each primer. PCR amplification was performed using the ABI 2720 Thermocycler under the following conditions: 10 min at 95℃; 35 cycles of 45 s at 94℃, 45 s at 58℃, and 45 s at 72℃; and a 5 min final extension at 72℃. The sizes of the PCR products were estimated according to a molecular size marker (GENESCAN 400 HD) on ABI PRISM 3730 sequencer (Applied Biosystems).

Table 2. Characteristics of microsatellite marker used for population analysis of *S. gottschei*

Locus	Forward primer sequence (5'→3')	Reverse primer sequence (5'→3')	Motif
Sc-02	TTGAGAGATACTCTGACACCCC	TAGGACCCACAGGACTAAACTG	(CA)7
Sc-06	GGAAAGTGATATGTGGAGGTGT	AACTTAACGTTCTGCTCTTTCG	(TG)8
Sc-19	AGCACCGAGGTAAAATGAGTTA	GGGTCTGGTCATGAGGTAAATA	(CT)7
Sc-26	TTGCAGAAAGAAATGTCAACAG	TAAAACAAGATGTGATGACCCA	(CT)9
Sc-27	CTAGCATTTGCTAATCTGACCC	AAGGTTGTTGCCTAACGTAGAA	(CAA)7
Sc-28	CATCAAAGAACACTCAAACATCA	ATACATGTGCGTAGTTGAATCG	(CAA)7
Sc-30	CTTGATCCACTACCCATGATCT	ACGAAATGAGTTCGAGGTTTTA	(TTG)7
Sc-36	CTGCTGGTGTTGTTTTCTGTTA	CCTGAAAGAGGCACATATTCTC	(TTG)9
Sc-53	TGAGTCTTATTTTCAAACGGAAA	TAGCAAGCCCCTTATGTGTAGT	(TGTG)8

3.4 Statistical analyses

The MS genotypes were determined using GeneMapper software (Applied Biosystems, UK). The number of alleles per locus, expected heterozygosity (*He*) and observed heterozygosity (*Ho*), inbreeding coefficient (F_{IS}), and departure from Hardy–Weinberg equilibrium (HWE) were estimated for each locus using GENEPOP software (Raymond and Rousset 1995) and allelic richness (AR) was calculated using FSTAT v 2.9.3.2 (Goudet 2001) to correct for the number of alleles. Micro-Checker v 2.2.3 was used to test for the presence of genotyping errors, including stuttering, allelic dropouts, and null alleles.

To assess the relationships between populations, principal component analysis (PCA) was performed using FSTAT (ver. 2.9.3.2), and ARLEQUIN 3.5 software (Excoffier & Lischer 2010) was used to calculate F_{ST} values. Values of F_{ST} were tested for significant deviation from zero using 10,000 permutations per run. Analysis of molecular variance (AMOVA) based on allele frequency was performed using ARLEQUIN 3.5.

Genetic differences and relationships among populations were examined with POPULATION 1.2.28 (http://www.cnrs-gif.fr/pge). Nei's genetic distances between all pairs of populations were calculated and a dendrogram was constructed using the unweighted pair group method with arithmetic mean (UPGMA) (1,000 bootstrap resampling across loci) (Langella 2007); the results were visualized using TreeView (http://rana.lbl.gov/EisenSoftware.htm).

Population genetic structure was assessed using STRUCTURE ver. 2.3.4 via a Bayesian model-based cluster analysis based on admixture and correlated allele frequencies models (Pritchard et al, 2000). For each value of K from 1 to 5, a

5,000-iteration burn-in period followed by 50,000 Markov Chain Monte Carlo (MCMC) was used in each run. The most appropriate number of genetic clusters was derived using delta K (Evanno et al, 2005). The geographical differentiation was explained using the assignment test, GeneClass2 (Piry et al, 2004).

4. Results

4.1 Genetic diversity of *S. gottschei* populations based on COI gene sequences

All 326 individuals were obtained 530 bp sequence of COI gene and a sequence similarity search performed using the GenBank BLASTn (http://www.ncbi.nlm.nih. gov/BLAST/). All individuals showed over 99% nucleotide sequence identity (Acc. No. KY675052.1) to *S. gottschei*. Sequence analysis of the COI gene from *S. gottschei* in Korea and China revealed 43 polymorphic site with sequence variation, which resulted in 40 haplotypes (Table 3). Haplotype analysis demonstrated that no SNP that could distinguish China from Korea, but the haplotype of the Korean and Chinese population appeared different.

Table 3. Variable nucleotide sites using COI analysis

	Polymorphic sites																																										
	16	52	73	76	80	93	106	107	115	121	127	136	148	149	154	164	175	199	208	220	235	245	247	252	253	256	272	289	292	295	298	304	325	328	349	368	399	412	427	430	433	451	454
Hap_1	T	T	C	C	C	T	A	G	C	T	G	A	T	C	T	C	C	T	A	C	A	G	T	C	A	T	G	T	A	G	C	G	C	C	A	T	T	T	T	T	G	T	T
Hap_2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_4	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
Hap_5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.
Hap_6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_7	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
Hap_8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.
Hap_9	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_10	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	T	C	.	.	.	.	.	.	.	.	.	.	.
Hap_12	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	C	.	T	.	.	.	.	.	.	.	C	.
Hap_13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
Hap_14	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
Hap_15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	C	.	.	.	.	.	.	.	.	.	.	.
Hap_16	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_17	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	T	C	.	.	G	.	.	.	.	.	.	.	.
Hap_18	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_19	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C
Hap_20	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_21	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_22	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	T	C	.	.	.	.	.	C	.	.	.	.	.
Hap_23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_24	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
Hap_25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.
Hap_26	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	C	T	.	.	.	.	.	.	.	.	.	.
Hap_27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	T	.	.	.	.	.	.	.	.	.	.	.	C
Hap_28	.	.	.	T	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
Hap_29	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	T	T	.	.	.	.	.	.	.	.	.	.	.	C
Hap_30	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
Hap_31	.	.	G	.	.	G	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	T	T	.	.	.	.	.	.	.	.	.	.	.	C
Hap_32	.	.	.	.	.	G	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	G	.	G	.	.	.	T	T	.	.	.	.	.	A	.	.	.	.	.	C
Hap_33	.	.	.	.	T	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	T	.	.	.	.	.	.	.	.	.	.	.	C
Hap_34	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	T	T	.	.	.	.	.	.	.	.	.	.	.	C
Hap_35	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
Hap_36	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	.
Hap_37	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	T	T	.	.	.	.	.	.	.	.	.	.	.	C
Hap_38	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	T	T	.	.	.	.	.	.	.	.	.	.	.	C
Hap_39	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	T	T	.	.	.	.	.	.	.	.	.	.	.	C
Hap_40	.	C	.	.	.	.	.	.	.	.	.	.	.	T	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	T	.	T	.	.	.	.	.	.	.	A	.	.

	Korea								China		
	IJ	YP	GP	JC	JRS	HD	NHG	HC	CH1	CH2	CH3
Hap_1	7	3	8	27	18	20	17	10			
Hap_2	1										
Hap_3	1										
Hap_4		9			1			1			
Hap_5		1									
Hap_6		1									
Hap_7		1	1								
Hap_8			1	1		1					
Hap_9			2				1				
Hap_10			3				4				
Hap_11			5		3	7		10			
Hap_12			2		1	1		3			
Hap_13			1								
Hap_14			1	2	1						
Hap_15			1					4			
Hap_16			1								
Hap_17				1				1			
Hap_18					1						
Hap_19					2	3		3			
Hap_20					2						
Hap_21					2			1			
Hap_22					1						
Hap_23						1	1				
Hap_24						1					
Hap_25							2	2			
Hap_26								1			
Hap_27									33	16	9
Hap_28									3		
Hap_29									1		1
Hap_30									3	3	7
Hap_31									1		
Hap_32									1		
Hap_33									5	3	2
Hap_34									1		
Hap_35											1
Hap_36											3
Hap_37											1
Hap_38											2
Hap_39										1	
Hap_40										1	
Number of polymorphic sites	2	6	13	6	10	9	4	9	18	13	10
Nucleotide diversity (%)	0.095	0.332	0.524	0.145	1.567	0.355	0.126	0.465	0.384	0.495	0.597
Number of haplotypes	3	5	11	4	10	7	5	10	8	6	8
Haplotype diversity	0.417	0.629	0.868	0.243	0.679	0.639	0.523	0.837	0.518	0.609	0.809

4.2 Genetic diversity of *S. gottschei* populations based on MS analysis

The nine polymorphic SSR markers were used to detect a total of 96 alleles among the 326 *S. gottschei* samples from both populations. All loci were highly polymorphic in most of the population. The genetic diversity and characterization of the SSR loci are presented in Table 4. The mean number of alleles was 5.25 (2–14), and the mean allelic richness value was 4.02 (2.0–7.4). The observed and expected heterozygosities ranged from 0.067 to 0.909 and from 0.250 to 0.826, with an average of 0.459 and 0.563, respectively. The mean genetic diversity was higher in the Korean groups (IJ, YP, GP, JC, JRS, HD, NHG, HC) than in the Chinese groups (CH1, CH2 and CH3) with H_e=0.583 and H_e=0.530, respectively. None of the loci deviated from HWE after Bonferroni correction (P=0.05/9). However, the H_e was greater than H_o at most SSR loci except in the HD and NHG populations, indicating excessive homozygosity at these loci. Also, null allele was presence at most of MS loci, especially the number of null allele of HD and NHG population was lower than other population. The F_{IS} value was high in most population except for NHG.

Table 4. Variability of alleles at 9 microsatellite markers

Population		Microsatellite loci									Mean all loci
		Sc-02	Sc-06	Sc-19	Sc-26	Sc-27	Sc-28	Sc-30	Sc-36	Sc-53	
IJ (N=9)	N	9	9	9	9	9	9	9	8	9	9
	Na	2	5	4	5	3	4	5	3	7	4
	AR	2.00	4.77	3.78	4.88	2.99	3.88	4.88	3.00	6.44	4.07
	R	199-203	204-216	245-253	261-271	146-155	194-212	291-309	208-214	134-190	
	Ho	0.222	0.444	0.667	0.556	0.222	0.222	0.333	0.500	0.778	0.438
	He	0.366	0.556	0.608	0.752	0.621	0.542	0.771	0.592	0.765	0.619
	FIS	0.407	0.210	-0.103	0.273	0.656	0.605	0.583	0.164	-0.018	0.309
	HW	0.3660	0.5556	0.6078	0.7516	0.6209	0.5425	0.7712	0.5917	0.7647	0.619
YP (N=15)	N	15	15	15	15	14	15	15	15	15	15
	Na	4	3	5	8	3	3	5	2	6	4
	AR	3.32	2.53	4.78	6.00	2.50	2.95	4.11	2.00	5.02	3.69
	R	195-203	206-248	241-263	261-291	146-152	197-206	294-336	211-214	134-186	
	Ho	0.333	0.267	0.333	0.467	0.071	0.333	0.067	0.533	0.667	0.341
	He	0.582	0.384	0.802	0.766	0.262	0.480	0.568	0.515	0.772	0.570
	FIS	0.435	0.313	0.593	0.399	0.735	0.314	0.886	-0.037	0.141	0.420
	HW	0.582	0.384	0.802	0.766	0.262	0.480	0.568	0.515	0.772	0.570
GP (N=26)	N	26	26	25	26	26	26	26	26	26	26
	Na	3	5	8	8	4	6	5	3	6	6
	AR	2.72	3.62	4.66	6.00	3.08	3.98	3.08	2.31	4.52	4.12
	R	195-203	206-214	241-263	261-279	140-155	191-224	291-309	211-214	134-162	
	Ho	0.346	0.577	0.560	0.577	0.115	0.346	0.308	0.423	0.654	0.448
	He	0.361	0.615	0.603	0.784	0.479	0.480	0.365	0.461	0.689	0.574
	FIS	0.043	0.063	0.073	0.268	0.763	0.282	0.160	0.083	0.052	0.229
	HW	0.361	0.615	0.603	0.784	0.479	0.480	0.365	0.461	0.689	0.574
JC (N=31)	N	31	31	31	31	31	31	31	31	31	31
	Na	3	7	9	12	6	5	6	4	9	7
	AR	2.68	4.83	5.66	6.62	4.28	3.83	3.51	2.90	6.30	4.51

Table 4. (*Continued*)

	R	195-203	204-218	239-263	261-289	146-161	191-224	291-309	208-217	126-174	
	Ho	0.323	0.645	0.581	0.742	0.194	0.581	0.129	0.484	0.710	0.487
	He	0.336	0.744	0.701	0.774	0.646	0.643	0.347	0.493	0.827	0.612
	FIS	0.040	0.134	0.174	0.042	0.704*	0.098	0.632*	0.020	0.144	0.221
	HW	0.496	0.224	0.226	0.431	0	0.104	0.00022	0.728	0.014	0.318
	N	32	32	32	32	32	32	32	32	32	32
	Na	3	4	9	10	6	4	5	4	10	6
	AR	2.23	3.50	5.32	5.63	4.18	2.88	3.10	2.69	5.35	3.88
JRS	R	195-203	206-214	237-263	261-291	146-161	191-203	291-309	202-214	134-174	
(N=32)	Ho	0.344	0.813	0.719	0.594	0.125	0.406	0.156	0.469	0.563	0.465
	He	0.335	0.649	0.705	0.732	0.527	0.372	0.310	0.499	0.749	0.542
	FIS	-0.027	-0.256	-0.019	0.191	0.766	-0.095	0.500*	0.061	0.252	0.153
	HW	1.000	0.481	0.455	0.076	0*	0.341	0.001	0.048	0.021	0.346
	N	34	34	34	34	34	34	33	34	34	34
	Na	4	5	9	14	7	3	4	5	11	7
	AR	2.52	4.23	5.44	7.33	3.77	2.71	2.49	2.85	6.89	4.25
HD	R	193-203	206-214	239-263	261-297	140-161	191-200	291-309	202-214	126-174	
(N=34)	Ho	0.294	0.794	0.706	0.794	0.206	0.294	0.909	0.235	0.794	0.559
	He	0.265	0.690	0.705	0.814	0.445	0.389	0.537	0.371	0.809	0.558
	FIS	-0.113	-0.154	-0.001	0.025	0.541*	0.247	-0.711*	0.368	0.019	0.025
	HW	1.000	0.675	0.771	0.223	0	0.103	0	0.005	0.737	0.502
	N	25	25	25	24	24	25	25	25	25	25
	Na	3	4	6	9	5	4	2	4	8	5
	AR	2.31	3.02	4.25	6.38	4.22	2.99	2.00	3.02	5.65	3.76
NHG	R	195-203	204-210	239-263	259-277	140-161	191-200	294-297	211-220	134-170	
(N=25)	Ho	0.320	0.720	0.720	0.833	0.333	0.360	1.000	0.480	0.720	0.610
	He	0.381	0.581	0.643	0.794	0.663	0.424	0.510	0.577	0.788	0.596
	FIS	0.163	-0.245	-0.122	-0.050	0.503*	0.155	-1.000*	0.171	0.088	-0.037
	HW	0.657	0.460	0.394	0.246	0.0006	0.215	0	0.132	0.533	0.377

Table 4. (*Continued*)

HC (N=36)	N	36	36	36	34	36	36	36	36	36	36
	Na	4	5	8	10	6	5	3	2	9	6
	AR	3.03	3.705	5.334	5.158	4.826	3.611	2.212	1.995	5.653	3.947
	R	193-203	206-214	241-263	261-281	140-161	191-224	291-297	211-214	126-174	
	Ho	0.500	0.444	0.583	0.559	0.361	0.472	0.306	0.389	0.833	0.494
	He	0.471	0.610	0.736	0.718	0.737	0.519	0.372	0.380	0.781	0.592
	FIS	-0.062	0.274	0.210	0.224	0.514*	0.092	0.181	-0.023	-0.069	0.149
	HW	0.040	0.029	0.053	0.109	0	0.144	0.480	1.000	0.033	0.236
CH1 (N=48)	N	48	48	48	48	48	48	47	48	48	48
	Na	3	3	6	7	6	4	3	3	9	5
	AR	2.70	2.97	4.92	6.76	5.66	3.21	2.49	3.00	7.40	4.34
	R	199-203	206-210	229-253	261-285	140-155	197-212	288-294	208-214	130-198	
	Ho	0.146	0.375	0.604	0.563	0.354	0.375	0.319	0.604	0.688	0.447
	He	0.139	0.350	0.616	0.774	0.625	0.469	0.330	0.577	0.670	0.506
	FIS	-0.051	-0.073	0.019	0.275	0.436*	0.202	0.034	-0.048	-0.026	0.085
	HW	1.000	0.842	0.310	0.002	0.000	0.144	1.000	0.526	0.624	0.556
CH2 (N=40)	N	40	40	40	40	39	40	40	40	40	40
	Na	3	5	7	4	5	3	3	3	5	4
	AR	2.98	4.50	5.90	3.97	4.56	2.58	2.82	3.00	4.89	3.88
	R	199-203	206-214	229-267	261-267	140-155	197-203	291-297	208-214	134-150	
	Ho	0.275	0.400	0.450	0.325	0.462	0.175	0.375	0.600	0.400	0.369
	He	0.250	0.497	0.663	0.332	0.618	0.517	0.499	0.566	0.426	0.497
	FIS	-0.103	0.197	0.324	0.022	0.256	0.664*	0.251	-0.061	0.061	0.240
	HW	1.000	0.243	0.004	0.513	0.012	0	0.007	0.855	0.405	0.319
CH3 (N=30)	N	30	30	30	29	30	30	28	30	30	30
	Na	3	3	7	7	4	3	3	3	5	5
	AR	2.99	3.00	6.43	6.67	4.00	3.00	3.00	2.77	4.95	4.35
	R	199-203	206-210	235-267	259-279	140-155	197-203	291-297	208-214	134-150	
	Ho	0.267	0.300	0.300	0.586	0.633	0.200	0.464	0.300	0.467	0.404

Table 4. (*Continued*)

	He	0.382	0.575	0.651	0.653	0.708	0.471	0.551	0.514	0.638	0.588
	FIS	0.306	0.482	0.544*	0.104	0.107	0.579*	0.159	0.420	0.272	0.335
	HW	0.036	0.000*	0	0.080	0.369	0.000	0.190	0.014	0.187	0.153
Mean all pop.	N	30	30	30	29	29	30	29	30	30	
	Na	3	4	7	9	5	4	4	3	8	
	AR	2.68	3.70	5.13	5.94	4.01	3.24	3.06	2.68	5.73	
	R	193-203	204-248	229-267	259-297	140-161	191-224	288-336	202-220	126-198	
	Ho	0.306	0.525	0.566	0.600	0.280	0.342	0.397	0.456	0.661	
	He	0.352	0.568	0.676	0.718	0.576	0.482	0.469	0.504	0.719	
	FIS	0.094	0.086	0.154	0.161	0.544	0.286	0.152	0.102	0.083	
	HW	0.594	0.451	0.423	0.362	0.349	0.284	0.483	0.443	0.435	

N_A, number of alleles; AR, allelic richness; H_O and H_E, observed and expected heterozygosities; F_{IS}, inbreeding coefficient; HWE, significant deviation from Hardy-Weinberg equilibrium after Bonferroni correction ($P < 0.005$).

4.3 Genetic relationships between the Korean and Chinese *S. gottschei* populations

A total of 49 alleles (42 in the Korean population and 7 in the Chinese population) were unique and distinguished the country of origin between the Korean groups (IJ, YP, GP, JC, JRS, HD, NHG, and HC) and Chinese groups (CH1, CH2, and CH3) (Table 5). The genetic distance and Pairwise F_{ST} for each pair of populations are shown in Table 6. All the Pairwise F_{ST} estimates were statically significant ($P < 0.05$). The pairwise F_{ST} ranged from -0.001 to 0.221, and genetic distance ranged from 0.040 to 0.431 among the 11 sample sites (Table 6). The mean difference in genetic distance between the eighth group in Korea and the third group in China was 0.301(0.203-0.431). This was a greater genetic distance than the mean for all Chinese groups, of 0.191 (0.068–0.123), and all Korean groups, of 0.109 (0.040–0.306).

Analysis of molecular variance (AMOVA) performed all sites which no grouping and regional group originating from Korea and China. The AMOVA analyzed all sites showed that of the total genetic variance, 10.78% was among sites and 74.79% was within sites. Also, the results of AMOVA grouping from regional population indicated the 12.83% was among groups and 16.78% was among sites within groups, and 70.4% was within sites (Table 7).

Table 5. Frequencies of each microsatellite allele in the Korean (IJ-YP-GP-JC-JRS-HD-NHG-HC) and Chinese (CH1-CH2-CH3) populations

Locus	Allele	China	Korea	Locus	Allele	China	Korea	Locus	Allele	China	Korea
Sc-02	193	0.000	0.010	Sc-26	259	0.009	0.002	Sc-30	288	0.004	0.000
	195	0.000	0.063		261	0.235	0.078		291	0.291	0.036
	197	0.000	0.005		263	0.483	0.390		294	0.678	0.679
	199	0.081	0.156		265	0.115	0.212		297	0.026	0.203
	201	0.055	0.000		267	0.094	0.020		300	0.000	0.024
	203	0.864	0.767		269	0.000	0.007		303	0.000	0.014
					271	0.013	0.073		306	0.000	0.005
Sc-06	204	0.000	0.010		273	0.000	0.073		309	0.000	0.024
	206	0.699	0.466		275	0.000	0.044		336	0.000	0.014
	208	0.195	0.067		277	0.004	0.027				
	210	0.089	0.353		279	0.009	0.032	Sc-36	202	0.000	0.007
	212	0.004	0.053		281	0.000	0.007		208	0.572	0.017
	214	0.013	0.041		283	0.021	0.007		211	0.318	0.655
	216	0.000	0.002		285	0.017	0.007		214	0.110	0.304
	218	0.000	0.005		289	0.000	0.010		217	0.000	0.014
	248	0.000	0.002		291	0.000	0.005		220	0.000	0.002
					295	0.000	0.002				
Sc-19	229	0.008	0.000		297	0.000	0.002	Sc-53	126	0.000	0.012
	235	0.008	0.000						130	0.025	0.000
	237	0.000	0.002	Sc-27	140	0.474	0.034		134	0.614	0.353

Table 5. (*Continued*)

239	0.297	0.019		143	0.038	0.002	138	0.072	0.257
241	0.191	0.043		146	0.090	0.597	142	0.131	0.144
243	0.000	0.005		149	0.013	0.029	146	0.097	0.060
245	0.004	0.012		152	0.312	0.189	150	0.034	0.046
247	0.428	0.488		155	0.073	0.109	154	0.008	0.043
249	0.008	0.215		158	0.000	0.010	158	0.004	0.000
251	0.000	0.010		161	0.000	0.029	162	0.000	0.026
253	0.038	0.097					166	0.000	0.007
255	0.000	0.046	Sc-28	191	0.000	0.087	170	0.000	0.019
257	0.000	0.002		194	0.000	0.005	174	0.000	0.024
263	0.000	0.060		197	0.347	0.161	186	0.000	0.005
267	0.017	0.000		200	0.606	0.697	190	0.004	0.002
				203	0.042	0.017	198	0.008	0.000
				206	0.000	0.012			
				212	0.004	0.005			
				221	0.000	0.002			
				224	0.000	0.014			

Table 6. Pairwise F_{ST} (above diagonal) and genetic distance (below diagonal) between 11 sample site of *S. gottschei*

D_{IS} F_{ST}	CH1	CH2	CH3	IJ	YP	GP	JC	JRS	HD	NHG	HC
CH1	0.000	0.096*	0.066*	0.212*	0.205*	0.170*	0.150*	0.180*	0.197*	0.178*	0.152*
CH2	0.123	0.000	0.043*	0.206*	0.221*	0.196*	0.158*	0.186*	0.212*	0.184*	0.158*
CH3	0.102	0.068	0.000	0.136*	0.161*	0.132*	0.106*	0.129*	0.140*	0.122*	0.104*
IJ	0.431	0.394	0.333	0.000	0.115	0.071*	0.061*	0.057*	0.085*	0.037*	0.073*
YP	0.373	0.397	0.347	0.306	0.000	0.066*	0.062*	0.058*	0.095*	0.077*	0.074*
GP	0.269	0.311	0.241	0.166	0.134	0.000	0.012	-0.001	0.027	0.035*	0.008
JC	0.261	0.268	0.220	0.182	0.147	0.041	0.000	0.009	0.042*	0.043*	0.006
JRS	0.290	0.291	0.232	0.138	0.118	0.022	0.034	0.000	0.032	0.037*	0.014
HD	0.337	0.359	0.263	0.188	0.181	0.056	0.086	0.061	0.000	0.018	0.031
NHG	0.314	0.316	0.248	0.122	0.169	0.074	0.102	0.076	0.048	0.000	0.027
HC	0.257	0.258	0.203	0.190	0.160	0.035	0.036	0.040	0.065	0.068	0.000

Table 7. AMOVA for 11 sample sites of *S. gottschei*

Analysis	Source of Variation	*d.f.*	Sum of squares	Variance components	Percentage of variation (%)
All site	Among sites	10	203.7	0.29	10.78
(No grouping)	Within sites	326	676	2.07	74.79
Regional grouping	Among groups	1	116.8	0.38	12.83
(Korea vs. China)	Among sites within groups	9	992.1	0.49	16.78
	Within sites	326	676	2.07	70.4

According to the UPGMA tree generated from the D_{ce} values, the CH1, CH2, and CH3 populations are grouped into a different clade than those of the other geographic populations (IJ, YP, GP, JC, JRS, HD, NHG, and HC) (Fig 1). The clustering pattern was similar to the Principal coordinate analysis (PCA) based on D_{ce} values (Fig 2). The PCA of the 11 snail populations showed that the genetic variation was expected to be 83.29% of the total for the top three principal coordinates (59.46% for the first principal coordinate and only 14.35% for the second principal coordinate This indicated a difference between the Korean and Chinese populations. PCR further confirmed the results exhibited by STRUCTURE outputs predicted K = 2 to be the most likely number of clusters (Fig 3).

In addition, the individual assignment test using GeneClass2 indicated that IJ, YP, GP, JC, JRS, HD, NHG, and HC are one group, and CH1, CH2, and CH3 are different groups, separating into two clusters at 96.9%. These results also suggest a genetic difference between snails originating in Korea and those originating in China.

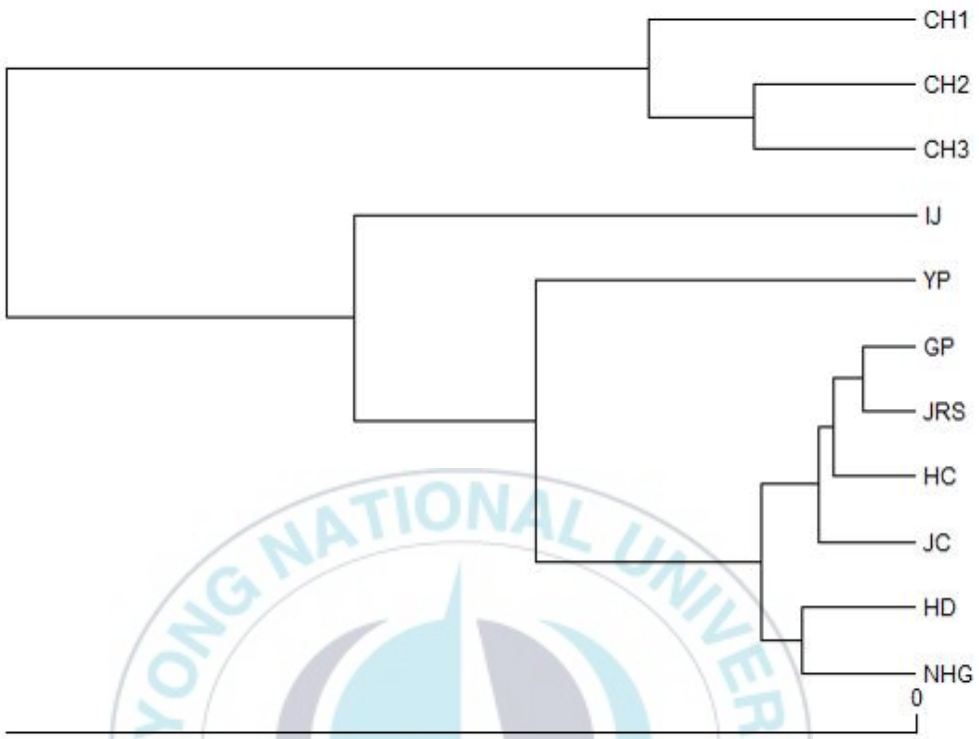

Figure 1. UPGMA tree based on matrices of Nei's distance of nine microsatellite markers.

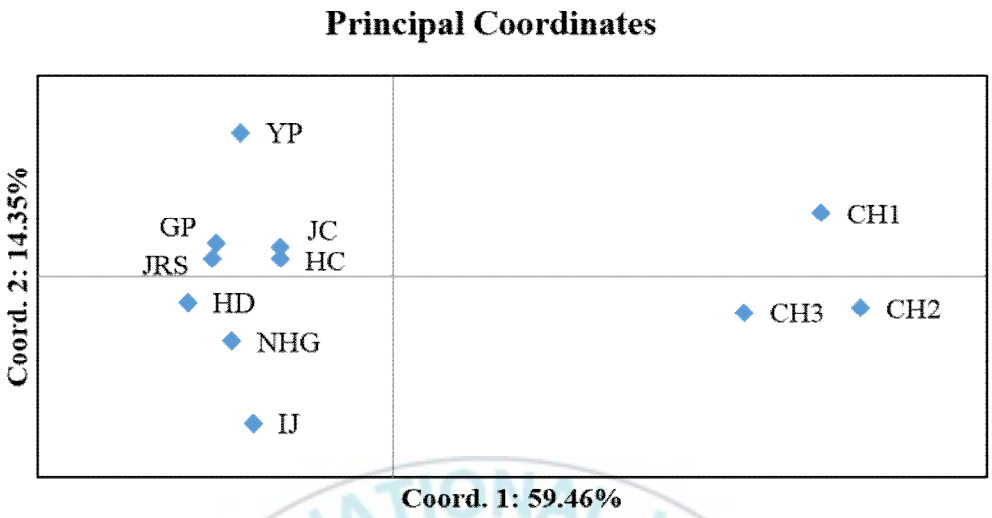

Figure 2. Pincipal coordinate analysis based on Nei distances using 9 MS loci from 11 populations.

K=2

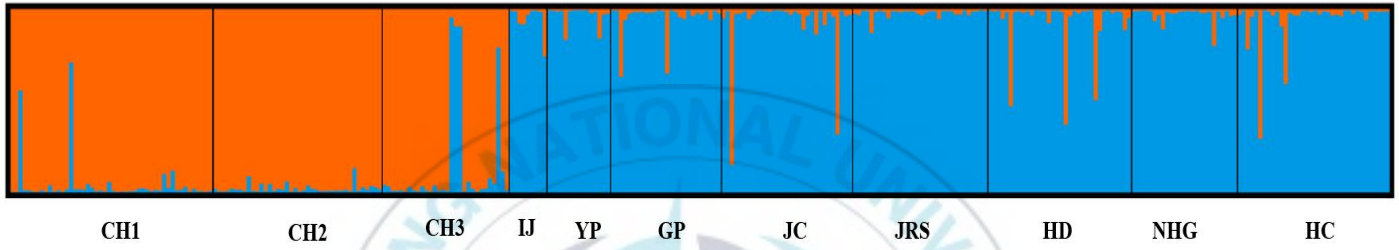

Figure 3. Population structure inferred by Bayesian clustering at *K*=2. The separation of the column into two colors represents the differences between each cluster.

5. Discussion

Although the Korean melania snail inhabits Korea, Japan, China, and Taiwan, phylogeographic analyses have only been performed in *Semisulcospira libertina* based on mtDNA sequences (Miura et al, 2013; Chiu et al, 2017). Since the Korean melania snail is imported mainly from China, a genetic analysis of Chinese samples distributed in Korea and those collected in Korea is needed. In this study, we aimed to assess genetic diversity and to distinguish the country of origin between Korean and Chinese *S. gottschei* populations using mtDNA and microsatellite markers. Haplotype diversity of the assembled COI gene sequences showed 29 haplotypes in the Korean population and 14 haplotypes in the Chinese population, for a total of 43 haplotypes (Table 3). No SNP that distinguished between the Korean and Chinese populations was found, but the Chinese haplotype group was distinct from the Korean haplotype group, and there was no gene flow between the Korean and Chinese populations. In a previous study of the phylogenic relationship of the 16s rRNA gene between Korean and Japanese samples of *S. libertina*, the Korean samples were found to be from a sister lineage to the Japanese congener (Miura et al, 2013). In addition, Chiu et al (2017) found *S. libertina* from Korea, Japan, China and Taiwan fell into three divergent phylogroups, which are distinct from the Korean and Chinese phylogroups. In the present study, the interspecific genetic distance of *S. gottschei* between the Korean and Chinese populations was approximately 2%. The mtDNA sequence divergence reported within mollusk species generally does not exceed 5% (Wise et al, 2004; Robinson and Dillon 2008). However, Chiu et al (2017) reported that the mean sequence divergence of *S. libertina* among three major phylogroups was 13.05% due to the inclusion of cryptic species within *S. libertina*. To further

assess genetic diversity in *S. gottschei* between the Korean and Chinese populations, a genetic population analysis based on microsatellites was conducted.

In the present study, analysis of the genetic structure of *S. gottschei* was performed to confirm gene flow between the Korean and Chinese populations using nine SSR loci developed for *S. coreana*. Heterozygosity is an important measure of population genetic diversity because it indicates a large range of genotypes for adaptive response to changing environmental conditions (Xu et al, 2001). No significant deviation from HWE was observed for any locus in this study (Table 4). However, the H_e value was greater than H_o value at all SSR loci, indicating excessive homozygosity at these loci. The significantly limited heterozygosity is associated with the existence of null alleles (Selkoe and Toonen 2006), which are common in freshwater mollusc (Gu at al, 2012). Also, these deficits were observed in several invertebrate species, such as *Lampsilis abrupta, Panopea generosa,* and *Ostrea edulis* (Vadopalas et al, 2004; Launey et al, 2002; Eackles and King 2002). The presence of the null allele might indicate a major limitation in the cross-species transfer of the MS marker (Barbara et al, 2007). In the present study, theses microsatellite loci showed cross-species transfer. In this case, the presence of the null allele might be caused by the character of the MS marker, rather than limited transferability of the MS locus. In addition, high F_{IS} estimate was found in most populations except for NHG with no null allele excluding Sc-27 loci. Moderate to high F_{IS} value were found in freshwater snails with poor mobility (Nguema et al, 2013), which was also found in this study.

A total of 49 unique alleles could be distinguished between the Korean groups ((IJ, YP, GP, JC, JRS, HD, NHG, and HC) and Chinese groups (CH1, CH2, and CH3) (Table 5). This unique allele, considered a population-specific marker of

gene flow, was used to differentiate the geographic populations (Slatkin 1985). It has been reported that marine invertebrates can be geographically distinguished based on this unique allele (Zhan et al, 2009; An et al, 2009), and this result was also found in the present study. In the Pairwise F_{ST} and AMOVA tests, the Korean and Chinese populations also showed significant differences (Table 6). The AMOVA analysis of all sites and groups of the Korean and Chinese populations showed that the total genetic variance was 10.78% and 12.83%, respectively, with higher genetic variance in the groups than at all sites (Table 7). Furthermore, the genetic distance was different between the eighth group in the Korean population and the third group in the Chinese population, with a mean difference of 0.301 (0.203–0.431). These results were further confirmed by the UPGMA dendrogram tree and PCA (Fig 1 and 2). Both of these genetic population analyses showed separation of the 11 populations into two clusters. The STRUCTURE results also revealed clear distinction between the Korean and Chinese populations (Fig 3). In this study, microsatellite analysis indicated that the Korean and Chinese melania snail populations were genetically different, which accords with mtDNA results. Miura et al. (2013) proposed that the freshwater snail migrated from Japan to Korea during the Pliocene and Pleistocene ice ages via a land bridge connecting East Japan to the Asian continent through present-day Korea (Lockwood 1979; Park et al, 2006). In addition, our data support that Korean *Semisulcospira* originated in Northeast Asia and is distinct from Chinese *Semisulcosapira,* which originated in South Asia.

To confirm the differentiation of *S. gottschei* between Korea and China, we found high genetic variation according to mtDNA and polymorphic MS analyses. Our study revealed that analysis of SSR loci in *S. coreana* is a powerful means to

monitor the genetic structure of *S. gottschei* populations. Nine MS markers were used to distinguish Korean from Chinese *S. gottschei* populations. The ability to differentiate these populations will help prevent mislabeling of melania snails in the domestic market. Also, these results provide valuable genetic information for management and further improvement of Korean domestic melania snails.

6. Reference

Davis G.M. (1969) A taxonomic study of some species of *Semisulcospira* in Japan (Mesogastropoda: Pleuroceridae). Malacologia., 7, 211-294.

Sydes M.A., Peakall R. (1998) Extensive clonality in the endangered shrub Haloragodendron lucasii (Haloragaceae) revealed by allozyme and RAPDs. Molecular Ecology., 7, 87-92.

Moon S.K., Kim I.S., Lim C.W., Yoon N.Y., Jeong B.Y. (2015) Proximate and Fatty Acid Compositions of Commercial Domestic and Imported Melania Snails *Semisulscospira* sp. Korean. J. Fish. Aquat. Sci., 48, 977-981.

Rasmussen R.S., Morrissey M.T. (2008) DNA-based methods for the identification of commercial fish and seafood species. Comprehensive reviews in food science and food safety, 7, 280-295.

Planes S., and Fauvelot C. (2002) Isolation by distance and vicariance drive genetic structure of a coral reef fish in the Pacific Ocean. Evolution, 56, 378-399.

Roe K.J., Hartfield P.D., Lydeard C. (2001) Phylogeographic analysis of the threatened and endangered superconglutinate-producing mussels of the genus *Lampsilis* (Bivalvia: Unionidae). Mol. Ecol., 10, 2225-2234.

Toro J.E., Ojeda J.A., Vergara A.M. (2004) The genetic structure of *Mytilus chilensis* (Hupe 1854) populations along the Chilean coast based on RAPDs analysis. Aquacul. Res., 35, 1466-1471.

Kong L. and Li Q. (2007) Genetic comparison of cultured and wild populations of the clam *Coelomactra antiquata* (Spengler) in China using AFLP markers. Aquaculture, 271, 152-161.

Gruenthal K.M., Acheson L.K., Burton R.S. (2007) Genetic structure of natural populations of California red abalone (*Haliotis rufescens*) using multiple genetic markers. Mar. Biol.,

152, 1237-1248.

Yamashita Y., Omura Y., Okazaki E. (2006) Distinct regional profiles of trace element content in muscle of Japanese eel *Anguilla japonica* from Japan, Taiwan, and China. Fish. Sci., 72, 1109–1113.

Rooker J.R., Secor D.H., Zdanowicz V.S., DeMetrio G., Relini L.O. (2003) Identification of northern bluefin tuna stocks from putative nurseries in the Mediterranean Sea and western Atlantic Ocean using otolith chemistry. Fish. Oceanogr., 12, 75–84.

Kang J.H., Noh E.S., Park J.Y., An C.M., Choi J.H., Kim J.K. (2015) Rapid origin determination of the northern mauxia shrimp (*Acetes chinensis*) based on allele specific polymerase chain reaction of partial mitochondrial 16S rRNA gene. Asian. Australas. J. Anim. Sci., 28, 568–572.

Park Y.J., Noh E.S., Park J.Y., Kim E.M., Noh J.K., Choi T.J., Kang J.H. (2018) Development of a simple, rapid multiplex PCR method using the 16S rRNA gene to determine the country of origin of brackish water bivalve *Corbicula japonica*. Food analytical methods, 11, 3034-3041.

Takashima Y., Morita T., Yamashita M. (2006) Complete mitochondrial DNA sequence of Atlantic horse mackerel *Trachurus trachurus* and molecular identification of two commercially important species *T. trachurus* and *T. japonicus* using PCR-RFLP. Fish. Sci., 72, 1054-1065.

Rahman M.T., Uddin M.S., Sultana R., Moue A., Setu M. (2013) Polymerase chain reaction (PCR): a short review. A.K.M.M.C. J., 4, 30-36.

Kang J.H., Kang H.S., Noh E.S., Park J.Y., An C.M. (2014) Isolation and characterization of novel microsatellite markers for the northern mauxia shrimp, *Acetes chinensis*, using pyrosequencing. Marine genomics, 18, 67-69.

Dong Y.H., Yao H.H., Sun C.S., Lv D.M., Li M.Q., Lin Z.H. (2016) Development of polymorphic

SSR markers in the razor clam (*Sinonovacula constricta*) and cross-species amplification. Genet. Mol. Res., 15(1).

Hauser L., Adcock G.J., Smith P.J., Ramírez J.H.B., Carvalho G.R. (2002) Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). Proceedings of the National Academy of Sciences, 99, 11742-11747.

An H., Lee Y., Park J., Lee C. (2009) Genetic characterization of four East Asian giant scallop (*Mizuhopecten yessoensis*) populations using microsatellite markers. Aquaculture Research, 40, 619-624.

Ball A.O., Sedberry G.R., Zatcoff M.S., Chapman R.W., Carlin J.L. (2000) Population structure of the wreckfish *Polyprion americanus* determined with microsatellite genetic markers. Marine Biology, 137, 1077-1090.

Gross R., Palm S., Koiv K., Prestegaard T., Jussila J., Paaver T., Geist J., Kokko H., Karjalainen A., Edsman L. (2013) Microsatellite markers reveal clear geographic structuring among threatened noble crayfish (*Astacus astacus*) populations in Northern and Central Europe. Conserv. Genet., 14, 809-821.

Allendorf F.W., England P.R., Luikart G., Ritchie P.A., Ryman N. (2008) Genetic effects of harvest on wild animal populations. Trends. Ecol. Evol., 23, 327-337.

Raymond M., Rousset F. (1995). GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. J. Hered., 86, 248-249.

Goudet J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2). Accessed 27 September, 2012.

Excoffier L., Lischer H.E. (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources, 10, 564-567.

Langella O. (2007) Populations 1.2.3.0: Population genetic software (individuals or population distances, phylogenetic trees). http://bioinformatics.org/~tryphon/populations/

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

Evanno G., Regnaut S., Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol., 14, 2611-2620.

Piry S., Alapetite A., Cornuet J.M., Paetkau D., Baudouin L., Estoup A. (2004) GENECLASS2: a software for genetic assignment and first-generation migrant detection. J. Hered, 95, 536-539.

Miura O., Köhler F., Lee T., Li J., Ó Foighil, D. (2013) Rare, divergent Korean *Semisulcospira spp.* mitochondrial haplotypes have Japanese sister lineages. J. Mollus. Stud., 79, 86-89.

Chiu Y.W., Bor H., Kuo P.H., Hsu K.C., Tan M.S., Wang W.K., Lin H.D. (2017) Origins of *Semisulcospira libertina* (gastropoda: semisulcospiridae) in Taiwan. Mitochondrial DNA Part A, 28, 518-525.

Wise J., Harasewych M.G., Dillon R.T. (2004) Population divergence in the sinistral whelks of North America, with special reference to the east Florida ecotone. Mar. Biol., 145, 1167-1179.

Robinson J.D. and Dillon R.T. (2008) Genetic divergence among sympatric populations of three species of oyster drills (*Urosalpinx*) in Cedar Key. Florida. Bull. Mar. Sci., 82, 19-31.

Xu Z., Primavera J.H., De La Pena L.D., Pettit P., Belak J., Alcivar-Warren A. (2001) Genetic diversity of wild and cultured black tiger shrimp (*Penaeus monodon*) in the Philippines using microsatellites. Aquaculture, 199, 13-40.

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

Gu Q., Xiong B., Zhu Y., Wang Q., Shi P. (2012) Isolation and characterization of polymorphic

microsatellites loci in the freshwater snail *Bellamya aeruginosa* (Mollusca, Gastropoda). Conservation. Genet. Resour., 4, 387-390.

Launey S., Ledu C., Boudry P., Bonhomme F., Naciri-Graven Y. (2002) Geographic structure in the European flat oyster (*Ostrea edulis* L.) as revealed by microsatellite polymorphism. J. Hered., 93, 331-351.

Eackles M.S. and King T.L. (2002) Isolation and characterization of microsatellite loci in *Lampsilis abrupta* (Bivalvia: Unionidae) and cross-species amplification within the genus. Mol. Ecol. Notes., 2, 559-562.

Vadopalas B., Leclair L.L., Bentzen P. (2004) Microsatellite and allozyme analyses reveal few genetic differences among spatially distinct aggregations of geoduck clams (*Panopea abrupta*, Conrad 1849). J. Shellfish. Res., 23, 693-706.

Barbara T., PALMA-SILVA C., Paggi G.M., Bered F., Fay M.F., Lexer C. (2007) Cross-species transfer of nuclear microsatellite markers: potential and limitations. Mol. Ecol., 16, 3759-3767.

Nguema R.M., Langand J., Galinier R., Idris M.A., et al. (2013) Genetic diversity fixation and differentiation of the freshwater snail *Biomphalaria pfeifferi* (Gastropoda Planorbidae) in arid lands. Genetica., 141, 171-184.

Slatkin M. (1985) Rare alleles as indicators of gene flow. Evolution, 39, 53-65.

Zhan A., Hu J., Hu X., Zhou Z., Hui M., Wang S., Peng W., Wang M., Bao Z. (2009) Fine-scale population genetic structure of Zhikong scallop (*Chlamys farreri*): do local marine currents drive geographical differentiation?. Mar. Biotechnol., 11, 223-235.

Lockwood J.G. (1979). The Antarctic Ice-Sheet: Regulator of Global Climates?. Rev. Geogr. J., 145, 469-471.

Park Y.C., Kitade O., Schwarz M., Kim J.P., Kim W. (2006) Intraspecific molecular phylogeny, genetic variation and phylogeography of *Reticulitermes speratus* (Isoptera:

Rhinotermitidae). Mol. Cells. 21, 59-103.

Acknowledgements

이 논문이 완성되기까지 많은 도움과 격려를 주신 모든 분들께 감사의 마음을 전합니다. 긴 학위 기간 동안 부족한 제가 스스로 깨우치고 배워나갈 수 있도록 이끌어 주시고, 따뜻하게 격려해 주신 최태진 교수님과 강정하 박사님께 깊은 감사와 존경의 마음을 드립니다. 또한, 논문의 심사를 위해 조언과 충고를 아끼지 않고 많은 관심을 가져주신 이명숙, 김군도 교수님과 김은미 박사님께도 감사드립니다. 여러 해 동안 함께 일하며 많은 힘이 되어주신 국립수산과학원 생명공학과의 여러 박사님들과 동료 연구원분들께도 감사의 마음을 전합니다. 또한, 저에게 관심과 격려를 아끼지 않고 많은 도움을 주신 모든 분들께 진심으로 감사의 마음을 전합니다. 마지막으로, 지금의 제가 있기까지 항상 옆에서 변함없는 사랑으로 지켜봐 주시고 든든한 버팀목이 되어준 우리 가족, 아버지 박우진님, 어머니 이해숙님 그리고 남편 손명우님께 깊은 감사와 사랑의 마음을 표합니다. 감사합니다.