Low diversity, little genetic structure but no inbreeding in a high-density island endemic pit-viper *Gloydius shedaoensis*

Guannan Wen\(^a\), Long Jin\(^b,\(^t\), Yayong Wu\(^c\), Xiaoping Wang\(^d\), Jinzhong Fu\(^a,\(^e\), and Yin Qi\(^a,\(^b\)

\(^a\)Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, China  
\(^b\)Key Laboratory of Southwest China Wildlife Resources Conservation (Ministry of Education), China West Normal University, Nanchong 637009, China  
\(^c\)College of Life Sciences and Food Engineering, Yibin University, Yibin 644007, China  
\(^d\)Nature Conservation of Snake Island and Laotieshan Mountain, Dalian 116041, China  
\(^e\)Department of Integrative Biology, University of Guelph, Guelph, Ontario N1G 2V1, Canada

**Abstract**

Island species and their ecosystems play an important role in global biodiversity preservation, and many vulnerable island species are conservation priorities. Although insular habitat likely facilitates the speciation process, it may also aggravate the fragility of these species with high risk of inbreeding. The Shedao pit-viper *Gloydius shedaoensis* is an island endemic species with an extremely high population density, which has been categorized as vulnerable in the IUCN (International Union for the Conservation of Nature and Natural Resources) Red List. We collected 13,148 SNP (Single Nucleotide Polymorphism) from across its genome and examined its genetic diversity and demographic history. The Shedao pit-viper has a low genetic diversity but shows no sign of inbreeding. Furthermore, population genetic structure analysis, including the neighbor-joining tree, principal coordinate analysis, clustering, and spatial autocorrelation, revealed a general lack of spatial structure. Only a large number of individuals, small size of the island, and the lack of population structure likely all contribute to the lack of inbreeding in this species. We also detected signs of male-biased dispersal, which likely is another inbreeding avoidance strategy. Historical demographic analysis suggested that the historical population size and distribution of the species are much larger than their current ones. The multiple transgressive–regressive events since the Late Pleistocene are likely the main cause of the population size changes. Taken together, our results provide a basic scientific foundation for the conservation of this interesting and important species.

**Keywords:** genetic diversity and structure, *Gloydius shedaoensis*, historical demography, inbreeding, insular species

Island species and their ecosystems are an important part of global biodiversity, and islands have often been regarded as ideal natural cradles for species diversification (Losos and Ricklefs 2009). The diverse properties of insular ecosystems have facilitated the bewildering diversity of forms for those species endemic to islands (Losos and Ricklefs 2009; Brown et al. 2013). However, islands are relatively impoverished, unsaturated, and disharmonic, and the small area and high population density likely force the island species facing the fates of fragility (Simberloff 2000). In particular, for those small islands, they usually have a lower overall richness of species assemblages but a higher density of a taxon than that in equivalent mainland habitats, sometimes even to an apparently excessive degree (“Excess density compensation,” MacArthur et al. 1972; but see Whittaker and Fernández-Palacios 2007). Such a situation has been reported in several faunal taxa, like birds (Crowell 1962), lizards (Case 1975), and geckos (Rodda et al. 2001; but see Rodda and Dean-Bradley 2002).

One of the deleterious consequences of over-crowding is more opportunities for inbreeding, which may lead to a decrease in genetic density, an accumulation of deleterious recessive alleles, or even worse, carrying the populations toward an imminent extinction (e.g., royal wolves; Robinson et al. 2019). Nevertheless, high costs of inbreeding (Keller and Waller 2002) can also lead to the evolution of inbreeding avoidance tactics, such as sex-biased dispersal in birds and mammals (Pusey 1987; but see Li and Kokko 2019), avoiding mating with relative in canids and bank voles (Geffen et al. 2011; Lemaître et al. 2012), kin recognition in zebra finches (Arct et al. 2010), and sperm selection in field cricket (Breitman et al. 2004; but see Blouin and Blouin 1988). Alternatively, inbreeding may not cause selection for...
inbreeding avoidance, because of a tradeoff between potential benefits of inbreeding and costs of inbreeding avoidance (Lehmann and Perrin 2003; Kokko and Ots 2006). Diverse patterns of inbreeding strategies have been reported, such as inbreeding tolerance in New Zealand robins (Jamieson et al. 2009) and bighorn sheep (Rioux-Paquette et al. 2010), or breeding preference in cichlid fish (Thünen et al. 2007; but see Szulkin et al. 2013).

The Shedao pit-viper Gloydius shedaoensis is an insular snake endemic to the Shedao Island in the Bohai Sea, off the eastern coast of China. The small island has a size of \( \sim 0.73 \) km\(^2\) and has an extremely high population density of the pit-viper. The island population size is estimated at approximately 20,000 (Li et al. 2007). The adult vipers are \( \sim 600 \) mm in length and they rely almost exclusively on seasonal migrating avian preys, which is different from other snakes (Shine et al. 2002; Martins et al. 2008). For most of the year, food resources on the island are extremely barren, and snakes typically remain in hibernation and aestivation (Huang 1989; Li 1995). Such extreme seasonal foraging conditions make Shedao pit-vipers very vulnerable to environmental changes. The species is currently listed as class 2 protected species in China and has been categorized as vulnerable in the IUCN Red List. To conserve this unique insular species, many studies have been carried out over the past several decades (Li 1995; Shine et al. 2002); however, its genetic diversity and demography history remain largely unknown.

In this study, we collected genome-wide SNP data using RAD (Restriction-site Associated DNA) sequencing and addressed several conservation genetic-related questions of the Shedao pit-viper (G. shedaoensis). First, we evaluate the genetic diversity of the species. Second, we detect population genetic structure, spatial genetic distribution, and factors that may impact gene flow of the population. Third, we reconstruct the historical population demography of the species. With a large amount of genetic data, we hope that our study will fill a significant knowledge gap and provide a scientific foundation for the conservation of this interesting and important insular species.

**Materials and Methods**

**Sampling and laboratory protocols**

A total of 71 individuals (36 females and 35 males) of G. shedaoensis from the Shedao Island were collected. Detailed location information is provided in Supplementary Figure S1 and Supplementary Table S1. A small piece of skin near the base of the tail was collected from each individual and all individuals were released on-site. Total DNA was extracted using a phenol–chloroform extraction protocol. The quality, integrity, and quantity of genomic DNA were examined using a fluorometer and real-time PCR (Polymerase Chain Reaction). Libraries were sequenced on an Illumina HiSeq2500 platform as paired-end, 150 bp reads at the Novogene Corporation (Tianjin, China).

**Data processing and quality control**

The process_radtags.pl from Stacks version 2.41 (Catchen et al. 2013) was used to clean and de-multiplex the raw reads. Reads with low-quality scores (average Phred33 score within 20-bp sliding window <10) were discarded. We preserved the tags with intact restriction enzyme cut sites and ambiguous barcodes (maximum 2 mismatches).

All 71 samples were genotyped. We ran ustacks with the following settings: minimum number to create a stack, \( m = 5 \); maximum distance allowed between stacks, \( M = 6 \); maximum number of stacks at a single de novo stacks = 3 (default). We also enabled the delevager algorithm to resolve over-merged tags. The maximum number of mismatches allowed between loci (\( n \)) was set to 6 in cstacks (Paris et al. 2017). We performed gapped alignments between stacks in ucslstacks by keeping the number of gaps allowed between stacks before merging at the default of 2, and the minimum length of aligned sequence in a gapped alignment at the default of 0.80. Then, we used isu2bam to associate the pair-end reads to each single-end locus and used gstacks to assemble the pair-end contigs.

SNP calling was conducted using populations in Stacks. We kept loci present in at least 85% of individuals within all \( r = 0.85 \). Loci with low minor allele frequency were excluded \( (\text{min}_\text{maf} 0.05) \). Loci present in all individuals were kept \( (p = 1) \). To minimize the effect of linkage disequilibrium, we only wrote out the first SNP per fragment (write_single_snp).

For the exported dataset, we first thinned the dataset by excluding SNPs unique to both libraries. Then we further filtered data by PLINK version 1.90b3.46 (Purcell et al. 2007) as following: 1) removing loci potentially in physical linkage by removing one locus from each pair of loci with \( r^2 > 0.85 \) (\( r^2 \)); 2) removing duplicated samples by excluding one individual from each pair of individuals with PI_HAT > 0.9 (genome). At last, the final mean coverage depth was calculated using VCFTools (depth; Danecek et al. 2011).

**Population genetic structure analysis**

A set of descriptive statistics, including nucleotide diversity (\( \pi \)), observed heterozygosity (\( H_o \)), and expected heterozygosity (\( H_e \)), were obtained from Stacks with populations (= fstacks). A genomic inbreeding coefficient \( F_{\text{UP}} \), which is based on the correlation between uniting gametes (Yang et al. 2011), was also obtained from individual SNP data.

We evaluated the population structure using 3 commonly employed methods for SNP data. A neighbor-joining (NJ) tree based on p-distance was constructed with Splitstree (Huson 1998). Additionally, a principal coordinate analysis (PCoA) using Euclidean distance between individuals was carried out with the R package “dattR” (Gruber et al. 2018; R Core Team 2020). The R package “ggplot2” (Wickham and Chang 2008) was used for plotting the results. Lastly, the genetic clustering analysis was conducted using Admixture (Alexander et al. 2009). Starting from \( K = 1 \) to \( K = 6 \), 10 independent runs were performed for each \( K \) with a tolerance of 1e-4. For each
converged K, we plotted results from 1 run with the lowest cross-validation (CV) error.

We also explored the fine-scale spatial patterns using a spatial autocorrelation analysis implemented in SPAGeDi version 1.5 (Hardy and Vekemans 2002). The individual pairwise relatedness coefficient (Moran’s I statistics; Hardy and Vekemans 1999) was first calculated, and all samples were assigned to distance classes. Distance classes were fine-tuned according to the criteria suggested by the program (%partic > 50%, CV patric ≤ 1, and # pairs > 100). For each distance class, a null distribution of the relatedness coefficient with a 95% confidence interval (CI) was obtained from 999 permutations. We further inferred significant autocorrelation if the average observed relatedness coefficient value within a distance class fell outside its CI. Lastly, we applied the Mantel correlation test for an overall relationship between the relatedness coefficients (Moran’s I statistics) and pairwise geographic distances (Euclidean) among individuals using the R package “vegan” (Oksanen et al. 2019) with 999 permutations. To ensure each individual receiving a unique spatial coordinate, only one individual from each collecting site was used. Consequently, we retained 66 individuals (35 females and 32 males) for this analysis. To test for sex-specific patterns, the analysis was separately conducted for all samples, females only, and males only.

We also mapped the variance of pairwise genetic dissimilarity with isolation by distance (IBD) as null model for detecting possible migration barriers or corridors in both sexes. The distribution of residual dissimilarity (DResD) procedure of Keis et al. (2013) was used. The relation of genetic dissimilarity and geographical distance was fitted through the points estimated by an IBD model. The residual values from the fitted model (IBD residuals) were used as a measure of genetic dissimilarity with the IBD effect accounted for. We used the pairwise genetic distance matrix calculated in SPAGeDi version 1.5 as data input and interpolated over a grid of points (steps of 10 m) covering the entire sampling area. We set 10 m as the minimum distance between points to exclude the potential effect of animal mobility within a home range. To maximize signals of any corridors or barriers, the analysis was applied to pairs of individuals with distance up to the genetic patch size calculated from the spatial autocorrelation analysis (female: 0.53 km, male: 0.24 km; see results). We run 1,000 random iterations to test whether the mean genetic dissimilarity at grid points significantly deviated from an IBD model, and 100 bootstrap iterations to estimate the post hoc statistical power of the randomization test. To test potential sex-biased dispersal, the analysis was separately conducted for females and males.

Demographic simulations
We inferred the demographic history of the island population using the stairway plot method (Liu and Fu 2015), which explores the population size changes over time using site frequency spectrum (SFS). The folded SFS of our data was computed using daemon (Gutenkunst et al. 2010). The mutation rate per site per generation (μ = 2.904e-9) was inferred from the transcriptome data of this species (Lu et al. 2020). The generation time of 3 years was used (Yang 1983).

Experimental Ethics
We have obtained the permits from Nature Conservation of Snake Island and Laotieshan Mountain. All activities were under permission from local conservation authorities and animal handling followed the approved protocols (number 2017005, Animal Ethics Committee of Chengdu Institute of Biology).

Results

Data
Approximately 24.6 million reads were obtained and 63% of them participated in SNP calling. A total of 70 individuals were retained, with 1 male being removed because of a large amount of missing data. After quality control, 13,159 unique loci shared by all individuals were kept. We further removed 11 loci from 11 pairs of loci due to possible physical linkage. The final dataset included 13,148 SNPs with a mean coverage of 19.1 (± 6.62) per individual. No duplicate sample was detected.

Genetic diversity and population structure
The Shedao pit-viper showed a rather low level of genetic diversity. Its nucleotide diversity (π) was 9.9e-4 (SD = 1e-5). The base population size per generation (N0) was estimated using the program DnaSP v 5.10.4 (Librado and Rozas 2009). The final dataset included 13,148 SNPs with a mean coverage of 19.1 (± 6.62) per individual. No duplicate sample was detected.

The population revealed very little structure. The NJ tree produced a “star” pattern with branches having similar lengths (Figure 1A). In consistent with the tree, PCA showed a dispersed pattern (Figure 1B), with the first 2 components explaining merely 4.23% of the total variation. Clustering analysis did not reveal any detectable subdivision among all sampled individuals, which showed the lowest CV error being at K = 1 (Figure 1C).

Fine-scale spatial analysis, however, did reveal several interesting patterns. Figure 2 presents the results from the spatial autocorrelation analysis with Moran’s I. The all-individual and females-only analysis revealed decreased relatedness with increased distance. The first 2 distance classes showed significantly higher values than expected by chance, as the last class showed a significantly lower value (P < 0.05). Furthermore, the lines intercepted the x-axis at 0.53 km (all samples) and at 0.47 km (females-only), respectively. The first x-intercept is often considered to represent the size of a genetic patch (Sokal 1979; Sokal and Wartenberg 1983; Diniz-Filho and Telles 2002), which is a genetic homogeneous region and is often defined as operational units for conservation in populations lack of sub-structure (Diniz-Filho and Telles 2002). Conversely, males revealed a largely uniform relatedness across space although its first distance class possessed a significant positive coefficient (Figure 2). For males, the first x-intercept was 0.24 km, less than half of that of females. Nevertheless, the overall correlation between relatedness and distance was weak; the Mantel test did not detect a significant global correlation between relatedness coefficient and geographical distance (P = 1).

The DResD analysis did not detect any significant barrier or corridor to gene flow for males or females. The entire population showed low genetic dissimilarity across the island; however, numerous areas had significantly lower values than expected from IBD alone (Figure 3), and most of these residual values had significant bootstrap support (Supplementary Figure S2). Furthermore, minor differences between sexes were noticed. A total of 48 areas showed significant negative IBD residuals in males, which is not only larger in quantity...
but also more scattered in distribution than the 33 such areas detected in females.

Demographical history
Because the population had little structure, all samples were treated as one population. The stairway plot revealed 3 effective population size reduction events within the past 800,000 years (Figure 4). The first population size decrease occurred at approximately 60,000 years ago, with the effective population size greatly reduced from \( \sim 2 \) to \( \sim 0.4 \) million individuals. The second reduction started approximately 15,000 years ago and reduced the population size to approximately 80,000 individuals. The most recent population decline occurred at approximately 3,000 years ago, which reduced the effective population size to the current approximately 20,000 individuals.

Discussion
Unsurprisingly, the Shedao pit-viper has low genetic diversity. Our analysis of genome-wide 13,148 SNPs consistently suggested that all individuals on the island are closely related and the population is lack of subdivision (NJ tree, PCoA, and clustering analysis; Figure 1), as well as the overall low levels of nucleotide diversity (\(< 0.001\)) and heterozygosity (\(0.3378\)). These are consistent with the results from microsatellite DNA and mtDNA data (Wang et al. 2015). The reduction in genetic variation is similar to other insular species (e.g., Bothrops jararaca complex, Podarcis erhardii, and Pelophylax nigromaculatus; Frankham 1997; Grazziotin et al. 2006; Hurston et al. 2009; Wang et al. 2014).

Even at a fine spatial scale, population genetic patchiness is hardly detectable. The gradual and continuous decrease in autocorrelation coefficients in the correlograms (Figure 3A,B) suggested a lack of discrete patches of genetic variability. Although strong genetic structure across space could also cause the pattern (Diniz-Filho and Telles 2002), it is unlikely considering results from all other analyses. Furthermore, the patch size of 0.53 km for all individuals across the island (Figure 3A) is quite large because the maximum length of the island is merely 1.5 km (Li 1995). The dispersal distance of Shedao pit-vipers revealed by a radio-tracking study is generally within 50 m, rarely >100 m (Shine et al. 2003). This size of home range is compatible with the estimation of patch size, as a genetic patch reflects the genetic relatedness of interrelated individuals rather than that of neighbor individuals.
Additionally, the DResD analysis did not detect any genetic barrier or corridor. We only detected some weak patchiness with the IBD residuals (Figure 3). Within numerous small areas, individuals are genetically more similar than predicted by IBD. Overall, the population is nearly panmictic and gene flow is evenly distributed across the island. The island population shows no sign of inbreeding. The observed and expected heterozygosities are nearly identical. Also, the average inbreeding $F$ was negative (Mean $F_{\text{uni}} = -0.0116$). This is consistent with early findings from microsatellite DNA data (Wang et al. 2015). The large number of individuals in the island population (approximately 20,000 individuals; Li et al. 2007), the small size of the island (0.73 km$^2$), and the consequent lack of population structure have likely all contributed to the lack of inbreeding in this species. Furthermore, the strong seasonal constraints on reproduction likely also play an important role (Yang 1983; Huang 1989). Lehmann and Perrin (2003) showed in model research that species under strong time constraints on reproduction could be tolerant of inbreeding. Nevertheless, the small dispersal distance observed from radiotelemetry (50 m; Shine et al. 2003) appears to be contradictory. We propose an additional hypothesis that male-biased dispersal is likely to also play a role in inbreeding avoidance in the Shedao pit-viper. Sex-biased dispersal is commonly employed in many organisms as a strategy for inbreeding avoidance (Pusey 1987).

As tolerance of inbreeding is often different between sexes (Parker 1979; Waser et al. 1986; Kokko and Ots 2006), sex-specific dispersal is often being selected (Bengtsson 1978; Perrin and Mazalov 2000; Guillaume and Perrin 2009). Our results suggest distinct male-biased dispersal patterns in the Shedao pit-viper. There is a clear contrast between the sexes in spatial autocorrelation (Figure 2), where females exhibited a higher level of relatedness within a larger geographical scale than males do (Banks and Peakall 2012). Also, the distribution of areas that deviated from IBD was more sporadic and fragmented in males (Figure 3 and Supplementary Figure S2). With extremely high population density on the island (Li et al. 2007), other strategies such as kin recognition coupled with mate choice could have also evolved (Blouin and Blouin 1988; Brown and Eklund 1994). Inbreeding is known to have detrimental effects on organisms’ fitness (Keller and Waller 2002), as such has important conservation implications. More studies on mate and resource competition and radiotrack/mark-recapture will be helpful to examine the association between kin recognition and sex-biased dispersal in G. shedaoensis.

The historical distribution of the Shedao pit-viper and its population size are much larger than those at present time. The species diverged from its closest relative, the black eyebrow pit-viper G. intermedius at approximately 2.5 million years ago (Lu et al. 2020) and the Shedao island was...
separated from the continent only approximately 1 million years ago (Li 1995). Furthermore, the Shedao Island is only 5.3 km from the nearest continent and it was reconnected to the continent during several episodes of sea-level drops (Li 1995). Demographic analysis suggested its historical effective population size was 100 times the current size at approximately 60,000 years ago (Figure 4). Considering the size of the population, the distribution range of the species was unlikely

Figure 3. DResD plots for females and males on the Shedao Island. The IBD residual values represent the difference in genetic dissimilarity between the true value of a particular sample pair and the IBD best-fit line. Blue lines indicate areas where the residual values are significant ($P < 0.05$).

Figure 4. Inferred effective population size changes for the *G. shedaoensis* population within the last 800,000 years. The bold line shows the median estimate; the dark gray and light gray lines represent 85% CI and 95% CI, respectively, based on 200 bootstrap replicates.
Acknowledgments
We are grateful to Dr W. Yang of the Lund University and Y. Gong of the Chengdu Institute of Biology for writing suggestions. Most of the analyses benefit from the computing platform funded by the Science and Technology Support Project of Sichuan Province, China (2016NZ0103).

Funding
This project was supported by Key Laboratory of Southwest China Wildlife Resources Conservation (Ministry of Education), China West Normal University (XNYB18-02), and Nature Conservation of Snake Island and Laotieshan Mountain, Dalian, China (Y8Y3041).

Availability of Data and Materials
Demultiplexed fastq files of ddRADseq sequences have deposited at NCBI SRA under Bioproject PRJNA684801 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA684801).

Authors’ Contributions
G.W. collected the data, conducted the analysis, and drafted the manuscript. L.J., Y.W., and X.W. collected the samples and provided important field and lab experiment conditions. J.F. optimized the premise and edited the language. Y.Q. conceived the project, and finalized the manuscript. All authors participated in project design and manuscript preparation.

Conflicts of Interest Statement
No conflict of interest statement to declare.

Supplementary Material
“Supplementary material can be found at https://academic.oup.com/cz.”

References
Alexander DH, Novembre J, Lange K, 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19:1655–1664.
Arct A, Rutkowski J, Martyka R, Drobiński SM, Cichoń M, 2010. Kin recognition and adjustment of reproductive effort in zebra finches. Biology Lett 6:762–764.
Banks SC, Peacock ROD, 2012. Genetic spatial autocorrelation can readily detect sex-biased dispersal. Mol Ecol 21:2092–2105.
Hefner R, Hersteinsson P, Angerbjörn A et al., 2011. Ecological genetics of pronghorn antelope (Antilocapra americana) – insights into the maintenance of genetic diversity. Mol Ecol 20:5348–5358.
Grazziotin FG, Monzel M, Echeverrigaray S, Bonatto SL, 2006. Phylogeography of the Bothrops jararaca complex (Serpentes: Viperidae): past fragmentation and island colonization in the Brazilian Atlantic Forest. Mol Ecol 15:3969–3982.
Weigelt P, Steinbauer MJ, Cabral JS, Kreft H, 2016. Late Quaternary climate change shapes island biodiversity. *Nature* 532:99–102.

Whittaker RJ, Fernández-Palacios JM, 2007. *Island Biogeography: Ecology, Evolution, and Conservation*. Oxford: Oxford University Press.

Wickham H, Chang W, 2008. ggplot2: an implementation of the grammar of graphics. R package version 0.7. Available from: http://CRAN.R-project.org/package=ggplot2.

Yang J, Lee SH, Goddard ME, Visscher PM, 2011. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 88:76–82.

Yang M, 1983. Sexual cycle of *Agkistrodon shedaoensis* Zhao. *Acta Herpetol* 2:17–25.

Zhao EM, Huang MH, Zong Y, 1998. *Fauna Sinica Reptilia, Squamata: Serpentes*. Vol. 3. Beijing, China: Science Press. 1–8.