Population Genetic Structure of a Viviparous Sand Lizard, the *Phrynocephalus forsythii* in the Tarim Basin, Xinjiang of China

Jiabao Duan ¹, Jingxian Nong ¹, Luoxue Jiang ¹, Tianying Chen ¹, Wen Zhong ¹, Wei Zhao ², Yue Qi ², Penghui Guo ¹,* and You Li ¹,³,*

¹ College of Life Science and Engineering, Northwest Minzu University, Lanzhou 730030, China; p202115153@stu.xbmu.edu.cn (J.D.); p210811925@stu.xbmu.edu.cn (J.N.); y210830268@stu.xbmu.edu.cn (L.J.); y220830394@stu.xbmu.edu.cn (T.C.); y192130516@stu.xbmu.edu.cn (W.Z.)
² Gansu Key Laboratory of Biomonitoring and Bioremediation for Environmental Pollution, School of Life Sciences, Lanzhou University, Lanzhou 730030, China; zhaowei@lzu.edu.cn (W.Z.); qiy20@lzu.edu.cn (Y.Q.)
³ Gansu Tech Innovation Center of Animal Cell, Biomedical Research Center, Northwest Minzu University, Lanzhou 730030, China
* Correspondence: skyguoph@xbmu.edu.cn (P.G.); 284142549@xbmu.edu.cn (Y.L.)

Abstract: The desert ecosystem occupies an important position in the composition of global biodiversity. The Tarim Basin is located in south Xinjiang of China and has the world’s second largest mobile desert, the Taklimakan Desert. As an endemic species in this region, *Phrynocephalus forsythii* has been demonstrated to have a potentially high extinction risk due to climate change. In order to understand the overall genetic status and provide accordant conservation strategies for the species, we investigated the genetic diversity and population structure of *P. forsythii* from 15 sites in the Tarim Basin using 21 highly polymorphic microsatellite markers. We found significant genetic structure across the study region. We also revealed generally low levels of gene flow between the 25 sites, suggesting individual dispersal and migration may be restricted within populations. In addition, geographical distance and ambient temperature might be important factors in explaining the observed genetic structure. Our results will provide a scientific basis for the future protection of *P. forsythii* in this area, as well as an important reference for the conservation and management of biodiversity in desert ecosystems.

Keywords: *Phrynocephalus forsythii*; microsatellite; genetic diversity; population structure

1. Introduction

Desert ecosystems are among the most dominant ecosystems and an essential element of the terrestrial landscape [1]. The natural environment of desert ecosystems is exceedingly fragile, and its formation and development are the consequence of arid climate, surface processes, and the evolution of vegetation [2,3]. Desertification is a major issue in arid and semi-arid areas, threatening approximately 41 percent of the world surface area and more than 38 percent of the population [4,5]. Global Desertification Vulnerability Index (GDVI) showed that surrounding regions are at high risk of desertification. Meanwhile, the Representative Concentration Pathways (RCPs) predicted an increased risk of desertification mainly in China and northern India [6,7]. As one of the most widespread ecosystems in China, desert ecosystems are also sensitive to global change [8]. On the other hand, this complex natural environment has nurtured a rich diversity of flora and fauna, and its organisms are unique compared to other ecosystems and occupy an important position in the composition of global biodiversity [9].

The Tarim Basin, located in south Xinjiang Province, is the largest inland basin in China. It nourishes the world’s second largest mobile desert, the Taklimakan Desert [10,11]. The Tarim Basin is 1500 km long from east to west and 600 km wide from north to south, covering an area of about 530,000 km² and surrounded by altitudes ranging from 800 to
It is characterized by a typical arid climate with low precipitation and high evaporation [12]. Due to the extreme harshness of the ecological environment, the process of aridification in the Taklimakan Desert and its climate changes have a significant impact on the population structure of species [13,14]. Currently, research in the Tarim Basin and the hinterland of the desert has mostly focused on plants, especially cherished and endangered plants [15,16], and has concentrated on the evolutionary history and phylogenetic processes [17–19]. However, research on the genetic diversity, structure, and differentiation of animal populations in the hinterland and surrounding areas of the Taklimakan Desert is relatively lacking. Zhang et al. (2022) revealed the genetic mechanism of *Ovis* and environmental adaptability of local sheep breeds in the Taklimakan Desert by analyzing the genome and transcriptome with different agro-geographical characteristics, providing a theoretical basis for the development and conservation of sheep breed germplasm resources in extreme desert environments [20]. The desert and semi-desert inhabited ungulate *Gazella subgutturosa* has declined sharply due to natural factors and habitat fragmentation. Studies on this species in Xinjiang showed that their population had a low level of genetic diversity, and there was a certain genetic differentiation among the populations, underscoring the urgent need to strengthen the conservation of genetic diversity of *G. subgutturosa* in Xinjiang [21].

As predominant species in desert ecosystems, reptiles have successfully survived in arid environments by improving morphological and physiological characteristics such as temperature regulation, water balance, and movement processes [22,23]. *Phrynocephalus forsythii*, belonging to the family Agamidae, is an endemic species of the Tarim Basin and is a viviparous lizard that inhabits sparse scrub or the Gobi area. Research on evolutionary history and phylogeny of *P. forsythii* is relatively mature [13,24], whereas studies on its genetic structure and population differentiation are still rare.

The amount of genetic diversity reflects the diversity of genetic factors and their combinations that determine biological traits. The richer the genetic diversity, the richer the morphological, behavioral, physiological, and other characteristics of the population and the stronger the adaptability to the environment. Therefore, the amount of genetic diversity can be used as an indicator of population adaptability to its environment [25]. Although *P. forsythii* was not currently identified as endangered, it was demonstrated with the potential for a high risk of extinction under climate change [26]. Sinervo et al. (2018) developed an eco-physiological model for predicting extinction risk under climate change for 20 *Phrynocephalus* lizards and revealed 12 of them, including *P. forsythii*, being at high extinction risk due to exceeded thermal limits. Additionally, by building a climate refugia map using locations of Nature Reserves and extinction areas of China, the author demonstrated that current distribution sites of *P. forsythii* were beyond the refugia map [26]. Baseline genetic data on genetic diversity and structure, therefore, is crucial for understanding the overall genetic status and providing accordant conservation strategies for the species. Our study specifically aimed to investigate the genetic diversity and population structure of *P. forsythii* from 15 sites in the Tarim Basin using 21 highly polymorphic microsatellite markers. We also incorporated geography distance data and climatic factors in the analyses to preliminarily infer the potential influence on genetic structure. The results are expected to complement and enrich the basic genetic data of *P. forsythii* in the Tarim Basin, providing the scientific and theoretical basis for the conservation of the species and contributing an important reference for the conservation and management of biodiversity in desert ecosystems.

2. Materials and Methods

2.1. Study Area and Sample Collection

Based on the reported distribution areas of *P. forsythii* and our own field observation [27,28], a total of 171 individuals were sampled from 15 sampling sites around the Tarim Basin region (Figure 1; Table 1). After measuring basic data such as body length, tail length, and head length, a piece of tissue (~5 mm$^2$) was cut from the posterior end of
the tail, and one toe was also clipped for individual identity. Tissues were preserved in a solution consisting of a volume ratio of 1:1 with anhydrous ethanol and normal saline for subsequent molecular analyses.

Figure 1. Map of sampling sites of P. forsythii in the Tarim Basin. Map was edited using data downloaded from the National Foundation Geographic Information System (NFGIS) (https://www.webmap.cn/ (accessed on 16 March 2023)).

Table 1. Sampling information and genetic diversity parameters.

<table>
<thead>
<tr>
<th>Site Abbreviation</th>
<th>Sampling Site</th>
<th>N</th>
<th>Ne</th>
<th>Ho</th>
<th>He</th>
<th>Fis</th>
<th>A</th>
<th>AR</th>
<th>IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF</td>
<td>Minfeng</td>
<td>12</td>
<td>5.946</td>
<td>0.594</td>
<td>0.765</td>
<td>0.224</td>
<td>184</td>
<td>2.540</td>
<td>0.344</td>
</tr>
<tr>
<td>CL</td>
<td>Cele</td>
<td>14</td>
<td>6.631</td>
<td>0.515</td>
<td>0.751</td>
<td>0.314</td>
<td>205</td>
<td>2.420</td>
<td>0.394</td>
</tr>
<tr>
<td>YT</td>
<td>Yutian</td>
<td>16</td>
<td>4.360</td>
<td>0.394</td>
<td>0.672</td>
<td>0.414</td>
<td>165</td>
<td>2.040</td>
<td>0.505</td>
</tr>
<tr>
<td>HT</td>
<td>Hetian</td>
<td>14</td>
<td>4.860</td>
<td>0.486</td>
<td>0.728</td>
<td>0.332</td>
<td>166</td>
<td>2.190</td>
<td>0.428</td>
</tr>
<tr>
<td>PS</td>
<td>Pishan</td>
<td>18</td>
<td>8.915</td>
<td>0.494</td>
<td>0.829</td>
<td>0.404</td>
<td>274</td>
<td>2.440</td>
<td>0.429</td>
</tr>
<tr>
<td>YC</td>
<td>Yecheng</td>
<td>13</td>
<td>6.092</td>
<td>0.418</td>
<td>0.759</td>
<td>0.449</td>
<td>178</td>
<td>2.010</td>
<td>0.496</td>
</tr>
<tr>
<td>YJS</td>
<td>Yingjisha</td>
<td>8</td>
<td>3.663</td>
<td>0.466</td>
<td>0.613</td>
<td>0.240</td>
<td>101</td>
<td>1.850</td>
<td>0.474</td>
</tr>
<tr>
<td>BC</td>
<td>Bachu</td>
<td>8</td>
<td>2.790</td>
<td>0.207</td>
<td>0.427</td>
<td>0.603</td>
<td>75</td>
<td>1.160</td>
<td>0.713</td>
</tr>
<tr>
<td>XH</td>
<td>Xinhe</td>
<td>3</td>
<td>1.341</td>
<td>0.135</td>
<td>0.268</td>
<td>0.819</td>
<td>30</td>
<td>0.810</td>
<td>0.740</td>
</tr>
<tr>
<td>RQ</td>
<td>Ruqiang</td>
<td>10</td>
<td>4.217</td>
<td>0.371</td>
<td>0.638</td>
<td>0.419</td>
<td>127</td>
<td>1.990</td>
<td>0.505</td>
</tr>
<tr>
<td>YLN</td>
<td>Yulinan</td>
<td>11</td>
<td>3.381</td>
<td>0.277</td>
<td>0.558</td>
<td>0.543</td>
<td>102</td>
<td>1.350</td>
<td>0.655</td>
</tr>
<tr>
<td>LT</td>
<td>Luntai</td>
<td>2</td>
<td>2.108</td>
<td>0.381</td>
<td>0.435</td>
<td>0.210</td>
<td>46</td>
<td>1.730</td>
<td>0.914</td>
</tr>
<tr>
<td>LTN</td>
<td>Luntainan</td>
<td>20</td>
<td>7.590</td>
<td>0.356</td>
<td>0.780</td>
<td>0.544</td>
<td>248</td>
<td>2.100</td>
<td>0.543</td>
</tr>
<tr>
<td>ALE</td>
<td>Alae</td>
<td>18</td>
<td>6.003</td>
<td>0.370</td>
<td>0.752</td>
<td>0.508</td>
<td>210</td>
<td>2.120</td>
<td>0.535</td>
</tr>
<tr>
<td>JS</td>
<td>Jiashi</td>
<td>4</td>
<td>2.680</td>
<td>0.381</td>
<td>0.489</td>
<td>0.290</td>
<td>68</td>
<td>1.830</td>
<td>0.468</td>
</tr>
</tbody>
</table>

Notes: N: Sample size, Ne: Effective number of alleles, Ho: Observed heterozygosity, He: Expected heterozygosity, Fis: Inbreeding coefficient, A: Number of alleles, AR: Allelic richness, IR: Internal relatedness.

2.2. DNA Extraction and PCR Amplification

The SteadyPure Universal Genomic DNA Extraction Kit (AGBIO, Changsha, China) was used to extract the whole genome in accordance with the manufacturer’s instructions. Twenty-one highly polymorphic fluorescently labelled microsatellite markers were ampli-
fied, including ten (PVMS11, PVMS12, PVMS15, PVMS18, PVMS20, PVMS32, PVMS35, PVMS38, and PVMS39) developed by Zhan and Fu (2009) [29] and eleven (Phr58 excluded) by Urquhart et al. (2005) [30]. PCR amplification was conducted using the MRT (Multiplex-Ready Technology) method [31]. PCR reactions were performed in a 12 µL total volume containing 2.4 µL of 5 × buffer (Mg²⁺ free), 0.06 µL of 5 U/µL Immolase DNA polymerase, 0.09 µL of 10 µmol/L fluorescently labeled upstream primer tag F (FAM, NED, VIC, PET), 0.09 µL of 10 µmol/L downstream primer tag R, 2.4 µL of 0.4 µmol/L locus-specific primers, 2 µL of template DNA, and the rest was supplemented with ddH₂O. Amplification was performed with cycling conditions as follows: pre-denaturation at 95 °C for 10 min; 5 cycles of denaturation at 92 °C for 60 s, annealing at 50 °C for 90 s, and extension at 72 °C for 60 s; 20 cycles of denaturation at 92 °C for 30 s, annealing at 63 °C for 90 s, and extension at 72 °C for 60 s; 40 cycles of denaturation at 92 °C for 15 s, annealing at 54 °C for 30 s, and extension at 72 °C for 30 s; and a final extension at 72 °C for 30 min, preserved at 4 °C. All PCR products were detected by 1.4% agarose gel electrophoresis and sequenced by capillary electrophoresis (GENEWIZ, Suzhou, China).

2.3. Genetics data Processing and Analysis

2.3.1. Genetic Diversity

Sequencing results were scored using GeneMarker® HID [32]. Based on 1000 bootstraps and a 95% confidence interval, the possibility for substantial allele dropout and stuttering was calculated using Micro-checker version 2.2.3 [33]. Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD) among loci were calculated by Genepop version 4.7.5 [34]. Sequential Bonferroni corrections were applied to adjust significance values for multiple comparisons [35]. GenAlEx version 6.51 [36] was used to generate allelic frequencies and to estimate genetic diversity parameters, including effective number of alleles (Ne), observed heterozygosity (Ho), and expected heterozygosity (He). The mean number of alleles per locus (A), inbreeding coefficient (Fis), and allelic richness (AR) for each population were obtained using the R package ‘diveRsity’ (R version 4.2.1) [37,38]. Internal relatedness (IR) was also calculated as an estimate of parental relatedness [39] in an R extension package, Rhh (R version 4.2.1) [38,40]. Fixation index (FST) was used to investigate patterns of differentiation among populations. Calculation of actual differentiation DEST [41] also was performed in an R package ‘DEMEtics’ (R version 4.2.1) [38,42], as FST may underestimate differentiation when using highly polymorphic markers such as microsatellites [41].

2.3.2. Genetic Differentiation and Population Structure

Population genetic structure was evaluated using Bayesian clustering analysis of individual genotypes in various ways. Initially, we implemented the Bayesian genetic clustering algorithm to determine population genetic structure in STRUCTURE version 2.3.4 [43]. We ran the analysis with a burn-in of 10,000 and 300,000 Markov Chain Monte Carlo (MCMC) steps after the burn-in, using an admixture model and correlated allele frequencies without contributing any information on geographic location. The value of K was tested from 1 to 15 due to the possibility of each location representing a distinct genetic cluster, and 10 replicates of each K were run. The best K was determined using STRUCTURE HARVESTER web version 0.6.94 [44], according to the AK proposal by Evanno et al. [45]. For the 10 runs of the best K, we used CLUMPP version 1.1.2 [46] to average the membership probabilities, and DISTRUCT 1.1 [47] to visualize the final results.

Discriminant Analysis of Principal Components (DAPC) is another widely used strategy for analyzing the clustering relationships [43,48]. When the level of genetic differentiation is low, it is a multivariate analysis that really is independent of the HWE hypothesis and provides a higher resolution of population structure [48]. We performed DAPC analysis and determined the best clustering K by using R package ‘adegenet’ (R version 4.2.1) [38,49]. The value of K was set from 1 to 15, and 10 replicates were run for each K. Using the Bayesian Information Criterion (BIC) parameter values for each K as reference, the best K
was determined by the trend in successive BIC values (BIC decreased sharply to decreased slightly or increased), according to Ward’s hierarchical clustering methods [50]. The number of principal components (PC) used in the DAPC analysis is critical, and the effect of different PC numbers on the results can be considerable. Consequently, cross-validation was employed to determine the best number of PCs to retain explain majority of the sources of variation.

2.3.3. Spatial Scales of Genetic Variation

The correlation between genetic distance and Euclidean geographic distance was evaluated using the Mantel test and Mantel correlogram to determine the impact of geographical distance on genetic differences in the population. The Mantel test was used to determine the effect of isolation by distance (IBD) using GenAlEx 6.51 [36]. Spatial autocorrelation analysis was also performed using 50 km and 100 km distance classes to further study the spatial distribution scale of genetic variation (statistical calculations of spatial autocorrelation were based on 95% confidence intervals defined by 1000 random permutations).

As a complement to the Mantel test, redundancy analysis (RDA) was used to analyze the relationship between genetic structure and explanatory variables. RDA is a multiple linear regression analysis with greater validity than the Mantel test in species–environment relationships with diversity [51]. The allele frequency at each microsatellite locus was used as the response variable in the RDA analysis, and four variables related to distance were used as explanatory variables: geographic coordinates, minimum distance to nearest population (DN), degree of isolation (DI, calculated as the average distance between a population and its three closest neighboring populations), and mean pairwise distance (MPD). The RDA analysis was performed in the “vegan” package of the R (R version 4.2.1) [38,52], with one distance variable as the main explanatory variable and all other distance variables as covariates.

2.3.4. Correlation of Climatic Factors with Genetic Structure

To evaluate the influence of climatic factors on the genetic differences of *P. forsythii*, we used RDA to analyze the connection between climatic factors and genetic structure. We downloaded climate data from the World Climate Website (WorldClim, https://www.worldclim.org/ accessed on 20 November 2023) and used ArcGIS to extract 19 climatic variables from the 15 sampling sites in the Tarim Basin (Table A1). After screening using the Pearson correlation analysis to exclude highly correlated factors, the chosen climatic factors were employed as explanatory variables, and the allele frequencies of each microsatellite locus were used as response variables for RDA analysis using the R package ‘vegan’ (R version 4.2.1) [38,52].

3. Results

3.1. Genetic Diversity

The micro-checker showed no evidence of scoring errors associated with stuttering, large allele dropout, or null alleles. When 21 microsatellite loci from 15 *P. forsythii* populations were evaluated individually, 120 HWE results out of 315 detections were still significantly deviated after sequential Bonferroni correction, whereas no deviation was observed for XH, LT, and JS populations. There were 3150 pairwise locus combinations of linkage disequilibrium test, and 78 were found. Except for YJS, BC, XH, LT, and JS, remaining populations had loci that exhibited linkage disequilibrium. However, all pairwise comparison results were not consistent for the 15 populations (no pairing comparison showed linkage disequilibrium for 15 populations simultaneously). These 21 loci were therefore considered as independently segregating loci.

A total of 21 microsatellite loci detected a total of 2179 alleles in 15 populations of *P. forsythii*. The number of alleles (*A*) ranged from 30 (XH) to 274 (PS) with an average of 145.267 (Table 1). The observed heterozygosity (*Ho*) and expected heterozygosity (*He*)
Diversity 2023, 15, x FOR PEER REVIEW 7 of 14

varied from 0.135 (XH) to 0.594 (MF) and from 0.268 (XH) to 0.829 (PS), respectively (Table 1). Inbreeding coefficient (\(F_{is}\)) varied from 0.210 (LT) to 0.819 (XH) with a mean number of 0.421 (Table 1). Allelic richness (AR) varied from 0.810 (XH) to 2.540 (MF), and internal relatedness (IR) varied from 0.344 (XH) to 0.914 (LT) (Table 1).

3.2. Genetic Differentiation and Population Structure

Paired \(F_{ST}\) values varied from 0 to 0.210, with 78 results of \(p\) values being significant after sequential Bonferroni correction (Table 2). Overall, there was a greater genetic divergence between populations (\(F_{ST} < 0.25\); Table 2). Paired \(D_{ST}\) values ranged from −10.079 (YJS-MF) to 0.788 (XH-PS), with 44 significant values after correction (Table 2). The \(D_{ST}\) values were generally higher than \(F_{ST}\), as expected (Table 2).

Table 2. Pairwise \(F_{ST}\) values (lower triangle) and pairwise \(D_{ST}\) values (upper triangle).

<table>
<thead>
<tr>
<th>Population</th>
<th>MF</th>
<th>CL</th>
<th>YT</th>
<th>HT</th>
<th>PS</th>
<th>YC</th>
<th>YJS</th>
<th>BC</th>
<th>XH</th>
<th>RQ</th>
<th>YLN</th>
<th>LT</th>
<th>LTN</th>
<th>ALE</th>
<th>JS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF</td>
<td>0.72</td>
<td>0.06</td>
<td>0.70</td>
<td>0.83</td>
<td>0.45</td>
<td>0.78</td>
<td>0.07</td>
<td>0.74</td>
<td>0.92</td>
<td>0.34</td>
<td>0.61</td>
<td>0.29</td>
<td>0.34</td>
<td>0.10</td>
<td>0.93</td>
</tr>
<tr>
<td>CL</td>
<td>0.13</td>
<td>0.72</td>
<td>0.17</td>
<td>0.30</td>
<td>0.40</td>
<td>0.07</td>
<td>0.06</td>
<td>0.23</td>
<td>0.11</td>
<td>0.08</td>
<td>0.06</td>
<td>0.16</td>
<td>0.27</td>
<td>0.05</td>
<td>0.28</td>
</tr>
<tr>
<td>YT</td>
<td>0.18</td>
<td>0.06</td>
<td>0.29</td>
<td>0.53</td>
<td>0.60</td>
<td>0.28</td>
<td>0.28</td>
<td>0.37</td>
<td>0.14</td>
<td>0.09</td>
<td>0.12</td>
<td>0.11</td>
<td>0.09</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>HT</td>
<td>0.05</td>
<td>0.08</td>
<td>0.27</td>
<td>0.46</td>
<td>0.63</td>
<td>0.03</td>
<td>0.14</td>
<td>0.27</td>
<td>0.11</td>
<td>0.16</td>
<td>0.15</td>
<td>0.16</td>
<td>0.16</td>
<td>0.04</td>
<td>0.17</td>
</tr>
<tr>
<td>PS</td>
<td>0.07</td>
<td>0.06</td>
<td>0.05</td>
<td>0.36</td>
<td>0.33</td>
<td>0.10</td>
<td>0.05</td>
<td>0.09</td>
<td>0.12</td>
<td>0.08</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>YC</td>
<td>0.13</td>
<td>0.08</td>
<td>0.11</td>
<td>0.14</td>
<td>0.19</td>
<td>0.07</td>
<td>0.05</td>
<td>0.16</td>
<td>0.14</td>
<td>0.05</td>
<td>0.06</td>
<td>0.11</td>
<td>0.11</td>
<td>0.06</td>
<td>0.16</td>
</tr>
<tr>
<td>YJS</td>
<td>0.09</td>
<td>0.12</td>
<td>0.13</td>
<td>0.13</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>BC</td>
<td>0.37</td>
<td>0.21</td>
<td>0.75</td>
<td>0.96</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>XH</td>
<td>0.48</td>
<td>0.35</td>
<td>0.75</td>
<td>0.96</td>
<td>0.98</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>RQ</td>
<td>0.67</td>
<td>0.52</td>
<td>0.84</td>
<td>0.92</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>YLN</td>
<td>0.82</td>
<td>0.66</td>
<td>0.94</td>
<td>0.96</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>LT</td>
<td>0.97</td>
<td>0.83</td>
<td>0.95</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>LTN</td>
<td>0.99</td>
<td>0.95</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>ALE</td>
<td>0.99</td>
<td>0.96</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>JS</td>
<td>0.99</td>
<td>0.97</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Notes: Significant \(p\) value bolded after Bonferroni correction.

Bayesian clustering analysis using STRUCTURE identified four distinct clusters (\(K = 4\), Figure 2) using the AK criterion. MF clustered individually as one blue group. CL, YT, and HT populations clustered into one pink group; PS, YC, YJS, and XH groups clustered as one green group; and the rest (BC, RQ, YLN, LT, LTN, ALE, and JS) clustered together as the red group.

Figure 2. STRUCTURE clustering diagram.

Cross-validation of DAPC results showed that PC = 26 had the highest percentage of correctly predicted subsamples with the lowest mean square error and was able to explain most of the sources of variation (42.6%). The optimal \(K\) value was determined as 5 based on continuous BIC value trends (Figure 3a), and the 15 P. forsythii populations were divided into five groups (Figure 3b). To facilitate comparison of the cluster assignments of the samples at different \(K\) values, the posterior probabilities for \(K\) values from 2 to 5 were output as bar charts (Figure 3c). The bar chart results indicated that the posterior probabilities of cluster assignment for the samples were all high (97.5%; Figure 3c). The results showed that MF and YT each clustered into one singular group (purple and green, respectively) (Figure 3b,c). ALE, BC, and JS clustered into one pink group, CL and HT clustered into one orange group, and the rest (LT, LTN, PS, RQ, XH, YC, YJS, and YLN) clustered into one dark-green group.
Figure 2. STRUCTURE clustering diagram.

Cross-validation of DAPC results showed that PC = 26 had the highest percentage of correctly predicted subsamples with the lowest mean square error and was able to explain most of the sources of variation (42.6%). The optimal $K$ value was determined as 5 based on continuous BIC value trends (Figure 3a), and the 15 $P. forsythii$ populations were divided into five groups (Figure 3b). To facilitate comparison of the cluster assignments of the samples at different $K$ values, the posterior probabilities for $K$ values from 2 to 5 were output as bar charts (Figure 3c). The bar chart results indicated that the posterior probabilities of cluster assignment for the samples were all high (97.5%; Figure 3c). The results showed that MF and YT each clustered into one singular group (purple and green, respectively) (Figure 3b,c). ALE, BC, and JS clustered into one pink group, CL and HT clustered into one orange group, and the rest (LT, LT N, PS, RQ, XH, YC, YJS, and YLN) clustered into one dark-green group.

Figure 3. Results of DAPC showing the genetic structure ((a): Plots showing BIC values with different $K$; (b): Results of DAPC; (c): Barplots of the posterior probabilities of group assignment for each sample, $K = 2–5$).

3.3. Spatial Scales of Genetic Variation

Isolation by distance (IBD) analysis showed that the genetic distance (GD) and geographical distance (GGD) of the $P. forsythii$ populations were significantly and positively correlated in the sampling area of the Tarim Basin ($R^2 = 0.0054$, $p = 0.021$), which indicated that the population was gradually accumulating genetic variation during the dispersal process. Meanwhile, the results of spatial autocorrelation (Figure 4) based on Mantel correlogram demonstrated that the genetic correlation values ($r$) among individuals in the population at 50 km distance class were significantly negative up to 100 km and positive above 100 km, and $r$ also remained stable around 0 after greater than 100 km (Figure 4a). The results of analyses using a larger distance class of 100 km showed a very similar pattern of correlations to that for the 50 km distance class (Figure 4b).

RDA analysis with the full model showed a significant correlation between genetic variation and spatial variables ($p = 0.006$, Bonferroni corrected $p = 0.004$). The total contribution of the four spatial variables to the genetic variation was 44.1% (13.0% after $R^2$ correction, Figure 5a). Controlling for the effects of the remaining spatial variables (DN, DI, and MPD were used as covariates), RDA analysis with the partial model revealed an explanation of 19.0% (1.5% after $R^2$ correction) when observing the effect of geographic coordinates on genetic variation. The partial full model for the three residual spatial variables (DN, DI, and MPD) showed explanations of 13.2%, 9.2%, and 8.1%, respectively (of 1.3%, 0.4%, and 0.2% after $R^2$ correction, respectively).
Figure 4. Correlograms showing genetic correlation ($r$) as a function of distance ((a): 50 km distance class; (b): 100 km distance). The 95% confidence intervals (dashed lines) were determined by 1000 permutations (U-upper level, L-lower level). Error bars of each estimate of $r$ bounding the 95% confidence intervals were determined by 1000 bootstraps.

Figure 5. Triplots of full model RDA analysis: (a) spatial factors; (b) climatic factors. The allele frequencies are displayed in red and the populations are denoted as black open circles. The blue vectors show how explainable variables fall along that RDA space. (x,y) = geographic coordinates; DN = minimum distance to nearest population; DI = degree of isolation; MPD = mean pairwise distance.
3.4. Correlation of Climatic Factors with Genetic Structure

In total, 12 climatic factors were retained (Bio2, Bio4, Bio5, Bio6, Bio8, Bio10, Bio11, Bio13, Bio14, Bio15, Bio16, and Bio17, see Table A1 for full names of the 12 climatic factors) for full model RDA analysis (Figure 5b), and the result showed a non-significant correlation ($p = 0.433$, Bonferroni-corrected $p = 0.432$). The 12 climate factors explained 86.2% of the total genetic variation (3.1% after $R^2$ correction), with Bio2 explaining the highest 11.0% (0.3% after $R^2$ correction), followed by Bio15 (5.3%, 0.2% after $R^2$ correction). Results of partial model when the effects of a factor on genetic variation were observed controlling for the remaining climatic variables (with the other 11 climate factors as covariates) are shown in Table A1.

4. Discussion

We used 21 microsatellite loci in the current study to investigate genetic structure and level of gene flow of the *P. forsythii* in the Tarim Basin region. We found significant genetic structure across the study region, with roughly consistent results from STRUCTURE and DAPC. Along with the results from IBD and RDA analyses, our study revealed generally low levels of gene flow in the study area, suggesting individual dispersal and migration may be restricted to within populations.

The results of HWE test showed that 120 tests deviated from HWE after Bonferroni correction, and no deviation was found in populations with small sample sizes (XH, LT, and JS). The HWE deviation mainly came from populations with relatively large sample sizes, which suggested that the effectiveness of HWE test had limited validity in populations with small samples. Similarly, the small sample size could be also the reason why the XH and LT populations showed significant differences in He. Micro-checker results demonstrated that there were no null alleles in the dataset. Therefore, the HWE deviation is likely to be caused by the Wahlund effect, with causal factors including inbreeding and possible genetic structure.

Regarding the genetic mechanism of inbreeding decline, researchers tend to support the dominant effect theory, which suggests that some recessive pure homozygotes have deleterious effects, and inbreeding increases the proportion of these harmful homozygotes, which will cause the phenotypic performances to exhibit a decrease [53]. The inbreeding coefficient was calculated based on HWE and reflected the relationship between $H_0$ and $H_e$. Our $Fis$ results showed that the inbreeding coefficients of 15 populations were higher than 0, indicating that there was a lack of heterozygotes in the population, which was consistent with the results of heterozygous defects in HWE detection, suggesting that the populations may be in the state of inbreeding depression [54]. Negative and a value closer to $-1.0$ of $IR$ means that there are fewer inbred individuals and less inbreeding in the population, whereas positive and closer to $+1.0$ means more close relatives and higher inbreeding [55]. The results showed that $IR$ was high (mean of 0.543) and there was no significant difference between populations, suggesting inbreeding may be present in various populations. On the other hand, genetic structure is also a driver of the Wahlund effect. In contrast to inbreeding, individuals in various groups within a large population with genetic structure are more likely to mate with unrelated individuals. If genetic structure is a major factor of the Wahlund effect, then the high $Fis$ and $IR$ in our study cannot be explained by inbreeding. In fact, high $Fis$ values not relating to inbreeding have also been reported in other species [56,57]. Presumably, therefore, high $Fis$ and $IR$ values in our study are likely the reflection of individual mating in a natural state of the species in the study region. This speculation, of course, needs further verification using *P. forsythii* populations across its whole distribution.

Genetic differentiation index is an important indicator of the degree of genetic differentiation among populations [58]. The greater the degree of differentiation, the more pronounced the genetic structure of the population. $F_{ST}$ and $D_{EST}$ of the 15 populations indicated that there were evident genetic differences among the populations, and the differences were mainly from ALE. The pairing comparison between this population and other
populations was almost always significant, and the results showed consistency. The LT population was discovered as the least distinct from the other 14 populations, possibly due to its small sample size. Both the IBD test and the RDA analysis of the four distance-related variables and the genetic variation in the population indicated that there was a significant isolation effect of distance in the study area, and geographical distance was significantly and positively correlated with genetic distance within 100 km. Therefore, it is suggested that the population gradually accumulates genetic variation during the process of dispersal. In addition, the dispersal ability of *P. forsythii* cannot be neglected when considering the geographical isolation effect. Based on previous studies on the relationship between genetic structure and dispersal in vertebrates, it is generally assumed that geographical factors limit species dispersal and thus limit gene flow [59]. The dispersal capabilities of *P. forsythii* might be crucial to explain the observed genetic structure. At present, however, detailed knowledge on the dispersal distance and home range of the species is still lacking. Further research needs to confirm the threshold dispersal distance and whether the species are capable of fulfilling long-distance movement.

Although the results of the RDA permutation test were non-significant, the mean daily temperature difference (Bio2) was higher than the rest climate factors, suggesting that environmental temperature changes may have a relatively larger effect on the population genetic variation. Evidently, changes in ambient temperature have a regulatory impact on behavior of poikilothermy. Studies have shown that climate warming increases the number of times that poikilothermic animals shuttle between moving into sunlight and drilling into caves to regulate body temperature, which results in shorter predation times, reduced chances of mating, and increased probability of being preyed on by predators, leading to increasing maternal mortality and reducing reproductive output [60,61]. These behavioral regulations may result in decreased effective mating and possibly reduced gene flow. Qi et al. [62] also found a relatively low genetic diversity of *P. forsythii* in the same region and suggested the high temperature being the main reason. However, detailed studies are needed to confirm the impact of temperature on the genetic variation observed in our dataset.

5. Conclusions

In our study, *P. forsythii* have some degree of genetic differentiation, with low levels of gene flow between populations, and individual dispersal and migration are largely restricted to within populations, showing a clear genetic structure overall. Geographical distance and environmental temperature may also be important factors affecting the observed genetic structure. Considering its potential extinction risk, we call for strengthening conservations of *P. forsythii*. Construction of reserves in the Tarim Basin, for example, would be an effective strategy for its protection.

**Author Contributions:** Conceptualization, Y.L. and P.G.; methodology, Y.L., W.Z. (Wei Zhao) and Y.Q.; validation, Y.L.; formal analysis, J.D.; investigation, L.J., T.C. and W.Z. (Wen Zhong); resources, Y.L.; writing—original draft preparation, J.D.; writing—review and editing, J.D., J.N. and Y.L.; data curation, J.N.; supervision, Y.L.; funding acquisition, Y.L. and P.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China, grant No. 32060311; the Gansu Provincial Science and Technology Plan Project, grant No. 20JR10RA124 and 23CXNA0045; the Gansu Province Higher Education Youth Doctoral Support Project, grant No. 2023QB-002; Longyuan Youth Innovation and Entrepreneurship Talent Project of Gansu Province, grant No. 202117; and the Fundamental Research Funds for the Central Universities of Northwest Minzu University, grant No. 31920210143, 31920190083, and 31920230027.

**Institutional Review Board Statement:** The animal study protocol was approved by the Experimental Animal Ethics Committee of Northwest Minzu University (Approval No.: xbum-sm-2018050). It was approved on 12 November 2018 and the term of validity is from 1 September 2018 to 31 December 2024.
Data Availability Statement: The data presented in this study are available in a publicly accessible repository (Figshare, DOI: 10.6084/m9.figshare.23700189.v1).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Results for partial models of twelve climatic environmental factors.

<table>
<thead>
<tr>
<th>Code</th>
<th>Climatic Environmental Factor</th>
<th>Variance Explained</th>
<th>Variance/%</th>
<th>$R^2$ Corrected Variance/%</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIO2</td>
<td>Mean Diurnal Range (Mean of monthly (max temp – min temp))</td>
<td>7.4</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>BIO4</td>
<td>Temperature Seasonality (standard deviation $\times$ 100)</td>
<td>6.6</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>BIO5</td>
<td>Max Temperature of Warmest Month ($^\circ$C)</td>
<td>8.8</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>BIO6</td>
<td>Min Temperature of Coldest Month ($^\circ$C)</td>
<td>8.7</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>BIO8</td>
<td>Mean Temperature of Wettest Quarter ($^\circ$C)</td>
<td>8.3</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>BIO10</td>
<td>Mean Temperature of Warmest Quarter ($^\circ$C)</td>
<td>6.4</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>BIO11</td>
<td>Mean Temperature of Coldest Quarter ($^\circ$C)</td>
<td>6.5</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>BIO13</td>
<td>Precipitation of Wettest Month (mm)</td>
<td>8.2</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>BIO14</td>
<td>Precipitation of Driest Month (mm)</td>
<td>12.6</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>BIO15</td>
<td>Precipitation Seasonality (Coefficient of Variation)</td>
<td>9.8</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>BIO16</td>
<td>Precipitation of Wettest Quarter (mm)</td>
<td>6.4</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>BIO17</td>
<td>Precipitation of Driest Quarter (mm)</td>
<td>10.4</td>
<td>1.7</td>
<td></td>
</tr>
</tbody>
</table>

References


24. Qi, Y.; Zhao, W.; Li, Y.; Zhao, Y. Environmental and geological changes in the Tarim Basin promoted the phylogeographic formation of Phrynocephalus forsythii (Squamata: Agamidae). Gene 2021, 768, 145264. [CrossRef] [PubMed]
28. Qi, Y.; Zhao, W.; Li, Y.; Zhao, Y. Environmental and geological changes in the Tarim Basin promoted the phylogeographic formation of Phrynocephalus forsythii (Squamata: Agamidae). Gene 2021, 768, 145264. [CrossRef] [PubMed]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.