Genetic variation and population structure analysis of leymus Chinensis from eurasian steppes using SSR makers.

Naseer Ahmed (nrchachar786@gmail.com)  
Chinese Academy of Agricultural Sciences Grassland Research Institute

Xianyang Hou  
Shanxi Agricultural University  https://orcid.org/0000-0003-3906-8902

Research Article

Keywords: Leymus chinensis, SSR marker, Genetic diversity, population structures

Posted Date: February 1st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1285080/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License. 
Read Full License
Abstract

*Leymus chinensis* is an essential perennial grass species extensively dispersed in the typical grassland communities of the Eurasian steppe region. It is relished by livestock due to its high quality and being a nutritionally valuable forage crop. *Leymus chinensis* has recently gained extensive consideration on its genetic diversity. However, genetic diversity studies on *L. Chinensis* using SSR markers is currently limited. In the present study, we investigate the genetic variation and population structure analysis of *Leymus chinensis* from Eurasian steppes using SSR makers. For the genetic diversity, nineteen SSR markers were used and a total of 133 alleles were identified across the 166 *L. Chinensis* plants. Our findings illustrated that the polymorphic rate for all SSR markers was greater than 80%, with the exception of SSR12i and SSR6c, which had polymorphism rates of 50% and 75%, respectively. The gene diversity (H) ranged from 0.0545 for SSR12i to 0.4720 for SSR25v, and the average was 0.3136. Furthermore, genetic diversity analysis indicated that the 166 samples could be grouped into five main population clusters based on their maximum membership coefficients which were assigned as Pop1 to Pop5. Among the five populations, the largest values of allele (total number of detected alleles), Ne (effective number of alleles) and Na (observed mean number of alleles) was found to be higher in Pop1, with values of 61, 1.461, and 1.977, respectively. Additionally, AMOVA showed that 13% of the total genetic variation occurred among the population and 87% genetic variation within the populations of the species. Whereas, the pairwise Fst specified the moderate genetic variation ranging from 0.0336 to 0.0731. Finally, the principal coordinate analysis revealed that the x-axis and y-axis explained 5.72% and 4.86% of the variation in molecular data, respectively. Taken together, these SSR markers provide new insights for a more precise understanding of the genetic diversity of *L. Chinensis* germplasm and could potentially enhance the breeding program of *L. Chinensis*.

1. Introduction

The genus *Leymus* belongs to the grass family (*Triticeae; Poaceae*). In the Eurasian steppe zone, leymus is one among the most distinctive grassland communities and it is widely spread throughout the Eurasia's eastern steppe, including Russia's outer Baikal region, western North Korea, Mongolia, the Northeast Plain, the Northern Plain, and China's Inner Mongolian Plateau [1]. In Asia, *Leymus chinensis* grasslands occupy a total area of about 420,000 km², and 220,000 km² of area is found in China. These grasslands contribute to the ecological civilization of China through water and soil conservation and provision of resources for livestock farming, particularly in northern China [2]. Due to its long and thick underground rhizomes, as well as adventitious roots at each node, *Leymus chinensis* significantly protects the environment.

It is a self-incompatible species that often requires outcrossing and may have expanded genetic diversity geographic distribution [3]. *Leymus chinensis* is well adapted to its environment and could grow under different soil and climate conditions. It is characterized as drought and salt tolerant, thus it can grow well during dry seasons even when soil moisture becomes less than 6%, Na₂CO₃ at concentrations of 175 mmol/L and NaCl at concentration of 600 mmol/L [4,5]. This species is an important animal husbandry
resource because of its high protein content, higher vegetative production, and good palatability [6]. Many researchers have been paying attention on how *Leymus chinensis* responds to global changes such as extreme temperatures, CO₂ doubling and drought from a macro perspective because of its significant function in environmental protection [7-10]. However, research on the genetic basis of *Leymus chinensis*’ environmental adaptations are lacking in the literature, owing to the species' limited genomic resources. To date, only 1,815 ESTs and 51 sequences of *Leymus chinensis* proteins have been stored in public databases [11]. Additionally, gene discovery is still behind the times, with just a few genes having been cloned and functionally validated [12, 13]. *Leymus chinensis* offers high protein content, high vegetative productivity, and good palatability, making it an important foodstuff for animal husbandry development [6]. In addition, *Leymus chinensis* is also an elite gene pool.

Studies of genetic diversity in plants are fundamental aspects of plant breeding research. The exploitation of phenotypic and genetic diversity enables the breeder for developing new and improved varieties. Some researchers use morphological characterization approaches to study diversity; however, it might be time consuming. Biochemical markers are also used for study of diversity although they are affected by developmental stages of plants and are limited in number [14]. As a result, molecular markers have been frequently employed in numerous studies for assessment of genetic diversity due to their associated varying merits such as cost-effectiveness, abundance in nature and are not affected by different plant developmental stages [15].

Various reports have documented the genetic diversity or relationship of *Leymus chinensis* at the molecular level. However, studies on genetic variation measurements at this level within and among identically named inbred lines sustained by different programs are lacking. Molecular fingerprinting approach complements phenotypic measures in the quantification of genetic variations, because it shows changes in DNA which cannot be expressed phenotypically. It is essential to determine the level of genetic diversity both among and within inbreed lines of *Leymus chinensis* by means of microsatellite markers. Comprehending the genetic variation and population structure offers the opportunity to determine the usefulness of emerging germplasm resources for breeding program and fundamental research. SSR markers are specific and highly polymorphous [16], but to design specific primers, they require the knowledge of the genomic sequence and are thus restricted mainly to economically important species. Therefore, current study was carried out using SSR markers to systematically investigate the population structure and genetic diversity in a set of 166 *L. Chinensis* plants. The objectives of the current research were to: (a) evaluate the levels of diversity among accessions collected from different regions of China, (b) evaluate the population structure of these accessions.

2. Results

2.1 Polymorphism of SSR markers

In the present study, nineteen SSR markers were used for genetic diversity and a total of 133 alleles were detected across the 166 *L. chinensis* individuals. The polymorphic rate was more than 80% for all the
SSR markers, except SSR12i and SSR6c which showed 50% and 75% polymorphism respectively. Remarkably, SSR7d, SSR8e, SSR9f, SSR10g, SSR11h, SSR17n, SSR19r, SSR20s, SSR22t, SSR24u and SSR25v displayed the highest polymorphic rate of 100% (Table 1). Additionally, the number of bands ranged between 4 and 10, while polymorphic bands were between 2 to 10. The gene diversity (H) ranged from 0.0545 at SSR12i to 0.4720 at SSR25v, with an average of 0.3136. The effective number of alleles (Ne) ranged from 1.0600 at SSR12i to 1.8990 at SSR25v with an average of 1.5199. Moreover, the Shannon's information index (I) ranged from 0.1203 at SSR12i to 0.6646 at SSR25v with an average 0.4761. The results revealed that majority of the examined SSR markers had relatively high polymorphism and were sufficient to further explore the genetic diversity of *L. chinensis* germplasm.

**Table 1.** Diversity parameters associated with the 166 accessions of *L. Chinensis*
| LOCUS NAME | Total Bands | Polymorphic Bands | Polymorphism (%) | Ne   | H    | I    |
|------------|-------------|-------------------|------------------|------|------|------|
| SSR3a      | 6           | 5                 | 83.33%           | 1.6271 | 0.3564 | 0.5245 |
| SSR4b      | 9           | 8                 | 88.89%           | 1.5693 | 0.3455 | 0.5240 |
| SSR6c      | 4           | 3                 | 75.00%           | 1.1819 | 0.1489 | 0.2741 |
| SSR7d      | 10          | 10                | 100%             | 1.5292 | 0.3369 | 0.5165 |
| SSR8e      | 7           | 7                 | 100%             | 1.6212 | 0.3608 | 0.5371 |
| SSR9f      | 8           | 8                 | 100%             | 1.3737 | 0.2644 | 0.4291 |
| SSR10g     | 10          | 10                | 100%             | 1.5300 | 0.3301 | 0.5056 |
| SSR11h     | 9           | 9                 | 100%             | 1.4515 | 0.3000 | 0.4725 |
| SSR12i     | 4           | 2                 | 50.00%           | 1.0600 | 0.0545 | 0.1203 |
| SSR13j     | 6           | 5                 | 83.33%           | 1.4109 | 0.2690 | 0.4175 |
| SSR14k     | 7           | 6                 | 85.71%           | 1.4778 | 0.3008 | 0.4600 |
| SSR15m     | 6           | 5                 | 83.33%           | 1.6928 | 0.3791 | 0.5458 |
| SSR17n     | 6           | 6                 | 100%             | 1.3778 | 0.2406 | 0.3833 |
| SSR18p     | 5           | 4                 | 80.00%           | 1.4298 | 0.2626 | 0.4014 |
| SSR19r     | 6           | 6                 | 100%             | 1.6516 | 0.3823 | 0.5669 |
| SSR20s     | 7           | 7                 | 100%             | 1.6429 | 0.3812 | 0.5666 |
| SSR22t     | 7           | 7                 | 100%             | 1.5606 | 0.3363 | 0.5096 |
| SSR24u     | 9           | 9                 | 100%             | 1.7903 | 0.4363 | 0.6267 |
| SSR25v     | 7           | 7                 | 100%             | 1.8990 | 0.4720 | 0.6646 |
| Average    | 7.00        | 6.53              | 91.03            | 1.5199 | 0.3136 | 0.4761 |
| Total      | 133         | 124               |                  |      |      |      |

Ne: effective number of alleles, H: gene diversity, I: Shannon's information index.

### 2.2 Population Structure of *Leymus chinensis*

To understand the population structure among 166 *Leymus chinensis* individuals, they were divided into 5 optimum clusters using an admixture model-based method, which was created from the STRUCTURE HARVESTER website with the largest log probability of the data (LnP(D)) derived $\Delta K$ value as shown in (Fig. 1a-c). The number of populations (K) was measured from 1 to 10 with ten replicates. Among these most congruent cultivars arrangement was offered by K=1 with different populations. The genetic diversity analysis revealed that 166 samples could be classified in to five main population clusters on the
basis of maximum membership coefficients, which were labeled as Pop1 to Pop5 (Fig. 2). The findings revealed a related group component in various clusters. The Pop1 cluster consisted of 61 samples, accounting for 39.75% of all samples, followed by Pop2 which consisted of 45 samples, while the least number of 12 samples were retained by Pop5 (Table 2). Our results showed that, despite all the proportions, majority of samples had relatively close genetic background, with interspecific gene flow to some extent.

**Table 2. Summary of the genetic diversity statistics of L. Chinensis populations.**

| Population | Size | Na   | Ne   | I    | H   | Ht  |
|------------|------|------|------|------|-----|-----|
| POP1       | 61   | 1.977| 1.461| 0.443| 0.286| 0.291|
| POP2       | 46   | 1.939| 1.497| 0.463| 0.303| 0.310|
| POP3       | 25   | 1.788| 1.435| 0.407| 0.265| 0.276|
| POP4       | 22   | 1.735| 1.433| 0.394| 0.259| 0.271|
| POP5       | 12   | 1.803| 1.495| 0.444| 0.294| 0.321|
| Average    | 33.2 | 1.848| 1.460| 0.430| 0.281| 0.294|
| Total      | 166  | 1.985| 1.383| 0.409| 0.254| 0.255|

**Ht:** Overall gene diversity, **I:** Shannon's information index, **H:** Gene diversity, **Ne:** Effective number of alleles and **Na:** Observed number of alleles.

The phylogenetic analysis of neighbor-joining (NJ) and the principal component analysis (PCA) were applied on 166 individuals to uncover the genetic relationship among them the based on the Manhattan index dissimilarity matrix. Both PCA plot and NJ dendrogram tree had clearly distinguished five clusters (Fig. 3). In addition, the x-axis and y-axis in the PCA plot illustrated 5.72% and 4.86% of the molecular data variance respectively.

### 2.3 Genetic diversity of L. Chinensis

Among the five populations (Table 2, Fig. 2), the largest value of total number of detected alleles, the Na and Ne were found to be higher in Pop1, which were 61, 1.977, and 1.461, respectively. Additionally, the Pop2 showed the largest H and I value 0.303 and 0.463 respectively. while highest Ht value of 0.321 was observed in Pop5.

Both the analysis of molecular variance (AMOVA) and pairwise Fst were used to access the genetic differences among the populations. The AMOVA analysis revealed 87% of the genetic variation was observed within the population while, that the 13% of the total genetic variation was seen among population (Table. 3), which is due to five population groups among accessions. Present analysis demonstrated a significant difference between individuals and within population groups. Additionally,
pairwise Fst showed moderate genetic differentiation which ranged from 0.0336 to 0.0731 (Table 4). The highest level of genetic difference was noted between Pop2 and Pop4, while Pop1 and Pop3 showed lowest genetic differences. Additionally, the PCA was performed GenAIEx version 6.5, based on the matrix of pairwise Nei’s unbiased genetic distance. The results revealed that the x-axis illustrated 5.72% of the variance of molecular data, while the y-axis explained 4.86% (Fig. 3).

Table 3. Analysis of molecular variance (AMOVA) showing genetic diversity in L. chinensis germplasm.

| Source        | df  | SS      | MS   | Est. Var. | %   |
|---------------|-----|---------|------|-----------|-----|
| Among Pops    | 4   | 443.691 | 110.923 | 2.971     | 13% |
| Within Pops   | 161 | 3116.845 | 19.359 | 19.359    | 87% |
| Total         | 165 | 3560.536 | 22.330 |           | 100%|

df: degrees of freedom, SS: Sum of squares, MS: Mean square, EST. Var: Estimated variance.

Table 4. The matrix of pairwise genetic differentiation among populations (Fst).

| Population | Pop1  | Pop2  | Pop3  | Pop4  | Pop5  |
|------------|-------|-------|-------|-------|-------|
| Pop1       | -     |       |       |       |       |
| Pop2       | 0.0341| -     |       |       |       |
| Pop3       | 0.0336| 0.0467| -     |       |       |
| Pop4       | 0.0481| 0.0731| 0.0593| -     |       |
| Pop5       | 0.0374| 0.0637| 0.0658| 0.0694| -     |

3. Discussion

Analyzing the diversity of collections of germplasm of several crop species has indicated substantial variation for a variety of traits [17]. The different types of phenotypic characteristics have proved to be an essential tool in plant classification and the ensuing information is usually of great importance to plant breeders because it aids in the production of plant species with suitable nutritional and agronomic characteristics [18].

To classify potential parent plants, identifying the genetic diversity of the improved and L. chinensis genotypes is an essential component required to characterize and conserve the germplasms. SSR markers are among the most extensively used DNA markers and are used for a variety of purposes (e.g., genome mapping, diversity and recognition of varieties) [19]. Unlike the biochemical and morphological, both the environment and growth strategies do not influence molecular markers [20]. The survival,
characterization, and germplasm reproduction are largely dependent on genetic diversity assessment. In classical breeding, genetic diversity is influenced through the selection of outcomes from various allele frequencies, and this results in loss of variation and beneficial effects. [21]. However, little is understood about the genetic diversity of *L. Chinensis*. In present study, the genetic variation of 166 *L. Chinensis* genotypes was tested using 19 microsatellite markers. The findings revealed a large degree of genetic variation among the genotypes used.

In present study, the 19 SSR markers used detected different number of alleles per locus. This could be attributed to the remnant heterozygosity in certain varieties and to the expected varietal heterogeneity in which landrace varieties are composed of pure line mixtures which contribute to their wide-ranging adaptation in traditional farming systems. These outcomes corroborate the findings by Shah et al. [22] where a different group of rice genotypes were investigated. By comparison, the average number of alleles per locus achieved in our study is comparatively lower as compared to those recorded in previous reports. For instance, Kuroda et al. [23] using 7 SSR loci, recorded an average value of 9.28 alleles per locus, while Rahman et al. [24] reported an average of 6.33 alleles per locus in 34 varieties using a small set of 3 SSR markers. The observed variations in average number of alleles per locus could be attributed to the choice of SSR markers with multiple scorable alleles and the diverse nature of genotypes used in the former and present studies.

The genetic diversity is the indicator of heterozygosity. The mean genetic diversity value of 0.3136 obtained in this study depicts relative heterozygosity in relation to the 166 genotypes of *L. chinensis* studied. The average value of gene diversity recorded in the present research is lower than 0.71 as reported in previous studies [25]. The higher mean genetic diversity value recorded by Lapitan et al. [25] might result from high genetic materials exchange rate among the rice genotypes mostly used for studies during their genetic improvement research. The relative heterozygosity identified in this research could have resulted from the mixing or exchange of genetic materials in various parental lines particularly during the adopted strategies for improving *L. chinensis*. Notably, hybrid variety in conventional rice breeding is the result of genetic material exchange between two lines.

In present study, the maximum genetic distance value found between pops and within pops showed a high genetic dissimilarity between them and indicates a high level of divergence. The chromosomal mutation and diverse geographical context potentially contribute to the genetic dissimilarity observed between the genotypes of the *L. chinensis* studied. Analysis of molecular variance revealed percentage variance in *L. chinensis* between populations and among accession genotypes used in the present study. The AMOVA results showed that 13% of the overall genetic variation appeared among populations, which is attributed to having five population groups among the accessions, whereas 87% genetic variation occurred within the populations of both species. The present research revealed significant differences between individuals and within population groups. For a given SSR, high genetic variation in a sample of population might be due to an increase in gene flow or the mutation of a number of repeats of a given genotype. Additionally, this high genetic variation could be due to the natural selection process within the studied genotypes of *Leymus chinensis*. In contrast, the comparatively low genetic variation between the
genotypes of *Leymus chinensis* could be due to the sharing of the same SSR profiles among accessions. The lower genetic variation between these genotypes may be related to the possibility of sharing a common ancestry, despite growing in different regions. Cao et al. [26] reported a similar difference in the level of variation between and among the rice genotypes groups studied using microsatellite markers.

In addition, pairwise Fst indicated a slight genetic differentiation, varying from 0.0336 to 0.0731. The maximum level of genetic difference was noted between Pop2 and Pop4, while the lowest is between Pop1 and Pop3. In addition, the PCA was performed based on the matrix of pairwise Nei's unbiased genetic distance. The observed genetic dissimilarity indicates a common ancestral origin or a high degree of interbreeding that occurs in the exchange of identical alleles in the *L. chinensis* genomes. These findings are similar to those recorded by Sajib et al. [27], but somewhat greater than those documented by Shah et al. [22]. On the other hand, Ravi et al. [28] used SSR markers among 40 rice varieties and achieved an average genetic similarity of 0.79. This genetic variation may be attributed to the use of different rice genotypes in different groups. The authors speculated that the large degree of similarity discovered may be due to intra-specific differences in the cultivars used.

From the cluster analysis conducted, it was observed that the genetic diversity of the 166 *L. chinensis* samples could be grouped into five major population clusters on the basis of their maximum membership coefficients, that were labeled as Pop1 to Pop5. The K was measured from 1 to 10 with ten replicates, and the most congruently arrangement of cultivars were offered by K=1 value with different populations. The findings displayed a similar group component in distinct clusters. Cluster Pop1 consisted of 61 samples, accounting for 39.75% of all the samples, followed by Pop2 which comprised of 45 samples while Pop5 retained the least with only 12 samples. In view of all the proportions, the findings demonstrated that majority of the samples possessed comparatively similar genetic background, with interspecific gene flow to some extent.

Genetic diversity is a vital characteristic in breeding programs for genetic improvement [29]. However, for *Leymus chinensis*, the information on genetic diversity is very limited. Moreover, for evaluating the species’ potential for the development of new germplasm resources, its evaluation of genetic diversity and population structure is vital. While the previous information of genetic diversity and pairwise relatedness could offer valuable insights on effectively utilizing huge collections of genetic resources [30–32]. Hence, in present research, 166 individuals of *Leymus Chinensis* were used for the genetic diversity assessment, which were collected from their native habitats across the Eurasian steppe. A total of 19 SSR markers harbored 133 alleles throughout the entire database. Medium level of genetic diversity of *L. Chinensis* was implied by Shannon's information index, effective number of allele and gene diversity of 0.4761, 1.5199, and 0.3136, respectively. The Bayesian model-based structure analysis is commonly employed in plant species to infer concealed population structure [33]. In current study, the structure analysis revealed that the optimum number of cluster for 166 *L. Chinensis* individuals is five. Both the principal component analysis (PCA) and the neighbor-joining (NJ) phylogenetic analysis validated the structural pattern as distinctive, revealing five main clusters.
4. Materials And Methods

4.1. Plant material and study site

The current study was carried out at the Institute of Grassland Research (IGR, Hohhot, China), Chinese Academy of Agricultural Sciences (CAAS), Beijing China. The *Leymus chinensis* germplasm accessions were acquired from five lands, namely Inner magnolia, National forage gene bank, Jilin, Shanxi, and Republic of magnolia. The screening of this unique collection was conducted using SSRs markers among these accessions.

4.2 Genomic DNA Extraction

A total of 166 individuals, including 44 Republic of magnolia germplasm and 122 cultivated germplasm, were collected from the different introduced areas of *Leymus chinensis* across China. Among the cultivated individuals, 61 came from National forage gene bank, 44 from Inner magnolia province, 19 from Jilin province, and 18 from Shanxi province. A nursery of germplasm collected was developed at the experimental station of Agro-pastoral ecotone located shaerqin research station (N40°34, E111°56′) Hohhot China. These accessions were cultivated at the *Leymus chinensis* germplasm station of Agro-pastoral ecotone. For DNA extraction, ten fresh young leaves from each individual were collected randomly from 30 days old plants, samples were sealed in plastic bags, immediately kept in the liquid nitrogen (-196 °C) and stored in deep freezer at -80 °C for DNA extraction. The Hi-DNAsecure Plant Kit (Tiangen, Beijing, China) was used for total genomic DNA extraction from each sample following instructions from the manufacturer. Moreover, the quantity and quality of extracted DNA was checked by Nano Drop 2000 (Thermo Scientific, Washington, DE, USA).

4.3 Polymerase Chain Reaction

The PCR reaction mixture consists of total volume of 10µl consisting of DNA template 50 ng, forward and reverse primer 0.5µl each, 4µl PCR master mix, and 4µl ddH2O model. The PCR amplifications were performed using the DNA Thermal Cycler (Bio Rad Laboratories, Inc, China) under the following conditions: Thermal profiling was setup with initial denaturation temperature of 94°C for 30 seconds, followed by denaturation with 35 cycles (94°C for 30 seconds), annealing (52-55°C for 30s), extension (72°C for 30s) and the final extension (72°C for 5 mins). Later the quality of PCR product was checked by gel electrophoresis by running 5 µl of PCR product on 1% agarose gel.

4.4 Electrophoresis and fragment detection

The fragment detection was performed using 10 µl of PCR amplification final product mixing with 3 µl of loading buffer 30% sucrose. Three (3) µl of the sample was loaded into 30% acrylamide-bisacrylamide gel in 1x TBE and electrophoresed for 90-110 minutes at 150 volts. The fragments were identified through silver staining technique as follows: the peeled gel was fixed with 0.5% ethanol containing 10% acetic acid solution for 20 mins. The fixed gel was subjected for staining by placing it in 0.1% AgNO3 staining solution for 15 mins. Then, the stained gel was washed with ddH2O twice, and finally to obtain visible
fragments, following the washing gel was placed in 37% formaldehyde solution containing 8% NaOH for 10 mins. Fragments of each allele were scored in base pairs using the Quantity One software version 4.6.2 (Bio-Rad), the size of the PCR products for each SSR markers was recorded in an Excel spreadsheet.

Table 5

| CODE | Forward prime                | Reverse primer                | Origin       |
|------|------------------------------|------------------------------|--------------|
| 3    | TGTTTCCTTCTTCTTGATGCC        | CATGGGATTCAATGGGTATA         | Genome       |
| 4    | GGGGGTTATCTCTCTTTTTTG        | CTGCTCTCCATGCATGTGTT         | Genome       |
| 5    | AGGGGCCAAAGAGAGAGAAAG        | TCACACATTACCCACACCT          | Genome       |
| 6    | AGCTAGCCCTCCTACCAAGAGAG     | CACGCTTTGTGTCTGTGTG          | Genome       |
| 7    | AAATGAGGGCTGTGCAAAAAA        | AGCTTTCTTTTTCCGAGAGG         | Genome       |
| 8    | TGTGCAGGTGACTTTGGAAG         | GTTTTTACAGCTTGAAGTGA         | Genome       |
| 9    | GTGGTCACACATCTCCAGAGG        | ACATGCATACTACTCCGCG          | Genome       |
| 10   | ACACATTGGTCCGCCATAC          | AGTCCAGCTTTGGCTACCAT         | Genome       |
| 11   | GCCCTGTTTTTGGTCGAGAG         | AGGACATGAACTCCTGAGG          | Genome       |
| 12   | GACCATGACTGAGACCAAA          | TGCACATGACGCCCTGTACCT        | Genome       |
| 13   | GCCGTGTGCTTTCTTCTTCC         | GAGCTGCATCGCTATCATCA         | Genome       |
| 14   | CGACCAAGGTAATAGGGCA          | GCCACCAATAGCTTGAAGT          | Genome       |
| 15   | GACCCGTCTTTCTTTCTTCTTCT       | CAGAGGAACTGGAACGAAAGGCGGCTGTA | Liu et al 1996 |
| 16   | ACTCTCTCCTCCTGCTTCC          | CTTTCTTGCTACATTGCC          | Genome       |
| 17   | ACACCTTCCCAGGACAATCCATTG     | AGCACGCAAGAGCACCGAAGAGGCGGCT | Liu et al 1996 |
| 18   | GCAATGAACACTGAATCCCTGCAGG    | GTGCAAGGAGAGAGGCGGCTGTA      | Liu et al 1996 |
| 19   | CCGCCGATGACCTTCTCTC          | ACCCAACGACCTATGGGAC          | Liu et al 1996 |
| 20   | GLTGTGCGACTACCTCACTCC        | CAGAGGAATGGAACAGACACTG       | Wang et al 2016 |
| 21   | GCAATGAACCTGAATCCGAG          | CGTGTGAGACCTCAGGTGGTACC      | Wang et al 2016 |
| 22   | CATCGTCACGGATGAAGAAGTTG     | AATGGGCAGCAGTATGTCAGTGA      | Wang et al 2016 |

4.5. Statistical Analysis
**Genetic Diversity.** Microsoft Excel 2013 was employed for analyzing total number of bands (TNB), number of polymorphic bands (NPB) and percentage of polymorphic bands (PPB). While the effective number of alleles (Ne), observed number of alleles (Na), Shannon's information index (I), Nei's gene diversity (H), and the overall gene diversity (Ht) were calculated and the estimate of genetic diversity level with five populations was analyzed using POPGENE version 1.2 software [34]. Moreover, GenAIEx v 6.5 was employed to calculate genetic differences among populations (FST) and pairwise Fst. The hierarchical was carried out using the Analysis of Molecular Variance (AMOVA) [35].

**Population structures.** The population structure of the 166 Leymus chinensis was investigated using the program STRUCTURE version 2.3.4 [36] by using the admixture model, correlated allele frequencies and a burn-in period of 100,000 iterations, followed by 1,000,000 Markov Chain Monte Carlo (MCMC) repetitions [EVANNO G et al. 2005]. The K value ranged from 1 to 9 with 30 independent runs. The optimum number of populations was identified by maximum likelihood (LnP9K)) and (ΔK) followed by Evanno's method [37]. The structure Harvester version 0.6.94 was used to examine the structure results [38] cluster analysis. DArwin program v 6.0.9 was used to construct unweighted neighbor-joining phylogenetic trees and principal component analysis based on the dissimilarity matrix measured with the Manhattan index [39, 40]. Additionally, GenAlEx v 6.5 was employed to estimate Principal Coordinate Analysis (PCS) for summarizing the variation patterns in multi-locus dataset, on the basis of matrix of pairwise Nei's genetic distance [35]. Finally, the correlation between geographic and genetic distance matrices was measured using SPSS version 22.0 software.

5. **Conclusions**

The overall aim of this study was to characterize genetic diversity in *L. chinensis* germplasm using molecular markers. In numerous studies, the SSR markers have been applied for characterization of genetic diversity and population structure in various crops including grasses and related species. However, genetic diversity studies on *L. chinensis* using SSR markers have so far been less studied. In this study, several polymorphic SSR markers were characterized for *L. Chinensis*, the microsatellite markers displayed high DNA polymorphism levels in the *Leymus Chinensis*. Thus, the markers can be applied in *L. Chinensis* genotypes for evaluating genetic diversity and relationship. The present investigation offers an overall population structure and genetic diversity assessment in 166 *L. Chinensis* genotypes. The genetic diversity and population structure analysis of *L. Chinensis* provides a greater understanding of its domestication and gene exchange in china. Additionally, these markers are often known as perfect methods for distinguishing such polymorphic cultivars. Taken together, the microsatellite markers offer new insights for comparatively accurate understanding of genetic diversity and relationship of *L. chinensis*, which is crucial for molecular marker-assisted selection for breeding programs as well as for identification, utilization and conservation of germplasm. The molecular markers employed in the present study should be used for constructing a genome database that will be useful in *L. chinensis* breeding programs and characterization of other *L. chinensis* germplasm. The SSR markers have proved to be highly effective for characterization of variation in our *L. chinensis* germplasm. Suggesting that increasing the number of SSR markers would enhance the likelihood of identifying
individuals within population groups. Our outcomes define broad scale patterns of population genetic structure in *L. chinensis* germplasm. These results will be useful for breeders wishing to exploit genetic diversity.

**Declarations**

**Conflicts of Interest:**

The authors declare no conflicts of interest.

**Funding:**

The present study was funded by

**Author Contribution:**

NA conducted research and wrote manuscript

**References**

1. Bai Y, Han X (2004) Wuj,Li. Ecosystem stability and compensatory effects in the Inner Mongolia grassland. Nature 431:181–184
2. Bai w, Xun F, Li Y (2010) Rhizome severing increases root lifespan of Leymus chinensis in a typical steppe of Inner Mongolia. PLoS One 5:8
3. Zhang WD, Chen SY, Liu GS, Jan (2004) CC. Seed-set and pollen–stigma compatibility in Leymus chinensis. Grass Forage Sci 59:180–185
4. MA H, Liang Z (2007) Effects of different soil pH and soil extracts on the germination and seedling growth of Leymus chinensis.181–188
5. Wang R, Chen L, Bai Y, Xiao C (2008) Seasonal dynamics in resource partitioning to growth and storage in response to drought in a perennial rhizomatous grass, Leymus chinensis. J Plant Growth Regul 27:39–48
6. Wang LJ, Li XF, Liu GS (2009) Improvement of Leymus chinensis for drought tolerance by expression of a LEA gene from wheat. Biotechnol Lett 31:313–319
7. Niu S, Liu W, Wan S (2008) Different growth responses of C3 and C4 grasses to seasonal water and nitrogen regimes and competition in a pot experiment. J Exp Bot 59:1431–1439
8. Wang J, Bao Y (2007) Evapotranspiration of Leymus chinensis steppe community at its different deteriorated phases. Chin J Ecol 26:1148–1152
9. Xu Z, Zhou G, Shimizu H (2009) Are plant growth and photosynthesis limited by pre-drought following rewatering in grass? J Exp Bot 60:3737–3749
10. Xu ZZ, Zhou GS (2006) Combined effects of water stress and high temperature on photosynthesis, nitrogen metabolism and lipid peroxidation of a perennial grass Leymus chinensis. Planta 224:1080–1090

11. Jin H, Plaha P, Park JY, Hong CP, Lee IS, Yang ZH, Kim YA (2006) Comparative EST profiles of leaf and root of Leymus chinensis, a xerophilous grass adapted to high pH sodic soil. Plant Sci 170:1081–1086

12. Ma P, Liu J Isolation and characterization of a novel plasma membrane intrinsic protein gene, LcPIP1, in Leymus chinensis that enhances salt stress tolerance in Saccharomyces cerevisiae. Applied biochemistry and biotechnology. 2012, 166:479–485

13. Xianjun P, Xiangyong M, Weihong F (2011) Improved drought and salt tolerance of Arabidopsis thaliana by transgenic expression of a novel DREB gene from Leymuschinensis. Plant Cell Rep 30:1493–1502

14. Winter P, Kahl G Molecular marker technologies for plant improvement. World Journal of Microbiology and Biotechnology. 1995, 11:438–448

15. McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Blair M (1997) Microsatellite marker development, mapping and applications in rice genetics and breeding, Oryza: From Molecule to Plant. Springer, pp 89–99

16. Jun TH, Van K, Kim MY, Lee SH (2008) WALKER D R. Association analysis using SSR markers to find QTL for seed protein content in soybean. Euphytica 162:179–191

17. Yang RC, Jana S, Clarke J (1991) Phenotypic diversity and associations of some potentially drought-responsive characters in durum wheat. Crop Sci 31:1484–1491

18. Maduakor H, Lal R Root system topgrowth and yield of cassava (Manihot esculenta Crantz) as affected by plant population in an arid soil. J. Root Crops. 1989, 15:115–122

19. DA SJ (2005) Molecular markers for phylogeny, breeding and ecology in agriculture. Genetic Resources and Biotechnology 3:221–256

20. Ovesana J, Polakova K, Leisova L DNA analyses and their applications in plant breeding. Czech Journal of Genetics and Plant Breeding. 2002, 38:29–40

21. Singh RK, Sharma RK, Singh AK, Singh VP, Singh NK, Tiwari SP, Mohapatra T (2004) Suitability of mapped sequence tagged microsatellite site markers for establishing distinctness, uniformity and stability in aromatic rice. Euphytica 135:135–143

22. Shah SM, Naveed SA, Arif M (2013) Genetic diversity in basmati and non-basmati rice varieties based on microsatellite markers. Pak J Bot 45:423–431

23. Kuroda Y, Sato YI, Bounphanousay C, Kono Y, Tanaka K (2007) Genetic structure of three Oryza AA genome species (O. rufipogon, O. nivara and O. sativa) as assessed by SSR analysis on the Vientiane Plain of Laos. Conserv Genet 8:149–158

24. Rahman MS, Molla MR, Alam MS, Lutfur R (2009) DNA fingerprinting of rice (Oryza sativa L.) cultivars using microsatellite markers. Aust J Crop Sci 3:122–128
25. Lapitan VC, Brar DS, Abe T, Redona E (2007) Assessment of genetic diversity of Philippine rice cultivars carrying good quality traits using SSR markers. Breeding Science 57:263–270

26. Cao Q, Lu BR, Xia H, Rong J, Sala F, Spada A, Grassi F (2006) Genetic diversity and origin of weedy rice (Oryza sativa f. spontanea) populations found in north-eastern China revealed by simple sequence repeat (SSR) markers. Ann Botany 98:1241–1252

27. Sajib AM, Hossain M, Mosnaz MJ, Hossain H, Islam M, Ali M, Prodhan SH (2012) SSR marker-based molecular characterization and genetic diversity analysis of aromatic landraces of rice (Oryza sativa L.). Journal of BioScience & Biotechnology, 1, 2.

28. Ravi MS, Geethanjali F, Sameeya F, Maheswaran M (2003) Molecular marker based genetic diversity analysis in rice (Oryza sativa L.) using RAPD and SSR markers. Euphytica 133:243–252

29. Liu F, Hong Z, Xu D, Jia H, Zhang N, Liu X, Lu M (2019) Genetic diversity of the endangered Dalbergia odorifera revealed by SSR markers. Forests 10:225

30. Ambreen H, Kumar S, Kumar A, Agarwal M, Jagannath A, Goel S (2018) Association mapping for important agronomic traits in safflower (Carthamus tinctorius L.) core collection using microsatellite markers. Front Plant Sci 9:402

31. Xu Q, Zeng X, Lin B (2017) A microsatellite diversity analysis and the development of core-set germplasm in a large hulless barley (Hordeum vulgare L.) collection. BMC Genet 18:102

32. Yang X, Song J, Todd J, Peng Z, Paudel D, Luo Z, Zhao Y (2019) Target enrichment sequencing of 307 germplasm accessions identified ancestry of ancient and modern hybrids and signatures of adaptation and selection in sugarcane (Saccharum spp.), a ‘sweet’crop with ‘bitter’genomes. Plant Biotechnol J 17:488–498

33. Porras-hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo A, Lareu M (2013) An overview of STRUCTURE: applications, parameter settings, and supporting software. Front Genet 4:98

34. Popgene (2017) Version 1.32; the User-Friendly Shareware for Population Genetic Analysis. Molecular Biology and Biotechnology Center, University of Alberta, Edmonton. Available online: http://www.ualberta.ca/~fyeh (accessed on 23 January 2019)

35. Peakall R, Smouse PE (2012) GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research— an update. Bioinformatics 28:2537–2539

36. Pritchard JK, Stephens M (2000) Inference of population structure using multilocus genotype data. Genetic 155:945–959

37. 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

38. Earl D, Vonholdt B (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4(2):359–361

39. Perrier X, Jacquemoud-Collet JP, DARwin (2006) Software. Available online: http://darwin.cirad.fr/darwin (Accessed on 30 January 2019)
40. Perrier X, Flori A, Bonnot F (2003) Data analysis methods. In: Hamon P, Seguin M, Perrier X, Glaszmann JC (eds) Genetic Diversity of Cultivated Tropical Plants. Enfield, Science Publishers: Montpellier, France., pp 43–76

Figures

Figure 1

Population structure analysis of 166 Leymus chinensis individuals based on 19 SSRs. (a) Population estimation using mean of estimated LnP(D) with K ranging from 1 to 10 (b) Population estimation using LnP(D) derived ΔK with K ranged from 1 to 10. (c) Clustering of 166 L. Chinensis individuals presented by different colors i.e. 1. Red, 2. Green, 3. Blue, 4. Yellow and 5. Pink inferred by Structure analysis. Each bar representing an individual, where different colors representing the estimated membership coefficient using the Q statistic.

Figure 2

Details of hierarchal population structure analysis of 166 L. Chinensis individuals Pop1, Pop2 Pop3, Pop4 and Pop5 population code.
Figure 3

The relationships among the 166 individuals of Leymus chinensis. (a) An unweighted neighbor joining (NJ) tree among 166 L. Chinensis individuals calculated using Manhattan index based on dissimilarity matrix. (b) The principal coordinate analysis (PCA) on the basis of dissimilarity matrix, first principal coordinate, the x-axis displayed 5.72% variation while the second principal coordinate the y-axis showed 4.86% variation.

Figure 4

Matrix of pairwise genetic differentiation between populations (Fst)