

Article

Not peer-reviewed version

Genetic Diversity and Population Structure of the Peruvian Andean Legume, tarwi (*Lupinus mutabilis*), With High Quality SNPs

[Amelia Huaranga-Joaquin](#) , [Carla L. Saldaña](#) , [David Saravia](#) , Sady García-Bendezú , [Pedro Rodríguez-Grados](#) , Wilian Salazar , Felix Camarena , Pedro Injante , [Carlos I. Arbizu](#) *

Posted Date: 19 January 2023

doi: 10.20944/preprints202301.0345.v1

Keywords: Fabaceae; bioinformatics; molecular markers; neglected crop; genomics



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Genetic Diversity and Population Structure of the Peruvian Andean Legume, Tarwi (*Lupinus mutabilis*), With High Quality SNPs

Amelia Huaranga-Joaquin ¹, Carla L. Saldaña ², David Saravia ^{1,2}, Sady García-Bendezú ¹, Pedro Rodríguez-Grados ³, Wilian Salazar ², Felix Camarena ¹, Pedro Injante ² and Carlos I. Arbizu ^{1,*}

¹ Facultad de Agronomía, Universidad Nacional Agraria La Molina, Av. La Molina s/n, Lima 15024, Peru; ahuaringa@lamolina.edu.pe (A.H.-J.); agricultura_precision@inia.gob.pe (D.S.); sjgarcia@lamolina.edu.pe (S.G.-B.); camafe@lamolina.edu.pe (F.C.)

² Dirección de Desarrollo Tecnológico Agrario, Instituto Nacional de Innovación Agraria (INIA), Av. La Molina 1981, Lima 15024, Peru; carla18317@gmail.com (C.L.S.); r_cambioclimatico@inia.gob.pe (W.S.); pinjante@inia.gob.pe (P.I.)

³ Facultad de Ciencias, Universidad Nacional José Faustino Sánchez Carrión, Av. Mercedes Indacochea Nro. 609, Huacho 15136, Peru; pmrg1711@gmail.com

* Correspondence: carbizu@inia.gob.pe

Abstract: *Lupinus mutabilis* Sweet (Fabaceae), “tarwi” or “chocho”, is an important grain legume in the Andean region. In Peru, studies on tarwi have been mainly focused on morphological features, however, they have not been molecularly characterized. Currently, it is possible to explore genetic parameters of plants with reliable and modern methods like genotyping-by-sequencing (GBS). We here for the first time used single nucleotide polymorphisms (SNPs) markers to infer the genetic diversity and population structure of 89 accessions of tarwi from nine Andean regions of Peru. A total of 5922 SNPs distributed along all chromosomes of tarwi were identified. STRUCTURE analysis revealed that this crop is grouped into two clusters. A dendrogram was generated using the UPGMA clustering algorithm and, similar to the principal coordinate analysis (PCoA), it showed two groups that correspond to the geographic origin of the tarwi samples. AMOVA showed a reduced variation between clusters (7.59%) and indicated that variability within populations is 92.41%. Population divergence (F_{st}) between clusters 1 and 2 revealed low genetic difference (0.019). We also detected a negative F_{is} for both populations, demonstrating that, similar to other *Lupinus* species, tarwi also depends on cross-pollination. SNPs markers were powerful and effective for the genotyping process in this germplasm. We hope that this information is the beginning of the path towards a modern genetic improvement and conservation strategies of this important Andean legume.

Keywords: Fabaceae; bioinformatics; molecular markers; neglected crop; genomics

1. Introduction

Lupinus mutabilis Sweet (Fabaceae), also known as “tarwi”, “tarhui” or “chocho” is an Andean legume cultivated in the Andean region in South America [1]. Tarwi has had importance in diet since pre-hispanic time [2]. It has important nutritional compounds, mainly due to its protein values that varies from 32% to 51.6% (rich in globulins, 43–45% and albumins, 8–9%) [3], high oil content (13–24%), crude fiber (6.2–11%), minerals such as iron, magnesium and phosphorus, and bioactive compounds such as isoflavones and phenols, with proven antioxidant capacity [4]. These nutritional levels present in tarwi are even better than soybeans [5]. Due to these qualities, it constitutes an alternative crop to reduce malnutrition, and is considered as the emerging protein crop for Europe and temperate climate zones [6]. Likewise, its adaptation to altitudes of 3100 to 3,850 m.a.s.l., temperate climate, and influence of the length of the day, makes it susceptible to low temperatures (–2 °C) in the initial stages, it requires around 350 to 800 mm of rain, and can grow for 240 to 300 days [7]. Also, for its cultivation on marginal lands, under drought stress, mostly without tilling the land

and implementing agronomic practices, it is considered a resilient species against the impact of climate change. Likewise, it possesses the ability to fix atmospheric nitrogen and mobilization of phosphorus that promotes agroecological production [8].

On the other hand, tarwi presents key domestication characteristics, including indehiscent pods and permeable seeds with tegument, which represents a locally important crop in several Andean areas [9]. *L. piurensis* is indicated as the probable wild progenitor of *L. mutabilis*. Atchison et al. [2] generated nextRAD sequence data for 212 accessions of Andean *Lupinus*, representing 63 species and resolved relationships between species that diverged over time and sheds light on the origins of domestication. It has been reported that the tarwi ecotypes from northern Peru are bushy and decumbent in growth and generally have a prominent stem; while the ecotypes of the central and southern zone have herbaceous and bushy growth with semi-erect growth and non-prominent stem. Also the vegetative period varies from 180 to 270 days [8]. Camillo et al. [10] evaluated accessions from 22 populations of 16 *Lupinus* species, and showed that *L. mutabilis* and *L. semperflorens*, among other 13 species, presented $2n = 48$ chromosomes; while in *L. bandelierae* presented $2n = 36$ chromosomes. They also suggested that cytologically Andean lupines are more closely related to North American species than those of South America [10]. The introduction of *L. mutabilis* in the Mediterranean area shows a wide intraspecific genetic variability in collections, which allows the establishment of conservation and improvement programs [6]. Guilengue et al. [6] also evaluated the associations between genome size and morphological characters using Spearman's correlation analysis for 23 accessions, finding that no individual morphological trait presented strong correlations with genome size [6].

Currently, *L. mutabilis* remains a poorly studied crop in the field of genetics. Chirinos-Arias et al. [11], indicated that the inter-accession genetic modification in *L. mutabilis*, according to the accessions and ISSR markers evaluated, is considerable. They reported that *L. mutabilis* is an autogamous plant with a considerable degree of allogamy. Ruiz-Gil et al. [12] carried out a morphometric analysis using the flower standards of *L. mutabilis*, *L. piurensis* and the population that presents intermediate characteristics. The analysis of canonical discriminants with data from a morphometric analysis showed the existence of three different groups: (1) *L. mutabilis*, (2) *L. piurensis* and (3) the population with intermediate characteristics. Allo and autopolyploidization events, along with other chromosomal rearrangements, during the evolution of this species could have led to duplication/or triplication of genome regions, as reported in the Old World species *L. angustifolius* [13].

Molecular data increased the understanding of plant systematics at various taxonomic levels [14]. The genetic similarity between genotypes can be assessed with DNA markers [15], also it can help to select accessions for establishing a core collection. Genotyping by sequencing (GBS) is one such sequence identification method variants using next generation sequencing technology, producing powerful and cost-effective genotyping process [16]. In addition, its genotyping can be easily replicated [17]. Its application has been reported in different crops such as dry bean [18], potato [19], reed canarygrass [20], lentils [21], maize [22], barley [23], rice [24], soybean [25], switchgrass [26], and wheat [27] Thus, the objective of this study were to characterize a collection of tarwi germplasm currently maintained by the Grain Legumes and Oilseeds Research Program of the National Agrarian University - La Molina (UNALM for its acronym in Spanish) to gain a better understanding of the genetic diversity and population structure of this legume by employing a NGS technique.

2. Materials and Methods

2.1. Plant Material and DNA Extraction

The experimental material of tarwi was obtained from the Germplasm Bank of the Grain Legumes and Oilseeds Research Program of UNALM. A total of 89 young leaves were collected in labeled paper envelopes and stored in plastic containers with silicone gel, for the preservation of samples during transport to the National Institute of Agricultural Innovation (INIA) for genomic DNA extraction. Further details of the samples examined in this work are in Table S1. Genomic DNA

was extracted using the CTAB method [28] adapted for this species. Previously, leaves were grounded with liquid nitrogen. Finally, 100 mg of ground tissue was used. DNA quantity and quality were evaluated by using the Qubit™4 Fluorometer (Invitrogen, Waltham, MA, USA) and agarose gel (1%), respectively.

2.2. Genotyping by Sequencing Data

Samples were sent to Biotechnology Center of DNA sequencing of University of Wisconsin-Madison. The GBS library will be subjected to a single run using the Illumina HiSeq 2500 platform, constructing 2×150 bp 100 bp pairwise sequenced read libraries. Quality control of the raw data was performed using FastQC software (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Reads were then trimmed by removing adapter and padding sequences using the `gbstrim.pl` script (<https://bitbucket.org/jgarbe/gbstrim/src/master/>). Then, the data was analyzed using the bioinformatics TASSEL-5 GBS v2 pipeline [29,30]. We used the *L. angustifolius* genome (NC_032009) [31] as reference. GBS sequence tags were identified in the FASTQ files, indicating that the restriction enzyme used was ApeKI and the barcode used for each sample for the preparation of libraries. Then, we used the `GBSSeqToTagDBPlugin` to convert GBS to a unique tag database with default parameter, minimum base quality score of 20. Only reads that had a complete barcode sequence were considered. Then, the `TagExportToFastqPlugin` was used and unique tags of each sequence were indexed to produce a tag count file for each sample in FASTQ text format (the `-t` argument was used), which was used as input to the Burrows-Wheeler Aligner (BWA) v.0.7.17 program [32] to align all the tags with the reference genome (NC_032009). The output file in SAM format was transformed into a binary file using the `SAMToGBSdbPlugin` pipeline. `DiscoverySNP CallerPluginV2` was used to identify SNPs from the aligned tags, position and allele data. Finally, quality score for SNP positions was identified. Data curation was conducted using `VCFTools` v.0.1.15 [33] with the following parameters: (i) minimum minor allele frequency of 0.1, (ii) maximum minor allele frequency of 1, (iii) number of alleles two, and (iv) maximum missing data of 0.9. The VCF file was used as input to the package `SNPRelate` v.1.20.1 [34] to another filtering process that included: (i) removal of multi-allelic, monomorphic SNP positions, and (ii) removal of SNPs with a linkage disequilibrium (R^2 value) more than 0.2.

2.3. Phylogenetic Analysis

We constructed a phylogenetic tree from a distance matrix using R software v.4.2.1 [35]. To calculate genetic distances, Provesti's coefficient [36] was considered, then a dendrogram was obtained with the unweighted pair group method with arithmetic mean (UPGMA) clustering algorithm with 1000 bootstrap replicates from `poppr` package v.2.9.2 [37]. Argument `dudi.pco` of `ade4` v.1.7-16 package [38] was employed in R to conduct a principal coordinate analysis (PCoA). Tree was viewed in `FigTree` v.1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

2.4. Population Structure and Genetic Diversity

We determined the population structure in 89 accessions of tarwi using all original filtered SNPs data sets comprising 10 % missing imputed data. To convert VCF data sets to STRUCTURE format, we used PLINK PED [39] with `-plink` option in `VCFTools` software and `PGSpider` [40]. Finally, we employed STRUCTURE software v.2.3.4 [41] with populations (K value) ranging from 1 to 15, replicated 10 times, with a burn-in length of 20,000 and 50,000 Monte Carlo iterations. The results produced by STRUCTURE were processed in STRUCTURE HARVESTER software [42], and we detecting the most probable number of clusters of our data using Delta K value [43]. Finally, to discriminate populations across all the ten iterations of the selected K values, we used the package `Pophelper` v.2.3.1 [44] in R software.

3. Results

3.1. Sequencing and Distribution of SNPs

After filtering out the raw reads, the total demultiplexed reads for all 89 genotypes were 579.02 M with the average reads per accession being 6.51 M. We followed Arbizu et al. (2016) [45] and Martínez-Flores et al. (2020) [46] approach and obtained 338,638 read tags; 36.06% of them uniquely aligned to the *L. angustifolius* reference genome [31]. A total of 35,760 raw SNPs were obtained, and filtering approach yielded 5922 high-quality SNPs across the 20 chromosomes of *Lupinus* with an average marker density of 79,52 kb (Table 1, Figure 1). The highest number of SNPs were physically mapped to chromosome four (7.26%, 430 SNPs). Chromosomes 11 (112.74 kb) and 14 (65.01 kb) presented the highest and lower marker densities, respectively (Table 1).



Figure 1. Genome-wide density plot of 5922 SNPs in the tarwi genome.

Table 1. Genomic distribution of 5,922 SNPs across 20 chromosomes of *L. Mutabilis*.

Chromosomes	No. of SNPs	%SNPs	Total length (Mb)	Density (Kb)
1	363	6.13	36.46	100.43
2	322	5.44	24.70	76.70
3	410	6.92	30.15	73.54
4	430	7.26	23.77	63.57
5	303	5.12	26.38	87.05

6	344	5.81	33.11	96.25
7	294	4.96	19.78	67.29
8	330	5.57	25.52	77.34
9	331	5.59	21.75	65.72
10	273	4.61	16.34	59.86
11	319	5.39	35.96	112.74
12	237	4.00	19.07	80.45
13	230	3.88	17.82	77.48
14	250	4.22	16.25	65.01
15	280	4.73	20.96	74.87
16	244	4.12	20.79	85.19
17	252	4.26	21.30	84.52
18	184	3.11	16.59	90.15
19	248	4.19	18.16	73.23
20	278	4.69	21.99	79.10

3.2. Population Structure and Genetic Relationships

We performed a population structure analysis using 5922 high-quality SNPs among the 89 accessions of tarwi. The Evanno method [43] indicated the best K value (number of populations) is two for our data set (Figure S1). STRUCTURE analysis showed admixture, except for some accessions (Figure 2). Tarwi accessions did not cluster according to the regions they belong (Figure S2) but were separated into the following clusters: (i) cluster 1 included 59 accessions, and (ii) cluster 2 included 30 accessions (Figure 2, Table 2). There is some degree of grouping when tarwi accessions are labelled according to the region they belong as two clusters are formed (cluster 1: center + south and cluster 2: north). However, few accessions are intermixed (Figure S3).

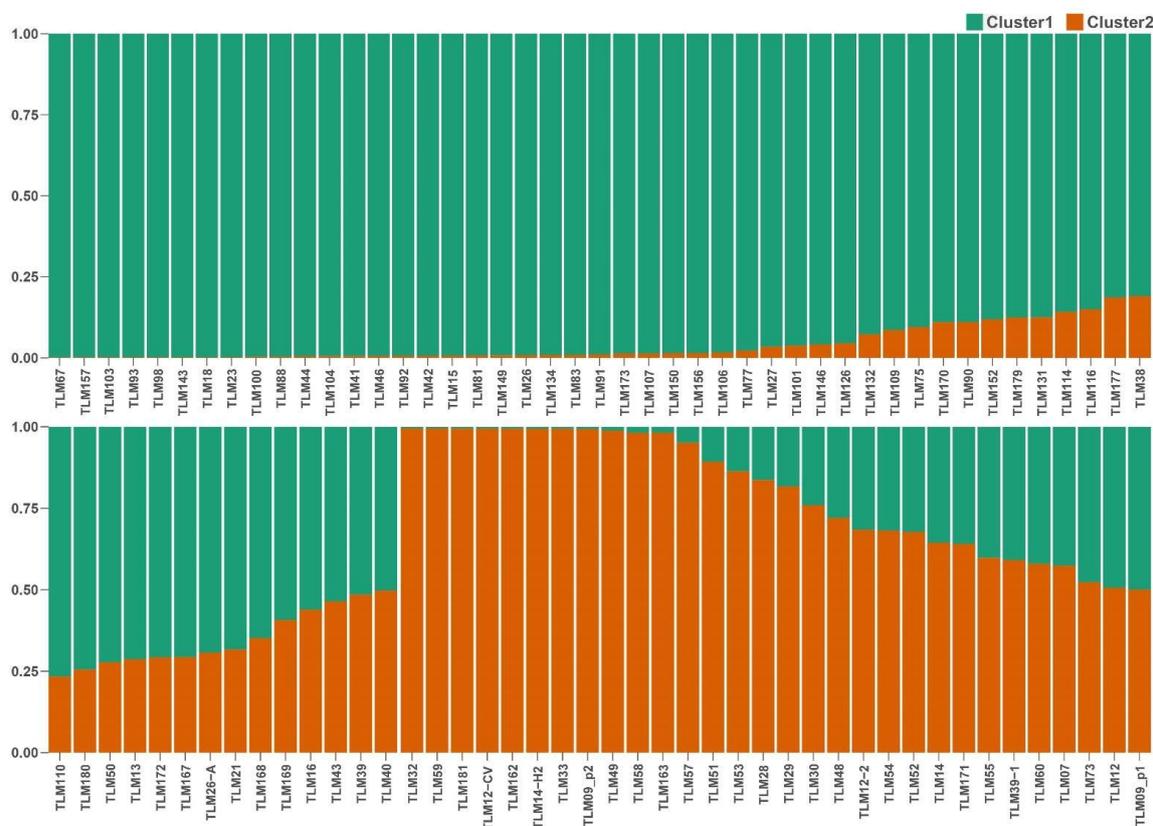


Figure 2. Population structure of 89 samples of tarwi inferred by the STRUCTURE analysis using 5922 SNPs markers.

Table 2. Origin of the tarwi germplasm collection among the two clusters inferred by the STRUCTURE analysis.

Department (Region)	Cluster 1	Cluster 2
Ancash (north)	1	13
Apurimac (south)	3	1
Cajamarca (north)	1	2
Cusco (south)	32	1
Huancavelica (center)	3	1
Huanuco (center)	1	2
Junin (center)	5	2
La Libertad (north)	0	5
Puno (south)	13	3
Total	59	30

Similar to the STRUCTURE analysis, the principal coordinate analysis (PCoA) based on the pairwise genetic distance matrix among all the 89 tarwi accessions also depicted two clusters (Figure 3). The first and second axis explained 5.82% and 4.18% of the variance, respectively. A UPGMA phylogenetic tree was constructed and two major clades were identified. Clade (cluster) 1 mainly contains accessions from the north and center regions of Peru while accessions from the north are mostly within clade (cluster) 2. The UPGMA was manually edited to show STRUCTURE grouping. Overall, there was a good agreement between these analyses (Figure 4).

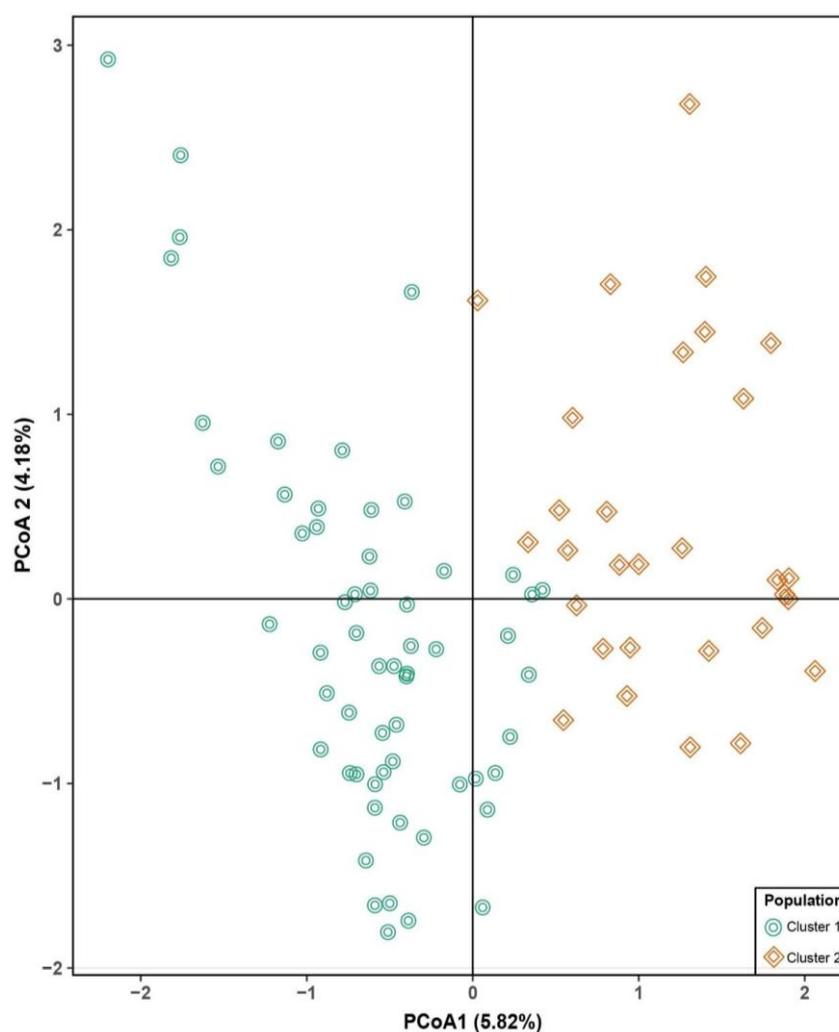


Figure 3. Principal coordinate analysis (PCoA) of 89 samples of a Peruvian germplasm of tarwi. Percentages of variance explained by each coordinate are noted in parentheses.

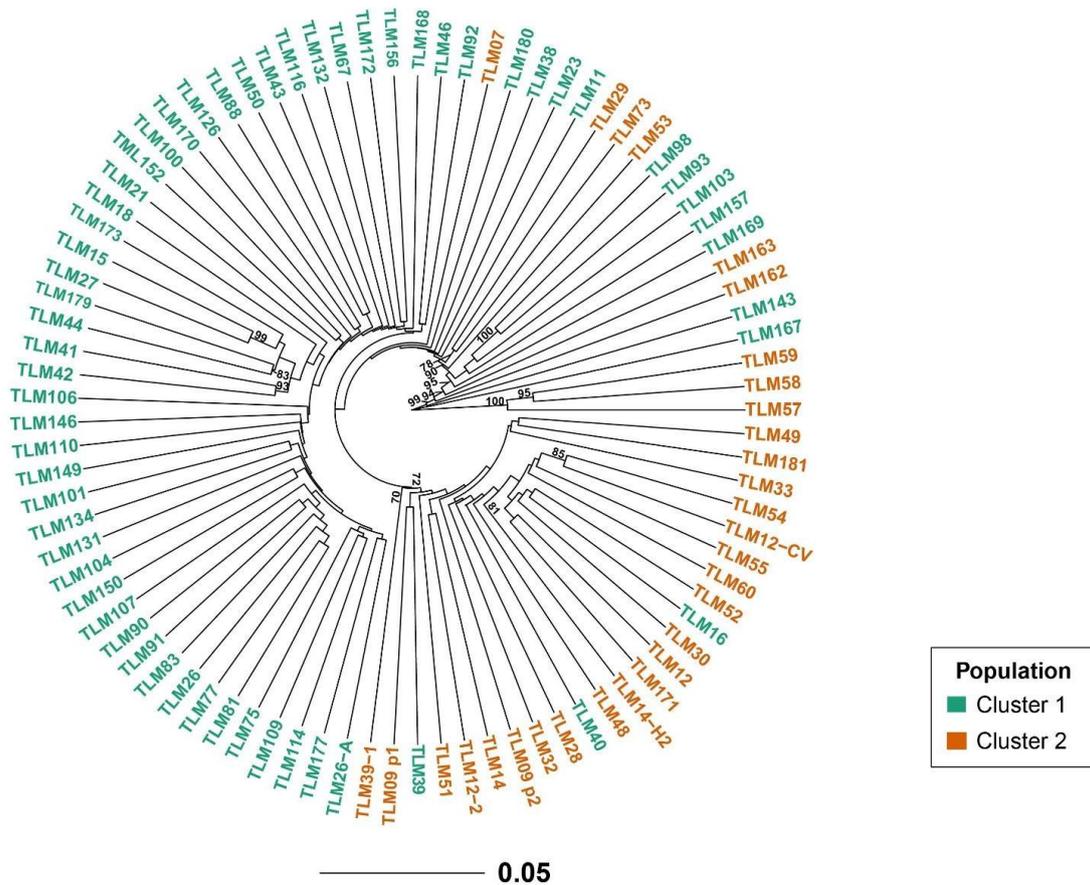


Figure 4. Dendrogram based on Provesti's genetic distance and the UPGMA clustering method of 89 accessions of tarwi using 5922 SNPs markers. Bootstrap values greater than 70% are shown.

3.3. Genetic Diversity of the Tarwi Collection

Diversity indices did not show significant differences among the two clusters identified by STRUCTURE (Table 2). The analysis of allelic patterns across all clusters revealed that the number of different alleles were similar. Allelic richness and observed heterozygosity among the two clusters did not varied greatly either. Genetic diversity (i.e. expected heterozygosity) was 0.421 and 0.433 for cluster 1 and 2, respectively. Cluster 1 presented a Shannon-Wiener's index of 4.08 while cluster 2 a value of 3.4, showing high diversity. Inbreeding coefficients for both clusters are negative, demonstrating that there is an excess of observed heterozygotes. In addition, the average coefficient of genetic differentiation among the two clusters was 0.019 (Table 2).

Table 2. Genetic diversity indices based on 5922 SNPs among two clusters of tarwi.

Cluster	Number of Accessions	N _a	A _R	H _o	H _e	H	F _{is}	F _{st}
1	59	2	1.997	0.641	0.421	4.08	-0.524	
2	30	1.999	1.998	0.657	0.433	3.4	-0.518	
	Mean	1.999	1.999	0.649	0.427	3.74	-0.521	0.019

N_a: number of different alleles, A_R: allelic richness, H_o: observed heterozygosity, H_e: expected heterozygosity, H: Shannon-Wiener index, F_{is}: inbreeding coefficient, F_{st}: gene differentiation coefficient.

We conducted an Analysis of Molecular Variance (AMOVA) in order to define the patterns of genetic variation, considering the two clusters identified by STRUCTURE. AMOVA revealed that the genetic variability among clusters was 7.59% while the rest (91.41%) was within clusters (Table 3). This confirms that there is great variation among the tarwi accessions.

Table 3. Analysis of molecular variance (AMOVA) of the genetic variation among and within two clusters of 89 accessions of tarwi.

Source of Variation	df	SS	MS	Est. Var.	%
Between clusters	1	4,960.23	4,960.23	95.48	7.59
Within clusters	87	101,150.54	1,162.65	1,162.65	92.41
Total	88	106,110.77	1,205.80	1,258.13	100

4. Discussion

Molecular markers represent an important component in the plant breeding area and are widely used today for multiple purposes. These markers are employed to deepen the knowledge of the diversity and population structure in plant genetic resources that help plant breeders to develop new and improved cultivars with favorable characteristics for farmers [47,48]. Knowledge of the genetic structure and diversity of germplasm collections is an important foundation for crop improvement [49]. Single nucleotide polymorphisms have gained popularity due to their abundance in the genomes and their amenability for high-throughput detection formats and platforms [50]. To date, limited studies were conducted with molecular markers to determine the genetic diversity of tarwi and other crops in Peru. We here for the first time employed genome-wide SNPs to infer the genetic diversity and population structure of Peruvian germplasm of tarwi.

Genetic diversity indices of lupin, based on SNPs, is high among the nine populations sampled across the Peruvian Andes, which is concordant for individuals that are landraces, as reported for other landraces of rye [51], pea [52], maize [53], rice [54,55], squash [56], bean [57], wheat [58]. The wealth and abundance among tarwi landraces can be explained due to their adaptation to local environments and diversity of grower's choice [49]. Unfortunately, genetic diversity indices and population structure for *L. mutabilis* have not been reported in detailed. On the other hand, these were inferred for other *Lupinus* species. Raman et al. [59] used simple sequence repeat (SSR) and DArT molecular markers and 94 Ethiopian accessions of white lupin (*L. albus*). They reported that those accessions represent a unique genepool with high level of genetic diversity. Similarly, with 11 SRAP primer pair combinations, El-Harty et al. [60] reported high genetic diversity for Egyptian white lupin genotypes. In addition, Atnaf et al. [61] used 15 SSR and 212 Ethiopian white lupin landraces, indicating that this germplasm possessed high genetic diversity. Their gene diversity (i.e. expected heterozygosity) (0.31) is close to the average value obtained in this study (0.427). A very similar gene diversity index was reported by Ji et al. [62] (0.476) for narrow-leafed lupin (*L. angustifolius*) using 76 SSR markers. On the contrary, Skorupski et al. [63] indicated that average heterozygosity of *L. nootkatensis* is 0.03. This reduced value may be explained by the isolation of this species in Iceland. Genetic studies on tarwi are scarce. Chirinos- Arias et al. [11] analyzed the genetic variability of 30 accessions of tarwi from the Andean Peruvian region with inter simple sequence repeat (ISSR) markers, indicating a broad genetic diversity among them. In a more recent study [6], a total of 23 tarwi accessions with six ISSR markers were employed and revealed important levels of diversity, but this is not related to phenotypic diversity, reflecting the recent domestication of tarwi.

Assessing population structure provides insights into the genetic diversity of the species under study and facilitates association mapping studies [64]. STRUCTURE analysis revealed that 89 samples of tarwi from the Peruvian Andes clustered in two well-defined groups associated with their geographic zone (center + south and north). Similar result were provided by PCoA. This clustering pattern meets our expectations as individuals from these two geographic zones differ on their morphology. Tarwi landraces from the center and south of Peru tend to be more compact with reduced branching and present early plant maturity. On the other hand, in northern Peru, tarwi landraces are more vigorous and possess more branching with late maturity. However, farmers cultivate tarwi under two forms of conditioning of the land i) zero tillage on fallow land. They make a hole to deposit the seeds in most of the localities in the northern regions; ii) in the north-center (Ancash, Huánuco and Huanavelica), farmers use the "yunta" for soil preparation in fallow land [65] and iii) in southern Peru, farmers have started preparing the soil by conventional tillage. That is, soil conditioning and the use of local varieties in each region influence the period of the crop cycle, which

is late when soil movement is involved [66,67]. Moreover, these differences may be explained by the latitude where these tarwi landraces are cultivated. Latitude affects plant growth significantly [68]. The higher the latitude, the shorter the growing season of these landraces with a more reduced size, as revealed in *Arabidopsis thaliana* [68] and *Ambrosia artemisiifolia* [69]. In addition, this clustering pattern may be due to the common process of exchanging tarwi seeds by growers living in close geographic areas like the center and south of Peru versus farthest places (northern area). Other lupin species such as *L. albus* [61], *L. angustifolius* [62] were also grouped into two populations.

A low degree of differentiation was exhibited among the two populations of tarwi, demonstrating that they share genetic material through high levels of breeding. Lupins are generally considered self-pollinating species [70], therefore, they tend to homozygosity. However, the negative F_{is} for both populations indicated an excess of heterozygotes for *L. mutabilis*, demonstrating that tarwi depends also on cross-pollination. These results are in agreement with Caliarì et al. [1] who indicated that outcrossing rates of *L. mutabilis* varied between 16.6% and 58.8%. Consequently, this crop should be treated as a cross-pollinated plant in breeding programs. Similarly, *L. albus* [71], *L. nootkatensis* [63] and *L. angustifolius* [62] depend also on cross-pollination. In the Andean Peruvian zone it is very common to observe populations of cultivated tarwi coexisting with its wild relative, *L. piurensis* [12], which could favor interbreeding. According to AMOVA, the greatest variation exist within accessions of tarwi (92.41%), which is explained by the sexual propagation of this species. In addition, low genetic variation between tarwi populations may be due to gene flow caused by the exchange of seeds, as depicted for *L. angustifolius* [62]. Similarly, Atnaf et al. [61] indicated that 92% of allelic variability was attributed to individuals within populations of *L. albus*.

The growing demand of novel sustainable protein sources (legumes, insects, others) [72] can be supplied by lupins, which are protein-rich legume crop but they are still limited for human consumption due to the presence of alkaloids [73]. Currently, lupin breeders only deal with a reduced part of the gene pool of this species, employing mainly low-alkaloid individuals to develop new cultivars [74]. In Peru, this study represents an initial step for breeding and conservation of this important legume also known as "lost crop of the Incas". However, further research is needed. For instance, NGS techniques should be employed to develop molecular tools for this crop, considering that its introduction in other continents will provide new source of proteins and biomass, while contributing to the improvement of poor soils [75]. Our next steps is to employ genome editing techniques for functional genomics and improvement of this neglected crop, aiming to alleviate poverty in rural communities of the Andean Peruvian region.

5. Conclusions

We here for the first time employed SNP markers distributed along all chromosomes of a neglected legume from the Andean region, tarwi, and demonstrated these markers were successful to infer the genetic diversity and population structure of this crop. As expected for a status of improvement of a landrace, different indices showed tarwi possessed high levels of genetic diversity. In addition, tarwi accessions were clustered into two populations according to their geographic zones. An excess of heterozygotes was detected, providing evidence that tarwi presents cross-pollination. Additional work should be conducted aiming to develop new tarwi cultivars by employing NGS techniques. We will also put efforts to use genome editing techniques in the near future.

Supplementary Materials: The following supporting information can be downloaded at Preprints.org, Figure S1: Plot of K ranging from 1 to 15. All K values were obtained from STRUCTURE analysis. Two populations were considered in a data set of 5922 SNPs markers and 89 samples of tarwi. Figure S2: Principal coordinate analysis (PCoA) of 89 samples of a tarwi based in geographic origin. Figure S3: Principal coordinate analysis (PCoA) of 89 samples of a tarwi based in geographic region. Table S1: Origin of the tarwi germplasm collection among the two clusters inferred by the STRUCTURE analysis.

Author Contributions: Conceptualization, A.H.-J., C.I.A, C.L.S and S.G.; methodology, A.H.-J., C.I.A, C.L.S and P. R.-G.; software C.I.A and C.L.S; validation, A.H.-J., C.I.A and C.L.S; formal analysis, A.H.-J., C.I.A and C.L.S; investigation, A.H.-J., C.I.A, C.L.S, D.S and P.I resources, A.H.-J., W.S., P.I. and C.I.A.; data curation, A.H.-J, C.L.S.

and C.I.A.; writing—original draft preparation, A.H.-J., C.I.A., C.L.S and D.S.; writing—review and editing, A.H.-J., C.I.A., C.L.S. and W.S.; visualization, A.H.-J., C.I.A and C.L.S.; supervision A.H.-J. and C.I.A.; project administration, A.H.-J.; funding acquisition, A.H.-J.

Funding: This research was funded by the project “Fondo para Regeneración de Germoplasma de Leguminosas” of the Ministry of Education (MINEDU) of the Peruvian Government. C.L.S, D.S., W.S., P.I. and C.I.A. were funded by project “Creación del servicio de agricultura de precisión en los Departamentos de Lambayeque, Huancavelica, Ucayali y San Martín 4 Departamentos” of the Ministry of Agrarian Development and Irrigation (MIDAGRI) of the Peruvian Government with grant number CUI 2449640. C.I.A. was also supported by PP0068 “Reducción de la vulnerabilidad y atención de emergencias por desastres”.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: All data generated during this study are included in this published article.

Acknowledgments: We thank Legumes and Oilseeds Research Program of UNALM and Santiago Antúnez de Mayolo University for providing the samples. Also we thank Robert Quiñones for his support with fieldwork. Finally, the authors Bioinformatics High-performance Computing server of Universidad Nacional Agraria (BioHPC-UNALM) for providing resources to perform the analyzes.

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

References

- Caligari, P.D.S.; Römer, P.; Rahim, M.A.; Huyghe, C.; Neves-Martins, J.; Sawicka-Sienkiewicz, E.J. The Potential of *Lupinus mutabilis* as a crop. **2000**, 569–573, doi:10.1007/978-94-011-4385-1_54.
- Atchison, G.W.; Nevado, B.; Eastwood, R.J.; Contreras-Ortiz, N.; Reynel, C.; Madriñán, S.; Filatov, D.A.; Hughes, C.E. Lost crops of the incas: Origins of domestication of the Andean pulse crop Tarwi, *Lupinus mutabilis*. *Am. J. Bot.* **2016**, *103*, 1592–1606, doi:10.3732/ajb.1600171.
- Neves Martins, J.M.; Talhinhos, P.; Bruno de Sousa, R. Yield and seed chemical composition of *Lupinus mutabilis* in Portugal. *Rev. Ciências Agrárias* **2016**, *39*, 518–525, doi:10.19084/rca16079.
- Carvajal-Larenas, F.E.; Linnemann, A.R.; Nout, M.J.R.; Koziol, M.; van Boekel, M.A.J.S. *Lupinus mutabilis*: Composition, Uses, Toxicology, and Debittering. *Crit. Rev. Food Sci. Nutr.* **2016**, *56*, 1454–1487, doi:10.1080/10408398.2013.772089.
- Borek, S.; Ratajczak, W.; Ratajczak, L. Regulation of storage lipid metabolism in developing and germinating lupin (*Lupinus* spp.) seeds. *Acta Physiol. Plant.* **2015**, *37*, 1–11, doi:10.1007/s11738-015-1871-2.
- Guilengue, N.; Alves, S.; Talhinhos, P.; Neves-Martins, J. Genetic and genomic diversity in a tarwi (*Lupinus mutabilis* sweet) germplasm collection and adaptability to mediterranean climate conditions. *Agronomy* **2020**, *10*, doi:10.3390/agronomy10010021.
- Adomas, B.; Galek, R.; Gas-Smerek, M.; Helios, W.; Hurej, M.; Kotecki, A.; Kozak, M.; Malarz, W.; Okorski, A.; Agnieszka, A.I.P.-C.; et al. *Adaptation of the Andean lupin (Lupinus mutabilis Sweet) to natural conditions of south-western Poland*; 2015; ISBN 9788377172353.
- Jacobsen, E.; Mujica, A. El tarwi (*Lupinus mutabilis* Sweet) y sus parientes silvestres. *Botánica Económica los Andes Cent. Univ. Mayor San Andrés, La Paz, Bolív.* **2006**, 458–482, doi:10.1038/hdy.2017.45.
- Martínez Flores, L.A.; Ruivenkamp, G.; Jongerden, J. Fitomejoramiento y racionalidad social: los efectos no intencionales de la liberación de una semilla de lupino (*Lupinus mutabilis* Sweet) en Ecuador. *Antípoda. Rev. Antropol. y Arqueol.* **2016**, 71–91.
- Camillo, M.F.; Pozzobon, M.T.; Schifino-Wittm, M.T. Chromosome numbers in South American andean species of *Lupinus* (Leguminosae). *Inst. Botánica del Nord. (IBONE). Bonplandia* **2006**.
- Chirinos-Arias, M.C.; Jiménez, J.E.; Vilca-Machaca, L.S. Analysis of Genetic Variability among thirty accessions of Andean Lupin (*Lupinus mutabilis* Sweet) using ISSR molecular markers. *Sci. Agropecu.* **2015**, *6*, 17–30, doi:10.17268/sci.agropecu.2015.01.02.
- Ruiz-Gil, P.J.; Paramo-Ortíz, E.E.; De La Luna-Bonilla, Ó.Á. ¿Entrecruces entre *Lupinus* domesticados y silvestres? El tarwi del Perú P. *Desde el Herb. CICY* **2019**, *11*, 195–200.
- Kroc, M.; Koczyk, G.; Świącicki, W.; Kilian, A.; Nelson, M.N. New evidence of ancestral polyploidy in the Genistoid legume *Lupinus angustifolius* L. (narrow-leaved lupin). *Theor. Appl. Genet.* **2014**, *127*, 1237–1249, doi:10.1007/s00122-014-2294-y.
- Soltis, P.S.; Soltis, D.E. Plant Molecular Systematics: Macromolecular Approaches: Inferences of Phylogeny and Evolutionary Processes. *Evol. Biol.* **1995**, *28*, 139–194, doi:10.2307/3243598.
- Melchinger, A.E.; Gumber, R.K. Overview of heterosis and heterotic groups in agronomic crops. *Concepts Breed. Heterosis Crop Plants* **2015**, 29–44, doi:10.2135/cssaspecpub25.c3.
- Elshire, R.J.; Glaubitz, J.C.; Sun, Q.; Poland, J.A.; Kawamoto, K.; Buckler, E.S.; Mitchell, S.E. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **2011**, *6*, 1–10, doi:10.1371/journal.pone.0019379.

17. van Inghelandt, D.; Melchinger, A.E.; Lebreton, C.; Stich, B. Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. *Theor. Appl. Genet.* **2010**, *120*, 1289–1299, doi:10.1007/s00122-009-1256-2.
18. Beebe, S.; Rengifo, E.; Gaitan, E.; Duque, M.C.; Tohme, J. Plant Genetic Resources. *Crop Sci.* **2001**, *854–862*, doi:10.4324/9781317753285-2.
19. Uitdewilligen, J.G.A.M.L.; Wolters, A.M.A.; D'hoop, B.B.; Borm, T.J.A.; Visser, R.G.F.; van Eck, H.J. A Next-Generation Sequencing Method for Genotyping-by-Sequencing of Highly Heterozygous Autotetraploid Potato. *PLoS One* **2013**, *8*, 141940, doi:10.1371/journal.pone.0062355.
20. Ramstein, G.P.; Lipka, A.E.; Lu, F.; Costich, D.E.; Cherney, J.H.; Buckler, E.S.; Casler, M.D. Genome-wide association study based on multiple imputation with low-depth sequencing data: Application to biofuel traits in reed canarygrass. *G3 Genes, Genomes, Genet.* **2015**, *5*, 891–909, doi:10.1534/g3.115.017533.
21. Wong, M.M.L.; Gujaria-Verma, N.; Ramsay, L.; Yuan, H.Y.; Caron, C.; Diapari, M.; Vandenberg, A.; Bett, K.E. Classification and characterization of species within the genus *lens* using genotyping-by-sequencing (GBS). *PLoS One* **2015**, *10*, 1–16, doi:10.1371/journal.pone.0122025.
22. Lu, F.; Romay, M.C.; Glaubitz, J.C.; Bradbury, P.J.; Elshire, R.J.; Wang, T.; Li, Y.; Li, Y.; Semagn, K.; Zhang, X.; et al. High-resolution genetic mapping of maize pan-genome sequence anchors. *Nat. Commun.* **2015**, *6*, doi:10.1038/ncomms7914.
23. Liu, H.; Bayer, M.; Druka, A.; Russell, J.R.; Hackett, C.A.; Poland, J.; Ramsay, L.; Hedley, P.E.; Waugh, R. An evaluation of genotyping by sequencing (GBS) to map the *Breviaristatum-e* (ari-e) locus in cultivated barley. *BMC Genomics* **2014**, *15*, 1–11, doi:10.1186/1471-2164-15-104.
24. Spindel, J.; Wright, M.; Chen, C.; Cobb, J.; Gage, J.; Harrington, S.; Lorieux, M.; Ahmadi, N.; McCouch, S. Bridging the genotyping gap: Using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. *Theor. Appl. Genet.* **2013**, *126*, 2699–2716, doi:10.1007/s00122-013-2166-x.
25. Jarquín, D.; Kocak, K.; Posadas, L.; Hyma, K.; Jedlicka, J.; Graef, G.; Lorenz, A. Genotyping by sequencing for genomic prediction in a soybean breeding population. *BMC Genomics* **2014**, *15*, 1–10, doi:10.1186/1471-2164-15-740.
26. Lipka, A.E.; Lu, F.; Cherney, J.H.; Buckler, E.S.; Casler, M.D.; Costich, D.E. Accelerating the switchgrass (*Panicum virgatum* L.) breeding cycle using genomic selection approaches. *PLoS One* **2014**, *9*, doi:10.1371/journal.pone.0112227.
27. Poland, J.; Endelman, J.; Dawson, J.; Rutkoski, J.; Wu, S.; Manes, Y.; Dreisigacker, S.; Crossa, J.; Sánchez-Villeda, H.; Sorrells, M.; et al. Genomic Selection in Wheat Breeding using Genotyping-by-Sequencing. *Plant Genome* **2012**, *5*, doi:10.3835/plantgenome2012.06.0006.
28. Doyle, J.J.; Doyle, J.L. Doyle_plantDNAextractCTAB_1987.pdf. *Phytochem. Bull.* **1987**, *19*, 11–15.
29. Bradbury, P.J.; Zhang, Z.; Kroon, D.E.; Casstevens, T.M.; Ramdoss, Y.; Buckler, E.S. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **2007**, *23*, 2633–2635, doi:10.1093/bioinformatics/btm308.
30. Glaubitz, J.C.; Casstevens, T.M.; Lu, F.; Harriman, J.; Elshire, R.J.; Sun, Q.; Buckler, E.S. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS One* **2014**, *9*, doi:10.1371/journal.pone.0090346.
31. Hane, J.K.; Ming, Y.; Kamphuis, L.G.; Nelson, M.N.; Garg, G.; Atkins, C.A.; Bayer, P.E.; Bravo, A.; Bringans, S.; Cannon, S.; et al. A comprehensive draft genome sequence for lupin (*Lupinus angustifolius*), an emerging health food: insights into plant–microbe interactions and legume evolution. *Plant Biotechnol. J.* **2017**, *15*, 318–330, doi:10.1111/pbi.12615.
32. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, *25*, 1754–1760, doi:10.1093/bioinformatics/btp324.
33. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158, doi:10.1093/bioinformatics/btr330.
34. Zheng, X.; Levine, D.; Shen, J.; Gogarten, S.M.; Laurie, C.; Weir, B.S. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **2012**, *28*, 3326–3328, doi:10.1093/bioinformatics/bts606.
35. Team, R.C. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing. *Vienna, Austria* **2020**, 326864.
36. Prevosti, A.; Ocaña, J.; Alonso, G. Distances between populations of *Drosophila subobscura*, based on chromosome arrangement frequencies. *Theor. Appl. Genet.* **1975**, *45*, 231–241, doi:10.1007/BF00831894.
37. Kamvar, Z.N.; Tabima, J.F.; Grünwald, N.J. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2014**, *2014*, 1–14, doi:10.7717/peerj.281.
38. Dray, S.; Dufour, A.B. The ade4 package: Implementing the duality diagram for ecologists. *J. Stat. Softw.* **2007**, *22*, 1–20, doi:10.18637/jss.v022.i04.

39. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; De Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575, doi:10.1086/519795.
40. Lischer, H.E.L.; Excoffier, L. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* **2012**, *28*, 298–299, doi:10.1093/bioinformatics/btr642.
41. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959, doi:10.1093/genetics/155.2.945.
42. Earl, D.A.; vonHoldt, B.M. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **2012**, *4*, 359–361, doi:10.1007/s12686-011-9548-7.
43. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620, doi:10.1111/j.1365-294X.2005.02553.x.
44. Francis, R.M. pophelper: an R package and web app to analyse and visualize population structure. *Mol. Ecol. Resour.* **2017**, *17*, 27–32, doi:10.1111/1755-0998.12509.
45. Arbizu, C.I.; Ellison, S.L.; Senalik, D.; Simon, P.W.; Spooner, D.M. Genotyping-by-sequencing provides the discriminating power to investigate the subspecies of *Daucus carota* (Apiaceae). *BMC Evol. Biol.* **2016**, *16*, 1–16, doi:10.1186/s12862-016-0806-x.
46. Martínez-Flores, F.; Crespo, M.B.; Simon, P.W.; Ruess, H.; Reitsma, K.; Geoffriau, E.; Allender, C.; Mezghani, N.; Spooner, D.M. Subspecies Variation of *Daucus carota* Coastal (“Gummifer”) Morphotypes (Apiaceae) Using Genotyping-by-Sequencing. *Syst. Bot.* **2020**, *45*, 688–702, doi:10.1600/036364420x15935294613527.
47. Govindaraj, M.; Vetriventhan, M.; Srinivasan, M. Importance of genetic diversity assessment in crop plants and its recent advances: An overview of its analytical perspectives. *Genet. Res. Int.* **2015**, *2015*, doi:10.1155/2015/431487.
48. Xu, Y.; Zhang, X.; Li, H.; Zheng, H.; Zhang, J.; Olsen, M.S.; Varshney, R.K.; Prasanna, B.M.; Qian, Q. Smart breeding driven by big data, artificial intelligence, and integrated genomic-enviromic prediction. *Mol. Plant* **2022**, *15*, 1664–1695, doi:10.1016/j.molp.2022.09.001.
49. Thomson, M.J.; Septiningsih, E.M.; Suwardjo, F.; Santoso, T.J.; Silitonga, T.S.; McCouch, S.R. Genetic diversity analysis of traditional and improved Indonesian rice (*Oryza sativa* L.) germplasm using microsatellite markers. *Theor. Appl. Genet.* **2007**, *114*, 559–568, doi:10.1007/s00122-006-0457-1.
50. Mammadov, J.; Aggarwal, R.; Buyyarapu, R.; Kumpatla, S. SNP markers and their impact on plant breeding. *Int. J. Plant Genomics* **2012**, *2012*, doi:10.1155/2012/728398.
51. Persson, K.; Von Bothmer, R. Genetic diversity amongst landraces of rye (*Secale cereale* L.) from northern Europe. *Hereditas* **2002**, *136*, 29–38, doi:10.1034/j.1601-5223.2002.1360105.x.
52. Martín-Sanz, A.; Caminero, C.; Jing, R.; Flavell, A.J.; Perez de la Vega, M. Genetic diversity among Spanish pea (*Pisum sativum* L.) landraces, pea cultivars and the World *Pisum* sp. core collection assessed by retrotransposon-based insertion polymorphisms (RBIPs). *Spanish J. Agric. Res.* **2011**, *9*, 166, doi:10.5424/sjar/20110901-214-10.
53. Bracco, M.; Lia, V. V.; Hernández, J.C.; Poggio, L.; Gottlieb, A.M. Genetic diversity of maize landraces from lowland and highland agro-ecosystems of Southern South America: Implications for the conservation of native resources. *Ann. Appl. Biol.* **2012**, *160*, 308–321, doi:10.1111/j.1744-7348.2012.00544.x.
54. Hour, A. ling; Hsieh, W. hsun; Chang, S. huang; Wu, Y. pei; Chin, H. shiuan; Lin, Y. rong Genetic Diversity of Landraces and Improved Varieties of Rice (*Oryza sativa* L.) in Taiwan. *Rice* **2020**, *13*, doi:10.1186/s12284-020-00445-w.
55. Aesomnuk, W.; Ruengphayak, S.; Ruanjaichon, V.; Sreewongchai, T.; Malumpong, C.; Vanavichit, A.; Toojinda, T.; Wanchana, S.; Arikitt, S. Estimation of the genetic diversity and population structure of thailand’s rice landraces using snp markers. *Agronomy* **2021**, *11*, 1–14, doi:10.3390/agronomy11050995.
56. Lorello, I.M.; Lampasona, S.C.G.; Peralta, I.E. Genetic diversity of squash landraces (*Cucurbita maxima*) collected in andean valleys of Argentina. *Rev. la Fac. Ciencias Agrar.* **2020**, *52*, 293–313.
57. Özkan, G.; Haliloglu, K.; Türkoglu, A.; Özturk, H.I.; Elkoca, E.; Poczai, P. Determining Genetic Diversity and Population Structure of Common Bean (*Phaseolus vulgaris* L.) Landraces from Türkiye Using SSR Markers. *Genes (Basel)*. **2022**, *13*, 1410.
58. Tehseen, M.M.; Tonk, F.A.; Tosun, M.; Istipliler, D.; Amri, A.; Sansaloni, C.P.; Kurtulus, E.; Mubarik, M.S.; Nazari, K. Exploring the Genetic Diversity and Population Structure of Wheat Landrace Population Conserved at ICARDA Genebank. *Front. Genet.* **2022**, *13*, 1–12, doi:10.3389/fgene.2022.900572.
59. Raman, R.; Cowley, R.B.; Raman, H.; Luckett, D.J. Analyses Using SSR and DArT Molecular Markers Reveal that Ethiopian Accessions of White Lupin (<i>Lupinus albus</i> L.) Represent a Unique Genepool. *Open J. Genet.* **2014**, *04*, 87–98, doi:10.4236/ojgen.2014.42012.
60. El-Harty, E.; Ashrie, A.; Ammar, M.; Alghamdi, S. Genetic variation among egyptian white lupin (*Lupinus albus* L.) genotypes. *Turkish J. F. Crop.* **2016**, *21*, 148–155, doi:10.17557/tjfc.95532.

61. Atnaf, M.; Yao, N.; Martina, K.; Dagne, K.; Wegary, D.; Tesfaye, K. Molecular genetic diversity and population structure of Ethiopian white lupin landraces: Implications for breeding and conservation. *PLoS One* **2017**, *12*, 1–20, doi:10.1371/journal.pone.0188696.
62. Ji, Y.; Liu, R.; Hu, J.; Huang, Y.; Wang, D.; Li, G.; Rahman, M.M.; Zhang, H.; Wang, C.; Li, M.; et al. Genetic diversity analysis for narrow-leaved lupin (*Lupinus angustifolius* L.) by SSR markers. *Mol. Biol. Rep.* **2020**, *47*, 5215–5224, doi:10.1007/s11033-020-05596-z.
63. Skorupski, J.; Szenejko, M.; Gruba-Tabaka, M.; Śmietana, P.; Panicz, R. Inferring population structure and genetic diversity of the invasive alien nootka lupin in Iceland. *Polar Res.* **2021**, *40*, 1–13, doi:10.33265/polar.v40.4536.
64. Eltaher, S.; Sallam, A.; Belamkar, V.; Emara, H.A.; Nower, A.A.; Salem, K.F.M.; Poland, J.; Baenziger, P.S. Genetic diversity and population structure of F3:6 Nebraska Winter wheat genotypes using genotyping-by-sequencing. *Front. Genet.* **2018**, *9*, 1–9, doi:10.3389/fgene.2018.00076.
65. Huaranga, A.; De la Cruz, N.; Camarena, F.; Caycho, P.; Mostacero, E.; Miguel, P. PHENOTYPIC EVALUATION AND PRELIMINARY YIELD OF LUPIN ECOTYPE (*Lupinus mutabilis* Sweet) GROWN IN CARHUAZ, ANCASH (Sweet) GROWN IN CARHUAZ, ANCASH. *Ancash. In. Abstract Book XV ILC 2019: Developing lupin crop into a modern and sustainable food and feed source*; 2019; ISBN 9789997430656.
66. Mujica, A.; Moscoso, G. *La planta del tarwi*. En Zavaleta (Ed): *Lupinus mutabilis* (Tarwi). *Leguminosa andina con gran potencial industrial*. Universidad Nacional Mayor de San Marcos, Universidad del Perú. Decana de América Fondo Editorial; 2018; ISBN 9789972466205.
67. MINAGRI Manejo agronómico. Prácticas de conservación de suelo, producción, comercialización y perspectivas de granos andinos. **2018**, 86.
68. Li, B.; Suzuki, J.I.; Hara, T. Latitudinal variation in plant size and relative growth rate in *Arabidopsis thaliana*. *Oecologia* **1998**, *115*, 293–301, doi:10.1007/s004420050519.
69. Zhou, L.; Yu, H.; Yang, K.; Chen, L.; Yin, W.; Ding, J. Latitudinal and Longitudinal Trends of Seed Traits Indicate Adaptive Strategies of an Invasive Plant. *Front. Plant Sci.* **2021**, *12*, 1–11, doi:10.3389/fpls.2021.657813.
70. Fijen, T.P.M.; Morra, E.; Kleijn, D. Pollination increases white and narrow-leaved lupin protein yields but not all crop visitors contribute to pollination. *Agric. Ecosyst. Environ.* **2021**, *313*, 107386, doi:10.1016/j.agee.2021.107386.
71. Kelemu, S.; Shiferaw, E.; Hailu, F. GENETIC DIVERSITY AND POPULATION STRUCTURE OF LUPINUS ALBUS (L.) FROM THE AMHARA REGION OF ETHIOPIA USING SEED STORAGE PROTEIN MARKERS. *J. Agric. Sci.* **2022**, *97*, 1–11, doi:10.17520/biods.2019016.
72. Nadathur, S.R.; Wanasundara, J.P.D.; Scanlin, L. *Proteins in the Diet: Challenges in Feeding the Global Population*; Elsevier Inc., 2017; ISBN 9780128027769.
73. Shrestha, S.; Hag, L. van t.; Haritos, V.S.; Dhital, S. Lupin proteins: Structure, isolation and application. *Trends Food Sci. Technol.* **2021**, *116*, 928–939, doi:10.1016/j.tifs.2021.08.035.
74. Vishnyakova, M.A.; Kushnareva, A. V.; Shelenga, T. V.; Egorova, G.P. Alkaloids of narrow-leaved lupine as a factor determining alternative ways of the crop's utilization and breeding. *Vavilovskii Zhurnal Genet. Seleksii* **2020**, *24*, 625–635, doi:10.18699/VJ20.656.
75. Gulisano, A.; Alves, S.; Rodriguez, D.; Murillo, A.; van Dinter, B.J.; Torres, A.F.; Gordillo-Romero, M.; Torres, M. de L.; Neves-Martins, J.; Paulo, M.J.; et al. Diversity and Agronomic Performance of *Lupinus mutabilis* Germplasm in European and Andean Environments. *Front. Plant Sci.* **2022**, *13*, 0–1, doi:10.3389/fpls.2022.903661.