Discover Food

RESEARCH Open Access

Assessing biodiversity of apricot landraces in Gilgit-Baltistan (Pakistan) using morphological and genetic traits

Iftikhar Alam[1], Muhammad Arslan Nawaz[2], Annette Braun-Lüllemann[3], Oliver Gailing[4,5], Markus Mueller[4,5], Andreas Buerkert[1] and Martin Wiehle[1,6*]

*Correspondence:
Martin Wiehle
wiehle@uni-kassel.de
[1]Organic Plant Production and Agroecosystems Research in the Tropics and Subtropics, University of Kassel, Steinstrasse 19, 37213 Witzenhausen, Germany
[2]W. Neudorff GmbH KG, An der Mühle 3, 31860 Emmerthal, Germany
[3]Pomologenverein, An der Kirche 5, 37318 Hohengandern, Germany
[4]Forest Genetics and Forest Tree Breeding, Georg-August University of Göttingen, Büsgenweg 2, 37077 Göttingen, Germany
[5]Center for Integrated Breeding Research (CiBreed), University of Göttingen, 37077 Göttingen, Germany
[6]Center for International Rural Development (Tropenzentrum), University of Kassel, Steinstrasse 19, 37213 Witzenhausen, Germany

Abstract

For centuries, apricots (*Prunus armeniaca* L.) were used as a staple food in Gilgit-Baltistan (G-B), northern Pakistan where their high nutritional and medicinal value is still highly appreciated by many inhabitants. Here, many locally recognized landraces (referring to mainly undescribed varieties or cultivars) of apricot are available of which some are at risk of extinction, because of current human preferences replacing landraces by modern cultivars and other crop species. To assess the distribution, diversity, management and use of apricots in G-B, we made use of the region's unique topography by employing the so-called 'Dead-end valley sampling' approach. It allowed to test the influence of market distance on morphological parameter classes. In 2021 and 2022, landrace richness, quantitative and qualitative morphological characteristics, and genetic diversity were investigated in 240 samples of several valleys. By employing 12 SSR markers we elaborated the identification of homonyms and synonyms of landraces. In addition, a total of 52 households from five valleys with three villages each as well as employees of two arboricultural research centres of Gilgit and Baltistan were interviewed. Among the sampled landraces, 120 different landraces were distinguished by farmers of which *Halman*, *Habi*, *Karfo chuli* and *Shikanda* were most frequent. The highest Simpson index (0.96) was found in Ishkoman, the most western valley studied. The high average genetic diversity (He = 0.729) across three identified genetic clusters was supported by high excess of heterozygotes, indicated by negative inbreeding coefficients. Across-valley β-diversity was 0.737. Correlation between phenotypic and genetic dendrograms was moderate ($r = 0.45$, $p = 0.001$) indicating genetically linked fruit traits, warranting further investigation. Due to its considerable genetic and phenotypic diversity, Gilgit-Baltistan constitutes an important reservoir of apricot germplasm for future breeding programs. Promoting local landraces—potentially even recognized as distinct varieties—could help farmers diversify production and income sources while preserving the region's historically significant bio-cultural heritage. To realize this potential, conservation and management measures combined with stakeholder-driven research, and controlled breeding initiatives for effective monitoring will be essential.

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Keywords α-Diversity, β-Diversity, Dead-end valley sampling, Livelihood, Microsatellite markers, Northern area, Tree management, Varietal richness

1 Introduction

Apricot (*Prunus armeniaca* L.), Rosaceae, is widely spread across temperate climates of the northern hemisphere, including the autonomous Gilgit-Baltistan (G-B) region in northern Pakistan (Fig. 1). With a total volume of 174546 tons on 13551 ha, the country's production ranks sixth after Turkey, Uzbekistan, Italy, Algeria, and Iran [19, 41]. In G-B, apricot trees are traditionally managed in small-holder agropastoral systems of the districts of Skardu, Ghanche, Hunza/Nagar, Astore, and Diamer (in descending order of quantity; Agribusiness Project—Sendal et al. [50]). Apricots of G-B are well-known for their unique colours, tastes, and flavours, sometimes even referred to as the "Hunza apricot" [53]. Apricot contributes to about 62% of the total fruit production and is thus the regions' most important fruit crop. Apricots are used as table fruit throughout the summer and dried during the winter season as a kind of staple fruit for soups and juices. Local export amounts to about 8600 t of fresh and 4374 t of dry fruits [15]. The extracted oil from bitter seeds is employed for cosmetic purpose (hair and skin care) and oil lamps, while oil from sweet seeds is used for local dishes such as "Hani-e-chzamik". Leftover kernel cake is fed to cattle and is also used by women for cleaning hair. Apricot is generally considered beneficial for stomach ailments as well as used against chronic diseases and joint pain whereby many more pharmacological uses are reported [14]. Major knowledge gaps on the morpho-genetic diversity of the local germplasm exist as the remoteness of many valleys and poor infrastructure hampered in-depth studies [21].

According to Groppi et al. [17], the genus of apricot can be divided into three major clusters: (i) the highland Central Asian group comprising germplasm from Xinjiang, Central Asia, Balochistan, Afghanistan, Pakistan and Kashmir, (ii) the Middle Eastern and Mediterranean group, and (iii) the European group, which is considered to stem from group ii. Apricots are supposed to have spread from Central Asia to the western Eurasian regions. Three main migration waves are suggested [20, 26]: (i) Traders and merchants of largely Anatolian origin, brought germplasm into the Levant and northern Mediterranean from China and Central Asia trespassing G-B along the Silk Road since about 2500 years. (ii) The Romans further distributed germplasm west on their expeditions in the 2nd century BC and gradually distributed it throughout the Mediterranean and the Middle East. (iii) Dried fruits and stones from Uzbekistan, Kyrgyzstan, Tajikistan and the Hindu Kush region were brought to Anatolia by the Persian army during campaigns by Alexander the Great around 334 BC followed by further material influx into Europe through the Romans around 1st BC. The human-induced distribution has thus likely been vital in shaping its genetic and regional diversity.

For thousands of years the intensive trade along a lateral branch of the Silk Road to the West has made G-B a corridor for plant genetic resources and the region can therefore be considered as one of the few connecting links between the crops original and newer gene pools. Hereby to our knowledge, G-B is not a centre of origin, but probably a centre of domestication and diversification. The continuous influx of germplasm (seeds and seedlings) in ancient times—as also suggested for similarly managed apricot populations in oasis systems of Tunisia [30]—is considered a major cause of apricot diversity in

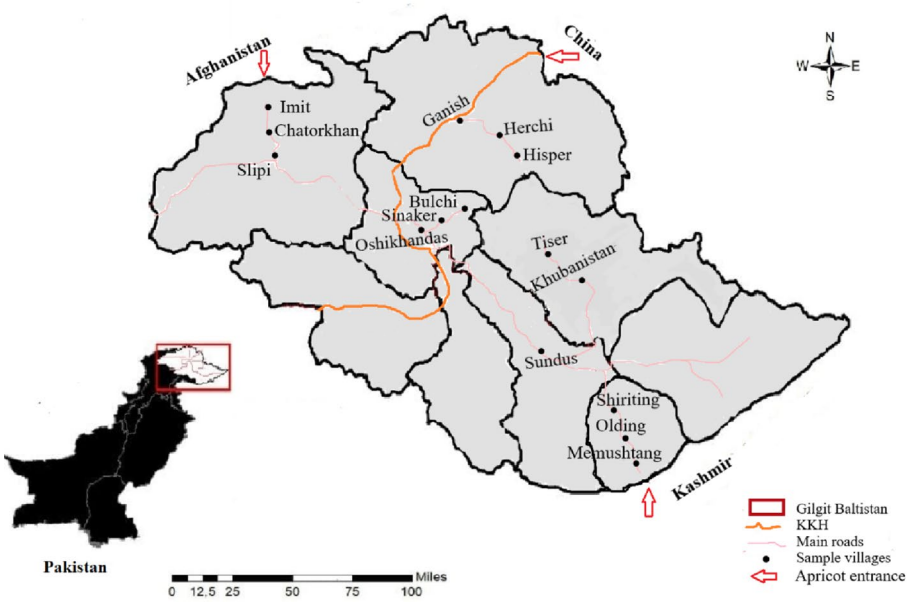


Fig. 1 Map of studied villages used for local apricot varieties' collection in five valleys of Gilgit-Baltistan, northern Pakistan.(sources: DICA-GIS, www.diva-gis.org/gdata. Accessed 10th November 2022)

G-B and beyond. In addition, the largely spontaneous growth of seedlings has resulted in a wide range of landraces with marketable fruit qualities, while vegetative propagation techniques have contributed to the creation of well recognized varieties [21, 36]. Moreover, the effects of intense UV radiation in this high altitude region might further result in mutant creation, although to our knowledge research in this discipline is limited [56].

In G-B, local landraces are distinguished by morphological characteristics with unique local names of the Baltistan region [3], such as *Alishah Kakas*, *Alman*, *Badam Zu*, *Balani*, *Burum Zu*, *Charmagzi*, *Dudor*, *Durkali*, *Ghamburi*, *Habi*, *Mamur*, *Mirza Khani Neli*, *Shikanda*, *Surya's Zu* of the Gilgit region and *Bro Chuli*, *Dospa*, *Halman*, *Kharfo Chuli*, *Khapouk*, *Kho Chuli*, *Marghulam*, *Saracat*, *Shakanda*, *Ta Chuli*, *Walfi Chuli*, *Yakkar*. However, due to the linguistic diversity in the region of G-B, it remains unclear until today if the named landraces are true-to-type, homonyms, or synonyms.

This rich set of yet unknown diversity may capture potentially valuable traits for future breeding [43]. Many indigenous apricot varieties are largely resistant to biotic and abiotic stress such as powdery mildew—Erysiphales [34] and Plum pox virus—Patatavirales [9], and are highly tolerant to drought and low relative air humidity [54]. In contrast, newly introduced and high-yielding cultivars often lack these features but supersede local landraces. Especially old individuals are successively logged for their valuable wood and to provide space for agricultural land and housing. Both processes have led to a massive uniformization of highly diverse inter- and intraspecific fruit tree orchards and the loss of likely century-old individuals, narrowing the traditional gene stock in G-B as predicted already more than 30 years ago [59]. The extent to which indigenous apricot landraces from Gilgit-Baltistan are preserved in germplasm banks locally or worldwide is still unknown. Thus, some rare landraces still grown in the field are likely in danger of extinction. Genetic narrowing has moreover made orchards very susceptible to diseases. Among the diseases, Gummosis (caused by fungi such as *Cytospora*) and Shot hole are most devastating on apricots. Their occurrence is partly caused by careless pruning and

grafting with unsterilized tools and by environmental changes [1]. Farmers on traditional agricultural land often do not take care of proper planting distances, grafting, pruning, application of fertilizers, and disease management. Moreover, non-standardized cleaning, grading and packaging processes as well as the employment of unskilled labour, the use of unmarketable landraces/varieties, and restricted transportation facilities in the remote region of G-B result in poor yields and fruit qualities with an estimated postharvest loss of 40% to 50% [2]. This in turn, affects local revenues, hampers farmers' willingness to cultivate local apricot landraces/varieties in the future, and may thus contribute to the gradual extinction of the rich pool of available germplasm. Lastly, farmers prefer growing wheat (*Triticum aestivum* L.) over apricot as the harvesting seasons coincide.

The points outlined above call for an assessment of apricot distribution, morphology, genetic diversity, management, and use in G-B. Usually, morphological characterization is used for cultivar identification and germplasm management with the aim of selecting superior genotypes for breeding programs [57]. Molecular markers such as Simple Sequence Repeat (SSR) markers in contrast, have been fundamental to evaluate the genetic diversity of apricot, even to the sub-species/varietal level [6, 55].

Given the above, our research aimed at (i) exploring quantitative and qualitative fruit characteristics, (ii) assessing the varietal (referred hereafter as landrace) richness and diversity within and across villages and valleys, and (iii) identifying genetic variation among landraces. We further wanted to determine the complex relationship between nomenclature and landrace identity. Such comprehensive knowledge will support the development of site-specific management plans to safeguard the diverse apricot germplasm and could inform future domestication and conservation strategies in the region and beyond.

2 Materials and methods

2.1 Study area

Gilgit-Baltistan, a 72,496 km² large area located between 35–37° N and 72–75° E, is geographically at the periphery of Central Asia and hosts a vast richness of wild and crop species due to its rugged topography with climatic and agricultural niche conditions [25]. Gilgit-Baltistan is distinguished by long summer daylight hours, low relative humidity, and a steady supply of glacier meltwater distributed through traditional *karez* irrigation systems. Despite the relatively small area, a number of diverse ethnic groups coexist in Gilgit-Baltistan, of which the largest tribes are the Burusho, Sheena, Balti, Wakhi, and Khuwar that can be differentiated by cultural practices and languages [29].

Reportedly, 96% of the local population directly depends on agriculture by practicing subsistence farming within an extended family system [2]. They receive up to 50% of their income from trees, through the sale of fresh and dry fruits [28]. Especially women contribute significantly to household incomes along the apricot value chain and in turn, benefit from its revenues [33].

2.2 Sampling design

In the summers of July to October 2021 and 2022 a field survey was conducted focusing on five valleys (Kharmang, Shiger, Nagar, Ishkoman, and Bagrot) across an altitudinal range of 1457 m to 3380 m a.s.l. (Table 1). We hereby used the so called 'Dead-end valley' sampling approach, assuming that the valleys can be only entered through one road

Table 1 Information on sample numbers and geo-spatial characteristics per region, dead-end valley, and village (highest first) in Gilgit-Baltistan, Northern Pakistan between July and Oct 2021–2022

Region	Valley	Village	No. of tree owners	Total trees	Trees with fruits	Coordinate N, E	Elevation (m)
Baltistan	Kharmang	Memushtang	3	13	13	34° 45′, 76° 04′	3380
		Olding	3	11	11	34° 44′, 76° 09′	2729
		Shiriting	3	11	11	35° 51′, 76° 11′	2533
	Shiger	Tisar	4	10	10	35° 40′, 75° 27′	2388
		Khubanistan	3	18	15	35° 28′, 75° 42′	2294
		Sundus	3	15	14	35° 17′, 75° 37′	2251
Gilgit	Nagar	Hisper	3	13	*	36° 10′, 74° 59′	3123
		Herchi	4	15	14	36° 16′, 74° 43′	2368
		Ganish	4	15	15	36° 18′, 74° 40′	2137
	Ishkoman	Imit	4	12	9	36° 30′, 73° 54′	2397
		Chatorkhan	4	25	19	36° 20′, 73° 35′	2110
		Slipi	4	13	*	36° 11′, 73° 46′	1890
	Bagrot	Bulchi	4	10	*	35° 01′, 74° 33′	2425
		Sinaker	3	25	18	36° 57′, 74° 30′	2085
		Oshikhandas	3	14	*	35° 35′, 74° 28′	1457
	Research centre	MARC Skardu	1	3	*	35° 17′, 75° 38′	2255
		MARC Jaglot	1	17	*	36° 57′, 74° 30′	2085
Total			54	240	149		

*Fruits were immature or already harvested/perished

and are thus characterised by market-close (down-valley) and market-distant (up-valley) villages [58], resulting in a total of 15 villages [27]. Per village, all landraces were sampled until no new landrace was identified [58] or at least not considered morphologically different from the specimens already sampled, resulting in a total of 220 samples. In addition, reference materials from the Mountain Agricultural Research Centre (MARC) in Skardu (varieties $n = 3$) and Jaglot (landraces $n = 15$, varieties $n = 2$) were collected, totalling 240 tree samples. The locations of individual trees were recorded using a handheld GPS device (Garmin eTrex 30; horizontal accuracy ± 2 m, GARMIN® Ltd., Southampton, UK). For field data collection, three pre-designed questionnaires and descriptor sets were applied: (i) a dendrometric traits questionnaire (A1) to assess tree performance and characteristics, and (ii) two fruit phenotypic trait descriptors, one qualitative (A2) and one quantitative (A3).

2.3 Local nomenclature of apricot in Gilgit-Baltistan

To unravel the complex relationship between landrace name, morphology and genetics, we defined identity categories as follows: (1) true-to-type = individuals with the same name given by the farmer and same genotype that matches the allelic composition to 100%; (2) homonym = same name, but different genotype (< 100% match); (3) synonym = different name, but same genotype; and (4) unique = landraces that were only mentioned once. Traditionally and occasionally received and introduced plant materials (seeds, seedlings, scions) are named after the (foremost ancient) person who brought it to the nursery or farm, for instance *Ali Shah*, *Khan Bahadur*, and *Marghulam*. Frequently, landraces have a second, sometimes even a third suffix according to fruit characteristics (such as taste, size, shapes, and colour) ending mainly with the local term for apricot itself (Table 2). During the field survey, landraces that could not be named

Table 2 Freely combinable expressions related to apricot landraces/traits in the five local languages present in Gilgit-Baltistan, Northern Pakistan between July and October 2021 and 2022

Languages	Apricot	Wild trees	Taste		Colour	Purpose/use		
			Sweet	Sour	White	Cake	Juice	Soup
Burushaski	*Zu*	*Jangali*	*Shikanda*	*Shuqour*	*Burum*	*Minna*	*Chamous*	*Dawodo*
Balti	*Chuli*	*Kho*	*Staa*	*Sikrmo*	*Karfo*	*Kounoq*	*Droug*	*Chafay*
Shina	*Jue*	*Galia*	*Mispa*	*Churko*	*Shee Jue*	*Peno*	*Chamos*	*Baleyai*
Whaki	*Chuwen*	*Tash*	*Khuzhg*	*Treshp*	*Rikhn*	*Manai*	*Chumos*	*Qaq Moch*
Khuwar	*Juli*	*Lalme*	*Sheerini*	*Shouth*	*Eshpayru*	*Piano*	*Chamborogh*	*Tushmach*

by farmers, but were morphologically different from the already screened landraces, received working titles with consecutive numberings (e.g. *Zoqi Zu 1*, *Zoqi Zu 2*; Table A4).

2.4 Dendrological and morphological characterization

The sampling of morphometric parameters was based on the approach of Krichen et al. [30]. Each selected tree was described for standard dendrological parameters, including tree height (m by interception theorems), crown diameter (m), and stem girth at breast height (cm), each measured with a measuring tape. Girth at breast height was later on recalculated into diameter at breast height (DBH). Additional biophysical and management-related data were collected, including tree vitality, presence of grafting marks, thinning practices, and visible symptomes of diseases or pest attacks. The area of cultivation was also briefly described and classified later on into cultivated or uncultivated land (outside farmer's field; questionnaire A1).

2.5 Fruit phenotypic description

Per tree, five representative fruits were randomly collected for phenotypic description, resulting in a total of 745 investigated fruits. Qualitative characters (fruit shape, fruit peel surface, pulp texture, and stone shape) were recorded based on the phenotypic fruit descriptor A2. Fruit colour was determined by using the Royal Horticultural Society (RHS) colour chart (sixth edition, 2015, RHS media, 80 Vincent-Square, London, England). Fragrance and taste were assessed by two of the authors independently each. Quantitative characters of fruits and stones (in three dimensions—length, width, depth) were measured in millimetres with the help of a digital vernier calliper. Fruit and stone weight were measured with the help of US-ACE Digital Grain Scale (accuracy ± 0.1 g; descriptor A3). To quantify the share of flesh and seed, we calculated the fruit flesh to stone weight ratio. Fruit and stone volumes were calculated assuming an ellipsoid form. For comparative and documentation reasons, fruits, leaves, and stones were photographed with a digital camera. For a more accurate determination of stone dimensions, colours, and shapes, the IRIS Visual Analyzer (Alpha MOS, Toulouse, France) was used to photograph stones and later analysed with the embedded picture analysis software AlphaSoft (Alpha MOS, Toulouse, France).

2.6 DNA isolation and SSR analysis

Per tree, 3–4 leaves were sampled, shade air-dried and stored in tea filter bags. From one leaf, ca. 0.25 cm^2 were taken and milled. Genomic DNA was extracted from 238 leaf samples (with 2 lost) using the DNeasy™ 96 Plant Kit (Qiagen GmbH, Hilden, Germany). A literature review identified the most commonly applied microsatellite (SSR)

markers, of which eighteen were selected for screening. Out of these, twelve amplifiable and highly polymorphic SSR markers were used (Table 3), which were also applied by Basile et al. [5] and Corrado et al. [8].

The amplification reaction was conducted in 1 µL dNTP (2.5 mM each dNTP) 0.3 µM of each forward (with M13-specific sequence and dye labelled, [31, 49] and reverse primer (PIG-tailed according to Brownstein et al. 1996), 0.2 µL (5U μL^{-1}) unit of *Taq* DNA polymerase (HOT FIREPol® DNA Polymerase, Solis BioDyne OÜ, Tartu, Estonia), 1.5 µL 10x reaction buffer B (Solis BioDyne OÜ, Tartu, Estonia), 1.5 µL $MgCl_2$ (25 mM), and 1 µl DNA (ca. 0.6 ng μL^{-1}).

The PCR reaction started with an initial denaturation at 95 °C for 15 min to fully denature the DNA template and activate the polymerase. This was followed by 10 touch-down cycles, each consisting of denaturation at 94 °C for 1 min, annealing for 1 min starting at 60 °C and decreasing by 1 °C per cycle, and extension at 72 °C for 1 min. After the touch-down phase, 25 standard cycles were performed, each with denaturation at 94 °C for 1 min, annealing at 50 °C for 1 min, and extension at 72 °C for 1 min. The program concluded with a final extension at 72 °C for 20 min to ensure complete product

Table 3 Comparison of genetic variation of apricot populations from 15 villages within five valleys of Gilgit-Baltistan, Pakistan, using 12 SSR markers

SSR ID	Linkage group	Primer sequence (5′-3′) M13: F + CACGACGTTGTAAAACGAC PIG-tailed: GTTTCTT + *R*	Size range (bp)	Na	Ho[1]	He[2]	F_{IS}
AMPA101[a]	7	F: CACGACGTTGTAAACGACCAGTTTGATTT R: GTTTCTTGATCCACCCTTTGCATAAAATC	190–210	8	0.771	0.742	−0.039 ***
AMPA119 [a]	1	F: CACGACGTTGTAAACGACGTGCCCACTTA R: GTTTCTTTCGACGATCAGACTTGCTACAG	104–108	6	0.617	0.586	−0.052 ***
pchgms5[b]	3	F: CACGACGTTGTAAACGACCGCCCATGAC R: GTTTCTTGTCAAGAGGTACACCAG	284–310	12	0.620	0.645	0.039
ssrPaCITA19[c]	5	F: CACGACGTTGTAAACGACGACAAATACA R: GTTTCTTGAACAGCTAGCCCCTTTGTCATAC	114	10	0.648	0.705	0.081 ***
UDAp-410[c]	5	F:CACGACGTTGTAAACGACTTGTTGACAAGA R: GTTTCTTCAACGGGTTGGTTTCAGAAG	116–146	9	0.826	0.815	−0.013 ***
UDAp-414[c]	1	F: CACGACGTTGTAAACGACCAAGCACAAGC R: GTTTCTTGGTGGTTTCTTATCCGATGC	150–214	11	0.786	0.789	0.004
UDAp-420[c]	6	F: CACGACGTTGTAAACGACTTCCTTGCTTCC R: GTTTCTTCCCAGAACTTGATTCTGACCA	154–262	9	0.825	0.770	−0.071
UDP96001[d]	–	F: CACGACGTTGTAAACGACAGTTTGATTTT R: GTTTCTTTGCCATAAGGACCGGTATGT	104	3	0.595	0.497	−0.197 *
UDP98406[d]	2	F: CACGACGTTGTAAACGACTCGGAAACTG R: GTTTCTTATGGGTCGTATGCACAGTCA	82–100	11	0.871	0.795	−0.096 **
UDP98-409[d]	8	F: CACGACGTTGTAAACGACGCTGATGGGTTT R: GTTTCTTCGGACTCTTATCCTCTATCAACA	118–161	17	0.825	0.838	0.015***
UDP98-411[d]	2	F: CACGACGTTGTAAACGACAAGCCATCCAC R: GTTTCTTCCAAAAACCAAAACCAAAGG	154–180	12	0.769	0.767	−0.002 ***
UDP98-412[d]	6	F: CACGACGTTGTAAACGACAGGGAAAGTT R: GTTTCTTGCTGAAGACGACGATGATGA	81–114	10	0.799	0.802	0.003 ***
Mean				9.8	0.746	0.729	− 0.027

Na—number of alleles; He—expected heterozygosity; Ho—observed heterozygosity; F_{IS}—fixation index, Exact test significant at P-value < 0.0001. Primers developed by: [a]Hagen et al. [18], [b]Sosinski et al. [51], [c]Lopes et al. [35], [d]Cipriani et al. [7]; primers contain M13 sequences (forward primers) or PIG-tail sequences (reverse primers) at their 5′ ends (see main text)

[1] He: < 0.3 (low), 0.3 ≤ He ≤ 0.6 (moderate), > 0.6 (high)

[2] Ho: < 0.2 (low), 0.2 ≤ Ho ≤ 0.5 (moderate), > 0.5 (high)

formation. Positive and negative controls were included and repeated in all reaction plates to guarantee reproducibility. The PCR fragments were then separated and sized using an ABI 3130xl Genetic Analyzer (Applied Biosystems Inc., Foster City, CA, USA) with the GS 500 ROX (Applied Biosystems Inc.) used as an internal size standard. Genotyping was conducted using the GeneMapper 4.1® software (Applied Biosystems Inc.).

2.7 Data analysis

For dendrometric, fruit and stone phenotypic traits, descriptive statistics including the mean coefficients of variation (CV%) were calculated for the available sample numbers (total trees n = 240, trees with fruits n = 149). Per village, the abundance and richness of landraces were assessed and landrace alpha (α)-calculated. Further α-diversity measures (Simpson's and Pielou's index) were estimated by using the 'vegan' package in R (version 4.2.2, R Core Team, 2021, Vienna, Austria [40]), and for the assessment of variation among the quantitative and qualitative traits Principal Component Analysis (PCA) was used. A hierarchical clustering was performed from landrace diversity data using Chord distance and a contiguity-constrained segmentation approach [32]. The beta (β)-diversity for each node of the dendrogram was computed by weighing all putative landraces inside the associated sub-cluster according to their total abundance [48, 58]. Each landrace's beta-diversity contribution (%β; range: 0–100%) was calculated per village.

A dissimilarity matrix was generated based on stone traits (area, roundness, width, and height) using quantitative and qualitative data for 149 individuals bearing fruits (Table 1). The Gower distance was applied via the 'daisy' package [24].

The number and efficacy of SSR markers were assessed using the probability of identity for unrelated (P_{ID}) and related (P_{IDsib}) individuals. Genetic diversity was calculated with GenAlEx 6.5 software [44] including the percentage of polymorphic loci, the number of alleles per locus (Na), the number of effective alleles (Ne), the number of private alleles, expected (He) and observed (Ho) heterozygosity for each region, valley, and village as well as isolation-by-distance (IBD, ~Mantel Test).

The genetic clustering analysis was conducted using the Bayesian hierarchical clustering algorithm implemented in the STRUCTURE 2.3.4 software [47]. An admixture model assuming correlated allele frequencies was employed, with a burn-in of 10,000 followed by 100,000 MCMC replicates. To determine the optimal number of genetic clusters (K), we performed 10 iterations for each of the 1 to 10 possible values of K. The most likely value of K was determined using the software STRUCTURE-HARVESTER [11] based on the highest value of ΔK and the lowest value of mean posterior probability (LnP(K)) of the data [12].

Additionally, allelic data was analyzed using the 'poppr' package [22] to estimate Rogers' Euclidean distance, which was used to construct a genetic dendrogram. This dendrogram was based on the genetic diversity and differentiation of 12 SSR markers (Table 3) that were applied to DNA samples from apricot trees in 15 villages in five valleys. The Ward D method [39] was employed to create the dendrogram for both 149 (fruit bearing trees) and 238 individuals. The genetic cluster information extracted from the dendrogram helped to quantify putative homonyms, synonyms, and unique accessions (Fig. 9).

The adequacy of both phenotypic and genetic distances (n = 149, each) was assessed using the tanglegram approach embedded in the 'dendextend' package [24]. Because the dendrograms were based on different distance measures, cophenetic correlation

matrices were generated. These matrices were constructed based on the dendrograms themselves by subjecting them to a Mantel test through the 'vegan' package [10]. The ensuing cophenetic correlation coefficient (range: −1 to 1) indicated the similarity between matrices and thus justified their further use for the tanglegram chart. In addition, the entanglement quality index (range: 0—full association to 1—full divergence) between both clusters was computed to unravel the meaningfulness of the obtained tanglegram. Finally, the number of clusters for both phenotypic and genetic dendrograms was determined using the elbow criteria method, which is incorporated in the 'factoextra' package [23].

3 Results

The study revealed a high diversity reflected in dendrological, morphological, and genetic traits assessed in the field and the lab.

3.1 Landrace richness and diversity

A total of 120 individual landraces were differentially named by farmers and a few more ($n = 20$) received a consecutive numbering due to obvious morphological distinctiveness (Table A1). Thus, almost half (47.5%) out of 240 collected individuals have been assigned a nomenclature across the sampling area, resulting in a large variation of α-diversity per region, valley, and village (Table 4). α-diversity was highest in the Gilgit region, particularly in the two valleys Ishkoman and Bagrot. Medium altitude villages Chatorkhan,

Table 4 Relative abundance and contribution of each apricot landrace/variety ($n = 240$) to overall α-diversity among villages, valleys, and regions of Gilgit-Baltistan, Pakistan, surveyed between July and October 2021 and 2022

Region	Valley	Village	Altitude	Richness	Simpson Index range: 0–1	Pielou's index range: 0–1
Baltistan	Kharmang	Memushtang	High	11	0.89	0.38
		Olding	Medium	11	0.90	0.38
		Shiriting	Low	9	0.88	0.40
	Shiger	Tisar	High	8	0.84	0.40
		Khubanistan	Medium	13	0.90	0.35
		Sundus	Low	10	0.88	0.38
Gilgit	Nagar	Hisper	High	10	0.89	0.39
		Herchi	Medium	11	0.87	0.38
		Ghanish	Low	12	0.88	0.35
	Ishkoman	Imit	High	8	0.82	0.39
		Chatorkhan	Medium	23	0.95	0.30
		Slipi	Low	10	0.89	0.39
	Bagrot	Bulchi	High	8	0.86	0.41
		Sinaker	Medium	15	0.92	0.34
		Oshikhandas	Low	10	0.87	0.38
	Research	MARC Skardu		17	0.67	0.33
	Centre	SMAR Jaglot		3	0.67	0.61
	Kharmang		Highest	22	0.93	0.30
	Shiger			22	0.93	0.30
	Nagar			27	0.94	0.29
	Ishkoman			37	0.96	0.27
	Bagrot		Lowest	26	0.94	0.29
Baltistan				38	0.95	0.26
Gilgit				93	0.98	0.22

Khubanistan, Olding, and Sinaker showed highest values except for Nagar Valley in which the up-valley village Hisper displayed the highest α-diversity.

Across-village β-diversities resulted in four distinct clusters (Fig. 2). The overall weighted β-diversity was 0.737, where landraces with beta-contribution (%β) values above 1 (*Habi*, *Halman*, *Karfo Chuli*, and *Shikanda*) contributed most to the cluster formation (Table A4). Only two landraces, namely *Chite Jue* and *Halman*, were shared between the two regions. Across valleys, a number of landraces were commonly shared such as *Halman*, *Kharfo chuli*, *Kho chuli*, and *Shakanda* (Baltistan: Kharmang and Shiger), *Halmas*, *Shikanda*, *Jangir*, *Habi* and *Burum Zu* (Gilgit: Bagrot and Nagar).

3.2 Dendrometry, management, and use

On uncultivated land (outside field margins), 62% of trees were grown (Table A2). According to farmer's statement, tree ages ranged from 2- to 150-years (78% of tree samples were between 2 and 40 years old), whereby most farmers did not know the exact age of their trees. The average tree height of 4.6 m ranged from 1 to 8 m (CV% = 46.4), while stem diameter averaged 0.3 m (0.3–1.1 m) with the highest CV% (59.8) across all dendrometric parameters assessed (Table A1). Crown diameter averaged 6.7 m (0.7 to 14.1 m, CV% = 44.9). The largest trees were found in the Chatorkhan village (Ishkoman).

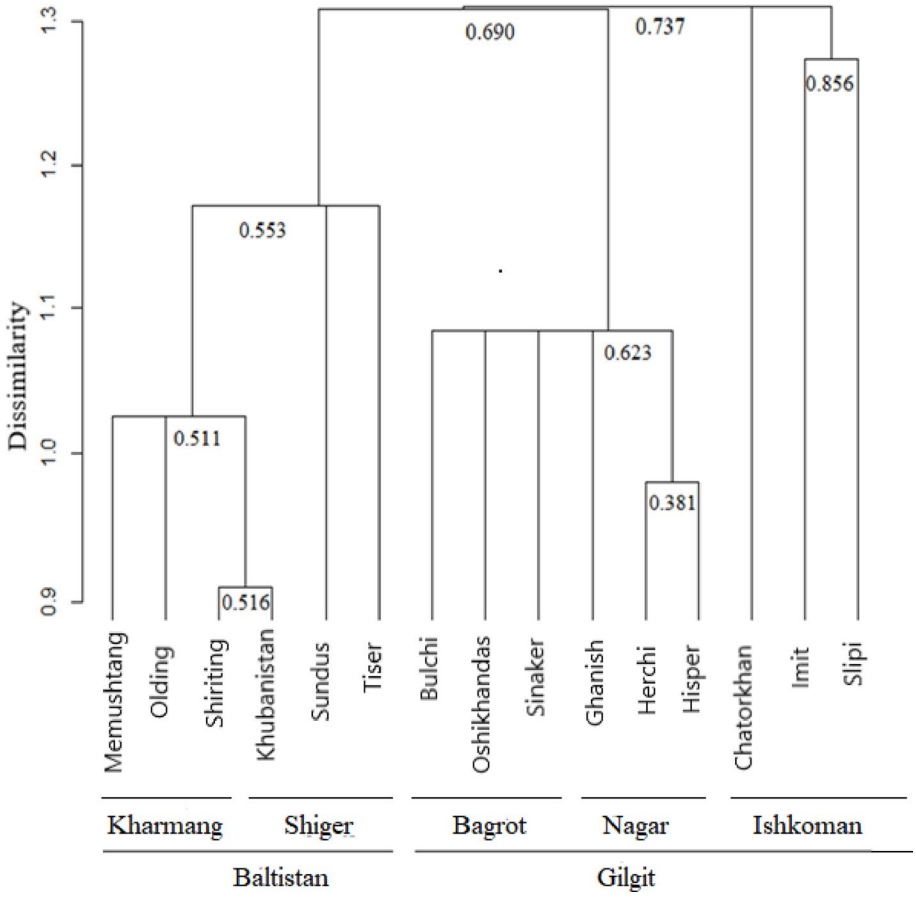


Fig. 2 Dendrogram based on contiguity-constrained abundance values and chord distance for 220 apricot individuals from 15 villages across five valleys of Gilgit-Baltistan, Pakistan, excluding reference material from MARC Skardu and Jaglot. β-diversity values are given at each node; total average weighted beta is 0.737

More than half (53%) of the plant material (scions and plants) was reportedly brought/planted by the farmers' ancestors (grandfathers and fathers), 29% of the trees were planted through seeds by the owner/farmer, 14% came from nurseries (governmental/private), while the origin of 4% was unclear (Table A2). The majority of farmers remembered that trees were grafted (73%), however, grafting spots were only visible on 11% of the trees. Scions were cleft grafted on locally grown root stocks (apricot seedlings), while 10% were grafted on plum or peach root stock seedlings. Knowledge on trimming and pruning was underrepresented; most farmers only removed dry and old branches from the trees and used them as fuelwood, while no arbori-cultural practices were witnessed.

Flowering of trees reportedly lasted from mid-March (1457 m, Oshikhandas, Gilgit region) to the last week of April (3380 m, Memushtang, Baltistan region) although flowering fluctuated between 10 and 15 days across years. During the first part of July, 79% of the fruits matured and were harvested manually by climbing trees and shaking branches, often with the help of sticks. Dropped fruits were collected from the ground by using crates/cartoons and local willow baskets (*Giran*). Half (49%) of the trees produced between 30 and 60 kg of fruits (Table A2). Trees were categorised into three groups regarding fruit usage into: dry (46%), fresh (16%), and both, dry and fresh (6%). A quarter of the landraces are popular for a number of traits, such as size, colour, taste, and suitability for drying. Traditionally, fruits are sun-dried, but more than half (51%) were additionally sulphured (sulphur fumigation) for increased quality and marketing demands (increased shelf life, improved optical features, reduced insect infestation). Over the last five years, reportedly 80% of the produce was sold dried and 4% as table fruits in nearby markets.

3.3 Fruit phenotype

3.3.1 Quantitative description

We found the highest fruit weight and volume in the Gilgit region, here particularly in the Ishkoman valley (Table 5 and Table A3 for landrace specific information). The smallest and largest fruit weight and volume differed by a magnitude of 5.5 and 4.9 times, respectively. For stone weight and volume, the same range of differences (4.2 and 4.4 times, respectively) was observed. Fruit flesh weight and weight of fruit: stone ratio was highest in Ishkoman and lowest in Shiger (Tables 5 and 6).

Principal component analysis (PCA) was performed using measurements of stone area, roundness, width, and height. The first two principal components, PC1 and PC2, accounted for 92% of the total variance (68% and 24%, respectively). The PCA biplot (Fig. 3) indicated that stone area, width, and height were positively correlated, whereas stone roundness was negatively associated with PC1.

3.3.2 Qualitative description

Out of the 149 fruit bearing trees (Table 1), 60% of the fruits were sweet (Table A2), while *Ali Shah Halman*, *Marghulam*, *Kakas*, *Ramazan Narmo*, and *Shakanda* were evaluated as very sweet. Most (80%) landraces had weak fragrance, while only a few such as *Sharacat* had an intensive fragrance. *Chuli*, *Dulkali*, *Karfo*, *Kurue Juli*, *Sharacat*, and *Washil Juli* had a strong flesh-to-stone adherence (6%). Most (77%) landraces were moderately juicy, while especially *Ghamburi*, *Karfo Chuli*, *Marghulam*, *Mirza Khani*, *Shreerini*, *Shikhanda*, and *Yarkhandi* were evaluated as very juicy. The absence of kernel

Table 5 Average fruit phenotypic traits of largest and smallest apricot landraces collected from five valleys in GB, Pakistan, between July and October 2021 and 2022

Regions/valleys group*	Landraces	Fruit		Landraces	Stone	
		Weight (FW) g	Volume (cm^3)		Weight (SW) g	Volume cm^3
Baltistan						
Kharmang						
L	*Batapa*	24.1	24.2	*Bato Shakanda*	2.6	3.7
S	*Khon Chuli*	7.7	7.2	*Fat Chuli*	0.9	1.0
Shiger						
L	*Marghulam*	37.1	28.0	*Sharacat*	3.8	4.4
S	*Kho Silbow Chuli*	11.2	8.7	*Fat Chuli*	1.2	1.6
Gilgit						
Nagar						
L	*Dala Shah*	26.7	23.9	*Dala Shah*	2.6	3.0
S	*Dudor*	10.4	8.4	*Dudor*	1.1	1.3
Ishkoman						
L	*Ghamburi*	41.6	35.1	*Ghamburi*	2.9	4.0
S	*Sakar Chuwan*	10.1	7.3	*Toshtako*	1.1	1.3
Bagrot						
L	*Neli Pewan*	31.0	30.0	*Neli Pewan*	2.4	3.7
S	*Lole Jue*	11.6	10.2	*Chite Jue*	1.3	1.5
Mean		21.2	18.3		2.0	2.6
STD		12.6	11.0		1.0	1.3
CV %		59.4	60.0		50.2	52.5

*L = largest, S = smallest

Table 6 Average fruit phenotypic traits of apricot varieties collected from five valleys in GB, Pakistan, between July and October 2021 and 2022

Region/valleys	Group	Landraces	Flesh thickness (mm)	Landraces	FW:SW ratio
Baltistan					
Kharmang	L	*Batapa*	10.4	*Chite Jue*	12.3
	S	*Khon Chuli*	5.3	*Khon Chuli*	5.7
Shiger	L	*Marghulam*	12.0	*Marghulam*	12.1
	S	*Kho Silbow Chuli*	6.7	*Kho Silbow Chuli*	5.2
Gilgit					
Nagar	L	*Dala Shah*	11.0	*Ghulam-E Zu*	13.6
	S	*Dudor*	6.7	*Shukhor Zu*	8.1
Ishkoman	L	*Sharakh*	12.5	*Sheerini*	19.4
	S	*Sakar Chuwan*	5.7	*Sakar Chuwan*	7.2
Bagrot	L	*Chusangi*	10.5	*Zoqi 7*	13.1
	S	*Loli Jue*	6.2	*Younch*	6.7
Mean			8.7		10.3
STD			2.8		4.5
CV %			32.4		43.6

* Letter L for largest and S for Smallest; FW = Flash weight, SW = Seed weight

bitterness was quantified in 81% of the samples, while 5% and 14% indicated a weak to strong bitterness, respectively. The six most common fruit color shades were strong orange-yellow (20%), brilliant yellow (15%), strong orange and light orange-yellow (each 12%), and light greenish-yellow and brilliant orange-yellow (9% each; Fig. 4). The PCA of all qualitative traits of fruits and stones showed overall low variations among villages and valleys (Figs. 5 and 6).

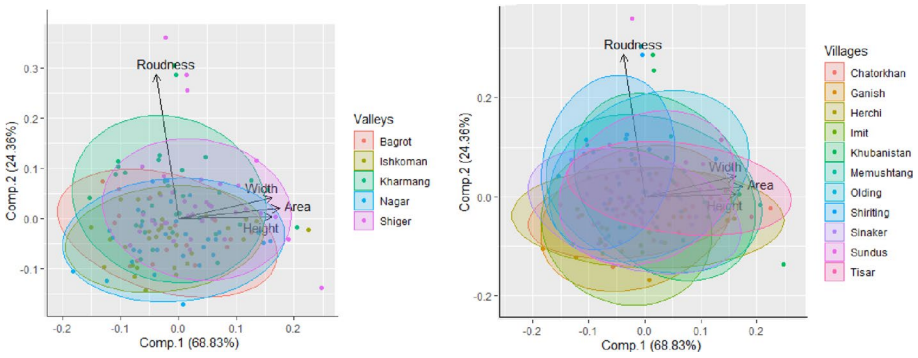


Fig. 3 Principal component analysis (PCA) of apricot stones quantitative parameters in valleys and villages of Gilgit-Baltistan, Pakistan

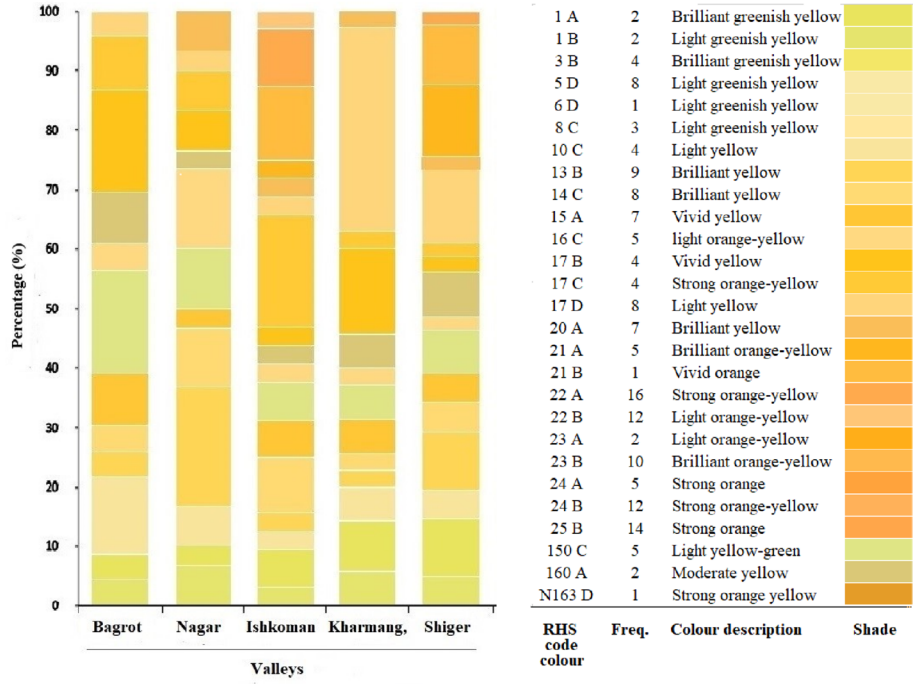


Fig. 4 Relative shades of fruit colours based on Royal Horticultural Society colour charts observed in Gilgit-Baltistan, Pakistan during survey in July-October 2021 and 2022

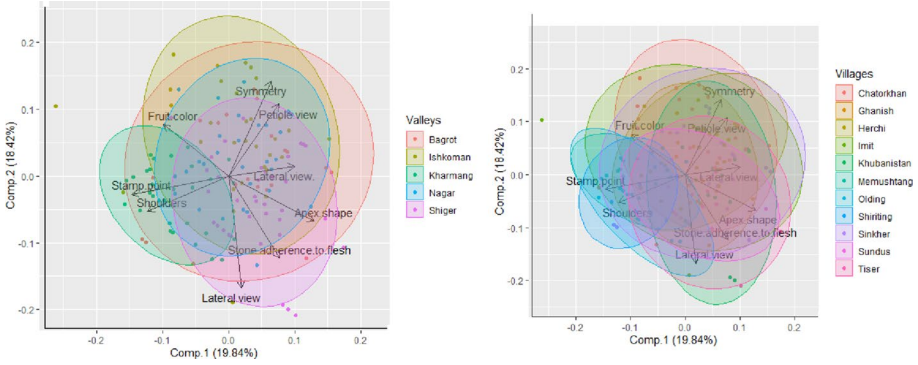


Fig. 5 Principal component analysis (PCA) of fruit qualitative parameters in valleys and villages of Gilgit-Baltistan, Pakistan

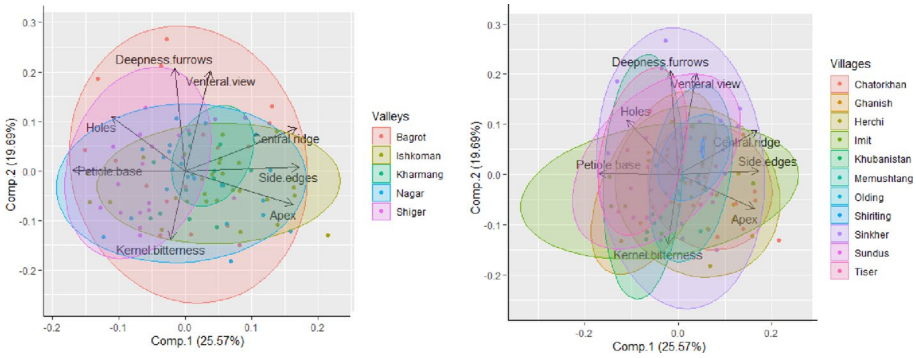


Fig. 6 Principal component analysis (PCA) of apricot stones qualitative parameters in valleys and villages of Gilgit-Baltistan, Pakistan

3.4 Genetic variation of SSRs

Of the 240 leaf samples, 238 were successfully amplified. Using the 12 SSR markers employed (Table 3), a total of 210 alleles were detected, averaging 18 alleles per locus (range: 12–29). Deviation from Hardy-Weinberg equilibrium was observed for all markers, except for UDAp_414, UDAp_420, and pchgms5. Expected (He) and observed heterozygosity (Ho) averaged moderate to high, respectively (Table 3). The average F_{IS} for SSR markers was negative. The probability of identity for P_{ID} was generally lower than for P_{IDsib}; while P_{ID} indicated that 3 marker combinations would be sufficient, P_{IDsib} indicated 6 markers to fully discriminate analysed individuals.

3.5 Genetic variation among regions, valleys, and villages

Region and valley-wise, not only landrace richness, but also genetic diversity parameters were highest for the Gilgit region, in particular for the Ishkoman and Bagrot valley. However, at the village level, the down-valley villages in the Kharmang and Shiger valleys of the Baltistan region exhibited high allelic richness and diversity values, while the valleys of the Gilgit region (Nagar, Ishkoman, and Bagrot) showed an opposite trend (Table 7).

AMOVA revealed high genetic variation within groups (range: 96 to 97%). Overall F-statistics showed a very low level of fixation, although F-values (F_{ST}, F_{IS}, F_{IT}) increased numerically from the regional to the village level (Table 8).

IBD exhibited a significant, moderate positive correlation across villages ($r = 0.65$, $p = 0.01$). The Bayesian clustering obtained by the STRUCTURE software, revealed the presence of substructure with the highest probability of having subpopulations at K = 3 (ΔK: 2.05, LnP(K): – 8502; for visual reasons, we also displayed K = 2), (Figs. 7 and 8). For both, K = 2 and K = 3, a clear separation between valleys still with high rates of admixture was shown. The valleys Bagrot and Nagar were most similar to the Gilgit region, while the Ishkoman valley reflected genetic materials available in the two research stations Jaglot and Skardu (Fig. 8).

3.6 Homonyms and synonyms

Based on morphology and local names given, there were different matching profiles. In total, 26 landraces (10.9%) were assigned as true-to-true type, 60 (25.2%) as homonyms, and 31 (13%) as synonyms. In addition, we had 61 unique (25.6%), while 60 (25.2%) were

Table 7 Average genetic variation and ($F_{IS\text{-inbreeding coefficient } F = 1 - Ho/He}$) in apricot landraces collected from 15 villages within 5 different valleys of two regions of Gilgit-Baltistan, Pakistan during July–Oct 2021 and 2022

Region	Valley	Village	Number of trees (*n*)	Na	Na Freq. ≥ 5%	Ne	I	No. of private alleles	He	uHe	F_{IS}
Baltistan	Kharmang	Memushtang	13	5.33	3.92	3.33	1.32	0	0.66	0.69	−0.144
		Olding	11	5.67	3.75	3.40	1.35	0	0.66	0.69	−0.172
		Shiriting	11	5.50	3.83	**3.62**	**1.40**	0	**0.69**	**0.73**	−0.060
	Shiger	Tiser	10	4.25	4.25	2.99	1.16	0	0.61	0.65	−0.232
		Khubanistan	18	5.25	4.33	3.18	1.27	0.08	0.64	0.66	−0.670
		Sundus	15	5.58	3.92	**3.39**	**1.34**	0.17	**0.67**	**0.69**	−0.125
Gilgit	Nagar	Hisper	13	5.50	4.25	**3.82**	**1.44**	0	**0.72**	**0.75**	−0.150
		Herchi	15	5.42	4.25	3.47	1.37	0	0.70	0.72	−0.125
		Ganish	15	4.83	3.58	2.92	1.22	0.08	0.64	0.66	−0.197
		Imit	12	6.17	4.42	**3.97**	**1.50**	0.08	**0.71**	**0.74**	−0.003
	Ishkoman	Chatorkhan	25	6.92	4.25	3.68	1.44	0.08	0.68	0.69	−0.002
		Slipi	13	6.17	5.08	3.96	1.50	0.08	0.70	0.73	0.080
	Bagrot	Bulchi	10	5.25	5.25	**3.62**	**1.37**	0.17	**0.69**	**0.73**	−0.163
		Sinaker	25	6.08	3.75	3.45	1.35	0.08	0.69	0.70	−0.171
		Oshikhandas	14	4.83	3.67	3.32	1.29	0	0.68	0.71	−0.145
Genebanks		Skardu research	3	3.83	3.83	3.27	1.22	0	0.66	**0.79**	−0.298
		Jaglot_research	17	7.58	5.42	**4.39**	**1.62**	0	**0.73**	0.75	−0.030
	Kharmang		35	7.00	4.33	3.63	1.45	0	0.69	0.70	−0.097
	Shiger		43	6.83	4.25	3.42	1.39	0.33	0.67	0.68	−0.072
	Nagar		43	6.83	3.83	3.52	1.43	0.08	0.70	0.71	−0.128
	Ishkoman		50	8.25	5.08	**4.14**	**1.57**	0.42	**0.71**	**0.72**	0.045
	Bagrot		49	7.17	3.92	3.64	1.43	0.33	0.70	0.71	−0.133
Baltistan			81	9.33	4.50	3.80	1.55	0.33	0.70	0.70	−0.061
Gilgit			159	10.50	4.50	**4.15**	**1.62**	**1.50**	**0.73**	**0.74**	−0.016

Bold marked richness and diversity values are the highest for the particular village

identified as ambiguous as they were neither clear homonyms nor synonyms. The most frequent name of landraces were *Habi*, *Halman*, *Marghulam* and *Shikanda* (Table A4, Fig. 9). The varieties *Gold Rich* and *Pisana*, which were collected as reference material from the two research centres, clustered as true-to-type each, in one clade each that shows correct assignment (Fig. 9).

The Mantel test run for the phenotypic and genetic distance matrices indicated a low correlation ($r = 0.17$, $p = 0.001$) supporting the construction of a tanglegram (Fig. 10). However, the cophenetic correlation itself was significantly positive ($r = 0.44$, $p = 0.001$). The entanglement index indicated a moderate cluster linkage (0.56). We found a total of four pairs, where two were true-to-type (*Bato Shakanda* and *Chusangi*), while two were homonyms (*Meli*, *Yakkar*, *Karfo Chuli*). For both genetic and phenotypic data, the elbow criterion suggested that the most likely number of clusters was K = 2 (Fig. 7).

4 Discussion

Apricot trees and their products are deeply imbedded in Gilgit-Baltistan's agricultural and food system. The rapid transformation of agriculture in G-B, which is increasingly cash crop-driven, puts this plant genetic resource under pressure. To understand the complex relationship between diversity and use, shaped by historical events and current transformation processes, we decided for a combined data approach on landrace nomenclature, dendrology, morphology and genetic data. This allows a holistic assessment of the status of the apricot cultivation and germplasm management system in a highly vulnerable environment. To our knowledge, no detailed scientific studies on landrace richness and diversity within and across regions of northern Pakistan as well as neighbouring countries are available yet.

4.1 Landrace richness and diversity

For most valleys, landrace α-diversity was concentrated in the middle villages. Traditionally, these villages are also the oldest settlements where local rulers (*Mir*s) resided until the middle of the 20th century. It is known that these local kingdoms were recipients and active collectors of rare and exotic goods, including plant material (fruits, seeds, seedlings) of various species [59]. Down-valley villages primarily served defensive functions against rioting and invading armed groups and were therefore less intensively used for agriculture. In contrast, mid-valley settlements were dedicated to the cultivation, accumulation, and propagation of crops, including fruit trees, and were thus effectively protected from theft. Up-valley villages could only be established with the permission of local rulers; they are therefore more recent settlements and are likely to harbour lower levels of diversity as compared to the mid valley villages. High beta-contribution (%β) values indicated that local landraces of apricot were and are vitally exchanged by farmers across villages and valleys, less between regions, although genetic setups indicate common clusters between Baltistan and Gilgit.

4.2 Dendrometry, management, and use

Across the region, no tree management, pruning/trimming, and propagation activities were witnessed, which has multiple causes. For apple (*Malus* × *domestica* Borkh.), Wiehle et al. [58] proposed a lack of facilities and awareness, eroding interest of the youth to work on-farm in combination with low-income options. These factors in combination

Table 8 Analysis of molecular variance (AMOVA) based on SSRs of apricot genotyped in 15 villages of 5 valleys from 2 regions of Gilgit-Baltistan, Pakistan

Grouping	Source	df	Sum of squares	Variance components	Estimated variation	%	F_{ST}	F_{IS}	F_{IT}
Regions	Among regions	1	16.53	16.53	0.06	0.01			
	Among individuals within regions	236	1024.34	4.34	0.00	0			
	Within regions	238	1062.00	4.46	4.46	0.99			
	Total	475	2102.87	0.00	4.52	1.00	0.013*	−0.014*	−0.001
Valleys	Among valleys	6	103.72	17.29	0.20	0.04			
	Among individuals within valleys	231	937.15	4.06	0.00	0			
	Within valleys	238	1062.00	4.46	4.46	0.96			
	Total	475	2102.87	0.00	4.66	1.00	0.045	−0.048	0.000
Villages	Among villages	16	154.72	9.67	0.20	0.04			
	Among individuals within villages	221	886.14	4.01	0.00	0			
	Within villages	238	1062.00	4.46	4.46	0.96			
	Total	475	2102.87	0.00	4.67	1.00	0.046	−0.053	−0.005

* Significant at 0.05

with site-specific environmental differences may contribute to the observed differences in tree sizes across villages. Also, the random selection of root stocks likely influences phenology and vegetative growth, affects edaphic responses, as well as biotic factors such as nematodes susceptibility [37, 38]. Some of the old-aged trees have reached stem diameters > 1 m (own measurements) which was also observed by [3]. The average age category of tree samples was young to moderately young (~ 40 years) with often large tree heights (> 4 m)—indicating that trees are economically unfavourable as productivity is lowered and the modes of harvesting aggravated when compared with mono-aged and small-sized trees of commercial Central European plantations. Farmers showed greater interest in modern cash crop cultivars, favouring traits such as early ripening, large fruit and kernel size, rapid drying, and high yield. In all valleys, farmers frequently cut down non-grafted or old plants and replaced them with modern apricot cultivars or shifted to other species (currently and foremost cherry). Local organisations like the Economic Transformation Initiatives (ETI) and the Governmental Agriculture Department push this process.

Yield data were largely based on estimates as famers did not weigh their harvest. Instead, they typically only provide shares of use such as the portion for home consumption as table and dry fruit or fruit sale at local markets. Nonetheless, yields were higher than those reported for South-East Spain (3 to 8 kg tree^{-1}, Pérez-Pastor et al. [45]), but lower compared with those of commercial Turkish varieties (50 to 60 kg tree^{-1}, Polat and Caliskan [46]). The high variability of tree yields within villages at least partly reflects differences in climatic conditions, alternate bearing habits as well as age and management of trees [52]. Farmers mentioned in general lower yields in 2021 as a likely result of heavy rain and snowfall events during the flowering season which reduced the number of flowers and the availability of pollinators. In addition, low fruit revenues were also linked to post harvest losses due to inappropriate harvesting and handling of fruits as well as reduced shelf life of most landraces. Similar handling and marketing constraints

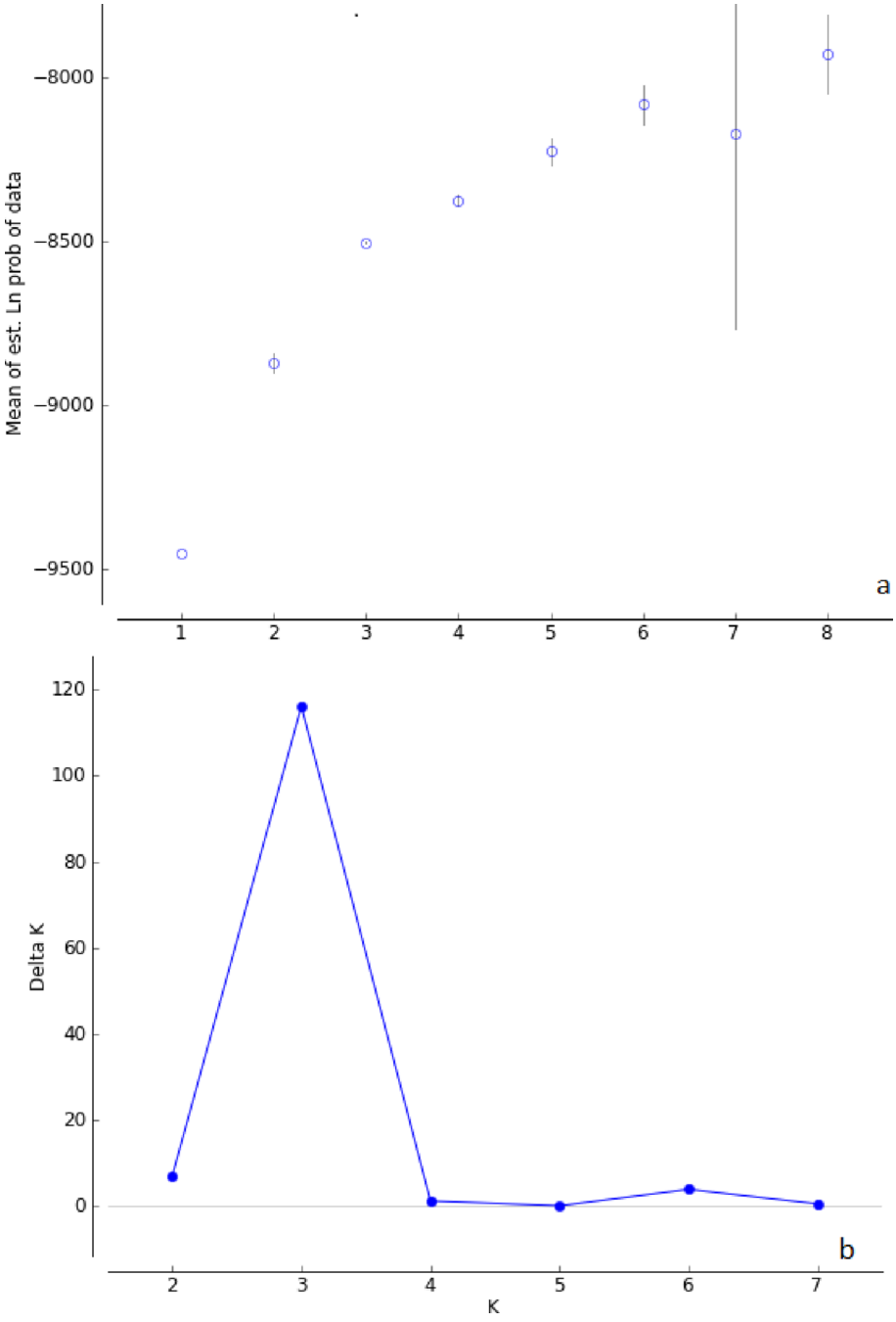


Fig. 7 **a** The Ln(K) ± s.d. for mean log-likelihood and **b** mean (|L″ (K)|)/s.d.(L(K) values for delta K were computed through 10 iterations using the STRUCTURE HARVESTER [11] software to identify the most probable number of groupings among the 238 apricot landraces collected from 15 villages situated in five valleys of the Karakoram Mountains in Gilgit-Baltistan, Northern Pakistan during the period of July to October 2021 and 2022

have been reported from the neighbouring apricot production areas in Kashmir and Quetta [13] which decreased the profitability of the produce.

4.3 Fruit phenotype

Quantitative fruit and stone traits such as weight and volume varied widely, given planting seeds from the best trees over an extended period of time, which resulted in a high

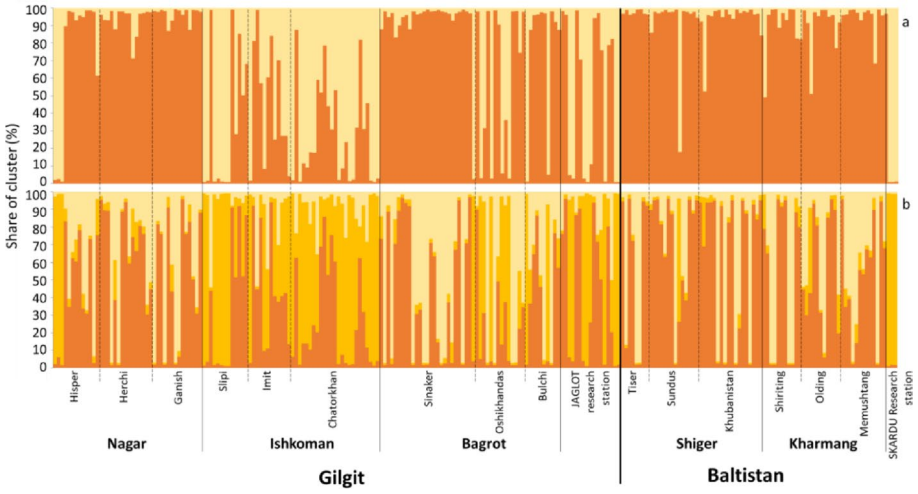


Fig. 8 Bayesian inference was used to determine the most likely number of genetic clusters among local apricot landraces from 15 villages across five valleys of Gilgit-Baltistan, Karakoram Mountains, northern Pakistan, sampled between July and October 2021 and 2022. Stacked STRUCTURE bar plots are shown for K = 2 (**a**) and K = 3 (**b**), with the three colours in the vertical bars representing the proportional probability of each individual genotype belonging to one of the three clusters. Jaglot and Skardu research stations' material was used as reference material

amount of variation [26]. This was also confirmed by a study investigating the local germplasm collection in Jaglot-Gilgit, Pakistan [55]. In general, fruits of grafted trees on cultivated land were larger-sized than fruits of non-grafted trees on uncultivated land, which likely reflects the quasi-intensified management (unintentional trimming through branch removal) and inputs (irrigation water, farmyard manure, chemical fertilizers). However, these observations are strongly influenced by various environmental and sample-related factors, which are difficult to distinguish in the field. Finally, we only considered kernel qualitative traits as they are presumingly the most conservative characters, which was tested and proven with other qualitative traits (shape, taste colour, sweetness, aroma) poorly differentiating among landraces.

The STRUCTURE data allowed to distinguish the western (Baltistan) from the eastern (Gilgit) region, indicating a pronounced separation of populations. However, the relatively heterogenous distribution within valleys and villages also showed a strong admixture. The research stations Jaglot and Skardu seemed also to contain germplasm from the Ishkoman valley confirming this valley as a hotspot of diversity likely caused by historic facts (Text A1).

4.4 Marker selection and quality

Despite the extensive utilization of SSRs in apricot diversity research, we recognize that there is still a significant demand for standardized SSR datasets (including adjusted SSR allele sizes), to conduct comparative analyses across regions and countries as well as to determine diversity in local gene banks. Efforts in this direction for apricot—and for pear (*Pyrus communis* L.)—are, for example, being supported and are ongoing at INRA (www.ecpgr.cgiar.org). Although we intended to use a commonly used set of SSRs through a literature screening, we had to select SSRs that produced amplifiable products. Thus, a simple comparison across studies must be treated with caution.

The high share of markers not being in Hardy-Weinberg-Equilibrium indicated non-random mating and vegetative reproduction, which is likely caused by the still abundant

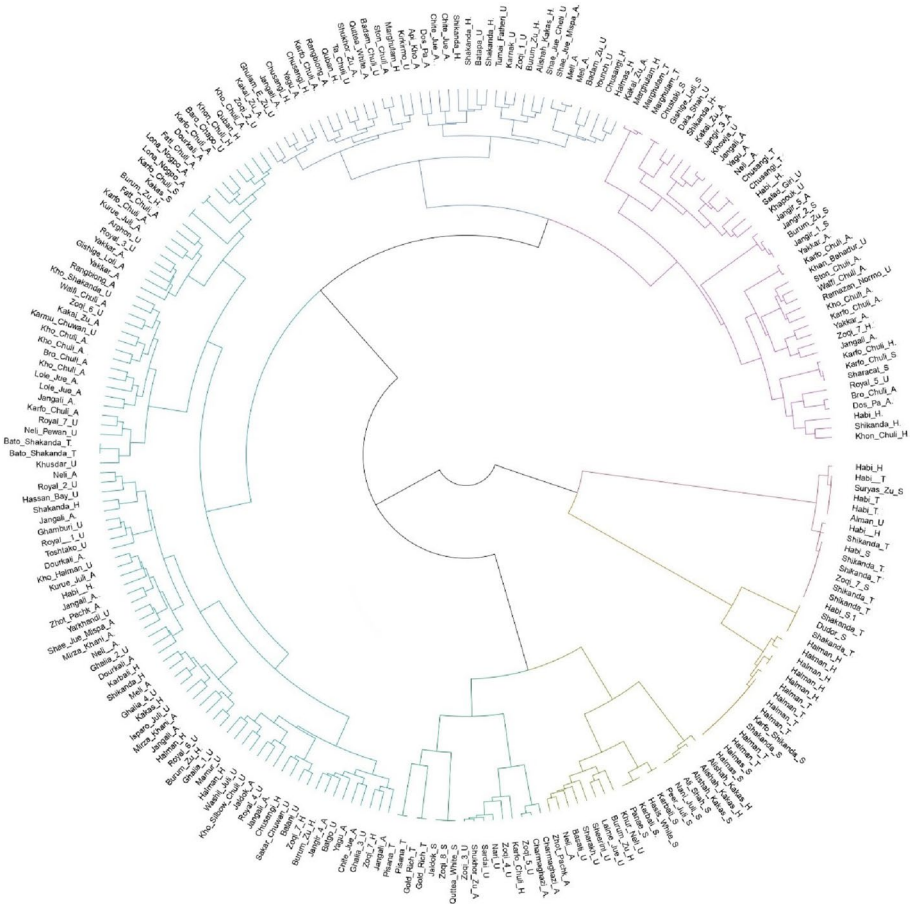


Fig. 9 Genetic analysis of 240 apricot individuals using 12 SSR markers was conducted on samples collected from 15 villages in two valleys of Gilgit-Baltistan, Karakoram Mountains, northern Pakistan, during July and October 2021 and 2022. Landraces were classified as true-to-type (T: same name, same genotype), synonyms (S: different names, same genotype), homonyms (H: same name, different genotype), ambiguous (A: sharing both homonym and synonym characteristics), and unique types (U: different names, different genotype)

grafting activity, as mentioned by 73% of farmers, which may foster sub-structuring of populations (see also Bayesian approach below). According to the obtained P_{ID} and P_{IDsib} values, the selection and decision of 12 SSRs passing the probability threshold of < 0.01 (Waits et al. 2001) after 3 and 6 markers, respectively, indicates a high efficacy of the used SSRs employed in distinguishing individuals.

4.5 Genetic variation in regions, valleys, and villages

The average allele count (n = 18) per locus in our investigation surpassed that of 12 apricot accessions in Gilgit (n = 9, [55]) and 29 wild Turkish varieties (n = 6, [42]). The presence of private alleles and their variation provides further evidence of unique genetic material within the collected individuals among villages and valleys. The substantial count of private alleles identified, and genetic diversity especially within the Gilgit region and among villages in Ishkoman valley and Hisper, is deemed crucial for maintaining the adaptive potentials of populations in response to forthcoming environmental changes [4]. The comparatively high genetic diversity (He = 0.72) identified underscores the importance of the recognized diversity and its interpretation for future breeding and conservation strategies. Further, we like to highlight that we purposely added validated

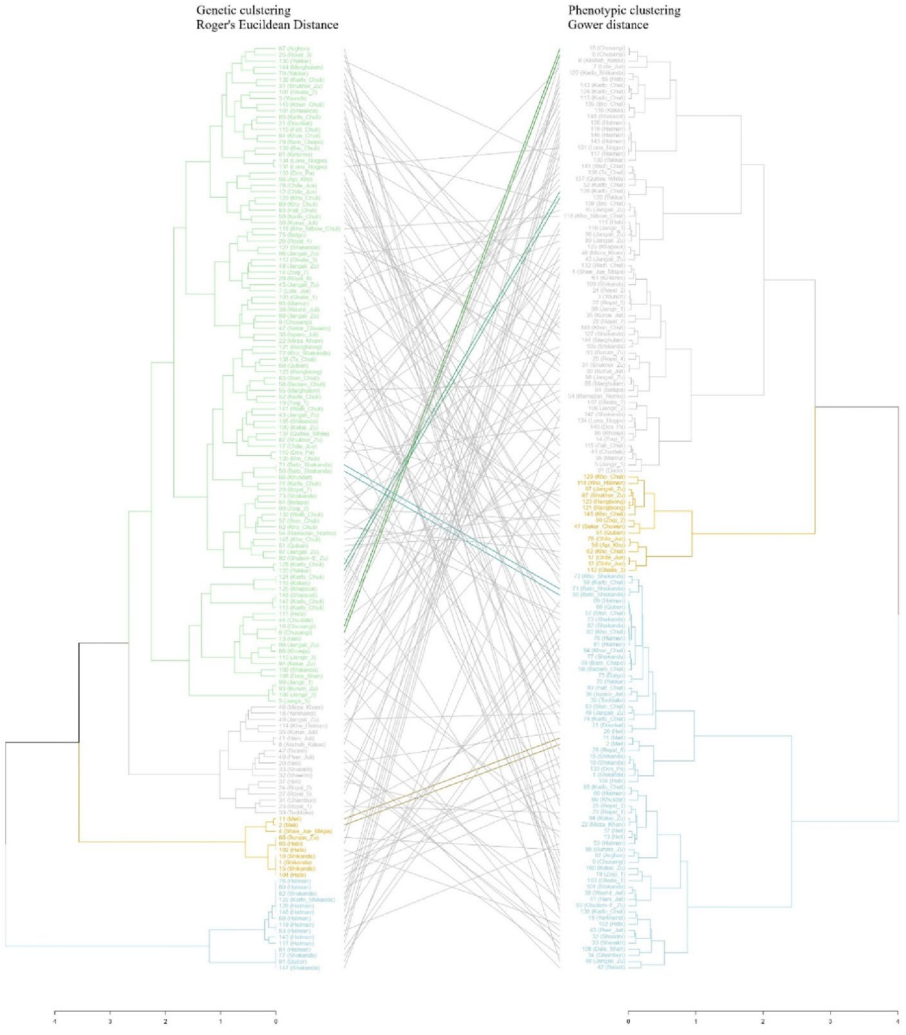


Fig. 10 Tanglegram comparing phenotypic and genetic distance dendrograms of 149 apricot landraces from Gilgit-Baltistan. Each line connects the same genotype in both dendrograms. Coloured lines (often appearing as two nearly parallel lines) indicate closely related genotype pairs that occupy corresponding positions in both dendrograms, suggesting agreement between phenotypic and genetic relationships. Grey lines indicate genotype pairs whose positions do not correspond between dendrograms, reflecting entanglement and a lack of clear association

reference materials from local research stations to verify morphological and genetic characters of well-known cultivars (*Gold Rich* and *Pisana*)—a circumstance that is lacking in many studies. It furthermore proofs the quality of lab analyses that accessions from farmers and research stations clustered together.

4.6 Homonyms and synonyms

More than 75% of apricots names were homonyms or synonyms, compared with 11% true-to-types found by cross-checking given names and genetic setup. This high imbalance shows that farmers either give different landrace names depending on local nomenclatures and language diversity in the region (synonyms) or do not know or remember their correct names. False determinations and assignments are typical and largely prevailing problems for cultigens worldwide such as in apple [58] and almond (*Prunus dulcis* (Mill.) D.A.Webb, [16] and particularly prevalent in multi-linguistic societies

as present in G-B [55]. However, a lack of interest and awareness may also significantly contribute to the unfamiliarity with the names of landraces and cultivars [27]. However, some well-known apricot landraces—such as *Shikanda*, *Halman*, *Shakanda*, *Habi*, *Ali Shah Kakas*, *Marghulam*, *Boto Shakanda*, *Meli*, and *Longpo*—are recognized in both the Balti (Shiger and Kharmang valleys) and Burushaski (Hunza Valley) languages and, consequently, in both regions (Baltistan and Gilgit). As a result, these landraces were more likely to be named correctly, leading to accurate identification as true-to-type.

Although the tanglegram was statistically robust, the alignment between the two cophenetic clusters was weak (Fig. 10), resulting in only a few clades with matching genetic and phenotypic individuals. Although this signals genetically linked traits, the tanglegram rather questions if assessed phenotypic markers have been effective in differentiating the individuals.

5 Conclusions

The two study regions, encompassing five valleys, harbour considerable phenotypic and genetic diversity in apricot, providing a valuable foundation for future breeding and conservation programs. However, the current low level of such activities poses a risk, as local markets are dominated by commercial cultivars, leading to the gradual replacement or loss of indigenous landraces. Addressing this challenge requires responding to emerging consumer demand for niche and value-added products, creating opportunities to promote selected local landraces. These could be further characterized, maintained, and developed into recognized apricot varieties. Although such landraces may have lower yields and limited market presence at present, higher prices for niche cultivars could offer farmers a compensatory benefit. Promoting local landraces and varieties not only allows for the diversification of farmers' produce and income sources but also facilitates the preservation of the historically significant biocultural heritage of Gilgit-Baltistan.

Thus, the targeted marketing of underutilized landraces, often part of old aged orchards could hence diversify farm outputs and foster conservation through use across G-B. To accomplish this, coordinated conservation efforts, coupled with stakeholder-driven research and controlled breeding initiatives, are essential.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s44187-026-00874-4.

Supplementary Material 1.

Acknowledgements

We thank Alexandra Dolynska, Department of Forest Genetics and Forest Tree Breeding, Georg-August-Universität Göttingen, for technical assistance during the laboratory analysis. We are grateful to Prof. Dr. Susanne Neugart, who granted access to her lab for high-throughput screening of seed dimensions. Last, but not least we are indebted to our respondents in Gilgit-Baltistan, our village hosts, farmer cooperative societies representatives and the Mountain Agriculture Research center Joglat and Skardu for sharing the nursery-maintained varieties and facilitating our study. We also thank MNS Agricultural University of Multan and the Government Agriculture Department Gilgit-Baltistan, Pakistan, for logistic and administrative providing support. Furthermore, we want to thank Mareike Köster, Franziska Lehr, Naila Mirbaz, Zulfiqar Ali Shah, and Muhammed Aziz Khan for their support to the research team.

Author contributions

Conceptualization: I.A., A. B.-L., A.B., M.W.; Methodology: I.A., A.B.-L., M.M., A.B., M.W.; Software: I.A., A.N.; Validation: I.A., A. B.-L., O.G., M.M., A.B., M.W.; Formal analysis: I.A., A.N., M.W.; Investigation: I.A., A.N., M.W.; Resources: A. B.-L., O.G., M.M., A.B.; Writing: I.A., A.N., M.W.; Visualization: I.A., A.N., M.W.; Supervision: A.B., M.W.; Project administration: A.B., M.W.; Funding acquisition: A.B., M.W. All authors reviewed the manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL. We thank DAAD for co-financing this research in the frame of the biodiversity+ project (ID 57569302) between MNS University of Agriculture, Multan, and University of Kassel as well as through the ATSAF-CGIAR++- Junior Scientist Program.

Data availability
The SSR genotyping data supporting this study are deposited in the DaKS repository (c) in accordance with FAIR data principles. EVA and other community repositories do not currently support this data type.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 14 August 2025 / Accepted: 6 February 2026
Published online: 08 June 2026

References

1. Abbas A. First report of gummosis disease of major fruits in Gilgit-Baltistan (GB) Pakistan. IJLSSR. 2018;4(3):1805–9. https://doi.org/10.21276/ijlssr.2018.4.3.11.
2. AKRSP. The horticulture study report of the Japan international Cooperation Agency–Gilgit-Baltistan. Aga Khan Rural Support Programme; 2010.
3. Ali S, Masud T, Abbasi KS. Physico-chemical characteristics of apricot (*Prunus armeniaca* L.) grown in Northern areas of Pakistan. Sci Hortic-Amsterdam. 2011;130(2):386–92. https://doi.org/10.1016/j.scienta.2011.05.040.
4. Allendorf FW, Luikart G. Conservation and the genetics of populations. Volume 1stst. Blackwell Publishing Ltd.; 2007.
5. Basile B, Rao R, Corrado G. Genotypic diversity and population structure of the apricot landraces of the campania region (Southern Italy) based on fluorescent SSRs. Genet Resour Crop Evol. 2023;70(1):125–34. https://doi.org/10.1007/s10722-022-01412-8.
6. Bourguiba H, Scotti I, Sauvage C, Zhebentyayeva T, Ledbetter C, Krška B, et al. Genetic structure of a worldwide germplasm collection of *Prunus armeniaca* L. reveals three major diffusion routes for varieties coming from the species' center of origin. Front Plant Sci. 2020. https://doi.org/10.3389/fpls.2020.00638.
7. Cipriani G, Lot G, Huang W-G, Marrazzo MT, Peterlunger E, Testolin R. AC/GT and AG/CT microsatellite repeats in peach (*Prunus persica* (L.) Batsch): isolation, characterisation and cross-species amplification in *Prunus*. Theo App Gene. 1999;99:65–72 https://doi.org/10.1007/s001220051209.
8. Corrado G, Forlani M, Rao R, Basile B. Diversity and relationships among neglected apricot (*Prunus armeniaca* L.) landraces using morphological traits and SSR markers: implications for agro-biodiversity conservation. Plants. 2021. https://doi.org/10.3390/plants10071341.
9. Decroocq Stéphane, Cornille A, Tricon D, Babayeva S, Chague Aurélie, Eyquard J-P, et al. New insights into the history of domesticated and wild apricots and its contribution to *Plum pox* virus resistance. Mol Ecol. 2016;25(19):4712–29. https://doi.org/10.1111/mec.13772.
10. Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci. 2003;14(6):927–30. https://doi.org/10.1111/j.1654-1103.2003.tb02228.x.
11. Earl DA, vonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour. 2012;4(2):359–61. https://doi.org/10.1007/s12686-011-9548-7.
12. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005;14(8):2611–20. https://doi.org/10.1111/j.1365-294X.2005.02553.x.
13. Farooq M, Burfat FM, Abdullah F, Muhammad G, Muhammad AF. The problems faced by fruit growers to market access: a case study of district neelum Azad Jammu and Kashmir. BEPLS. 2016;5(2):1–10.
14. Fatima T, Bashir O, Gani G, Bhat TA, Jan N. Nutritional and health benefits of apricots. Int J Unani Integ Med. 2018;2(2):5–9. https://doi.org/10.33545/2616454X.2018.v2.i2a.25.
15. GoGB. Gilgit-Baltistan agriculture statistics survey Report-2014. Agriculture Statistics Cell Directorate of Agriculture, Government of Gilgit-Baltistan; 2014.
16. Gouta H, Ksia E, Buhner T, Moreno MÁ, Zarrouk M, Mliki A, Gogorcena Y. Assessment of genetic diversity and relatedness among Tunisian almond germplasm using SSR markers. Hereditas. 2010;147(6):283–92. https://doi.org/10.1111/j.1601-5223.2009.02147.x.
17. Groppi A, Liu S, Cornille A, Decroocq Stéphane, Bui QT, Tricon D, et al. Population genomics of apricots unravels domestication history and adaptive events. Nat Comm. 2021. https://doi.org/10.1038/s41467-021-24283-6.
18. Hagen LS, Chaib J, Fady B, Decroocq V, Bouchet JP, Lambert P, et al. Genomic and cDNA microsatellites from apricot (*Prunus armeniaca* L.). Molecular Eco Notes. 2004;4:742–5. https://doi.org/10.1111/j.1471-8286.2004.00802.x.
19. Izhar H, Mubarik A, Aqsa Y. Apricot cluster feasibility and transformation study. (Ed.): Cluster development based agriculture transformation plan Vision-2025. Project No. 131(434)PC/AGR/CDBAT-120/2018. Pakistan: Rawalpindi; 2020. Ali Mubarik.
20. Janick J, Moore JN, editors. Advances fruit breeding. West Lafayette: Purdue University; 1975.

21. JICA. The project for the promotion of value-added fruit products in Gilgit-Baltistan final report. Japan International Cooperation Agency; 2012.
22. Kamvar ZN, Tabima JF, Grünwald NJ. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ. 2014; 2:e281. https://doi.org/10.7717/peerj.281
23. Kassambara A. Practical guide to principal component methods in R. 1st. Scotts Valley: CreateSpace Independent Publishing Platform; 2017.
24. Kaufman L, Rousseeuw PJ. Finding groups in data. Wiley; 1990.
25. Khan B, Abdukadir A, Qureshi R, Bot G, Mustafa H. Medicinal uses of plants by the inhabitants of Khunjerab National Park, Gilgit, Pakistan. Pak J Bot. 2011; 43(5): 2301–10.
26. Khan M, Maghuly F, Borroto-Fernandez EG, Pedryc A, Katinger H, Laimer M. Genetic diversity and population structure of apricot (*Prunus armeniaca* L.) from Northern Pakistan using simple sequence repeats. Silvae Genet. 2008;57(3):157–64. https://doi.org/10.1515/SG-2008-0024.
27. Köster M, Alam I, Rana J, Wiehle M, Buerkert A. A stony track towards innovation in remote Highland regions: agricultural intensification in the apricot sector of Northern Pakistan. Agric Food Secur. 2024. https://doi.org/10.1186/s40066-024-00475-3.
28. Kousar R, Makhdum MS, Abbas A, Nasir J, Naseer MA. Issues and impacts of the apricot value chain on the upland farmers in the Himalayan range of Pakistan. Sustain (Switzerland). 2019;11(16):1–13. https://doi.org/10.3390/su11164482.
29. Kreutzmann H. Linguistic diversity in space and time: a survey in the Eastern Hindukush and Karakoram. H Ling. 2014. https://doi.org/10.5070/H94022918.
30. Krichen L, Audergon JM, Trifi-Farah N. Assessing the genetic diversity and population structure of Tunisian apricot germplasm. Sci Hortic-Amsterdam. 2014;172:86–100. https://doi.org/10.1016/J.SCIENTA.2014.03.038.
31. Kubisiak TL, Nelson CD, Staton ME, Zhebentyayeva T, Smith C, Olukolu BA, et al. A transcriptome-based genetic map of Chinese chestnut (*Castanea mollissima*) and identification of regions of segmental homology with Peach (*Prunus persica*). Tree Gen Genomes. 2013;9(2):557–71. https://doi.org/10.1007/s11295-012-0579-3.
32. Legendre P, Legendre L. Numerical Ecology. Third english edition. Amsterdam: Elsevier (Developments; 2012. in environmental modelling, 24).
33. Lehr F. The status of female empowerment in the apricot sector in Gilgit-Baltistan, Pakistan. Universität Göttingen and Kassel; 2024. https://doi.org/10.17170/kobra-2024090610800
34. Li Y, Ma Y, Zhang T, Bi Y, Wang Y, Prusky D. Exogenous polyamines enhance resistance to *Alternaria alternata* by modulating redox homeostasis in apricot fruit. Food Chem. 2019; 301:125303. https://doi.org/10.1016/j.foodchem.2019.125303
35. Lopes MS, Sefc KM, Laimer M, Da Cââmara Machado A. Identificationof microsatellite loci in apricot (*Prunus armeniaca* L.). Molecular Eco Notes. 2002;2:24–26
36. Martín C, Herrero María, Hormaza Josél. Molecular characterization of apricot germplasm from an old stone collection. PLoS ONE. 2011;. https://doi.org/10.1371/journal.pone.0023979.
37. Mendelné Pászti E, Bujdosó Géza, Mendel Ákos. Vegetative characteristics of three apricot cultivars grafted on six different rootstocks. Horticulturae. 2022;8(11):1004. https://doi.org/10.3390/horticulturae8111004.
38. Mir MM, Iqbal U, Mir SA. Production technology of stone fruits. Springer Singapore (Production Technology of Stone Fruits); 2021.
39. Murtagh F, Legendre P. Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? J Classif. 2014;31(3):274–95. https://doi.org/10.1007/s00357-014-9161-z.
40. Oksanen J. Vegan: ecological diversity; 2022.
41. Özen M, Gül Mevlüt. Marketing structure of apricot production and analysis of its problems: a case of mut district in Mersin Province. Int J Agric Life Sci. 2020;4(1):79–86. http://dergipark.gov.tr/ijafls.
42. Ozrenk K, Ozkan G, Akin M, Orhan E, Eyduran SP, Ercisli S. Molecular characterization and genetic relationships of wild apricots (*Prunus armeniaca* L.) from Turkey using SSR markers. CR Acad Bulg Sci. 2020;73(5):641–50. https://doi.org/10.7546/CRABS.2020.05.06.
43. Peace CP. DNA-informed breeding of rosaceous crops: promises, progress and prospects. Hort Res. 2017;4:17006. https://doi.org/10.1038/hortres.2017.6.
44. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinf (Oxford England). 2012;28(19):2537–9. https://doi.org/10.1093/bioinformatics/bts460.
45. Pérez-Pastor A, Ruiz-Sánchez MªC, Domingo R. Effects of timing and intensity of deficit irrigation on vegetative and fruit growth of apricot trees. Agr Water Manage. 2014;134:110–8. https://doi.org/10.1016/j.agwat.2013.12.007.
46. Polat AA, Caliskan O. Yield and fruit characteristics of various apricot cultivars under subtropical climate conditions of the Mediterranean region in Turkey. Int J Agr. 2013. https://doi.org/10.1155/2013/687345
47. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155(2):945–59. https://doi.org/10.1093/genetics/155.2.945.
48. Ricotta C. Of beta diversity, variance, evenness, and dissimilarity. Ecol Evol. 2017;7(13):4835–43. https://doi.org/10.1002/ece3.2980.
49. Schuelke M. An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol. 2000;18(2):233–4. https://doi.org/10.1038/72708.
50. Sendall A, Mir M, Khabir A. The agribusiness project (TAP) apricot value chain assessment final report for the agribusiness project; 2013. www.cnfa.org
51. Sosinski B, Gannavarapu M, Hager LD, Beck LE, King GJ, Ryder CD, et al. Characterization of microsatellite markers in peach (*Prunus persica*(L.) Batsch). Theo App Gene.2000;101:421–8.https://doi.org/10.1007/s001220051499.
52. Spies M. Northern Pakistan. High mountain farming and socionatures. Vanguard Books; 2019.
53. Thompson M. Bio-geographical survey and collection of temperate fruit and nut genetic resources in Northern Pakistan; 1988. https://scholar.google.de/scholar?hl=en&as_sdt=0%2C5&q=Thompson%2C+M.%2C+%281988%29.+Bio-geographical+survey+and+collection+of+temperate+fruit+and+nut+genetic+resources+in+Northern+Pakistan.+Trip+Report%2C+International+Board+for+Plant+Genetic+Resources%2C+IBPGR%2F88%2F55%2C+pp.+105.&btnG=#d=gs_cit&u=%2Fscholar%3Fq%3Dinfo%3ABE5pFxeZaf8J%3Ascholar.google.com%2F%26output%3Dcite%26scirp%3D0%26hl%3Den
54. Torrecillas Melendreras A, Galego R, Pérez Pastor A, Ruiz Sánchez MaríaC. Gas exchange and water relations of young apricot plants under drought conditions. 1999; 132(4):445–52. https://doi.org/10.1017/S0021859699006577

55. Ullah S, Muhammad A, Hussian I, Hyder H-U-R, Muhammad Zeeshan. Geneticalysis of economically important apricot cultivars in Gilgit Baltistan based on SSR molecular markers. Rom Biotech Lett. 2017;22(2):12456–63.
56. Wang X, Liu S, Zuo H, Zheng W, Zhang S, Huang Y, et al. Genomic basis of high-altitude adaptation in Tibetan *Prunus* fruit trees. Curr Biol. 2021;31(17):3848–e38608. https://doi.org/10.1016/j.cub.2021.06.062.
57. Wani AA, Zargar SA, Malik AH, Kashtwari M, Nazir M, Khuroo AA et al. Assessment of variability in morphological characters of apricot germplasm of Kashmir, India. Sci Hortic-Amsterdam. 2017; 225:630–7. https://doi.org/10.1016/j.scienta.2017.07.029
58. Wiehle M, Nawaz MA, Dahlem R, Alam I, Khan AA, Gailing O, et al. Pheno-genetic studies of apple varieties in Northern Pakistan: a hidden pool of diversity. Sci Hortic-Amsterdam. 2021. https://doi.org/10.1016/j.scienta.2021.109950.
59. Whiteman PT. Mountain oases. A technical report of agricultural studies (1982–1984) in Gilgit District, Northern Areas, Pakistan; 1985.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.