RESEARCH ARTICLE

Population genetic structure of a timberline oak (*Quercus semecarpifolia* Sm.) of western Himalayas and conservation implications

Harish S. Ginwal[1] · Anita Rawat[1] · Chander Shekhar[1] · Maneesh S. Bhandari[1] · Haripriya Kavidayal[1] · Rajeev Shankhwar[1] · Abhishek Yadav[1] · Rama Kant[1] · Santan Barthwal[1] · Rajendra K. Meena[1]

Abstract

Forests are considered as one of the most complex terrestrial ecosystems due to their high level of biodiversity, including genetic diversity. Understanding the genetic diversity of keystone species at a population level is vital to forest managers and policymakers for the conservation and sustainable utilization of forest genetic resources. *Quercus semecarpifolia*, commonly known as brown oak, is a keystone species of climax community thriving in the alpine zone of the Himalayas, which is presently experiencing population decline and range shift under the changing climate. In the present study, a landscape genetic approach was employed for deciphering the population genetic structure of *Q. semecarpifolia* in the western Himalayas using nuclear simple sequence repeat (SSR) markers. By analysing 718 individuals of 24 populations at 10 SSR loci, a high gene diversity (expected heterozygosity, He = 0.72; Allelic richness, Ar = 8.37) was recorded with a moderate genetic differentiation (F_{ST} = 0.16; P < 0.001). Genetic clustering and STRUCTURE analysis have displayed two major gene pools which appear to be primarily differentiated by the landscape and ecological constraints rather than the linear geographical distances. The hierarchical AMOVA analysis further supports the regional genetic divergence with a substantial proportion of genetic variation detected among the regions. Diversity maps generated by spatial interpolation elucidated the distribution pattern of genetic diversity and structure across the range, and aid in the demarcation of the diversity hotspots for conservation implications. To the best of our knowledge, this is the first comprehensive genetic study carried out in any Himalayan oaks, and the information generated herein is novel and of paramount importance in guiding conservation decisions.

Introduction

Forests are considered one of the most complex terrestrial ecosystems due to the high biodiversity in terms of genetic resources, species and habitat ecology (Geburek and Konrad 2008). However, the last century witnessed the ruination and vanishing losses of wild forests and biodiversity.

Overall, biodiversity is measured at three levels, namely genetic diversity, species diversity, and ecological diversity. Intra-specific genetic diversity is crucial for the long-standing survival of a species in an ecosystem (Gapare 2014). Adaptation, evolution and survival of a species in the long-term depend mainly on the subsistence of adequate genetic variability both within and among the populations, to accommodate new selection pressures brought by the ecological changes and demographic load (Reed and Frankham 2003). Considering the importance of genetic diversity, Convention on Biological Diversity has recognized it as a crucial component of biodiversity and included it among the 23 global targets for 2030 in the 15th Conference of Parties to the UN Convention on Biological Diversity (CBD 2022). Hence, investigating the genetic diversity of keystone

✉ Harish S. Ginwal
  ginwalhs@icfre.org

✉ Rajendra K. Meena
  rajnrcpb@gmail.com

1 Division of Genetics and Tree Improvement, ICFRE-Forest Research Institute, Dehradun, Uttarakhand 248195, India

species can be of immense importance in guiding conservation programmes and reducing the risk of biodiversity loss (Souto et al. 2015).

Oaks (*Quercus* L.) are keystone species for many ecosystems, conferring a wide range of ecological functions and socio-economic values. For instance, oaks are considered by many societies as sacred trees, symbols of strength and endurance, with high cultural and historical value. In addition, they cater important ecological services like carbon sequestration, a reservoir of biodiversity, soil and water protection, etc., and other services of economic and cultural importance (QUERCUS PORTAL; https://quercusportal.pierroton.inra.fr/). Worldwide, a total of 450 *Quercus* species have been reported, across a wide range of habitats, including temperate deciduous forests, temperate and subtropical evergreen forests, subtropical and tropical savannah, subtropical woodland, oak-pine forest, etc. (Nixon 2006; https://powo.science.kew.org/). The highest species diversity of oak is recorded in Central America and South Asia. Over 35 *Quercus* spp. have been reported to occur in the Indian Himalayan Region (Negi and Naithani 1995) at an altitudinal range between 2200 and 3900 m. The Indian Himalayan Region is a global biodiversity hotspot found within the Himalayas with exceptionally rich biodiversity and endemism (Negi et al. 2019). Based on altitudinal gradient, six climatic zones were recognized in this region, such as warm temperate (900–1800 m), cool temperate (1800–2400 m), cold zones (2400–3000 m), alpine zone (3000–4000 m), glacier zone (4000–4800 m) and perpetually frozen zone (above 4800 m) (Uttarakhand Forest Statistics 2014–2015). The region is also regarded as the 'land of the goddess' due to the presence of many monumental temples, and the state provides hospitality to about 36.9 million pilgrims or tourists annually (Uttarakhand Tourism Statistics 2018). Most of these religious places are in the mountains covered with beautiful, lush green forests, and pose a lot of anthropogenic pressure over the forest vegetation. Ecologically and geologically, Uttarakhand Himalayas are highly fragile and have witnessed several natural calamities in the past like cloud bursts, landslides, erratic rainfall, flood, etc., which have severely affected the population structure of forestry species. This prompts us to generate baseline reference data for ecologically important keystone species to quantify the impact of such events in future. The temperate broad-leaved forests of the western Himalayas are mostly dominated by oaks, namely *Quercus floribunda*, *Q. glauca*, *Q. lanata*, *Q. leucotrichophora* and *Q. semecarpifolia*,

Quercus semecarpifolia, commonly known as 'brown oak' or 'kharsu', is one of the late successional evergreen broad-leaved tree species of climax community occupying highest altitudinal range up to 3700 m in the western Himalayas (Shekhar et al. 2022). Beside the Indian Himalayan Region, it has also been reported in Afghanistan, Bhutan, China, Myanmar, Nepal and Pakistan at the altitude ranging from ~2500 to ~4000 m (Oaks of the world database; http://oaks.of.the.world.free.fr/quercus_semecarpifolia.htm). Despite the dominating distribution, it has been observed to be associated with several broad leaved and conifer tree species of Himalayas, viz., *Abies* spp., *Acer* spp., *Betula utilis*, *Cupressus torulosa*, *Fraxinus micrantha*, *Picea* spp., *Pinus wallichiana*, *Rhododendron* spp., *Taxus wallichiana*, etc. (Shekhar et al. 2022). *Quercus semecarpifolia* also serves as an excellent source of quality fodder, fuel-wood, timber, wood and tannin, which has led its over-extraction from natural forest (Singh et al. 2010; Shrestha 2003). Along with the anthropogenic pressure, inherently slow growth rate and inadequate regeneration are other factors causing population deterioration in the Himalayas (Vetaas 2000; Bisht 2001; Shrestha 2003). Reduced regeneration in *Q. semecarpifolia* has been associated with several factors, viz., high moisture content and short viability of seeds (vivipary), poor seed crop every year, desiccation and frost sensitivity of seeds, edibility of seeds by wild animals, etc. (Singh et al. 2011; Bisht 2001; Tashi 2004). The ecological cost of oak forest degradation is perhaps more important and damage is irreversible. Further, it has been reported to be vulnerable to climate change showing an altitudinal shift or upslope movement in response to the future climate change scenario (Bisht et al. 2013; Shekhar et al. 2022).

Noteworthily, the life-history traits of species and interaction of environmental variables may disrupt the habitat connectivity, and influence the genetic structure spatially as well as temporally (Gómez et al. 2005; Yang et al. 2018). Spatial overlaying of the genetic diversity over landscape elements could provide significant cartographic outputs for guiding conservation and management of forest genetic resources. Consequently, this present study was undertaken to understand the spatial genetic structure and gene diversity in *Q. semecarpifolia* populations of western Himalayas to inform future conservation efforts. Specifically, we attempted to understand the key ecological and biological questions of conservation importance, such as what level of genetic diversity exists in the populations in situ? How the individual populations are genetically differentiated and structured? What is the spatial pattern of genetic admixture and allelic diversity? How is the genetic variation segregated within and among the populations? What could be the possible causes associated with interpopulation genetic divergence? How the geographic information systems (GIS)-based tools could be employed in guiding the conservation programme? To unscramble these questions, we used a multidisciplinary approach combining SSR-based population genetic analysis with GIS-based tools to decipher the spatial genetic structure of *Q. semecarpifolia* in relevance to

the landscape features. With the best of our knowledge, this is a pioneer study on population genetics of brown oak of Indian Himalayas. Hence, the information generated herein is novel and immensely useful in guiding conservation and management plans of Himalayan oak forest.

Materials and methods

Study area and sample collection

Present study was carried out in the western Himalayas under Uttarakhand state (India) which is geographically divided into two regions, Garhwal and Kumaon. The entire distribution range of *Q. semecarpifolia* in both the regions was intensively surveyed from 2016 to 2020. Leaf samples were collected from 718 individual trees belonging to 24 natural populations (Table 1). Each population was represented by 28 to 30 individuals by keeping 100 to 300 m distance between them. The sampling was carried out randomly with linear transect method and each sampled individual was tagged with their geo-coordinates using global positioning system (GPS). To avoid sample degradation and fungal development, leaf samples were desiccated using silica gel and completely dried samples were stored at -80 ˚C.

Genomic DNA extraction and SSR genotyping

Genomic DNA was extracted using double DNA extraction protocol standardized by modifying protocol of Doyle and Doyle (1987). In brief, 0.5 g ground leaf tissue was incubated with 1 mL pre-chilled lysis buffer containing tris base (100 mM), EDTA (20 mM), NaCl (1.42 M), ascorbic acid (5 mM), PVP (3% w/V) and β-mercaptoethanol (5 µL) at 4 °C for 30 min. After centrifugation, the homogenate was again incubated with 1 mL lysis buffer containing 3% CTAB at 60 °C for 60 min. The samples were emulsified with chloroform:isoamyl alcohol (24:1), and the cell debris was separated out by centrifugation. Further, genomic DNA present in upper aqueous phase was precipitated by overnight incubation with chilled isopropanol, followed by 96% ethanol containing 3 M sodium acetate. After washing with 70% ethanol, dried DNA pellet was re-suspended in 100 µL TE buffer. Finally, genomic DNA was quantified and diluted to prepare a final working concentration of 10 ng µL^{-1} for use in polymerase chain reaction (PCR).

Polymorphic SSRs identified through cross-transferability from other closely related *Quercus* spp. (Mishima et al. 2006; Ueno et al. 2008; Ueno and Tsumura 2008) were used in PCR-based genotyping as per Shekhar et al. (2021). A PCR amplicon of each selected primer pair was sequenced using Sanger's dideoxy method and verified for the presence of desired repeat motifs. Ten polymorphic SSRs were selected

Table 1 Geographical details of sampled populations of *Q. semecarpifolia*

Pop code	Geographic locations	Districts	Forest divisions	No. of samples	Latitude (N)	Longitude (E)	Altitude (m)
QS01	Kanchula Kharak	Chamoli	Kedarnath Wildlife Division	30	30˚27'11.50"	79˚14'29.90"	2577
QS02	Chopta	Chamoli	Kedarnath Wildlife Division	30	30˚28'51.90"	79˚11'52.30"	2937
QS03	Deoban	Dehradun	Chakrata	30	30˚44'52.40"	77˚51'58.30"	2818
QS04	Bhujkoti	Dehradun	Chakrata	30	30˚46'11.40"	77˚54'31.20"	2828
QS05	Lokhandi	Dehradun	Chakrata	30	30˚45'30.40"	77˚49'57.80"	2750
QS06	Rudranath	Chamoli	Kedarnath Wildlife Division	30	30˚29'34.00"	79˚18'40.10"	3135
QS07	Auli	Chamoli	Nanda Devi National Park	28	30˚31'26.30"	79˚33'53.80"	2944
QS08	Yamunotri	Uttarkashi	Upper Yamuna Barkot	30	30˚59'42.60"	78˚27'36.40"	2942
QS09	Radi Top	Uttarkashi	Upper Yamuna Barkot	30	30˚45'20.20"	78˚12'32.40"	2589
QS10	Chaurangi Khal	Uttarkashi	Uttarkashi	30	30˚38'53.40"	78˚30'17.60"	2800
QS11	Bhukkitop	Uttarkashi	Uttarkashi	30	30˚50'28.40"	78˚39'37.20"	2689
QS12	Dudhatoli	Pauri	Garhwal	30	30˚03'07.12"	79˚06'05.65"	2315
QS13	Naina Peak	Nainital	Nainital	30	29˚24'25.38"	79˚26'05.08"	2500
QS14	Badhanital Forest	Rudraprayag	Rudraprayag	30	30˚29'49.08"	78˚56'64.01"	2470
QS15	Munsiyari	Pithoragarh	Pithoragarh	30	30˚03'50.07"	80˚12'37.06"	3140
QS16	Chorikhal	Pauri	Garhwal	30	30˚02'18.12"	78˚59'42.33"	2582
QS17	Ghes	Chamoli	Badrinath	30	30˚08'27.14"	79˚43'18.54"	2393
QS18	Balcha	Dehradun	Tons	30	30˚93'33.00"	77˚07'52.80"	2806
QS19	Mundhola	Dehradun	Chakrata	30	31˚05'00.00"	77˚09'50.40"	2735
QS20	Nag Tibba	Dehradun	Mussoorie	30	30˚35'08.20"	78˚08'29.49"	2639
QS21	Pinswar	Tehri	Kedarnath Wildlife Division	30	30˚35'08.20"	78˚67'26.70"	2829
QS22	Kunjkharak	Nainital	Nainital	30	29˚29'57.71"	79˚19'54.31"	2393
QS23	Narayan Ashram	Pithoragarh	Pithoragarh	30	29˚58'43.40"	80˚39'18.40"	2879
QS24	Himkola	Pithoragarh	Pithoragarh	30	30˚01'53.70"	80˚38'51.30"	3088

for further genotyping in *Q. semecarpifolia* (Table 2). The PCR reactions were performed in thermal cycler machine (Mater cycler gradient Nexus; Eppendorf) in a 15 µL reaction volume containing 15 ng template DNA, 1× *Taq* buffer, 2.4 mM MgCl$_2$, 0.2 mM dNTPs, 0.2 µM each of forward and reverse primers, 0.65 U *Taq* DNA polymerase and nuclease free sterile water. The PCR was run with the following cycling parameters: first a denaturation step at 94 °C for 3 min, followed by 40 cycles of 94 °C for 45 s, 48–55 °C for 45 s, 72 °C for 45 s and a final extension at 72 °C for 7 min. Afterwards, fragment length polymorphism was analysed using automated capillary electrophoresis in LabChip GX Touch 24 Nucleic Acid Analyzer (Perkin Elmer, USA) along with an internal size standard, and the allelic data was extracted through Gene Reviewer software (Perkin Elmer). The allelic data showing deviations from the expected periodicity of the repeats were adjusted through allele binning software Tandem v1.07 (Matschiner and Salzburger 2009). Also, the marker data were analysed for the presence of null alleles and large allele dropout using software MICRO-CHECKER v2.2 (Van Oosterhout et al. 2004) with 1000 randomizations and a 95% confidence interval. The SSR loci with null allele frequency ranging from 0.1 to 0.3 in more than 30% of the populations were considered for further analysis as per Bagnoli et al. (2009). Deviation from Hardy-Weinberg equilibrium (HWE) was estimated by χ^2 test using software Arlequin v3.1 (Schneider et al. 2000).

Genetic diversity and differentiation

Diversity measures, such as number of alleles per locus (Na), effective number of alleles (Ne), Shannon's information index (I), observed heterozygosity (Ho), expected heterozygosity (He) and coefficient of inbreeding (F_{IS}) were calculated using the program GenAlEx v6.5 (Peakall and Smouse 2012). The number of alleles in a population is an imperative measure of genetic variation but can be difficult to use in comparisons if sample size varies. Therefore, we aimed to maintain a uniform sample size across the sampled populations. Furthermore, allelic diversity was calculated as allelic richness (Ar) and private allelic richness (PAr) using a rarefaction method implemented in software HP-Rare v1.0 (Kalinowski 2005).

Genetic relationship was studied by constructing dendrogram based on pairwise genetic distances (D_A; Nei et al. 1983) using UPGMA (unweighted pair group method with arithmetic mean) algorithm implemented in software POP-TREE v2 (Takezaki et al. 2010). The robustness of dendrogram topologies was further tested by bootstrap resampling (n = 1000). Analysis of molecular variance (AMOVA), Wright's F statistical measures of genetic differentiation, and gene flow, were calculated using software GenAlEx with 1000 permutations. The covariance matrix of pairwise F_{ST} was further used in multivariate principal coordinate analysis (PCoA) and the Mantel test (Mantel 1967). Isolation by distance was tested through the Mantel test using software GenAlEx in which pairwise genetic distances between the population were correlated with pairwise geographic?al distances (vertical and horizontal). Both the analyses were

Table 2 Details of SSR primer pairs validated and used in *Q. semecarpifolia*

Sl. No.	Locus name	Repeat motif	Primer sequence (5' – 3')	Ta	Product size
1.	CR627959[a]	(GCA)$_4$	F: GCTCCCTGGTAGTCGGCTAAAGA R: CAATTGGGACAACATGGAAGCAT	55	228–288
2.	DN950717[a]	(GGA)$_3$	F: TAGTTTTCCCAGCGAATCCAACA R: CTTCTTGAAGGGACTGACCCCAT	55	215–275
3.	DN950726[a]	(GAT)$_9$	F: GCAAGAAGCATGCAGATGGAGAT R: GCATGGCCGTCATTAGCATTAAG	55	324–393
4.	bcqm76[b]	(GT)$_8$	F: ATTAGTTTGCCTAGCTCTACCATG R: AGTCGCAGTCCCAATAGTAAAG	57	139–259
5.	QmC00419[c]	(AG)$_3$	F: TTTTTGGTGTTTCGGTGGTTTGG R: AAGGATGAATTGAGGGTTTCGGA	55	151–267
6.	QmC00716[c]	(TC)$_{11}$	F: AAGAGAACCCATTCCATCCCTGA R: TTTCCCGAACAGTGGTTTCTTGA	55	158–326
7.	QmC00898[c]	(AG)$_6$	F: AGGCAGCCGTAAGAGTTGAGCTT R: CTTTCTTCCTTTGCTGCCCTTGT	55	78–182
8.	QmC01368[c]	(CCA)$_3$	F: TAGCAGAATCATCGGAAGCGAAC R: CCGAGCTTGATGAACTGCGTACT	55	137–194
9.	QmC02269[b]	(TC)$_{10}$	F: CGGGTTCTAGCTGCTGTTGGTTT R: GATTATGGGTGCGTTCATGGTGT	55	113–257
10.	bcqm96[b]	C$_{10}$	F: CATATCTCAATTTCTGTGTTCTTAGTAG R: TATGTTAATGTGTTGGCAATCC	57	63–199

[a] Ueno and Tsumura (2008); [b] Mishima et al. (2006); [c] Ueno et al. (2008)

performed in all the 24 populations together as well as in different groups independently.

Genetic structure

The population genetic structure was deciphered using Bayesian model-based clustering approach implemented in the program STRUCTURE v2.3.4 (Pritchard et al. 2000). The program was run with ancestry model with admixture under assumption of correlated allele frequencies. The simulations were run with 10 iterations for each K value (in our case 1–10) with 500,000 Markov Chain Monte Carlo (MCMC) sampling runs after a burn-in period of 500,000 iterations. The optimal K value was determined using web-based program StructureSelector (Li and Liu 2018) which employs method developed by Evanno et al. (2005) as well as other alternative methods (MedMedK, MedMeanK, MaxMedK, and MaxMeanK) developed by Puechmaille (2016). The resultant data of replicated STRUCTURE runs were further collated into a matrix (the Q-matrix) of membership coefficients using CLUMPP v1.1.2 (Jakobsson and Rosenberg 2007) and graphically displayed as bar plot using DISTRUCT v1.1 (Rosenberg 2004). Owing to strong genetic structure, each major cluster was further investigated individually for the nested structuring by repeating the PCoA and STRUCTURE analyses.

Spatial mapping of allelic diversity and genetic disjuncture

Genetic and geo-spatial data were organized in a geo-database of ESRI ArcGIS v9.3 (ESRI, Redlands, CA, USA). Particularly, allelic diversity (Ar and PAr) was spatially overlaid over the species distribution map using inverse distance weighted (IDW) interpolation function implemented in ArcGIS (Shepard 1968; Hengl 2009; Chiocchini et al. 2016). For overlaying, the current eco-distribution map of *Q. semecarpifolia* generated through maximum-entropy (MaxEnt) approach of modelling by Shekhar et al. (2022) was used as a surface map. The IDW interpolation determines cell values using a linearly weighted combination of a set of sample points, considering that the local influence of each single point decreases with a distance. Similarly, inferred ancestry of each population calculated by Bayesian analysis was used to draw pie-charts and spatially overlayed over distribution maps manually.

Results

Allele frequencies and gene diversity

Cross-amplified SSRs were first verified for the presence of expected repeat motifs by sequencing their PCR amplicons, and polymorphic SSRs with desired repeat motifs were used for genotyping. In total, 332 alleles were generated in 718 individuals of 24 populations across the 10 SSR loci. Each SSR was found to be highly polymorphic, exhibiting on average 10 alleles per locus ranging from 4 (DN950717) to 18 (QmC02269) alleles. Based on the frequency distribution, 60 alleles were considered abundant with frequency >0.05, 187 were rare with frequency <0.05, and 85 were unique to one of the 23 populations. Among the analyzed populations, the number of different alleles and effective number of alleles varied from 8 (Auli) to 13 (Munsiyari) and 4 (Yamunotri) to 7 (Bhukkitop), respectively.

Our experimental populations demonstrated a high allelic diversity with a mean allelic richness of 8.37 ranging from 6.71 (Auli) to 9.77 (Munsiyari). Based on the private alleles, populations from Munsiyari (QS15) and Nagtibba (QS20) showed the highest genetic distinctness, having the maximum number of private alleles. Overlaying allelic diversity (Ar) to the MaxEnt-derived distribution map enabled to demarcate the populations or the regions with maximum diversity (Fig. 1). Likewise, overlaying of private alleles highlighted the populations with unique genetic constitution (Fig. S1). Coincidentally, the populations demonstrating high allelic richness at Bhukkitop (QS11), Munsiyari (QS15), Mundhola (QS19), and Nag Tibba (QS20), contained a significant proportion of private alleles, and hence, these could be considered at top priority in the conservation programme. Perversely, the populations of Kedarnath Wildlife Sanctuary (QS02 and QS06), Chakrata forest division (QS03, QS04, QS05), Nanda Devi National Park (QS07) have displayed low allelic diversity. Other key diversity measures, such as observed heterozygosity (Ho), expected heterozygosity (He), and Shannon information index have also depicted a significant level of gene diversity in our experimental populations with their mean values recorded as 0.55, 0.72, and 1.75, respectively (Table 3). Among the populations, values of 'Ho' and 'He' were varied in a range of 0.38 (Kunjkharak; QS22) to 0.67 (Chopta; QS02) and 0.63 (Lokhandi; QS05) to 0.81 (Himkhola; QS24), respectively. The mean inbreeding coefficient was high ($F_{IS} = 0.26$), indicating an excess of homozygotes. All the analyzed populations were conspicuously divided into two different groups based on the F_{IS} values. The first group was constituted by seven high-altitude populations (upper Himalayan) showing no inbreeding ($F_{IS} = 0.006$) whereas
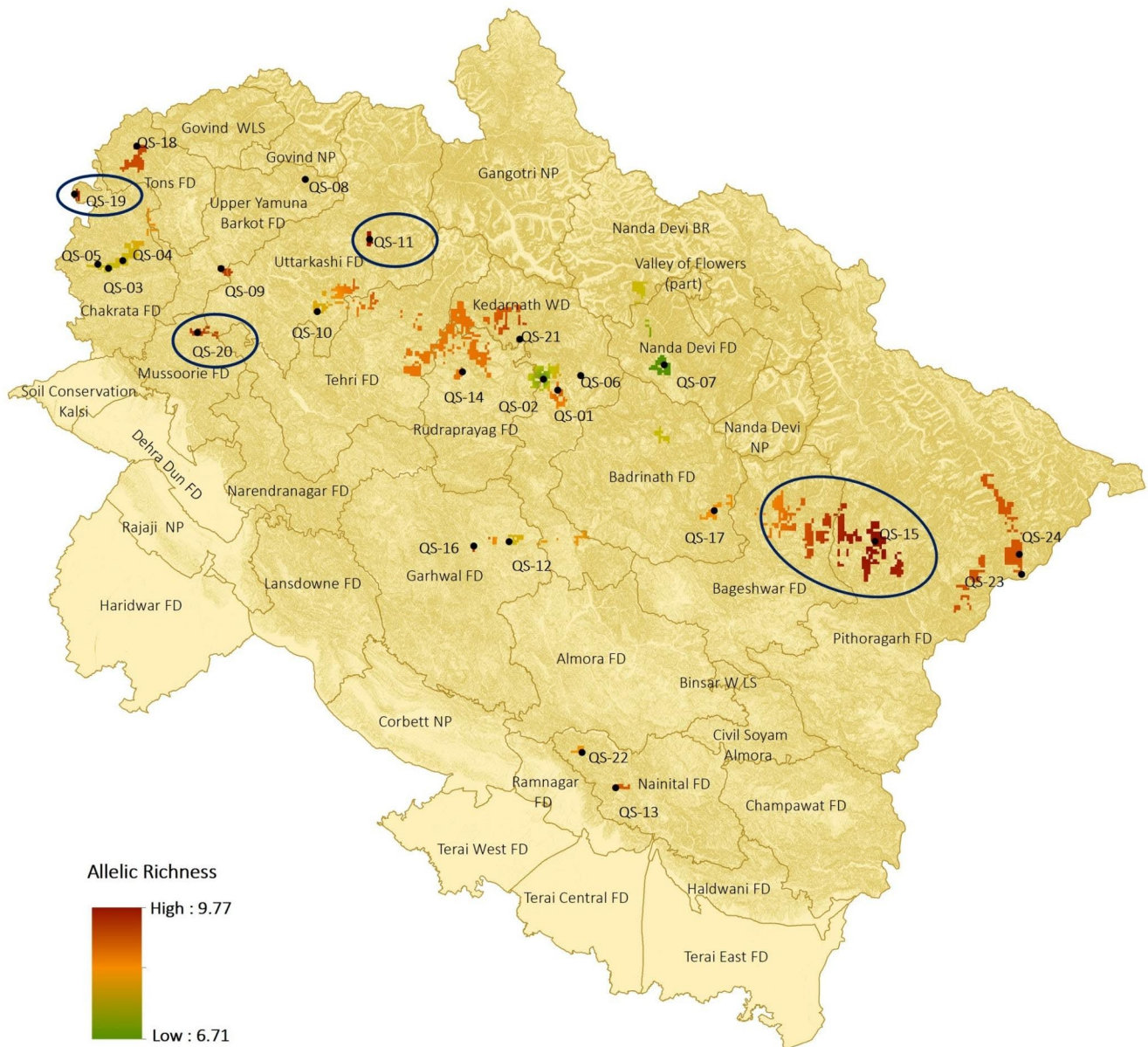
Fig. 1 Spatial overlaying of allelic richness over distribution map of *Q. semecarpifolia.* The encircled areas depict the populations of high conservation importance

remaining seventeen populations demonstrated very high inbreeding levels ($F_{IS} = 0.345$).

Genetic relationship among populations

The UPGMA dendrogram grouped the populations into two major clusters with a strong bootstrap support (Fig. 2a). Conspicuously, first major group (Cluster I) comprised the populations of upper Himalayan range, viz., Chopta (QS02), Deoban (QS03), Bhujkoti (QS04), Lokhandi (QS05), Rudranath (QS06), Auli (QS07), and Yamunotri (QS08). Whereas, all other populations of middle or lower Himalayan range were grouped in the second major group (Cluster

II) irrespective of their geographic position. Relatively, the populations under Cluster II exhibited higher gene diversity (He = 0.75) than that of Cluster I (He = 0.65). Overall topology of the UPGMA dendrogram was also consistent with the spatial clustering obtained in the PCoA where both the groups were evidently separated with considerable genetic variance (65.53%) accounted for by the first principal coordinate (Fig. 3). Further, Mantel test showed that the relationship of genetic distances between populations was non-significant with horizontal geographical distances (Fig. S2a). Whereas a weak but significant correlation was observed with altitudinal distances (r = 0.033; P < 0.024) (Fig. S2b). It signifies that the vertical geographical

Table 3 Genetic diversity statistics calculated in studied populations using ten SSR loci

Population code	Na	Ne	I	Ho	He	Ar	PAr	F_{IS}
QS 01	11	6	1.82	0.57	0.75	8.78	0.35	0.30
QS 02	8	4	1.57	0.67	0.69	6.95	0.32	0.07
QS 03	9	5	1.60	0.67	0.67	7.60	0.25	0.02
QS 04	9	5	1.61	0.67	0.66	7.71	0.22	0.05
QS 05	10	5	1.56	0.62	0.63	7.68	0.32	0.04
QS 06	9	5	1.58	0.67	0.67	7.29	0.07	0.05
QS 07	8	4	1.46	0.67	0.64	6.71	0.05	-0.04
QS 08	8	4	1.45	0.65	0.63	6.78	0.05	-0.02
QS 09	12	7	1.90	0.49	0.77	9.13	0.43	0.42
QS 10	9	5	1.73	0.47	0.74	7.89	0.25	0.41
QS 11	12	7	1.97	0.54	0.79	9.65	0.52	0.36
QS 12	10	5	1.62	0.45	0.68	7.77	0.23	0.34
QS 13	11	6	1.83	0.47	0.75	8.75	0.31	0.44
QS 14	10	5	1.74	0.51	0.72	8.46	0.30	0.34
QS 15	13	7	1.91	0.56	0.75	9.77	0.69	0.32
QS 16	12	6	1.80	0.55	0.70	9.26	0.20	0.22
QS 17	10	6	1.71	0.56	0.70	8.17	0.28	0.25
QS 18	11	6	1.86	0.46	0.75	8.96	0.26	0.44
QS 19	12	7	2.01	0.54	0.81	9.59	0.48	0.36
QS 20	12	6	1.82	0.59	0.73	9.18	0.69	0.24
QS 21	11	6	1.93	0.48	0.79	9.29	0.42	0.44
QS 22	10	5	1.72	0.38	0.72	8.11	0.16	0.45
QS 23	11	6	1.84	0.54	0.76	8.69	0.33	0.34
QS 24	10	6	1.95	0.45	0.81	8.78	0.53	0.47
Mean	10	6	1.75	0.55	0.72	8.37	0.32	0.26

Definitions: Number of observed alleles (Na), Effective number of alleles (Ne), Shannon's information index (I), Observed heterozygosity (Ho), Expected heterozygosity (He), Allelic richness (Ar), Private allelic richness (PAr), Coefficient of inbreeding (F_{IS})

distances are more crucial in genetic divergence and sub-structuring of *Q. semecarpifolia* populations rather than the horizontal distances. However, the correlation was observed as non-significant for both the geographical distances when the upper and lower Himalayan groups were analysed independently (Fig. S2c-f).

Genetic differentiation and spatial genetic structure

AMOVA without hierarchical structuring revealed considerable genetic variation (84%) explained within the populations while only 16% of variation was explained among the populations (Table 4a-b). Further, AMOVA was also performed assuming hierarchical structuring with three levels, i.e., within populations, among populations and among groups, where the groups were defined as per the UPGMA dendrogram. The analysis revealed that 76% of the genetic variation existed within the populations and only 6% was among the populations. Remarkably, a substantial amount of genetic variation (18%) was detected between the groups indicating a high genetic divergence among groups. Variance estimates were based on 999 permutations. The difference between the individuals within the populations was statistically significant (P < 0.001). By considering each

group as an independent unit, the partitioning of genetic variance was further re-examined for the populations of both groups individually. Interestingly, most genetic variance in both the groups was confined within the populations with the negligeable genetic differentiation (Table 4c-d).

The results were further supported with the Wright's fixation index (F_{ST}) which has indicated a moderate level of genetic differentiation among studied populations. The value of F_{ST} was recorded as 0.16 when calculated without hierarchical structuring. Further, it has been re-calculated for different hierarchical levels where the overall fixation index due to populations and groups was recorded as 0.24 but a significant proportion of it was explained by the groups (F_{RT} = 0.18). Pairwise F_{ST} values indicate the genetic relatedness between two sampled populations, which ranged from 0.01 to 0.22. Consequently, highest genetic distance was observed between the populations Lokhandi (QS5) and Dudatoli (QS12) while lowest was recorded for Bhujkoti (QS4) and Lokhandi (QS5). Gene flow (Nm) is another important factor associated with genetic divergence among populations which has been recorded as 2.24 for all the analysed populations. Moreover, the gene flow was remarkably high among populations of both the groups, when analysed
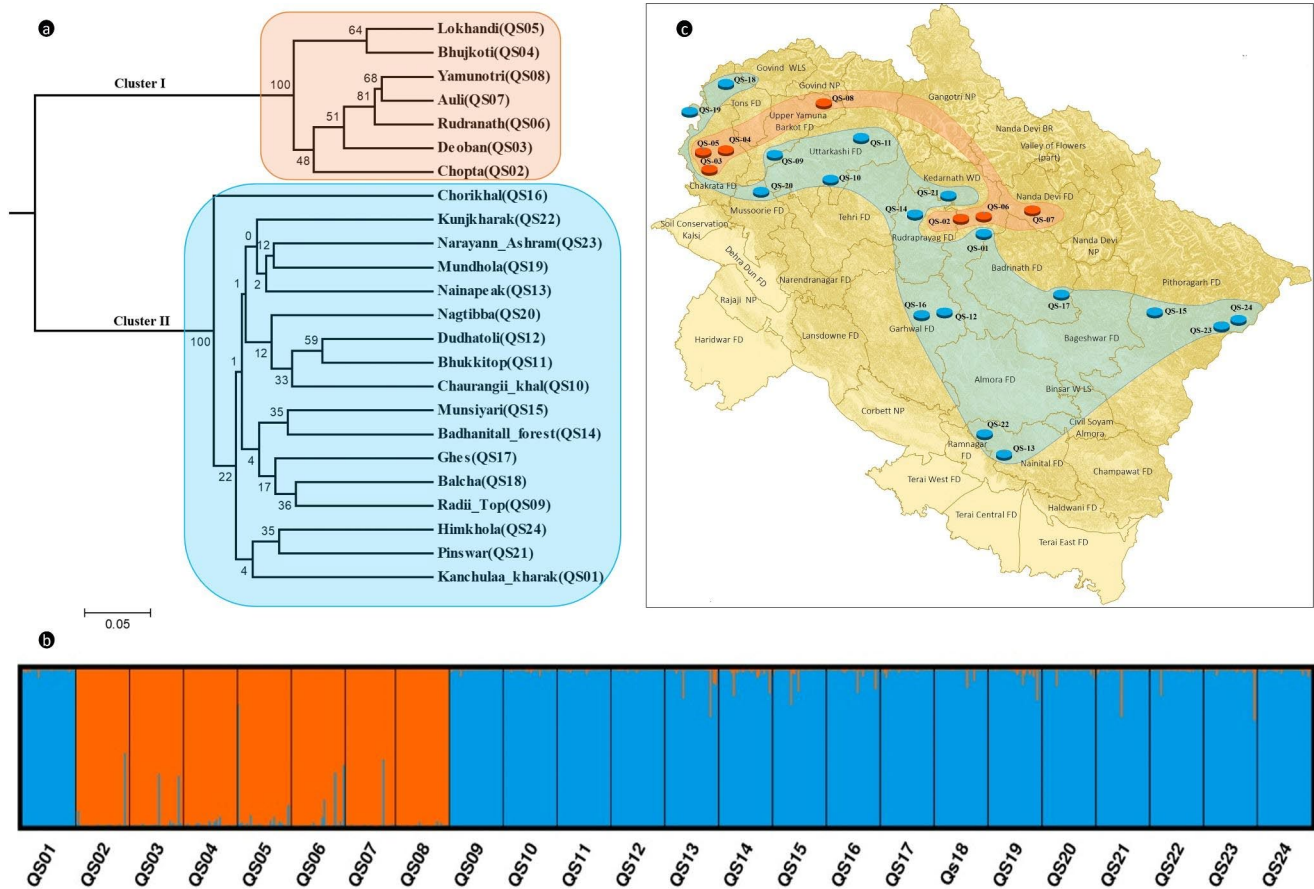
Fig. 2 Genetic clustering and spatial genetic structure in *Q. semecarpifolia*: (**a**) UPGMA dendrogram among sampled populations in which two distinct genetic clusters are highlighted by red and blue coloured box, (**b**) Bar plot showing pattern of genetic admixture among individual genotypes and populations at K = 2 in which each population is separated by a vertical line and inferred ancestry of individuals are represented by coloured bars and (**c**) Spatial overlaying of inferred ancestry of individual populations where the populations grouped under two cluster are highlighted with different colour shades corresponded to UPGMA dendrogram

independently. For instance, Nm values of Cluster I and II were recorded as 6.90 and 3.12, respectively.

Based on ΔK plot (Fig. S3), optimal K value was determined as two, indicating the presence of two major gene pools in the *Q. semecarpifolia* metapopulation. The clusters defined by Bayesian analysis were observed to be in perfect agreement with the pattern generated by other methods, viz., UPGMA and PCoA. As per the inferred ancestries (Q-matrix), all the populations were clearly defined by two genetic clusters with a membership coefficient value greater than threshold (Q ≥ 0.80). The pattern of genetic admixture is shown as a bar plot in which seven populations of the upper Himalayan region were defined by cluster I (dark orange) while other seventeen populations were assigned to cluster II (blue) (Fig. 2b and c). Based on the clustering and structure analysis, it is conspicuously evident that the *Q. semecarpifolia* populations of the western Himalayas form two distinct gene pools with least genetic exchange across the regions. However, genetic admixing was adequately high among the populations within their respective gene pools irrespective of their geographical distance.

The genetic relationship was further examined for nested clustering and sub-structuring within both the major groups individually. It was observed that the sub-structuring in the Cluster I was more pronounced than the Cluster II. As revealed by the nested clustering, seven populations of Cluster 1 were conspicuously sub-grouped into two sub-clusters and the clustering topology appeared in accordance to their geographic position. However, the sub-clustering in another major group (Cluster II) was not much conspicuous. The results were also supported by the PCoA (Fig. S4) and STRUCTURE analysis (Fig. S5 and S6) done in both the clusters, independently.
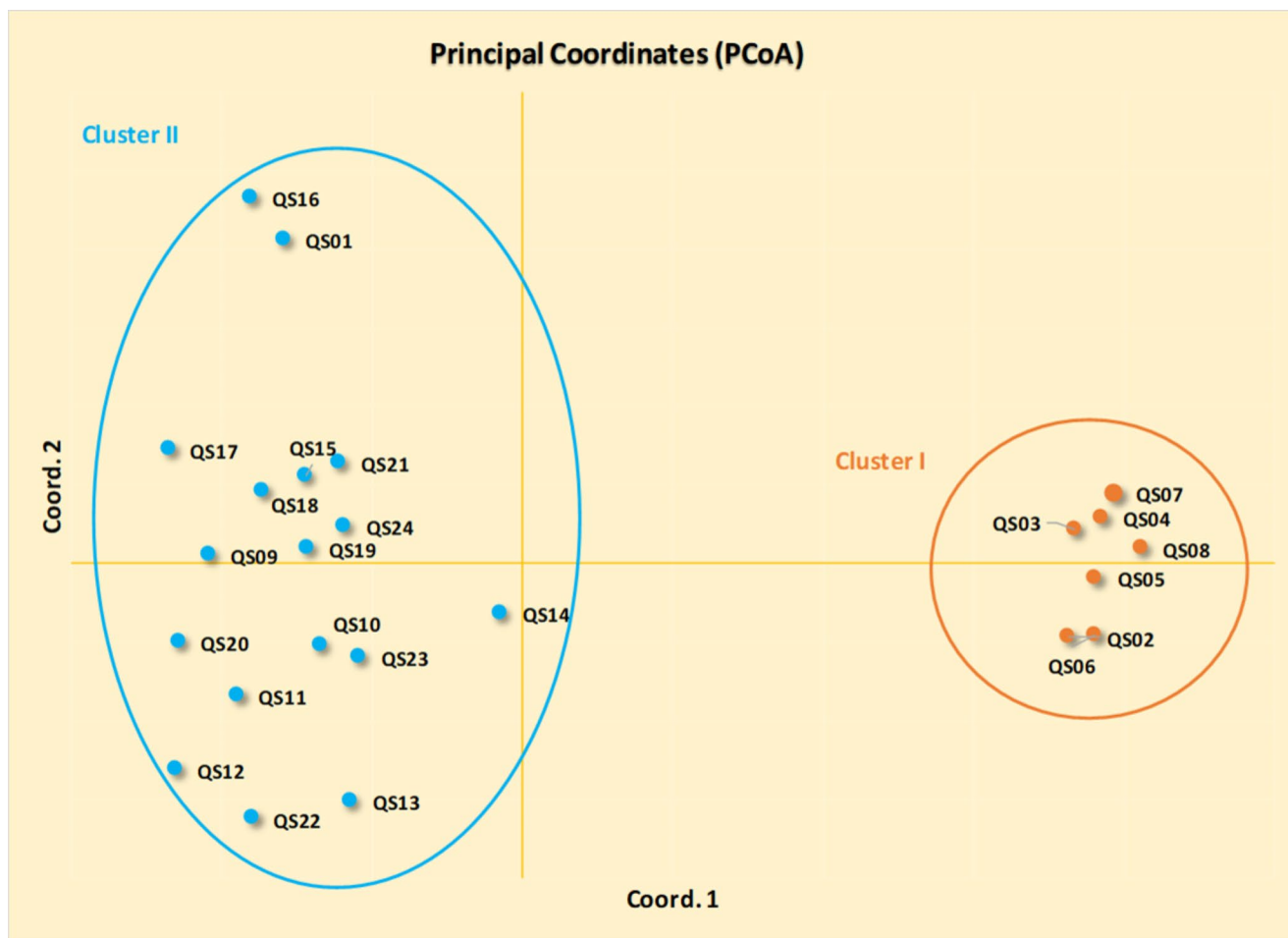
Fig. 3 Principal coordinate analysis (PCoA) showing spatial genetic clustering of *Q. semecarpifolia* populations with most genetic variance (65.53%) explained by first coordinate

Table 4 Partitioning of genetic variance in *Q. semecarpifolia* populations

Source	Degree of freedom	Sum of squares	Estimated variance	% Variation	Genetic differentiation	P (rand > = data)
a) Partitioning of the variance in all 24 populations assuming no hierarchical structure						
Among populations	23	1028.129	0.685	16	$F_{ST} = 0.156$	0.001
Within populations	1412	5238.921	3.710	84		
Total	1435	6267.051	4.395	100		
b) Partitioning of the variance in all 24 populations assuming hierarchical structure as defined by UPGMA and PCoA clustering						
Among Groups	1	529.788	529.788	18	$F_{RT} = 0.176$	0.001
Among populations	22	498.345	22.652	6	$F_{SR} = 0.079$	0.001
Within populations	1412	5238.921	3.710	76	$F_{ST} = 0.240$	0.001
c) Partitioning of the variance estimated in Cluster I						
Among populations	6	76.656	0.159	5%		
Within populations	409	1364.755	3.337	95%	$F_{ST} = 0.045$	0.001
Total	415	1441.411	3.496	100%		
d) Partitioning of the variance estimated in Cluster II						
Among populations	16	421.684	0.375	9%		
Within populations	1003	3874.167	3.863	91%	$F_{ST} = 0.088$	0.001
Total	1019	4295.851	4.237	100%		

Discussion

Understanding intraspecific genetic variability is important to unravel the adaptive or evolutionary potential of a tree species against the prevailing environmental changes and anthropogenic pressure (Templeton et al. 1995), and immensely important for guiding species conservation programme. This present study unveiled a first baseline information of population genetics in a timberline oak (*Q. semecarpifolia*) of the western Himalayas. In accordance with the questions asked, level of gene diversity, genetic divergence and population genetic structure were determined for the experimental populations of western Himalayas. The spatial distribution of allelic diversity and genetic structure was further elucidated by overlaying them over the distribution map. The distribution of gene diversity was also analyzed in relation to the horizontal and vertical geographical distances.

Across all the loci, we found a good level of polymorphism exhibiting 4 to 18 alleles per locus, which was also in congruence to the earlier report by Ueno et al. (2008) in *Q. mongolica*. The calculated measures of gene diversity (Ho and He) in *Q. semecarpifolia* populations of the Himalayan region are found comparable to the earlier studies carried out in different oak species (Table S1). Nonetheless, the deviations are observed in measures of genetic differentiation and inbreeding among different oak species. For instance, negligeable genetic differentiation and low inbreeding had been observed in most oak species but significant inbreeding had been reported in *Q. glauca* ($F_{IS} = 0.29$; Lee et al. 2006), *Q. petrea* ($F_{IS} = 0.39$; Lupini et al. 2019), and *Q. oglethorpensis* ($F_{IS} = 0.23$; Spence et al. 2021). Similar to this study, high levels of genetic diversity were also reported in three Mexican oaks (*Q. candicans, Q. crassifolia*, and *Q. castanea*) but significant inbreeding was detected in about 40% populations (Oyama et al. 2018). Earlier study by Spence et al. (2021) demonstrated that the threatened oak species with a narrow distribution range possess lesser genetic diversity than the other widely distributed oaks. Conclusively, high gene diversity with varied level of genetic differentiation and inbreeding reported in various oak species may be ascribed to the range size, habitat degradation, threat status, environment heterogeneity, anthropogenic disturbances, and life history traits like open pollinated mating behavior of this genus, masting events, viviparous seed germination, etc.

However, this situation may further deteriorate if local and landscape-scale anthropogenic pressures are not checked. This is because the gradual loss of private alleles and increasing inbreeding, as found in the present case, can affect the sustainability and existence of *Q. semecarpifolia* populations, if kept small and isolated for many generations. Gene flow is a key process involved in the distribution of gene diversity within and across spatially separated populations and counteracts the genetic differentiation. In general, Nm value greater than 1 indicates optimal gene flow (or little differentiation) among the populations (McDermott and McDonald 1993), and movement of at least a single individual per generation can prevent significant divergence between the populations (Wright 1969). Despite a substantial level of gene flow (Nm = 2.24) observed in our experimental populations, a moderate level of genetic differentiation was also detected which was further investigated for the two inferred groups independently. Groupwise partitioning of molecular variance revealed that the gene diversity was mostly detected within the populations of each group, and the genetic differentiation among populations of each group was negligible. Also, substantial levels of gene flow detected in both the clusters, viz., Cluster I (Nm = 6.90) and Cluster II (Nm = 3.12), had indicated that the populations within the groups were well genetically connected but the gene flow was restricted between populations of two groups. Further, the extent of gene flow depends on the geographic range of a species and breeding behaviour. For instance, the broader the distribution ranges, the greater the chance of allele dispersal as well as reunion during fertilization, and open-pollinated taxa demonstrate higher gene flow than the self-pollinated one (Hamrick and Godt 1990, 1996). Thus, the substantial genetic diversity with high gene flow recorded in *Q. semecarpifolia* could be attributed to its wide distribution range and open pollinated reproductive system.

The oak forest consists of gregarious patches of microhabitats which unceasingly affected by factors driven by anthropogenic activities and climate change, such as recurrence of forest fires, reduced regeneration, frost, tourism, pilgrimage, illicit felling, collection of high valued caterpillar fungus (*Ophiocordyceps sinensis*), etc. For example, several millions of collectors stay in alpine meadows and tree line areas of Uttarakhand, Nepal and Tibet each summer, and dig up soil to collect caterpillar fungus from plant roots. They not only trample ground vegetation but also collect firewood from the nearby tree line populations (Singh 2018). These disturbances may affect the species regeneration, dispersal, successional status, and inbreeding in the long-term. This temporal heterogeneity is especially strong in temperate forests (Wright Jr 1976). Being a later successional species with poor colonizing habit, achieving good regeneration has remained a matter of concern in *Q. semecarpifolia* (Negi and Negi 2021; Rawat et al. 2022). Besides, monoecious pollination, masting event (long fruiting cycle of 8 to 10 years), low seed viability, and precocious germination (Negi and Naithani 1995; Singh et al. 2011), are other important causes affecting population dynamics and genetics, which

may potentially be influenced by the climate change in the Himalayas (Chakraborty et al. 2018).

Strikingly, high inbreeding detected in this study indicated some important evolutionary changes adopted by the populations. The actual causes associated with this are not well known and necessitate further investigation. But based on the literature and field observations, it could be explained by various factors, such as degradation of natural population, poor regeneration, monoecious pollination, limited seed dispersal, masting events, etc. In addition to common evolutionary drivers, such as gene flow and selection, positive assortative mating (i.e., preferential mating among genetically or phenotypically close relatives) caused by spatial isolation and asynchronized flowering among individuals of a location may result in significant deviation of homo-and-heterozygote frequency (Lemes et al. 2003; Kremer and Hipp 2020). Also, the effective population size is not maintained at several pockets of distribution range due to overexploitation and habitat destruction. Moreover, *Q. semecarpifolia* is a viviparous oak in which the acorns begin to germinate before or during their deposition on the ground (Tewari et al. 2019). Consequently, most of the seed dispersal occurs between nearest-neighbour populations during mast seed year due to their large size and precocious germination. Consequently, the individuals within a population may undergo preferential mating among their close relatives and led inbreeding in long-term. In case of SSR-based allelic data, excess of homozygotes could also be aroused due to the presence of null alleles as exemplified in *Q. glauca* (Lee et al. 2006). This study also detected null alleles for some of the SSR loci, and hence, their effects could not be ruled out.

Although the rear edge populations are often disproportionately important for the survival and evolution of biota (Hampe and Petit 2005), it is always useful to conserve all the populations irrespective of their genetic constitution because the populations with low allelic diversity may still contain important "unique" alleles. However, the conservation programme could be prioritized first for the populations which are rich in gene diversity. The geospatial interpolation of genetic diversity enabled the demarcation of conservation units for in situ conservation. Populations or regions capturing higher allelic diversity as well as private alleles were identified and designated as diversity hotspots. For instance, the populations of eastern region of Uttarakhand under Pithoragarh forest division displayed a great level of allelic diversity and it may be considered as a key conservation unit. Besides, four genetically diverse populations have been recognized into both the geographic region of the state, i.e., Bhukkitop, Mundhola, and Nag Tibba in Garhwal and Munsiyari in Kumaon Himalayas, which may be prioritized in the conservation programmes. Viewing the spatial distribution of genetic diversity, the main center of genetic diversity of Himalayan brown oak is located in the eastern region of Uttarakhand.

The UPGMA clustering has distinguished the population in two well defined clusters with high bootstrap support where the marginal populations of the upper Himalayan region were separated from the remaining populations. The consistency of the dendrogram was further confirmed by the principal coordinate analysis. Genetic divergence among natural populations of a tree species is known to be influenced by the geographic distances and discrete geographic barriers by limiting the seed and pollen dispersal (Wright 1943). Moreover, ecological isolation may also be driven by habitat heterogeneity in nearby populations without any geographic barrier, which may promote divergence by local adaptation and drift (Misiewicz and Fine 2014). In the same way, altitudinal gradients may also play a crucial role in shaping genetic diversity and structure of populations, particularly those marginally distributed (Reis et al. 2015). Astonishingly, Mantle's test conducted in our sampled population has revealed no significant relationship of genetic distance with the horizontal geographic distance. Besides, a weak but significant correlation was detected with the vertical altitudinal distance, whereas the relationship with both the geographic distances was found to be non-significant when analyzed in both the groups independently. It suggests that the altitudinal variation is apparent as crucial in determining genetic divergence of *Q. semecarpifolia* populations at metapopulation level, but the spatial distance plays insignificant role at local scale. In congruence, similar observations were also depicted in the Mongolian oak (Ohsawa et al. 2007; Ueno et al. 2008), where no significant isolation by distance was observed among the populations along geographic span.

Literature indicates that the forests in upper areas of western and central Himalayan region are being vulnerable to the projected impacts of climate change (Joshi et al. 2012; Shrestha et al. 2012). The potential habitat of *Q. semecarpifolia* is predicted to shrink by 40% and 76% with 1 and 2 °C increase in temperature, respectively (Saran et al. 2010). There is a prevalent assumption that geographically peripheral populations harbor lower genetic diversity and higher genetic differentiation than the core/ buffer zone populations as a result of higher genetic drift, fragmentation, and isolation (Lesica and Allendorf 1995; Eckert et al. 2008; Pandey and Rajora 2012). In congruence to this, the peripheral high-altitude populations distinguished as Cluster I forms a sub-alpine timber line in the western Himalayas and demonstrated lesser gene diversity than the core populations occurring in the broadleaved mixed forest at lower altitudinal range.

The analysis of genetic structure deciphered the admixed ancestry among individuals and populations, where two major gene pools have been recognized throughout the distribution range in western Himalayas. Interestingly, the sub-structuring was observed across the altitudinal gradient rather than the horizontal spatial distance. In agreement to the clustering depicted by UPGMA and PCoA, the populations from the upper Himalayan region were separated from other populations and clearly defined by their respective clusters with Q values ≥ 9. The results of the genetic clustering and STRUCTURE analysis have indicated towards the obvious genetic constraint aroused by life-history traits, geographic barriers, clinal variation, and ecological heterogeneity between populations (Loveless and Hamrick 1984; Morente-López et al. 2018), which need to be studied in-depth through environment association analysis using gene based markers. Re-analyzing data of both the clades separately gives an important clue that the differences of inbreeding in the populations of both the groups would have led to this clustering pattern. As evident in Table 3, most high-altitude populations grouped in Cluster I showed negligible inbreeding while significant inbreeding was detected in the core populations which are exposed to various anthropogenic disturbances. Independent Bayesian analysis of both the groups further revealed a strong sub-structuring in Cluster I, indicating existence of multiple gene pools but the sub-structuring in Cluster II remained insignificant. It showed that the populations of both the groups responded differently against the prevailing evolutionary forces. Topographic features such as high mountain ranges, perennial rivers, grasslands, etc. disconnect the populations to several hundred miles and cause hindrances in the smooth genetic exchange via pollen. Similarly, extent of seed dispersal is restricted by own endogenous physiological characteristics and geographic barrier. Immigration of alien genes changes the genetic composition of the recipient populations by constituting novel allelic combinations (Milgroom 2015), which may further contribute to boost the adaptive and evolutionary potential of populations against the changing environment. Being an extreme environmental condition, peripheral populations are at great risk of diversity loss and likely to be constrained in their ability to tolerate rapid climate change. Thus, conservation plans should include populations found both near the center and the periphery of a species' distribution, where conservation of peripheral populations may allow to continue the evolutionary process that are likely to generate future evolutionary diversity (Lesica and Allendorf 1995).

Conservation implications and conclusions

Preservation of genetic diversity is one of the main objectives of conservation programmes (Frankham 2010; Allendorf et al. 2013; Oldenbroek 2017), and is aimed at maximizing either expected heterozygosity or allelic diversity. In fact, maximization of allelic diversity is considered to be more efficient in upholding genetic diversity of subdivided populations than maximization of expected heterozygosity because the former maintains a larger number of alleles and better control of inbreeding (López-Cortegano et al. 2019). Hence, the populations with higher allelic diversity can be prioritized for conservation either in situ or ex situ (Petit et al. 1998). Present study revealed a good level of gene diversity and significant genetic differentiation (among groups) in *Q. semecarpifolia* populations of the western Himalayas, with a center of diversity predicted at forest areas under Pithoragarh division, the eastern region of Uttarakhand extending towards Nepal. Spatial genetic clustering has exemplified two distinct gene pools of *Q. semecarpifolia* in western Himalayas, which are separated by numerous geographic and ecological barriers. Compared to horizontal geographic distance, altitudinal variation, environment heterogeneity, and landscape features play a significant role in shaping the distribution of genetic diversity.

The gene diversity is relatively high in the core populations of lower altitudinal range than the peripheral populations at high-altitude. Augmented gene flow from these genetically diverse and distinct populations needs to be considered as a way of increasing fitness and the adaptive potential of populations. In order to make best possible use of the high genetic diversity populations, it will be important to harvest acorns from these populations and used them to infuse the diversity in the populations with a narrow genetic base. Also, the highly diverse populations could serve as a source for the seed or planting material for establishment of ex situ field gene banks. Conversely, alleles of the smaller and degraded population could also be rescued by broadcasting their acorns into the large sink populations with broad genetic base.

Data availability This manuscript did not generate any genetic resource in the form of nucleic acid or protein sequences, SNPs, expression data, etc. which could be deposited in the databases. The SSR sequences used here for the genotyping work are already in the public domain and their references have been cited in the manuscript. Further, all the data generated herein are presented either in the manuscript or supplementary material.

Declarations

Conflicts of interest/Competing interests I hereby affirm that the manuscript has neither published elsewhere nor submitted simultaneously for publication elsewhere. The article has been read by each author and there is no conflict of interest.

Ethics approval All necessary permissions were obtained for sample collection from concerned state forest department.

References

Allendorf FW, Luikart GH, Aitken SN (2013) Conservation and the genetics of populations, 2nd edn. Wiley-Blackwell

Bagnoli F, Vendramin GG, Buonamici A et al (2009) Is *Cupressus sempervirens* native in Italy? An answer from genetic and palaeobotanical data. Mol Ecol 18:2276–2286. https://doi.org/10.1111/j.1365294X.2009.04182.x

Bisht H (2001) Physiobiochemical aspects of seed viability in Quercus semecarpifolia smith: a possible recalcitrant seed. Ph D Thesis 45–48. https://doi.org/10.3923/rjss.2012.63.70. H.N.B. Garhwal University, Srinagar (Garhwal), 246174, Uttarakhand, India

Bisht VK, Kuniyal CP, Nautiyal BP, Prasad P (2013) Spatial distribution and regeneration of *Quercus semecarpifolia* and *Quercus floribunda* in a subalpine forest of western Himalaya, India. Physiol Mol Biol plants 19:443–448. https://doi.org/10.1007/s12298-013-0189-z

CBD (2022) Nations Adopt Four Goals, 23 Targets for 2030 In Landmark UN Biodiversity Agreement. 15th Conference of Parties to the UN Convention on Biological Diversity, 19th December 2022, Montreal. https://www.cbd.int/article/cop15-cbd-press-release-final-19dec2022. Accessed 1 June 2023

Chakraborty A, Saha S, Sachdeva K, Joshi PK (2018) Vulnerability of forests in the himalayan region to climate change impacts and anthropogenic disturbances: a systematic review. Reg Environ Chang 18:1783–1799. https://doi.org/10.1007/s11027-010-9257-7

Chiocchini F, Mattioni C, Pollegioni P et al (2016) Mapping the genetic diversity of *Castanea sativa*: exploiting spatial analysis for biogeography and conservation studies. J Geogr Inf Syst 8:248. https://doi.org/10.4236/jgis.2016.82022

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographic ranges: the central–marginal hypothesis and beyond. Mol Ecol 17:1170–1188

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Frankham R (2010) Where are we in conservation genetics and where do we need to go? Conserv Genet 11:661–663. https://doi.org/10.1007/s10592-009-0010-2

Gapare WJ (2014) Merging applied gene conservation activities with advanced generation breeding initiatives: a case of study of *Pinus radiata* D. Don. New For 45:311–331. https://doi.org/10.1007/s11056-013-9398-0

Geburek T, Konrad H (2008) Why the conservation of forest genetic resources has not worked. Conserv Biol 22:267–274. https://doi.org/10.1111/j.1523-1739.2008.00900.x

Gómez A, Vendramin GG, González-Martínez SC, Alía R (2005) Genetic diversity and differentiation of two Mediterranean pines (*Pinus halepensis* Mill. And Pinus pinaster Ait.) Along a latitudinal cline using chloroplast microsatellite markers. Divers Distrib 11:257–263. https://doi.org/10.1111/j.1366-9516.2005.00152.x

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. Ecol Lett 8:461–467. https://doi.org/10.1111/j.1461-0248.2005.00739.x

Hamrick JL, Godt MJW (1990) Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding, and genetic resources. Sinauer Associates Inc., Sunderland, pp 43–63

Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. Philos Trans R Soc B 351:1291–1298. https://doi.org/10.1098/rstb.1996.0112

Hengl T (2009) A practical guide to geostatistical mapping. University of Amsterdam, Amsterdam

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Joshi PK, Rawat A, Narula S, Sinha V (2012) Assessing impact of climate change on forest cover type shifts in western himalayan eco-region. J For Res 23:75–80. https://doi.org/10.1007/s11676-012-0235-7

Kalinowski ST (2005) hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. Mol Ecol Notes 5:187–189. https://doi.org/10.1111/j.1471-8286.2004.00845.x

Kremer A, Hipp AL (2020) Oaks: an evolutionary success story. New Phytol 226:987–1011. https://doi.org/10.1111/nph.16274

Lee Y-J, Hwang S-Y, Ho K-C, Lin T-P (2006) Source populations of *Quercus glauca* in the last glacial age in Taiwan revealed by nuclear microsatellite markers. J Hered 97:261–269. https://doi.org/10.1093/jhered/esj030

Lemes MR, Gribel R, Proctor J, Grattapaglia D (2003) Population genetic structure of mahogany (*Swietenia macrophylla* King, Meliaceae) across the brazilian Amazon, based on variation at microsatellite loci: implications for conservation. Mol Ecol 12:2875–2883. https://doi.org/10.1046/j.1365-294X.2003.01950.x

Lesica P, Allendorf FW (1995) When are peripheral populations valuable for conservation? Conserv Biol 9:753–760. https://doi.org/10.1046/j.1523-1739.1995.09040753.x

Li YL, Liu JX (2018) StructureSelector: a web-based software to select and visualize the optimal number of clusters using multiple methods. Mol Ecol Resour 18:176–177. https://doi.org/10.1111/1755-0998.12719

López-Cortegano E, Pouso R, Labrador A et al (2019) Optimal management of genetic diversity in subdivided populations. Front Genet 10:843. https://doi.org/10.3389/fgene.2019.00843

Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. Annu Rev Ecol Syst 65–95. https://doi.org/10.1146/annurev.es.15.110184.000433

Lupini A, Aci MM, Mauceri A et al (2019) Genetic diversity in old populations of sessile oak from Calabria assessed by nuclear and chloroplast SSR. J Mt Sci 16:1111–1120. https://doi.org/10.1007/s11629-018-5335-1

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

Matschiner M, Salzburger W (2009) TANDEM: integrating automated allele binning into genetics and genomics workflows. Bioinformatics 25:1982–1983. https://doi.org/10.1093/bioinformatics/btp303

McDermott JM, McDonald BA (1993) Gene Flow in Plant Pathosystems. Annu Rev Phytopathol 31:353–373. https://doi.org/10.1146/annurev.py.31.090193.002033

Milgroom MG (2015) Migration and population structure. In: Population Biology of Plant Pathogens: Genetics, Ecology, and Evolution. The American Phytopathological Society, pp 119–146. https://doi.org/10.1094/9780890544525.006

Mishima K, Watanabe A, Isoda K et al (2006) Isolation and characterization of microsatellite loci from Quercus mongolica var. Crispula. Mol Ecol Notes 6:695–697. https://doi.org/10.1111/j.1471-8286.2006.01313.x

Misiewicz TM, Fine PVA (2014) Evidence for ecological divergence across a mosaic of soil types in an amazonian tropical tree: Protium subserratum (Burseraceae). Mol Ecol 23:2543–2558. https://doi.org/10.1111/mec.12746

Morente-López J, García C, Lara-Romero C et al (2018) Geography and environment shape landscape genetics of Mediterranean alpine species Silene ciliata Poiret. (Caryophyllaceae). Front Plant Sci 1698. https://doi.org/10.3389/fpls.2018.01698

Negi SS, Naithani HB (1995) Oaks of India, Nepal and Bhutan. International Book Distributors, Dehradun

Negi M, Negi VS (2021) Temporal changes in Oak forests over last three decades in western Himalaya, India. Trees For People 6:100146. https://doi.org/10.1016/j.tfp.2021.100146

Negi VS, Pathak R, Rawal RS, Bhatt ID, Sharma S (2019) Long-term ecological monitoring on forest ecosystems in indian Himalayan Region, Criteria and indicator approach. Ecol Indic 102:374–381. https://doi.org/10.1016/j.ecolind.2019.02.035

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data. J Mol Evol 19:153–170. https://doi.org/10.1007/BF02300753

Nixon KC (2006) Global and neotropical distribution and diversity of oak (genus Quercus) and oak forests. In: Kappelle M (ed) Ecology and conservation of neotropical montane oak forests. Springer, Berlin, Heidelberg, pp 3–13

Ohsawa T, Tsuda Y, Saito Y et al (2007) Altitudinal genetic diversity and differentiation of Quercus crispula in the Chichibu Mountains, central Japan. Int J Plant Sci 168:333–340. https://doi.org/10.1086/510413

Oldenbroek JK (2017) Genomic management of animal genetic diversity. Wageningen Academic Publishers, Wageningen

Oyama K, Ramírez-Toro W, Peñaloza-Ramírez JM et al (2018) High genetic diversity and connectivity among populations of Quercus candicans, Quercus crassifolia, and Quercus castanea in a heterogeneous landscape in Mexico. Trop Conserv Sci 11. https://doi.org/10.1177/19400829187661

Pandey M, Rajora OP (2012) Genetic diversity and differentiation of core vs. peripheral populations of eastern white cedar, Thuja occidentalis (Cupressaceae). Am J Bot 99:690–699. https://www.jstor.org/stable/41415610

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an

update. Bioinformatics 28:2537e2539. https://doi.org/10.1093/bioinformatics/bts460

Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of genetic markers. Conserv Biol 12:844–855. https://doi.org/10.1111/j.1523-1739.1998.96489.x

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959. https://pubmed.ncbi.nlm.nih.gov/10835412/

Puechmaille SJ (2016) The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. Mol Ecol Resour 16:608–627. https://doi.org/10.1111/1755-0998.12512

Rawat B, Rawat JM, Purohit S et al (2022) A comprehensive review of Quercus semecarpifolia Sm.: an ecologically and commercially important himalayan tree. Front Ecol Evol 10:961345. https://doi.org/10.3389/fevo.2022.961345

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. Conserv Biol 17:230–237. https://doi.org/10.1046/j.1523-1739.2003.01236.x

Reis TS, Ciampi-Guillardi M, Bajay MM et al (2015) Elevation as a barrier: genetic structure for an atlantic rain forest tree (Bathysa australis) in the Serra do marmountain range, SE Brazil. Ecol Evol 5:1919–1931. https://doi.org/10.1002/ece3.1501

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. Mol Ecol Notes 4:137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Saran S, Joshi R, Sharma S et al (2010) Geospatial modeling of Brown oak (Quercus semecarpifolia) habitats in the Kumaun Himalaya under climate change scenario. J Indian Soc Remote Sens 38:535–547. https://doi.org/10.1007/s12524-010-0038-2

Schneider S, Roessli D, Excoffier L (2000) Arlequin: a software for population genetics data analysis. User Man ver 2:2496–2497. https://journals.sagepub.com/doi/pdf/10.1177/117693430500100003

Shekhar C, Rawat A, Bhandari MS et al (2021) Cross-transferability-based identification and validation of simple sequence repeat (SSR) markers in oaks of western Himalayas. Silvae Genet 70:108–116. https://doi.org/10.2478/sg-2021-0009

Shekhar C, Ginwal HS, Meena RK et al (2022) Spatio-temporal distribution of broad-leaved Quercus semecarpifolia indicates altitudinal shift in northwestern Himalayas. Plant Ecol 223:671–697. https://doi.org/10.1007/s11258-022-01240-x

Shepard D (1968) A two-dimensional interpolation function for irregularly-spaced data. In: Proceedings of the 1968 23rd ACM national conference. Association for Computing Machinery, New York, pp 517–524

Shrestha BB (2003) Quercus semecarpifolia Sm. In the himalayan region: Ecology, exploitation and threats. Himal J Sci 1:126–128. https://doi.org/10.3126/hjs.v1i2.212

Shrestha UB, Gautam S, Bawa KS (2012) Widespread climate change in the Himalayas and associated changes in local ecosystems. PLoS ONE 7:e36741. https://doi.org/10.1371/journal.pone.0036741

Singh SP (2018) Research on indian Himalayan Treeline Ecotone: an overview. Trop Ecol 59(2):163–176. http://www.tropecol.com/pdf/open/PDF_59_2/01%20Surender%20P%20Singh.pdf

Singh G, Rawat GS, Verma D (2010) Comparative study of fuelwood consumption by villagers and seasonal "Dhaba owners" in the tourist affected regions of Garhwal Himalaya, India. Energy Policy 38:1895–1899. http://www.ias.ac.in/currsci

Singh G, Rai ID, Rawat GS (2011) The year 2010 was' mast seed year'for the Kharsu oak (Quercus semecarpifolia Sm.) In the western Himalaya. Curr Sci 100:1275–1276. http://www.ias.ac.in/currsci

Souto C, Mathiasen P, Acosta M (2015) Identifying genetic hotspots by mapping molecular diversity of widespread trees: when

commonness matters. J Hered 106:537–545. https://doi.org/10.1093/jhered/esv023

Spence ES, Fant JB, Gailing O et al (2021) Comparing genetic diversity in three threatened Oaks. Forests 12:561. https://doi.org/10.3390/f12050561

Takezaki N, Nei M, Tamura K (2010) POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. Mol Biol Evol 27:747–752. https://doi.org/10.1093/molbev/msp312

Tashi S (2004) Regeneration of *Quercus semecarpifolia* Sm. in an old growth oak forest under Gidakom FMU-Bhutan. M. Sc. Dissertation, Forest Ecology and Forest Management. Department of Forestry, Wageningen University and Research Centrum, the Netherlands

Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: a cladistic analysis of the geographic distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. Genetics 140:767–782. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1206651/

Tewari A, Shah S, Singh N et al (2019) Acorn maturation and regeneration problem in *Quercus semecarpifolia* Sm. Himal tree line Int J Sci Technol Res 8:3781–3787

Ueno S, Tsumura Y (2008) Development of ten microsatellite markers for *Quercus mongolica* var. *Crispula* by database mining. Conserv Genet 9:1083–1085. https://doi.org/10.1007/s10592-007-9462-4

Ueno S, Taguchi Y, Tsumura Y (2008) Microsatellite markers derived from *Quercus mongolica* var. *Crispula* (Fagaceae) inner bark expressed sequence tags. Genes Genet Syst 83:179–187. https://doi.org/10.1266/ggs.83.179

Uttarakhand at a Glance 2020-21 (2021) Directorate of Economics & Statistics, Government of Uttarakhand, Dehradun, India. 31 pages. https://des.uk.gov.in/pages/display/61-uttarakhnad-at-a-glance

Uttarakhand Forest Statistics 2014–2015 (2015) Uttarakhand Forest Department, Dehradun, India. 181 pages. http://www.uafdc.in/Acts&Rules/Forest_Statistical_2014-15.pdf

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Vetaas OR (2000) The effect of environmental factors on the regeneration of *Quercus semecarpifolia* Sm. In central Himalaya, Nepal. Plant Ecol 146:137–144. https://doi.org/10.1023/A:1009860227886

Wright S (1943) Isolation by distance. Genetics 28:114–138. https://doi.org/10.1093/genetics/28.2.114

Wright S (1969) Evolution and the genetics of populations, volume 2: the theory of gene frequencies. University of Chicago Press, Chicago

Wright HE Jr (1976) The Environmental setting for Plant Domestication in the Near East: wild cereal grains may not have entered the Near East until the end of the Pleistocene 11,000 years ago. Sci (80-) 194:385–389. https://doi.org/10.1126/science.194.4263.385

Yang J, Vázquez L, Feng L et al (2018) Climatic and soil factors shape the demographic history and genetic diversity of a deciduous oak (*Quercus liaotungensis*) in Northern China. Front Plant Sci 9:1534. https://doi.org/10.3389/fpls.2018.01534