Genet Resour Crop Evol
https://doi.org/10.1007/s10722-024-01988-3

RESEARCH ARTICLE

Genetic diversity and population structure analysis in tamarind (*Tamarindus indica* L.) using SCoT and SRAP markers

C. Kanupriya · G. Karunakaran · Pritee Singh · R. Venugopalan · Deepa Samant · D. C. Lakshmana Reddy · K. Prakash

Received: 8 August 2023 / Accepted: 6 April 2024

Abstract Tamarind (*Tamarindus indica* L.) is a large and long-lived tropical tree. Although the distinctly acidic pulp of its fruit finds multiple applications in nutrition, medicine, and industry, the genetic diversity of tamarind and the identification of unique genotypes suitable for different applications have received little attention. To address this gap, molecular genetic diversity of 91 genotypes of tamarind from different regions in India was assessed using start-codon-targeted (SCoT) markers and sequence-related amplified polymorphism (SRAP) markers. Of the 46 markers tested, 10 SCoT and 10 SRAP markers were polymorphic, and the polymorphic information content values ranged from 0.38 to 0.45 with a mean of 0.40 for both the primers. Genetic relationships among the genotypes investigated using a neighbor-joining tree showed the 91 genotypes to be composed of four distinct clusters, and analysis of population structure showed them to be corresponding to four pure-population groups (the probability score was 0.80 or higher). Principal component analysis of the molecular variance showed it to be along three axes: axis 1 explained 31.83% of the variation; axis 2, 12.40%; and axis 3, 9.99%. The analysis of molecular variance indicated that 2% of the variance was observed among populations, and 10% of the variance was due to the differences among individual plants. Most (88%) of the variation was found within individuals. The study demonstrated significant genetic diversity among the genotypes of tamarind, which can help in targeted breeding not only for overall improvement of the crop but also for the enhancement of specific traits. The identified markers can serve as valuable genomic resources for the conservation and utilization of tamarind.

Keywords *Tamarindus indica* L. · Natural population · SCoT · SRAP · Molecular diversity · Cluster analysis · PCoA · Structure analysis

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10722-024-01988-3.

C. Kanupriya · G. Karunakaran (✉) · P. Singh · R. Venugopalan · D. C. Lakshmana Reddy
ICAR-Indian Institute of Horticultural Research, Hessarghatta, Bengaluru, Karnataka 560 089, India
e-mail: ganesan.karunakaran@icar.gov.in

D. Samant
ICAR-IIHR-Central Horticultural Experiment Station, Bhubaneswar, Odisha 751 019, India

K. Prakash
ICAR-Indian Agricultural Statistics Research Institute, New Delhi 110 012, India

Introduction

Tamarindus is a monotypic genus (family Fabaceae) represented solely by *T. indica* L., or tamarind. The species is native to tropical Africa but is widely grown and naturalized in South East Asia and South

Published online: 04 May 2024

America, being well-adapted to arid and semi-arid climates. Tamarind is widely distributed in India, and every part of the sprawling and long-lived tree—pods, pulp, shell, seeds, leaves, flowers and bark—finds applications in the food industry, traditional medicine, cosmetics, and various other industries (Dorni et al. 2017).

The characteristically tangy pulp of tamarind pods is a key ingredient in Indian cooking, especially in southern India, and used in a variety of dishes including chutneys, sauces, and beverages. Besides culinary uses, the fruit has multiple medicinal and industrial applications because of its rich profile of antioxidants, dietary fibre, and vitamins (Sandesh et al. 2014). The hard but brittle outer shell of tamarind pods is used as fuel or as a component in producing biochar, which is used as a soil amendment. Tamarind seeds can be processed to prepare tamarind kernel powder (TKP), which is rich in a high-molecular-weight polysaccharide known as natural gum, also finds numerous applications because of its unique properties: the gum is used as a stabilizer, a bulking agent, as a coating, as a thickening agent, and as an emulsifier in various food products. On industrial scale, TKP is used in textiles and paper as an adhesive and as sizing (Thombare et al. 2014). Tamarind seed oil is used in soaps, cosmetics, and as a base for various pharmaceutical formulations (Chacón-Fernández et al. 2019). Tamarind leaves and flowers contain essential oils, free and conjugated fatty acids, as well as flavonoids. These components have been traditionally used in medicine due to their potential health benefits and are believed to be antimicrobial, anti-inflammatory, and anthelmintic and a source of antioxidants (Raja et al. 2022). Tamarind leaf extract has been used to reduce metal salts into nanoparticles, a process deployed in the textile industry (Mansingh et al. 2021). Lastly, tamarind bark is rich in tannins and polyphenols, which contribute to its medicinal properties, and tannins from the bark are also used in manufacturing ink and for fixing dyes (Bhadoriya et al. 2011).

India leads the world in acreage and production of tamarind—41,631 hectares and 156,289 tonnes, respectively—in 2020/21. The major contributors are the southern Indian states of Karnataka, Tamil Nadu, Kerala, Andhra Pradesh, Telangana, and Maharashtra (Spice Board 2023). India also exports tamarind pulp, concentrate, and powder, and processed tamarind-based foods, to countries across the globe. In the coming years, the demand for tamarind in the international market is expected to be driven by the product's culinary versatility, beneficial health effects, use in vegan diet and products, industrial applications, and its association with the health and wellness industry (Chinnadurai et al. 2018). To meet the growing demand, crop improvement efforts have focused on developing superior cultivars with traits such as high pulp yield, drought tolerance, enhanced nutritional and fruit quality, and adaptability to different environments.

Molecular markers play a crucial role in plant breeding: they offer valuable tools for assessing genetic diversity, characterizing and conserving germplasm, facilitating marker-assisted selection, conducting parentage analysis, identifying cultivars, and supporting breeding programmes. However, in tamarind, molecular markers such as random amplified polymorphic DNA or RAPD (Shi-yu et al. 2008; Nyadoi et al. 2010; Yahya 2010; Algabal et al. 2011a; Gangaprasad et al. 2013; Kumar et al. 2015; Mayavel et al. 2020), inter simple sequence repeat or ISSR (Sarmiento et al. 2017), and amplified fragment length polymorphism or AFLP (Algabal et al. 2011b) have not received adequate attention, thus constituting a potential research gap. To the best of our knowledge, markers for start-codon-targeted (SCoT) polymorphism and sequence-related amplified polymorphism (SRAP) have not been used for analysing diversity in tamarind.

Both SCoT and SRAP markers are PCR-based techniques that offer several advantages. They are highly reproducible, designed to target specific DNA sequences, cost effective, and require minimal amounts of template DNA. Similar to AFLP markers, SCoT and SRAP markers exhibit high levels of polymorphism, enabling the detection of genetic variations within and between populations without the need for highly skilled personnel. Markers for SCoT, developed by Collard and Mackill (2009), are derived from short conserved sequences flanking the ATG start codon and use a single primer that serves as both the forward and the reverse primer. On the other hand, SRAP markers, developed by Li and Quiros (2001), amplify open reading frames (ORFs) and can use a wide range of primer combinations through unique pairings of forward and reverse primers. These markers have been effectively applied in breeding programmes for marker-assisted

selection, genetic mapping, and quantitative trait loci (QTL) analysis in fruit crops such as date palm (Saboori et al. 2020), mango (Zhou et al. 2020), sweet orange (Juibary et al. 2021), peach (Li et al. 2023), and olive (Partovi et al. 2020). Additionally, these markers play a crucial role in conservation genetics by helping to evaluate the genetic diversity and structure of wild populations and, in turn, to formulate effective strategies to conserve them. In our laboratory (Division of Fruit Crops, Indian Institute of Horticultural Research, Bengaluru), we have validated many such markers on several fruit crops including pomegranate (Shwetha et al. 2020), karonda (Kanupriya et al. 2019), avocado (Tripathi et al. 2020), and bael (Chaturvedi et al. 2023).

In an earlier study (Kanupriya et al. 2024), we examined phenotypic diversity in *T. indica* on a large scale—422 genotypes collected from eight states across India—using multivariate analysis and the following sixteen quantitative traits: total pod mass, pulp mass, shell mass, fibre mass, seed mass, pod length, pod breadth, pulp percentage, real value of pulp, titratable acidity, reducing and total sugar. All the traits differed significantly among the states. In the present study, we employed a combination of SRAP and SCoT markers to enhance the genetic analysis of tamarind. By using these two marker systems simultaneously, we sought to broaden the assessment of genetic diversity so as to obtain a more comprehensive understanding of the extent of genetic variation, population structure, and relationships among different tamarind genotypes. Because each of the two markers is based on a different underlying principle, the markers provide complementary information about different aspects of the genome, such as coding regions, promoter regions, and conserved non-coding sequences. Another advantage of using two independent marker systems is that if they produce consistent results, the results can be considered more reliable and accurate. Using different marker systems simultaneously makes it possible to compare their performance in terms of polymorphism, reproducibility, and ease of use and thus to select the most suitable markers for further research or breeding programmes, potentially saving time and resources by focusing on markers that provide the most informative and reliable results.

Materials and methods

Plant material

A total of 91 samples of *T. indica* were collected from major tamarind-growing states in India, including the southern states of Karnataka, Tamil Nadu, Telangana, and Andhra Pradesh; western states of Gujarat and Maharashtra; the central state of Madhya Pradesh; and the north-eastern state of Mizoram. Of the 91 genotypes, 63 were chosen based on in-situ evaluation of trees through a survey. The numbers of samples, the state and district (a district is an administrative unit within a state), and geographical coordinates for each sample are shown in Table 1. The remaining samples were collected as secondary collections from different institutions that maintain tamarind germplasm (Table 2). Because tamarind is highly heterozygous, only scions were collected and grafted on rootstock using softwood grafting to ensure that each genotype in the collection was true to type. Due to the unavailability of scion wood for 'Thailand,' seeds were used instead. Similarly, seeds of 'Lakshamana' were utilized to raise an open-pollinated progeny population. The collected genotypes were planted in a field repository at the ICAR- Indian Institute of Horticultural Research in Lakshmana, Bengaluru, Karnataka, India. The plants were spaced 5 × 5 m in the field (Fig. 1).

Molecular analysis

Extraction of genomic DNA and estimating its purity

Fresh, young, and healthy leaves were collected for DNA extraction. A total of 0.5 g of leaf tissue from each genotype was taken; the DNA was extracted using a modification of the CTAB (cetyl trimethyl ammonium bromide) method described by Doyle and Doyle (1990); and the concentration of the extracted DNA was determined using an ultraviolet (UV) absorption spectrophotometer (Eppendorf, Germany). The quality of the extracted DNA was assessed using gel electrophoresis (0.8% agarose gel), and the concentration of the DNA was adjusted to 50 ng/μL for the PCR (polymerase chain reaction) assay.

Table 1 Samples forming the primary collection (63 genotypes) of tamarind (*Tamarindus indica* L.)

Population	Sample no	ID	District	State	Geographic coordinates	
					Latitude	Longitude
MIZ	Z8	1	Aizawl	Mizoram	23°27′25.2″	92° 24′53.28″
MIZ	Z9	2	Aizawl	Mizoram	N 23° 27′ 25.2	E 92° 24′ 52.92
MIZ	Z10	3	Aizawl	Mizoram	N 23° 27′ 25.2	E 92° 24′ 52.92
MIZ	Z15	4	Aizawl	Mizoram	N 23° 27′ 25.2	E 92° 24′ 52.92
MIZ	Z16	5	Aizawl	Mizoram	N 23° 27′ 4.8	E 92° 24′ 39.6
MIZ	Z22	6	Aizawl	Mizoram	N 23° 27′ 6.84	E 92° 24′ 11.52
MIZ	Z23	7	Aizawl	Mizoram	N 23° 28′ 58.4	E 92° 23′ 32.4
MIZ	Z25	8	Aizawl	Mizoram	N 23° 27′ 6.48	E 92° 24′ 20.16
MIZ	Z26	9	Aizawl	Mizoram	N 23° 2′ 44.88	E 92° 24′ 20.16
MIZ	Z30	10	Aizawl	Mizoram	N 23° 27′ 46.08	E 92° 24′ 20.04
MIZ	Z31	11	Aizawl	Mizoram	N 23° 27′ 46.08	E 92° 24′ 20.04
MIZ	Z36	12	Aizawl	Mizoram	N 23° 27′ 46.08	E 92° 24′ 20.04
MIZ	Z40	13	Aizawl	Mizoram	N 23° 28′ 57.36	E 92° 18′ 32.76
MIZ	Z42	14	Aizawl	Mizoram	N 23° 27′ 46.08	E 92° 24′ 20.04
MIZ	Z43	15	Aizawl	Mizoram	N 23° 48′ 9.36	E 92° 45′ 31.68
MIZ	Z44	16	Aizawl	Mizoram	N 23° 48′ 9.72	E 92° 45′ 23.4
MIZ	Z45	17	Aizawl	Mizoram	N 23° 48′ 9	E 92° 45′ 48.24
MIZ	Z49	18	Aizawl	Mizoram	N 23° 48′ 8.28	E 92° 45′ 40.68
KAR	D11	39	Devanhalli	Karnataka	N 13° 14′ 37.32	E 77° 43′ 1.92
KAR	D10	40	Devanhalli	Karnataka	N 13° 14′ 37.32	E 77° 43′ 1.92
KAR	D8	41	Devanhalli	Karnataka	N 13° 14′ 37.32	E 77° 43′ 1.92
KAR	D6	42	Devanhalli	Karnataka	N 13° 14′ 37.32	E 77° 43′ 1.92
KAR	D5	43	Devanhalli	Karnataka	N 13° 14′ 37.32	E 77° 43′ 1.92
KAR	D4	44	Devanhalli	Karnataka	N 13° 14′ 37.32	E 77° 43′ 1.92
KAR	LK (Lakshamana)	45	Tumkuru	Karnataka	N 13° 19′ 59.16	E 77° 6′ 52.32
KAR	H1	29	Tumkuru	Karnataka	N 13° 19′ 58.08	E 77° 6′ 50.76
KAR	H2	30	Tumkuru	Karnataka	N 13° 19′ 59.4	E 77° 6′ 58.56
KAR	OP1	52	Tumkuru	Karnataka	N 13° 20′ 16.44	E 77° 7′ 2.28
KAR	OP2	53	Tumkuru	Karnataka	N 13° 20′ 16.44	E 77° 7′ 2.28
KAR	OP3	54	Tumkuru	Karnataka	N 13° 20′ 16.44	E 77° 7′ 2.28
KAR	OP4	55	Tumkuru	Karnataka	N 13° 20′ 16.44	E 77° 7′ 2.28
KAR	OP5	56	Tumkuru	Karnataka	N 13° 20′ 16.44	E 77° 7′ 2.28
KAR	OP6	57	Tumkuru	Karnataka	N 13° 20′ 16.44	E 77° 7′ 2.28
MAH	P17	58	Pune	Maharashtra	N 18° 12′ 34.2	E 74° 49′ 15.6
MAH	P18	59	Pune	Maharashtra	N 18° 12′ 34.92	E 74° 7′ 4.8
MAH	P2	60	Pune	Maharashtra	N 18° 12′ 32.76	E 74° 6′ 44.28
MPH	R8	61	Ratlam	Madhya Pradesh	N 23° 26′ 3.84	E 74° 35′ 42.72
MPH	R25	62	Ratlam	Madhya Pradesh	N 23° 26′ 40.92	E 74° 0′ 23.4
TEL	T9	63	Mancherial	Telangana	N 19° 14′ 42.72	E 79° 13′ 51.6
TEL	T10	64	Mancheria	Telangana	N 16° 43′ 28.92	E 79° 13′ 51.12
TEL	T13	65	Mahbubnagar	Telangana	N 16° 43′ 20.98	E 78° 27′ 16.56
TEL	T16	66	Mancheria	Telangana	N 19° 52′ 44.76	E 78° 33′ 46.8
TEL	T18	67	Mancheria	Telangana	N 19° 52′ 44.76	E 78° 33′ 46.8
TEL	T20	68	Mancheria	Telangana	N 19° 52′ 44.76	E 78° 33′ 46.8

Table 1 (continued)

Population	Sample no	ID	District	State	Geographic coordinates	
					Latitude	Longitude
TEL	T21	69	Mancheria	Telangana	N 19° 52′ 44.76	E 78° 33′ 46.8
TEL	T30	70	Adilabad rural	Telangana	N 19° 24′ 53.64	E 78° 17′ 39.12
TEL	T40	71	Adilabad rural	Telangana	N 19° 20′ 35.88	E 78° 17′ 36.6
TEL	T41	72	Adilabad rural	Telangana	N 19° 18′ 17.28	E 78° 22′ 22.8
TEL	T42	73	Adilabad rural	Telangana	N 19° 18′ 17.16	E 78° 22′ 15.6
TEL	T45	74	Mahbubnagar	Telangana	N 19° 13′ 31.08	E 78° 28′ 26.4
TEL	T34	75	Mahbubnagar	Telangana	N 16° 24′ 46.44	E 77° 32′ 45.6
TEL	T39	76	Mahbubnagar	Telangana	N 19° 13′ 31.08	E 78° 28′ 26.4
CHA	BG1	79	Bastar	Chhattisgarh	N 19° 15′ 23.04	E 81° 45′ 11.52
CHA	BG2	80	Bastar	Chhattisgarh	N 19° 13′ 30.72	E 81° 46′ 0.12
CHA	BG3	81	Bastar	Chhattisgarh	N 19° 13′ 1.56	E 81° 27′ 57.6
CHA	B1	82	Bastar	Chhattisgarh	N 19° 6′ 18.72	E 81° 57′ 50.04
CHA	B5	83	Bastar	Chhattisgarh	N 19° 8′ 7.08	E 81° 48′ 34.92
CHA	B12	84	Bastar	Chhattisgarh	N 19° 0′ 36.72	E 81° 29′ 36
CHA	B18	85	Bastar	Chhattisgarh	N 18° 33′ 55.08	E 82° 0′ 11.88
CHA	B22	86	Bastar	Chhattisgarh	N 19° 3′ 45.36	E 82° 5′ 31.2
CHA	B13	87	Bastar	Chhattisgarh	N 18° 31′ 25.32	E 81° 2′ 25.8
CHA	B19	88	Bastar	Chhattisgarh	N 18° 33′ 46.8	E 82° 1′ 18.48
CHA	B21	89	Bastar	Chhattisgarh	N 18° 34′ 52.32	E 82° 2′ 45.24

PCR amplification

A total of 26 SCoT (Collard and Mackill 2009) and 20 SRAP primers (Li and Quiros 2001) were initially screened for polymorphism in tamarind genotypes. After screening, 10 each of the SCoT and 10 SRAP primers exhibiting the desired characteristics in terms of banding patterns and polymorphism were chosen for further analysis. The specific details of the selected markers are given in Supplementary Table 1. Standard conditions for PCR assays of the markers were maintained as suggested by Collard and Mackill (2009) for SCoT and by Li and Quiros (2001) for SRAP. The resulting product of the PCR was analysed on 2% agarose gel with ethidium bromide staining in 1×TBE (tris–borate-EDTA) buffer, and a 100 bp DNA ladder (ThermoFisher Scientific) was used as the marker of fragment size. To make the DNA bands visible, the agarose gel was observed under UV light using a gel documentation system (UVI-TEC, Cambridge). Each SCoT primer or each SRAP primer pair was considered as one genetic marker, and data sheets were prepared by filling up the matrix data, namely the presence or absence of a band, recorded as (1) or (0) respectively. Faint or unclear bands were excluded from the analysis. To estimate band size, a medium-range DNA ruler (500 bp) was run alongside the amplified products.

Analysis of markers

The polymorphic information content (PIC) value for each SCoT or SRAP marker was determined using the following formula given by Roldan-Ruiz et al. (2000): $PICi = 2fi\ (1 - fi)$, in which fi represents the frequency of the marker bands present whereas $(1 - fi)$ represents the frequency of the marker bands absent. Serrote et al. (2020) have proposed the following classification on the informativeness for dominant markers, based on PIC values: low (0–0.10), medium (0.10–0.25), high (0.30–0.40), and very high (0.40–0.50). Additionally, other parameters such as expected heterozygosity (H) and discriminating power (D) were calculated using the Marker Efficiency Calculator (iMEC) software, as described by Amiryousefi et al. (2018).

Principal component analysis (PCA) and phylogenic and Bayesian cluster analysis were performed

Table 2 Samples forming the secondary collection (28 genotypes) of tamarind (*Tamarindus indica* L.)

Population	Sample name	ID	Institute	District	State	Unique features
TN	GTBR1	19	Institute of Forest Genetics and Tree Breeding	Coimbatore	Tamil Nadu	Red coloured pulp during unripe stage, high anthocyanin content
TN	GTBR2	20	Institute of Forest Genetics and Tree Breeding	Coimbatore	Tamil Nadu	Red coloured pulp during unripe stage, high anthocyanin content
TN	GTBR3	21	Institute of Forest Genetics and Tree Breeding	Coimbatore	Tamil Nadu	Red coloured pulp during unripe stage, high anthocyanin content
TN	GTBR4	22	Institute of Forest Genetics and Tree Breeding	Coimbatore	Tamil Nadu	Red coloured pulp during unripe stage, high anthocyanin content
TN	GTBS1	23	Institute of Forest Genetics and Tree Breeding	Coimbatore	Tamil Nadu	Pulp has high sucrose content
TN	GTBS2	24	Institute of Forest Genetics and Tree Breeding	Coimbatore	Tamil Nadu	Pulp has high sucrose content
TN	GTBS3	25	Institute of Forest Genetics and Tree Breeding	Coimbatore	Tamil Nadu	Pulp has high sucrose content
TN	GTBS4	26	Institute of Forest Genetics and Tree Breeding	Coimbatore	Tamil Nadu	Pulp has high sucrose content
TN	PKM1	27	Tamil Nadu Agricultural University	Periyakulam	Tamil Nadu	High pulp recovery, high in tartaric (17%) and ascorbic acid
TN	URI (Urigam)	28	Tamil Nadu Agricultural University	Periyakulam	Tamil Nadu	High yield (265 kg/tree)
KAR	N10	31	Forest Department Nursery	Dharwad	Karnataka	High pod and pulp yield
KAR	D1	32	Forest Department Nursery	Dharwad	Karnataka	Cluster bearing, high yields
KAR	SMG4	33	Forest Department Nursery	Dharwad	Karnataka	Cluster bearing, straight pods, heavy yielder
KAR	DS2 (Dharwad Sel 2)	34	Forest Department Nursery	Dharwad	Karnataka	Long and straight pods, high yield
KAR	DS1 (Dharwad Sel1)	35	Forest Department Nursery	Dharwad	Karnataka	Long and curved pods, high yield
KAR	NTI84	36	Forest Department Nursery	Dharwad	Karnataka	Small pods, high pulp recovery
KAR	NTI32	37	Forest Department Nursery	Dharwad	Karnataka	High pod and pulp yield
KAR	NFN10	38	Forest Department Nursery	Dharwad	Karnataka	High pod and pulp yield
KAR	D2	46	Forest Department Nursery	Dharwad	Karnataka	High pod and pulp yield
KAR	NFN8	47	Forest Department Nursery	Dharwad	Karnataka	High pod and pulp yield
KAR	NFN3	48	Forest Department Nursery	Dharwad	Karnataka	High pod and pulp yield
KAR	NFN6	50	Forest Department Nursery	Dharwad	Karnataka	High pod and pulp yield
KAR	NFN5	51	Forest Department Nursery	Dharwad	Karnataka	High pod and pulp yield

Table 2 (continued)

Population	Sample name	ID	Institute	District	State	Unique features
KAR	GKVK Sel	49	Department of Forestry, GKVK, UAS	Bengaluru	Karnataka	High pod and pulp yield
AP	AR (Anantha Rudhira)	77	Horticulture Research Station	Anantapur	Andhra Pradesh	Red coloured pulp during unripe stage, high anthocyanin content
AP	AA (Anantha Amalika)	78	Horticulture Research Station	Anantapur	Andhra Pradesh	High pod and pulp yield
GUJ	GP (Goma Pratik)	90	Central Institute of Arid Horticulture Regional Centre	Godhra	Gujrat	Long pod size, TSS is 71 ^{0}Brix and yield of 58.50 kg/tree (9th year of planting)
THAI	THAI (Thailand)	91			Thailand	Sweet pulp, low acidity

Fig. 1 Sites from which scion wood was collected site (**a**–**e**, respectively for Madhya Pradesh, Mizoram, Chhattisgarh, Telangana, and Karnataka); field gene bank for tamarind at IIHR, Bengaluru **f**; close-up of a genotype in field gene bank **g**

using the software package R ver. 4.0.4. For analysing the structure of the population, the data on the genotypes as obtained from both the markers were fed into the software package STRUCTURE ver. 2.3.4 (Pritchard et al. 2000). The data were evaluated for potential populations (K = 1 – 10) with three replications. The burn-in period was set to 10,000 lengths, followed by 100,000 Markov Chain Monte Carlo (MCMC) replicates (Pritchard et al. 2000). The most appropriate number of subpopulations (K) was determined using the log-probability [LnP(D)] estimate for each K, and the number of subpopulations was inferred based on K values (Evanno et al. 2005). The final population structure was determined using the web-based program 'Structure Harvester' and plotting the Ln P(D) values against the ΔK values to select

the best K value (Evanno et al. 2005). A membership probability of 0.80 or higher was considered adequate for assigning a given accession to one of the pure-population groups, whereas accessions with scores less than 0.80 were considered to be admixtures.

Results

Molecular diversity and cluster analysis based on SCoT loci

Of the 26 SCoT markers screened initially, 10 were chosen (28.5%) that produced clear, reproducible polymorphic bands. Molecular analysis revealed a total of 60 bands, from 3 (in SCoT- 3 and SCoT-33) to 9 (in SCoT- 31). The average number of polymorphic bands was 6, and polymorphism was 100% (Table 3). Genetic diversity within the studied population was assessed using PIC values. Among the 10 chosen primers, SCoT-24 recorded the highest PIC value (0.45), followed by SCoT- 3 and SCoT- 31, with a PIC value of 0.43 each. On the other hand, SCoT-16, SCoT-18, and SCoT-32 recorded the lowest PIC value (0.38; Table 4). The mean PIC value of all the 10 primers was 0.40. The average values of H (expected heterozygosity) and D (discriminating power) were 0.44 and 0.54, respectively. The highest H value (0.50) was seen in SCoT-16 and SCoT-18 and the lowest (0.32), in SCoT-24. On the other hand, the D value of primers, which represents the efficiency of a marker to distinguish between closely related genotypes, ranged from 0.14 (in SCoT-24) to 0.83 (in SCoT-32).

The data generated by SCoT markers were used to assess the dissimilarity matrix and to construct a dendrogram using the unweighted neighbor-joining (N-J) method. Based on genetic similarity, the N-J tree grouped the genotypes into four major clusters, numbered I to IV (Fig. 2), each of which was subdivided into two sub-clusters, namely a and b. Cluster I comprised 41 members: sub-cluster Ia consisted of 38 of genotypes, including the genotype from Thailand and commercial cultivars such as Goma Prateek (from Gujarat), Urigam (from Tamil Nadu), and GKVK selection and NFN8 (from Karnataka); sub-cluster Ib consisted of 3 genotypes, including NFN5 from Karnataka. Cluster II comprised 22 members: sub-cluster IIa consisted of 20 genotypes, including 6 secondary collections from Tamil Nadu, one of which was PKM1, the most popular commercial cultivar of tamarind in India, and 3 from Karnataka; sub-cluster IIb consisted of 2 secondary collections from Karnataka and 1 from Andhra Pradesh. Cluster III comprised 10 genotypes and cluster IV, 23 genotypes; the latter included the farmers' selection LK (Lakshamana) from Tumkur district of Karnataka.

Table 3 Details of polymorphism shown by start-codon-targeted (SCoT) markers used in the study

Primer code	Total number of bands	Number of polymorphic bands	Polymorphic information content	Expected heterozygosity H	Discriminating power D
SCOT3	3	3	0.43	0.37	0.44
SCOT9	8	8	0.40	0.45	0.57
SCOT13	8	8	0.39	0.47	0.60
SCOT16	7	7	0.38	0.50	0.40
SCOT18	5	5	0.38	0.50	0.74
SCOT20	6	6	0.39	0.47	0.62
SCOT24	6	6	0.45	0.32	0.14
SCOT31	9	9	0.43	0.40	0.39
SCOT32	5	5	0.38	0.48	0.83
SCOT33	3	3	0.39	0.47	0.62
Mean	6	6	0.40	0.44	0.54
Range	3–9	3–9	0.38–0.45	0.32–0.50	0.14–0.83

All the ten primers showed 100% polymorphism

Table 4 Details of polymorphism shown by sequence-related amplified polymorphism (SRAP) markers used in the study

Primer code	Total number of bands	Number of polymorphic bands	Polymorphic information content	Expected heterozygosity H	Discriminating power D
me3-em3	3	3	0.38	0.5	0.33
me3-em6	3	3	0.43	0.37	0.94
me6-em3	3	3	0.40	0.47	0.27
me6-em8	4	4	0.40	0.44	0.56
me8-em4	3	3	0.40	0.45	0.89
me10-em6	3	3	0.44	0.34	0.33
me10-em7	4	4	0.38	0.49	0.67
me11-em6	8	8	0.44	0.37	0.28
me12-em9	5	5	0.45	0.32	0.96
me13-em4	3	3	0.40	0.45	0.31
Mean	3.9	3.9	0.41	0.42	0.55
Range	3–8	3–8	0.38–0.45	0.32–0.50	0.27–0.96

All the ten primers showed 100% polymorphism

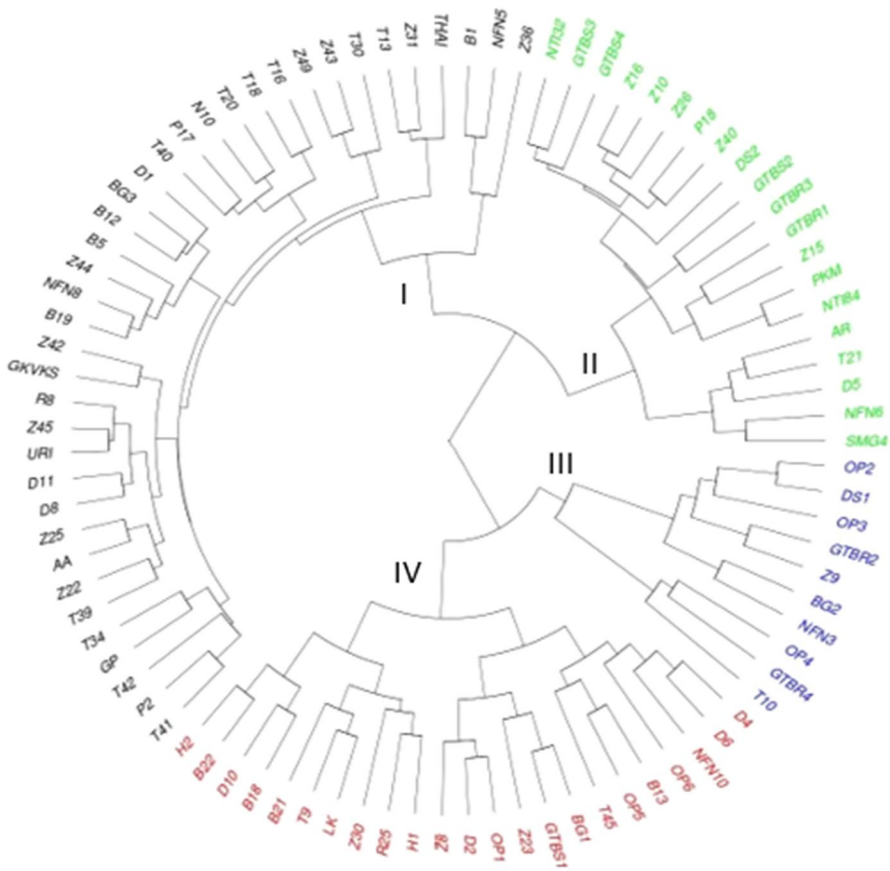

Fig. 2 Dendrogram based on 10 start-codon-targeted (SCoT) primers for 91 genotypes of tamarind

Molecular diversity and cluster analysis based on SRAP loci

Similarly, out of 20 SRAP primer combinations (PC) screened, 10 PCs producing clear, polymorphic bands were selected. A total of 39 DNA fragments were amplified from the 91 samples that were tested using the 10 SRAP PCs. The number of amplified bands ranged from 3 to 8, averaging 3.9 polymorphic band per sample, accounting for 100% polymorphism for all PCs tested in the study. Among all the PCs implemented, the combination me12-em9 exhibited the highest PIC value of 0.45, closely followed by 2 PCs, me11-em6 and me10-em6, with PIC of 0.44 each. Two primers, me10-em7 and me3-em3, registered lowest PIC value of 0.38 (Table 3). The mean PIC value across all 10 PCs was 0.41. The average values for H (expected heterozygosity) and D (discriminating power) were 0.42 and 0.55 respectively. The highest H value (0.50) was revealed by the PC me3-em3, whereas the PC me12-em9 had the lowest H value (0.32). The D value of PCs ranged from 0.27 (me6-em3) to 0.96 (me12-em9).

The dendrogram for 10 SRAP PCs (Fig. 3) also consisted of four major clusters, which were numbered and labelled as mentioned above. Cluster I comprised 8 members: sub-cluster Ia consisted of 6 genotypes, including commercial cultivars such as Goma Prateek (Gujarat), Urigam (Tamil Nadu), and Anantha Amalika (Andhra Pradesh); sub-cluster Ib consisted of 2 genotypes. Cluster II comprised 25 members: sub-cluster IIa consisted of 23 genotypes, including 5 secondary collections from Tamil Nadu and Anatha Rudhira (a commercial cultivar, with red pulp, from Andhra Pradesh); sub-cluster IIb consisted of 2 collections from Mizoram. Cluster III comprised 12 members and cluster IV, 46 members, including

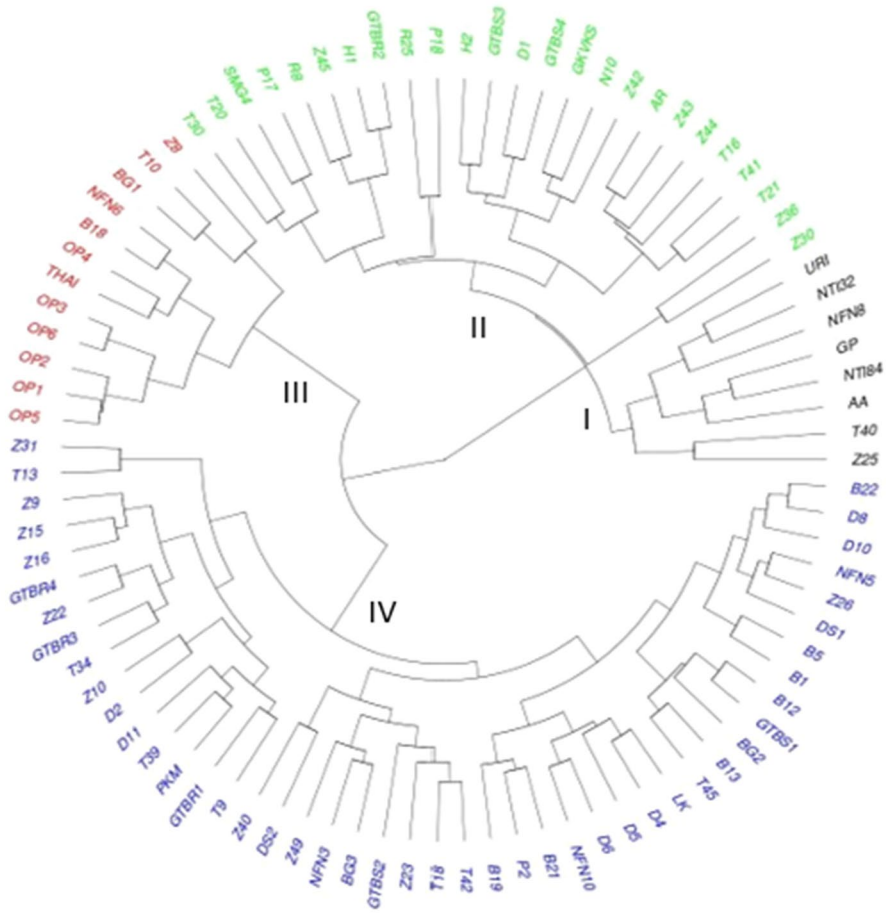

Fig. 3 Dendrogram based on 10 sequence-related amplified polymorphism (SRAP) primers for 91 genotypes of tamarind

the farmers' selection LK (Lakshamana) from Karnataka and PKM1 from Tamil Nadu.

Analysis of population structure

Combined binary data on SCoT and SRAP markers were used for determining the population structure and associations among the 91 genotypes. The Bayesian-based model program STRUCTURE was used to gain a better understanding of these associations. The analysis considered a range of population numbers, from K = 1 to K = 9. The highest ad hoc measure of delta K was achieved at K = 4 (Fig. 4a); accordingly, the 91 genotypes were divided into four distinct populations, P1, P2, P3, and P4, each of which was subdivided into two groups, pure and admixture, based on the membership fraction of each accession: accessions with a probability score of 0.80 or higher were placed in the pure group and those with a score below 0.80, into the admixture group. Of the total, 39 were placed in the former group and 52, in the latter. Specifically, P1 comprised 10 pure genotypes and 18 admixtures: Urigam, GKVK Selection, and Goma Pratik were the pure genotypes and Anantha Amalika and SMG4 were admixtures. In P2, comprising 8 pure types and 10 admixtures, PKM1, Dharwad Selection 2, and Anantha Rudhira were the pure genotypes. Similarly, P3 and P4 comprised 12 and 9 pure genotypes, along with 11 and 13 admixtures, respectively. The genotype from Thailand was included in P3 as a pure genotype, and P4 included Dharwad Selection 1, which was a pure genotype, and farmers' selection Lakshamana (Kanupriya et al. 2020), which was an admixture (Fig. 4b).

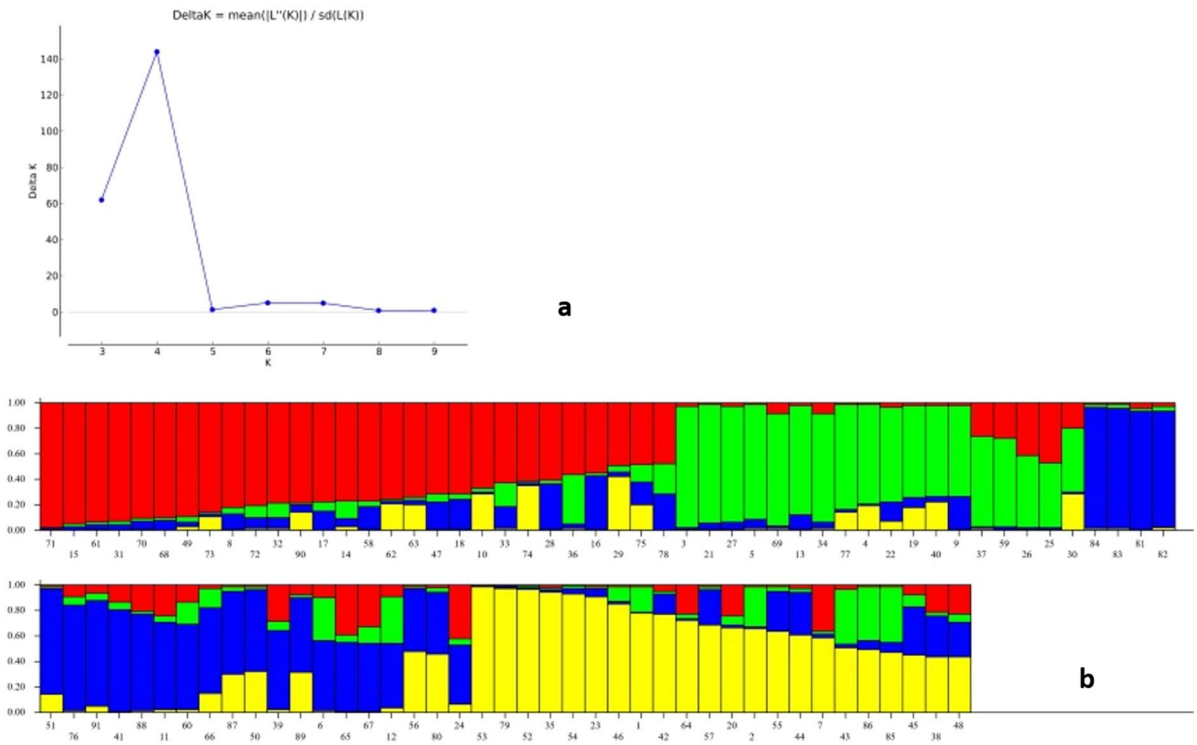

Fig. 4 **a** Estimation of population using LnP(D)-derived Δk for K from 1 to 10 for different populations of tamarind using combined start-codon-targeted (SCoT) and sequence-related amplified polymorphism (SRAP) marker data. The statistic Δk is used for determining the optimal number of populations (K) in the data set; it measures the rate of change in the log probability of the data between successive values of K. **b** Model-based clustering of 91 genotypes into four populations using combined marker data: Population 1, red bars; Population 2, green bars; Population 3, blue bars; Population 4, yellow bars. The numbers below each set of bars correspond to the names of genotypes listed in Tables 1 and 2

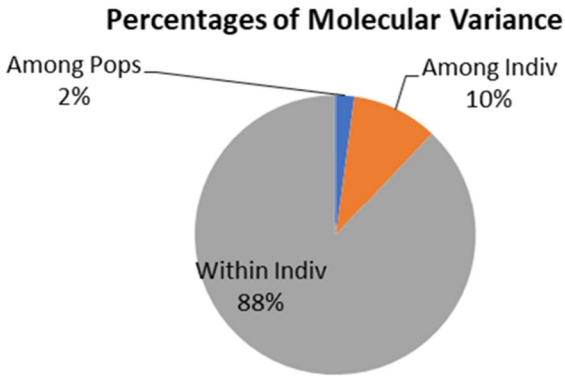

Fig. 5 Shares (%) of difference sources of molecular variance seen in 91 tamarind genotypes based on combined data from start-codon-targeted (SCoT) and sequence-related amplified polymorphism (SRAP) primers

Analysis of molecular variance and principal component analysis

The analysis of molecular variance based on combined data on SCoT and SRAP markers attributed only 2% of the genetic variation to differences between populations; 10%, to differences among individuals; and as much as 88%, to differences among individual genotypes. These findings are summarized in Supplementary Table 2 and illustrated in Fig. 5. The cumulative variation explained by the top three principal components (PCs) and their corresponding eigenvalues are presented in Supplementary Table 3. Together, the three PCs accounted for 54.2% of the total variation. The primary principal component had the highest influence and explained 31.8% of the variation, followed by the second principal component, which explained 12.4% of the variation. The eigenvalues associated with these components were 22.52 for the first axis, 8.77 for the second axis, and 7.06 for

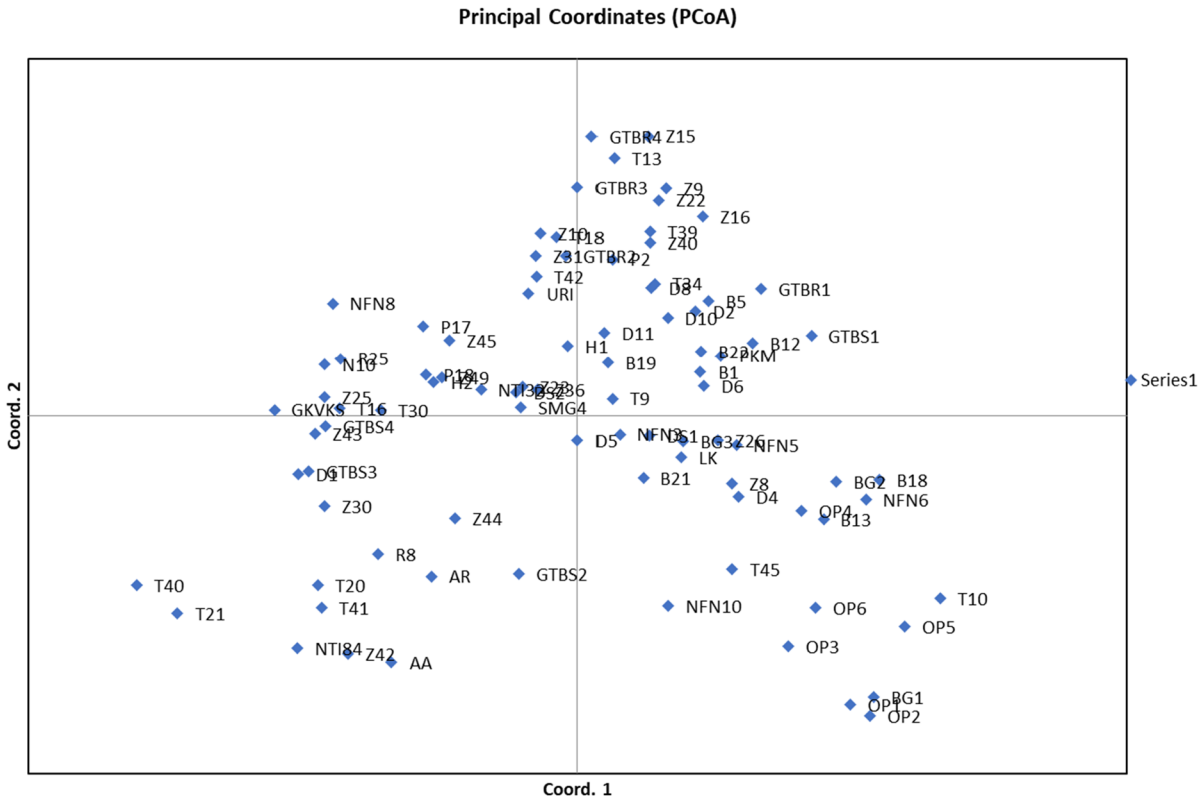

Fig. 6 Clustering in 91 tamarind genotypes into four groups based on principal coordinate analysis of combined data on start-codon-targeted (SCoT) and sequence-related amplified polymorphism (SRAP) primers

the third axis. Based on the PCA biplot analysis, all 91 tamarind genotypes were grouped into four clusters, consistent with the results obtained from the N-J tree-based clustering (Fig. 6).

Discussion

Despite the economic and ecological importance of tamarind, comprehensive research of its genetic diversity is rare, and that lack of information hinders greater exploitation of this crop for various applications, including crop improvement, sustainable cultivation, and the development of targeted products. Although tamarind did not originate in India, the country not only hosts a rich diversity of tamarind genotypes but also enjoys a rich heritage of traditional knowledge of the many uses of this versatile tree and its cultivation. Over centuries, farmers and local communities in different regions of India have selectively grown and conserved diverse tamarind genotypes, resulting in wide variations in tree morphology, fruit characteristics, and adaptability to various ecological conditions. Such selection, combined with diverse agro-climatic zones of the country ranging from arid to semi-arid and tropical regions, has favoured the evolution and adaptation of tamarind genotypes with distinct traits. These variations relate to fruit size, pulp quality, taste, colour, and seeds and have been well documented. However, it is important to acknowledge that phenotypic traits can be influenced significantly by environmental factors, and morphological traits alone will be seldom adequate for assessing genetic diversity. In response to this limitation, diversity analysis based on molecular markers has emerged as a more accurate and reliable approach (Idrees and Irshad 2014).

In the field of plant germplasm resources, researchers commonly use either SRAP or SCoT; however, we used both and thereby benefitted from their complementary effects. Of the 46 combined markers we screened, 20 displayed good amplification and showed 100% polymorphism, making them on par with 100 and 95.56% polymorphism recorded by SCoT and SRAP markers respectively in 33 accessions of *Paris polyphylla* (Zhao et al. 2020). Similar polymorphism values for tamarind were reported earlier: 90% with 12 ISSR markers on 32 trees by Sarmiento et al. (2017); 89% with 12 AFLP markers on 36 genotypes by Algabal et al. (2011a, b); 79 with 12 RAPD markers on 36 genotypes by Algabal et al. (2011a, b); 88 with 13 RAPD on 28 genotypes by Gangaprasad et al. (2013); and 82% with 15 RAPD on 9 genotypes by Mayavel et al. (2020). A quantitative measure of genetic variation within a population or among different genotypes is PIC: the higher the PIC value, the greater the degree of polymorphism and the larger the pool of genetic diversity as well as the informativeness and discriminatory power of a specific marker locus. In the present study, PIC values for both SCoT and SRAP markers ranged from 0.38 to 0.45 (mean of 0.40), higher than the value of 0.25 obtained by Sarmiento et al. (2017) using ISSR markers and that of 0.31 obtained by Mayavel et al. (2020) using RAPD markers. Four SCoT primers and eight SRAP primers recorded PIC values greater than 0.40, indicating them to be efficient and robust markers. Among the two types, SRAP markers outperformed SCoT markers in capturing a wide range of genetic diversity within the population, suggesting that SRAP markers are more useful for studying population diversity or genetic relationships among tamarind genotypes. A high discrimination power (D) of a marker indicates its ability to differentiate between different genotypes or individuals in a population: again, three SRAP markers and one SCoT marker expressed D values greater than 0.8, making them valuable for applications such as identification of individuals or genetic fingerprinting.

Despite considerable differences in their geographical origin, the 91 genotypes fell into four distinct clusters, thereby underscoring the robustness and significance of genetic factors, compared to geographical proximity, in determining population structure. Similar findings were reported by Algabal et al. (2011a, b), Gangaprasad et al. (2013), and Mayavel et al. (2020) using RAPD markers and by Sarmiento et al. (2017) using ISSR markers. The similarities are attributed to the highly cross-pollinated nature of the crop, gene drift, seed dispersal by animals and human intervention in dispersal, and cultivation of tamarind.

Valuable insights could be obtained by incorporating molecular markers and model-based approaches into the genetic make-up and admixture patterns of populations. Such insights are particularly crucial to preserving the genetic diversity of perennial woody trees (Ferreira et al. 2016). In the present investigation, the combined analysis of

SCoT and SRAP markers using STRUCTURE provided valuable insights into the population structure of 91 genotypes, which fell into four distinct populations, each with varying proportions of pure genotypes and admixtures. Of these, 39 were identified as pure genotypes with a probability score of 0.80 or higher, whereas the remaining 59 were admixtures (probabilities lower than 0.80). It is noteworthy that the study did not find any clear correspondence between geographical origins of the genotypes and the observed population structure—a finding consistent with that reported by Ganesan et al. (2014), who classified 300 accessions of moringa, the drumstick tree (*Moringa oleifera*), into five groups that showed no strong geographical correlation. Similar results have also been reported in other species, indicating that genetic divergence should not be attributed solely to geographical distance (Nosil and Feder 2012). Overall, the present study showed low genetic divergence—amounting to only 2%—among the tamarind populations, suggesting that genetic differences between the populations were relatively small compared to the overall genetic variation across the entire sample. On the other hand, individual trees within each population showed significant genetic differences (10%). This within-population variation could be due to various factors such as genetic drift, mutation, and recombination. However, the largest share (approximately 88%) of genetic variation was found within individuals, implying that the greatest source of genetic diversity was neither different populations nor even individuals within a given population but individual trees. These percentages provide valuable insights into the genetic structure and diversity of tamarind populations, underscoring the importance of recording and conserving within-population differences while also acknowledging the differences at the population level. Most cross-pollinated species typically exhibit a pattern of greater genetic diversity within populations and relatively low diversity among populations, as reported in moringa (Popoola et al. 2019), mango (Ravishankar et al. 2015), apple (Kumar et al. 2019), etc. In the present study, the first three PCs together explained 54% of the total variation, thereby corroborating the N-J-based cluster analysis, which divided the 91 genotypes into four clusters irrespective of the geographical origins of the genotypes.

Conclusion

Genetic diversity of 91 tamarind genotypes sourced from locations across India was assessed. Using two different marker systems, SCoT and SCAR, distinct clusters and differentiated populations were identified, highlighting the genetic uniqueness present among tamarind genotypes. The present research provides crucial foundational information for deploying molecular approaches to conservation and sustainable utilization of tamarind, a valuable tree crop.

Acknowledgements We are thankful to the Indian Council of Agricultural Research and to the Director, ICAR-IIHR, for funding this study and are particularly grateful to the following entities for assistance in conducting surveys and for sparing valuable germplasm: Forest Departments of the states of Chhattisgarh and Karnataka, National Bureau of Plant Genetic Resources (Hyderabad Centre), KVKs of Aizawl in Mizoram and Ratlam in Madhya Pradesh, BAIF Development Research Foundation, Pune, and Institute of Forest Genetics and Tree Breeding, Coimbatore.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability We confirm that all data pertaining to the paper has been included in the manuscript, ensuring transparency and reproducibility of the findings.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Algabal A, Narayanaswamy P, Simon L (2011a) Estimation of genetic variability in tamarind (*Tamarindus indica*) using RAPD markers. Int J Plant Breeding 5(1):10–16

Algabal A, Narayanaswamy P, Simon L (2011b) Amplified fragment length polymorphism marker-based genetic diversity in tamarind (*Tamarindus indica*). Int J Fruit Sci 11:1–16

Amiryousefi A, Hyvönen J, Poczai P (2018) iMEC: online marker efficiency calculator. Appl Plt Sci 6(6):1–5

Bhadoriya SS, Ganeshpurkar A, Narwaria J, Rai G, Jain AP (2011) *Tamarindus indica*: extent of explored potential. Pharmacognosy Rev 5(9):73

Spice Board (2023) accessed via http://indianspices.com/about-us/overview-about-us.html

Chacón-Fernández MG, Hernández-Medel MR, Bernal-González M, Durán-Domínguez-de-Bazúa MC, Solís-Fuentes JA (2019) Composition, properties, stability and

thermal behavior of tamarind (*Tamarindus indica*) seed oil. Grasasy Aceites 70(4):e333–e333
Chaturvedi K, Kumara U, Sane A, Singh P, Kumar P, Tripathi PC (2023) Exploring the genetic diversity of *Aegle marmelos* (L.) Correa populations in India. Plt Genet Res 21(2):107–114
Chinnadurai M, Varadharaj S, Chitra R (2018) production, marketing and value chain of tamarind in India. Tamarind science and technology. Scientific Publishers, UK, pp 116–132
Collard BC, Mackill DJ (2009) Start codon targeted (SCoT) polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. Plt Mol Bio Reporter 27:86–93
Dorni AC, Amalraj A, Gopi S, Varma K, Anjana SN (2017) Novel cosmeceuticals from plants—An industry guided review. J Appl Res Med Aromatic Plts 7:1–26
Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15
Evano G, Regnuat S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14(8):2611–2620
Ferreira V, Ramos-Cabrer AM, Carnide V, Pinto-Carnide O, Assunção A, Marreiros A, Rodrigues R, Pereira-Lorenzo S, Castro I (2016) Genetic pool structure of local apple cultivars from Portugal assessed by microsatellites. Tree Genet Genomes 12(3):36
Ganesan S, Singh R, Choudhury DR, Bharadwaj J, Gupta V, Singode A (2014) Genetic diversity and population structure study of drumstick (*Moringa oleifera* Lam.) using morphological and SSR markers. Indust Crops Prod 60:316–325
Gangaprasad S, Rajkumar R, Ravikumar RL, Savita MH, Krishnamurthy K, Hittalamani S (2013) Genetic diversity analysis in tamarind (*Tamarindus indica* L.). J Spices Aromatic Crops 22(1):55–61
Idrees M, Irshad M (2014) Molecular markers in plants for analysis of genetic diversity: a review. Eur Acad Res 2(1):1513–1540
Juibary PL, Seyedmehdi FS, Sheidai M, Noormohammadi Z, Koohdar F (2021) Genetic structure analysis and genetic finger printing of sweet orange cultivars (*Citrus sinensis* (L.) Osbeck) by using SCoT molecular markers. Genet Res Crop Evol 68:1645–1654
Kanupriya C, Tripathi PC, Singh P, Venugopalan R, Radhika V (2019) Analysis of morphological, biochemical and molecular diversity in karonda (*Carissa carandas* L.) germplasm. Fruits 74(3):130–140
Kanupriya C, Karunakaran G, Singh P (2020) A promising new tamarind selection-Lakshamana: linking biodiversity with livelihood. J Hort Sci 15(2):221–224
Kanupriya C, Karunakaran G, Singh P, Venugopalan R, Samant D, Prakash K (2024) Phenotypic diversity in Tamarindus indica L. sourced from different provenances in India. Agrofor Syst 98(2):477–490
Kumar M, Ponnuswami V, Rajamanickam C, Preethi TL (2015) Assessment of genetic diversity in tamarind (*Tamarindus indica* L.) using random amplified polymorphic DNA markers. SAARC J Agriculture 13(1):27–36
Kumar C, Singh SK, Singh R, Pramanick KK, Verma MK, Srivastav M, Choudhury DR (2019) Genetic diversity and population structure analysis of wild Malus genotypes including the crabapples (M. baccata (L.) Borkh. & M. sikkimensis (Wenzig) Koehne ex C. Schneider) collected from the Indian Himalayan region using microsatellite markers. Genetic Res Crop Evol 66:1311–1326
Li G, Quiros CF (2001) Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica. TAG 103:455–461
Li W, Ma Y, Kou Y, Zeng Z, Qiu D, Ma C (2023) Analysis of genetic diversity and relationships between late-mature peach (*Prunus persica* L.) varieties assessed with ISSR and SRAP markers. Fruit Res 3:36
Mansingh BB, Binoj JS, Sai NP, Hassan SA, Siengchin S, Sanjay MR, Liu YC (2021) Sustainable development in utilization of *Tamarindus indica* L. and its by-products in industries: a review. Current Res Green Sustain Chem 4:100207
Mayavel A, Vikashini B, Bhuvanam S, Shanthi A, Kamalakannan R, Kim KW, Kang KS (2020) RAPD polymorphism and genetic distance among phenotypic variants of *Tamarindus indica*. J Korean Soc Forest Sci 109(4):421–428
Nosil P, Feder JL (2012) Genomic divergence during speciation: causes and consequences. Philos Trans R Soc b: Biol Sci 367(1587):332–342
Nyadoi P, Jamnadass R, Okori P, Okullo JBL, Obua J, Nasoro M, Burg K (2010) *Tamarindus indica* tropical populations genetic structure. Gene Conserve 9(37):152–166
Partovi R, Iranbakhsh A, Sheidai M, Ebadi M (2020) Population genetic studies in wild olive (Olea cuspidata) by molecular barcodes and SRAP molecular markers. Caryologia 73(1):125–132
Popoola J, Lgwe D, Jegede O, Iwu V, Adegbite A, Omonhinmin C (2019) Agronomic practices, genetic diversity and population structure of *Moringa oleifera* (Lam.) in Nigeria. Biochem Molec Biol 45:365. https://doi.org/10.20944/preprints201910.0209.v1
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2):945–959
Raja MM, Poonguzhali S, Rao BV, Choudhury A, Devarajan A (2022) A scientific evidence-based review of Tamarind usage in Indian Folklore Medicine. J Nat Rem 2022:347–362
Ravishankar KV, Bommisetty P, Bajpai A, Srivastava N, Mani BH, Vasugi C, Rajan S, Dinesh MR (2015) Genetic diversity and population structure analysis of mango (*Mangifera indica*) cultivars assessed by microsatellite markers. Trees 29:775–783
Roldan-Ruiz I, Dendauw J, Van Bockstaele E, Depicker A, De Loose MA (2000) AFLP markers reveal high polymorphic rates in ryegrasses (Lolium spp.). Mol Breeding 6:125–134
Saboori S, Noormohammadi Z, Sheidai M, Marashi S (2020) SCoT molecular markers and genetic fingerprinting of date palm (*Phoenix dactylifera* L.) cultivars. Genet Res Crop Evol 67:73–82
Sandesh P, Velu V, Singh RP (2014) Antioxidant activities of tamarind (*Tamarindus indica*) seed coat extracts using in vitro and in vivo models. J Food Sci Technol 51:1965–1973. https://doi.org/10.1007/s13197-013-1210-9

Sarmiento L, Pérez-Almeida I, Díaz B, Álvarez H, Viera W (2017) Molecular marker-based characterization of Ecuadorian dry forest tamarind plus trees. Bioagro 29(3):153–162

Serrote CML, Reiniger LRS, Silva KB, dos Santos Rabaiolli SM, Stefanel CM (2020) Determining the Polymorphism Information Content of a molecular marker. Gene 726:144175

Shi-yu Y, Cheng-zhong H, Yi-he ZH, Juan L, Bing-yi W (2008) Analysis of the genetic diversity of sweet Thai tamarind introduced from Thailand based on RAPD markers. 林业科学研究 21(1): 60–63.

Shwetha HK, Chaturvedi K, Rao AVD, Murthy BNS, Radhika V (2020) Analysis of genetic diversity in pomegranate using SRAP (Sequence-related amplified polymorphism) markers. Indian J Plt Genet Res 33(2):161–171

Thombare N, Srivastava S, Chowdhury AR (2014) Multipurpose application of tamarind seed and kernel powder. Sci Reporter 2014:32–33

Tripathi PC, Kanupriya C, Karunakaran G, Shamina A, Singh P, Radhika V (2020) Variability studies in avocado (Persia americana Mill.) using physico-chemical properties. Fruits 75(3):115–122

Yahya AQA (2010) RAPD and AFLP Markers Based DNA Fingerprinting and Diversity Studies in Tamarind (*Tamarindus indica* L). University of Agricultural Sciences, Banglore

Zhao X, Zou G, Zhao J, Hu L, Lan Y, He J (2020) Genetic relationships and diversity among populations of Paris polyphylla assessed using SCoT and SRAP markers. Phy Mol Biol Plt 26:1281–1293

Zhou L, He XH, Yu HX, Chen MY, Fan Y, Zhang XJ, Luo C (2020) Evaluation of the genetic diversity of mango (*Mangifera indica* L.) seedling germplasm resources and their potential parents with start codon targeted (SCoT) markers. Genet Res Crop Evol 67:41–58