PLANT SCIENCE TODAY
ISSN 2348-1900 (online)
Vol 11(sp3): 280-288
https://doi.org/10.14719/pst.4909

HORIZON
e-Publishing Group
HePG

RESEARCH ARTICLE

Comprehensive phytochemical exploration of red, sweet and sour tamarind genotypes through GC-MS analysis

Amaravel M[1], Sugitha Thankappan[3], Asish K. Binodh[2], Bagath Singh C[1], Akshayasri[1], Radha Krishnan G[1], Nagarajan B[1] & Mayavel A[1]*

[1]Division of Genetics and Tree Improvement, ICFRE-Institute of Forest Genetics and Tree Breeding, Coimbatore 641 002

[2]Centre for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore 641 003

[3]School of Agricultural Sciences, Karunya Institute of Technology and Sciences, Karunya Nagar, Coimbatore 641 114

*Email: mayavelscientist@gmail.com

OPEN ACCESS

ARTICLE HISTORY
Received: 01 September 2024
Accepted: 11 November 2024
Available online
Version 1.0 : 29 December 2024
Version 2.0 : 10 February 2025

Check for updates

Additional information
Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

CITE THIS ARTICLE
Amaravel M, Sugitha Thankappan, A.K. Binodh ,C. Bagathsingh, Akshayasri, Radhakrishnan G, Nagarajan B, Mayavel A. Comprehensive phytochemical exploration of red, sweet and sour tamarind genotypes through GC-MS analysis. Plant Science Today. 2024; 11(sp3): 280-288. https://doi.org/10.14719/pst.4909

Abstract

Tamarind fruit pulp is a significant spice and flavouring agent used in various cuisines worldwide. Tamarind pulp has potential therapeutic value due to the presence of numerous bioactive components and widely utilized in preparation of different Ayurvedic medicines for treating the different illness. The present investigation aimed to evaluate the phytochemical constituents in the fruit pulp of different phenotypic variants of red, sweet and sour Tamarind clones. The Tamarind fruits were collected from the Tamarind germplasm bank at the ICFRE-Institute of Forest Genetics and Tree Breeding and subjected to phytochemical and GC-MS analysis. The Tamarind clones IFGTBTI-4, IFGTBRT-4 and IFGTBST-5 were subjected to GC-MS analysis. The methanol extract of fruit pulp was analysed by GCMS for identify the bioactive component present in the different phenotypic variants of Tamarind. The GCMS analysis revealed 22 components in sour Tamarind, 18 components in red Tamarind and 22 components in sweet Tamarind. The most important bioactive compounds present in all the Tamarind types are *myo-inositol, 4-C-methyl-, L-(+)-ascorbic acid 2,6-dihexadecanoate* and *2-furancarboxaldehyde, 5-(hydroxymethyl)* and also the components such as α-*calacorene, gamma-sitosterol* and *levoglucosenon*. Therapeutic potential, including anti-inflammatory, antioxidant, antiviral and anticancer properties. The findings contribute to the pharmacological validation of Tamarind extracts and to supporting their utility in pharmaceutical, cosmetic and food industries. This study highlights Tamarind's role as a source of natural bioactive compounds with significant health benefits, emphasizing its potential in the development of new therapeutic agents and supplements based on sustainable, bio-based chemicals.

Keywords

GCMS analysis; phytochemical analysis; tamarind

Introduction

Tamarindus indica L., commonly known as Tamarind, belongs to the tropical Fabaceae family and the Caesalpiniaceae subfamily, thriving in the regions of Africa and Southern Asia. Its composition mainly consists of 30-50% of pulp, 11-30% of shell and 25-40% of seeds (1). Based on the variations in the fruit color, acidity and total sugars, Tamarind is delineated into red, sweet and sour types, which is attributed to the presence of various phyto-compounds. The medicinal applications of Tamarinds are widely utilized in many traditional medical systems for its effectiveness in treating various ailments

(2). The fruit pulp is laxative and effective for treating fever, scurvy, nausea and bile-related disorders (3). Tamarind leaves are used for their anthelmintic and vermifuge properties to inhibit the growth of intestinal parasites and were widely adopted in ethnomedicinal practices across Africa, Asia and Latin America (4). Moreover, the seeds are traditionally used to manage diabetes, fever and gastrointestinal infections (5). The diverse pharmacological actions of *T. indica* are attributed to its array of phytochemicals, including flavonoids, saponins, alkaloids, tannins, polyphenols and steroids, making every part of the plant rich in therapeutic compounds (6). While the health and economic benefits of *T. indica* are recognized worldwide, detailed insights into the biologically active constituents of its methanol extract, particularly regarding its antioxidant and antibacterial properties, remain scarce (7). This study aims to bridge this gap by exploring the bioactive phyto-constituents of *T. indica* fruit through GC-MS analysis and assessing its potential for antibacterial and antioxidant activities, reflecting the ongoing interest in sourcing novel pharmaceuticals from natural products through advanced selection methodologies.

Materials and Methods

Sample collection and preparation of crude extract

Ripened fruits from five sour Tamarind trees were collected from the IFGTB VMG at the Forest Campus, Coimbatore (11° 01'N; 76°94'E) and from each of the red and sweet Tamarind varieties, 5 clones were collected from the National Tamarind Germplasm Bank for Red and Sweet Tamarind at IFGTB-FRS in Salem (11°09'N; 76°84'E). In GC-MS analysis, 20 g of pulp of ripened Tamarind fruits of each accession were weighed to prepare aqueous and methanolic extracts of the fruit samples. The sample extracts were centrifuged and filtered by Whatman filter paper. The sample extracts were stored in sterile centrifuge tubes and refrigerated at 4 ^{0}C for further analysis.

The crude extract of the fruit samples re-suspended with 100 mg/mL (w/v) of methanol, filtered with Whatmann filter paper no. 1, evaporated to dryness followed by mixed with 1 mL of methanol and impurities in the extract were removed using a Varian Bond Elute C18 solid phase extraction column.

Untargeted profiling of tamarind fruit crude extract

Exactly 1 µL of the purified samples were injected into a coupled Varian 4000 GC/MS interfaced with a GC-MS/MS Chemstation and NBS 49 K mass spectral library containing mass spectra of over 100000 compounds. A fused silica capillary column (10 m x 0.2 mm) with a cross-linked methyl silicon phase utilized. Helium served as the transporter gas. The range of temperature was set between 40 °C and 250 °C, with a rise of 5 °C per min and a 2 min solvent delay. The injector transfer line and ion source temperatures were maintained at 230 °C and 220 °C respectively. The mass spectral data observed during the assay were related with the compounds of mass spectra of in the ChemStation NIST library.

Metabolite analysis

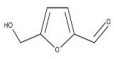
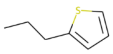
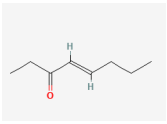
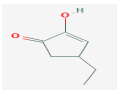
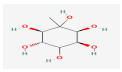
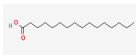
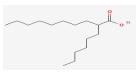
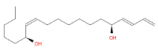
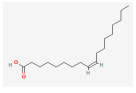
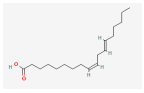
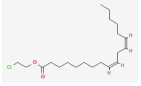
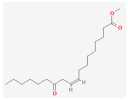
The metabolites profiled from the GC-MS spectra were selected based on the retention time (RT) and the maximum possible hit values. Before analyzing the data, a data integrity check is carried out to verify that all required information has been collected, which is followed by row-wise normalization. The logarithmically transformed values were auto scaled before conducting analysis. The metabolites were analyzed using the online tool Metaboanalyst 6.0 using R programming at the backend (24). The statistical and machine learning data analysis includes one-way ANOVA, Multivariate analysis (PCA), hierarchical clustering, Self-organizing map (SOM-unsupervised) and Random Forest analysis (RF-supervised). SOM and RF algorithms are used to identify the trends in high dimensional data analysis using the packages R som and Random Forest packages respectively in R programming.

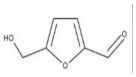
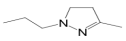
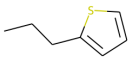
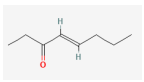
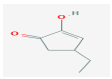
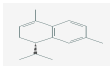
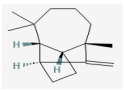
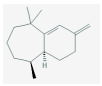
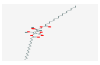
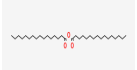
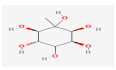
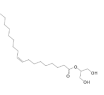
Results

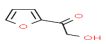
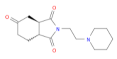
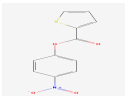
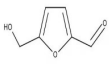
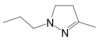
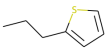
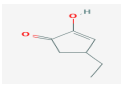
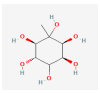
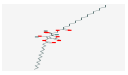
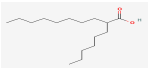
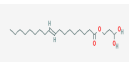
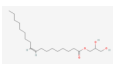
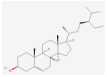
The GC-MS analysis of methanol extracts from IFGTBST-5, IFGTBTI-4 and IFGTBRT-4 Tamarind clone revealed the presence of 22 components in IFGTBST- 5, 22 components in IFGTBRT-4 and 18 components in IFGTBTI-14 (Table 1). The most significant compounds identified in the IFGTBRT-4, IFGTBTI-4 and IFGTBST-5 Tamarind clones were 2-Furancarboxaldehyde, 5-(hydroxymethyl) (53.33%: 53.27%: 55.40%), Myo-Inositol, 4-C-methyl- (63.50%: 23.54%: 20.66%) and L-(+)-Ascorbic acid 2,6-dihexadecanoate (58.48%: 59.44%: 54.11%). In the sweet Tamarind clone IFGTBST-5, 2-Furancarboxaldehyde 5-(hydroxymethyl), showed a high amount of 55.4%. The red Tamarind clone IFGTBRT-4 exhibited high levels of Myo-Inositol,4 -C-methyl (63.50%) and L-(+)-Ascorbic acid-2,6-dihexadecanoate (58.48%). The component 9,12-Octadecadienoic acid (Z,Z) was found in 22.76% of the sour Tamarind clone IFGTBTI-4. In the red Tamarind clone IFGTBRT-4, α-Calacorene was found at 69.74%. The sweet Tamarind clone IFGTBST-5 showed a high amount of Gamma-Sitosterol (67.78%), a phytosterol and Levoglucosenone was present at 66.01%. The component Octadecanoic acid was found in 52.68% of the IFGTBST-5 clone.

These component signals were collected from the NIST database using the “MALDI quant” package in R program. Analysis of Principal component analysis (PCA) was conducted following data normalization. PCA was conducted for Permutational multivariate analysis of variance (PERMANOVA) to identify variations in metabolite production and explore their interrelationships among genotypes. Principal components with eigenvalues greater than 1 were considered. The figure presents the scatter plot of the PC scores along with their corresponding loading plots. PC1 and PC2 together explained 97.58% of the total variance, effectively representing all variables (Fig. 1). PC1 accounted for 63.75% of the variance and was associated with the following volatile compounds:2-Furancarboxaldehyde,1-H-Pyrazole 4,5-dihydro-3-methy1-1-propyl Thiophene-2-propyl, Pentadecanoic acid, Palmitic anhydride, Furyl hydroxymethyl ketone, Levoglucosenone, Isoindole1,3,5-trione,perhydro-2-[2-(1-piperidyl) ethyl],2-Thiophene carboxylic acid,4-nitrophenylester,1,2-Naphthalenedione 3,8 dimethyl-(1-methylethyl),2,3-Dihydroxypropyl elaidate,9-Octadecenoic

Table 1. GC-MS analysis of the methanol pulp extract of tamarind clones

Sl. No.	RT	Lead compound	Mol. Formula	Mol. Structure	Mol. Wt	%	Biological properties
				IFGTBTI-4			
1	6.023	2-Furancarboxaldehyde, 5-(hydroxymethyl)	$C_6H_6O_3$		126	53.27	Inhibit the formation of sickled cells in blood, Antimicrobial
2	6.023	1H-Pyrazole,4,5-dihydro-3-methyl-1-propyl	$C_7H_{14}N_2$		126	12.36	Antioxidant, anticancer, antiulcer, antihypertensive
3	6.023	Thiophene, 2-propyl	$C_8H_{14}O$		126	7.98	-
4	6.023	4-Octen-3-one	$C_7H_{10}S$		126	4.11	Flavoring agent or adjuvant
5	6.023	4Ethyl-2-hydroxycyclopent-2-en-1-one	$C_7H_{10}O_2$		126	2.6	Antibacterial
6	13.21	Myo-Inositol, 4-C-methyl	$C_7H_{14}O_6$		194	23.54	anti-inflammatory
7	14.914	l(+)Ascorbic acid 2,6-dihexadecanoate	$C_{38}H_{68}O_8$		652	59.44	Antioxidant
8	14.914	n-Hexadecanoic acid	$C_{16}H_{32}O_2$		256	16.2	Anti-HIV potential, antimicrobial, flavoring agent
9	14.914	Pentadecanoic acid	$C_{15}H_{30}O_2$		242	4.94	Flavoring Agents
10	14.914	Palmitic anhydride	$C_{32}H_{62}O_3$		494	3.88	Emulsifier (allows mixing oil and water), opacifying agent
11	14.914	2-Hexyldecanoic acid	$C_{16}H_{32}O_2$		256	2.89	Cleansing; Emulsifying; Surfactant
12	16.65	9Octadecenoic acid(Z)-, 2-hydroxy-1-(hydroxymethyl) ethyl ester	$C_{21}H_{40}O_4$		356	7.72	Antagonist
13	16.65	E,E,Z-1,3,12-Nonadecatriene-5,14-diol	$C_{19}H_{34}O_2$		294	6.52	Antagonist
14	16.65	Oleic acid	$C_{18}H_{34}O_2$		282	4.47	Antimicrobial
15	16.991	Heptadecanoic acid	$C_{17}H_{34}O_2$		270	2.78	Antagonist
16	18	9,12-Octadecadienoic acid (Z,Z)	$C_{18}H_{32}O_2$		294	22.76	Agonist
17	18.572	2-Chloroethyl linoleate	$C_{20}H_{35}ClO_2$		342	9.86	-
18	19.433	9-Octadecenoic acid, 12-hydroxy-, methyl	$C_{19}H_{34}O_3$		342	3.99	_

IFGTBRT-4							
1	6.201	2-Furancarboxaldehyde, 5-(hydroxymethyl)-	$C_6H_6O_3$		126	53.33	Inhibit the formation of sickled cells in blood, Antimicrobial
2	6.201	1H-Pyrazole,4,5-dihydro-3-methyl-1-propyl	$C_7H_{14}N_2$		126	8.76	Antioxidant, anticancer, antiulcer, antihypertensive
3	6.201	Thiophene, 2-propyl-	$C_7H_{10}S$		126	6.01	-
4	6.201	4-Octen-3-one	$C_8H_{14}O$		126	3.76	Flavoring agent or adjuvant
5	6.201	4-Ethyl-2-hydroxycyclopent-2-en-1-one	$C_7H_{10}O_2$		126	1.93	Antibacterial
6	9.646	Isoledene	$C_{15}H_{24}$		204	2.06	Cytotoxic
7	9.986	Alpha-Calacorene	$C_{15}H_{20}$		200	69.74	Anti-viral
8	9.986	Cadala-1(10),3,8-triene	$C_{15}H_{22}$		202	13.66	Anti-fungal
9	9.986	Pyrimido[1,6-a]indole,1,2,3,4-tetrahydro-2,5-dimethyl-	$C_{13}H_{16}N_2$		200	3.06	Pharmacological uses
10	9.986	Naphthalene, 1,2-dihydro-1,1,6-trimethyl-	$C_{13}H_{16}$		172	1.84	Cancer treatment, flavoring
11	11.826	Longifolene-(V4)	$C_{15}H_{24}$		204	5.36	Flavoring agent
12	11.826	10s,11s-Himachala-3(12),4-diene	$C_{15}H_{25}$		204	3.98	Anti-bacterial, anti-allergens
13	11.826	Hinesol	$C_{15}H_{26}O$		222	3.67	Anti-hepatotoxic activity
14	14.951	l(+)Ascorbic acid 2,6-dihexadecanoate	$C_{38}H_{68}O_8$		652	58.48	Antioxidant
15	14.951	n-Hexadecanoic acid	$C_{16}H_{32}O_2$		256	16.25	Anti-HIV potential ,antimicrobial, flavoring agent
16	14.951	Pentadecanoic acid	$C_{15}H_{30}O_2$		242	4.87	Flavoring Agents
17	14.951	Palmitic anhydride	$C_{32}H_{62}O_3$		494	3.92	emulsifier (allows mixing oil and water), opacifying agent or emollient
18	14.951	2-Hexyldecanoic acid	$C_{16}H_{32}O_2$		256	2.85	Cleansing; Emulsifying; Surfactant
19	14.23	Myo-Inositol, 4-C-methyl-	$C_7H_{14}O_6$		194	63.50	anti-inflammatory
20	17.505	9Octadecenoic acid (Z), 2-hydroxy-1-(hydroxymethyl)ethyl ester	$C_{21}H_{40}O_4$		356	8.55	Antagonist
21	17.505	Octadec-9-enoic acid	$C_{18}H_{34}O_2$		282	5.82	-
22	17.505	E,E,Z1,3,12-Nonadecatriene-5,14-diol	$C_{19}H_{34}O_2$		294	4.94	Antagonist

			IFGTBST-5				
1	3.863	Furyl hydroxymethyl ketone	$C_6H_6O_3$		126	54.26	-
2	4.305	Levoglucosenone	$C_6H_6O_3$		126	66.01	Anti-inflammatory, Cytotoxic
3	4.305	Isoindole-1,3,5-trione,perhydro-2-[2-(1 -piperidyl)ethyl]-	$C_{15}H_{22}N_2O_3$		278	16.92	-
4	4.665	2-Thiophenecarboxylic acid, 4-nitrophenyl ester	$C_{11}H_7NO_4S$		249	27.5	-
5	6.058	2-Furancarboxaldehyde, 5-(hydroxymethyl)-	$C_6H_6O_3$		126	55.4	Inhibit the formation of sickled cells in blood, Antimicrobial, Preservative
6	6.058	1H-Pyrazole,4,5-dihydro-3-methyl-1-propyl-	$C_7H_{14}N_2$		126	13.06	Antioxidant, anticancer, antiulcer, antihypertensive
7	6.058	Thiophene, 2-propyl-	$C_7H_{10}S$		126	9.74	-
8	6.058	4-Ethyl-2-hydroxycyclopent-2-en-1-one	$C_7H_{10}O_2$		126	2.55	Antibacterial
9	10.886	1,2-Naphthalenedione, 3,8-dimethyl-5-(1methylethyl)	$C_{15}H_{16}O_2$		244	2.98	Antibacterial
10	13.178	Myo-Inositol, 4-C-methyl-	$C_7H_{14}O_6$		194	20.66	-
11	14.911	l(+)Ascorbic acid 2,6-dihexadecanoate	$C_{38}H_{68}O_8$		652	54.11	Anti-oxidant
12	14.911	n-Hexadecanoic acid	$C_{16}H_{32}O_2$		256	14.75	Anti-HIV potential, antimicrobial, flavoring agent
13	14.911	Pentadecanoic acid	$C_{15}H_{30}O_2$		242	6.71	Flavoring Agents
14	14.911	Palmitic anhydride	$C_{32}H_{62}O_3$		494	5.01	emulsifier (allows mixing oil and water), opacifying agent or emollient
15	14.911	2-Hexyldecanoic acid	$C_{16}H_{32}O_2$		256	3.33	Cleansing; Emulsifying; Surfactant
16	16.648	2,3-Dihydroxypropyl elaidate	$C_{21}H_{40}O_4$		356	13.95	Flavoring agent
17	16.648	9Octadecenoic acid (Z), methyl ester	$C_{19}H_{36}O_2$		296	7.71	Flavoring agent
18	17.419	9-Octadecenoic acid(Z), 2,3-dihydroxypropyl ester	$C_{21}H_{40}O_4$		356	7.17	Flavoring agent
19	17.687	Octadecanoic acid	$C_{18}H_{36}O_2$		284	52.68	Flavoring agent
20	17.687	Octadecanoic acid, 2-(2hydroxyethoxy) ethyl ester	$C_{22}H_{44}O_4$		372	11.02	Emulsifier
21	17.687	L-Ascorbic acid, 6-octadecanoate	$C_{24}H_{42}O_7$		442	5.46	Anti-oxidant
22	35.878	Gamma.-Sitosterol	$C_{29}H_{50}O$		414	67.78	Hypolipidemic agents Anti-proliferative activity

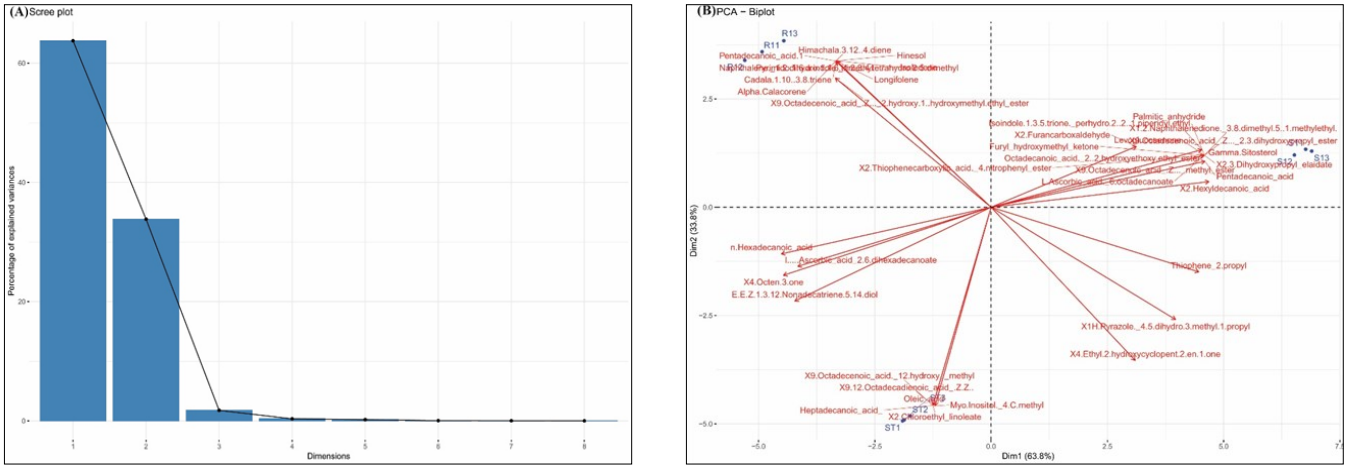


Fig. 1. Principal component analysis for metabolomics of sour, red and sweet Tamarind (Scree plot (A) and PCA biplot (B))

acid (Z) methyl ester,9-Octadecenoic acid (Z),2,3-dihydroxypropyl ester, Octadecanoic acid 2-(2-hydroxyethoxy) ethyl ester, L-Ascorbic acid 6 octadecanoate and Gamma-Sitosterol. Similarly, PC2 explained 33.82% of the total divergence and showed strong correlations with compounds of volatile Alpha-Calacorene, Cadala-1(10),3,8-triene, Pyrimido -[1,6-a] indole,1,2,3,4-tetrahydro-2,5-dimathyl, Naphthalene,1,2-dihydro-1,1,6-trimethyl, Longifolene, Himachala-3(12),4-diene, Hinesol and Pentadecanoic acid.

From the heat map cluster analysis (Fig. 2), the IFGTBRT -4 Tamarind clone had a row z-score >1, which was observed in metabolites such as 4-Ethyl-2-hydroxycyclopent-2-en-1-one,Myo-Inositol, 4-C-methyl, Oleic acid, Pentadecanoic acid, 9,12Octadecadienoic acid (Z,Z), 2-Chloroethyl linoleate and 9 Octadecenoic acid, 12hydroxy, methyl. The IFGTBRT-4 Tamarind clone had a row z-score>1, which was observed in Alpha-Calacorene,Cadala-1(10),3,8-triene,Pyrimido [1,6-a] indole, 1,2,3,4-tetrahydro-2,5-dimathyl, Naphthalene, 1,2-dihydro-1,1, 6-trimethyl, Longifolene, Himachala-3(12), 4-diene,Hinesol and Pentadecanoic acid. The IFGTBST-5 Tamarind clone had Levoglucosenone, Isoindole-1,3,5-trione, perhydro-2-[2-(1-piperidyl)ethyl], 2-Thiophenecarboxylicacid, 4-nitrophenyl-ester, 1,2 Naphthalenedione, 3,8-dimethyl-5 (1methylethyl), 2, 3-Dihydroxypropyl elaidate, 9-Octadecenoic acid(Z), methyl ester, 9Octadecenoic acid(Z), 2,3-dihydroxypropyl ester, Octadecanoic acid, 2-(2-hydroxyethoxy) ethyl ester, L-Ascorbic acid, 6-octadecanoate and Gamma-Sitosterol.

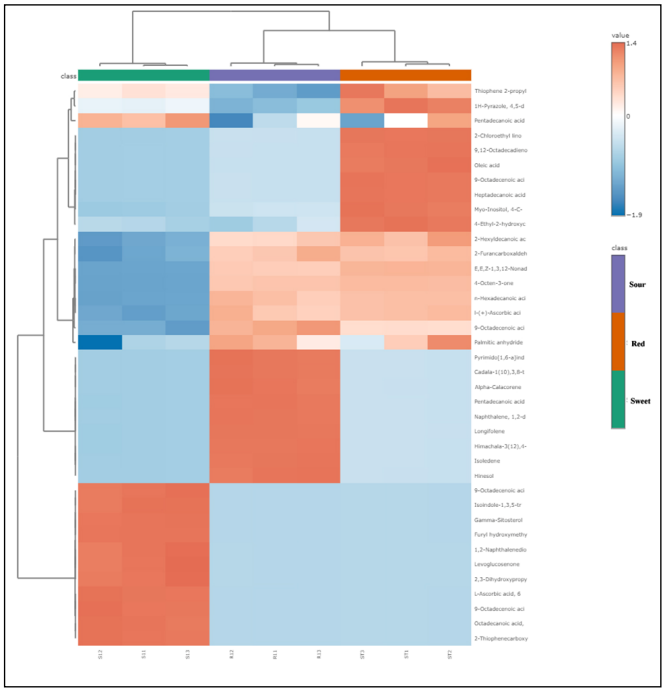


Fig. 2. Heat map clustering for metabolomes of sour, red and sweet Tamarind

An unsupervised neural network algorithm of self-organizing map (SOM) was used to automatically detect the key trends in the metabolomic profiles of tamarind clones. It is based on interconnected nodes in a grid, with each node serving as a model. Accordingly, the SOM clustering revealed only two matrices cluster (0,1) with IFGTBRT-4(R11; R12; R13) and IFGTBST-5 (ST1; ST2; ST3) and cluster (O,2) including IFGTBTI-4 (S11; S12; S13). (Fig. 3) Table 2. Similarly, the RF

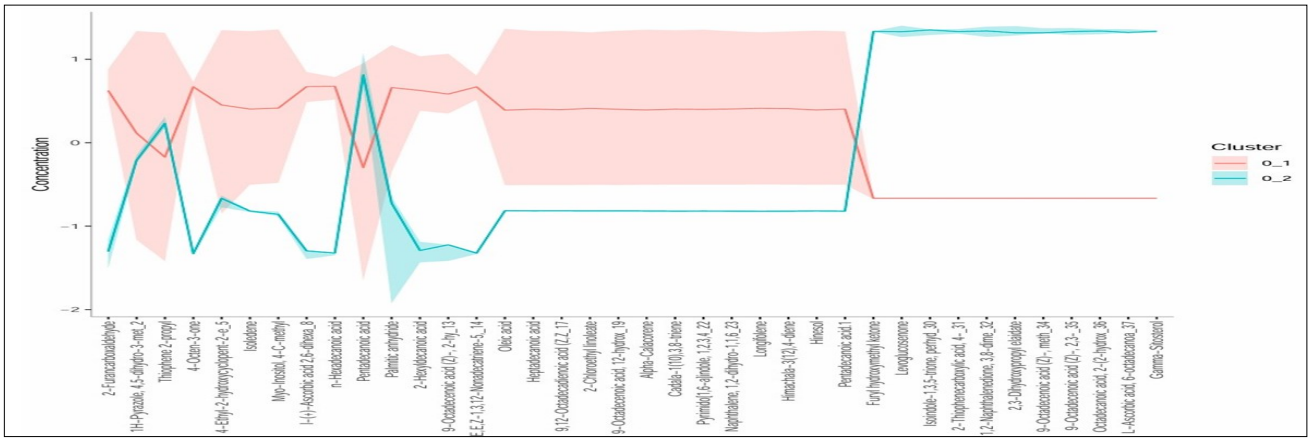


Fig. 3. SOM cluster analysis for Tamarind metabolites. The x-axes are components and y-axes are intensities. The blue lines denote median intensities of respective clusters

Table 2. Clustering result using SOM for Tamarind clones

Cluster	Sample in each cluster
Clsuter(0,0)	
Clsuter(0,1)	R11 (IFGTBRT-4), R12 (IFGTBRT-4), R13 (IFGTBRT-4), ST1 (IFGTBST-5), ST2 (IFGTBST-5) and ST3 (IFGTBST-5)
Clsuter(0,2)	S11 (IFGTBTI-4), S12 (IFGTBST-5) and S13 (IFGTBST-5)

analysis is a supervised. The machine learning algorithm utilizes an ensemble of classification trees, with each tree being constructed through random feature selection from a bootstrap sample at each branch. Random Forest (RF) also offers additional valuable information, including out-of-bag (OOB) error, variable importance measures and outlier assessments. The class error value was 0.00 for all the three genotypes studied. The significant features identified are graded by mean decrease accuracy in categorization accuracy when they are permuted. MDA value for 1H-Pyrazole, 4 is high in sour type whereas less in red type. The MDA values for most of the feature's compounds exhibited very low intensity in the sweet type. (Fig. 4).

Discussion

GC-MS analysis of methanol extracts from the IFGTBST-4, IFGTBTI-5 and IFGTBRT-4 Tamarind clones revealed the presence of multiple components in each clone. The most significant compounds identified in these clones include 2-Furancarboxaldehyde, 5-(hydroxymethyl)-, Myo Inositol,4-C methyl- and l (+) Ascorbic_acid 2,6 dihexadecanoate. These compounds are known to possess therapeutic benefits, including anti-inflammatory, antimicrobial and antioxidant properties. The observed bioactivity aligns with findings from previous studies (8, 9, 22, 23), which also reported the identification of biological compounds in methanol extracts of Tamarind. These studies highlighted peak area percentages exceeding 1%, suggesting significant biological relevance. The high levels of these compounds in the analysed clones underscore their potential applications in pharmaceutical, nutraceutical and cosmetic industries.

In the red Tamarind clone IFGTBRT-4, 2 Furancarboxaldehyde, 5-(hydroxymethyl), a compound known for its pharmacological properties, including anti-inflammatory, antioxidant and anti-hypoxic effects, was notably abundant (9). The United States Food and Drug Administration (FDA) has reported that this compound has gained attention for its therapeutic potential in the treatment of sickle cell disease (10). In the red Tamarind clone IFGTBRT-4, Myo-Inositol, 4-C-methyl-, a derivative of myo-inositol, plays a crucial role in cellular processes, potentially impacting brain health, mood, cognitive function, and metabolic disorders (11). The component L-(+) -Ascorbic acid 2,6dihexadecanoate, a fat-soluble vitamin C derivative, is highly valuable in pharmaceuticals, cosmetics and the food industries due to its improved stability and ability to integrate into lipid-based systems, offering sustained antioxidant protection.

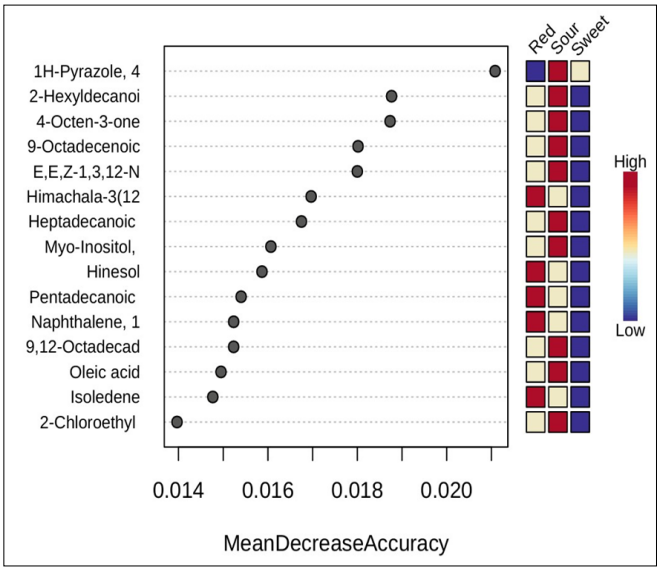


Fig. 4. Significant features of Tamarind metabolites identified by Random Forest

The component 9,12 Octadecadienoic acid (Z,Z), commonly known as α-linolenic acid, identified in the sweet Tamarind clone IFGTBTI-4 is associated with significant health benefits, including decreased risk of mortality for coronary heart disease and cardiovascular diseases, despite a potential slight increase in cancer-related mortality risk (12). Additionally, α-linolenic acid may offer the added advantage of diminishing blood clots formation, further contributing to its cardiovascular benefits (13, 14). In the red Tamarind clone IFGTBRT-4, α-calacorene, a sesquiterpene, was prominent and is known for its antimicrobial, anti-inflammatory, antiviral and antifungal properties, highlighting its importance in both pharmaceutical development and traditional herbal medicine (15, 16). Gamma-sitosterol, a phytosterol observed in the same clone, offers various health benefits, including lowering cholesterol levels, improving urinary tract health, and potential anticancer properties, emphasizing its value in nutrition, dietary supplementation and pharmaceutical research (17, 18). Levoglucosenone, found in the Tamarind clone IFGTBST-5, is recognized for its unique structure, derived from the pyrolysis of cellulose. It holds promise as a versatile building block for synthesizing pharmaceuticals, agrochemicals, and advanced materials, showcasing the significance of sustainable, bio-based chemicals (19, 20). Octadecanoic acid, present in IFGTBTI-4, is a versatile saturated fatty acid widely used across the cosmetics, food and pharmaceutical industries for its stabilizing, emollient, and lubricant properties. Its unique characteristics distinguish it from other saturated fats in dietary considerations (21).

The unsupervised machine learning algorithm SOM classified the genotypes into two clusters, of which cluster (0,1) encompass a pooled metabolites of both red and sweet, whereas cluster (0,2) contained only sour type. This indicates that most of the metabolites with median intensities are fall under sour type, represented by the which is indicated by line in the Fig. 3. As explained above the compounds like gamma sitosterol, Ascorbic acid, Octadecenoic acid and isoindole and Levoglucosenone are highly significant for the sour type in determining its genetic traits. However, the metabolites 2-Furancarboxaldehyde, 2-Hexyldecanoic acid, Hinesol,

Pentadecanoic acid.1 and 4-Octen-3-one could be a visible metabolite marker for clustering red and sweet types. Here few metabolites are shared by both of them.

The supervised RF analysis identified 15 signature features based on mean decrease accuracy values. Unlike to SOM, the sweet genotype exhibited lower intensities for most metabolites, except for 1H-Pyrazole,4, which was less in red genotype. The sour type displayed medium to low intensities of all the 15 features, suggesting its richness of metabolite diversity. This diversity in the sour genotype highlights its potential for further exploration using targeted profiling studies.

Conclusion

The metabolomic analysis of methanol extracts from IFGTBST-5, IFGTBTI-4 and IFGTBRT-4 Tamarind clones identified 22 metabolites in both IFGTBRT-4 and IFGTBST-5 and 18 metabolites in IFGTBTI-4, with significant variability in their chemical profiles. Notably, key metabolites such as 2-Furancarboxaldehyde, 5-(hydroxymethyl)-, Myo-Inositol,4-C-methyl- and L(+)Ascorbic acid 2,6-dihexadecanoate were dominant across the clones. The red Tamarind clone IFGTBST-5 exhibited high levels of Myo-Inositol,4-C-methyl- and L (+) Ascorbic acid 2,6-dihexadecanoate, while the sweet Tamarind clone IFGTBRT-4 was characterized by a high content of 2-Furancarboxaldehyde, 5-(hydroxymethyl)- and α-Calacorene. The results from PCA indicated that PC1 and PC2 together explained 97.58% of the total variance, with PC1 strongly associated with a range of volatile metabolites, including Gamma-Sitosterol, while PC2 correlated with Alpha-Calacorene and other aromatic metabolites. The heat map cluster analysis further revealed distinct metabolic profiles for each clone, highlighting the unique composition of volatile metabolites in the IFGTBRT-4, IFGTBTI-4 and IFGTBST-5 clones. This metabolomic analysis underscores the substantial metabolic diversity among Tamarind clones, providing a crucial biochemical foundation for identifying metabolomic superior genotypes. By elucidating the distinct metabolic profiles, this study enhances our understanding of the biochemical traits associated with desirable characteristics, thus offering valuable insights for breeding programs aimed at improving Tamarind varieties.

Acknowledgments

The authors express gratitude to the CAMPA Authority, Ministry of Environment Forest and Climate Change, Government of India, for the financial support received under the All India Coordinated Research Project.

Authors' contributions

AM, TS and MA planned and coordinated this study. CB, MA (M. Akshayasri) and GR and carried out the performed the lab experiment. BN and AB carried out the grammatical correction. TS and MA participated in the design of the study and performed the statistical analysis. AM, TS and MA revised the paper for final publication. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Ethical issues: None

References

1. El-Siddique K, Gunasena HP, Prasa BA, Pushpakumara DK, Ramana KV, Vijayanand P, et al. Tamarind–*Tamarindus indica* L. fruits for the future. Southampton Centre for Underutilized Crops, Southampton. UK; 2016. p.188
2. Menezes PPA, Trevisan CCS, Barbalho SM, Guiguer LE. *Tamarindus indica* L. a plant with multiple medicinal purposes. J Pharmaco and Phytochem. 2016;5(3):50-54.
3. Raimondi L, Lodovici M, Guglielmi F, Banchelli G, Ciuffi M, Boldrini E, et al. The polysaccharide from *Tamarindus indica* (TS-polysaccharide) protects cultured corneal□derived cells (SIRC cells) from ultraviolet rays. J Pharma and Pharmaco. 2003; 55:333 -38. https://doi.org/10.1211/002235702630
4. Pamplona-Roger GD. Encyclopedia of medicinal plants education and health library; Madrid, Spain. 1999;2536.
5. Gupta C, Prakash D, Gupta S. Studies on the antimicrobial activity of Tamarind (*Tamarindus indica*) and its potential as a food bio-preservative. Inter Food Res J. 2014; 21(6):2437-41.
6. Komakech R, Kim YG, Matsabisa GM, Kang Y. Anti-infammatory and analgesic potential of *Tamarindus indica* Linn. Integra Med Res. 2019;8(3):181-86. https://doi.org/10.1016/j.imr.2019.07.002 (2019)
7. Kaur R, Mudgal R, Jose J, Kumar P, Tomar S. Glycan-dependent chikungunya viral infection divulged by the antiviral activity of NAG-specifc chi-like lectin. Virology. 2019; 526:91-98. https://doi.org/10.1016/j.virol.2018.10.009 (2019)
8. Senthilkumar N, Sumathi R, Lenora DLM, Divya G. Phytochemical properties and antioxidant activity of natural colourant extracted from Red Tamarind *Tamarindus indica* var. rhodocarpa in Tamilnadu. Res J Chem and Environ. 2021;25(3):79-88.
9. Farag MR, Alagawany M, Bin-Jumah M, Othman SI, Khafaga AF, Shaheen HM, Abd El-Hack ME. The toxicological aspects of the heat-borne toxicant 5 hydroxymethylfurfural in animals. Molecules. 2020;25(8):1941. https://doi.org/10.3390/molecules25081941
10. Cronin RM, Lin CJ, Chiang C, MacEwan SR, DeBaun MR, Hyer JM. The use of FDA-approved medications for preventing vaso-occlusive events in sickle cell disease. Blood Advances. 2023;7 (13):3114.https://doi.org/10.1182/bloodadvances.2022008965
11. Anupama N, Madhumitha G, Rajesh KS. Role of dried fruits of *Carissa carandas* as anti-inflammatory agents and the analysis of phytochemical constituents by GC-MS. BioMedical Res Inter. 2014;2014(1):512369. https://doi.org/10.1155/2014/512369
12. Naghshi S, Aune D, Beyene J, Mobarak S, Asadi M, Sadeghi O. Dietary intake and biomarkers of alpha linolenic acid and risk of all cause, cardiovascular and cancer mortality: systematic review and dose-response meta-analysis of cohort studies. Bmj. 2020;375:n2213. https://doi.org/10.1136/bmj.n2213
13. Brzosko S, De Curtis A, Murzilli S, De Gaetano G, Donati MB, Iacoviello L. Effect of extra virgin olive oil on experimental thrombosis and primary hemostasis in rats. Nutri, Metabol and Cardiovascular Dis: NMCD. 2002;12(6):337-42. https://doi.org/10.1111/j.1538-7836.2003.tb05555.x
14. Yang Q, Cao W, Zhou X, Cao W, Xie Y, Wang S. Anti-thrombotic effects of α-linolenic acid isolated from *Zanthoxylum bungeanum*

Maxim seeds. BMC Complementary and Alternative Med. 2014;14 (1):348. https://doi.org/10.1186/1472-6882-14-348

15. Umroni A, Rianawati H, Siswadi, Pujiarti R. GC-MS analysis of essential oil of Kayu Papi (*Exocarpus latifolius* R. Br., Fam: Santalacea from East Nusa Tenggara. IOP Conference Series: Earth and Environmental Science; 2021. 762(1):012032. https://doi.org/10.1088/1755-1315/762/1/012032
16. Vukovic N, Milosevic T, Sukdolak S, Solujic S. Antimicrobial activities of essential oil and methanol extract of *Teucrium montanum*. Evidence-Based Complementary and Alternative Med. 2007;4:17-20. https://doi.org/10.1093/ecam/nem108
17. Sonia S, Singh SK. Phytoconstituents of *Ziziphus nummularia* (Burm. f.) Wight and Arn. leaves extracts using GC-MS spectroscopy. Res and Rev: J Life Sci. 2019;9:109-18.
18. Singh S, Kaushal S, Hunjan MS, Sharma P. Chemical composition and molecular docking studies of *Cyamopsis tetragonoloba* (L.) Taub. leaves and fruits extracts as antifungal agents against rice phytopathogenic fungi. Arch Phytopath and Plant Protect. 2022;55 (10):1179-200. https://doi.org/10.1080/03235408.2022.208/769
19. Camp JE, Greatrex BW. Levoglucosenone: Bio-based platform for drug discovery. Front in Chem. 2022;10:902239.
20. Tan XJ, Li HM, Peng QX, Zhou H, Chen YS, Lu YY, et al. Enzymatic acylation of lutein with a series of saturated fatty acid vinyl esters and the thermal stability and anti-lipid oxidation properties of the acylated derivatives. J Food Sci. 2021;86(12):5240-52. https://doi.org/10.1111/1750-3841.15966
21. Oluwasina OO, Awonyemi IO. Citrus peel extract starch-based bioplastic: Effect of extract concentration on packed fish and bioplastic properties. J Polymers and the Environ. 2021;29 (6):1706-16. https://doi.org/10.1007/s10924-020-01990-7
22. Fagbemi KO, Aina DA, Adeoye-Isijola MO, Naidoo KK, Coopoosamy RM, Olajuyigbe OO. Bioactive compounds, antibacterial and antioxidant activities of methanol extract of *Tamarindus indica* Linn. Scientific Reports. 2022;12(1):9432. https://doi.org/10.21203/rs.3.rs-1238883/v1
23. Muthu SE, Nandakumar S, Rao U. The effect of methanolic extract of *Tamarindus indica* Linn. on the growth of clinical isolates of *Burkholderia pseudomallei*. Indian J Med Res. 2005;122:525-28. https://doi.org/10.54085/ap.2024.13.2.54
24. Pang Z, Lu Y, Zhou G, Hui F, Xu L, Viau C, et al. MetaboAnalyst 6.0. towards a unified platform for metabolomics data processing, analysis and interpretation. Nucleic Acids Res. 2024;52(1):398-406. https://doi.org/10.1093/nar/gkae253