

Utility of ITS and rbcL Sequences for Resolving Intraspecific Divergence in *Saccharum spontaneum* L

Ramasamy Rajesh

Pasumpon Thiru Muthuramalinga Thevar Memorial College

Karuppaiah Thirumalainayagam Selvavinayagam

Pasumpon Thiru Muthuramalinga Thevar Memorial College

Koilpillai Dharmar

dharmarsk@gmail.com

Pasumpon Thiru Muthuramalinga Thevar Memorial College

Research Article

Keywords: Wild sugarcane, Dharpa, Kans grass, Phylogeny, DNA barcoding

Posted Date: April 2nd, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4165134/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Saccharum spontaneum L., commonly known as wild sugarcane or kans grass, is a close wild relative of sugarcane, representing the majority of terrestrial environments from the open ground to the shoreline. Due to its diverse applications as food, fodder, food ingredients, medicine, household products, pulp material, and even religious offerings, understanding the genetic relationships among populations is crucial. The present study attempted to understand the phylogenetic and evolutionary relationships of six field samples and two commercial samples from the chosen population of *S. spontaneum* in Tamil Nadu. The genomic material was extracted and amplified using nuclear ribosomal DNA internal transcribed spacer (nrDNA-ITS) and ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) sequences and compared with those obtained from the NCBI database. The evolutionary history was inferred using MEGA11 software. The results showed that the presence of highly conserved sites in the rbcL region led to a close phylogenetic relationship falling under a common clade, which could be useful for molecular identification of this species. On the other hand, ITS markers were found to be useful for determining the phylogeny of *S. spontaneum* via both phenetic and cladistic approaches. The ITS region in this wild sugarcane has better discrimination ability despite the highly conserved sites with few rapidly evolving sites. Four of the collected specimens showed a close relationship, along with two commercial specimens, suggesting shared adaptations or a common geographic origin. Additionally, two other populations were found to have diverged from the rest of the specimens, exhibiting low genetic distance, indicating recent evolutionary divergence.

Introduction

The *Saccharum* (sugarcane) genus belongs to the Poaceae or Gramineae. Currently, the genus includes sixteen taxonomically accepted species. Among these, *Saccharum spontaneum* L., classified as 'wild sugarcane' along with *S. robustum*, boasts impressive resilience against cold, drought, barren land, insects, and diseases (Bhuiyan et al. 2019; Dharshini et al. 2020; Kasirajan et al. 2021; Lai et al. 2022; Luo et al. 2023). The term '*spontaneum*' in the species epithet denotes 'spontaneous', likely indicating the plant's rapid growth and its capacity to easily establish itself in widespread environments. It is a rhizomatous geophyte, and in cultivation, asexual propagation through stem cuttings is the preferred mode of reproduction (Sreenivasan 1975).

In ancient Indian scriptures, *S. spontaneum* was mentioned as 'Kasha' grass used for weaving mats (Rig Vedha verse 10.100.10) and for making whips (Rig Veda verses 1.22.3, 1.37.3 and 8.33.11). In Satapatha-Brahmana, it has been mentioned as 'Ashvavala' (horse-tail) grass and is often identified as *S. bengalense*, *Desmostachya bipinnata*, *Eragrostris ciliaris* and *Imperata cylindrica* and is collectively named "Dharpa", which are used for mats and sacrificial offerings during the Hindu rituals, such as agnihotra, in various parts of India (Sheshadri 2013).

Economically, it is often used in the production of biofuels (Vaid et al. 2022; Dhungana et al. 2023), co-digesters in biogas production (Mahardhika et al. 2022), pulp production (Ferdous et al. 2020), the

preparation of biosorbents (Prens et al. 2022; Bharadwaj et al. 2023) and supercapacitors (Samantray et al. 2021). Although the sugar content in this plant is lower than that in domesticated sugarcanes, this plant has other potential traits that are important for plant breeding. It is a potential C-4 plant that offers superior photosynthetic and water-use efficiency, which is vital for ecosystem resilience (Bhattacharya 2019). In addition, the species has traits such as increased tillering, ratoonability, vigour and hardiness (Paterson et al. 2012).

As the germplasm contains important resistance traits, superior varieties in the modern day have been developed using this species as a predecessor for interspecific hybridization. Thus, *S. spontaneum* is still an important parental species (10–20% modern cultivar chromosomes; 80–90% from *S. officinarum*) in sugarcane breeding (Moore et al. 2013; Li et al. 2022). As early as 1921, sugarcane breeders employed the disease resistance characteristics of *S. spontaneum*, crossing tropical species with Java *S. spontaneum* ($2n = 112$) and developing a superior variety exhibiting strong resistance due to inheritance (Aitken et al. 2018). Its valuable traits have made it a crucial resource for enhancing cultivated sugarcane varieties. Notably, its hybridization with the noble cane *S. officinarum* led to the development of economically important cultivars (Paterson et al. 2012). The plant exhibits polyploidy, and the chromosome numbers of wild *S. spontaneum* varieties vary widely ($2n = 40–128$, euploidy/aneuploidy/polyploidy). In India, the tropical species *S. spontaneum* ($2n = 64$) is used to produce innovative parental species with three consanguinities, Co281, Co290 and Co213; this represented a second breakthrough in sugarcane breeding in the 20th century. In China, a series of new lines were bred using Yacheng *S. spontaneum* ($2n = 80$) and Yunnan *S. spontaneum* 75-2-11 ($2n = 64$), which are important parental species for sugarcane breeding (Yu et al. 2019). A chromosomal diversity study revealed increased allelic and genetic variation in addition to increased gene flow among different ploidy levels of *S. spontaneum* in India (Sobhakumari 2020; Dinesh Babu et al. 2022).

To date, only a few karyotypes of *S. spontaneum* have been successfully identified and utilized. Due to its advantages over conventional phenotypic classification, molecular marker technology, including RAPD, ISSR, TRAP, SSR, and AFLP, has been established as a valuable tool for sugarcane progeny identification and genetic diversity evaluation in wild sugarcane germplasm resources (Mary et al. 2006; Alwala et al. 2006; Tew and Pan 2010; Aitken et al. 2018; Govindaraj et al. 2021; Shahzad et al. 2022). Understanding the taxonomic evolution, interspecific diversity and intraspecific divergence is important for the future prospects of this important wild sugarcane. With this background, the current study was designed to evaluate the evolutionary taxonomy of selected populations of *S. spontaneum* present at major harvesting sites in Tamil Nadu, India.

Materials and Methods

Sample Collection and DNA Isolation

A total of eight plant samples were collected; among these, six were from the banks of three major rivers in Tamil Nadu: the Vaigai River (Manamadurai, Paramakudi), the Thamirabarani River (Papanasam,

Srivaikundam) and the Cauvery River (Aaduthurai, Kallanai). The locations of sample collection were chosen where the population was found to be abundant and often harvested for commercial utility purposes. Additionally, two samples were obtained from commercial vendors located in Kamuthi and Salem (Fig. 1; Table 1).

Table 1
Details of *S. spontaneum* sample collection

Specimen Id	Location Name	Latitude (Decimal Degrees)	Longitude (Decimal Degrees)	Elevation (m)
PTMTMC 002	Manamadurai	9.69827500	78.44971111	91
PTMTMC 102	Paramakudi	9.55898611	78.57350278	63
PTMTMC 005	Aaduthurai	11.02028333	79.48852778	42
PTMTMC 103	Kallanai	10.83386944	78.82637778	79
PTMTMC 008	Papanasam	8.71195556	77.36778056	98
PTMTMC 105	Srivaikundam	8.63154444	77.90851111	45
Kamuthi 001*	Kamuthi	9.41140833	78.36702778	61
Salem 001*	Salem	11.63971111	78.15591667	305
*- Commercial sample				

Approximately 200 g of fresh plant material from six field samples and predried material from two commercial samples were used for DNA extraction using CTAB method (Doyle and Doyle 1987). The extracted DNA was purified using a spin column kit to ensure high-quality genomic material. The genomic regions selected for the present study included the nuclear ribosomal DNA internal transcribed spacer (nrDNA-ITS) region (White et al. 1990) and the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL) region (Bafeel et al. 2011) and were amplified with suitable primers (Table 2).

Table 2
Primers used in the genomic profiling of *S. spontaneum*

Genomic Target	Primers	Primer Sequence 5'-3'	Reference
Internal Transcribed Spacer Region	ITS1	TCCGTAGGTGAACCTGCGG	White et al. 1990
	ITS4	TCCTCCGCTTATTGATATGC	
Barcoding Regions	rbcLa-F	ATGTCACCACAAACAGAGACTAAAGC	Bafeel et al. 2011
	rbcLa-R	GTAAAATCAAGTCCACCRGC	

The amplification process was performed using a PCR Thermal Cycler Program (Eppendorf ProS, Hamburg, Germany). The PCR thermal cycle for the ITS and rbcL markers consisted of one cycle of 5 min at 94°C; 35 cycles of 30 sec at 94°C, 30 sec at 58°C, and 60 sec at 72°C; and a final cycle of 10 min at 72°C. Each PCR mixture contained 10 µl of Taq premix, 4 µl of water, 1.5 µl of forward primer, 1.5 µl of reverse primer and 3 µl of DNA. The amplicons were visualized in agarose gel-separated (1X TAE buffer; 60 V for 1.30 h) bands previously stained with ethidium bromide under UV illumination.

Genomic Sequencing and Analyses

The extracted DNA was quantified using a Qubit4 Fluorometer (Thermo Fisher Scientific) with an Invitrogen standard kit, and the amplified products were sequenced with a MinION Mk1C (Oxford Nanopore Technologies) Nanopore Sequencer. A sequence similarity search was performed with the basic local alignment search tool (BLAST) from the NCBI database (Altschul et al. 1990), resulting in 100 sequences related to the ITS and rbcL regions. The eight genomic sequences from each ITS and rbcL region were aligned with 100 BLAST sequences using the minimum evolution method and evolutionary analyses (MEGA v11) software (Tamura et al. 2021).

Phenetic analysis was performed to calculate the evolutionary distance between the sequences. The similarity of sequences was evaluated with 1000 bootstrap replications with gamma distribution rates (G). Pairwise deletions were performed to remove the ambiguous positions in each sequence pair. The unweighted pair group method with arithmetic mean (UPGMA) method was employed as a simple clustering algorithm using the Tamura et al. (2004) substitution model. The transition/transversion ratio (R) was calculated, where R tends to be 0.5 when there is no bias towards either substitution. To evaluate the clustering further, the neighbour-joining (NJ) method was also employed using the K2P model substitution according to Kimura 1980.

The cladistic approach involves a maximum likelihood (ML) method to construct the most probable phylogenetic tree, which helps in understanding taxonomic evolution (Tamura et al. 2004). With an automatically constructed initial tree, the most probable ML tree was constructed using the nearest-neighbour-interchange (NNI) heuristic method with 1000 bootstraps.

The evolutionary relationships of taxa (ERT) were inferred using the minimum evolution (ME) method (Rzhetsky and Nei 1992). The evolutionary distances in terms of the number of base substitutions per

site were computed using the maximum composite likelihood method (Tamura et al. 2004). The ME tree was searched using the close-neighbour-interchange (CNI) algorithm at a search level of 1 (Nei and Kumar 2000).

Results

Phenetic and Cladistic analyses of the ITS region of *S. spontaneum*

The sequence lengths of the ITS region in the collected samples ranged between 516 and 652 bp. A total of 108 sequences of *S. spontaneum* were aligned in MEGA11, resulting in 657 positions after pairwise deletion. The phenetic approach modelled with a gamma distribution to estimate the rate variation among sites resulted in an overall pairwise distance (d) of 0.2030. This shows that there is a moderate level of diversity among the species analysed.

The simple clustering algorithm UPGMA was used for the calculated sequence pairs via the Tamura-Nei model with a gamma distribution shape parameter of one. The UPGMA tree showed closely related clusters of two commercial specimens and four collected specimens, Kamuthi 001, Salem 001, and PTMTMC 005, 102, 103 and 104. However, PTMTMC 002 was grouped in another cluster with a relatively close distance of 0.0209, and PTMTMC 108 was grouped in another cluster with a distance of 0.0443 from the six-taxon clade. The NJ clustering algorithm resulted in an optimal tree after 1000 bootstrap replications (Fig. 2). The evolutionary distances in this NJ tree were calculated using the K2P model in units of the number of base substitutions per site. In the NJ clustering tree, six specimens, Kamuthi 001, Salem 001, PTMTMC 005, 102, 103 and 104, were grouped in a single clade with a close relationship. However, PTMTMC 002 was grouped in a close clade with a distance of 0.0039. Similarly, PTMTMC 008 was also grouped in another clade at a distance of 0.0164 from the six-taxon clade.

The cladistic approach was also employed for constructing a maximum likelihood tree with a gamma distribution and was scaled with the number of substitutions per site. The Tamura-Nei statistical model was used to construct ML tree with superior topology, where nearest-neighbour-interchange (NNI) was used as a heuristic search method. The tree had a maximum log likelihood of -2778.37 with a gamma distribution (+ G) of 5 categories of 0.5778, indicating good fitness of the model.

The evolutionary relationships of taxa (ERT) were inferred using the minimum evolution (ME) method (Rzhetsky and Nei 1992). The optimal tree was drawn to scale, where the Tamura-Nei statistical model with the NJ method was used to compute the evolutionary distances in units of the number of base substitutions per site (Fig. 3). The ME tree was searched using the close-neighbour-interchange (CNI) algorithm. Seven of the collected specimens fell into a single clade of ERT with a close relationship; among these, PTMTMC 104 seemed to have evolved most recently (collected from the Srivaikundam, Thamirabarani River). This finding suggested a common ancestor for all these specimens and a distinct evolutionary history from that of other subpopulations. As in the phenetic analyses, PTMTMC 008 (collected from Papanasam, Thamirabarani River) diverged and clustered in a closely related neighbouring clade.

The transition-transversion bias (R) was estimated, where the substitution pattern and rates were modelled under the Tamura-Nei method (+ G). Previously, the initial tree for this tv/tr bias was constructed with the gamma-distributed ML statistical method. The discrete gamma distribution was used to model the evolutionary rate among the top 5 categories of sites (0.03, 0.20, 0.52, 1.15, 3.10 substitutions per site), resulting in a + G value of 0.5787. The maximum log likelihood for this computation was - 2777.390. The ITS nucleotide composition used in this study was more biased towards C and G, as the composition was A = 19.27%, T = 16.72%, C = 32.41%, and G = 31.60%. An NJ tree constructed via the ML method was used to estimate the substitution matrix. The substitution pattern and rates showed that transitions (78.24%) were more likely than transversions (21.76%) in this dataset, as a result of the ML substitute matrix estimation (Table 3).

Table 3
Maximum Likelihood Estimate of Substitution Matrix for the
ITS region

	A	T	C	G
A	-	1.82 [#]	3.52 [#]	13.93 [*]
T	2.10 [#]	-	36.82 [*]	3.44 [#]
C	2.10 [#]	19.00 [*]	-	3.44 [#]
G	8.49 [*]	1.82 [#]	3.52 [#]	-
Substitutions marked with *=transitional; #=transversional				

In summary, these results align with expectations for plant ITS regions, and the genomic region is suitable for exploring the genetic diversity, phylogenetic relationships, and evolutionary adaptations of *S. spontaneum*.

Phenetic and Cladistic analyses of the rbcL region of *S. spontaneum*

rbcL was chosen as the barcoding region for the evolutionary and divergence studies. There were 85 nucleotide sequences with a total of 590 positions, including the 1st, 2nd, 3rd and noncoding regions of rbcL. The phenetic approach involved a maximum likelihood estimate of the gamma parameter for site rates, and the resulting shape parameter for the discrete gamma distribution was 0.1. Substitution patterns and rates were estimated under the Jukes–Cantor model (+ G) (Jukes and Cantor 1969). The overall pairwise distance (d) was 0.00004, which infers poor divergence and intraspecific diversity.

This was further confirmed by the UPGMA and NJ phylogenetic trees constructed using the Jukes-Cantor Model to identify the evolutionary distances among the tested sequences. According to both clustering algorithms, all the tested specimens were inclusively clustered in a single clade with an absolute absence of diversification (Fig. 4). Similarly, the cladistic approach using the same JC model resulted in a

phylogenetic tree with the highest log likelihood (-826.86) that also did not reveal any significant diversification among the tested sequences.

The average nucleotide composition of the tested sequences was more balanced, with T = 29.09, C = 21.57, A = 26.20, and G = 23.14. Substitution patterns and rates were estimated under the Kimura-2-parameter model (+ G). The estimated transition/transversion bias (R) for the rbcL region was 2819.98. This revealed a strong bias towards transitions over transversions and was supported by the substitution matrix of transition and transitional substitutions in a 1:3 ratio (Table 4).

Table 4
Maximum Likelihood Estimate of the Substitution Matrix for the rbcL Region

	A	T	C	G
A	-	8.33 [#]	8.33 [#]	8.33 [*]
T	8.33 [#]	-	8.33 [*]	8.33 [#]
C	8.33 [#]	8.33 [*]	-	8.33 [#]
G	8.33 [*]	8.33 [#]	8.33 [#]	-
Substitutions marked with *=transitional; #=transversional				

Discussion

Internal transcribed spacer (ITS) is the spacer DNA situated between the small- and large-subunit rRNA genes either in the chromosome or in the transcript. The superior properties of this small nrDNA region make it a more reliable molecular marker in angiosperm systematics (Baldwin et al. 1995; Yang et al. 2016; Packia Lekshmi et al. 2024). Previously, the ITS region alone or in combination with other molecular markers was found to be a reliable indicator of phylogenetic relationships in *Saccharum* species, including *S. spontaneum* and *S. officinarum* (Yang et al. 2016). The gamma distribution values suggested that some sites in this ITS region are more conserved, while some are rapidly evolving with high substitution rates. It is often recommended that regions with higher evolutionary rates are suitable for better phylogenetic inference (Bacci et al. 2001). Some sites in the ITS region showing a high rate of evolutionary change ensure the suitability of this region for phylogenetic studies of *S. spontaneum*, along with other subtribes (Hodkinson et al. 2002).

Both the NJ and UPGMA clustering algorithms showed a similar grouping of six specimens (two commercial and four collected) in the present study, with a close relationship evidenced by the evolutionary distances. This close relationship may be attributed to potentially shared adaptations or geographic origins of the riverine ecosystems on the plains. However, in both clustering analyses of the ITS region, specimen PTMTMC 008 was found to have distinctly diverged from its common ancestor.

This is possibly due to specific mutations or selection pressures unique to this specimen collected from Papanasam (the Thamirabarani River) and recent diversification within the studied specimens. This recent diversification of PTMTMC 008 (Srivaikundam) possibly features interspecific diversity despite its collection from the same riverine sites (the Thamirabarani River) as PTMTMC 104. The evolutionary distance between these two specimens was 0.038, which might be attributed to several factors. Microenvironmental differences among these collection sites could drive genetic drift—evidenced by a high rate of evolution at a single site—leading to a specific evolutionary trajectory.

The ITS region along with *rbcL* was chosen for molecular barcoding to understand the taxonomic evolution and intraspecific diversification of *S. spontaneum*. Barcoding analyses with a combined approach involving coding and noncoding regions were found to be more effective than those involving separate methods (Kress 2017; Algarni 2022). Earlier, the *rbcL* region was found to undergo varying rates of evolutionary change; thus, the selection of *rbcL* for phylogenetic studies requires exploratory studies, especially in *Saccharum* species (Bousquet et al. 1992; Yun-Wu et al. 2002). In the present study, the evolutionary rate of the *rbcL* region of *S. spontaneum* was very slow, and the sites were highly conserved. Moreover, *rbcL* barcoding cannot efficiently discriminate the interspecific diversity of *S. spontaneum*, whereas it can serve as a valuable candidate marker for the molecular identification of *S. spontaneum*.

Conclusion

These results show the distinction between ITS and *rbcL* as a combined approach of using coding and noncoding regions as molecular markers; their utility can vary greatly depending on the research objective, whether it is species identification, understanding evolutionary relationships, or exploring genetic diversity. Although some regions, such as the ITS region, can provide detailed insights into evolutionary relationships and species differentiation, *rbcL*, despite its limited phylogenetic resolution, is valuable for species identification due to its conserved nature. These findings emphasize the importance of marker selection in genetic studies, tailored to the specific objectives of understanding evolutionary dynamics or achieving accurate species identification. However, the phylogeny of the *rbcL* region identified in the present study cannot be utilized for molecular taxonomy or molecular evolution studies but can be used for molecular identification because of the highly conserved sites of this region in *S. spontaneum*.

Declarations

Data Accessibility

GenBank Accession Numbers

ON692525, PP106511, PP106512, PP106513, PP106514, PP106515, PP106516, PP106517

Author Contribution

All the authors contributed equally to the idea of the project concept, production, assembly, analysis and writing of the manuscript.

Acknowledgements

The authors gratefully acknowledge the Department of Science & Technology, Government of India, and thank Mr. K. Kalaiselvan and Mr. R. Ramar for their assistance in sample collection.

Funding

This study was funded under the DST-SHRI project by the Department of Science & Technology (DST), Government of India (Grant number: DST/TDT/SHRI-27/2021).

Conflict of interest

The authors declare no actual or potential conflicts of interest.

Ethics approval

No approval is required for this study.

References

1. Aitken K, Li J, Piperidis G, Qing C, Yuanhong F, Jackson P (2018) Worldwide genetic diversity of the wild species *Saccharum spontaneum* and level of diversity captured within sugarcane breeding programs. *Crop Sci* 58:218–229. <https://doi.org/10.2135/cropsci2017.06.0339>
2. Algarni AA (2022) Molecular identification and phylogenetic analysis of *Aloe shadensis* from Saudi Arabia based on matK, rbcL and ITS DNA barcode sequence. *Saudi J Biol Sci* 29:1125–1133. <https://doi.org/10.1016/j.sjbs.2021.09.053>
3. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
4. Alwala S, Suman A, Arro JA, Veremis JC, Kimbeng CA (2006) Target region amplification polymorphism (TRAP) for assessing genetic diversity in sugarcane germplasm collections. *Crop Sci* 46:448–455. <https://doi.org/10.2135/cropsci2005.0274>
5. Bacci Jr M, Miranda VFO, Martins VG, Figueira AVDO, Lemos MV, Pereira JO, Marino CL (2001) A search for markers of sugarcane evolution. *Genet Mol Biol* 24:169–174. <https://doi.org/10.1590/S1415-47572001000100023>
6. Bafeel SO, Arif IA, Bakir MA, Khan HA, Farhan AH, Homaidan AA, Ahamed A, Thomas J (2011) Comparative evaluation of PCR success with universal primers of maturase K (matK) and ribulose-1, 5-bisphosphate carboxylase oxygenase large subunit (rbcL) for barcoding of some arid plants. *Plant Omics* 4:195–198. <https://search.informit.org/doi/10.3316/informit.310811888981920>

7. Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ (1995) The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Ann Missouri Bot Gard* 82:247–277. <https://doi.org/10.2307/2399880>
8. Bharadwaj B, Dutta S, Qaiyum MA, Samal PP, Dey B, Dey S (2023) Pristine wild sugarcane (*Saccharum spontaneum*) as a biosorbent for removal of methylene blue from wastewater: isotherm, kinetics and regeneration studies. *Int J Phytoremediation*. <https://doi.org/10.1080/15226514.2023.2260002>
9. Bhattacharya A (2019) Effect of High Temperature on Protein Metabolism in Plants. In: Bhattacharya A (ed) *Effect of High Temperature on Crop Productivity and Metabolism of Macro Molecules*. Academic Press, Elsevier, London, pp 217–309. <https://doi.org/10.1016/B978-0-12-817562-0.00003-3>
10. Bhuiyan SA, Garlick K, Piperidis G (2019) *Saccharum spontaneum*, a novel source of resistance to root-lesion and root-knot nematodes in sugarcane. *Plant Dis* 103:2288–2294. <https://doi.org/10.1094/PDIS-02-19-0385-RE>
11. Bousquet J, Strauss SH, Doerksen AH, Price RA (1992) Extensive variation in evolutionary rate of *rbcl* gene sequences among seed plants. *Proc Natl Acad Sci* 89:7844–7848. <https://doi.org/10.1073/pnas.89.16.7844>
12. Dharshini S, Manoj VM, Suresha GS, Narayan JA, Padmanabhan TS, Kumar R, Meena MR, Manickavasagam M, Ram B, Appunu C (2020) Isolation and characterization of nuclear localized abiotic stress responsive cold regulated gene 413 (*SsCor413*) from *Saccharum spontaneum*. *Plant Mol Biol Rep* 38:628–640. <https://doi.org/10.1007/s11105-020-01224-z>
13. Dhungana P, Prajapati B, Bhatt P, Regmi D, Yadav M, Maharjan S, Lamsal U, Kathariya S, Chaudhary P, Joshi J (2023) Production of bioethanol from *Saccharum spontaneum* by simultaneous saccharification and electro-fermentation using mixed culture of microbes. *Biofuels* 14:191–199. <https://doi.org/10.1080/17597269.2022.2126083>
14. Dinesh Babu KS, Janakiraman V, Palaniswamy H et al (2022) A short review on sugarcane: its domestication, molecular manipulations and future perspectives. *Genet Resour Crop Evol* 69:2623–2643. <https://doi.org/10.1007/s10722-022-01430-6>
15. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19:11–15.
16. Ferdous T, Quaiyyum MA, Salam A, Jahan MS (2020) Pulping of bagasse (*Saccharum officinarum*), kash (*Saccharum spontaneum*) and corn stalks (*Zea mays*). *Curr Res Green Sustain Chem* 3:100017. <https://doi.org/10.1016/j.crgsc.2020.100017>
17. Govindaraj P, Gowri R, Mohanraj K, Amalraj VA (2021) SSR marker based molecular genetic diversity analysis among *Saccharum spontaneum* (L.) collected from North Western region of India. *Sugar Tech* 23:730–740. <https://doi.org/10.1007/s12355-021-00956-w>
18. Hodkinson TR, Chase MW, Lledó DM, Salamin N, Renvoize SA (2002) Phylogenetics of *Miscanthus*, *Saccharum* and related genera (Saccharinae, Andropogoneae, Poaceae) based on DNA sequences

- from ITS nuclear ribosomal DNA and plastid *trnL* intron and *trnL-F* intergenic spacers. J Plant Res 115:381–392. <https://doi.org/10.1007/s10265-002-0049-3>
19. Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: Munro HN (Ed), Mammalian protein metabolism, Volume III Academic Press, New York, pp. 21–132.
 20. Kasirajan L, Valiyaparambath R, Velu J, Hari H, Srinivasavedantham V, Athaiappan S (2021) Gene expression studies of *Saccharum spontaneum*, a wild relative of sugarcane in response to salinity stress. Biotechnol Appl Biochem 68:288–296. <https://doi.org/10.1002/bab.1923>
 21. Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120. <https://doi.org/10.1007/BF01731581>
 22. Kress WJ (2017) Plant DNA barcodes: Applications today and in the future. J Syst Evol 55:291–307. <https://doi.org/10.1111/jse.12254>
 23. Lai YH, Lin KH, Liu CP, Liao TS, Huang MY, Wang CW, Chen CI (2022) Photosynthetic responses of *Eulophia dentata*, *Bletilla formosana*, and *Saccharum spontaneum* under various photosynthetic photon flux density conditions. Photosynthetica 60:539–548. <https://doi.org/10.32615/ps.2022.050>
 24. Li X, Guo Y, Huang F, Wang Q, Chai J, Yu F, Wu J, Zhang M, Deng Z (2022) Authenticity identification of *Saccharum officinarum* and *Saccharum spontaneum* germplasm materials. Agron 12:819. <https://doi.org/10.3390/agronomy12040819>
 25. Luo T, Li CN, Yan R, Huang K, Li YR, Liu XY, Lakshmanan P (2023) Physiological and molecular insights into the resilience of biological nitrogen fixation to applied nitrogen in *Saccharum spontaneum*, wild progenitor of sugarcane. Front Plant Sci 13:1099701. <https://doi.org/10.3389/fpls.2022.1099701>
 26. Mahardhika EA, Hikmah VM, Sutaryo S, Purnomoadi A (2022) The effect anaerobic co-digestion Kans grass (*Saccharum spontaneum*) and dairy cow manure on biogas digester performance. Livest Res Rural Dev 34. <http://www.lrrd.org/lrrd34/12/34115soet.html>
 27. Mary S, Nair NV, Chaturvedi P, Selvi A (2006) Analysis of genetic diversity among *Saccharum spontaneum* L. from four geographical regions of India, using molecular markers. Genet Resour Crop Evol 53:1221–1231. <https://doi.org/10.1007/s10722-005-2433-z>
 28. Moore PH, Paterson AH, Tew T (2013) Sugarcane: the crop, the plant, and domestication. In: Moore PH, Botha FC (eds) Sugarcane: physiology, biochemistry, and functional biology, John Wiley & Sons, Inc., Iowa, USA, pp 1–17. <https://doi.org/10.1002/9781118771280.ch1>
 29. Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press. New York
 30. Packia Lekshmi NCJ, Brindha JR, Mahamuni D, Rajesh R (2024) Genomic profiling reveals intraspecific divergence and interspecific diversity in *Allium cepa* L.. Genet Resour Crop Evol. <https://doi.org/10.1007/s10722-024-01870-2>
 31. Paterson AH, Moore PH, Tew TL (2012) The gene pool of *Saccharum* species and their improvement. In: Paterson A (ed) Genomics of the Saccharinae. Plant Genetics and Genomics: Crops and Models, vol 11. Springer, New York, pp 43–71. https://doi.org/10.1007/978-1-4419-5947-8_3

32. Prens J, Kurt Z, James Rivas AM, Chen J (2022) Production and characterization of wild sugarcane (*Saccharum spontaneum* L.) biochar for atrazine adsorption in aqueous media. *Agron* 13:27. <https://doi.org/10.3390/agronomy13010027>
33. Rzhetsky A, Nei M (1992) A simple method for estimating and testing minimum-evolution trees. *Mol Biol Evol* 9:945–967. <https://doi.org/10.1093/oxfordjournals.molbev.a040771>
34. Samantray R, Karnan M, Subramani K, Anjeline CJ, Mishra SC, Sathish M (2021) A facile approach to fabricate *Saccharum spontaneum*-derived porous carbon-based supercapacitors for excellent energy storage performance in redox active electrolytes. *Sustain Energy Fuels* 5:518–531. <https://doi.org/10.1039/D0SE01420F>
35. Shahzad S, Saeed M, Perveen S et al (2022) Genetic diversity and population structure assessed through simple sequence repeat markers in *Saccharum* spp. hybrids from Pakistan and Sri Lanka. *Genet Resour Crop Evol* 69:2889–2900. <https://doi.org/10.1007/s10722-022-01411-9>
36. Sheshadri KG (2013) Grasses and their varieties in Indian literature. *Asian Agri-History* 17:325–334.
37. Sobhakumari VP (2020) Evolutionary dynamics of chromosomal diversity in *Saccharum spontaneum* population from Punjab and Haryana, India. *Cytologia* 85:313–318. <https://doi.org/10.1508/cytologia.85.313>
38. Sreenivasan TV (1975) Cytogenetical studies in *Saccharum spontaneum*. *Proc Indian Acad Sci*: 81:131–144. <https://doi.org/10.1007/BF03050754>
39. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc Natl Acad Sci* 101:11030–11035. <https://doi.org/10.1073/pnas.0404206101>
40. Tamura K, Stecher G, Kumar S (2021) MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol* 38:3022–3027. <https://doi.org/10.1093/molbev/msab120>
41. Tew TL, Pan YB (2010) Microsatellite (Simple Sequence Repeat) Marker-based Paternity Analysis of a Seven-Parent Sugarcane Polycross. *Crop Sci* 50:1401–1408. <https://doi.org/10.2135/cropsci2009.10.0579>
42. Vaid S, Sharma S, Dutt HC, Mahajan R, Bajaj BK (2022) An eco-friendly novel approach for bioconversion of *Saccharum spontaneum* biomass to biofuel-ethanol under consolidated bioprocess. *Bioresour Technol* 363:127784. <https://doi.org/10.1016/j.biortech.2022.127784>
43. White TJ, Bruns T, Lee S, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) *PCR Protocols: A Guide to Methods and Applications*, Academic Press, Inc., New York, pp. 315–322
44. Yang CF, Yang LT, Li YR, Zhang GM, Zhang CY, Wang WZ (2016) Sequence characteristics and phylogenetic implications of the nrDNA internal transcribed spacers (ITS) in Protospecies and Landraces of sugarcane (*Saccharum officinarum* L.). *Sugar Tech* 18:8–15. <https://doi.org/10.1007/S12355-014-0355-9>
45. Yu XH, Wang XH, Yang QH (2019) Genetic diversity and phylogenetic relationship of *Saccharum spontaneum* L. with different ploidy levels based on SRAP markers. *Sugar Tech* 21:802–814.

46. Yun-Wu Z, Huo-Sheng L, Yuan-Hong F, Yong-Gang Y, Qing C, Ya-Ping Z (2002) Sequence variation of *rbcL* gene and evolution of *Saccharum* and related species. *Acta Bot Yunnan* 24:29–36.

Figures

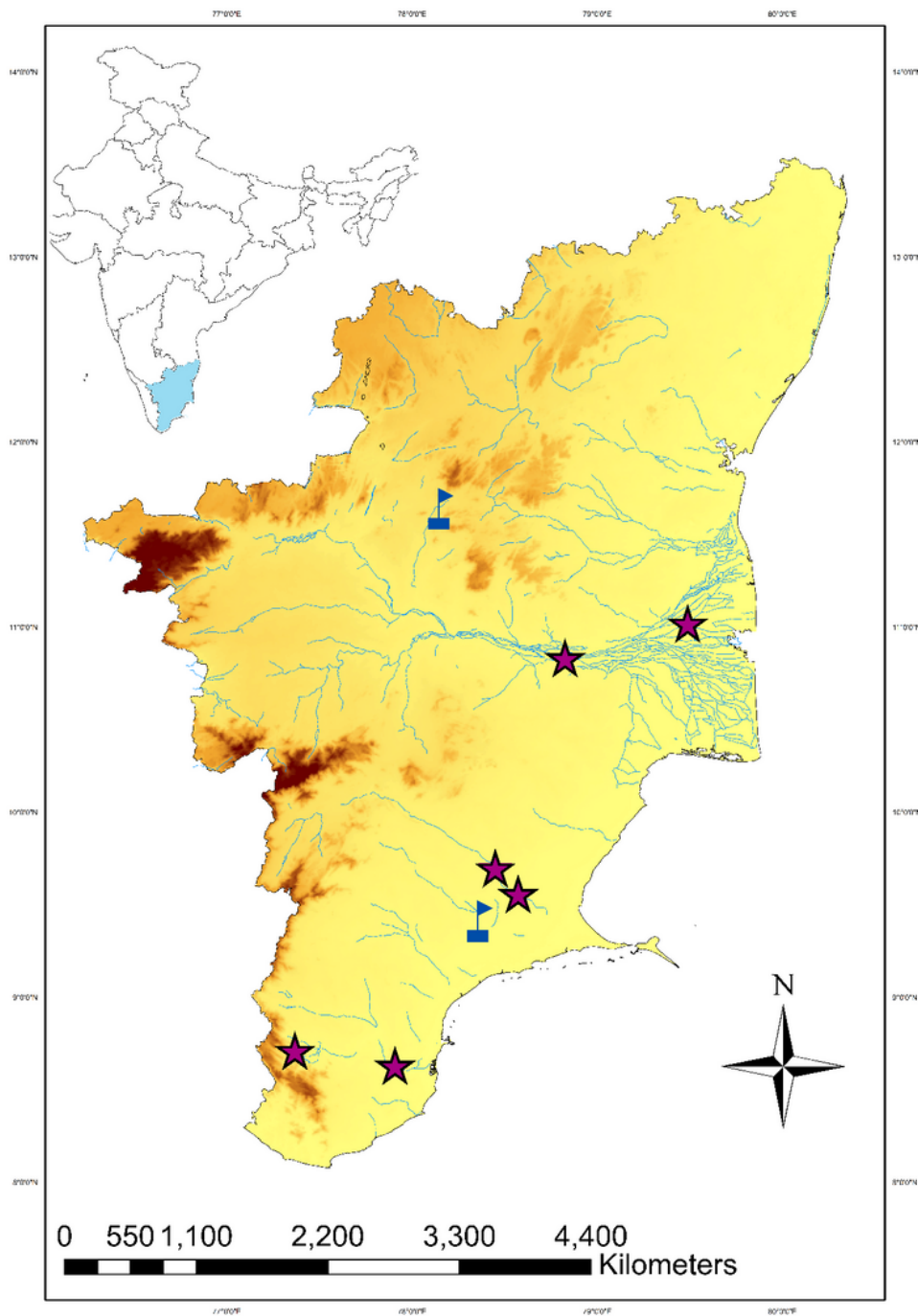


Figure 1

*S. spontaneum*samplingsites: *-field samples; ►- commercial sample

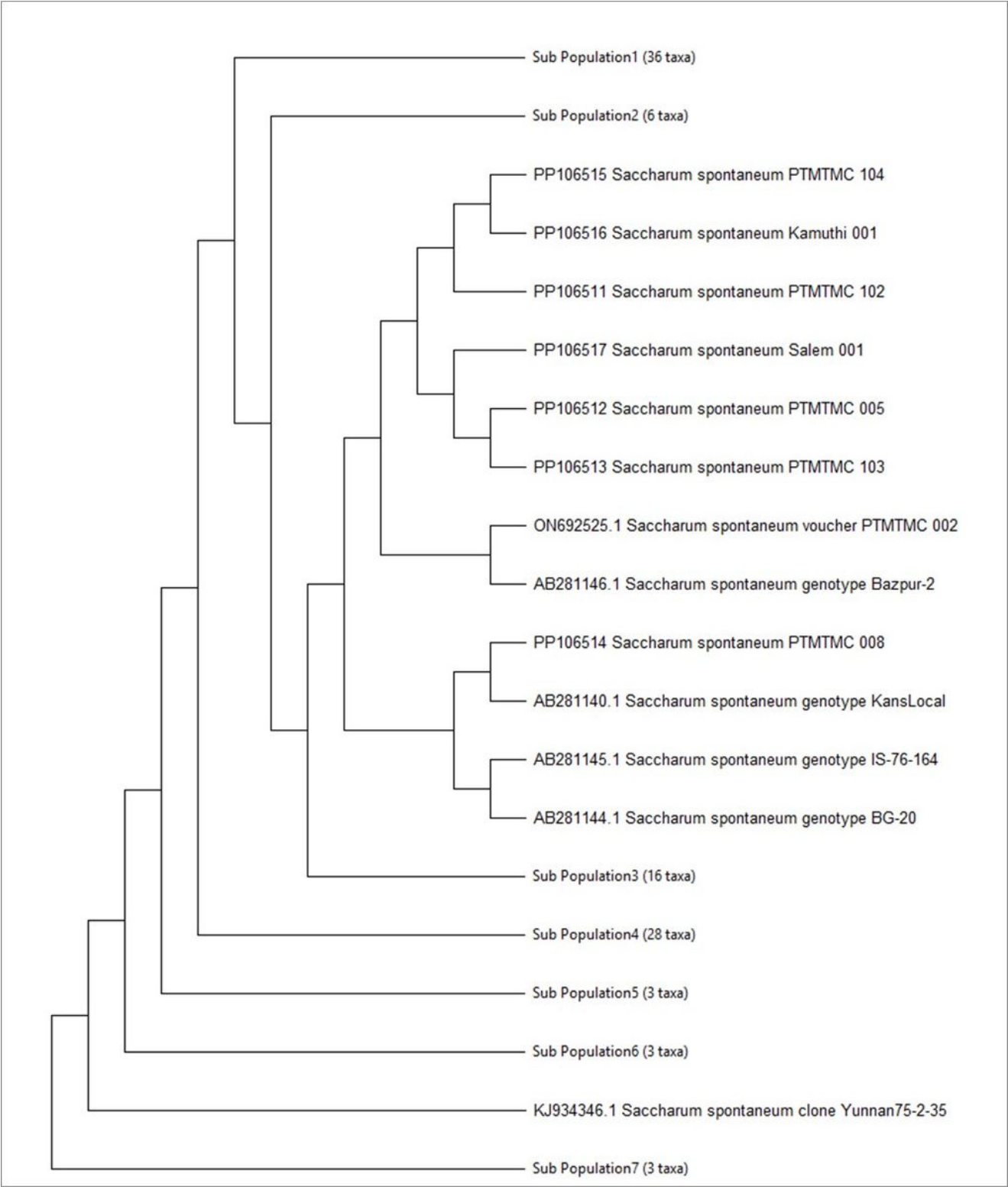


Figure 2

Neighbour joining (NJ) tree of the ITS region obtained via phenetic analysis

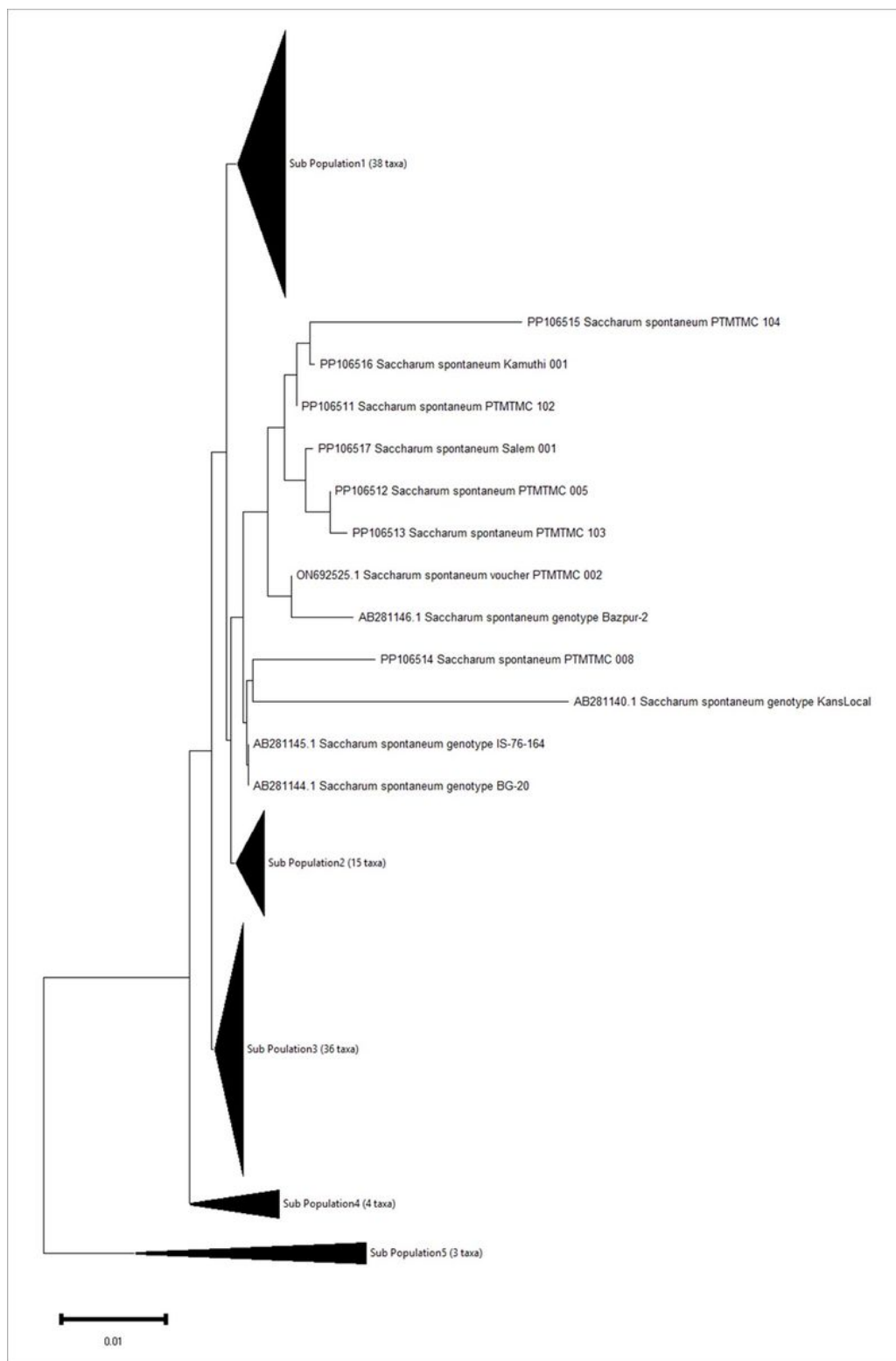


Figure 3

Cladistic analysis of the evolutionary relationships of taxa (ERT) in the ITS region

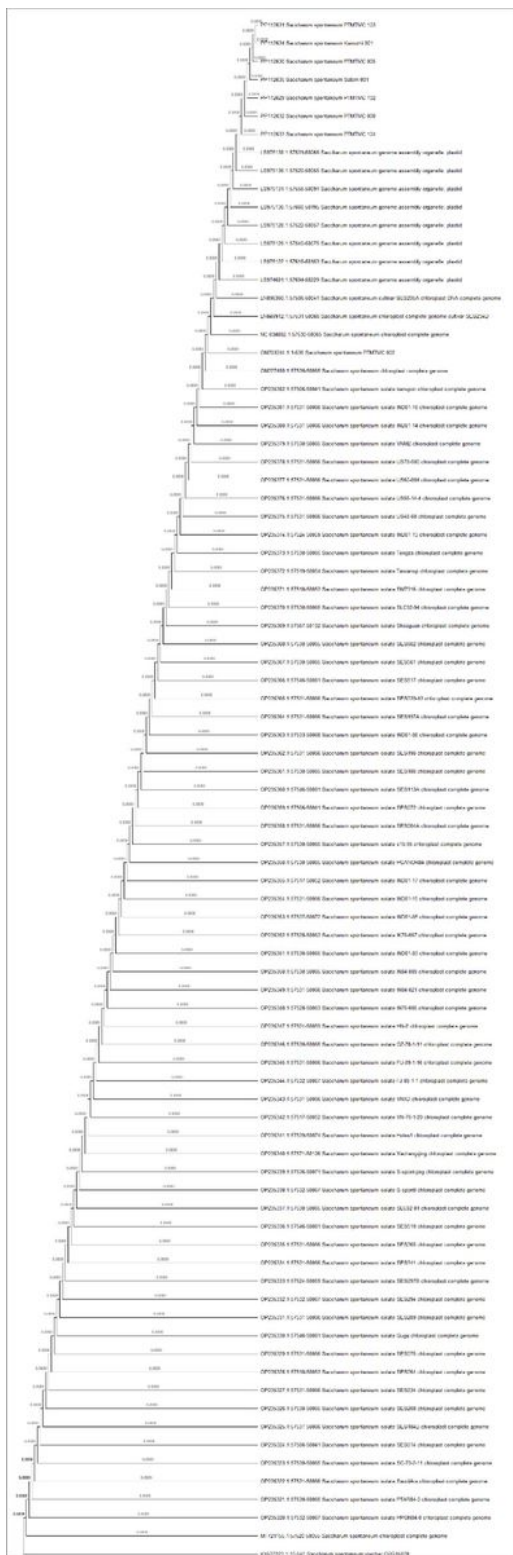


Figure 4

Neighbour joining (NJ) tree of the genetic analysis of the rbcL region