RESEARCH PAPER

Molecular characterization of *Pavetta bourdillonii* Sivad. & N. Mohanan using the chloroplast matK barcode

Ezhumalai Umamageshwari[1], Rajagopalan Rajendran[1*]

[1]Department of Botany, Annamalai University, Cuddalore, Tamil Nadu, India.

*** (Corresponding Author) Email: drrajendranau@gmail.com**

Abstract

Pavetta bourdillonii Sivad. & N. Mohanan is a Rubiaceae species found in Kerala's Agasthyamala Hills. The species was first discovered in the Attayar region of the Agasthyamala Hills and is associated with the Southern Western Ghats evergreen forest habitat. The chloroplast matK DNA barcode region was used to molecularly characterize *P. bourdillonii* in this investigation. Fresh young leaf material was collected from the Bonacaud forest region of the Agasthiyar Hills in Thiruvananthapuram District, Kerala and genomic DNA was extracted using a modified CTAB technique. The matK region was effectively amplified through polymerase chain reaction (PCR), yielding distinct amplicon of around 800 bp. A 431 bp consensus sequence was obtained using bidirectional Sanger sequencing followed by sequence editing. The sequence was submitted to the NCBI GenBank database under the accession number PZ240881. BLASTn analysis revealed 99.3% sequence identity to *Pavetta hongkongensis*, with 100% query coverage. A collection of 50 homologous sequences was then collected for comparative sequencing analysis and phylogenetic analysis. The study provides a molecular sequence for *P. bourdillonii* and adds to the DNA barcode reference data for the same.

Keywords: DNA barcoding, GenBank, matK, molecular characterization, *Pavetta bourdillonii*, Rubiaceae, Southern Western Ghats.

How to cite this article: Umamageshwari E, Rajendran R., Molecular Characterization of Pavetta bourdillonii Sivad. & N. Mohanan Using the Chloroplast matK Barcode. Int J Drug Deliv Technol. 2026;16(90s): 1594-1603; DOI: 10.25258/ijddt.16.90s.162

1. Introduction

Pavetta L. belongs to the Rubiaceae family, specifically the Pavetteae tribe of the Ixoroideae subfamily. The genus includes shrubs and small trees found primarily in tropical and subtropical parts of the Old World. *Pavetta bourdillonii* Sivad. & N. Mohanan was reported as a separate species from the Agasthyamala Hills in Thiruvananthapuram District, Kerala based on morphological traits that distinguished it from related taxa such as *P. concanica*, *P. laeta* and *P. travancorica* (Sivadasan and Mohanan, 1999). The species was discovered in the Attayar region at an elevation of 700-750 m and was described as a little evergreen understory tree linked with evergreen woodland along the Attayar river.

Pavetta's taxonomic history has included discrepancies in species delimitation and interpretation of morphological features. Kumar and Mohanan (2014) reviewed *P. bourdillonni*'s taxonomic status, highlighting historical variances in species conceptions within the genus *P. bourdillonii* is officially recognized as an acceptable name in Kerala, India, while different taxonomic treatments have been connected with the species. Such taxonomic issues highlight the significance of combining morphological data with genetic evidence for defining species with limited ranges.

DNA barcoding is a molecular method for species identification that involves an analysis of specified genomic regions. Chloroplast markers such as rbcL and matK have been widely studied in plants because they can provide valuable sequencing information for taxonomic identification and comparative investigations. The matK region, found inside the chloroplast genome, has been extensively employed in plant molecular systematics due to its comparatively high sequence variation when compared to other chloroplast coding regions (CBOL Plant Working Group,2009; Kress and Erickson, 2007). Recent studies also emphasized the use of plant DNA barcoding for species identification, taxonomy, biodiversity assessment and conservation research (Antil *et al*., 2023; Letsiou *et al*., 2024).

Within Rubiaceae, molecular data have made significant contributions to understanding taxonomic and tribal relationships. Chloroplast based phylogenetic studies of the subfamily Ixoroideae have identified *Pavetta* within the tribe Pavetteae and proved the utility of plastid sequence data for exploring group relationships (Ly *et al*., 2020). However, species level DNA barcode information is not equally accessible to all members of the genus.

Therefore, the present study was aimed to generate molecular barcode information for *Pavetta bourdillonii* using the chloroplast matK region. The goals were to produce a high-quality matK sequence from authenticated plant material, compare it to available database record using BLASTn, deposit the sequence in GenBank and contribute a reference dataset for future phylogenetic research.

2. Materials and Methods

2.1. Plant collection and Authentication

Fresh young *Pavetta bourdillonii* leaves were collected from the Bonacaud forest, Agasthiyar Hills, Thiruvananthapuram District, Kerala, India. The plant material was authenticated by the Institute of Forest Genetics and Tree Breeding (IFGTB) in Coimbatore, India. Fresh leaves were collected with sterile scissors and forceps, placed in labeled Zip-lock bags and refrigerated at -20 °C for DNA extraction



Figure 1: *Pavetta bourdillonii* Sivad. & N. Mohanan

2.2. Taxonomic Classification

Kingdom : Plantae
Phylum : Streptophyta
Class : Equisetopsida
Order : Gentianales
Family : Rubiaceae
Genus : *Pavetta*
Species : *bourdillonii*

2.3. DNA isolation

Genomic DNA was extracted from 200 mg of fresh leaf tissue using a modified CTAB approach described by Doyle and Doyle (1987). The leaf tissue was ground using liquid nitrogen and mixed with 500-1000 µL of CTAB extraction buffer. Sample was incubated at 65 °C for 45-50 minutes before being treated with chloroform alcohol (24:1). The DNA was precipitated in cold isopropanol, washed with 70% ethanol, air dried and dissolved in 30 µL of nuclease free water. The isolated DNA was kept at -20 °C until further investigation.

2.4. DNA quality and PCR amplification

The DNA quality was assessed using 1% agarose gel electrophoresis and a nanodrop spectrophotometer. The chloroplast matK region was amplified with universal matK primers. A 30 µL reaction mixture was used for PCR, which included the master mix, forward and reverse primers, template DNA and nuclease free water.

The PCR conditions were as follows: an initial denaturation at 95 °C for 5 minutes, followed by 34 cycles of denaturation at 95 °C for 1 minute, annealing at 48.5 °C for 1 minute and extension at 72 °C for 1 minute, with a final extension at 72 °C for 10 minutes. Electrophoresis on a 1% agarose gel verified the presence of PCR products.

2.5. Sequencing and sequence analysis

The amplified matK product was sequenced in both directions using Sanger sequencing at Bio Kart India Private limited in Bangalore. To generate a consensus sequence, forward and reverse sequences were edited and assembled with Bio Edit.

The consensus sequence was submitted to GenBank through the NCBI system and assigned accession number PZ240881. The similarity of sequences was determined by using BLAST (Altschul *et al*., 1990). Based on sequence similarity and coverage, GenBank selected around 50 similarly comparable matK sequences.

The sequences were aligned using MUSCLE, which is implemented into MEGA11(Felsenstein, 1985). The Jukes-Cantor model was used with 1000 bootstrap replicates to construct a Maximum Likelihood (ML) phylogenetic analysis (Tamura *et al*., 2021)

3. Results

3.1. DNA isolation and matK amplification

Using the modified CTAB approach, genomic DNA was successfully isolated from fresh, young *Pavetta bourdillonii* leaves. The extracted DNA has sufficient quality for PCR amplification, as evidenced by the DNA quality assessment and successful target region amplification. On 1% agarose gel electrophoresis, the chloroplast matK region produced a clear and distinct PCR amplicon about 800 bp. There were no notable non-specific bands detected, indicating that the target matK region was successfully amplified.

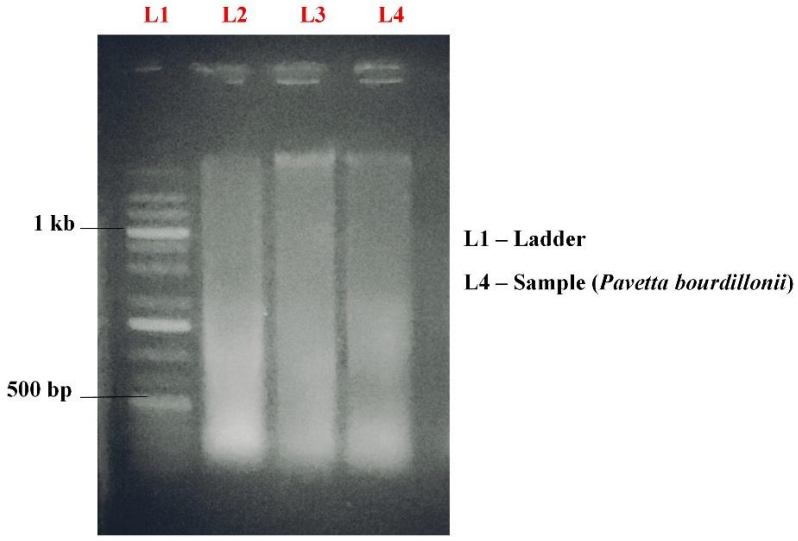

Figure 2: DNA isolation of *Pavetta bourdillonii* verified by agarose gel electrophoresis

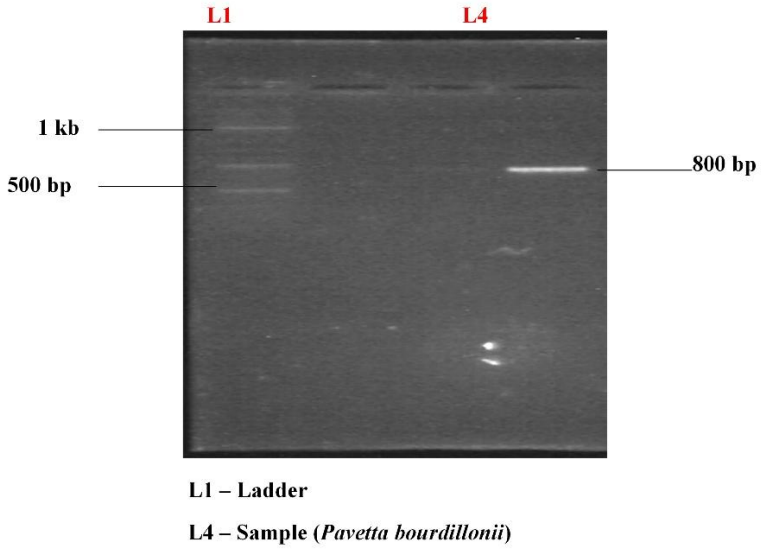

Figure 3: PCR amplification of matK gene in *Pavetta bourdillonii*

3.2. Sequencing and consensus sequence

The amplified matK product underwent bidirectional Sanger sequencing. The forward and reverse sequences were evaluated, edited and removed low-quality terminal ends, and then assembled to form a consensus sequence. *Pavetta bourdillonii* has a high-quality consensus sequence of 431 bp. The sequence was then submitted to the NCBI GenBank database and assigned the accession number PZ240881.This sequence was generated to give a molecular reference record for *P. bourdillonii*. Sequence information from authorized material can contribute to the available barcode data for the genus *Pavetta*.

DNA sequence of *Pavetta bourdillonii*

GATCCTTTTTCCACGAATATTGGAATTGGAATACT
CTTATTGATAGAAAGAAACCCAGTTTTGATTTTTC
ACCAAAAAGAAATCAAAAATTATTCTTCTTGTTAT
ATAATTCTCATGTATGTGAAAACGAATCCATTTTC
TTCTTTCTCCGTAAGCAATCTTCTCATTTGCGATC
AACATCTTTTGGAGTCTTTCTTGAACGGATATATT
TCTATGAAAAAAAAGAACGTCTTGTAGGAGTCTT
TACTAAGGATTTTCAAGCCAGTCTATGGTTGTTCA
AAGATCCTTTCATACATTATGTTAGGTATCAAGGA
AAATCCATTCTGGTTTCAAATGGCACGCCTCTTTT
GATGAATAAATGGAAATCTTACCTTGTCAATTTTT
GGCAATGTCATTTTGGCATTTGGTTTCACTCCAG
AAGGGGCTATATA

3.3. BLAST sequence similarity analysis

The 431 bp consensus sequence was compared to the NCBI database using BLASTn. The analysis revealed a maximum sequence similarity of 99.3% to *Pavetta hongkongensis*, with 100% query coverage. The high percentage of sequence identity, together with complete query coverage showed a strong sequence similarity between the obtained *P. bourdillonii* matK sequence and the corresponding *Pavetta* reference sequence.

The BLAST result offered a database-based molecular comparison for the newly produced sequence, supporting its placement within the *Pavetta* genus. However, the BLAST results were considered together with the phylogenetic analysis rather than being utilized alone to determine species identity.

Table 1: Study species BLAST results

Species name	Closest taxon	Accession number	Identity	Query coverage	E-value
Pavetta bourdillonii	*Pavetta hongkongensis*	KR531339.1	99.3%	100%	0.0

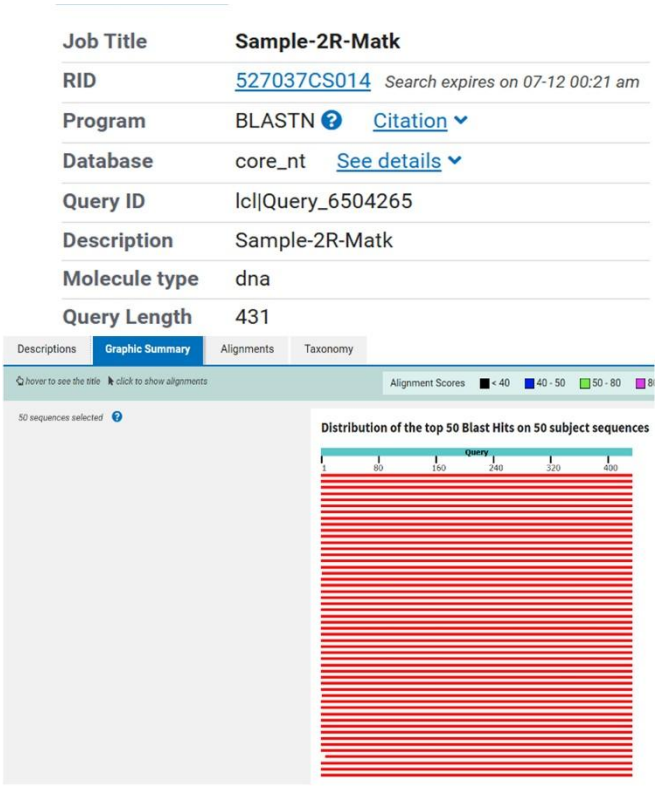

Figure 4: Blast format and graphical summary of *Pavetta bourdillonii*

3.4. Multiple sequence alignment

For comparative and phylogenetic analysis, 50 homologous matK sequences were selected from the NCBI database based on sequence similarity and coverage. The obtained *P. bourdillonii* sequence was combined with the retrieved reference sequences and aligned using MUSCLE in MEGA11.

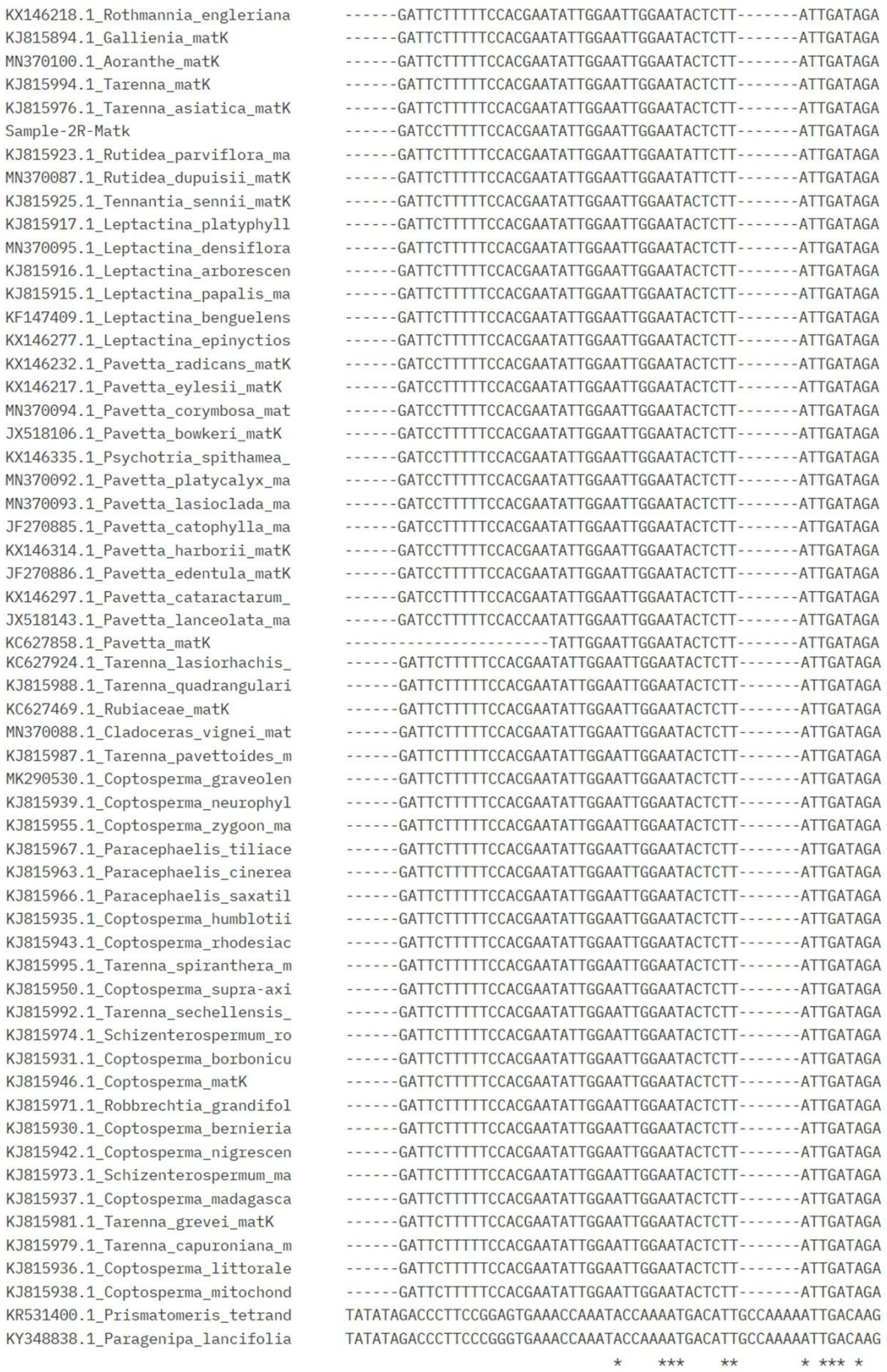

Figure 5: Multiple sequence alignment of *Pavetta bourdillonii* matK sequence with other related taxa

3.5. Phylogenetic tree construction

The aligned dataset was then subjected to maximum likelihood (ML) analysis with the Jukes-Cantor model. 1000 bootstrap replicates were used to assess the reliability of the inferred phylogenetic relationships. The constructed ML tree provided a basis for comparing *P. bourdillonii* to the selected *Pavetta* and related reference sequences.

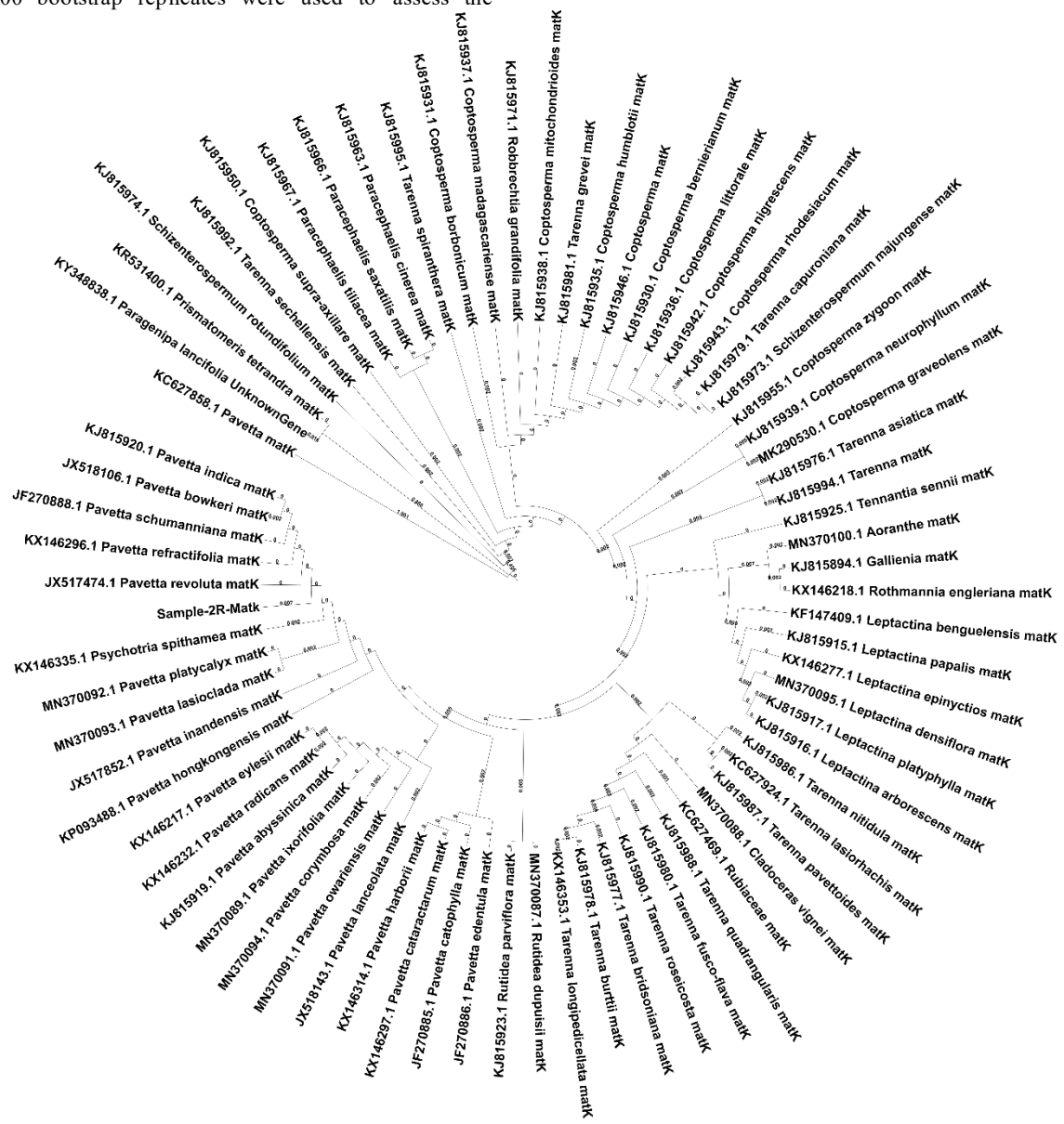

Figure 6: *Pavetta bourdillonii*- Maximum Likelihood- phylogenetic tree construction

4. Discussion

The present study successfully generated a 431 bp consensus matK sequence from morphologically validated *Pavetta bourdillonii* collected from Kerala's Agasthiyar Hills, which was deposited in GenBank under accession number PZ240881. The sequence recovery found in this investigation is comparable to recent matK based plant identification studies. Ramesh *et al.* (2022) analyzed 26 accessions representing seven *Momordica* species from the Indian subcontinent and found that matK sequences included enough nucleotide diversity to distinguish numerous taxa and reconstruct their evolutionary relationships. However, their analysis discovered that matK alone could not distinguish *M. dioica* and *M. sahyadica*. Thus, the current recovery of a usable matK sequence is consistent with the successful use of this marker in other plant groups, while the difficulty in obtaining an unambiguous species level match in *P.bourdillonii* is also consistent with matK's variable discriminatory performance among closely related taxa.

The BLAST result revealed that the *P. bourdillonii* sequence had 99.3% identity and 100% query coverage with *Pavetta hongkongensis*, which was the most significant finding in this investigation. Although

the closest database sequence did not belong to the same species as the morphologically described material, this strong sequence similarity shows that the generated sequence is closely connected to the known *Pavetta* reference sequences. Similar drawbacks of single-locus barcoding have been documented in other plant families, where closely related species might not exhibit enough sequence divergence for accurate species identification (Teklemariam *et al*., 2023). The 99.3% similarity found for *P. bourdillonii* in comparison to that study shows strong sequence affinity, but it should not be taken as a definitive species identification because a high similarity value can occur between closely related species when species level sequence divergence is low or when the correct taxon is not well represented in reference database.

When compared to recent large scale barcode library investigations, the current BLAST result is also significant. Le *et al*. (2025) produced 591 matK sequences and built a regional DNA barcode reference library with 2864 native seed plant species from the Guangdong-Hong Kong-Macao Greater Bay Area. According to their study, using the best match\best close-match approach, matK alone yielded a species level identification rate of 63.29%, however the combination of rbcL+matK+ITS2 gave significantly higher resolution in tree-based analyses. They also underlined how crucial it is to have well curated local reference libraries because insufficient taxonomy database can lower the accuracy of sequence-based identification. Although the sequence showed 100% query coverage and 99.3% similarity to *P. hongkongensis*, the absence of a matching *P. bourdillonii* sequence among the closest reference record limits BLAST's ability to provide a species-specific assignment. As a result, the value of PZ240881 is not only its resemblance to existing *Pavetta* sequences, but also its inclusion in the reference database for future identification of the species.

The molecular ambiguity reported in the current study can also be seen in the context of *Pavetta*'s taxonomic complexity. In a systematic study of *Pavetta* in Thailand, Chantaranothai (2021) identified 14 species and 15 taxa while also making various nomenclatural revisions, such as reducing several previously recognized names to synonymy and adding new regional records. The study demonstrated that species borders and proper name application within *Pavetta* necessitate a thorough analysis of morphological characteristics and specimen data. This taxonomy complexity serves as a significant foundation for interpreting the current molecular outcome. In our research, the morphologically identified *P. bourdillonii* did not produce *P. bourdillonii* as its closest BLAST match, but rather exhibited the strongest similarity to another *Pavetta*. As a result, the genetic finding does not contradict the morphological identification; rather, it illustrates that a single matK sequence may not contain enough diversity to distinguish all closely related *Pavetta* taxa. As a result, the molecular data supplement rather than replace the morphological evidence employed in identification.

The performance of matK demonstrated in this study can be compared to the results of recent species level barcoding studies. Li *et al*. (2021) studied chloroplast barcode regions in Orchidaceae and discovered that while chloroplast markers could distinguish several closely related orchid taxa, their discriminatory ability differed by genus and species. Similarly, Ramesh *et al*. (2022) discovered that matK successfully separated various Momordica species but failed to discriminate between *M. dioica* and *M. sahyadica*. These findings show that matK can give useful species level variation when there is enough nucleotide variation, but its resolving capacity is taxon-dependent. The 99.3% similarity between the *Pavetta* sample and *P. hongkongensis* implies that the matK region has limited diagnostic variation between the sampled taxon and at least some of its close relatives. In contrast to examples where matK generates species-specific sequence differences, the current study demonstrates a case in which matK gives strong evidence of close association but inadequate evidence for definitive species discrimination.

The maximum Likelihood analysis adds to the evidence available for interpreting the BLAST results. In the current tree, the *P. bourdillonii* sequence is included inside the large *Pavetta* assemblage, close to the *P. revoluta – P. refractifolia* region of the tree. This genus level placement is consistent with current phylogenetic treatments of Pavetteae, but the resolution obtained from multiple genomic markers. Using a dataset of 655 sequences representing 115 species and four plastid and two nuclear markers, De Block *et al*. (2026) found strong support for the monophyly of Pavetteae; however, individual marker analyses were reported to be largely unresolved; combining plastid and nuclear datasets significantly improved topology and node support. Additionally, three significant Pavetteae clades whose relationship were still unclear were found in their investigation. The current matK based tree effectively places the sample inside *Pavetta* when compared to their multilocus finding, however the sample's tight placement among individual *Pavetta* species should be considered provisional rather than a conclusive species relationship.

Therefore, the current result differs from studies where a single matK barcode accurately distinguishes species, but it is consistent with research demonstrating that barcode resolution declines in closely related or taxonomically complex groupings. Zhang *et al*. (2023) examined morphologically similar *Rubia* species using nuclear ITS and eight chloroplast markers. They found distinct sectional ties but were unable to fully clarify species relationship within the *Rubia* sect. oligoneura. They came to the conclusion that a more efficient method for authenticating complicated species was to combine morphological, molecular and chemical evidence. Similar patterns are seen in the current *Pavetta* study, the matK sequence offers a close reference match and helpful molecular location, but it is unable to independently determine the species boundary. Therefore, rather than being proof of an inaccurate morphological identification, the difference between the closest molecular match and

the morphological species name should be seen as a finding that needs more research.

Single chloroplast markers may not be sufficient to properly resolve species level relationships, according to recent phylogenomic research. According to Thureborn *et al*. (2024), there is a discrepancy between the nuclear and plastid phylogenies in the Rubiaceae, suggesting that the evolutionary history of a species may not always be fully reflected in chloroplast relationships. Therefore, *P. bourdillonii*'s chloroplast affinity is supported by its placement within the *Pavetta* assemblage; nevertheless, further independent markers are required for species level resolution.

Overall, the current investigation shows that matK was able to provide a molecular sequence for *P. bourdillonii*; nevertheless, the 99.3% similarity with *P. hongkongensis* indicates that there was little species level distinction. The sample's position within the *Pavetta* assemblage is consistent with recent phylogenomic research, and the unresolved species level relationship emphasizes the importance of multilocus methods. Thus, *P. bourdillonii* and related *Pavetta* species can be compared in the future using the GenBank sequence PZ240881 as a future molecular reference.

5. Conclusion

The present study provides the first molecular barcode record for *Pavetta bourdillonii* based on the chloroplast matK region. A 431 bp high-quality sequence was produced from authenticated material collected from Agasthiyar Hills of Kerala and deposited in GenBank as accession number PZ240881. The sample has 99.3% similarity and 100% query coverage to *Pavetta hongkongensis*, and phylogenetic analysis placed the sample into the *Pavetta* branch. These findings give the first publicly available matK sequence for *P. bourdillonii*, establishing a genetic reference for a species previously represented solely by morphological and taxonomic information. Although the single matK locus demonstrated modest resolution among closely related *Pavetta* taxa, the newly generated sequence provides a solid platform for future species authentication, taxonomic assessment and phylogenetic research of *P. bourdillonii* and related taxa.

Acknowledgement

Finally, I gratefully acknowledge the department of Biotechnology, Bharathiar University, Coimbatore, for providing me the laboratory facilities. I also thank Mr. Arjun and Mr. Arunraj for their assistance during field collection in Agasthya Hills. I would like to express my sincere gratitude to my research supervisor for his valuable guidance, encouragement and continuous support throughout this study. I also extend my heartfelt thanks to the UGC-SAP and DST-FIST for providing the instrumental facilities available in the Department of Botany, Annamalai University.

References

- **Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J.** (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), 403-410. 10.1016/S0022-2836(05)80360-2
- **Antil, S., Abraham, J. S., Sripoorna, S., Maurya, S., Dagar, J., Makhija, S., ... & Toteja, R.** (2023). DNA barcoding, an effective tool for species identification: a review. *Molecular biology reports*, *50*(1), 761-775. https://doi.org/10.1007/s11033-022-08015-7
- **CBOL Plant Working Group.** (2009). A DNA barcode for land plants. *Proceedings of the National academy of Sciences of the United States of America*, 12794-12797. https://www.jstor.org/stable/40484604
- **Chantaranothai, P.** (2021). Taxonomic Notes on Pavetta L. (Rubiaceae) from Thailand and New Records for Cambodia, Laos and Myanmar. *Tropical Natural History*, *21*(1), 146-166. https://doi.org/10.58837/tnh.21.1.246811
- **De Block, P., Razafimandimbison, S. G., Rakotonasolo, F., & Janssens, S. B.** (2026). Dismantling the genus Tarenna (Rubiaceae: Pavetteae) in Africa and Madagascar, with the description of three new genera. *Taxon*, *75*(1), e70031. **https://doi.org/10.1002/tax.70031**
- **Doyle, J. J., & Doyle, J. L.** (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical bulletin*.
- **Felsenstein, J.** (1985). Confidence limits on phylogenies: an approach using the bootstrap. *evolution*, *39*(4), 783-791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x
- **Kress, W. J., & Erickson, D. L.** (2007). A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS one*, *2*(6), e508. https://doi.org/10.1371/journal.pone.0000508
- **Kumar, E. S. S., & Mohanan, N.** (2014). Taxonomic status of Pavetta bourdillonii (Rubiaceae). *TAPROBANICA: The Journal of Asian Biodiversity*, *7*(2), 96-97.
- **Le, T. T. M., Jin, L., Wang, R. J., Deng, Y. F., Yan, H. F., & Ge, X. J.** (2025). A DNA barcode reference library of native seed plants in the Guangdong-Hong Kong-Macao Greater Bay Area. *Scientific Data*, *12*(1), 1505. https://doi.org/10.1038/s41597-025-05833-9

- **Letsiou, S., Madesis, P., Vasdekis, E., Montemurro, C., Grigoriou, M. E., Skavdis, G., ... & Tzakos, A. G.** (2024). DNA barcoding as a plant identification method. *Applied Sciences*, *14*(4), 1415. **https://doi.org/10.3390/app14041415**

- **Li, H., Xiao, W., Tong, T., Li, Y., Zhang, M., Lin, X., ... & Guo, X.** (2021). The specific DNA barcodes based on chloroplast genes for species identification of Orchidaceae plants. *Scientific Reports*, *11*(1), 1424. https://doi.org/10.1038/s41598-021-81087-w

- **Ly, S. N., Garavito, A., De Block, P., Asselman, P., Guyeux, C., Charr, J. C., ... & Guyot, R.** (2020). Chloroplast genomes of Rubiaceae: Comparative genomics and molecular phylogeny in subfamily Ixoroideae. *PLoS One*, *15*(4), e0232295. https://doi.org/10.1371/journal.pone.0232295

- **Ramesh, G. A., Mathew, D., John, K. J., & Ravisankar, V.** (2022). Chloroplast gene matK holds the barcodes for identification of Momordica (Cucurbitaceae) species from Indian subcontinent. *Horticultural Plant Journal*, *8*(1), 89-98. 10.1016/j.hpj.2021.04.001

- **Ramesh, G. A., Mathew, D., John, K. J., & Ravisankar, V.** (2022). Chloroplast gene matK holds the barcodes for identification of Momordica (Cucurbitaceae) species from Indian subcontinent. *Horticultural Plant Journal*, *8*(1), 89-98. 10.1016/j.hpj.2021.04.001

- **Sivadasan, M., & Mohanan, N.** (1999). Pavetta bourdillonii (Rubiaceae), a new species from India. *Botanical Bulletin of Academia Sinica*, *40*.

- **Tamura, K., Stecher, G., & Kumar, S.** (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, *38*(7), 3022-3027. https://doi.org/10.1093/molbev/msab120

- **Teklemariam, D. M., Gailing, O., Siregar, I. Z., Amandita, F. Y., & Moura, C. C.** (2023). Integrative taxonomy using the plant core DNA barcodes in Sumatra's Burseraceae. *Ecology and evolution*, *13*(4), e9935. **https://doi.org/10.1002/ece3.9935**

- **Thureborn, O., Wikström, N., Razafimandimbison, S. G., & Rydin, C.** (2024). Plastid phylogenomics and cytonuclear discordance in Rubioideae, Rubiaceae. *PLoS One*, *19*(5), e0302365. https://doi.org/10.1371/journal.pone.0302365

- **Zhang, X., Yang, L. E., Hu, Y., Wu, X., Wang, Z., Miao, Y., ... & Tan, N.** (2023). Integrating morphology, molecular phylogeny and chemotaxonomy for the most effective authentication in Chinese Rubia with insights into origin and distribution of characteristic Rubiaceae-type cyclopeptides. *Industrial Crops and Products*, *191*, 115775. 10.1016/j.indcrop.2022.115775