DNA BARCODING, GENETIC VARIABILITY AND PHYLOGENETICS OF TWO *ADIANTUM* SPECIES FROM SWAT VALLEY

M. Akhtar*[1], Z. Khan[1], A. Razaq[2], M. U. F. Awan[1] and M. Waheed[1]

[1]Department of Botany, Government College University, Lahore, Pakistan
[2]Department of Biological Sciences, University of Veterinary and Animal Sciences, Ravi Campus Pattoki, Pakistan
*Corresponding author's email: uzair.akhtar042@gmail.com; pen_friend_pu@yahoo.com
http://orcid.org/0009-0004-8631-9529

ABSTRACT

Swat Valley, located in Koh Hindu Kush-Himalayan region of Pakistan, is rich in floral diversity. However, climate change is impacting and putting plant diversity at risk, leading to elimination and even extinction of many vulnerable plant species. Effective conservation and monitoring therefore require reliable tools for species identification. Ferns are a key component of the valley's temperate vegetation; however, despite extensive morphological studies, their molecular characterization and phylogenetic relationships remain poorly investigated. Ferns constitute an important component of the valley's temperate vegetation, yet despite extensive morphological work, their phylogenetic relationships and molecular characterization remain largely unexplored and sketchy. This study was undertaken to DNA-barcode two *Adiantum* species and to assess their genetic variability and phylogenetic relationships. Two loci, *rbcLa* and *psbA-trnH* were amplified through polymerase chain reaction (PCR) and sequenced using Sanger method. Morphological diagnosis, combined with BLASTn homology searches, K2P genetic distance analyses, and phylogenetic reconstruction, consistently identified the fern species as *Adiantum capillus-veneris* L. and *A. venustum* D. Don. The results demonstrated the complementary effectiveness of *rbcLa* and *psbA-trnH* for accurate identification and monitoring of regional *Adiantum* species and revealed two likely misannotated sequences in GenBank (MF694655.1; MF694942.1) and one in BOLD systems (DBMPP-195-14). The study also suggested further morphological and multilocus investigation of *A. capillus-veneris* L. and *A. ogasawarense* Tagawa, as two markers used in this study appear to be highly conserved between them. The validated sequences have been deposited in NCBI database and accessions assigned to support future taxonomic, phylogenetic, and conservation studies of the genus.

Key words: *Adiantum capillus-veneris*, *A. venustum*, *rbcLa*, *psbA-trnH*, Phylogenetics, Genetic variability, Swat Valley, Fern diversity, Misannotations

https://doi.org/10.36899/JAPS.2026.4.0098 Published first online May 13, 2026

INTRODUCTION

Global climate change is a menace that affects floral diversity and food security of the world. Swat Valley of Pakistan, located in the northern Koh Hindu Kush-Himalaya temperate region of the country, is rich in floral diversity and the plant populations of immense medicinal, food, furniture and fuelwood importance. Climate change, however, is affecting and causing imbalance in plant community compositions and is a threat to the plant biodiversity (Chettri *et al.*, 2023; Shrestha *et al.*, 2026). Taxonomic identification, characterization and floral record of the plant population and the plant species are of utmost importance and prerequisite to monitor, conserve and restore the flora and vegetation fall of the area against climate change.

Ferns are an essential component of floral community composition of the temperate zones and function as sensitive indicator of loss in floral diversity and plant population of the area against climate shifts. Besides contributing to ecosystem health by carbon sequestration, soil stabilization, nutrient cycling and water purification, ferns are medicinally valuable source of impotent drugs for treatment of many human diseases (Rastogi *et al.*, 2018; Kumar *et al.*, 2023). The fern genus *Adiantum* belongs to subfamily Vittarioideae of Pteridaceae and comprises of approximately 252 species (Schoch *et al.*, 2020) widely distributed across tropical and temperate regions of the world (Pan *et al.*, 2011). The genus has been frequently growing in temperate zone of northern areas of Pakistan. Five species of *Adiantum* have been identified and characterized using classical taxonomy and electron microscopy (Attaullah *et al.*, 2019; Irfan *et al.*, 2021; Khan *et al.*, 2023).

Polymerase chain reaction (PCR), the molecular biology technique targeting amplification and gene sequencing of specific and conserved regions of nuclear, mitochondrial and plastid DNA has frequently been implied for genetic characterization and DNA barcoding of fern species. DNA barcoding offers a potential method to verify previously identified species, detect misidentifications or adulterants, and uncover novel or cryptic taxa (Pryer *et al.*, 2010; Nitta and Chambers, 2022). Standardized DNA sequences or gene markers typically of 400-800 bp long are used for amplification and gene sequencing. However, no single DNA marker has proven sufficient to delimit the fern species. Consequently, the combination of two or more plastid loci is recommended to resolve phylogenetic relationship and genetic divergence of the fern species (Zhao *et al.*, 2025). Several plastid loci *rbcL*, *psbA-trnH* and *trnL-F* along with nuclear *ITS2* and *IBR3_2* are widely used for fern barcoding (Wang *et al*., 2016; Nitta *et al.*, 2020; Kuo *et al.*, 2024). Among these, *rbcLa* (portion of *rbcL* gene) and the intergenic spacer, *psbA-trnH* have shown strong potential to discriminate Asian species of *Adiantum* (Ma *et al.*, 2010; Ebihara, 2011; Morajkar *et al.*, 2022).

NCBI's BLASTn (Basic Local Alignment Search Tool for nucleotides) tool remains the most widely used platform for aligning DNA sequences and searching match through reference sequence databases. However, its accuracy and authenticity rely heavily on the quality of both the query sequence and the reference database. Interpretation of results depends upon metrics such as percentage identity, maximum score, query coverage, and E-value thresholds. Misannotations in public databases can compromise results necessitating careful curation and cross-validation of top hits (Nilsson *et al.*, 2006). Hence, while BLASTn is a cornerstone of sequence similarity analysis, its reliability is contingent upon database post-analysis validation. For higher confidence in sequence identification, it is often complemented with tools such as MEGA-12 (Hall, 2013). Beyond species identification, DNA barcoding also contributes to understanding evolutionary relationships through molecular phylogenetics. Phylogenetic analyses reveal lineage divergence, monophyly, and cryptic speciation. Phylograms constructed using multiple sequence alignments (MSA) dataset aid in visualizing species placement and testing taxonomic boundaries (Ebihara, 2011; Lu *et al.*, 2012; Janakiraman *et al.*, 2022).

Morphology-based fern records are available in Pakistan, but molecular validation is lacking across taxa and localities. In this context, the present study used sequences of *rbcLa* and *psbA-trnH* markers for molecular identification and phylogenetic placement of *A. capillus-veneris* L. and *A. veneustum* D. Don from Swat Valley, Pakistan. By generating two locus barcode records and assessing sequence variability and divergence, this work contributes to floristic documentation and advances the molecular taxonomy of South Asian *Adiantum* species.

MATERIALS AND METHODS

Site of plant collection: The sampling sites, Marghazar (34°66'33"N and 72°34'47"E) and Malam Jabba Forest (35°03'12"N and 72°33'39"E), are in the district of Swat, Khyber Pakhtunkhwa (KP), Pakistan. Both regions are enriched with Sino-Japanese type of vegetation determined by subtropical and moist temperate conditions. The soil type is predominantly sand to sandy loam formed by river deposition and mountain erosion. It is characterized by moderate to high fertility, largely attributed to organic matter from natural vegetation (Khan *et al.*, 2022).

Sample collection: The sampling areas flourish with species of ferns during month of March-July after period of snow melting. Five spatially separated individuals of each species were collected at each site in July 2024 for voucher preparation and DNA barcoding. Fresh pinnules were removed from young fronds and immediately placed in airtight polythene zip-lock bags containing silica gel at an approximate tissue:silica ratio of 1:10 (w/w) for DNA extraction. The plant samples intended for morphological identification and DNA barcoding were collected from the field during July. Fresh pinnules were removed from young fronds and immediately placed in airtight polythene zip-lock bags containing silica gel at an approximate tissue : silica ration of 1:10 (w/w) for DNA extraction. Samples were kept in ice box during field work and transported to the laboratory within 24 hours. In parallel an additional frond from each plant was collected, placed between newspapers, pressed, and after drying, mounted on herbarium sheets for documentation and voucher reference. The identified specimens of *Adiantum capillus-veneris* L. and *Adiantum venustum* D. Don, were deposited in the Dr. Sultan Ahmed Herbarium, Department of Botany, Government College University, Lahore Pakistan, under voucher numbers GC. Herb.Bot-4377 and GC. Herb.Bot-4378, respectively.

Figure 1. Morphological characteristics of two species of *Adiantum*. *Adiantum capillus-veneris*: (a) adaxial side of a frond; (b) abaxial side of frond showing marginal sori covered by false indusia along the pinnule margins. *Adiantum venustum*: (c) Habitat; (d) Abaxial surface of a pinna with two false indusia along margin of each pinnule.

DNA Extraction: Because the preliminary attempts using a mixer mill and liquid nitrogen grinding resulted in DNA degradation, the dry pinnule (30 mg) were finely shredded into small particles (~1-2 mm) using sterilized scissors to facilitate cell lysis while minimizing DNA damage. During the lysis step, gentle grinding with a microtube pestle was performed while the samples were incubated in lysis buffer in a water bath. Total genomic DNA was extracted using the Qiagen DNeasy Plant Mini Kit (Catalogue No. 51304), following the recommended protocol, and DNA was eluted in 50 uL of elution buffer and stored at -20°C**.** DNA yield (ng/uL) and quality (A260/A280) were determined spectrophotometrically. The presence of chloroplast DNA in total extracted DNA was evaluated through agarose gel electrophoresis, following protocol described by Lee *et al.* (2012). The DNA samples (5 µL) were mixed with 1uL 6X loading dye and loaded onto a 1% agarose gel prepared in 1X TAE buffer, containing ethidium bromide (5ul/100ml of gel solution). A 1 kb standard DNA ladder (Invitrogen, Thermo Fisher Scientific) was used as a molecular size marker. The gel was run at 100V for 45min, and DNA bands were visualized under a UV transilluminator.

PCR amplification and sequencing: Forward and reverse primers for *rbcLa* and *psbA-trnH* (Table 1) were synthesized by Macrogen Inc., (Seol, Republic of Korea). PCR amplification of the chloroplast barcodes was performed using Thermo Fisher PCR Master Mix (K1081) in a final reaction volume of 35 uL containing 18 uL Master Mix, 10 uL nuclease-free water, 1.5 uL forward primer, 1.5 uL reverse primer and 4 uL of eluted DNA. Thermal cycling conditions followed Trujillo-Argueta *et al.* (2021) with some modification (Table 1). The PCR amplification products were visualized through gel electrophoresis on 1.5% agarose gel. Amplicons were purified using polyethylene glycol precipitation method, following protocol of Kusukawa *et al.* (1990). Ten specimens were collected per species (n=10; total specimen N=20). PCR was attempted for all collected specimens; however, two representative specimens per species/marker were selected for sequencing. Finally, purified PCR products of specimens were submitted for Sanger sequencing to Macrogen Inc., Seoul, Republic of Korea. Accordingly, *rbcLa* and *psbA-trnH* sequences passed quality control (visual inspection of chromatograms) and were used in downstream analysis.

Table 1. Primer sequences for *rbcLa* and *psbA-trnH* regions of chloroplast DNA and PCR temperature cycling programs for amplification.

Marker name	Primer name	Primer sequence	Temperature cycling program
rbcLa	rbcLa_F rbcLa-R	ATGTCACCACAAACAGAGACTAAAGC GTAAAATCAAGTCCACCACG	94℃ for 4min; 35 cycles of 94℃ for 30 sec, 55 ℃ for 30 sec, 72 ℃ for 1min; final extension 72℃ for 10 min
psbA-trnH	trnHf_05 psbA3_f	CGCGCATGGTGGATTCACAATCC GTTATGCATGAACGTAATGCTC	94 ℃ for 4 min; 35 cycles of 94℃ for 45 sec, 53℃ for 45 sec, 72 ℃ for 1min; final extension 72℃ for 10 min

Bioinformatic tools and analyses: The CHROMAS, version 2.6.6, MEGA (Molecular Evolutionary Genetics Analysis), version 12.1 and EXCEL software were used for bioinformatic analyses.

Sequence reads: The chromatograms and FASTA files of *rbcLa* and *psbA-trnH* sequences of samples were received from Macrogen (Korea). Chromatograms were visually inspected using CHROMAS software, version 2.6.6, and sequences were trimmed to remove noise. The clean sequences were copied, saved in FASTA format, and imported in MEGA v12.1. All corresponding *rbcLa* and *psbA-trnH* sequences of two species showed no difference and four sequences named SW01_rbcLa, SW01_psbA-trnH, and similarly SW29_rbcLa and SW29_psbA-trnH from sample SW01 and SW29 were selected for downstream analysis.

NCBI-BLASTn search: The two sequences (SW01_rbcLa, SW01_psbA-trnH) generated from the sample SW01 and two (SW29_rbcLa and SW29_psbA-trnH) from sample SW29 were queried against BLASTn database. The program megablast was opted to search closely related sequences with higher percentage identity ≥ 95%. Each query sequence search showed a list of 100 matches by default in the description panel. The names of top ten species (having homologous sequences) and their alignment statistics to each of the four query sequences are recorded in four separate tables (3-6). BLAST hits representing high percent identity, query cover, max score, and E-value (zero) were selected for downstream analyses. In order to select the matches for streamline analysis, we prioritized, Max score, higher query cover and percentage identity and low E-value (Pearson, 1999; Madden, 2013).

BOLDSYSTEMS searches: The *rbcLa* and *psbA-trnH* sequences of both specimens i.e., SW01 and SW29 were also queried in BOLDSYSTEMS database.

Phylogenetic tree reconstruction: For reconstruction of phylogenetic trees, we picked up the accession numbers of the most relevant and representative sequences from both: BLASTn hits and published articles (Bouma *et al.*, 2010; Lu *et al.*, 2012; Wang *et al.*, 2016; Huiet *et al.*, 2018). The sequences corresponding to the matched accession numbers were retrieved separately for *rbcLa* and *psbA-trnH* datasets from GenBank using web tool of MEGA, version 12.1 (Kumar *et al.*, 2024). Sequentially, the query sequences from two FASTA files, one containing SW01_rbcLa and SW29_rbcLa, and other containing SW01_psbA-trnH and SW29_psbA-trnH were imported in their respective sequence dataset for multiple sequence alignment (MSA). The MSA was performed using MUSCLE algorithm (Edgar, 2004; Thompson *et al.*, 2011) in

MEGA 12.1 and two multiple sequence alignment datasets were prepared for both markers: an *rbcLa* dataset including newly generated sequences of both *Adiantum* species and their closest BLASTn matches, and a *psbA-trnH* dataset comprising the same. Moreover, *rbcLa* and *psbA-trnH* datasets were concatenated to form a third dataset. Maximum Likelihood method was employed, due to its superior accuracy, in reconstructing evolutionary relationships (Huelsenbeck and Hillis, 1993).

Maximum Likelihood tree of *psbA-trnH* dataset was constructed in MEGA (version 12.1) software using Tamura (1992) nucleotide substitution model, Tamura 3-parameter (T92). A set of 26 sequences was analyzed, and the dataset was filtered using partial deletion option with a 95% site coverage threshold, resulting in an alignment of 401 positions. Analysis utilized up to three parallel computing threads for performance optimization. Evolutionary analyses were conducted in MEGA 12.1 (Kumar *et al.*, 2024).

The *rbcLa* dataset based Maximum likelihood (ML) phylogram was developed using Kimura 2-Parameter model (K2P) of nucleotide substitution and evolutionary rate was modeled using Gamma distribution across 5 categories (+G parameter = 0.2258). The analysis covered 28 sequences, and the partial deletion option was used to ignore less than 95% site coverage. The final dataset contained 456 positions. Both trees were supported by bootstrap value calculated from 1000 replicates. The concatenated tree (Maximum likelihood) was constructed by using T92 model (Tamura, 1992) of nucleotide base substitutions. The selected substitution model included a proportion of invariable sites (I= 0.5352) indicating that 53.52% of alignment positions were evolutionarily conserved across 857 positions of nucleotides in a dataset of 26 concatenated sequences. The bootstrap test (1000 replicates) was used to assess branch support.

Statistical analysis: All alignment datasets were analyzed in MEGA 12.1. The summary sequence statistics including aligned length, GC content (%), number of conserved, variable, parsimony informative and singleton sites were calculated for each dataset. Kimura 2-parameter model was used to estimate pairwise genetic distances; between species and within species distances were summarized as mean ± SD (Standard deviation). Barcode performance was evaluated by calculating (i) the strict barcode gap (minimum interspecific distance minus maximum intraspecific distance), (ii) the mean-based barcode gap, and (iii) by visualizing the distributions of K2P distances using histograms. Phylogenetic analyses were performed using Maximum Likelihood with 1000 bootstrap iterations and partial deletion (95% site coverage). The *psbA-trnH* tree used Tamura 3-parameter (T92), the *rbcLa* phylogram used Tamura 2-parameter model of nucleotide substitution and Gamma (G) with 1000 bootstrap iterations; branch lengths were estimated under same models. While concatenated analyses used T92 + I and resulting topology was compared with single-locus trees for concordance.

RESULTS

Morphological contrast: Morphological investigation of *Adiantum capillus-veneris* and *A. venustum* reveals distinct differences in both vegetative and reproductive traits. Although multiple characteristics separate the two species, several key features can be used for reliable identification. These include differences in frond size, stipe colour, lamina division, pinnae and pinnule shape, as well as number and morphology of false-indusia (**Error! Reference source not found.**;

Table ***2***).

Table 2. Differences in morphological characteristics of *A. capillus-veneris* and *A. venustum* collected from Swat Valley.

Species	Frond	Stipe	Lamina	Pinna	Pinnule	Pseudoindusia
Adiantum capillus-veneris L.	Small, 15-25 cm	Shiny black, slender, and wiry	8-18cm, bi to unipennate acropetally	Fan-shaped, thin, papery texture.6-10cm	Irregular lobes with dentate margin	4-6 at abaxial side of pinnules
Adiantum venustum D. Don	Large 30-40cm	Shiny reddish-brown, thick, not wiry	3-Pinnate mostly	Oblong, lanceolate, thick,4-8cm long	Wedge-shape, slightly serrated	1-2 at abaxial side of pinnules

PCR amplification and sequencing: The *rbcLa* and *psbA-trnH* regions of chloroplast DNA were successfully amplified using *rbcLa* and *psbA-trnH* primers. The PCR amplicons were sequenced using Sanger sequencing by Macrogen, Korea.

We received Chromatograms and corresponding FASTA files for all four samples. After careful inspection and noise removal in CHROMAS software, the two high-quality sequences from specimen SW01 (SW01_rbcLa and SW01_psbA-trnH) two from specimen SW29 (SW29_rbcLa and SW29_psbA-trnH) of specimen SW29 were subjected to downstream analyses.

BLASTN Searches: The 512bp long sequence SW01_rbcLa was queried through BLASTn tool of NCBI. All top ten hits showed 100% identity (PI) and full query coverage (QC) with a consistent maximum score of 946 indicating exceptionally strong matches (BOLDSYETEMS produced same overall outcome as BLASTn, with an ambiguous assignment for specimen SW01 (*rbcLa*) and correctly identified remaining queries where reference sequences were available.

Table *3*). Notably *Adiantum capillus-veneris* dominated the results with eight accessions while *Adiantum venustum* also appeared in two entries (MF694942, MF694655). The uniformity across these high scoring matches underscores *Adiantum capillus-veneris* as the most probable identity for SW01_rbcLa and a robust candidate for phylogenetic reconstruction. The query sequence was submitted to GenBank, and the accession number PV983596.1 was assigned.

BLASTn analysis of the SW29_*rbcLa* sequence returned the top ten hits including *Adiantum venustum* (six accessions) and *A. fengianum*, *A. bonatianum*, *A. fimbriatum* and *A. roborowskii*, ranked by E-value, query coverage and percent identity (Table 4). The query sequence showed high similarity to *A. venustum*, with multiple top BLASTn hits displaying ≥99.82% identity and full-length query coverage (99–100%). All alignments yielded E-values = 0, indicating highly significant matches. The accessions EF452136.1 and PP801083.1 of *A. venustum* showed a strong match to our query sequence with 99.82% identity and 100% query coverage. This pattern, supported by multi-accession validation confirmed the molecular identity of the query as *A. venustum*. The top BLASTn hit, accession number PV983595.1 for *Adiantum venustum* (Table 4) corresponds to the sequence assigned by GenBank to our query sequence.

The BLASTn search results of SW01_psbA-trnH sequence are summarized in (Table 5). The query sequence matched robustly with multiple reference sequences of *A. capillus-veneris* L. and with the single available sequence of *A. ogasawarense* Tagawa. Thus, the BLASTn output provided an ambiguous identification for sample SW01, indicating that the query sequence is either from *A. capillus-veneris* or *A. ogasawarense*, likely because the *psbA-trnH* spacer is highly conserved between these two taxa. The GenBank accession PX021940.1 was assigned to our query sequence SW01_*psbA-trnH* upon its submission (Table 5). Finally, the BLASTn query of sequence SW29_psbA-trnH demonstrated the highest similarity (PI>99% and QC >90%) with two available accessions (GQ248238.1 and EF590668.1) of *A. venustum* (Table 6). Other matching sequences include *A. monochlamys A. davidii*, *A. bonatianum* and *A. fimbriatum* with percent identity 98.6 –100% but lower query coverage (82-83%). Overall, the presence of *A. venustum* among the top hits, together with high identity and substantial coverage, supported the molecular identification of SW29 as *A. venustum* and indicated close affinity with other *Adiantum* species. The accession PX021939.1 corresponds to the SW29_psbA-trnH sequence generated in this study. BOLDSYETEMS produced same overall outcome as BLASTn, with an ambiguous assignment for specimen SW01 (*rbcLa*) and correctly identified remaining queries where reference sequences were available.

***Table* 3. The BLASTn statistical metrics and scientific names of top ten species whose subject sequences are highly similar with 512bp long query sequence, SW01_rbcLa**

Scientific Name	Max Score	Query Cover	E-value	Per. Ident	Acc. Len	Accession
***Adiantum** capillus-veneris*	946	100%	0	100	594	MN207128.1
***Adiantum** capillus-veneris*	946	100%	0	100	150568	NC004766.1
***Adiantum** venustum*	946	100%	0	100	552	MF694655.1
***Adiantum** capillus-veneris*	946	100%	0	100	1273	JF935320.1
***Adiantum** capillus-veneris*	946	100%	0	100	620	JN114780.1
***Adiantum** capillus-veneris*	946	100%	0	100	1268	JF935318.1
***Adiantum** capillus-veneris*	946	100%	0	100	525	PV410590.1
***Adiantum** venustum*	946	100%	0	100	553	MF694942.1
***Adiantum** capillus-veneris*	946	100%	0	100	512	PV983596.1
***Adiantum** capillus-veneris*	946	100%	0	100	1380	LC004119.1

Table 4. The BLASTn statistical metrics and scientific names of top ten species whose subject sequences are highly similar with 541bp long query sequence, SW29_rbcLa

Scientific Name	Max Score	Query Cover	E value	Per. ident	Acc. Len	Accession

Adiantum venustum	1000	100%	0	100	541	PV983595.1
Adiantum venustum	994	100%	0	99.82	1308	EF452136.1
Adiantum venustum	994	100%	0	99.82	1287	PP801083.1
Adiantum fengianum	989	100%	0	99.63	1309	MH019568.1
Adiantum bonatianum	989	100%	0	99.63	1309	MH019542.1
Adiantum venustum	985	99%	0	100	553	GQ248541.1
Adiantum venustum	985	99%	0	100	552	MF694656.1
Adiantum venustum	983	100%	0	99.45	1309	MH019645.1
Adiantum fimbriatum	983	100%	0	99.45	1309	MH019569.1
Adiantum roborowskii	977	100%	0	99.26	1309	MH019627.1

Table 5.The BLASTn statistical metrics and scientific names of top ten species whose subject sequences are highly similar with 490bp long query sequence, SW01_psbA-trnH.

Scientific Name	Max Score	Query Cover	E value	Per. ident	Acc. Len	Accession
Adiantum capillus-veneris	905	100%	0	100	150568	NC004766.1
Adiantum capillus-veneris	905	100%	0	100	490	PX021940.1
Adiantum ogasawarense	904	100%	0	100	533	AB575454.1
Adiantum capillus-veneris	904	100%	0	100	535	MG947187.1
Adiantum capillus-veneris	904	100%	0	100	550	AB575450.1
Adiantum sp. US:3515464	894	99%	0	100	511	MF785864.1
Adiantum capillus-veneris	878	97%	0	100	524	PP816217.1
Adiantum capillus-veneris	773	85%	0	100	418	KT427111.1
Adiantum flabellulatum	765	100%	0	94.95	152063	NC_064144.1
Adiantum flabellulatum	684	92%	0	94.09	498	AB575452.1

Table 6.The BLASTn statistical metrics and scientific names of top ten species whose subject sequences are highly similar with 517bp long query sequence, SW29_psbA-trnH.

Scientific Name	Max Score	Query Cover	E value	Per. ident	Acc. Len	Accession
Adiantum venustum	955	100%	0	100	517	PX021939.1
Adiantum venustum	905	95%	0	99.8	524	GQ248238.1
Adiantum venustum	872	91%	0	100	472	EF590668.1
Adiantum monochlamys	826	98%	0	96.07	559	AB575453.1
Adiantum davidii	797	83%	0	100	431	KT427133.1
Adiantum davidii	791	83%	0	99.77	431	KT427141.1
Adiantum fimbriatum	774	83%	0	99.3	428	KT427153.1
Adiantum bonatianum	767	83%	0	98.62	436	KT427102.1
Adiantum davidii	758	83%	0	98.61	426	KT427140.1
Adiantum bonatianum	758	82%	0	98.61	431	KT427103.1

Genetic variability: Comparative sequence characteristics of the two chloroplast markers are summarized in Table 7. For *rbcLa*, two datasets were examined: the full set of 24 *rbcLa* sequences (a), which included two discordant sequences (MF694655.1 and MF694656.1), and a dataset of 22 sequences (b) after removing these discordant entries. In both cases, the final alignment length was 456 bp, comprising 396 conserved sites and 60 variable sites, of which 59 were parsimony-informative and one was a singleton. In contrast, the *psbA-trnH* region showed higher variability, with an alignment length of 487 bp containing 360 conserved sites and 93 variable sites, all of which were parsimony-informative and with no singletons. The GC content was also lower in *psbA-trnH* (40.40%) compared with *rbcLa* (46.90%), reflecting the more variable and AT-rich nature of this intergenic spacer.

Table 7. Comparative sequence characteristics of the two chloroplast markers *rbcLa* and *psbA-trnH*

Marker	No. of sequences	Aligned length (bp)	Conserved sites	Variable sites	Parsimony informative sites	Singleton sites	GC (%)
(a) rbcLa	24	456	396	60	59	1	46.90
(b) *rbcLa*	22	456	396	60	59	1	46.90
psbA-trnH	22	487	360	93	93	None	40.40

Dataset for rbcLa (a) includes 24 sequences with two discordant entries, while (b) rbcLa represents the 22-sequence set without discordant haplotypes. The psbA-trnH dataset analyzed includes 22 sequences.

Genetic distance and barcode gap analyses: The strict barcode gap —defined as the minimum interspecific genetic distance minus the maximum intraspecific distance —and the mean-based barcode gap were calculated for *rbcLa* and *psbA-trnH* markers. The multiple sequence dataset of 24 sequences of rbcLa produced negative barcode gap (-0.0656) because of two discordant sequences (accessions; MF694655.1 and MF694656) present in dataset. However, after removing these two discordant sequences, the remaining 22 sequences in rbcLa dataset showed complete separation between intraspecific and interspecific distances, yielding a positive strict barcode gap of 0.0573. The *psbA-trnH* region exhibited the widest barcode gap of the two markers, with a mean interspecific divergence of 0.1166 and a strict barcode gap of 0.0024, reflecting its higher variability and stronger species-level discriminatory power.

The histogram based on K2P genetic distances for the 24-sequence *rbcLa* dataset (**Error! Reference source not found.**) showed overlap between intraspecific (orange) and interspecific (blue) distances. After excluding two misannotated sequences, this overlap disappeared and all intraspecific distances were 0 (**Error! Reference source not found.**). Likewise, the 22-sequence psbA–trnH dataset produced a histogram in which intraspecific distances were 0 and interspecific distances were non-zero (**Error! Reference source not found.**). In all histograms, a small fraction of interspecific pairs occurred near zero, indicating limited overlap and a narrow barcode gap in a subset of comparisons. Overall, *psbA–trnH* showed a wider spread of interspecific K2P distances than rbcLa, suggesting stronger species-level discriminatory power.

Table 8. Barcode gap analysis of *rbcLa* and *psbA-trnH* sequence datasets.

Chloroplast DNA Marker	Mean Interspecific Distance (X)	Mean Intraspecific Distance (Y)	Mean-based Barcode Gap (X-Y)	Minimum Interspecific Distance (A)	Maximum Intraspecific Distance (B)	Strict Barcode Gap (A-B)
(a) *rbcLa*	0.0561±0.03	0.0177±0.03	0.0384	0.0022	0.0678	-0.0656
(b) *rbcLa*	0.0573±0.03	0.0000±0.00	0.0573	0.0022	0.0000	0.0022
psbA-trnH	0.1164±0.07	0.0000±0.00	0.1164	0.0024	0.000	0.0024

(a) *rbcLa* based 24-sequence dataset including two discordant sequences (MF694655.1 and MF694656.1) shows negative strict barcode gap. (b) *rbcLa* dataset without two discordant sequences shows positive strict barcode gap. The *psbA-trnH* exhibits wider barcode gaps between two markers.

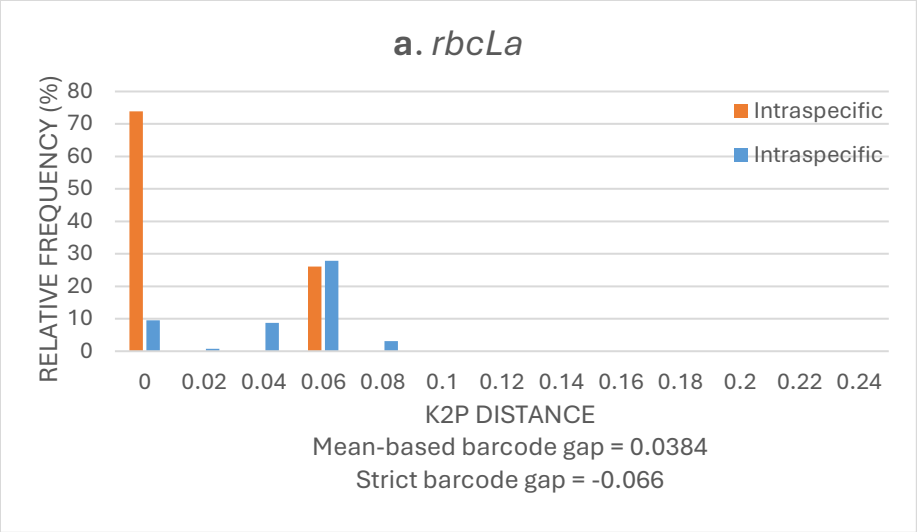

Figure 2. Distribution of intra and inter-specific K2P genetic distances for *rbcLa*.

A total of 276 pairwise comparisons were generated (23 intraspecific and 253 interspecific). The part of intraspecific pairs (orange bar) at 0.06 genetic distance represents two discordant accessions (MF694655.1 and MF694656.1).

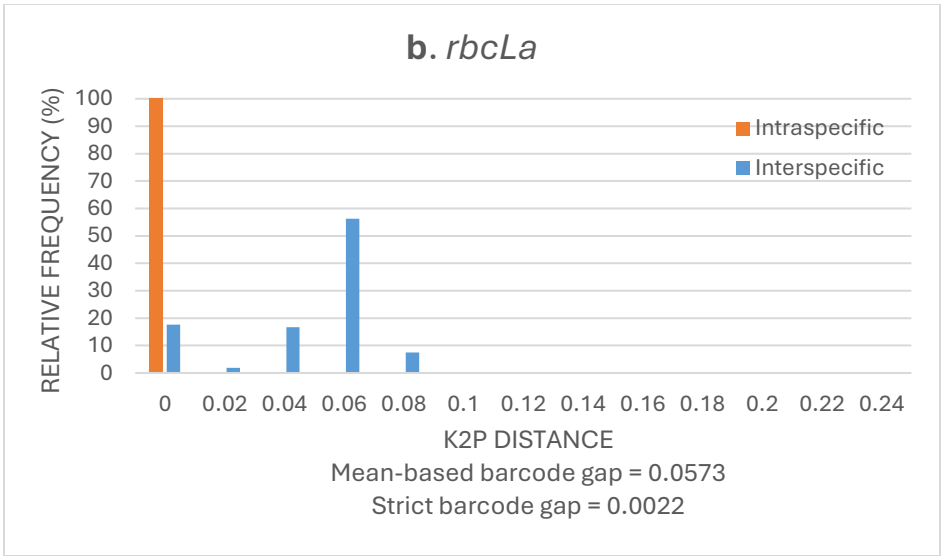

Figure 3. Distribution of intra and inter-specific K2P genetic distances for *rbcLa*

After removal of two discordant sequences (MF694655.1 and MF694656.1) the remaining 22 sequences generated a total of 231 pairwise comparisons (16 intraspecific and 215 interspecific). The part of intraspecific pairs (orange bar) at 0.06 genetic distance disappeared.

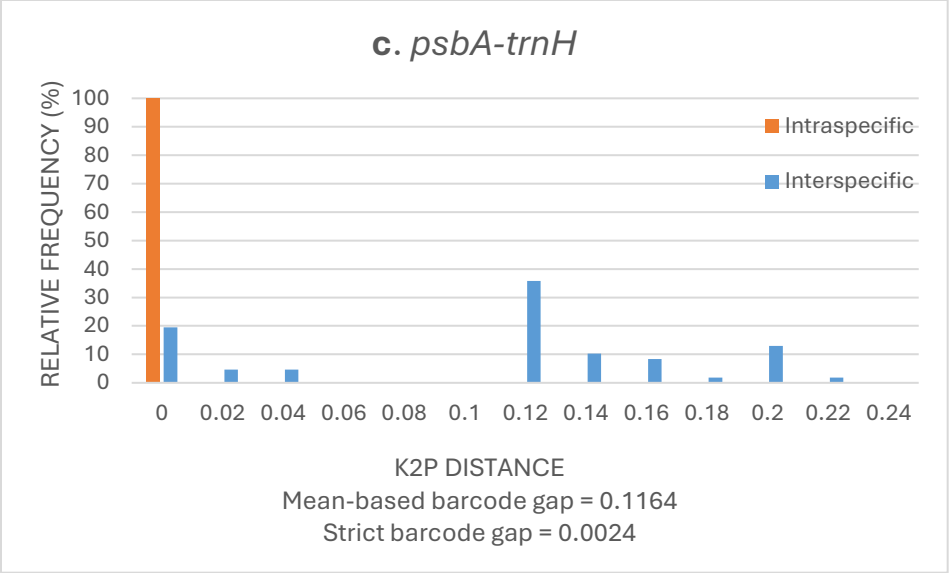

Figure 4. Distribution of intra and inter-specific K2P genetic distances for *psbA-trnH*

A total of 22 sequences generated 231 pairwise comparisons (intraspecific 16 and interspecific 215).

Phylogenetic placements: The phylograms based on Maximum Likelihood analyses of *rbcLa* and *psbA-trnH* resolved the sequences of each dataset into three major clades: the hispidulum, cap-ven and venustum clades with outgroup taxa. The SW01_rbcLa sequence (SW01) was placed in cap-ven clade alongside multiple sequences of *Adiantum capillus-veneris* and one sequence of *Adiantum ogasawarense* (Figure 5); two sequences of *A. venustum* also occurred in this clade, forming polytomy. The SW29_rbcLa sequence (SW29) grouped within the venustum clade and joined with sequence of *Adiantum venustum*, clearly separating from other species present in the clade (Figure 6). The *psbA-trnH* phylogram showed similar pattern i.e., the sequences SW01_psbA-trnH (SW01) and SW29_psbA-trnH (SW29) appeared in cap-ven and venustum clades, respectively. Sequences of *Adiantum* occurred in both cap-ven and venustum clades thus representing polyphyletic nature of *Adiantum venustum*.

The concatenated ML tree (Figure 7) based on *rbcLa* + *psbA-trnH* dataset well resolved the evolutionary relationship among species of genus *Adiantum* with strong bootstrap support across major clades: cape-ven, venustum and hispidulum, stabilized by outgroup taxa. The venustum clade formed a strongly supported monophyletic group comprising *A. venustum*, *A. davidii*, *A. fengianum* and *A. fimbriatum*. The cap-ven clade included *A. induratum*, *A. flabbellulatum* , A. *capillus-veneris*, and *A. ogasawarense*, all clustering with high confidence. The hispidulum clade formed a separate lineage showing evolutionary relationship between *A. hispidulum* and *A. diaphanum*. The query SW01 consistently grouped with *A. capillus-veneris showing its identity as A. capillus-veneris while* query SW29 firmly clustered A. *venustum* lineage thus representing its identification as *A. venustum*. The sister clades i.e., cap-ven clade and venustum clade changed their position in case of single marker dataset while concatenated ML tree recovered a well resolved topology.

DISCUSSION

The results of our study validate molecularly the nativeness of *A. capillus-veneris* and *A. venustum* in Swat Valley and are consistent with previous taxonomic studies (Irfan *et al*., 2022; Khan *et al*., 2023). The newly generated sequences of *rbcLa* and *psbA-trnH* regions is a contribution to incomplete data of the "The Flora of Pakistan." The *rbcLa* marker —characterized by its high amplification rate and low variability —is often coupled with more variable *psbA-trnH* intergenic spacer to improve the molecular diagnosis of fern species from underexplored regions (Trujillo-Argueta *et al.*, 2021). This variability and complementary use of markers is evident in our study for identification of *Adiantum* species of Swat Valley. Furthermore, the couplet of markers is an excellent choice for determining evolutionary relationships among ferns species using maximum likelihood (ML) tree-based methods (Wang *et al.*, 2016).

Our findings are generally consistent with previous barcoding studies. The *rbcLa* is recommended for barcoding *Adiantum* species as it is a reliable marker for amplification and broad phylogenetic placement, while *psbA-trnH* contributes stronger species level discrimination in closely related taxa; this matches our observation that variable spacer loci improved resolution where coding plastid loci were conservative (Wang *et al.*, 2016). Recent fern barcoding efforts in

underexplored regions similarly showed that ML-based plastid barcode analyses can recover coherent species clusters and are effective for regional inventory building, in agreement with our tree-based identifications (Trujillo-Argueta *et al.*, 2021).

At a broader scale, the new Asian fern barcode reference dataset (voucher-verified, 956 species) compiled by Kuo *et al.* (2024) confirms that incomplete reference coverage has been a major bottleneck in fern identification and that expanding curated libraries improves assignment reliability. This is consistent with the ambiguous matches observed in our study for underrepresented taxon, such as *A. ogasawarense*, for which only a single database entry was available for each marker (*rbcL* and *psbA-trnH*).

DNA Barcoding, which reconciles molecular and bioinformatic methods, is a robust and heuristic technique for species identification and is globally accepted by researchers (Zheng *et al.*, 2015; Kuo *et al.*, 2024; Nithaniyal *et al.*, 2025). Rather than contradicting conventional taxonomy, it complements morphological diagnosis by providing an independent line of evidence (Hajibabaei *et al.*, 2007). In a standard workflow, the specimens to be identified are examined morphologically, their target barcode regions are amplified and sequenced, and the generated sequences are analyzed bioinformatically. The most common methods of bioinformatic analyses are: BLASTn similarity searches, genetic distance and barcode gap evaluation, and phylogenetic reconstruction based on monophyly. In the present study the results of all three bioinformatic approaches consistently validated the taxonomic identity of specimens SW01 as *Adiantum capillus-veneris* and SW29 as *A. venustum*. However, the species identification was hampered by ambiguous BLASTn outcomes, very small and overlapping barcode gap, and polyphyly of taxa.

For specimen SW01, BLASTn results of *rbcLa* sequence (512bp) displayed its strong homology with multiple sequences of *Adiantum capillus-veneris*, authenticating its identification (

Table ***1***). However, the ambiguity in result for this sample was the presence of two sequences of *Adiantum venustum* (MF694655 and MF694642) occurring in a thorough list of BLASTn results, dominated with *A. capillus-veneris*. The barcode gap analysis and phylogenetic reconstruction resolved this ambiguity. The fall of intraspecific pairs at 0.06 value of K2P distance (orange bar), large value of negative barcode gap ((a) rbcLa based 24-sequence dataset including two discordant sequences (MF694655.1 and MF694656.1) shows negative strict barcode gap. (b) rbcLa dataset without two discordant sequences shows positive strict barcode gap. The psbA-trnH exhibits wider barcode gaps between two markers.

) and disappearance of both signals after excluding these two accessions from *rbcLa* sequence dataset (

) suggested probable misannotations in database. Consistent with earlier studies, our phylogenetic trees showed that genus *Adiantum* is monophyletic and *Adiantum venustum* and *Adiantum capillus-veneris* form sister clades (Wang *et al.*, 2016; Huiet *et al.*, 2018;). Notably, two disputed "*Adiantum venustum*" accessions clustered with species of cap-ven clade rather than within their venustum clade (Figure 5), demonstrated that these two sequences likely belong to *Adiantum capillus-veneris* not to *Adiantum venustum.* This pattern supports their likely reassignment pending voucher-based re-evaluation.

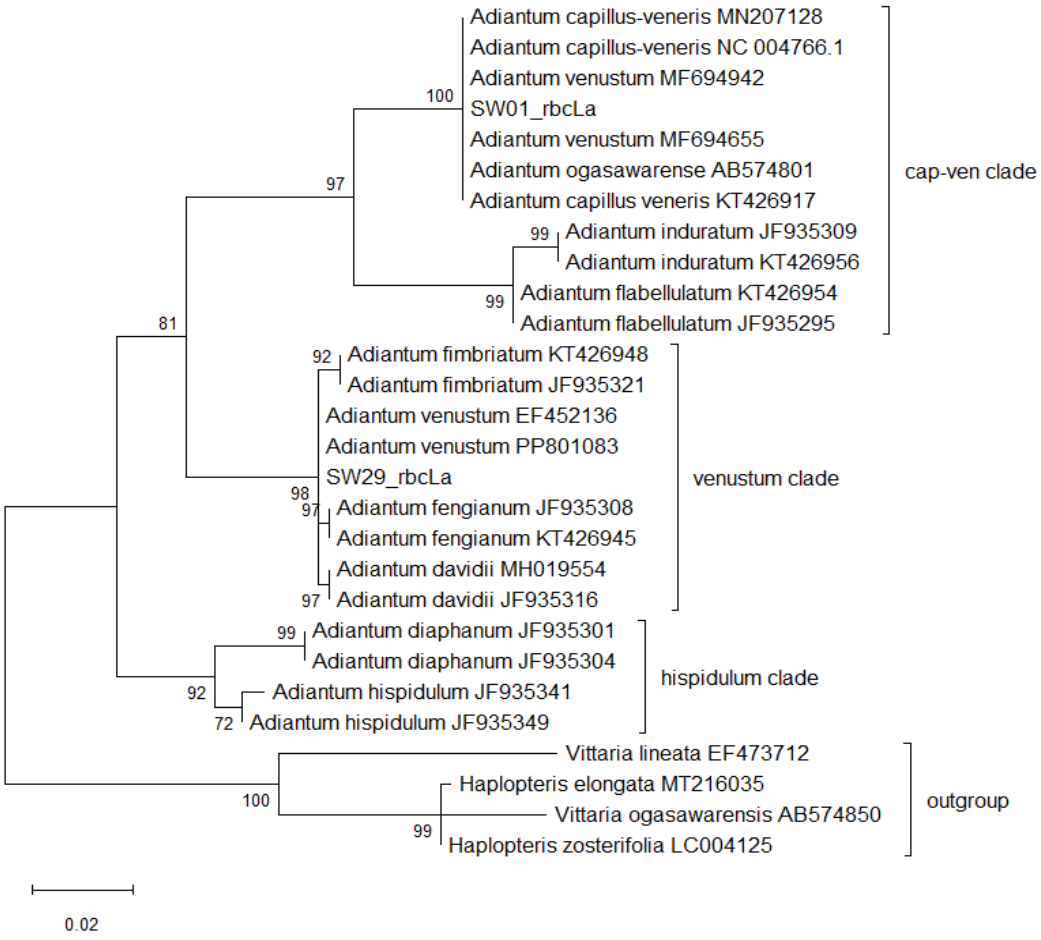

Figure 5. The *rbcLa* sequence dataset based phylogram of 24 sequences of genu *Adiantum* and 4 sequences of outgroup taxa. Our query sequences SW01 and SW29 were placed with their conspecific sequences from *Adiantum capillus-veneris* and *Adiantum venustum*, respectively.

The *psbA-trnH* sequence (490 bp) of the same specimen produced another ambiguous result, matching both *Adiantum capillus-veneris* and *A.* ogasawarense (

Table ***4***; **Error! Reference source not found.**; Figure 6). To investigate this, we retrieved the full 533 bp region of *A. ogasawarense* (accession AB575454) and aligned it with multiple sequences of *Adiantum capillus-veneris*; all were identical over the comparable region, indicating that *psbA-trnH* segment is conserved between *Adiantum capillus-veneris* and *Adiantum ogasawarense* and therefore has limited discriminatory power in this case. Cross-checking with *rbcL* further showed that only single available *A. ogasawarense* reference (AB574801; 1205bp) is identical to multiple published *Adiantum capillus-veneris* sequences (e.g., LC004119, DQ432659, JF935345, NC004766), strengthening the interpretation that current public records do not clearly separate these taxa at this locus. Given concordant molecular evidence and morphology, SW01 is best treated here as *A. capillus-veneris*, while targeted chloroplast + nuclear multilocus revision is recommended for definitive delimitation of *A. capillus-veneris* and *A. ogasawarense*.

Publicly available reference sequence databases are indispensable but are not error free and can lead to misinterpretations and ambiguity in results (Collins and Cruickshank, 2013). Misidentifications and misannotations in databases may arise due to different causes. First in ferns the plant body is simple, diagnostic characters are limited and plastic leading to confusion in delimiting species boundaries (Sundue and Rothfels, 2014). Consequently, if a specimen is initially misidentified, sequenced, and analyzed barcode is submitted (especially in groups lacking reliable conspecific reference sequences) the error can enter the database and propagate. Secondly specimens' mix-ups during collection, handling or

laboratory processing can result in incorrect sequences being associated with the wrong taxa. Once such sequences are uploaded to databases, they may remain undetected and can mislead taxonomic, phylogenetic and biodiversity studies (Duntsch *et al.*, 2022). When such errors enter public DNA database such as GenBank or BOLD, the consequences extend far beyond a single incorrect sequence. These errors can distort biodiversity research by inflating or deflating species estimates, misleading phylogenetic and evolutionary inference, compromised DNA-based species identification, and errors across studies.

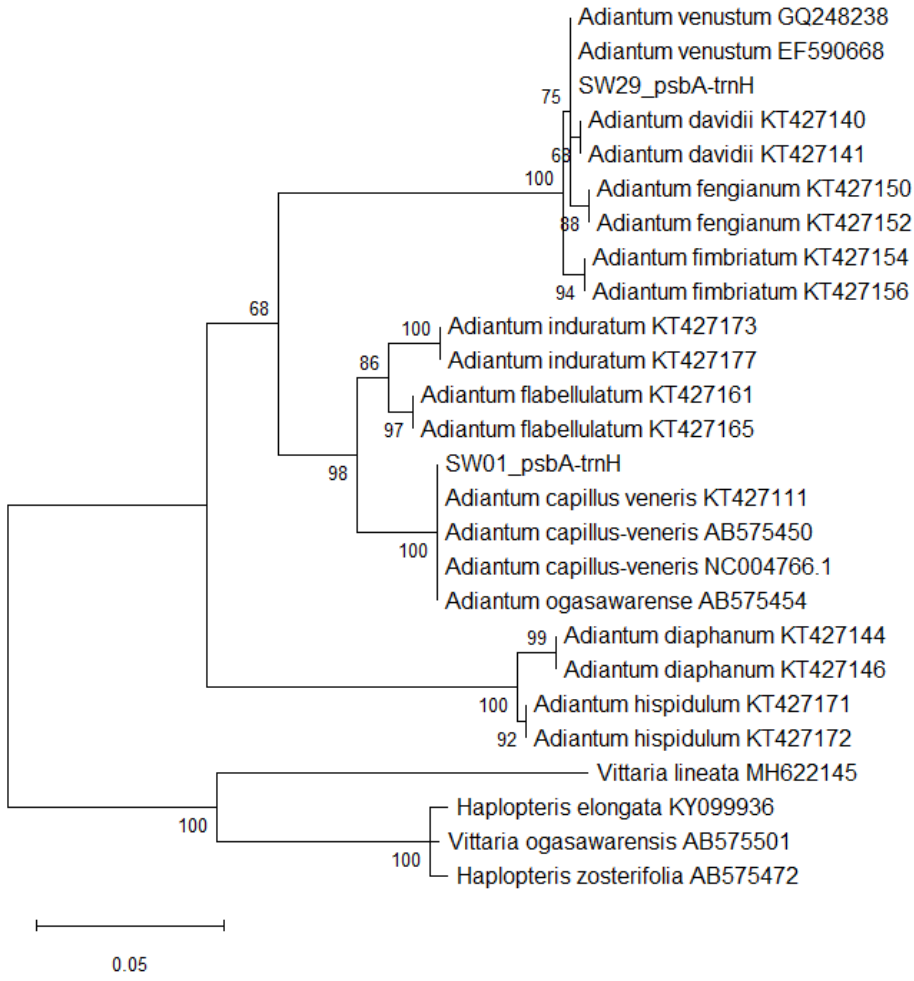

Figure 6. The phylogram constructed using 24 *psbA-trnH* sequences of *Adiantum* species and 4 sequences of outgroup taxa. The query samples SW01 and SW29 clustered with their respective conspecifics, *Adiantum capillus-veneris* and *Adiantum venustum*, confirming their species identity.

Altogether, the occasional BLAST hits, localized zero-distance interspecific pairs, negative or vary small barcode-gap values, and topological instability and polyphyly should not be interpreted solely as poor marker performance. In ferns, such patterns may arise from (i) misidentified public reference sequences (Schneider and Schuettpelz, 2006; Burgess *et al.*, 2011; Abdullah, 2017); (ii) incomplete taxon coverage in databases (Trujillo-Argueta *et al.*, 2021) (iii) hybridization and polyploidy, and (iv) Chloroplast capturing. After accounting for these factors, our integrated analyses consistently support SW01 as *A. capillus-veneris* and SW29 as *A. venustum*. Further work with broader geographic sampling, multiple individuals per species, and additional chloroplast and nuclear nuclei will further reinforce species boundaries in this complex.

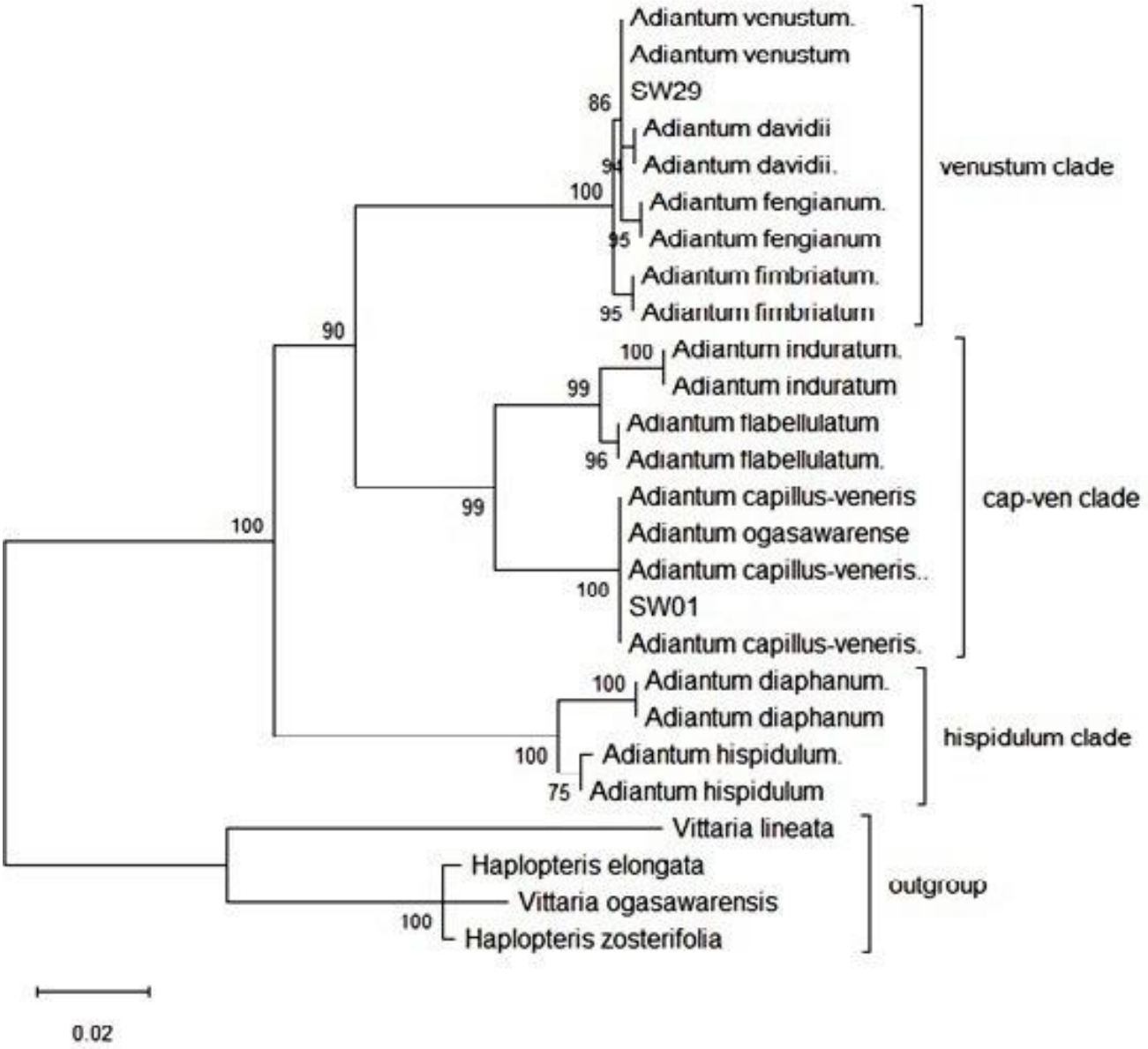

Figure 7. Maximum-likelihood phylogram inferred from a concatenated chloroplast markers datasets (rbcLa + psbA-trnH) comprising 26 terminal taxa. Node numeric indicate bootstrap support values (%) based on 1000 replicates. The scale bar (0.02) represents nucleotide substitutions per site.

Conclusion: This study provides four validated contributions: (i) new voucher-linked chloroplast barcode records (*rbcLa* and *psbA-trnH*) for *Adiantum* populations from Swat Valley; (ii) congruent species identification of sampled taxa as *A. capillus-veneris* and *A. venustum* using integrated approach (morphology, BLASTn, K2P distances, and ML phylogeny); (iii) evidence that selected GenBank entries (MF694942.1 and MF694655.1) are likely misannotated (iv) although *rbcL* and *psbA-trnH* are best markers for identification of ferns species, but their divergence is limited between *A. capillus-veneris* and *A. ogasawarense,* while, the latter remains poorly represented in public resources (single reference sequence and limited associated data), highlighting the need for continued database curation and expansion of reference libraries. This study also has clear limitations: restricted sampling breadth (two species and limited localities), sample size/species

and use of only two plastid loci (without nuclear markers). Therefore, conclusions are strongest for species authentication of sampled populations, while references on fine-scale population structure, endemism, and deeper phylogeography require broader and multilocus follow-up.

Acknowledgements: The authors gratefully acknowledge the support and facilities provided by the Government of the Punjab, particularly the Higher Education Department, for enabling the successful completion of this research. We thank Prof. Dr. Tehreema, Head of the Department of Botany, GCU Lahore. We are specially indebted to Dr. Sohaib and Muhammad Wasim for their moral support and motivation. We also acknowledge the use of GenBank and BOLD Systems for access to reference sequence data. Special thanks are extended to the field team for assistance with sample collection in the Swat Valley region.

Data availability statement: The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request. The four newly generated sequences have been deposited in GenBank database and accessions (PV983595.1, PX021939.1; PV983596.1, PX021940.1) have been assigned. Additional data, including sequence chromatograms and FAST files, alignment files and phylogenetic trees, are available upon request.

Conflict of Interest: The authors declare that there is no conflict of interest regarding the publication of this manuscript.

REFERENCES

Abdullah, M. T. (2017). Conserving the biodiversity of Kuwait through DNA barcoding the flora. PhD Thesis, University of Edinburg, UK. 1-303 pp. http://hdl.handle.net/1842/28786

Attaullah, Z. Khan and A. Midrarullah (2019). Taxonomic studies of pteridophytes of Utror Valley, District Swat, Khyber Pakhtunkhwa, Pakistan. Pak. J. Bot. 52(6): 2059-2065. http://dx.doi.org/10.30848/PJB2020-6(1)

Bouma, W. L., P. Ritchie and L. R. Perrie (2010). Phylogeny and generic taxonomy of the New Zealand Pteridaceae ferns from chloroplast *rbcL* DNA sequences. Aust. Syst. Bot. 23(3): 143-151. https://doi.org/10.1071/SB09047

Burgess, K. S., A. J. Fazekas, P. R. Kesanakurti, S. W. Graham, B. C. Husband, S. G. Newmaster, D. M. Percy, M. Hajibabaei and S. C. Barrett (2011). Discriminating plant species in a local temperate flora using the *rbcL+ matK* DNA barcode. Methods Ecol. Evol. 2(4): 333-340. https://doi.org/10.1111/j.2041-210X.2011.00092.

Chettri, N., B. Adhikari, S. Chaudhary and K. Wangchuk (2023). Changing discourses in the third pole: A systematic review of climate change impact on biodiversity in the Hindu Kush Himalaya. Ecol. Indic. 155: 111046. https://doi.org/10.1016/j.ecolind.2023.111046

Collins, R. and R. Cruickshank (2013). The seven deadly sins of DNA barcoding. Mol. Ecol. Resour. 13(6): 969-975. https://doi.org/10.1111/1755-0998.12046

Duntsch, L., P. Brekke, J. G. Ewen and A. W. Santure (2022). Who are you? A framework to identify and report genetic sample mix-ups. Mol. Ecol. Resour. 22(5): 1855-1867. https://doi.org/10.1111/1755-0998.13575

Ebihara, A. (2011). *rbcL* phylogeny of Japanese pteridophyte flora and implications on infrafamilial systematics. Bull. Natl. Mus. Nat. Sci. 37: 63-74. https://api.semanticscholar.org/CorpusID:55552293

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5): 1792-1797. https://doi.org/10.1093/nar/gkh340

Hajibabaei, M., G. A. Singer, P. D. Hebert and D. A. Hickey (2007). DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics? Trends Genet. 23(4): 167-172. https://doi.org/10.1016/j.tig.2007.02.001

Hall, B. G. (2013). Building phylogenetic trees from molecular data with MEGA. Mol. Biol. Evol. 30(5): 1229-1235. https://doi.org/10.1093/molbev/mst012

Huelsenbeck, J. P. and D. M. Hillis (1993). Success of phylogenetic methods in the four-taxon case. Syst. Biol. 42(3): 247-264. https://doi.org/10.1093/sysbio/42.3.247

Huiet, L., F. Li, T. Kao, J. Prado, A. R. Smith, E. Schuettpelz and K. M. Pryeri (2018). A worldwide phylogeny of *Adiantum* (Pteridaceae) reveals remarkable convergent evolution in leaf blade architecture. Taxon. 67(3): 488-502. https://doi.org/10.12705/673.3

Irfan, M., G. Jan, F. G. Jan and W. Murad (2021). Taxonomy and spore morphology of genus *Adiantum* (Vittarioideae; Pteridaceae) from Pakistan. Microsc. Res. Tech. 84(11): 2727-2736. https://doi.org/10.1002/jemt.23835

Irfan, M., G. Jan, F. G. Jan and W. Murad (2022). Floristic diversity and chorotype analysis of the pteridophytes of Pakistan. J. Anim. Plant Sci. 32(1): 1-10. https://doi.org/10.36899/JAPS.2022.1.0395

Janakiraman, N., M. Narayani and M. Johnson (2022). A review on molecular phylogeny of Pteridophytes using DNA barcoding. In: Ferns propagation, medicinal uses and environmental regulation. (eds. J. Marimuthu, H. Fernandez, A. Kumar and S. Thangaiah), Springer; Singapore. 39-58 pp. https://doi.org/10.1007/978-981-16-6170-9_3

Khan, M., Nauman, S. Aali, S. A. Razak, A. Zaman, M. Iqbal and S. N. Shah (2022). Assessment of floristic diversity in the mountain ecosystem of Marghazar Valley, Hindukush Range, Swat, Pakistan. Biodivers. J. 23(2): 1000-1013. https://doi. org/ 10.13057/biodiv/d230243

Khan, S., F. Hadi, U. Rabnawaz, F. Akhtar, A. Ullah and N. Uddin (2023). Floristic composition and biological spectrum of Baranai Valley, Madyan District, Swat Pakistan. J. Xi'an Shiyou Univ. Nat. Sci. Ed**.** 19(12): 1025-1046. https://www.xisdxjxsu.asia/V19I12-83.pdf

Kumar, S., G. Stecher, M. Suleski, M. Sanderford, S. Sharma and K. Tamura (2024). MEGA12: Molecular evolutionary genetic analysis (version 12) for adaptive and green computing. Mol. Biol. Evol. 41(12): 1-9. https://doi.org/10.1093/molbev/msae263

Kumar, V., V. Charde, S. B. Prasad, Y. Gandhi, S. K. Mishra, H. Rawat, A. Thakur, S. K. Shakya, T. Ansari and G. Babu (2023). Therapeutic potential of evergreen maiden hair fern *Adiantum venustum* D. Don: a comprehensive review. Food Chem. Adv. 3: 100439. https://doi.org/10.1016/j.focha.2023.100439

Kuo, L., S. Tang, Y. Huang, P. J. Xie, C. Chen, Z. Chang, T. Hsu, Y. Chang, Y. Chao and C. Chen (2024). A DNA barcode reference of Asian ferns with expert-identified voucher specimens and DNA samples. Sci. Rep. 11(1): 1-7. https://doi.org/10.1038/s41597-024-04161-8

Kusukawa, N., T. Uemori, K. Asada and I. Kato (1990). Rapid and reliable protocol for direct sequencing of material amplified by the polymerase chain reaction. Biotechniques. 9(1): 66-68. https://pubmed.ncbi.nlm.nih.gov/2393575

Lee, P. Y., J. Costumbrado, C. Y. Hsu and Y. H. Kim (2012). Agarose gel electrophoresis for the separation of DNA fragments. J. Vis. Exp. (62): 3923. doi: 10.3791/3923

Lu, J., J. Wen, S. Lutz, Y. Wang and D. Z. Li (2012). Phylogenetic relationships of Chinese *Adiantum* based on five plastid markers. J. Plant Res. 125: 237-249. https://doi.org/10.1007/s10265-011-0441-y

Ma, X. Y., C. X. Xie, C. Liu, J. Y. Song, H. Yao, K. Luo, Y. J. Zhu, T. Gao, X. H. Pang, J. Qian and S. L. Chen (2010). Species identification of medicinal pteridophytes by a DNA barcode marker, the chloroplast *psbA-trnH* intergenic region. Biol. Pharmace. Bull. 33(11): 1919-1924. https://doi.org/10.1248/bpb.33.1919

Madden, T. (2013). The BLAST sequence analysis tool. In: The NCBI handbook. 2nd ed. (eds. National Center for Biotechnology Information), NCBI; Bathesda (USA). 361-370 pp. https://www.ncbi.nlm.nih.gov/books/NBK143764

Morajkar, S., C. Suneetha, T. Harish Kumar and S. Hegde. (2022). Molecular markers in Pteridophytes. In: Ferns biotechnology, propagation, medicinal uses and environmental regulation. (eds. J. Marimuthu, H. Fernandez, A. Kumar and S. Thangaiah). Springer; Singapore. 99-134 pp. https://doi.org/10.1007/978-981-16-6170-9_5

Nilsson, R. H., M. Ryberg, E. Kristiansson, K. Abarenkov, K. H. Larsson and U. Kõljalg (2006). Taxonomic reliability of DNA sequences in public sequence databases: a fungal perspective. PloS One. 1(1): 1-4. https://doi.org/10.1371/journal.pone.0000059

Nithaniyal, S., B. Asir, K. Sarkar, R. Dayal and D. Jesubalan (2025). Molecular identification and phylogenetic analysis of endemic ferns of Dryopteridaceae from the Western Ghats of India. Preprint, SSRN. http://dx.doi.org/10.2139/ssrn.5174078.

Nitta, J. H. and S. M. Chambers (2022). Identifying cryptic fern gametophytes using DNA barcoding: A review. Appl. Plant Sci. 10(2): 1-10. https://doi.org/10.1002/aps3.11465

Nitta, J. H., A. Ebihara and A. R. Smith (2020). A taxonomic and molecular survey of the pteridophytes of the Nectandra cloud forest reserve, Costa Rica. PloS one, 15(11): 1-14. https://doi.org/10.1371/journal.pone.0241231

Pan, C., Y. Chen, X. Ma, J. Jiang, F. He and Y. Zhang (2011). Phytochemical constituents and pharmacological activities of plants from the genus *Adiantum*: A review. Trop. J. Pharm. Res. 10(5): 681-692. http/doi.org/10.4314/tjpr.v10i5.18

Pearson, W. R. (1999). Flexible sequence similarity searching with the FASTA3 program package. Methods Mol. Biol. 132: 185-219. https://doi.org/10.1385/1-59259-192-2:185

Pryer, K. M., E. Schuettpelz, L. Huiet, A. L. Grusz, C. J. Rothfels, T. Avent, D. Schwartz and M. D. Windham (2010). DNA barcoding exposes a case of mistaken identity in the fern horticultural trade. Mol. Ecol. Resour. 10(6): 979-985. https://doi.org/10.1111/j.1755-0998.2010.02858.x

Rastogi, S., M. M. Pandey and A. K. S. Rawat (2018). Ethnopharmacological uses, phytochemistry and pharmacology of genus *Adiantum*: A comprehensive review. J. Ethnopharmacol. 215: 101-119. https://doi.org/10.1016/j.jep.2017.12.034

Schneider, H. and E. Schuettpelz (2006). Identifying fern gametophytes using DNA sequences. Mol. Ecol. Notes, 6(4): 989-991. https://doi.org/10.1111/j.1471-8286.2006.01424.x

Schoch, C. L., S. Ciufo, M. Domrachev, C. L. Hotton, S. Kannan, R. Khovanskaya, D. Leipe, R. Mcveigh, K. O'Neill and B. Robbertse (2020). NCBI Taxonomy: a comprehensive update on curation, resources and tools. Database. 2020: 1-21. https://doi.org/10.1093/database/baaa062

Shrestha, U. B., S. Maharjan, A. Tiwari, Y. Luo, Phuentsho, S. K. Ghimire and B. B. Shrestha (2026). Divergent effects of climate change on the potential habitats of two medicinally important *Aconitum* species in the Hindu Kush Himalaya. Ecol. Evol. 16(1): 1-15. https://doi.org/10.1002/ece3.72965

Sundue, M. A. and C. J. Rothfels (2014). Stasis and convergence characterize morphological evolution in eupolypod II ferns. Ann. Bot. 113(1): 35-54. https://doi.org/10.1093/aob/mct247

Tamura, K. (1992). Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+ C-content biases. Mol. Biol. Evol. 9(4): 678-687. https://doi.org/10.1093/oxfordjournals.molbev.a040752

Thompson, J. D., B. Linard, O. Lecompte and O. Poch (2011). A comprehensive benchmark study of multiple sequence alignment methods: current challenges and future perspectives. PloS One. 6(3): 1-14. https://doi.org/10.1371/journal.pone.0018093

Trujillo-Argueta, S., R. F. Del Castillo, D. Tejero-Diez, C. A. Matias-Cervantes and A. Velasco-Murguía (2021). DNA barcoding ferns in an unexplored tropical montane cloud forest area of southeast Oaxaca, Mexico. Sci. Rep. 11(1): 22837. https://doi.org/10.1038/s41598-021-02237-8

Wang, F.H., J.M. Lu, J. Wen, A. Ebihara and D.Z. Li (2016). Applying DNA barcodes to identify closely related species of ferns: A case study of the Chinese *Adiantum* (Pteridaceae). PloS One. 11(9): 1-21. https://doi.org/10.1371/journal.pone.0160611

Zhao, J., J. G. Wang, Y. P. Hu, C. J. Huang, S. L. Fang, Z. Y. Wan, R. J. Li, H. Yu, Z. R. He and X. M. Zhou (2025). Phylogenetic Inferences and Historical Biogeography of Onocleaceae. Plants. 14(4): 1-18. https://doi.org/10.3390/plants14040510

Zheng, S. H., W. G. Ren, Z. H. Wang and L. F. Huang (2015). Use of chloroplast DNA barcodes to identify *Osmunda japonica* and its adulterants. Plant Syst. Evol. 301: 1843-1850. https://doi.org/10.1007/s00606-015-1197-y