



## Evaluating the discriminatory power of CBOL-Recommended DNA barcodes in the genus *Setaria* (Poaceae) from Taiwan

Chih-Hui CHEN<sup>1</sup>, Tzu-Yun LEE<sup>1</sup>, Hsien-Chun LIAO<sup>1</sup>, Pei-Luen LU<sup>2,\*</sup>

1. Taiwan Biodiversity Research Institute, 552005 No. 1 Minsheng East Road, Jiji Town, Nantou County, Taiwan. 2. Department of Life Science, National Taitung University, Taitung 950309, Taiwan. \*Corresponding author's email: peiluen@nttu.edu.tw

(Manuscript received 27 July 2025; Accepted 28 February 2026; Online published 23 March 2026)

**ABSTRACT:** DNA barcoding has emerged as a powerful tool for species identification, yet its efficacy can vary significantly across plant taxa. This study evaluates the discriminatory power of five DNA barcode loci—*rbcL*, *matK*, *trnH-psbA*, *trnL-F*, and ITS (internal transcribed spacer)—recommended by the Consortium for the Barcode of Life (CBOL), using *Setaria* (Paniceae, Panicoideae, Poaceae) species in Taiwan as a case study. Our results demonstrate that ITS and *trnL-F* possess the highest discriminatory capacity due to their high sequence variability and PCR success rates. In contrast, *trnH-psbA*, although easily amplified, provided minimal discriminatory power among species. Core barcodes *rbcL* and *matK* showed moderate effectiveness. Notably, some closely related species, such as *S. italica*, *S. viridis*, and *S. verticillata*, could not be reliably distinguished using any single marker. A combined approach using ITS and *trnL-F* provided markedly improved discriminatory power among species. These findings highlight the importance of taxon-specific marker selection and support the use of multi-locus barcoding strategies for effective plant identification, particularly in taxonomically complex genera like *Setaria*.

**KEY WORDS:** ITS, plant molecular taxonomy, phylogenetic resolution, Poaceae, *Setaria*, species discrimination, *trnL-F*.

### INTRODUCTION

The concept of DNA barcoding was formally introduced by Hebert *et al.* (2003), who proposed the use of short, standardized DNA regions to facilitate species identification. One year later, the Consortium for the Barcode of Life (CBOL) was established to promote global barcode efforts and to accelerate species discovery and taxonomic research. DNA barcoding was expected to provide a scalable solution to increasing taxonomic workloads and the global shortage of trained taxonomists.

While the mitochondrial cytochrome c oxidase subunit I (*CO1*) has been widely successful for distinguishing animal taxa (Hebert *et al.*, 2003), identifying a universal barcode for plants has proven far more challenging due to hybridization, introgression, polyploidy, and relatively low mutation rates in plastid genomes. Several candidate loci were proposed over the past two decades (Stoeckle, 2003; Kress *et al.*, 2005; Cowan *et al.*, 2006; Newmaster *et al.*, 2006; Chase *et al.*, 2007; Shaw *et al.*, 2007; Lahaye *et al.*, 2008), and the CBOL Plant Working Group eventually recommended *rbcL* and *matK* as core plant barcodes, with ITS and *trnH-psbA* as supplementary markers (CBOL Plant Working Group, 2009; China Plant BOL Group, 2011). Nevertheless, the performance of these markers varies considerably across plant lineages, and no single locus or locus combination has achieved universal applicability.

To address these limitations, additional plastid and nuclear loci—including *ycf1*, *rpoB*, *rpoC1*, *atpF-atpH*, *psbK-psbI*, and *trnL-F*—have been proposed as supplementary markers for resolving species in various

plant groups (Ferri *et al.*, 2009; Hart *et al.*, 2009; Dong *et al.*, 2015; Caetano Wyler and Naciri, 2016). The *trnL-F* region in particular has shown strong amplification success across angiosperms due to highly conserved primer sites (Taberlet *et al.*, 1991), and several studies have demonstrated its ability to improve species discrimination within Poaceae when used in multi-locus barcoding frameworks (Kress and Erickson, 2007; López-Alvarez *et al.*, 2012; Saarela *et al.*, 2017). In addition, *trnL-F* has extensive representation in GenBank, facilitating comparative analyses and phylogenetic inference. These characteristics provide a clear rationale for evaluating *trnL-F* as a supplementary barcode for *Setaria*.

Grasses (Poaceae) are one of the most taxonomically challenging plant families to identify based on morphology alone, due to reduced floral structures, phenotypic plasticity, and frequent character overlap among species (Saadullah *et al.*, 2016). Consequently, numerous DNA barcoding studies have attempted to resolve grass species, but results remain inconsistent, and no universal locus has been effective across the family (Supplementary Table S1). Multi-gene approaches using combinations of plastid and nuclear markers have been suggested to improve resolution within problematic grass genera (Chase *et al.*, 2007; Saadullah *et al.*, 2016; Wang *et al.*, 2017).

Gene availability in GenBank is also an important criterion when selecting barcodes for grasses. Hodkinson *et al.* (2007) noted that nrITS, *matK*, *rbcL*, *ndhF*, *trnL*, *gbss1*, *phB*, *rps16*, *rpoA*, and *rpoC2* showed the highest database coverage. Advances in massively parallel sequencing (MPS) and long-read technologies (Bleidorn, 2016; Hodkinson, 2018) now permit efficient sequencing



of multiple loci, including *trnL-F*. Importantly, *trnL-F* has been shown to amplify reliably across divergent Poaceae taxa using universal primers and has demonstrated strong discriminatory power in previous grass barcoding studies (Kress and Erickson, 2007; Saarela *et al.*, 2017). These attributes make *trnL-F* a valuable supplementary marker and justify its inclusion in this study.

The genus *Setaria* comprises nearly 100 species distributed worldwide, including agronomic crops, wild model organisms, and weedy species of economic concern. *Setaria italica* (foxtail millet) is a domesticated C4 cereal crop, whereas its wild ancestor *S. viridis* has become an important model for studying C4 photosynthesis, weed evolution, and abiotic stress tolerance (Brutnell *et al.*, 2010, 2015; Li and Brutnell, 2011; Huang *et al.*, 2016). Other species, including *S. faberi* and *S. verticillata*, are widespread weeds in agricultural landscapes. Although these species are morphologically distinguishable, previous studies have shown that commonly used plastid and nuclear markers often fail to resolve them genetically (Doust *et al.*, 2007; Kellogg *et al.*, 2009; Layton and Kellogg, 2014; Kim *et al.*, 2015). This highlights the need for a systematic evaluation of alternative barcode loci for the genus.

Benabdelmouna *et al.* (2001) used genomic in situ hybridization (GISH) to investigate polyploidy and genome composition in the genus *Setaria*. Their results showed that both *S. italica* and its putative wild ancestor *S. viridis* are diploids sharing the same A-genome haploid complement, whereas *S. adhaerans* is also diploid but possesses a distinct B-genome haploid set. In addition, *S. faberi* and *S. verticillata* were identified as allotetraploids, each carrying an AABB genome constitution. Zhao *et al.* (2013) corroborated these earlier findings and further characterized several additional genome types within the genus, including the C genome in diploid *S. grisebachii*; the D genome in both tetraploid and octoploid *S. glauca* as well as tetraploid *S. parviflora*; the E genome in tetraploid *S. plicata* and hexaploid *S. palmifolia*; and the F genome in hexaploid *S. arenaria*. Notably, they also observed that the B genome in *S. faberi* is distinct from the B genome in *S. verticillata*, suggesting independent evolutionary origins of the two tetraploids. Consistent with this, Kellogg (2017) indicated that tetraploid *S. faberi* may have arisen from two A-like progenitors rather than a conventional A + B hybrid origin.

In Taiwan, nine species were included in the Flora of Taiwan (Liu, 2000), including the cultivated foxtail millet, *S. italica*, and its wild relatives, *S. faberi*, *S. viridis*, and *S. verticillata*. Later, two naturalized species were added, namely *Setaria sphacelata* (Schumach.) Moss ex Stapf & C.E. Hu (Chen and Kuoh, 2006) and *Setaria barbata* (Lam.) Kunth. (Jung *et al.*, 2012). Among the 11 species, *Setaria pallide-fusca* (Schumach.) Stapf & C. E. Hubb. was treated as synonym of *Setaria parviflora* (Poiret) Kerguelen by Chen *et al.* (2014). The cultivation of

foxtail millet, *S. italica*, in Taiwan could be traced back to 5000 years ago (Tsang *et al.*, 2017). This study aims to assess the discrimination ability and operability of core and supplementary barcodes of plants, *rbcL*, *matK*, ITS, *trnH-psbA*, and supplemented with *trnL-F*, for the *Setaria* species in Taiwan, especially *S. italica* and its wild relatives.

Despite its ecological, agricultural, and cultural importance, the *Setaria* flora of Taiwan remains taxonomically challenging. Species identification is hindered by high morphological similarity, overlapping diagnostic characters, and frequent misidentifications in herbarium collections and floristic surveys. Closely related taxa such as *S. italica*, *S. viridis*, *S. faberi*, and *S. verticillata* are particularly prone to confusion, resulting in uncertain species boundaries. At present, no DNA barcode system has been systematically evaluated for Taiwanese *Setaria*, leaving a lack of molecular tools for reliable identification.

Given these taxonomic difficulties and the absence of prior barcoding assessments for Taiwanese *Setaria*, this study aims to evaluate the discriminatory power and practical applicability of core (*rbcL*, *matK*) and supplementary (ITS, *trnH-psbA*, *trnL-F*) plant DNA barcodes. By examining multiple plastid and nuclear loci, we seek to identify an effective barcode combination for accurate species identification in Taiwan. Establishing a robust molecular framework will facilitate biodiversity documentation and provide practical support for floristic research and weed management in Taiwanese agroecosystems, where several *Setaria* species are widespread and economically significant.

## MATERIALS AND METHODS

### Taxa sampling

A total of 124 samples of *Setaria* species were collected in the field of Taiwan and preserved in silica-gel for this study. They were one of *S. barbata*, three of *S. faberi*, two of *S. italica*, 14 of *S. pallidefusca*, six of *S. palmifolia*, six of *S. parviflora*, six of *S. plicata*, 22 of *S. pumila*, three of *S. sphacelata*, 32 of *S. verticillata*, and 29 of *S. viridis*. In addition to the above 124 *Setaria* samples, Five additional samples representing outgroup taxa were included: *Cenchrus echinatus* (1), *Eriochloa procer*a (1), *Panicum maximum* (1), *Pennisetum alopecuroides* (1), and *Pennisetum purpureum* (1). All vouchers of the above samples were deposited in the herbarium of Taiwan Biodiversity Research Institute (TAIE). Some accessions from GenBank were also included for both OTUs and outgroups in phylogenetic analyses and their accession number and identity were shown in phylogenetic trees.

### DNA extraction, PCR amplification and DNA sequencing

In this study, both coding- and noncoding regions from both nucleus and chloroplast genome were included.



In addition to the core sequences of plant barcoding, namely ITS, *rbcL*, *matK*, and *trnH-psbA*, we add *trnL-F* into this study in comparison with the core and supplementary DNA barcodes for plants in order to figure out their discrimination ability for *Setaria* species. The DNA sequences used for this study and the detailed information about the primers used for their PCR amplification were listed in supplementary data Table S2.

Total genomic DNAs were isolated using taco total DNA extraction kit-320 (GeneReach Biotechnology Corp., Taichung, Taiwan). Dried leaf materials were taken in about 1 cm square and ground using a mortar and pestle with liquid nitrogen. The powder was mixed with 600  $\mu$ L lysis buffer and shaken for one hour. The mixture was then transferred to 96-well extraction plate. The plate loaded with reagents provided in the kit and samples were installed in the taco nucleic acid automatic extraction system. The system then extracted total genomic DNAs automatically.

PCR amplifications were performed in total 20  $\mu$ L reaction volume containing 5  $\mu$ L genome DNAs extraction (20 $\times$  diluted), 2  $\mu$ L 10 $\times$  PCR buffer, 3.2  $\mu$ L 2.5 mM dNTP, 1.2  $\mu$ L 25 mM MgCl<sub>2</sub>, 1  $\mu$ L DMSO, 0.2  $\mu$ L Supertherm Taq, 1  $\mu$ L each of 2  $\mu$ M primers and 5.4  $\mu$ L ddH<sub>2</sub>O. The thermal cycles were performed with 95 $^{\circ}$ C for 5 min, followed with 35 cycles of 95 $^{\circ}$ C for 30 s, 52 $^{\circ}$ C for 30 s, 72 $^{\circ}$ C for 30 s, and final extension in 72 $^{\circ}$ C for 7 min. The PCR products were cleaned up with spin column of the kit. The Sanger Sequence reaction were done using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Carlsbad, CA, USA). Sequencing primers were the same as those used for PCR. The sequencing products then were analyzed by DNA analyzer 3730xl (Applied Biosystems).

### Sequence alignment and phylogenetic analysis

We examined the sequencing success rate of the five barcodes for all 124 samples of *Setaria*. Nevertheless, due to the different success rate of different loci, only the successfully amplified sequences from 72 (67 *Setaria* genus samples one *Cenchrus* genus samples one *Eriochloa* genus samples one *Panicum* genus samples two *Pennisetum*) samples were used to conduct further phylogenetic analyses finally. The detailed information and registered accession numbers on GenBank of successfully amplified loci were listed in supplementary data Table S3.

PCR amplified sequences were aligned using MUSCLE function of MEGA 11 (Kumar *et al.* 2021) and edited manually using BioEdit version 7.2.5 (Hall, 1999) for correction and trimming. In this study we utilize tree-based methods to figure out the discrimination ability. Bayesian inference (BI) and maximum likelihood (ML) phylogenies were reconstructed using MrBayes v3.2 (Ronquist *et al.*, 2012) and PhyML v3.1 (Guindon and Gascuel, 2003), respectively. The alignment matrix was

applied to jModelTest 2.1.10 (Guindon and Gascuel, 2003; Darriba *et al.*, 2012) to determine the substitution model used for reconstructing BI and ML phylogenies.

BI phylogeny was analyzed with four Markov chain Monte Carlo (MCMC) simulation for 2,000,000 generations and sampled one per 1000 generations. The first 800 trees of sample trees were discarded as burn-in before the probability was calculated. ML phylogeny was reconstructed under the setting of fixed proportion of invariable sites and four substitution rate categories. Internal support was assessed by 1000 bootstrap replicates (Felsenstein, 1985). Bootstrap (BS) values of 90%–100% were interpreted as strong support, 70%–89% as moderate, and under 69% as weak. Phylogenetic trees were visualized with TreeView (Page, 1996).

## RESULTS

### Sequence Characteristics of the Five Barcode Loci

The sequence characteristics of the five evaluated DNA barcode loci for Taiwanese species of the genus *Setaria* are summarized in Table 1. These include PCR success rates, total aligned sequence length, number of variable sites, number of parsimony-informative sites, number of indels, and average GC content.

Despite the relatively limited sample size, the PCR success rates provide meaningful insight into the practical usability of each marker. The highest success rates were observed for *trnH-psbA* and ITS (both >95%), followed by *rbcL* (92.7%) and *trnL-F* (85.4%). In contrast, *matK* exhibited the lowest PCR success rate (77%), suggesting lower amplification efficiency or sequencing difficulties.

Among the plastid loci, *matK*, *rbcL*, and *trnL-F* displayed aligned sequence lengths greater than 800 bp, with *matK* being the longest (871 bp). In contrast, *trnH-psbA* was notably shorter (486 bp). The nuclear ITS region, with an intermediate length of 569 bp, exhibited the highest sequence variation, containing 198 variable sites and 159 parsimony-informative sites. *trnL-F* showed the second-highest variability, whereas *trnH-psbA* had the lowest, with only 24 variable sites and 18 parsimony-informative sites.

Indel frequency varied among the loci. The ITS region contained seven indels, while the core plastid barcodes (*matK*, *rbcL*, and *trnH-psbA*) contained none. Notably, *trnL-F* showed the greatest number of indels among the plastid markers.

The GC content analysis revealed that the nuclear ITS region had the highest GC content (57.98%), while the four plastid loci showed lower values, ranging from 32.81% to 40.92%.

### Tree-Based Phylogenetic Analyses

Model selection results from jModelTest for phylogenetic analyses are also presented in Table 1. While the TIM model could not be directly implemented

**Table 1.** Molecular characteristics and parameters used in Bayesian analyses of the five loci evaluated for the species of genus *Setaria* in Taiwan.

	ITS	<i>matK</i>	<i>rbcL</i>	<i>psbA-trnH</i>	<i>trnL-F</i>
Sequencing success (%)	95.2 (118/124)	77.0 (94/122)	92.7 (114/123)	96.5 (110/114)	85.4 (105/123)
Total aligned characters (bp)	569	871	809	486	818
Variable sites	193	89	52	24	104
Informative sites	159	43	42	18	63
No. of Indels (outgroups excluded)	7	0	0	0	95
Average GC content (Standard error)	57.98 (1.06)	32.81 (0.21)	40.92 (0.16)	37.48 (0.25)	33.02 (0.33)
best model identified by jModelTest	GTR+I+G	TIM+I	TPM1uf+I+G	TIM2	TPM1uf+G
Substitution model used in MrBayes	6+I+G	6+I	6+I+G	6 (equal)	6+G
Substitution model used in PhyML	GTR	GTR	GTR	GTR	GTR

in MrBayes or PhyML, equivalent models were used where possible. MrBayes employed the best-fitting models for each locus, while PhyML applied the GTR model (Tavaré, 1986) uniformly across all loci. As no topological conflicts were detected between Bayesian Inference (BI) and Maximum Likelihood (ML) analyses, only the BI trees are presented and discussed below (Figures 1–5).

The phylogenetic analyses of the five individual loci revealed varying levels of phylogenetic resolution among Taiwanese *Setaria* species. The ITS phylogeny (Figure 1) showed a strongly supported polytomous clade comprising *S. italica* and its wild relatives (*S. faberi*, *S. viridis*, *S. verticillata*), reflecting their close evolutionary relationship. *S. plicata* appeared polytomous with a well-supported clade of *S. palmifolia*, while the discrimination power for *S. barbata* remained uncertain due to the availability of only a single successful sequence. Other species formed distinct monophyletic clades with strong support. The *matK* phylogeny (Figure 2) exhibited a similar topology to ITS, with the exception of an unresolved relationship between *S. palmifolia* and *S. pallide-fusca*, which formed a polytomous clade. Notably, *S. barbata* was well resolved in the *matK* tree. In the *rbcL* phylogeny (Figure 3), *S. italica* and its wild relatives again clustered into a strongly supported polytomous clade, while *S. pumila*, *S. sphacelata*, *S. pallide-fusca*, and *S. palmifolia* were successfully resolved. However, *S. palmifolia* and *S. plicata* were not clearly distinguished.

In contrast, the *trnH-psbA* phylogeny (Figure 4) showed the weakest resolution, grouping most *Setaria* species along with outgroup taxa (*Cenchrus* and *Pennisetum*) into a broad, strongly supported polytomous clade. Only *S. italica* and its wild relatives formed a moderately supported subclade, although some anomalous placements were observed—such as samples of *S. chondrachne*, *S. sphacelata*, and *S. pumila* nested within the *S. italica* clade. Interestingly, *S. verticillata* sequences from GenBank formed a separate, well-supported monophyletic clade. The *trnL-F* phylogeny (Figure 5) provided relatively good species-level resolution. While *S. italica*, *S. viridis*, and *S. verticillata* remained unresolved, *S. faberi* was clearly distinguished

as a monophyletic clade. However, similar to ITS, *trnL-F* failed to separate *S. palmifolia* from *S. plicata*, which together formed a polytomous group. Aside from these exceptions, most species were well resolved as distinct monophyletic lineages.

### Combined Loci Analysis

To determine whether increasing the number of informative sites could improve phylogenetic resolution among species, we reconstructed phylogenetic trees using concatenated datasets. However, combining all five loci did not improve discrimination among *Setaria* species. Owing to the low amplification success of *matK* and incomplete ITS sequences, the concatenated five-locus matrix contained substantial missing data and failed to resolve several species complexes (Figure 6). In particular, *S. italica*, *S. viridis*, and *S. verticillata* remained clustered in an unresolved polytomy, and overall resolution was not superior to analyses based on single loci.

In contrast, the targeted two-locus combination ITS + *trnL-F* produced the highest species-level resolution among all tested concatenations (Figure 7). This dataset successfully distinguished *S. plicata* and *S. palmifolia*, which were unresolved in the individual-locus trees, and clearly separated *S. faberi* from *S. italica* and its close relatives. Nonetheless, the *S. italica*–*S. viridis*–*S. verticillata* complex remained unresolved even with this combined dataset. Overall, ITS + *trnL-F* represented the most effective and operationally feasible multi-locus barcode combination for *Setaria*.

## DISCUSSION

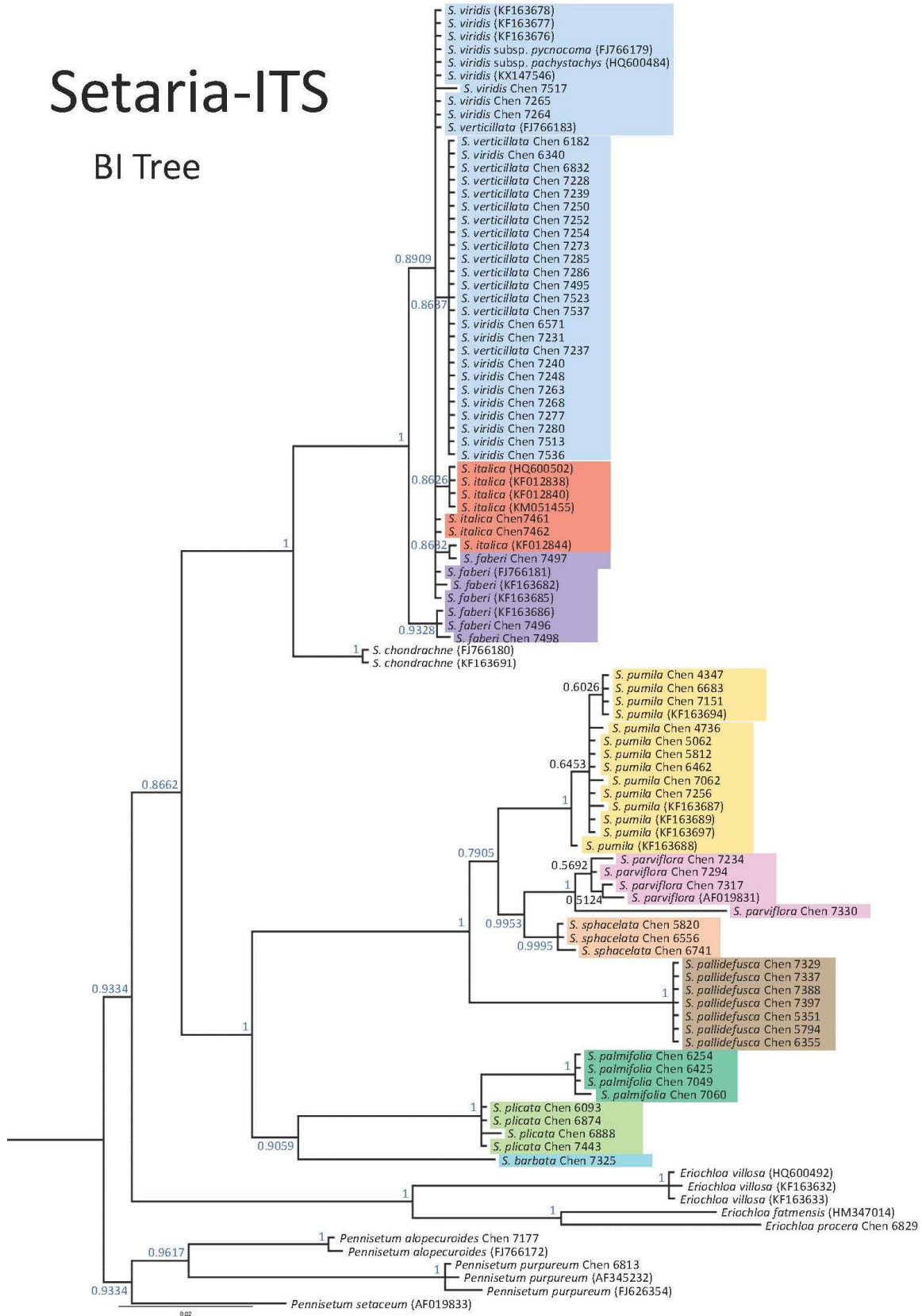
### Marker performance and taxon-specific barcode utility

The search for a single universal DNA barcode capable of reliably discriminating all grass species has been a long-standing objective in plant molecular systematics. However, accumulating evidence indicates that no single locus consistently achieves high discriminatory power across diverse taxa, as amplification efficiency and sequence variability differ substantially among markers (Hodkinson, 2018). In line with this broader pattern, the five loci evaluated in this



# Setaria-ITS

## BI Tree



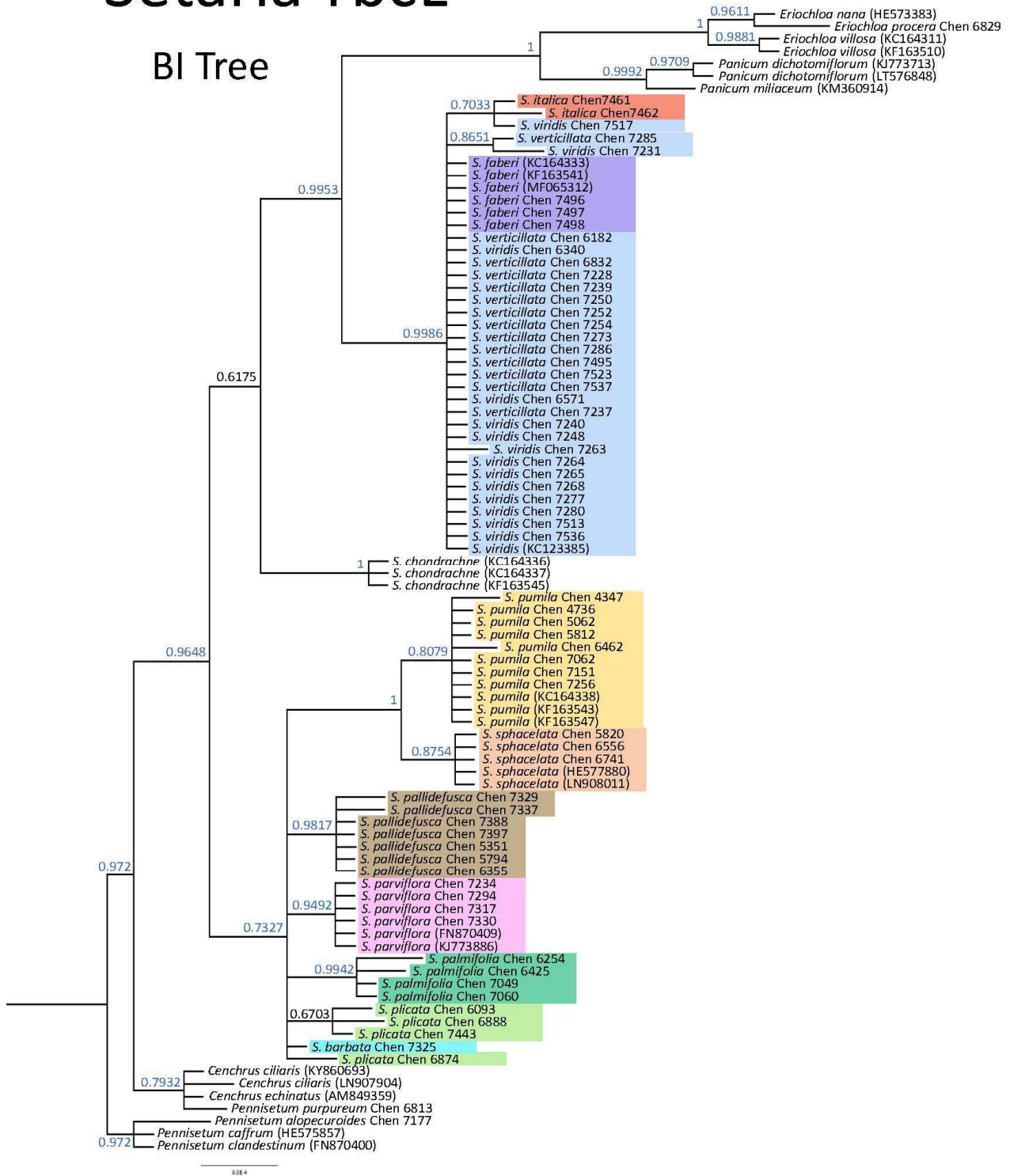
**Fig. 1.** Results of Bayesian analysis of the ITS data set illustrating relationships among *Setaria* species, rooted using six closely related species. Numbers above the branches represent posterior probabilities values above 0.5.





# Setaria-rbcl

BI Tree



**Fig. 3.** Results of Bayesian analysis of the *rbcl* data set illustrating relationships among *Setaria* species, rooted using six closely related species. Numbers above the branches represent posterior probabilities values above 0.5.



# Setaria-psbA-trnH

## BI Tree

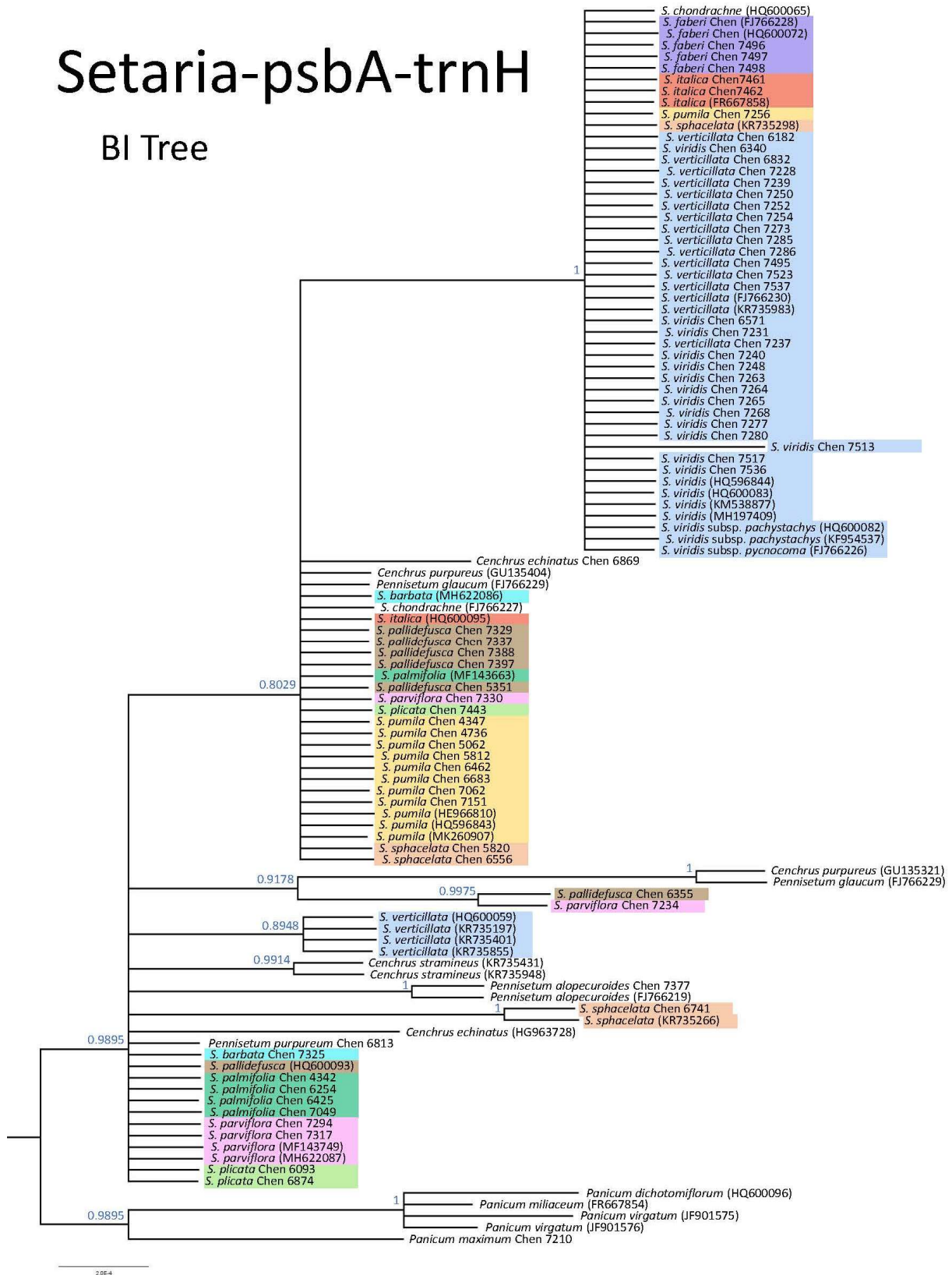
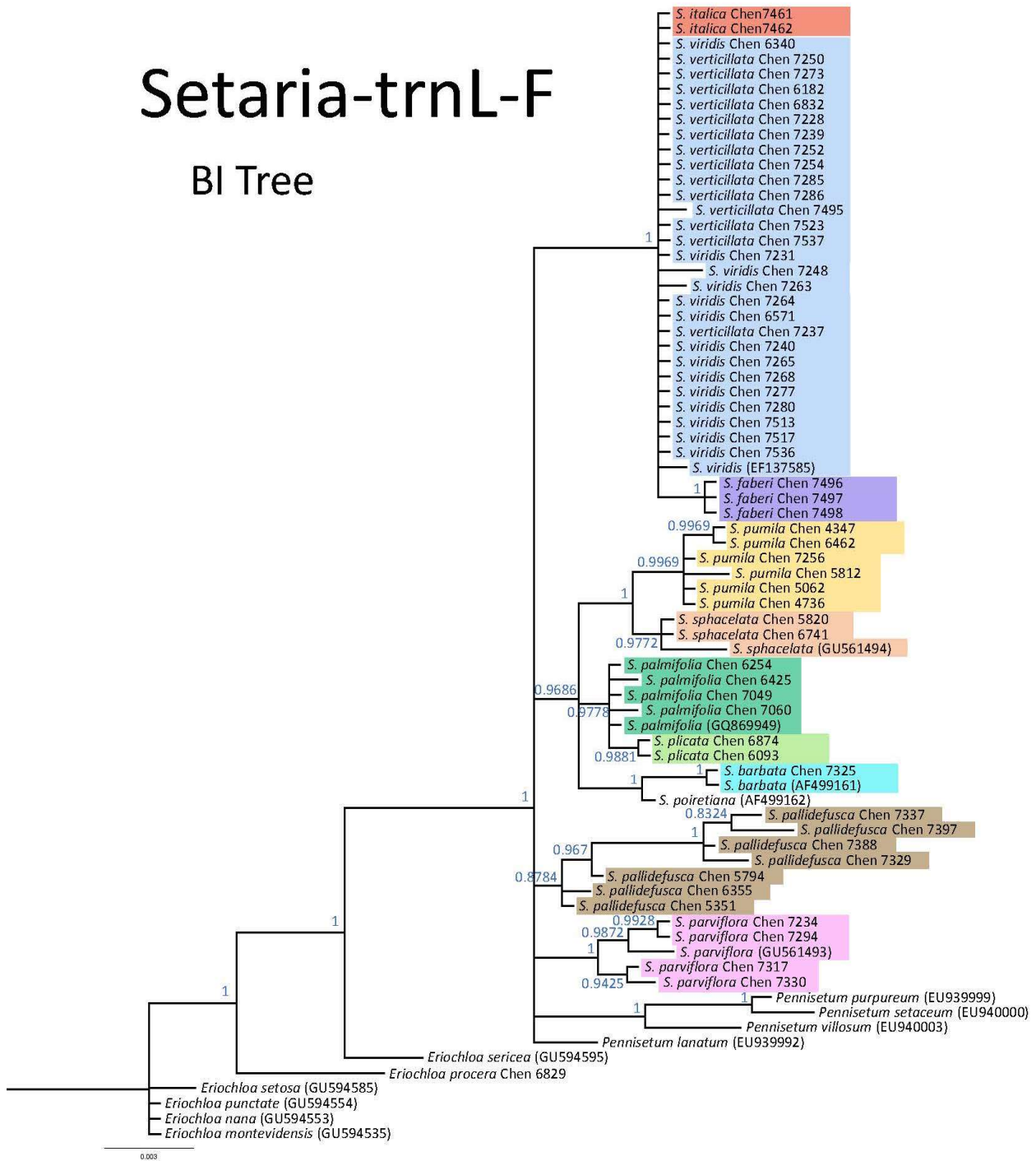


Fig. 4. Results of Bayesian analysis of the *psbA-trnH* data set illustrating relationships among *Setaria* species, rooted using four closely related species. Numbers above the branches represent posterior probabilities values above 0.5.



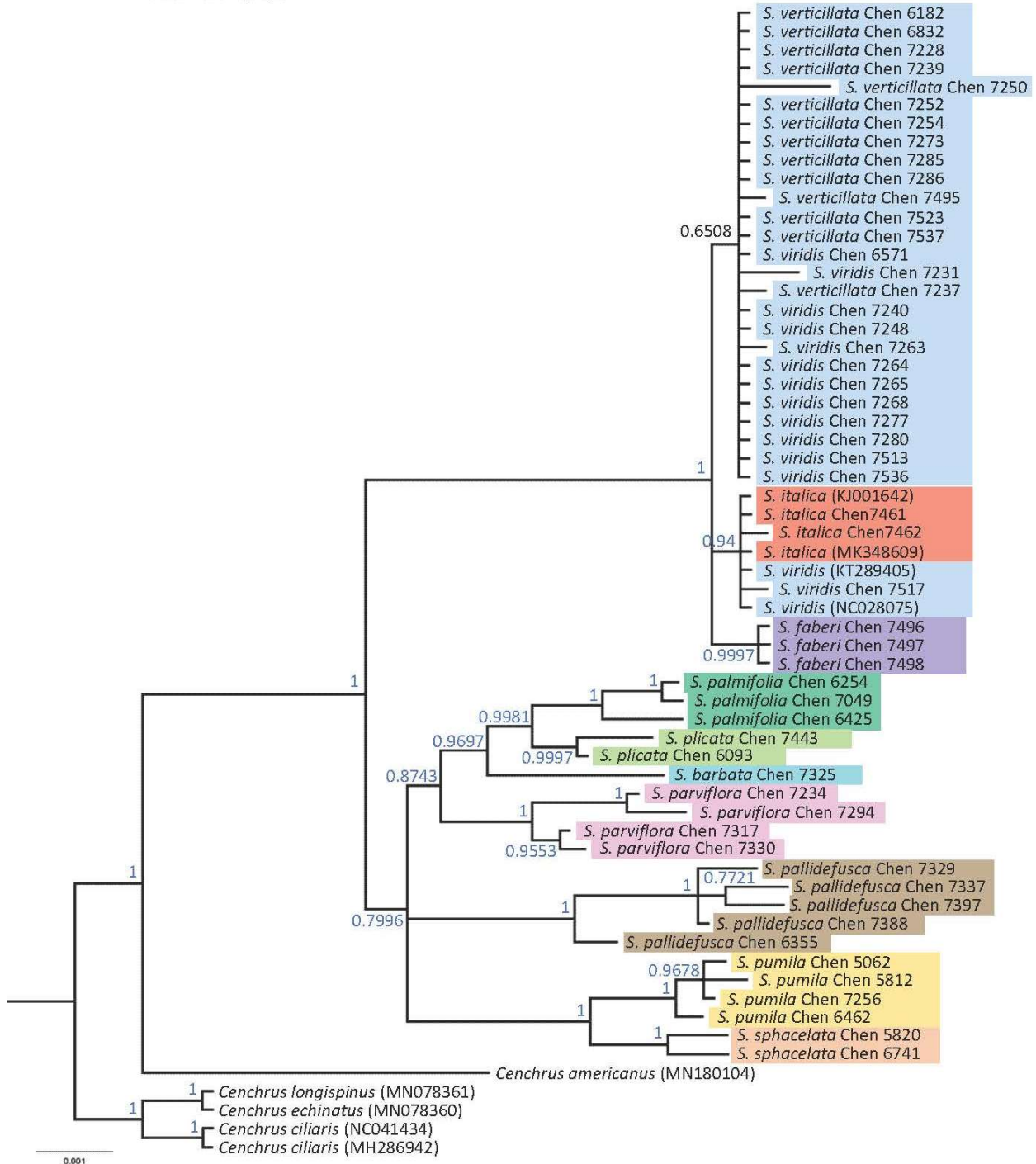
**Fig. 5.** Results of Bayesian analysis of the *trnL*-F data set illustrating relationships among *Setaria* species, rooted using ten closely related species. Numbers above the branches represent posterior probabilities values above 0.5.



# Setaria-Chloroplast

matK, trnL-F, rbcL, psbA-trnH

## BI Tree

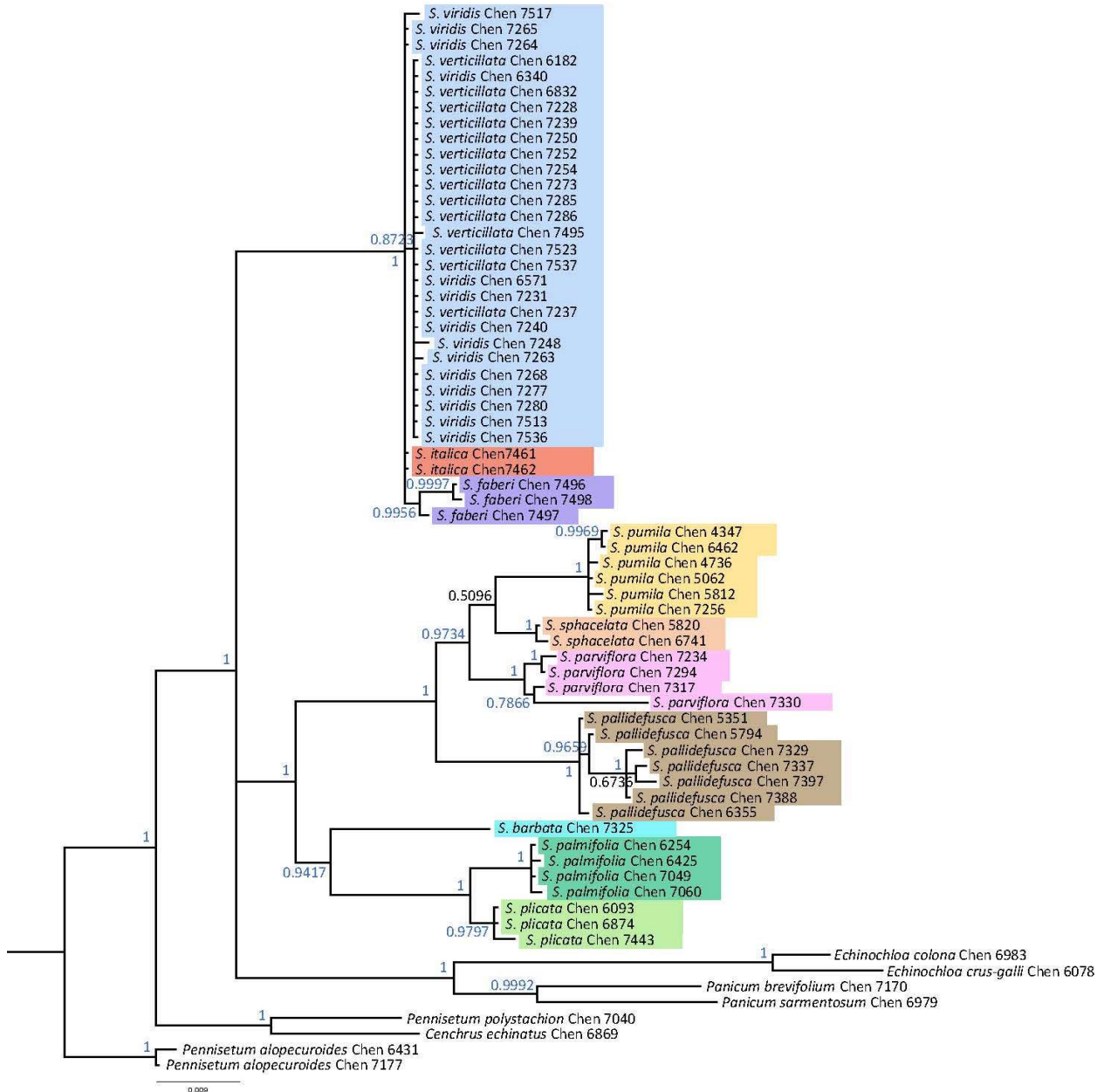


**Fig. 6.** Results of Bayesian analysis of the all chloroplast makers data set in this study illustrating relationships among *Setaria* species, rooted using four closely related species. Numbers above the branches represent posterior probabilities values above 0.5.



# Setaria-ITS+trnLF

## BI Tree



**Fig. 7.** Results of Bayesian analysis of the ITS + trnL-F data set illustrating relationships among *Setaria* species, rooted using seven closely related species. Numbers above the branches represent posterior probabilities values above 0.5.

study exhibited heterogeneous PCR success rates and varying levels of phylogenetic resolution for *Setaria* species in Taiwan.

Among the tested markers, ITS and *trnH-psbA* achieved the highest PCR success rates (>95%), suggesting that they are relatively easy to amplify and

sequence. By contrast, *matK* showed a notably lower success rate (77%), likely reflecting primer limitations or secondary structure issues that hinder amplification. These results are broadly consistent with previous studies reporting variable amplification success among standard barcode loci (Letsiou *et al.*, 2024).



Among the five loci, ITS contained the highest number of informative sites and demonstrated the strongest individual discriminatory capacity. *trnL-F* provided complementary plastid variation and uniquely resolved certain taxa that ITS alone did not distinguish. In contrast, *rbcL* and *matK* showed moderate performance, consistent with their generally conserved evolutionary rates in grasses. These results collectively indicate that locus variability and evolutionary rate, rather than PCR success alone, determine practical barcoding utility in *Setaria*.

This finding is consistent with most previous studies (Ragupathy *et al.*, 2009; Drumwright *et al.*, 2011; Wang *et al.*, 2017, 2022) but contrasts with some others (Syme *et al.*, 2013; Liu *et al.*, 2014; Saadullah *et al.*, 2016; Raveendar *et al.*, 2019; Skuza *et al.*, 2019; Loera-Sánchez *et al.*, 2020). While *rbcL* evolves slowly and is generally unsuitable for species-level discrimination, it is relatively easy to amplify and sequence (Letsiou *et al.*, 2024). Interestingly, its performance in this study exceeded initial expectations. Drumwright *et al.* (2011) suggested that the Consortium for the Barcode of Life (CBOL) recommended loci, *matK* and *rbcL*, could enable Poaceae species identification with 95% accuracy.

Among the five loci analyzed, *trnH-psbA* performed the worst. This result aligns with most previous studies (Syme *et al.*, 2013; Liu *et al.*, 2014; Saadullah *et al.*, 2016; Raveendar *et al.*, 2019; Skuza *et al.*, 2019; Loera-Sánchez *et al.*, 2020) but contradicts some others (Ragupathy *et al.*, 2009; Drumwright *et al.*, 2011; Wang *et al.*, 2017; Wang *et al.*, 2022). *trnH-psbA* has been recognized as one of the most variable non-coding plastid loci, with an intergenic spacer that often provides a high level of species discrimination (Chase *et al.*, 2007; Hollingsworth *et al.*, 2011). Although it exhibited a high PCR success rate in this study, its discriminatory power was poor. Pang *et al.* (2012) observed that the effectiveness of *trnH-psbA* varies across plant families, noting that its success rate in discriminating closely related Poaceae species was below 30%. While this locus has been successfully employed for species discrimination in many plant groups, it proved ineffective for *Setaria* species in this study, suggesting that its usefulness may be taxon-specific.

#### **Multi-Locus strategy: Complementarity over concatenation**

Combining loci proved more effective than relying on single markers, but only when the selected markers provided complementary phylogenetic signal. The ITS + *trnL-F* combination yielded the most stable and informative framework for species discrimination in Taiwanese *Setaria*. ITS contributed high nuclear variability, while *trnL-F* supplied additional chloroplast-based resolution. This result is consistent with several previous studies (Syme *et al.*, 2013; Wang *et al.*, 2022) but contrasts with others reporting limited phylogenetic utility for these loci (Liu *et al.*, 2014; Wang *et al.*, 2017;

Raveendar *et al.*, 2019). The two markers were selected for combined analysis because they offer complementary strengths in species delimitation within *Setaria* and exhibit relatively high PCR success rates. Together, they improved species-level discrimination beyond what either locus achieved independently.

In contrast, concatenating all five loci did not enhance resolution. The five-locus dataset contained substantial missing data due to incomplete recovery of *matK* and partial ITS sequences, resulting in unstable topologies and no improvement in resolving the *S. italica*–*S. viridis*–*S. verticillata* complex. These results parallel earlier studies (Kellogg *et al.*, 2009; Li *et al.*, 2012; Layton and Kellogg, 2014; Kim *et al.*, 2015), underscoring the challenge of genetically separating these taxa even when morphological differences are pronounced. These findings demonstrate that targeted marker selection based on empirical performance is more effective than indiscriminate multi-locus concatenation. A focused, lineage-informed strategy therefore provides a more practical framework for barcoding applications in *Setaria*.

#### **Limited genetic divergence in the *S. italica* complex**

A persistent limitation across all datasets was the inability to distinguish *S. italica*, its wild progenitor *S. viridis*, and the related *S. verticillata*. Despite morphological differentiation, these taxa share highly similar barcode sequences. This pattern likely reflects several interacting evolutionary processes documented in the genus (Doust and Kellogg, 2002; Bennetzen *et al.*, 2012).

First, *S. italica* represents a relatively recent domesticated derivative of *S. viridis*, leaving limited time for diagnostic mutations to accumulate, particularly in slowly evolving plastid loci (Li and Brutnell, 2011). Second, ongoing gene flow and introgression between cultivated and wild populations may homogenize barcode regions (Shi *et al.*, 2008; Fukunaga and Kawase, 2024). Third, incomplete lineage sorting in this young complex may preserve ancestral polymorphisms across species boundaries. Together, these processes constrain the resolving power of standard barcode markers.

Resolving this complex will likely require higher-resolution nuclear loci, such as ETS or low-copy genes, or genome-wide SNP datasets derived from genome skimming or RADseq approaches (Jia *et al.*, 2013). These methods can better capture recent divergence and reticulate evolutionary histories that are not detectable using conventional barcode regions.

#### **Species identification despite incomplete phylogenetic resolution**

DNA barcoding and phylogenetic reconstruction serve related but distinct purposes. While phylogenetic analyses aim to infer evolutionary relationships, barcoding focuses on accurate taxonomic assignment. Consequently,



incomplete resolution of deeper nodes does not necessarily preclude reliable species identification.

In this study, several *Setaria* species formed well-supported monophyletic clades across loci, enabling consistent identification even when relationships among clades remained ambiguous. These results highlight that diagnostic power—rather than complete phylogenetic resolution—is the primary criterion for evaluating barcode effectiveness. Nonetheless, the persistent ambiguity within recently diverged lineages underscores the limits of standard barcode loci for resolving complex evolutionary scenarios.

### Polyploidy and genomic complexity

The lack of detectable B-genome copies in *S. verticillata* and *S. faberi*, both recognized as allotetraploids with AABB genomic constitutions (Benabdelmouna *et al.*, 2001; Zhao *et al.*, 2013), warrants explanation. For nuclear markers such as ITS, the absence of divergent homeologs is most likely attributable to concerted evolution, a common phenomenon in polyploid plants. Ribosomal DNA arrays often undergo homogenization through unequal crossing-over and gene conversion, resulting in the dominance of a single sequence type across subgenomes. As a consequence, ITS frequently behaves as a single-locus barcode even in known allopolyploid species, masking underlying genome duplications.

Primer specificity may also contribute to this pattern. Standard ITS primers are known to preferentially amplify certain rDNA variants, potentially biasing PCR toward A-genome copies while failing to recover more divergent B-genome homeologs. For plastid loci, the situation is clearer: because plastids are maternally inherited, all chloroplast barcodes (*rbcL*, *matK*, *trnH-psbA*, *trnL-F*) reflect only the maternal A- or B-genome donor and cannot retrieve the biparental subgenomic history of allotetraploids. Therefore, the apparent absence of B-genome variants in our plastid datasets is expected.

Those factors indicate that standard barcode loci—particularly multicopy nuclear ITS and maternally inherited plastid markers—are inherently limited in their ability to capture the genomic complexity of allotetraploid *Setaria* species. Higher-resolution nuclear markers (e.g., low-copy genes, ETS) or genome-wide approaches (e.g., genome skimming, Hyb-Seq, RADseq) will likely be necessary to reveal the full extent of A- and B-genome divergence in *S. faberi* and *S. verticillata*.

### Taxonomic clarification of *Setaria pallide-fusca*

Our analyses provide clear molecular evidence supporting the distinctiveness of *Setaria pallide-fusca* in Taiwan. This taxon was previously treated as a synonym of *S. parviflora* (Chen *et al.*, 2014), based primarily on morphological assessment. In contrast, all samples identified as *S. pallide-fusca* in the present study

consistently formed a well-supported monophyletic clade distinct from *S. parviflora*. These findings confirm that *S. pallide-fusca* represents an independently evolving lineage and should be reinstated as a valid species in the Taiwan flora. Although secondary to the primary objective of barcode evaluation, this clarification resolves a significant taxonomic ambiguity and provides a molecular basis for future systematic work.

### Enhancing species resolution in *Setaria*

Our results show that although the CBOL-recommended barcodes provide a useful baseline for plant identification, they are not universally effective across all plant groups. In the case of *Setaria* in Taiwan, a taxon-specific strategy incorporating ITS and *trnL-F* achieved the highest and most consistent species-level resolution, outperforming both single-locus datasets and the full multi-locus concatenated approach. These findings underscore the growing consensus that barcode selection must be empirically validated within each lineage, rather than applied uniformly.

Looking forward, additional nuclear and plastid markers—such as ETS, *ycf1*, or broader genomic approaches like genome skimming—may help resolve persistent species complexes, particularly within the *S. italica*–*S. viridis*–*S. verticillata* group. Expanding population-level sampling and incorporating geographic data would further clarify patterns of gene flow, hybridization, and cryptic diversification, all of which are common in grasses and likely contribute to the limited discriminatory power of standard barcodes in this genus. Continued refinement of marker choices and sampling strategies will enhance the reliability of DNA barcoding for *Setaria* and improve its broader utility in floristic, ecological, and agricultural research.

## CONCLUSION

This study highlights the taxon-specific nature of DNA barcode performance within *Setaria* and reinforces the broader principle that barcode selection must be empirically validated within a defined phylogenetic and geographic context. Rather than supporting a universal solution, our findings demonstrate that effective species discrimination depends on strategic marker combination tailored to lineage-specific evolutionary dynamics.

The results further emphasize the importance of integrative approaches in resolving species boundaries in morphologically similar and taxonomically complex groups. In Taiwanese *Setaria*, combining molecular evidence with careful morphological reassessment proved essential for clarifying ambiguous taxa and re-evaluating historical classifications. At the same time, the limited resolution observed in certain lineages underscores the constraints of standard barcode loci for disentangling recent divergence or reticulate evolutionary histories.



Future investigations should incorporate broader geographic sampling and higher-resolution genomic data to resolve remaining uncertainties in species relationships and to clarify patterns of diversification within the genus. The development of lineage-informed, data-driven barcoding frameworks will further improve species identification accuracy and enhance their application in biodiversity assessment, floristic research, and agroecosystem management.

## ACKNOWLEDGMENTS

This work was supported by the Ministry of Agriculture, Executive Yuan, Republic of China. The help from staff of herbaria TAI, TAIE, TAI, TNM and PPI was also deeply appreciated.

## LITERATURE CITED

- Benabdelmouna, A., Shi, Y., Abirached-Darmency, M., Darmency, H. 2001 Genomic in situ hybridization (GISH) discriminates between the A and the B genomes in diploid and tetraploid *Setaria* species. *Genome* **44**(4): 685–690.
- Bennetzen, J.L., Schmutz, J., Wang, H., Percifield, R., Hawkins, J., Pontaroli, A.C., Estep, M., Feng, L., Vaughn, J.N., Grimwood, J., Jenkins, J., Barry, K., Lindquist, E., Hellsten, U., Deshpande, S., Wang, X., Wu, X., Mitros, T., Triplett, J., Yang, X., Ye, C.-Y., Mauro-Herrera, M., Wang, L., Li, P., Sharma, M., Sharma, R., Ronald, P.C., Panaud, O., Kellogg, E.A., Brutnell, T.P., Doust, A.N., Tuskan, G.A., Rokhsar, D., Devos, K.M. 2012 Reference genome sequence of the model plant *Setaria*. *Nat. Biotechnol.* **30**(6): 555–561.
- Bleidorn, C. 2016 Third generation sequencing: technology and its potential impact on evolutionary biodiversity research. *Syst. Biodivers.* **14**(1): 1–8.
- Brutnell, T.P., Bennetzen, J.L., Vogel, J.P. 2015 *Brachypodium distachyon* and *Setaria viridis*: Model Genetic Systems for the Grasses. *Annu. Rev. Plant Biol.* **66**: 465–485.
- Brutnell, T.P., Wang, L., Swartwood, K., Goldschmidt, A., Jackson, D., Zhu, X.-G., Kellogg, E., Van Eck, J. 2010 *Setaria viridis*: A Model for C4 Photosynthesis. *The Plant Cell* **22**(8): 2537–2544.
- Caetano Wyler, S., Naciri, Y. 2016 Evolutionary histories determine DNA barcoding success in vascular plants: seven case studies using intraspecific broad sampling of closely related species. *BMC Evol. Biol.* **16**(1): 103.
- CBOL Plant Working Group: Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M.W., Cowan, R.S., Erickson, D.L., Fazekas, A.J., Graham, S.W., James, K.E., Kim, K.-J., Kress, W.J., Schneider, H., van AlphenStahl, J., Barrett, S.C.H., van den Berg, C., Bogarin, D., Burgess, K.S., Cameron, K. M., Carine, M., Chacón, J., Clark, A., Clarkson, J.J., Conrad, F., Devey, D.S., Ford, C.S., Hedderson, T.A.J., Hollingsworth, M. L., Husband, B.C., Kelly, L.J., Kesanakurti, P.R., Kim, J.S., Kim, Y.-D., Lahaye, R., Lee, H.-L., Long, D.G., Madriñán, S., Maurin, O., Meusnier, I., Newmaster, S.G., Park, C.-W., Percy, D.M., Petersen, G., Richardson, J.E., Salazar, G.A., Savolainen, V., Seberg, O., Wilkinson, M.J., Yi, D.-K., Little, D.P. 2009 A DNA barcode for land plants. *PNAS* **106**(31): 12794–12797.
- Chase, M., Cowan, R., Hollingsworth, P., van den Berg, C., Madriñán, S., Petersen, G., Seberg, O., Jorgensen, T., Cameron, K., Carine, M., Pedersen, N., Hedderson, T., Conrad, F., Richardson, J.E., Hart, M., Barraclough, T., Kelly, L., Wilkinson, M. 2007 A proposal for a standardised protocol to barcode all land plants. *Taxon* **56**(2): 295–299.
- Chen, C.-H., Kuoh, C.-S. 2006 *Setaria sphacelata* (Schumach.) M.B. Moss ex Stapf & C.E. Hubb., a grass newly naturalized to Taiwan. *Endemic Species Research* **8**: 83–86.
- Chen, C.-H., Lin, C.-Y., Kuoh, C.-S. 2014 Grass Flora of Taiwan (2 of 3): Paniceae, Panicoideae. *Endemic Species Research Institute, Jiji, Taiwan*. 156 pp.
- Chen, C.-H., Veldkamp, J.F., Kuoh, C.-S. 2011 *Microstegium glabratum* (Brongn.) A. Camus (Poaceae, Andropogoneae), a new record for Japan, China, and Taiwan, based on morphological and molecular evidence. *Taiwania* **56**(2): 111–117. (see supplementary)
- Chen, C.-H., Veldkamp, J.F., Kuoh, C.-S., Tsai, C.-C., Chiang, Y.-C. 2009 Segregation of *Leptatherum* from *Microstegium* (Andropogoneae, Poaceae) confirmed by Internal Transcribed Spacer DNA sequences. *Blumea* **54**(1): 175–180. (see supplementary)
- Chiang, T.-Y., Schaal, B.A., Peng, C.-I. 1998 Universal primers for amplification and sequencing a noncoding spacer between the *atpB* and *rbcL* genes of chloroplast DNA. *Bot. Bull. of Acad. Sin.* **39**: 245–250. (see supplementary)
- China Plant BOL Group: Li, D.-Z., Gao, L.-M., Li, H.-T., Wang, H., Ge, X.-J., Liu, J.-Q., Chen, Z.-D., Zhou, S.-L., Chen, S.-L., Yang, J.-B., Fu, C.-X., Zeng, C.-X., Yan, H.-F., Zhu, Y.-J., Sun, Y.-S., Chen, S.-Y., Zhao, L., Wang, K., Yang, T., Duan, G.-W. 2011 Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *PNAS* **108**(49): 19641–19646.
- Cowan, R.S., Chase, M.W., Kress, W.J., Savolainen, V. 2006 300,000 species to identify: problems, progress, and prospects in DNA barcoding of land plants. *Taxon* **55**(3): 611–616.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D. 2012 jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **9**(8): 772.
- Dong, W., Xu, C., Li, C., Sun, J., Zuo, Y., Shi, S., Cheng, T., Guo, J., Zhou, S. 2015 *ycf1*, the most promising plastid DNA barcode of land plants. *Sci. Rep.* **5**: 8348.
- Doust, A. N., Kellogg, E. A. 2002 Phylogeny and character evolution in the genus *Setaria* (Poaceae). *Am. J. Bot.* **89**(8): 1203–1222.
- Doust, A.N., Penly, A.M., Jacobs, S.W.L., Kellogg, E.A. 2007 Congruence, Conflict, and Polyploidization Shown by Nuclear and Chloroplast Markers in the Monophyletic "Bristle Clade" (Paniceae, Panicoideae, Poaceae). *Syst. Bot.* **32**(3): 531–544.
- Drumwright, A.M., Allen, B.W., Huff, K.A. Ritchey, P.A., Cahoon, A.B. 2011 Survey and DNA Barcoding of Poaceae in Flat Rock Cedar Glades and Barrens State Natural Area, Murfreesboro, Tennessee. *Castanea* **76**(3): 300–310.
- Felsenstein, J. 1985 Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**(4): 783–791.



- Ferri, G., Al, M., Corradini, B., Beduschi, G. 2009 Forensic botany: species identification of botanical trace evidence using a multigene barcoding approach. *Int. J. Leg. Med.* **123**(5): 395–401.
- Fukunaga, K., Kawase, M. 2024 Crop evolution of foxtail millet. *Plants* **13**(2): 218.
- Guindon, S., Gascuel, O. 2003 A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Syst. Biol.* **52**(5): 696–704.
- Hall, T.A. 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic acids symposium series* **41**: 95–98.
- Hart, M., Clark, A., Forrest, L., Richardson, J.E., Pennington, R., Long, D., Cowan, R., Chase, M., Gaudeul, M., Hollingsworth, P. 2009 Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants: DNA barcoding. *Mol. Ecol. Resour.* **9**(2): 439–457.
- Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R. 2003 Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B Biol. Sci.* **270**(1512): 313–321.
- Hodkinson, T.R., Bouchenak-Khelladi, Y., Kinney, M.S., Savolainen, V., Jacobs, Surrey W.L., Salamin, N. 2007 Supersizing: progress in documenting and understanding grass species richness. In: Hodkinson, T.R., Parnell, J.A.N. (eds.), *Reconstructing the Tree of Life: Taxonomy and Systematics of Species Rich Taxa*, pp.276–290. CRC Press, Boca Raton, FL.
- Hodkinson, T.R. 2018 Evolution and Taxonomy of the Grasses (Poaceae): A Model Family for the Study of Species-Rich Groups. *Annual Plant Reviews Online* **1**: 255–294.
- Hollingsworth, P.M., Graham, S.W., Little, D.P. 2011 Choosing and Using a Plant DNA Barcode. *Plos One* **6**(5): e19254.
- Huang, P., Shyu, C., Coelho, C.P., Cao, Y., Brutnell, T.P. 2016 *Setaria viridis* as a Model System to Advance Millet Genetics and Genomics. *Front. Plant Sci.* **7**: 1–9.
- Jia, G., Huang, X., Zhi, H., Zhao, Y., Zhao, Q., Li, W., Chai, Y., Yang, L., Liu, K., Lu, H., Zhu, C., Lu, Y., Zhou, C., Fan, D., Weng, Q., Guo, Y., Huang, T., Zhang, L., Lu, T., Feng, Q., Hao, H., Liu, H., Lu, P., Zhang, N., Li, Y., Guo, E., Wang, S., Wang, S., Liu, J., Zhang, W., Chen, G., Zhang, B., Li, W., Wang, Y., Li, H., Zhao, B., Li, J., Diao, X., Han, B. 2013 A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nat Genet.* **45**(8): 957–961.
- Jung, M.-J., Chu, M.-K., Hsu, T.-C., Kao, R.-C., Dai, S.-S. 2012 Four newly naturalized grasses and rushes in Taiwan. *Taiwania* **57**(4): 426–433.
- Kellogg, E.A. 2017 Evolution of *Setaria*. In: Doust, A., Diao, X. (eds.) *Genetics and Genomics of Setaria*. Springer, Switzerland. 3–28.
- Kellogg, E.A., Aliscioni, S.S., Morrone, O., Pensiero, J., Zuloaga, F. 2009 A Phylogeny of *Setaria* (Poaceae, Panicoideae, Paniceae) and Related Genera Based on the Chloroplast Gene *ndhF*. *Int. J. Plant Sci.* **170**(1): 117–131.
- Kim, S., Kim, C.-S., Lee, J., Lee, I.-Y., Chung, Y.-J., Cho, M.-S., Kim, S.-C. 2015 Phylogenetic relationships among species of *Setaria* (Paniceae; Panicoideae; Poaceae) in Korea: insights from nuclear (ITS and *kn1*) and chloroplast DNA sequence data. *Plant Syst. Evol.* **301**(2): 725–736.
- 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.
- Kress, W.J., Wurdack, K.J., Zimmer, E.A., Weigt, L.A., Janzen, D.H. 2005 Use of DNA barcodes to identify flowering plants. *PNAS* **102**(23): 8369–8374.
- Kumar, K., Stecher, G., Tamura, S. 2021 MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Mol. Biol. Evol.* **38**(7): 3022–3027.
- Lahaye, R., van der Bank, M., Bogarin, D., Warner, J., Pupulin, F., Gigot, G., Maurin, O., Duthoit, S., Barraclough, T.G., Savolainen, V. 2008 DNA barcoding the floras of biodiversity hotspots. *PNAS* **105**(8): 2923–2928.
- Layton, D.J., Kellogg, E.A. 2014 Morphological, phylogenetic, and ecological diversity of the new model species *Setaria viridis* (Poaceae: Paniceae) and its close relatives. *Am. J. Bot.* **101**(3): 539–557.
- Letsiou, S., Madesis, P., Vasdekis, E., Montemurro, C., Grigoriou, M.E., Skavdis, G., Moussis, V., Koutelidakis, A.E., Tzakos, A.G. 2024 DNA barcoding as a plant identification method. *Appl. Sci.* **14**(4): 1415.
- Li, P., Brutnell, T.P. 2011 *Setaria viridis* and *Setaria italica*, model genetic systems for the Panicoid grasses. *J. Exp. Bot.* **62**(9): 3031–3037.
- Li, W., Zhi, H., Wang, Y.-F., Li, H.-Q., Diao, X.-M. 2012 Assessment of Genetic Relationship of Foxtail Millet with Its Wild Ancestor and Close Relatives by ISSR Markers. *J. Integr. Agric.* **11**(4): 556–566.
- Lindberg, C.L. 2016 Evolution of growth rates in Pooideae (Poaceae). Master's thesis, Norwegian University of Life Sciences, Ås, Norway. (see supplementary)
- Liu, H.-Y. 2000 Paniceae. In: Boufford, D. E., Hsieh, C.-F., Huang, T.-C., Kuoh, C.-S., Ohasi, H. Su, H.-J. (eds), *Flora of Taiwan*, 2nd ed. **5**: 453–521. Editorial Committee of the Flora of Taiwan, Dept. Bot. National Taiwan University, Taipei, Taiwan.
- Liu, J., Zhou, S., Li, C., Dong, W., Yang, R. 2014 DNA barcoding of *Leymus* (Poaceae). *JRAAS* **2**: 1–11.
- Loera-Sánchez, M., Studer, B., Kölliker, R. 2020 DNA barcode *trnH-psbA* is a promising candidate for efficient identification of forage legumes and grasses. *BMC Res. Notes* **13**(1): 35.
- López-Alvarez, D., López-Herranz, M.L., Betekhtin, A., Catalán, P. 2012 A DNA Barcoding Method to Discriminate between the Model Plant *Brachypodium distachyon* and Its Close Relatives *B. stacei* and *B. hybridum* (Poaceae). *Plos One* **7**(12): e51058.
- Newmaster, S., Fazekas, A., Ragupathy, S. 2006 DNA barcoding in land plants: Evaluation of *rbcL* in a multigene tiered approach. *Can. J. Bot.* **84**(3): 335–341.
- Page, R.D. 1996 TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**(4): 357–358.
- Pang, X., Liu, C., Shi, L., Liu, R., Liang, D., Li, H., Cherny, S.S., Chen, S. 2012 Utility of the *trnH-psbA* intergenic spacer region and its combinations as plant DNA barcodes: a meta-analysis. *Plos One* **7**(11): e48833–e48833.
- Ragupathy, S., Newmaster, S.G., Murugesan, M., Balasubramaniam, V. 2009 DNA barcoding discriminates a new cryptic grass species revealed in an ethnobotany study



- by the hill tribes of the Western Ghats in southern India. *Mol. Ecol. Resour.* **9**(s1): 164–171.
- Raveendar, S., Lee, G.-A., Lee, K.J., Shin, M.-J., Kim, S.H., Lee, J.-R., Cho, G.-T., Hyun, D.Y.** 2019 DNA Barcoding for Efficient Identification of *Triticum* Subspecies: Evaluation of Four Candidate Loci on Phylogenetic Relationships. *Plant Breed. Biotech.* **7**(3): 220–228.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Largetr, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P.** 2012 MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**(3): 539–542.
- Saadullah, Khan, Z.-u.-d., Ashfaq, M., Z.-u.-N.** 2016 Identification of the grass family (Poaceae) by using the plant DNA barcodes *rbcL* and *matK*. *J. Bio. Env. Sci.* **8**: 175–186.
- Saarela, J.M., Bull, R.D., Paradis, M.J., Ebata, S.N., Peterson, P.M., Soreng, R.J., Paszko, B.** 2017 Molecular phylogenetics of cool-season grasses in the subtribes Agrostidinae, Anthoxanthinae, Aveninae, Brizinae, Calothecinae, Koeleriinae and Phalaridinae (Poaceae, Pooideae, Poaceae, Poaceae chloroplast group 1). *Phytokeys* **87**: 1–139.
- Shaw, J., Lickey, E.B., Schilling, E.E., Small, R.L.** 2007 Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *Am. J. Bot.* **94**(3): 275–288.
- Shi, Y., Wang, T., Li, Y., Darmency, H.** 2008 Impact of transgene inheritance on the mitigation of gene flow between crops and their wild relatives: the example of foxtail millet. *Genetics* **180**(2): 969–75.
- Skuz, L., Szuko, I., Filip, E., Adamczyk, A.** 2019 DNA Barcoding in Selected Species and Subspecies of Rye (Secale) Using Three Chloroplast Loci (*matK*, *rbcL*, *trnH-psbA*). *Not. Bot. Horti Agrobot. Cluj-Na.* **47**(1): 54–62.
- Stoeckle, M.** 2003 Taxonomy, DNA and the bar code of life. *BioScience* **53**(9): 2–3.
- Syme, A.E., Udovicic, F., Stajsic, V., Murphy, D.J.** 2013 A test of sequence-matching algorithms for a DNA barcode database of invasive grasses. *DNA Barcode* **1**: 19–26.
- Taberlet, P., Gielly, L., Pautou, G., Jean, B.** 1991 Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* **17**(5): 1105–1109.
- Tavaré, S.** 1986 Some probabilistic and statistical problems on the analysis of DNA sequences. In: Miura, R.M. (ed.), *Lectures in Mathematics in the Life Sciences*. pp. 57–86. Providence (RI): American Mathematical Society.
- Tsang, C.-H., Li, K.-T., Hsu, T.-F., Tsai, Y.-C., Fang, P.-H., Hsing, Y.-I.C.** 2017 Broomcorn and foxtail millet were cultivated in Taiwan about 5000 years ago. *Bot. Stud.* **58**(1): 3–3.
- Wang, A., Gopurenko, D., Wu, H., Lepschi, B.** 2017 Evaluation of six candidate DNA barcode loci for identification of five important invasive grasses in eastern Australia. *Plos One* **12**(4): e0175338.
- Wang, A., Gopurenko, D., Wu, H., Stanton, R., Lepschi, B.J.** 2014 DNA barcoding for identification of exotic grass species present in eastern Australia. In 19th Australasian Weeds Conference. M. Baker, editor. Tasmanian Weed Society Hobart, Tasmania, Australia. 444–447 (see supplementary).
- Wang, J., Yan, Z., Zhong, P., Shen, Z., Yang, G., Ma, L.** 2022 Screening of universal DNA barcodes for identifying grass species of Gramineae. *Front. Plant Sci.* **13**: 998863.
- Zhang, X., Liu, Y., Gu, X., Guo, Z., Li, L., Song, X., Liu, S., Zang, Y., Li, Y., Liu, C., Wei, S.** 2014 Genetic diversity and population structure of *Rheum tanguticum* (Dahuang) in China. *Chin. Med.* **9**(1): 1–9 (see supplementary).
- Zhao, M., Zhi, H., Doust, A.N., Li, W., Wang, Y., Li, H., Jia, G., Wang, Y., Zhang, N., Diao, X.** 2013 Novel genomes and genome constitutions identified by GISH and 5S rDNA and knotted1 genomic sequences in the genus *Setaria*. *BMC Genomics* **14**(1): 244.

Supplementary materials are available from Journal Website