BIODIVERSITAS
Volume 26, Number 1, January 2025
Pages: 46-54

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d260105

Development and validation of novel microsatellite markers of a potentially invasive fern *Dicranopteris linearis* var. *linearis*

NURSHANARIAH BALKIS KAMARUDIN, MUSABBIHAH ABDUL AZIZ, AHMAD SOFIMAN OTHMAN, RUSLY ROSAZLINA♥

School of Biological Sciences, Universiti Sains Malaysia. 11800 Minden, Penang, Malaysia. Tel./fax.: +604-6536203, ♥email: rosazlinarusly@usm.my

Manuscript received: 3 July 2024. Revision accepted: 3 January 2025.

Abstract. *Kamarudin NB, Aziz MA, Othman AS, Rosazlina R. 2025. Development and validation of novel microsatellite markers of a potentially invasive fern* Dicranopteris linearis *var.* linearis. *Biodiversitas 26: 46-54. Dicranopteris linearis* is a forked fern native to Peninsular Malaysia. Recently, it has been recognized for its invasive potential in the ecosystem of Peninsular Malaysia due to its abundance, rapid spread, aggressive growth pattern, and competition for resources available to other native plants and crops. Despite its abundance and invasive potential, little is known about the potential domination mechanism and influence of genetic diversity on its establishment. Addressing this knowledge gap requires population genetic analysis, yet the fern has limited genomic information and molecular markers available. Hence, we aimed to develop novel microsatellite markers from *D. linearis* var. *linearis* using whole genome sequencing (WGS) data for preliminary population genetic analysis. Our analysis identified 879 simple sequence repeats (SSRs), predominantly trinucleotide repeats (AAC/GTT). Out of 462 primer pairs designed, 20 were randomly selected for validation, leading to eight markers verified as polymorphic across 30 tested individuals from the Batu Ferringhi population. These markers showed a moderate level of genetic variation (Na= 3.5, He= 0.4884) between individuals. Furthermore, these markers achieved successful amplification in six related species with transferability rates between 75% and 100%. The SSR markers developed in this study can be applied for future population-level studies to ascertain factors that may be driving diversity that enables the survival, adaption, and domination of the potentially invasive fern and its related species in the ecosystem of Peninsular Malaysia.

Keywords: Cross-amplification, *Dicranopteris linearis* var. *linearis*, forked fern, microsatellite markers

INTRODUCTION

Dicranopteris Bernh. is a genus of forked ferns in the family of Glecheniaceae (Filicopsida), with about 20 species distributed in the pantropical regions (PPG I 2016; Wei et al. 2022). Species in the genus can reproduce sexually via spores and vegetatively via clones (Yang et al. 2021). The genus is characterized by its pseudo-dichotomous branching pattern, producing a forked structure and the presence of buds on older branches, resulting in characteristics of indeterminate leaf growth (Marpaung and Susandarini 2021). Among the *Dicranopteris* species, *D. linearis* (Burm.f.) Underw. is the most common species, widely distributed in areas with humid climates, including the Pacific, Asia, Polynesia, Africa, and Europe (Russell et al. 1998; Baharuddin et al. 2021). It is recognized as a species complex with a high morphological variation and different chromosome numbers of n = 39, n = 40, n = 78, and n = 80 (Lima et al. 2021; Wei et al. 2022). About 13 varieties have been found in Southeast Asia, with at least seven varieties reported in Peninsular Malaysia, namely var. *alternans* (Mett.) Holtt., var. *altissima* (Holtt.) Holtt., var. *inaequalis* (Rosenst.) Holtt., var. *linearis*, var. *montana* (Holtt.) (Holtt.) Holtt., var. *subpectinata* (H.Christ) Holtt., and var. *subspeciosa* Holtt. (Go et al. 2012).

D. linearis is a sun-loving fern that grows abundantly in open and disturbed habitats such as road cuts, landslides, and degraded forest areas (Mai et al. 2019). It is actively studied for its phytochemical research and is considered a potential wild medicinal plant with anti-cancer properties (Baharuddin et al. 2021). Since most studies focus on phytochemical research, concerns about its widespread distribution and the impact of its dominance on other native plant species and crop plantations are often overlooked. In Peninsular Malaysia, the fern has been identified as potentially weedy due to its abundance in forest floors and disturbed sites (Zakaria and Akomolafe 2019). The domination of native ferns has also been reported to affect forest regeneration and soil edaphic properties negatively (Liyanage et al. 2021; Takeshige et al. 2023). In addition, it is on the list of noxious weeds that need to be eradicated due to the fierce competition for nutrients and water sources with crops (Ruzlan and Hamdani 2020).

The fern possessed characteristics such as indeterminate clonal growth type, fierce competition for available resources, the ability to thrive in nutrient-poor soils, and the production of an allelopathic effect to suppress the growth of other species (Russell et al. 1998; Kato-Noguchi 2015; Ruzlan and Hamdani 2020; Liyanage et al. 2021). These characteristics of ferns are commonly observed in invasive species. Therefore, there is concern that these characteristics could contribute to ferns becoming invasive in the future, as native plants are also capable of developing adaptive strategies to cope with current and future environmental changes that have led to their invasiveness (Zakaria and Akomolafe 2019). Knowledge of the genetic diversity of invasive species

is crucial for gaining insight into their invasion history and evolution, as the extent of genetic diversity largely determines their ability to adapt and evolve to the changing environment (Dlugosch and Parker 2008; Ward et al. 2008). In the case of *D. linearis*, despite its occurrence in the ecosystem of Peninsular Malaysia, little is known about the potential domination mechanism and the influence of genetic diversity on its establishment. Additionally, there are limited studies on the genetics of the fern population, particularly in Peninsular Malaysia. This is mainly due to the challenges associated with fern traits, which tend to have high chromosome numbers, resulting in larger genome sizes that are difficult to work with (Leitch 2012; Clark et al. 2016; Pelosi and Sessa 2021). Consequently, limited genomic information, reference genomes, and molecular markers are available for this species.

Previously, molecular markers such as isoenzyme and inter-simple sequence repeat (ISSR) have been used to evaluate the genetic diversity of the species and its related taxa (Russell et al. 1999; Marpaung and Susandarini 2022). However, these two markers were considered less informative than simple sequence repeat (SSR) markers in terms of detecting genetic variation at the population level. SSRs or microsatellites are short, repeated DNA sequences (1-6 bp) that are widely distributed in eukaryotic genomes (Karaca 2015). The markers are often used as molecular markers in population genetics due to the high level of polymorphism, codominant, locus-specific, and more reproducible (Vieira et al. 2016). Recently, the markers have been used in the studies of invasive populations due to their hypervariability, which is beneficial for tracing the introduction pathway and assessing their genetic variation (Wu et al. 2018; Hanjalić et al. 2021). Information on genetic variation can help practitioner managers evaluate invasive potential, analyze how genetic diversity is distributed throughout populations of the species, and predict its ability to spread in the ecosystem for effective planning of management controls (He et al. 2024). However, the development of SSR markers can be challenging, especially for non-model plant species, due to their high cost and need for more information on reference genomes. With the growing accessibility and affordability of Next-Generation Sequencing (NGS) technologies, it is now feasible to efficiently identify a large number of SSRs from non-model plants with little or no reference genome (Antunes et al. 2022). Whole genome sequencing (WGS) is one of the powerful approaches that NGS technologies offer for the large-scale discovery of SSR (Taheri et al. 2018). Hence, the aims of our study were (i) to develop a set of novel SSR markers from one of the *D. linearis* common varieties, *D. linearis* var. *linearis,* using WGS data and (ii) to validate the usability of the selected markers based on the PCR amplification, polymorphism assessment across 30 individuals of the fern species from Batu Ferringhi, and their transferability between related species. Given the limited availability of molecular markers for this species, the SSR markers developed in this study would be valuable molecular resources. Notably, the use of these SSR markers will be beneficial for future assessment of the genetic variation pattern of potentially invasive ferns at the population level, opening up exciting possibilities for predicting the possible route of introduction and primary means of its domination in the ecosystem Peninsular Malaysia.

MATERIALS AND METHODS

Plant materials and genomic DNA extraction

Fresh young leaves of *D. linearis* var. *linearis* were collected from Bukit Limau, Malaysia (5° 7' 19.4"N 100° 33' 10.7"E) and cleaned with 70% ethanol before drying with silica gels in a zip lock. The silica-dried samples of *D. linearis* var. *linearis* were stored at -20°C before DNA extraction. The total genomic DNA of the sample was extracted from approximately 80 mg of dried leaves sample using a modified CTAB procedure (Doyle and Doyle 1990) as described by Jasim et al. (2024).

The quality and the quantity of the genomic DNA sample were then assessed using NanoDrop1000 spectrophotometry (Thermo Scientific) and by electrophoresis on 0.8% agarose gel. Approximately 6 µg of genomic DNA pooled from six replicates of the same individual was sent to Macrogen Inc. (Seoul, Korea) for library preparation and genome sequencing.

Additionally, 30 individuals of *D. linearis* var. *linearis* and six other taxa namely *D. curranii*, *D. linearis* var. *alternans*, *D. linearis* var. *inaequalis*, *D. linearis* var. *subpectinata*, *D. linearis* var. *tetraphylla,* and *D. pubigera* were collected from Batu Ferringhi, Malaysia (5° 27' 39.8"N 100° 14' 33.2"E) for SSR markers validation and cross-transferability test, respectively. Next, to avoid sampling the same clone, two individuals of *D. linearis* var. *linearis* were collected at least 3 m apart. The same DNA extraction was performed as previously described, and the extracted DNA samples were diluted to approximately 20 ng/µL for SSR markers validation and cross-transferability test.

Whole genome sequencing

For whole genome sequencing, library preparation and paired-end sequencing were provided by Macrogen Inc. (Seoul, Korea) using an Illumina NovaSeq 6000 platform. The NGS library with 150-bp inserts was prepared using the TruSeq Nano DNA library prep kit (Illumina, USA) and the library was sequenced using paired-end reads following the manufacturers' specifications. The quality of raw data reads was analyzed using FastQC (Andrews et al. 2010). Reads from the raw data that contained adapters and low-quality bases (below Phred quality score Q20) were trimmed using Trimmomatic (Bolger et al. 2014). FastQC was also used to check the quality of Q20 (the sequencing error rate of 1%), Q30 (the sequencing error rate of 0.1%), and the GC content of the trimmed data. Trimmed reads were then, de novo assembled using the Platanus-allee version 2.2.2 (Kajitani et al. 2019) using default parameters to generate contigs.

SSR mining and primer design

Potential SSR markers were mined and identified among the de novo assembled contigs using the MicroSAtellite (MISA) tool (Beier et al. 2017). SSRs with motifs ranging from dinucleotides to hexanucleotides in size were identified. Mononucleotide repeats were not included. The

minimum of repeat units was set as follows: six repeats for dinucleotides and four for tri-, tetra-, penta-, and hexa-nucleotides (Gupta et al. 2020). Primer pairs were designed using Primer3 (Untergasser et al. 2012) with the following criteria: 100-500 bp amplicon sizes, 20 bp optimum primer length, 54-60°C melting temperature (T_m), and 50% optimum GC content. Compound SSRs were not considered for the primer design.

SSR markers validation and polymorphism

Twenty primer pairs were randomly selected from SSRs and synthesized at First BASE Laboratories, Apical Scientific Sdn. Bhd. (Selangor, Malaysia). These 20 primer pairs were used to amplify ten individuals of *D. linearis* var. *linearis* collected from the Batu Ferringhi population. Polymerase chain reactions (PCR) were performed in a 25 µL reaction volume containing 12.5 µL Taq II PCR Master Mix (2X), 2 µL (10 mg/mL) bovine serum albumin (BSA), 1.25 µL dimethyl sulfoxide (DMSO), 2 µL (20 ng) DNA template, 0.5 µL (0.2 µM) each primer, and 6.25 µL sterile distilled water. The PCR reaction was performed using Biometra TAdvance PCR thermocycler programmed at 95°C for 2 min, followed by 35 cycles of 95°C for 30 s, 50-60°C for 1 min, 72°C for 1 min, and a final extension of 72°C for 10 min followed by a holding temperature of 4°C.

All PCR products were first analyzed on 2% agarose gel, then detected on 6% non-denaturing polyacrylamide gel electrophoresis (PAGE) and viewed under UV transilluminator after staining with Redsafe™**.** Only those potentially polymorphic primers were evaluated for polymorphism across 30 individuals of *D. linearis* var. *linearis* using forward primers labeled with the fluorescent dyes FAM and HEX. This was done in a multiplex PCR of five primer pairs per group. All the PCR products were then sent to First BASE Laboratories for fragment analysis and sequenced using ABI 3730XL automated sequencer (Applied Biosystems).

Cross-species transferability

The transferability of the newly developed primers was also tested on the following six taxa, namely *D. curranii*, *D. linearis* var. *alternans*, *D. linearis* var. *inaequalis*, *D. linearis* var. *subpectinata*, *D. linearis* var. *tetraphylla*, and *D. pubigera*. The same DNA extraction and PCR conditions were performed as previously described to amplify the loci of these six taxa. The PCR products were viewed on 2% agarose gel stained with the Redsafe™, and if a clear band of the expected size was observed, the amplification was considered successful.

Data analysis

Peakscanner version 1.0 (Applied Biosystems) was used for DNA fragment genotyping using GeneScan-500 LIZ as the standard size. The number of alleles (*Na*), observed heterozygosity (*Ho*), expected heterozygosity (*He*), fixation index (*F*), test of deviation from Hardy-Weinberg equilibrium (*HWE*), and linkage disequilibrium of each locus was determined by using Arlequin version 3.1 (Excoffier et al. 2005). CERVUS version 3.0.3 (Kalinowski et al. 2007) was used to calculate the polymorphic information content (*PIC*) of each SSR marker. The occurrence of null alleles, allelic dropout, and possible genotyping errors were checked using MICROCHECKER (Van Oosterhout et al. 2004). If present, the frequency of null alleles (*NAF*) was estimated using the Brookfield 1 method.

RESULTS AND DISCUSSION

Whole genome sequencing

In this study, a total of 75,881,688 raw reads were generated from the whole genome sequencing of *D. linearis* var. *linearis*. After quality check and data filtering, a total of 68,182,786 high-quality clean reads were obtained with the Q20 (the sequencing error of 1%), Q30 (the sequencing error of 0.1%), and GC content percentages of 98.95%, 95.44%, and 41.76% respectively. After the de novo assembly, a total of 35,550 contigs were obtained. The filtered raw data obtained was deposited in the NCBI database (SRA accession: PRJNA1026717).

SSR mining and primer design

All 35,550 assembled contig sequences were used to mine out potential SSRs, and a total of 781 sequences containing 879 SSRs were identified. Of these, 66 sequences containing more than one SSR and five SSRs were present in a compound formation (Table 1). Among the 879 SSRs identified, trinucleotides were the most common repeat motif (578; 65.76%), followed by dinucleotides (151; 17.18%), hexanucleotides (121; 13.77%), tetranucleotides (15; 1.71%), and pentanucleotides (14; 1.59%) (Figure 1). Furthermore, AG/CT (73; 8.3%) was the common motif among the dinucleotide repeats, followed by AC/GT (53; 6.03%) and AT/AT (25; 2.84%) (Figure 2). Among the trinucleotide repeats, AAC/GTT (194; 22.07%), followed by AAG/CTT (173; 19.68%), ATC/ATG (126; 14.33%), AGC/CTG (39; 4.44%), AGG/CCT (31; 3.53%), AAT/ATT (8; 0.91%), ACC/GGT (5; 0.57%), and ACT/AGT (2; 0.23%) (Figure 2). Tetra-, penta-, and hexanucleotide repeats accounted for 17.06% of the total SSRs (Figure 2). With the five compound SSRs removed, 874 SSRs remain for primer design. Using the Primer3 tool with the defined parameters, 462 SSR primer pairs were successfully designed, while the 412 remaining loci were not considered due to the lack of flanking sequences. From these, we randomly selected 20 primer pairs for SSR marker validation.

SSR marker validation and polymorphism

Thirteen out of the 20 primer pairs successfully amplified the DNA samples with the expected PCR product size, while the remaining seven primer pairs were discarded because they failed to amplify. The 13 SSR sequences were deposited in the GenBank (Accession number: PP972062-PP972074). After the preliminary testing of polymorphism on the 6% non-denaturing PAGE, ten of the 13 primer pairs were selected for further screening of polymorphism in the 30 individuals of *D. linearis* var. *linearis*. Of the ten primer pairs sent for the fragment analysis, two were monomorphic, leaving only eight for the final analysis. The characteristics

of these eight polymorphic SSR markers are summarized in Table 2.

Cross-species transferability

The degree of cross-amplification success was determined by estimating the transferability rate of the eight loci to all six taxa. Our results showed the estimated transferability rate of the set of loci in the six taxa: 87.5% in *D. curranii*, 75% in *D. linearis* var. *alternans*, 87.5% in *D. linearis* var. *inaequalis*, 100% in *D. linearis* var. *subpectinata*, 75% in *D. linearis* var. *tetraphylla*, and 87.5% in *D. pubigera* (Table 3). This suggests that the use of these SSR loci should be useful for assessing genetic diversity in other related species.

Genetic characterization of SSR loci

A total of 28 alleles were detected in 30 individuals of *D. linearis* var. *linearis* based on the eight SSR loci. The number of alleles (*Na*) ranged from 2 (DL11) to 5 (DL2), with a mean value of 3.5 alleles per locus (Table 2). The observed (*Ho*) and expected (*He*) heterozygosity ranged from 0.1333 (DL11 and DL4) to 1.0000 (DL6) and from 0.1266 (DL11) to 0.7141 (DL2), with a mean value of 0.6417 and 0.4884, respectively (Table 2). Seven loci (DL1, DL2, DL5, DL6, DL7, DL11, and DL16) deviate significantly ($P<0.05$) from *HWE* in the Batu Ferringhi population except for one locus (DL4), while some pairs of the loci were in significant ($P<0.05$) linkage disequilibrium. The fixation index (*F*) ranged from -0.7865 (DL6) to 0.2564 (DL2), with a mean value of -0.2869 (Table 2). MICROCHECKER results suggested that no allelic dropout and possible genotyping error were found across all loci. The null allele frequency (*NAF*) estimated using Brookfield method 1 ranged from -0.2839 (DL6) to 0.0992 (DL2), with a mean value of -0.1040 (Table 2). The polymorphic information content (*PIC*) of each SSR loci ranged from 0.117 (DL11) to 0.654 (DL2), with a mean value of 0.418 (Table 2).

Discussion

Non-model plant species often lack genetic information and genomic resources, making it difficult to develop the molecular markers for the species (Unamba et al. 2015). This has become particularly problematic for invasive species since 55% of them still lack reference genomes (Matheson and McGaughran 2022). However, with the advent of NGS technologies, it is now possible to efficiently sequence and develop molecular markers of such species. The Illumina NovaSeq 6000 is one of the latest sequencing platforms that can generate large-scale sequencing data with high quality at a reasonable cost (Jones et al. 2023). Recently, it has been employed in the whole genome sequencing (WGS) of plants (Vats et al. 2023; Kim et al. 2024).

Table 1. Summary of SSR mining identified from *D. linearis* var. *linearis*

Parameter	Number
Total number of sequences	35,550
Total size of examined sequences (bp)	7,919,725
Total number of identified SSRs	879
Number of SSR-containing sequences	781
Number of sequences containing more than 1 SSR	66
Number of SSRs present in compound formation	5

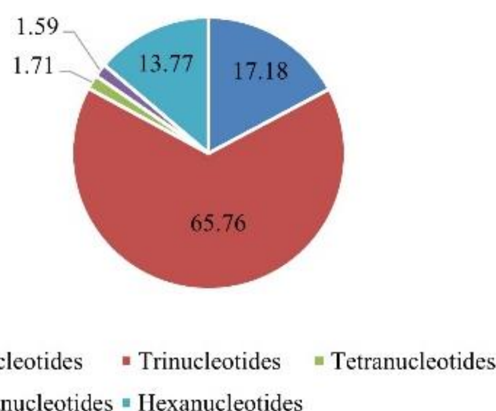

Figure 1. Distribution of SSR repeat types identified from *D. linearis* var. *linearis*

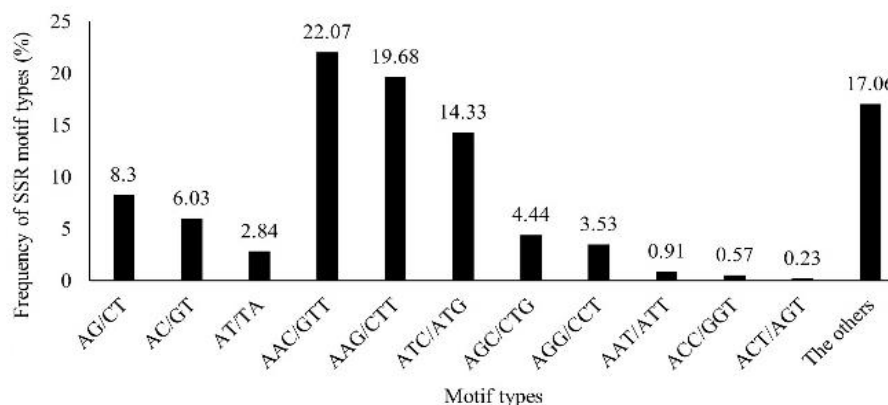

Figure 2. Frequency of SSR motif types identified from *D. linearis* var. *linearis*

Table 2. Characteristics of eight loci isolated from *D. linearis* var. *linearis* across 30 individuals obtained from Batu Ferringhi with locus name (ID), fluorescent-dyed forward (F) and reverse (R) primer sequences, SSR motif, annealing temperature (Ta), allele size range in base pairs (bp), GenBank accession number, number of alleles (*Na*), observed (*Ho*) and expected (*He*) heterozygosity, polymorphic information content (*PIC*), fixation index (*F*), and null allele frequency (*NAF*) from Brookfield method 1

ID	Primer sequence	SSR motif	Ta (°C)	Allele size range (bp)	GenBank	*Na*	*Ho*	*He*	*PIC*	*F*	*NAF*
DL1*	F: FAM-TCAAGGGTCACAAGGTTCAGG R: CCAAGGTGCGTTTTCTTGTGT	$(AAATCA)_4$	55	191-200	PP972062	4	0.9667	0.6254	0.538	-0.5603	-0.2178
DL2*	F: FAM-AAAAATCGCATCGGGCAAGG R: CCTCTGCATCAAGCTCCTCA	$(CAA)_5$	51	194-206	PP972063	5	0.5333	0.7141	0.654	0.2564	0.0992
DL4	F: FAM-CCACATATGTTGCAAGGGTTGA R: AGGCATCATCAGTGCACTTCA	$(TTCT)_4$	51	157-179	PP972065	3	0.1333	0.1283	0.121	-0.0404	-0.0064
DL5*	F: HEX-GTCCCGTTGTATCCAGGAGG R: GGAGGGGGATCTAGGTGAGA	$(GTT)_4$	55	236-253	PP972066	4	0.6333	0.5780	0.495	-0.0976	-0.0414
DL6*	F: HEX-GAGGCAGCGCAAAGAAGAAG R: CCGTGAAGGAGCTCGATACC	$(GAGTGA)_6$	51	171-183	PP972067	3	1.0000	0.5672	0.460	-0.7865	-0.2839
DL7*	F: HEX-TGGTGTGGTAGGACTTATGGC R: CCTCAGCCAACAAGCTCTGA	$(GTT)_5$	51	189-206	PP972068	3	0.8667	0.6266	0.535	-0.3924	-0.1550
DL11*	F: FAM-ATGGGCCTGTTCACCACTTT R: GTGAGGAGCATTGCAGATGC	$(ATC)_4$	55	203-231	PP972071	2	0.1333	0.1266	0.117	-0.0546	-0.0079
DL16*	F: FAM-TGAACAAGAAGGGAGATTCCA R: TCTGATTTGTACATCTTCATGCG	$(TCTCAC)_5$	55	156-168	PP972073	4	0.8667	0.5407	0.422	-0.6198	-0.2187
					Mean	3.5	0.6417	0.4884	0.418	-0.2869	-0.1040

Note: *Not in Hardy-Weinberg equilibrium ($P<0.05$)

Table 3. Transferability across six taxa of *Dicranopteris* species using the eight SSR markers validated in this study

Loci	*D. curranii*	*D. linearis* var. *alternans*	*D. linearis* var. *inaequalis*	*D. linearis* var. *subpectinata*	*D. linearis* var. *tetraphylla*	*D. pubigera*
DL1	-	+	-	+	-	-
DL2	+	+	+	+	+	+
DL4	+	-	+	+	-	+
DL5	+	+	+	+	+	+
DL6	+	+	+	+	+	+
DL7	+	-	+	+	+	+
DL11	+	+	+	+	+	+
DL16	+	+	+	+	+	+
Transferability rate (%)	87.5	75	87.5	100	75	87.5

Note: A positive (+) sign indicates a successful locus amplification and a negative (-) sign indicates failed locus amplification in the species

In this study, we used the Illumina NovaSeq 6000 platform to sequence the WGS data of *D. linearis* var. *linearis* for the development of SSR markers. To the best of our knowledge, this is the first report using this approach to develop SSR markers for the forked fern species. Until recently, the fern lineage received little attention in genomic research due to its genome complexity (Pelosi and Sessa 2021). This subsequently led to a lack of available genomic information on ferns, thus limiting our comparisons of WGS data with other fern species. After thorough quality control and data filtering, high-quality clean reads were obtained based on the high-quality scores of Q20 and Q30. The high-quality scores obtained indicate the high accuracy and reliability of the WGS data sequenced for de novo assembly (Xiong et al. 2021).

After de novo assembly of WGS data, a total of 879 SSRs were identified in *D. linearis* var. *linearis* with five different SSR repeat types, including di-, tri-, tetra-, penta-, and hexanucleotide. Since mononucleotides have a low polymorphism rate and can be the result of sequencing errors, they were not considered in the mining process (Zhao et al. 2015; Wang et al. 2018). Among these repeat types, dinucleotides and trinucleotides were the major repeat types, with trinucleotide being the most abundant in *D. linearis* var. *linearis* (Figure 1). Our result was consistent with a study by Li-ting et al. (2016), in which dinucleotides and trinucleotides were identified as the most common repeat types, with trinucleotides as the most abundant in the transcriptome assembly of another *Dicranopteris* species, *D. dichotoma*. The predominant motif in the fern was AG/CT for dinucleotide repeats and AAC/GTT for trinucleotide repeats (Figure 2). Victoria et al. (2011) reported that AG/CT was among the predominant motifs in the dinucleotide repeats of various plant species. Some studies also reported AAC/GTT as the predominant motif of trinucleotide repeats in other vascular plants, such as *Paeonia suffructicosa* and *Vigna radiata* (Gao et al. 2013; Wang et al. 2015). These findings suggest that these major repeat types and motifs are not unique to ferns but are common across different plant species. Since trinucleotides were the major repeat types with the most abundance in *D. linearis* var. *linearis*, their chances for amplification were expected to be higher. However, the major repeat types can still vary depending on the species. For instance, in another invasive species, *Medicago polymorpha*, mononucleotides were identified as the most abundant repeat motifs (Ren et al. 2023), which is inconsistent with our result. These differences may be due to the size of the dataset used for each species and the parameters set for SSR mining, which may affect the abundance of SSR repeat types found.

In the initial screening, only 13 primer pairs were successfully amplified with high-quality amplicons of the expected size. The seven remaining primer pairs could not be amplified, which may be due to the possible formation of a secondary structure in the long repeat motif length and replication slippage during the PCR amplification (Hosseinzadeh-Colagar et al. 2016; Bhattarai et al. 2021). After polymorphism screening, only eight loci were identified as polymorphic. The set of polymorphic SSR loci was used to verify its usability in the Batu Ferringhi population based on several genetic parameters.

The Hardy-Weinberg Equilibrium (*HWE*) test was used as a quality control measure to ensure that SSR loci accurately reflect genetic variation and provide reliable data for the population genetic analysis. Most loci in this study were found to deviate significantly from *HWE* in the Batu Ferringhi population ($P<0.05$). However, deviations from *HWE* are common in wild populations and have been reported in invasive species (Hernández-Espinosa et al. 2020; Kim et al. 2020; Sapkota et al. 2021; Jungová et al. 2023). One of the reasons for the significant deviation from *HWE* may be due to the presence of linkage disequilibrium detected in some loci pairs ($P<0.05$). In addition, an excess or deficiency of heterozygosity can also lead to deviation from *HWE* (Chew et al. 2021). In this study, seven loci (DL1, DL4, DL5, DL6, DL7, DL11, and DL16) were detected in heterozygote excess with negative fixation values, and only one locus (DL2) was detected in heterozygote deficiency with a positive value. The presence of null alleles could explain the heterozygosity deficiency detected in the locus. Our result suggested the possible presence of null alleles at the DL2 locus, with a moderate *NAF* (0.0992). However, no allelic dropout and possible genotyping error were detected across all loci. According to Mohd Rodzik et al. (2023), *NAF* with a frequency greater than or equal to 0.05 but less than 0.2 ($0.05 \leq NAF < 0.2$) were considered moderate frequencies, and therefore, all loci were retained and used in the present study. Since no allelic dropouts and genotyping errors were detected, the loci set was considered reliable for the genetic diversity analysis.

The informativeness of the SSR loci was evaluated using the polymorphic information content (*PIC*). It is used to demonstrate its usefulness in genetic diversity analysis. Zhong et al. (2023) classified SSR loci as highly polymorphic if the *PIC* value was greater than 0.5 ($PIC>0.5$), moderately polymorphic if it was between 0.25 to 0.5 ($0.25<PIC<0.5$), and as less polymorphic if it was less than 0.25 ($PIC<0.25$). In this study, three loci (DL1, DL2, and DL7) were highly polymorphic, three loci (DL5, DL6, and DL16) were moderately polymorphic, and only two loci (DL4 and DL11) were less polymorphic. The mean *PIC* (0.418) across all loci suggested that the reported loci set is moderately informative for assessing the genetic diversity of the species.

Furthermore, the polymorphic loci set was able to detect alleles ranging from two to five per locus. Compared to another study of the same species using isoenzyme loci (Russell et al. 1999), only two to three alleles per locus were detected. In addition, Li-ting et al. (2016) reported that only five polymorphic EST-SSR loci were detected in the transcriptome assembly of *D. dichotoma*. The higher number of alleles and polymorphic loci detected by the genomic SSR markers in this study can be explained by their codominant and hypervariable nature, which allows the detection of a higher number of alleles and mutations that alter the number of repeats (Jewell et al. 2006; Al-Faifi et al. 2016). This result demonstrates the ability of the genomic SSR markers to detect a higher number of alleles per locus and more polymorphic loci compared to previous tested codominant markers, therefore, making them better markers for assessing genetic variations in the species.

Additionally, the polymorphic loci set detected a moderate level of genetic variation among the individuals from the Batu Ferringhi (*He*= 0.4884). The mean of the observed heterozygosity (*Ho*) was higher than the mean of the expected heterozygosity (*He*), suggesting that the fern population has a higher degree of heterozygosity and thus genetic variation among the individuals increases. The result was further supported by the overall negative fixation values (*F*), suggesting that the fern population in Batu Ferringhi exhibits excessive heterozygosity. These results may indicate that the fern individuals undergo sexual reproduction, which could contribute to the detected genetic variation and potentially improve their ability to survive, adapt and dominate in the study area. In addition, the abundance of mature sporophytes with sporangia was observed throughout the study area, further indicating that sexual reproduction is not uncommon in *D. linearis* var. *linearis*. However, the result of our study contrasted with a study by Russel et al. (1999) using isoenzyme markers; a low level of genetic variation was detected in similar fern species in Hawaii, and vegetative growth was identified as the primary means of its domination. This contradiction may be explained by the differences in the environment, which can influence the mode of reproduction and its genetic diversity. Nevertheless, further studies are required to assess the genetic variation patterns at the population level and confirm the possible introduction route and domination mechanism of the potentially invasive fern in the ecosystem of Peninsular Malaysia.

One of the key advantages of SSR markers is the ability to cross-amplify, which is useful for identifying plant species with limited genetic information (Geetha and Siril 2022). Our result revealed a high level of transferability rate in *D. linearis* var. *subpectinata* (100%). However, the transferability rate of the set of loci decreased in the remaining taxa, ranging from 75% to 87.5%. The genomic SSRs are known to be more polymorphic, but their less conserved nature may contribute to lower transferability efficiency (Aiello et al. 2020). Nevertheless, the set of loci's overall transferability rate is still high, making it useful for evaluating related species.

In conclusion, the identification of 879 SSRs from the de novo assembly of *D. linearis* var. *linearis* contigs, using whole genome sequencing, is a significant step forward in our understanding of genetic diversity. The fact that trinucleotide repeats (AAC/GTT) are the most abundant motifs is a promising discovery. The set of eight polymorphic loci developed was able to detect a moderate level of genetic variation among the individuals, possibly resulting from the sexual reproduction of the fern. The high levels of transferability of the eight loci tested in six related taxa further underscore the potential of these SSR markers. With limited molecular markers available for this species, these SSR markers will be a valuable molecular resource for studying the genetic diversity of this potentially invasive fern and its related species. However, comprehensive population-level studies are urgently needed to fully exploit the potential of these markers and investigate possible factors that could drive diversity and improve the survival, adaptation and dominance of the potentially invasive fern in the ecosystem of Peninsular Malaysia.

ACKNOWLEDGEMENTS

The authors thank Assoc. Prof. Dr. Rahmad Zakaria (School of Biological Sciences, Universiti Sains Malaysia) and Dr. Akomolafe Gbenga Festus (Department of Plant Science and Biotechnology, Federal University of Lafia, Nigeria) for guidance on sampling and sample identification from the Batu Ferringhi and Bukit Limau, Malaysia. This project was funded by the Ministry of Higher Education Malaysia for Fundamental Research Grant Scheme (FRGS) under the grant number: FRGS/1/2021/WAB11/USM/02/3.

REFERENCES

Aiello D, Ferradini N, Torelli L, Volpi C, Lambalk J, Russi L, Albertini E. 2020. Evaluation of cross-species transferability of SSR markers in *Foeniculum vulgare*. Plants 9 (2): 175. DOI: 10.3390/plants9020175.

Al-Faifi SA, Migdadi HM, Algamdi SS, Khan MA, Ammar MH, Al-Obeed RS, Al-Thamra MI, El-Harty EH, Jakse J. 2016. Development, characterization and use of genomic SSR markers for assessment of genetic diversity in some Saudi date palm (*Phoenix dactylifera* L.) cultivars. Electronic J Biotechnol 21: 18-25. DOI: 10.1016/j.ejbt.2016.01.006.

Andrews S, Krueger F, Segonds-Pichon A, Biggins L, Krueger C, Wingett S. 2010. FastQC. A quality control tool for high throughput sequence. Data 370.

Antunes AM, Stival JGN, Targueta CP, de Campos Telles MP, Soares TN. 2022. A pipeline for the development of microsatellite markers using next generation sequencing data. Curr Genomics 23 (3): 175-181. DOI: 10.2174/1389202923666220428101350.

Baharuddin AA, Roosli RA, Zakaria ZA, Tohid SFM. 2021. *Dicranopteris linearis*: A potential medicinal plant with anticancer properties. Boletín Latinoam Y Del Caribe De Plantas Medicinales Y Aromáticas 20: 28-37. DOI: 10.37360/blackcpma.21.20.1.2.

Beier S, Thiel T, Münch T, Scholz U, Mascher M. 2017. MISA-web: A web server for microsatellite prediction. Bioinformatics 33 (16): 2583-2585. DOI: 10.1093/bioinformatics/btx198.

Bhattarai G, Shi A, Kandel DR, Solís-Gracia N, da Silva JA, Avila CA. 2021. Genome-wide simple sequence repeats (SSR) markers discovered from whole-genome sequence comparisons of multiple spinach accessions. Sci Rep 11: 9999. DOI: 10.1038/s41598-021-89473-0.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30 (15): 2114-2120. DOI: 10.1093/bioinformatics/btu170.

Chew PC, Christianus A, Zudaidy JM, Ina-Salwany MY, Chong CM, Tan SG. 2021. Microsatellite characterization of Malaysian mahseer (*Tor* spp.) for improvement of broodstock management and utilization. Animals 11 (9): 2633. DOI: 10.3390/ani11092633.

Clark J, Hidalgo O, Pellicer J, Liu H, Marquardt J, Robert Y, Christenhusz M, Zhang S, Gibby M, Leitch IJ. 2016. Genome evolution of ferns: Evidence for relative stasis of genome size across the fern phylogeny. New Phytologist 210 (3): 1072-1082. DOI: 10.1111/nph.13833.

Dlugosch KM, Parker IM. 2008. Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. Mol Ecol 17 (1): 431-449. DOI: 10.1111/j.1365-294X.2007.03538.x.

Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. Focus 12 (1): 13-15.

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): An integrated software package for population genetics data analysis. Evol Bioinformatics 1: 117693430500100000. DOI: 10.1177/117693430500100003.

Gao Z, Wu J, Liu Z, Wang L, Ren H, Shu Q. 2013. Rapid microsatellite development for tree peony and its implications. BMC Genomics 14: 886. DOI: 10.1186/1471-2164-14-886.

Geetha CM, Siril EA. 2022. Cross-species transferability of genomic SSR markers and genetic diversity among *Asparagus racemosus* Willd. accessions. Plant Gene 31: 100361. DOI: 10.1016/j.plgene.2022.100361.

Go R, Yen CL, Maideen H, Omar H, Mustafa M, Yusof UK. 2012. The distribution of the ferns Gleicheniaceae in Peninsular Malaysia. Acta Biologica Malaysiana 1 (1): 18-25. DOI: 10.7593/abm/1.1.18.

Gupta YM, Tanasarnpaiboon S, Buddhachat K, Peyachoknagul S, Inthim P, Homchan S. 2020. Development of microsatellite markers for the house cricket, *Acheta domesticus* (Orthoptera: Gryllidae). Biodiversitas 21 (9): 4094-4099. DOI: 10.13057/biodiv/d210921.

Hanjalić J, Ušanović L, Lasić L, Kalajdžić A, Stroil BK, Ahmić A, Pojskić N. 2021. Applicability of microsatellite loci designed for cosmopolitan species for the investigation of endemic species: A case study of *Silene sendtneri* Boiss. Glasnik Šumarskog Fakulteta Univerziteta u Banjoj Luci 31: 55-63. DOI: 10.7251/GSF2131055H.

He ZZ, Stotz GC, Liu X, Liu JJ, Wang YG, Yang J, Li LF, Zhang WJ, Nan P, Song ZP. 2024. A global synthesis of the patterns of genetic diversity in endangered and invasive plants. Biol Conserv 291: 110473. DOI: 10.1016/j.biocon.2024.110473.

Hernández-Espinosa R, González-Astorga J, Espinosa de los Monteros A, Cabrera-Toledo D, Gallego-Fernández JB. 2020. Transferability of microsatellite markers developed in *Oenothera* spp. to the invasive species *Oenothera drummondii* Hook. (Onagraceae). Diversity 12 (10): 387. DOI: 10.3390/d12100387.

Hosseinzadeh-Colagar A, Haghighatnia MJ, Amiri Z, Mohadjerani M, Tafrihi M. 2016. Microsatellite (SSR) amplification by PCR usually led to polymorphic bands: Evidence which shows replication slippage occurs in extend or nascent DNA strands. Mol Biol Res Commun 5 (3): 167-174.

Jasim JH, Othman AS, Nordin FA, Talkah NSM. 2024. High-quality genomic DNA extraction methods of Yellow *Spathoglottis* Blume complex for next-generation sequencing. Biodiversitas 25 (2): 654-663. DOI: 10.13057/biodiv/d250224.

Jewell E, Robinson A, Savage D, Erwin T, Love CG, Lim GAC, Li X, Batley J, Spangenberg GC, Edwards D. 2006. SSRPrimer and SSR taxonomy tree: Biome SSR discovery. Nucleic Acids Res 34 (2): 656-659. DOI: 10.1093/nar/gkl083.

Jones A, Stanley D, Ferguson S, Schwessinger B, Borevitz J, Warthmann N. 2023. Cost-conscious generation of multiplexed short-read DNA libraries for whole-genome sequencing. PLoS One 18 (1): e0280004. DOI: 10.1371/journal.pone.0280004.

Jungová M, Jurasová VM, Čepková PH, Svobodová LL, Svoboda P, Hejcman M. 2023. Origin and genetic variability of populations of the invasive plant *Rumex alpinus* L. in the Giant (Krkonoše) Mountains. Ecol Evol 13 (6): e10145. DOI: 10.1002/ece3.10145.

Kajitani R, Yoshimura D, Okuno M, Minakuchi Y, Kagoshima H, Fujiyama A, Kubokawa K, Kohara Y, Toyoda A, Itoh T. 2019. Platanus-allee is a de novo haplotype assembler enabling a comprehensive access to divergent heterozygous regions. Nat Commun 10 (1): 1702. DOI: 10.1038/s41467-019-09575-2. 274.

Kalinowski ST, Taper ML, Marshall TC. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol Ecol 16 (5): 1099-1106. DOI: 10.1111/j.1365-294X.2007.03089.x.

Karaca M. 2015. Microsatellites for genetic and taxonomic research on thyme (*Thymus* L.). Turkish J Biol 39: 147-159. DOI: 10.3906/biy-1406-20.

Kato-Noguchi H. 2015. Involvement of allelopathy in the formation of monospecific colonies of ferns. Nat Prod Commun 10 (5): 811-814. DOI: 10.1177/1934578X1501000526.

Kim J, Lim J, Kim M, Lee YK. 2024. Whole-genome sequencing of 13 Arctic plants and draft genomes of *Oxyria digyna* and *Cochlearia groenlandica*. Sci Data 11: 793. DOI: 10.1038/s41597-024-03569-6.

Kim S, Kim H, Cho M, Kim C, Kim S. 2020. Development and characterization of 17 microsatellite markers for *Sonchus oleraceus*. Appl Plant Sci 8 (3): e11329. DOI: 10.1002/aps3.11329.

Leitch IJ. 2012. Plant Genome Diversity Volume 2: Physical Structure, Behaviour and Evolution of Plant Genomes. In: Barker MS. Karyotype and genome evolution in pteridophytes. Springer Science and Business Media, Vienna.

Li-ting LIU, Qiang WEN, Xiao-chun H, Qi-jing LIU. 2016. De novo sequencing and characterization of juvenile sporophyte transcriptome of a fern, *Dicranopteris dichotoma*. For Res 29 (4): 500-507.

Lima LV, Sousa S, Almeida TE, Salino A. 2021. State of the art in cytogenetics, insights into chromosome number evolution, and new C-value reports for the fern family Gleicheniaceae. Anais Da Academia Brasileira de Ciências 93: e20201881. DOI: 10.1590/0001-3765202120201881.

Liyanage R, Kariyawasam T, Wijewickrama T, Bandara J, Madawala S. 2021. Some edaphic properties in *Dicranopteris linearis* dominated landscape: A perspective for restoration success. Ceylon J Sci 50 (1): 17-18. DOI: 10.4038/cjs.v50i1.7843.

Mai NT, Nguyen NH, Tsubota T, Shinogi Y, Dultz S, Nguyen MN. 2019. Fern *Dicranopteris linearis*-derived biochars: Adjusting surface properties by direct processing of the silica phase. Colloids Surf A: Physicochem Eng Aspects 583: 123937. DOI: 10.1016/j.colsurfa.2019.123937.

Marpaung AA, Susandarini R. 2021. Variation on morphology and spore characters of *Dicranopteri*s and *Sticherus* (Gleicheniaceae) from Rokan Hulu District, Riau, Indonesia. Biodiversitas 22 (10): 4475-4486. DOI: 10.13057/biodiv/d221041.

Marpaung AA, Susandarini R. 2022. Genetic diversity of *Dicranopteris* and *Sticherus* from Rokan Hulu, Riau based on ISSR marker. J Trop Biodiver Biotechnol 7 (1): 66552. DOI: 10.22146/jtbb.66552.

Matheson P, McGaughran A. 2022. Genomic data is missing for many highly invasive species, restricting our preparedness for escalating incursion rates. Sci Rep 12: 13987. DOI: 0.1038/s41598-022-17937-y.

Mohd Rodzik FF, Sudirman NA, Teh CK, Ong AL, Heng HY, Yaakop S, Mohd-Assaad N, Ong-Abdullah M, Ata N, Amit S. 2023. Development of nuclear DNA markers for applications in genetic diversity study of oil palm-pollinating weevil populations. Insects 14 (2): 157. DOI: 10.3390/insects14020157.

Pelosi JA, Sessa EB. 2021. From genomes to populations: A meta-analysis and review of fern population genetics. Intl J Plant Sci 182 (5): 325-343. DOI: 10.1086/713442.

PPG I. 2016. A community-derived classification for extant lycophytes and ferns. J Syst Evol 54 (6): 563-603. DOI: 10.1111/jse.12229.

Ren H, Wei Z, Zhou B, Chen X, Gao Q, Zhang Z. 2023. Molecular marker development and genetic diversity exploration in *Medicago polymorpha*. PeerJ 11: e14698. DOI: 10.7717/peerj.14698.

Russell AE, Raich JW, Vitousek PM. 1998. The ecology of the climbing fern *Dicranopteris linearis* on windward Mauna Loa, Hawaii. J Ecol 86: 765-779. DOI: 10.1046/j.1365-2745.1998.8650765.x.

Russell AE, Ranker TA, Gemmill CEC, Farrar DR. 1999. Patterns of clonal diversity in *Dicranopteris linearis* on Mauna Loa, Hawaii 1. Biotropica 31: 449-459. DOI: 10.1111/j.1744-7429.1999.tb00387.x.

Ruzlan KABC, Hamdani MSA. 2020. Occurrence and management of resistant weed species in FGV plantation in Malaysia: A review. Plant Archives 20 (1): 3057-3062.

Sapkota S, Boggess SL, Trigiano RN, Klingeman WE, Hadziabdic D, Coyle DR, Olukolu BA, Kuster RD, Nowicki M. 2021. Microsatellite loci reveal genetic diversity of Asian callery pear (*Pyrus calleryana*) in the species native range and in the North American cultivars. Life 11 (6): 531. DOI: 10.3390/life11060531.

Taheri S, Lee AT, Yusop MR, Hanafi MM, Sahebi M, Azizi P, Shamshiri RR. 2018. Mining and development of novel SSR markers using next generation sequencing (NGS) data in plants. Molecules 23 (2): 399. DOI: 10.3390/molecules23020399.

Takeshige R, Imai N, Aoyagi R, Sawada Y, Ong R, Kitayama K. 2023. Influences of fern and vine coverage on the above-ground biomass recovery in a Bornean logged-over degraded secondary forest. J For Res 28: 260-270. DOI: 10.1080/13416979.2023.2187682.

Unamba CIN, Nag A, Sharma RK. 2015. Next generation sequencing technologies: The doorway to the unexplored genomics of non-model plants. Front Plant Sci 6: 164495. DOI: 10.3389/fpls.2015.01074.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3-new capabilities and interfaces. Nucleic Acids Res 40 (15): e115-e115. DOI: 10.1093/nar/gks596.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4 (3): 535-538. DOI: 10.1111/j.1471-8286.2004.00684.x.

Vats S, Kumar V, Mandlik R, Patil G, Sonah H, Roy J, Sharma TR, Deshmukh R. 2023. Reference guided de novo genome assembly of transformation pliable *Solanum lycopersicum* cv. Pusa Ruby. Genes 14 (3): 570. DOI: 10.3390/genes14030570.

Victoria FC, da Maia LC, de Oliveira AC. 2011. In silico comparative analysis of SSR markers in plants. BMC Plant Biol 11: 15. DOI: 10.1186/1471-2229-11-15.

Vieira MLC, Santini L, Diniz AL, Munhoz CF. 2016. Microsatellite markers: What they mean and why they are so useful. Genet Mol Biol 39 (3): 312-328. DOI: 10.1590/1678-4685-GMB-2016-0027.

Wang H, Lei Y, Yan L, Wan L, Cai Y, Yang Z, LJ, Zhang X, Xu C, Liao B. 2018. Development and validation of simple sequence repeat markers from *Arachis hypogaea* transcript sequences. Crop J 6 (2): 172-180. DOI: 10.1016/j.cj.2017.09.007.

Wang LX, Elbaidouri M, Abernathy B, Chen HL, Wang SH, Lee SH, Jackson SA, Cheng XZ. 2015. Distribution and analysis of SSR in mung bean (*Vigna radiata* L.) genome based on an SSR-enriched library. Mol Breed 35: 1-10. DOI: 10.1007/s11032-015-0259-8.

Ward SM, Gaskin JF, Wilson LM. 2008. Ecological genetics of plant invasion: What do we know?. Invasive Plant Sci Manag 1 (1): 98-109. DOI: 10.1614/IPSM-07-022.1.

Wei Z, Xia Z, Shu J, Shang H, Maxwell SJ, Chen L, Zhou X, Xi W, Adjie B, Yuan Q, Cao J, Yan Y. 2022. Phylogeny and taxonomy on cryptic species of forked ferns of Asia. Front Plant Sci 12: 748562. DOI: 10.3389/fpls.2021.748562.

Wu CA, Hakkenberg AD, Beauchamp VB. 2018. Characterization of polymorphic microsatellite loci for invasive wavyleaf basketgrass, *Oplismenus undulatifolius* (Poaceae). Appl Plant Sci 6 (1): e1016. DOI: 10.1002/aps3.1016.

Xiong Y, Lei X, Bai S, Xiong Y, Liu W, Wu W, Yu Q, Dong Z, Yang J, Ma X. 2021. Genomic survey sequencing, development and characterization of single-and multi-locus genomic SSR markers of *Elymus sibiricus* L. BMC Plant Biol 21: 3. DOI: 10.1186/s12870-020-02770-0.

Yang L, Huang Y, Lima LV, Sun Z, Liu M, Wang J, Liu N, Ren H. 2021. Corrigendum: Rethinking the ecosystem functions of *Dicranopteris*, a widespread genus of ferns. Front Plant Sci 11: 581513. DOI: 10.3389/fpls.2020.581513.

Zakaria R, Akomolafe GF. 2019. Taxonomic diversity of ferns of two recreational forests in Kedah, Malaysia. Malaysian J Sci 38 (3): 1-11 DOI: 10.22452/mjs.vol38no3.1.

Zhao YM, Zhou T, Li ZH, Zhao GF. 2015. Characterization of global transcriptome using illumina paired-end sequencing and development of EST-SSR markers in two species of *Gynostemma* (Cucurbitaceae). Molecules 20: 21214-21231. DOI: 10.3390/molecules201219758.

Zhong X, Xu M, Li T, Sun R. 2023. Development of EST-SSRs based on the transcriptome of *Castanopsis carlesii* and cross-species transferability in other *Castanopsis* species. PLoS One 18 (7): e0288999. DOI: 10.1371/journal.pone.0288999.9.