Development and characterization of 20 novel EST-SSR markers for *Pteroceltis tatarinowii*, a relict tree in China

Mengyuan Zhang1,2*, Yunyan Zhang1,2*, Guozheng Wang1, Jingbo Zhou1, Yongjing Tian1, Qifang Geng1,2* and Zhongsheng Wang1

**PREMISE:** *Pteroceltis tatarinowii* (Ulmaceae), the only species of the genus *Pteroceltis*, is an endangered tree in China. Here, novel expressed sequence tag–simple sequence repeat (EST-SSR) markers were developed to illuminate its genetic diversity for conservation and assisted breeding.

**METHODS AND RESULTS:** Based on Illumina transcriptome data from *P. tatarinowii*, a total of 70 EST-SSR markers were initially designed and tested. Forty-eight of 70 loci (68.6%) were successfully amplified, of which 20 were polymorphic. The number of alleles per locus ranged from two to six, and the levels of observed and expected heterozygosity ranged from 0.018 to 0.781 and from 0.023 to 0.702, respectively. Additionally, cross-amplification was successful for 17 loci in two related species, *Ulmus gaussenii* and *U. chenmoui*.

**CONCLUSIONS:** These new EST-SSR markers are valuable transcriptomic resources for *P. tatarinowii* and will facilitate population genetics and molecular breeding of this species and its relatives in Ulmaceae.

**KEY WORDS** endangered tree; microsatellite; *Pteroceltis tatarinowii*; transcriptome; transferability; Ulmaceae.

---

*Pteroceltis tatarinowii* Maxim. (Ulmaceae), the only species of *Pteroceltis* Maxim., is an endangered Tertiary relict deciduous tree endemic to China (Li et al., 2013). Its bark (phloem fiber) has been famously utilized as a raw material to make the traditional Chinese Xuan paper (Cao, 1993; Li et al., 2012). *Pteroceltis tatarinowii* can adapt well to high-calcium soil and barren drought conditions, which makes it a popular tree in ecological restoration (You et al., 2018; Zhang et al., 2019). However, in recent decades, the distribution and size of wild *P. tatarinowii* populations have dramatically decreased due to fragmentation of native habitats and overexploitation for Xuan paper. This species is categorized as a National Class III Key Protected Species in China (Zhang et al., 2019). Therefore, there is an urgent need to perform population genetic studies that will assist in the development of conservation strategies.

With the advent of high-throughput transcriptome sequencing technologies, expressed sequence tag–simple sequence repeat (EST-SSR) markers have been rapidly mined in recent years. Compared to anonymous genomic SSRs, EST-SSRs have many advantages such as linking to species’ functional traits, high polymorphism, more transfer across taxonomic boundaries, and less susceptibility to null alleles and homoplasy (Varshney et al., 2005; Ellis and Burke, 2007; Yoichi et al., 2016). Previous population genetic studies of *P. tatarinowii* have utilized 12 genomic SSRs composed of dinucleotide repeat and imperfect SSR types (Li et al., 2015; Fan et al., 2019). However, the lack of EST-SSR markers developed from a transcriptional data set and the insufficient SSR repeat motif types of *P. tatarinowii* restrict the investigation of population genetics and molecular breeding using abundant information sites. Moreover, EST-SSR markers are relatively easy and inexpensive to develop, and are more accessible to studies based on functional traits than other types of genomic SSR markers (Hinchliffe et al., 2011; Ukoskit et al., 2018). In this study, we developed novel polymorphic EST-SSR markers for *P. tatarinowii* from an Illumina transcriptome data set to elucidate its population genetic diversity and test the transferability of the studied loci in two phylogenetically related Ulmaceae species, *Ulmus gaussenii* W.C. Cheng and *U. chenmoui* W.C. Cheng.

**METHODS AND RESULTS**

Fresh young leaf tissue from one *P. tatarinowii* tree from Nanjing Botanical Garden, Memorial Sun Yat-sen in Jiangsu Province, China (Appendix 1), was sampled for transcriptome sequencing. TRIzol reagent (Invitrogen Life Technologies, Carlsbad,
### TABLE 1. Characteristics of 20 EST-SSR markers developed for *Pteroceltis tatarinowii*.

| Locus | Primer sequences (5′–3′) | Repeat motif | Allele size range (bp) | $T_a$ (°C) | Fluorescent dye | BLASTX top hit description | E-value | GenBank accession no. |
|-------|--------------------------|--------------|------------------------|------------|----------------|---------------------------|---------|----------------------|
| E2    | F: AAACGGTTTCGGTTTCGCTTG | (CCAATA)$_5$ | 218–248               | 58         | FAM            | Hypothetical protein [L484_008773] [Morus notabilis] | 2.E-11  | JZ980980             |
|       | R: TCTATTCTGAGCTGGGCTTG  |              |                        |            |                |                           |         |                      |
| E3    | F: AGGTTGATGTTGATCGAGGA  | (GATGAG)$_6$ | 115–127               | 58         | HEX            | Serine/threonine protein kinase [Trema orientale] | 4.E-13  | JZ980981             |
|       | R: ATCCCTAGACAAGAGAGAT   |              |                        |            |                |                           |         |                      |
| E4    | F: CCGAGGATCAGCTGGGAGAA | (AGACAA)$_5$ | 154–178               | 58         | TAMRA          | Splicing factor-like protein [Parasponia andersonii] | 2.E-69  | JZ980982             |
|       | R: ATCCCTCCTACGACAAATGC  |              |                        |            |                |                           |         |                      |
| E5    | F: GTATGACCGGAGAATCGTA   | (GGTTGC)$_5$ | 254–272               | 58         | ROX            | Splicing factor-like protein [Trema orientale] | 3.E-64  | JZ980983             |
|       | R: ATGGAGGGTTTCTGGCTTTTT |              |                        |            |                |                           |         |                      |
| E6    | F: TCATTTCCGGCAACACGAAC  | (CAATT)$_5$ | 248–272               | 58         | FAM            | Zinc finger protein [Trema orientale] | 7.E-31  | JZ980984             |
|       | R: GATAGCAGTGCTGGTAACCG  |              |                        |            |                |                           |         |                      |
| E15   | F: AGGATGTGTTCTGTGTCGCC  | (GGTA)$_6$  | 201–221               | 60         | TAMRA          | Glycyl transferase [Trema orientale] | 9.E-14  | JZ980985             |
|       | R: TCGGCGATGATGATGACGTT  |              |                        |            |                |                           |         |                      |
| E16   | F: AGGTTGTTTGTTGCCCTTG  | (CCTT)$_5$  | 245–257               | 60         | HEX            | No significant similarity found | —       | JZ980986             |
|       | R: TGGGCTTTCTCCACCGAGTA  |              |                        |            |                |                           |         |                      |
| E19   | F: ATCCAGGCTAATGTCTCTCT | (ATGT)$_5$  | 110–125               | 60         | TAMRA          | MYC/MYB transcription factor [Parasponia andersonii] | 6.E-50  | JZ980987             |
|       | R: CAGGAATTGGACACATTTCA  |              |                        |            |                |                           |         |                      |
| E21   | F: CACCGATTGGAAATAAATGTATTCA | (GAAA)$_5$ | 208–220               | 60         | FAM            | Mitogen-activated protein kinase [Trema orientale] | 1.E-21  | JZ980988             |
|       | R: GTTTTTCTCAAGCTCTCTCGT  |              |                        |            |                |                           |         |                      |
| E23   | F: CCAGTACATCCAGCAACACA  | (AGAA)$_5$  | 185–191               | 60         | HEX            | No significant similarity found | —       | JZ980989             |
|       | R: TGGCAAACTCTCCCCTGTGTCAGT | (ATGA)$_5$ | 112–132               | 60         | TAMRA          | Golgi SNAP receptor complex, subunit [Trema orientale] | 4.E-18  | JZ980990             |
| E24   | F: CCTTTCTTTTATGCCCTCCA | (TTGT)$_5$  | 196–202               | 60         | HEX            | Rab-GTRase-TRC domain-containing protein [Trema orientale] | 3.E-03  | JZ980991             |
|       | R: TTGGGCTGAGCTTTGCTCAT  |              |                        |            |                |                           |         |                      |
| E36   | F: TGGGCGAGGAAAATGTTGCC  | (ATGA)$_5$  | 215–233               | 60         | FAM            | Chromatin structure-remodeling complex protein BSH [Morus notabilis] | 3.E-63  | JZ980992             |
|       | R: GTTTGCTGACCTTTGACGAT  |              |                        |            |                |                           |         |                      |
| E42   | F: AAATGCGAGCTATGGAGTTT  | (AGA)$_5$   | 111–114               | 60         | HEX            | Hypothetical protein [Prudu_020836] [Pruinus dulcis] | 2.E-30  | JZ980993             |
|       | R: GATGCTATATGTGCCCTGTC  |              |                        |            |                |                           |         |                      |
| E44   | F: AAGCGAGACATGGGCTTG    | (ACC)$_5$   | 264–279               | 60         | TAMRA          | Chlorophyll A-B binding protein [Parasponia andersonii] | 2.E-24  | JZ980994             |
|       | R: CTCGCTAATCTGCTGCCGGTC |              |                        |            |                |                           |         |                      |
| E48   | F: AAGGGAGGAAGGAGGCGAT   | (GGA)$_5$   | 255–267               | 60         | TAMRA          | 43-kDa postsynaptic protein [Trema orientale] | 5.E-73  | JZ980995             |
|       | R: GATGACAGGGCAGCTCGAACAT | (GTG)$_5$ | 238–256               | 60         | ROX            | Plastid division protein PDV [Parasponia andersonii] | 8.E-57  | JZ980996             |
| E53   | F: ATCCAGGCACTCGAGAAACGS | (TCT)$_5$  | 243–279               | 55         | ROX            | DYW domain-containing protein [Trema orientale] | 2.E-54  | JZ980997             |
|       | R: ATGGGACATTGAGCGCTTG   |              |                        |            |                |                           |         |                      |
| E65   | F: AAATGACCCAGCAACAGAC  | (AAAG)$_5$  | 259–279               | 60         | Protein DA1-related 2 isoform X1 [Sesamum indicum] | 3.E-31  | JZ980998             |
|       | R: ATCCACAAACACACGCTAT  |              |                        |            |                |                           |         |                      |
| E69   | F: AAGATCGCCGCAACAGAGAC  | (CCT)$_5$   | 253–257               | 60         | ROX            | 43-kDa postsynaptic protein [Parasponia andersonii] | 4.E-53  | JZ980999             |
|       | R: ACCACATGGCTCGAAATCAC  |              |                        |            |                |                           |         |                      |

*Note: $T_a$ = annealing temperature. Fluorescent dye used to tag the 5′ of each forward primer.*
TABLE 2. Genetic diversity statistics for five populations of *Pteroceltis tatarinowii* and two related taxa based on 20 newly developed EST-SSR markers.*

| Locus | *Pteroceltis tatarinowii* | Ulmus gaussenii *(N = 10)* | Ulmus chenmoui *(N = 10)* |
|-------|----------------------------|-----------------------------|-----------------------------|
|       | LYS *(N = 32)*             | JX *(N = 21)*               | TMS *(N = 24)*              | YZJ *(N = 34)* | SD *(N = 36)* |       |
|       | A | H_0 | H_e | A | H_0 | H_e | A | H_0 | H_e | A | H_0 | H_e | A | H_0 | H_e |
| E2    | 6 | 0.844 | 0.789 | 3 | 0.571 | 0.550 | 4 | 0.667 | 0.576 | 4 | 0.647 | 0.520 | 5 | 0.528 | 0.593 |
| E3    | 3 | 0.438 | 0.398 | 3 | 0.571 | 0.543 | 3 | 0.583 | 0.624 | 2 | 0.382 | 0.344 | 3 | 0.500 | 0.457 |
| E4    | 5 | 0.469 | 0.511 | 3 | 0.268 | 0.346 | 4 | 0.625 | 0.601 | 4 | 0.353 | 0.364 | 3 | 0.583 | 0.481 |
| E5    | 4 | 0.219 | 0.556 | 4 | 0.048 | 0.529 | 4 | 0.208 | 0.605 | 3 | 0.235 | 0.211 | 4 | 0.444 | 0.639 |
| E6    | 2 | 0.188 | 0.219 | 3 | 0.143 | 0.135 | 4 | 0.375 | 0.438 | 2 | 0.412 | 0.327 | 1 | 0.000 | 0.000 |
| E9    | 3 | 0.375 | 0.456 | 2 | 0.619 | 0.482 | 3 | 0.625 | 0.555 | 2 | 0.029 | 0.029 | 3 | 0.444 | 0.439 |
| E15   | 4 | 0.719 | 0.667 | 4 | 0.714 | 0.659 | 3 | 0.750 | 0.625 | 4 | 0.706 | 0.625 | 4 | 0.917 | 0.674 |
| E16   | 3 | 0.125 | 0.250 | 2 | 0.000 | 0.091 | 3 | 0.125 | 0.227 | 2 | 0.206 | 0.375 | 3 | 0.417 | 0.333 |
| E19   | 3 | 0.563 | 0.471 | 3 | 0.571 | 0.490 | 3 | 0.292 | 0.317 | 3 | 0.471 | 0.514 | 4 | 0.639 | 0.522 |
| E21   | 2 | 0.125 | 0.264 | 2 | 0.095 | 0.091 | 2 | 0.083 | 0.080 | 2 | 0.500 | 0.465 | 2 | 0.222 | 0.239 |
| E23   | 2 | 0.281 | 0.242 | 3 | 0.476 | 0.421 | 2 | 0.542 | 0.457 | 2 | 0.029 | 0.029 | 1 | 0.000 | 0.000 |
| E24   | 2 | 0.469 | 0.460 | 5 | 0.524 | 0.689 | 2 | 0.458 | 0.430 | 4 | 0.324 | 0.590 | 3 | 0.278 | 0.509 |
| E36   | 3 | 0.156 | 0.200 | 2 | 0.238 | 0.217 | 3 | 0.333 | 0.288 | 2 | 0.364 | 0.397 | 3 | 0.083 | 0.081 |
| E38   | 3 | 0.031 | 0.061 | 1 | 0.000 | 0.000 | 1 | 0.000 | 0.000 | 2 | 0.029 | 0.029 | 2 | 0.028 | 0.027 |
| E42   | 3 | 0.563 | 0.510 | 4 | 0.667 | 0.687 | 3 | 0.375 | 0.555 | 4 | 0.412 | 0.386 | 5 | 0.500 | 0.482 |
| E44   | 6 | 0.906 | 0.744 | 7 | 0.762 | 0.688 | 5 | 0.750 | 0.670 | 5 | 0.765 | 0.681 | 7 | 0.722 | 0.728 |
| E48   | 2 | 0.625 | 0.498 | 3 | 0.429 | 0.441 | 3 | 0.500 | 0.469 | 3 | 0.706 | 0.501 | 2 | 0.833 | 0.500 |
| E53   | 3 | 0.250 | 0.271 | 3 | 0.143 | 0.357 | 3 | 0.125 | 0.192 | 2 | 0.412 | 0.327 | 2 | 0.056 | 0.105 |
| E68   | 2 | 0.531 | 0.488 | 2 | 0.429 | 0.482 | 2 | 0.333 | 0.278 | 2 | 0.441 | 0.489 | 2 | 0.500 | 0.500 |
| E70   | 2 | 0.344 | 0.390 | 2 | 0.714 | 0.482 | 3 | 0.375 | 0.430 | 2 | 0.265 | 0.271 | 2 | 0.472 | 0.500 |
| Average | 3.15 | 0.411 | 0.422 | 3.10 | 0.400 | 0.416 | 3.00 | 0.396 | 0.421 | 2.80 | 0.384 | 0.374 | 3.05 | 0.408 | 0.391 |

Note: A = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; N = number of individuals sampled.

*Significant departure from Hardy–Weinberg equilibrium at P < 0.05.
temperature (Table 1). After amplification, we analyzed the multiple fluorescent dyes on an ABI 3730XL DNA Analyzer (Applied Biosystems, Foster City, California, USA) with GeneScan 500 LIZ Size Standard as an internal reference. GeneMarker version 2.2.0 (SoftGenetics, State College, Pennsylvania, USA) was then used to score the electrophoresis peaks and identify polymorphisms. Of the 48 candidate EST-SSR markers, 20 (41.7%) exhibited polymorphism in *P. tatarinowii*. All 20 EST-SSR sequences were deposited in GenBank (Table 1) and were used for cross-amplification tests. Furthermore, the corresponding sequences of these 20 EST-SSRs were BLASTed against the GenBank nonredundant database using BLASTX (Altschul et al., 1997) (Table 1). The characteristics of the 28 monomorphic EST-SSR markers are provided in Appendix 2.

Genetic diversity parameters (number of alleles and levels of expected and observed heterozygosity) were calculated in the five wild populations of *P. tatarinowii* using GenAlEx 6.5 (Peakall and Smouse, 2012). Significant deviation from Hardy–Weinberg equilibrium for each population and linkage disequilibrium for each primer pair were examined using GENEPOP version 4.2 (Rousset, 2008) using a Bonferroni correction. For each population of *P. tatarinowii*, levels of observed and expected heterozygosity varied from 0.000 to 0.917 (mean = 0.402) and from 0.000 to 0.789 (mean = 0.405), respectively (Table 2). The Hardy–Weinberg equilibrium test indicated that three primer pairs (E3 in the SD population, E4 in the JX and YZJ populations, and E5 in the LYS and TMS populations) deviated significantly from expectations after applying a Bonferroni correction (*P < 0.05; Table 2*), which might be caused by the Wahlund effect. No significant linkage disequilibrium was observed for any pair of loci after applying a Bonferroni correction.

Cross-species amplification tests of the 20 loci in *U. gaussenii* and *U. chenmou* followed the PCR procedures mentioned above. PCR products were detected using 2% agarose gels, and amplification was considered successful when one clear band was visible in the expected size range. We further analyzed the successful amplification products on an ABI 3730XL DNA Analyzer (Applied Biosystems) with GeneScan 500 LIZ Size Standard as an internal reference. GeneMarker version 2.2.0 (SoftGenetics) was then employed to score the electrophoresis peaks and identify polymorphisms. Overall, all loci were successfully amplified and exhibited polymorphisms, except for locus E36 for *U. chenmou* and loci E16 and E53 for *U. gaussenii* (Table 2).

**CONCLUSIONS**

Using high-throughput RNA sequencing data, we developed 20 polymorphic EST-SSR markers for *P. tatarinowii*. Most of these markers showed high transferability in two related species, suggesting that they may contribute to the population genetics and molecular breeding of other Ulmaceae species.

**ACKNOWLEDGMENTS**

The authors thank Jie Yang and Xin Zhuang for their assistance in collecting plant materials. This study was supported by the National Natural Science Foundation of China (grant no. 30970512).
DATA AVAILABILITY

Raw reads have been deposited into the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA accession number: SRR10158849). Primer sequences have been deposited to NCBI’s GenBank database; accession numbers are listed in Table 1.

LITERATURE CITED

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids* 25: 3389–3444.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: A flexible trimer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.

Cao, T. S. 1993. Xuan paper of China. China Light Industry Press, Beijing, China.

Ellis, J. R., and J. M. Burke. 2007. EST-SSRs as a resource for population genetic analyses. *Heredity* 99: 125–132.

Fan, J. J., X. P. Zhang, K. Liu, H. J. Liu, L. Zhang, X. P. Wang, and X. H. Li. 2019. The population genetic diversity and pattern of *Pteroceltis tatarinowii*, a relic tree endemic to China, inferred from SSR markers. *Nordic Journal of Botany* 2019: e01922.

Grabherr, G. M., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644–652.

Hinchliffe, D. J., R. B. Turley, M. Naoumkina, H. J. Kim, Y. Tang, K. M. Yeater, P. Li, and D. D. Fang. 2011. A combined functional and structural genomics approach identified an EST-SSR marker with complete linkage to the Ligon lintless-2 genetic locus in cotton (*Gossypium hirsutum* L.). *BMC Genomics* 12: 445.

Li, X. H., J. W. Shao, C. Lu, X. P. Zhang, and Y. X. Qiu. 2012. Chloroplast phylogeography of a temperate tree *Pteroceltis tatarinowii* (Ulmaceae) in China. *Journal of Systematics and Evolution* 50: 325–333.

Li, X. H., H. Zhang, D. Y. Wang, L. Zhang, J. W. Shao, and X. P. Zhang. 2013. The population genetic structure of endemic plant *Pteroceltis tatarinowii* by ISSR markers. *Acta Ecologica Sinica* 33: 4892–4901.

Li, X. H., X. P. Zhang, K. Liu, H. J. Liu, and J. W. Shao. 2015. Efficient development of polymorphic microsatellite loci for *Pteroceltis tatarinowii* (Ulmaceae). *Genetic and Molecular Research* 14: 16444–16449.

Pealkki, R., and P. E. Smouse. 2012. GenALEX 6.5: Genetic analysis in Excel. Population genetic software for teaching and research. An update. *Bioinformatics* 28: 2537–2539.

Pertea, G., X. Q. Huang, F. Liang, V. Antonesc, R. Sultana, S. Karamycheva, Y. D. Lee, et al. 2003. TiGR Gene Indices clustering tools (TGICL): A software system for fast clustering of large EST datasets. *Bioinformatics* 9: 651–652.

Rousset, F. 2008. GENEPOP’007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.

Rozen, S., and H. Skaltsky. 1999. Primer3 on the WWW for general users and for biologist programmers. In S. Misener and S. A. Krawetz [eds.], Methods in molecular biology, vol. 132: Bioinformatics: Methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

Thiel, T., W. Michalek, R. K. Varshney, and A. Graner. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 106: 411–422.

Ukoskit, K., G. Posudsavang, N. Pongsiripat, P. Chatwachirawong, P. Klomsaard, P. Poopimaj, and S. Tragoonrung. 2018. Detection and validation of EST-SSR markers associated with sugar-related traits in sugarcane using linkage and association mapping. *Genomics* 111: 1–9.

Varshney, R. K., A. Graner, and M. E. Sorrells. 2005. Genic microsatellite markers in plants: Features and applications. *Trends in Biotechnology* 23: 48–55.

Yoichi, W., S. Sakaguchi, S. Ueno, N. Tomaru, and K. Uehara. 2016. Development and characterization of EST-SSR markers for the genus *Rhododendron* section *Brachyclayx* (Ericaceae). *Plant Species Biology* 32: 455–459.

You, H. M., F. Kazue, and Q. Tang. 2018. Phytosociological study of *Pteroceltis tatarinowii* forest in the deciduous-forest zone of eastern China. In A. Greller, K. Fujiwara, and F. Pedrotti [eds.], Geographical changes in vegetation and plant functional types, 239–280. Springer, Cham, Switzerland.

Zhang, Y. Y., G. Z. Wang, J. H. Zhou, X. P. Zhou, P. F. Li, and Z. S. Wang. 2019. The first complete chloroplast genome sequence of *Pteroceltis tatarinowii* (Ulmaceae), an endangered tertiary relic tree endemic to China. *Mitochondrial DNA Part B* 4: 487–488.

APPENDIX 1. Locality and voucher information for populations of *Pteroceltis tatarinowii*, *Ulmus gaussenii*, and *U. chenmoui* used in this study.

| Species                  | Voucher specimens* | Population code | N  | Collection locality | Geographic coordinates |
|--------------------------|--------------------|-----------------|----|---------------------|------------------------|
| *P. tatarinowii* Maxim.  | WGG 20180421       | LYS             | 32 | Langya Mountain, Anhui, China | 32°17′N, 118°17′E       |
| *P. tatarinowii*         | ZMY 20180428       | JX              | 21 | Jingxian, Anhui, China | 30°39′N, 118°23′E       |
| *P. tatarinowii*         | ZMY 20180429       | TMS             | 24 | Tianmu Mountain, Zhejiang, China | 30°19′N, 119°26′E       |
| *P. tatarinowii*         | ZMY 20180430       | YJZ             | 34 | Swallow Promontory, Jiangsu, China | 32°09′N, 118°49′E       |
| *P. tatarinowii*         | ZMY 20180504       | SD              | 36 | Shidu, Beijing, China | 39°40′N, 115°31′E       |
| *U. gaussenii* W. C. Cheng | ZYX 20170102       | ZYY             | 10 | Langya Mountain, Anhui, China | 32°17′N, 118°17′E       |
| *U. chenmoui* W. C. Cheng | ZYX 20180421       | LYY             | 10 | Langya Mountain, Anhui, China | 32°17′N, 118°17′E       |

Note: N = number of individuals sampled.

*Vouchers were deposited in the Herbarium of Zhejiang University (HZU), Hangzhou, Zhejiang Province, China.*
APPENDIX 2. Characteristics of 28 monomorphic EST-SSR markers developed from the transcriptome of *Pteroceltis tatarinowii*.

| Locus | Primer sequences (5′–3′) | Repeat motif | \(T_a\) (°C) | Allele size range (bp) |
|-------|--------------------------|---------------|---------------|------------------------|
| E7    | F: TTCTCGAGAAGGCTCAGTTG  | (AACCTG)_5   | 58            | 118                    |
|       | R: CGAGGTTCCGTTGTTGTTTT  |               |               |                        |
| E8    | F: TTGATCTGATGTTGTTGAT   | (TGTTTG)_5   | 58            | 117                    |
|       | R: CAAGATTCGAAAAAGGCCG   |               |               |                        |
| E11   | F: TAACGGCCAGTAGAATCCA   | (ATATC)_6    | 60            | 275                    |
|       | R: GCCAGGTTATGATAGAGG     |               |               |                        |
| E14   | F: AGAGGTTGACTCTCCTGAT   | (TTCC)_5     | 60            | 250                    |
|       | R: AAAAGGAAGTGTGATGATATG  |               |               |                        |
| E22   | F: CAGTGAGCCAACAGAGTGGG  | (TTTC)_5     | 58            | 172                    |
|       | R: AAAGGCCCCAGTGGACAAAC  |               |               |                        |
| E25   | F: CCTTGAGTCTCCGAAAGTT   | (GGTT)_5     | 60            | 251                    |
|       | R: ATGGGCATGGATTGCCCTAAT |               |               |                        |
| E27   | F: CGAGGGCCAGAACAGTTAAG  | (ATTA)_5     | 55            | 176                    |
|       | R: CCAACAAAATCCTTAAGGATA  |               |               |                        |
| E28   | F: GGAGACGAAATTGTTGTTGG  | (CTT)_5      | 58            | 219                    |
|       | R: TGTCGAAATCTTTCCACCTC  |               |               |                        |
| E29   | F: GCTTGATGCTCCAAAGTT    | (AAAC)_5     | 60            | 251                    |
|       | R: CCAAAAGGAAATAAATAGGCA  |               |               |                        |
| E35   | F: TGATGGAGGCCAGATTTGATA | (TTAA)_5     | 60            | 251                    |
|       | R: TTCCTCAGTCACCTAGACA   |               |               |                        |
| E37   | F: TGTTGGTCTAGCTACCTTC   | (TTTG)_5     | 60            | 224                    |
|       | R: CCCCAGACATCTGCACACA   |               |               |                        |
| E39   | F: AAGGAATGATCCCTGCCTCA  | (CCG)_5      | 60            | 280                    |
|       | R: CTCCCTAACACCCAGGACCCA |               |               |                        |
| E41   | F: AATGGGCATTTGAAAGTGG   | (ATG)_5      | 60            | 271                    |
|       | R: CACCCTCTGCTCTTTAACAC |               |               |                        |
| E43   | F: AAGTGCCAAATGGGCTATG   | (CAG)_5      | 60            | 251                    |
|       | R: CCAGAGCTTGACTCTGTTGG  |               |               |                        |
| E45   | F: AATGGCGAATTGAGATGG    | (GAA)_5      | 60            | 241                    |
|       | R: TTTGGTTCTCTGTAATGGG   |               |               |                        |
| E47   | F: AACCCACACTCGCCTCACT   | (GGT)_5      | 60            | 219                    |
|       | R: CTCCCTCTCCACTTCAGCTC  |               |               |                        |
| E50   | F: AACAACACCCAACACAAAA   | (CTT)_5      | 55            | 208                    |
|       | R: TGGATGACTCTCCATCCAT    |               |               |                        |
| E51   | F: AAGGACGACAAGGTTTTTG   | (GGC)_5      | 60            | 206                    |
|       | R: TATTGGCTCGTAGACCGAGG  |               |               |                        |
| E52   | F: AATGAGGGGAAGGTTCTAG   | (CAA)_5      | 60            | 199                    |
|       | R: CCATCAACACCCAGTTACC   |               |               |                        |
| E54   | F: ACCCCCTCCCTTCTTCTGGT  | (TTA)_5      | 60            | 191                    |
|       | R: GACAAAGGTGCTCTCCCTG   |               |               |                        |
| E56   | F: AAATTCACGACGGAGTAAGA  | (GAA)_5      | 60            | 172                    |
|       | R: CTCCGCGCTTTCCCTCTCTCT |               |               |                        |
| E59   | F: AAACCCCAAGGAAATGTATTTAAA | (ATA)_6   | 55            | 162                    |
|       | R: TGTCCTTTGGGCTCTAGAT   |               |               |                        |
| E60   | F: AAGGAGGAGGAAATTGTTG   | (AAT)_5      | 60            | 155                    |
|       | R: CAAAATCTGCTTGAACGAGA  |               |               |                        |
| E61   | F: AACTCACTGACCCACACAC   | (CCT)_5      | 60            | 152                    |
|       | R: GAGGGCATGTTGAAAGAGAG  |               |               |                        |
| E62   | F: ACAAGGGCTGCTGAATGTC   | (ATG)_5      | 60            | 149                    |
|       | R: GAGACCTAGCTGACGCTCTCA |               |               |                        |
| E64   | F: AAGAGAGGAGGAGATGTTGG  | (GAG)_5      | 60            | 134                    |
|       | R: TGCTCGGACAGACTCCCATG  |               |               |                        |
| E65   | F: AATCAGTGATCGGAAACCG   | (TCT)_5      | 60            | 132                    |
|       | R: ATGGGGAACGTGTAACGCTTG |               |               |                        |
| E69   | F: AAAGAAGGCGGAGAAAAGTCA | (CTT)_5      | 60            | 112                    |
|       | R: ATCCCCCAAACACCGGATA   |               |               |                        |

Note: \(T_a\) = annealing temperature.