Development of microsatellite markers for a monotypic and globally endangered species, *Glyptostrobus pensilis* (Cupressaceae)

Gang-Tao Wang1,2,3, Zheng-Feng Wang1,2,4,5, Rui-Jiang Wang1,2,4, Dan Liang1,2, and Guo-Bin Jiang1,2,3

PREMISE OF THE STUDY: Microsatellite markers were developed to facilitate studies of genetic diversity and structure in *Glyptostrobus pensilis*, a critically endangered and monotypic conifer species.

METHODS AND RESULTS: Using restriction site–associated DNA sequencing (RAD-Seq), we developed 10 polymorphic and 27 monomorphic microsatellite markers. Polymorphism was characterized using 333 individuals from nine populations. The number of alleles per locus ranged from one to 14 at the population level. The levels of observed and unbiased expected heterozygosities varied from 0.058 to 0.844 and 0.219 to 0.583, respectively. Nine of these 10 polymorphic markers were successfully cross-amplified in *Taxodium distichum*, the species most closely related to *G. pensilis*.

CONCLUSIONS: These microsatellite markers can be used to reveal the genetic diversity in existing populations of *G. pensilis*, enabling its conservation and restoration.

KEY WORDS: Cupressaceae; endangered species; genetic diversity; genetic markers; *Glyptostrobus pensilis*; RAD-Seq.
### Table 1: Characteristics of 37 microsatellite markers developed in Glyptostrobus pensilis

| Locus | Primer sequences (5′–3′) | Repeat motif | Allele size range (bp) | GenBank accession no. |
|-------|--------------------------|--------------|------------------------|-----------------------|
| GP_19 | F: GCCAGCAGATTATACCCAG   | (GT),       | 314–338                | MH236836              |
|       | R: GGGCCACCAAGAGATGC     |              |                        |                       |
| GP_43 | F: AGGTGCTTGGTCAAAATCCC | (AC),       | 153–161                | MH236849              |
|       | R: GTCTAATTTGGAAGGGCACAAC|              |                        |                       |
| GP_46 | F: AAGGGTGCTCATCCTCACAG | (GAA),      | 152–156                | MH236852              |
|       | R: TCTAGCATTGAAACACATGCG|              |                        |                       |
| GP_57 | F: TATATATCTCAGGGGTCTCC | (GT),       | 207–212                | MH236855              |
|       | R: TGGCCAGGATAATTTGCGG   |              |                        |                       |
| GP_58 | F: AGAGGTAACTCCATCATGTC | (TC),       | 288–374                | MH236856              |
|       | R: GTCACTACCTATCGAAGATGCG|              |                        |                       |
| GP_71 | F: ACCTAGAAGGCAATAGGCC   | (AC),       | 199–201                | MH236858              |
|       | R: AGGAGAAGCATCTACACTAAGG|              |                        |                       |
| GP_75 | F: TGGTAGACTATGGTCAGATCT| (GA),       | 149–153                | MH236862              |
|       | R: TCAGCATTCTACCAAGATGTC|              |                        |                       |
| GP_80 | F: TGGTAGACCACTCAAGGCC  | (CA),       | 145–147                | MH236864              |
|       | R: AGAAGACAGCTACAGGCC   |              |                        |                       |
| GP_89 | F: ACACCTACATCATCTAGGTC  | (GT),       | 332–338                | MH236868              |
|       | R: ATGCACTATTGACATGGTTC |              |                        |                       |
| GP_94 | F: AGCATTTGGAACATCAAGGCC| (AG),       | 130–172                | MH236871              |
|       | R: ATGCTCTCACGTCAGGCCC   |              |                        |                       |
| GP_7  | F: TGGGCTGATAGATGGGCCC  | (GT),       | 332                    | MH236832              |
|       | R: TCTCTAGGCTGATGGGCCC  | (GT),       |                        |                       |
| GP_8  | F: ATCTCCATGCGTGACCACC  | (CTT),      | 224                    | MH236833              |
|       | R: AGTGGGTGTTACATGCTGTC |              |                        |                       |
| GP_9  | F: CGACTCTATGCGTGACCACC | (AT),       | 343                    | MH236834              |
|       | R: CATCCTAGATGGCAGATATGC| (AT),       |                        |                       |
| GP_17 | F: AATGGAGACAGGCAATAGG  | (GA),       | 190                    | MH236835              |
|       | R: GCCCTAGCCATTAGTACACC|              |                        |                       |
| GP_22 | F: AAGAGGGTTGGCAGTGTC   | (GGA),      | 156                    | MH236837              |
|       | R: GCCCTGCGTATAACATAGC  |              |                        |                       |
| GP_26 | F: ACATGTTTACAAATCTCAGTGC| (CT),      | 156                    | MH236839              |
|       | R: GAGGGAATTGCGCCTCCCT  |              |                        |                       |
| GP_28 | F: ACAACTCTTTGGAATGTGTC | (AT),       | 179                    | MH236840              |
|       | R: GGGTCTGAATATCAGATGTC |              |                        |                       |
| GP_29 | F: GATATGTGCAAAAAGGCC   | (AC),       | 370                    | MH236841              |
|       | R: TCTTCAAGAACAGATCAGAC | (AC),       |                        |                       |
| GP_31 | F: CGGTGATCCCTGCCTGCC   | (AC),       | 394                    | MH236842              |
|       | R: ACCAGCTCAGAAATTTGCC  |              |                        |                       |
| GP_32 | F: AGGTGATACAGGGTGAGGCC | (CT),       | 192                    | MH236843              |
|       | R: GGTGAAGAAGTGCAACTCGAC|              |                        |                       |
| GP_35 | F: GAACCTTTAATGGTGGAAGAGG| (GA),      | 251                    | MH236844              |
|       | R: GCCATGAAAGAAGAAATAGCC|              |                        |                       |
| GP_36 | F: TGCTTATCTCAGTCATCCCT | (AT),       | 207                    | MH236845              |
|       | R: CCCATTATGGAGACCGTGC  |              |                        |                       |
| GP_37 | F: TCTGCTCCTCCACAGGAAATAGGC| (CT), | 194                    | MH236846              |
|       | R: TGAACAACTTGGTGCTCTAAC|              |                        |                       |
| GP_39 | F: TGGAGAAGATATCTGATGGTCC| (GT),      | 153                    | MH236847              |
|       | R: TATGGAAATTTTGGCTCTACAG| (GT),     |                        |                       |
| GP_41 | F: AACTTGGAAAGGTAATGGG  | (GT),       | 175                    | MH236848              |
|       | R: ATCTTCTGCTTACCTAGCAC|              |                        |                       |
| GP_44 | F: TCAAGCCAGCTCAACCC    | (AC),       | 185                    | MH236850              |
|       | R: TCAAGACCTTACCTTGCGG  |              |                        |                       |
| GP_47 | F: AACTTGTTGCTTCACTGACC| (AC),       | 176                    | MH236853              |
|       | R: ATGTGGAAGAGTGGACACCCAG| (CT),    |                        |                       |
| GP_56 | F: TGGATCTTATGGGTGATGAC | (AC),       | 213                    | MH236854              |
|       | R: GCTTTGACATGCGAGGGTG  |              |                        |                       |
| GP_64 | F: TGCTTCACCTAGTGCGGAC  | (AC),       | 184                    | MH236857              |
|       | R: TGTGGAAGGTGTTGACCATGAG| (AC),     |                        |                       |
| GP_72 | F: CGGTGATGGGATCCATGTGC | (GT),       | 167                    | MH236859              |
|       | R: AAGGTGTTGCTGCGGCC    |              |                        |                       |
a modified cetyltrimethylammonium bromide (CTAB) method (Doyle, 1991).

Restriction site–associated DNA sequencing (RAD-Seq; Baird et al., 2008) was used to obtain partial genomic DNA sequences of *G. pensilis*. The microsatellites were then selected and developed based on these sequences. Two samples, one from the South China Botanical Garden and the other from Conghua District, Guangzhou Province, China, were used to construct the RAD-Seq libraries with the restriction enzyme EcoRI (Promega Corporation, Madison, Wisconsin, USA), followed by 150-bp paired-end sequencing using a HiSeq X Ten genetic analyzer (Illumina, San Diego, California, USA). From the two samples, 35,615,442 and 35,297,882 raw sequences were obtained, respectively. The raw sequence data are available in the National Center for Biotechnology Information (NCBI) Sequence Read Archive database (accession no. SRR7133729 and SRR7133728).

These raw sequences were subsequently combined and re-assembled by CAP3 (Huang and Madan, 1999), resulting in 3,285,999 contigs with a total length of 787,094,171 bp. The minimum and maximum lengths of the contigs were 80 bp and 2016 bp, respectively. The raw sequence data are available in the National Center for Biotechnology Information (NCBI) Sequence Read Archive database (accession no. SRR7133729 and SRR7133728).

Microsatellites with dinucleotide and trinucleotide motifs with minimum lengths of the contigs were 80 bp and 2016 bp, respectively. The raw sequence data are available in the National Center for Biotechnology Information (NCBI) Sequence Read Archive database (accession no. SRR7133729 and SRR7133728).

**Conclusions**

In this study, 10 polymorphic and 27 monomorphic microsatellite markers were developed for *G. pensilis*. The cross-amplification test indicated that nine of the 10 polymorphic markers can be successfully amplified in six *T. distichum* individuals. These markers will offer valuable tools for future investigations of genetic diversity and structure, level of gene flow, and conservation genetic studies in these two species.
TABLE 2. Genetic diversities of 10 polymorphic microsatellites developed for
in eight Chinese populations and one Vietnamese population.a

| Locus | A  | H_o | uH_o | F   | Adjusted P value |
|-------|----|-----|------|-----|-----------------|
| GP_19 | 4  | 0.833 | 0.773 | −0.087 | 0.526 |
| GP_43 | 1  | 0.000 | 0.000 | —   | —               |
| GP_46 | 2  | 0.833 | 0.530 | −0.667 | 0.242 |
| GP_57 | 3  | 0.750 | 0.679 | −0.125 | 0.571 |
| GP_58 | 6  | 0.800 | 0.844 | −0.059 | 0.863 |
| GP_71 | —  | —   | —    | —   | —               |
| GP_75 | 2  | 1.000 | 0.545 | −1.000 | 0.069 |
| GP_80 | 2  | 1.000 | 0.545 | −1.000 | 0.069 |
| GP_89 | 1  | 0.000 | 0.000 | —   | —               |
| GP_94 | 2  | 1.000 | 0.545 | −1.000 | 0.069 |
| Overall | — | — | — | — | — |

Note: A = number of alleles; F = fixation index; H_o = observed heterozygosity; uH_o = unbiased expected heterozygosity.
*aSee Appendix 1 for locality and voucher information.
*a indicates a significant deviation from Hardy–Weinberg equilibrium after Holm’s sequential Bonferroni correction (P < 0.05).

ACKNOWLEDGMENTS

The authors thank Z. Wang, X. J. Liu, and B. Chen for their field assistance in collecting samples. This study was supported by the Guangzhou Wild Life Conservation and Management Office (SYZFCG-[2017]032, Guangzhou Water Pine Germplasm Resource Conservation Program), Guandong Forestry Department Program for Rare and Endangered Plant Conservation, Botanical Gardens Conservation International (BGCI) G. pensilis Conservation Program, and the STS Program of the Chinese Academy of Sciences (KFJ-3W-No1-1).

AUTHOR CONTRIBUTIONS

R.J.W. conceived and designed the project. R.J.W., G.T.W., and D.L. carried out the field collection. G.T.W., Z.F.W., and G.B.J. carried out the laboratory procedures. G.T.W. and Z.F.W. analyzed the data. All authors read and approved the final version of the manuscript.

DATA ACCESSIBILITY

The microsatellites and raw sequences developed in this article have been deposited in the National Center for Biotechnology Information (NCBI). The GenBank accession numbers for the microsatellites are provided in Table 1, and the accession numbers for the raw sequences in the NCBI Sequence Read Archive are SRR7133729 and SRR7133728.

LITERATURE CITED

Averyanov, L. V., K. L. Phan, T. H. Nguyen, S. K. Nguyen, T. V. Nguyen, and T. D. Pham. 2009. Preliminary observation of native Glyptostrobus pensilis (Taxodiaceae) stands in Vietnam. TAIWANIA 54: 191–212.

Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3: e3376.
LePage, B. A. 2007. The taxonomy and biogeographic history of Glyptostrobus Endlicher (Cupressaceae). *Bulletin of the Peabody Museum of Natural History* 48: 359–426.

Li, F. G., and N. H. Xia. 2004. The geographical distribution and cause of threat to *Glyptostrobus pensilis* (Taxodiaceae). *Journal of Tropical and Subtropical Botany* 12: 13–20.

Nguyen, M. T., D. D. Vu, T. T. X. Bui, and M. D. Nguyen. 2013. Genetic variation and population structure in Chinese water pine (*Glyptostrobus pensilis*): A threatened species. *Indian Journal of Biotechnology* 12: 499–503.

Peakall, R., and P. Smouse. 2012. GenAlEx 6.5 (version 6.5): Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28: 2537–2539.

Wu, Z. Y., J. F. Liu, W. Hong, D. M. Pan, and S. Q. Zheng. 2011. Genetic diversity of natural and planted *Glyptostrobus pensilis* populations: A comparative study. *Chinese Journal of Applied Ecology* 22: 873–879.

**APPENDIX 1.** Locality information for the *Glyptostrobus pensilis* and *Taxodium distichum* samples used in this study.\(^a\)

| Species                     | Population code | N  | Collection locality                  | Voucher no.  |
|-----------------------------|-----------------|----|-------------------------------------|--------------|
| *Glyptostrobus pensilis*    |                 |    |                                     |              |
| (Staunton ex D. Don) K. Koch| JX              | 59 | Shangrao, Jiangxi Province, China    | IBSC799028   |
|                            |                 |    | Yingtan, Jiangxi Province, China    | IBSC799072   |
|                            |                 |    | Zixing, Hunan Province, China       | IBSC799035, 799034, 799082 |
|                            |                 |    | The Chinese University of Hong Kong, China | IBSC799085  |
|                            | HK              | 6  | Ningde, Fujian Province, China      | IBSC799064   |
|                            |                 |    | Sanming, Fujian Province, China     | IBSC799019   |
|                            |                 |    | Quanzhou, Fujian Province, China    | IBSC799016, 799075 |
|                            |                 |    | Fuzhou, Fujian Province, China      | IBSC799068   |
|                            | GD              | 74 | Guangzhou, Guangdong Province, China| IBSC799061, 799020, 799014, 799078, 799079, 799041, 799042, 799054, 799083, 799084 |
|                            |                 |    | Zuhuai, Guangdong Province, China   | IBSC799080, 799022 |
|                            |                 |    | Huaiji, Guangdong Province, China   | IBSC799056   |
|                            |                 |    | Meizhou, Guangdong Province, China  | IBSC799021, 799018, 799032 |
|                            |                 |    | Huizhou, Guangdong Province, China  | IBSC799066, 799057, 799031, 799030 |
|                            |                 |    | Tianteng, Guangxi Province, China   | IBSC799047   |
|                            |                 |    | Qinzhou, Guangxi Province, China    | IBSC799048   |
|                            |                 |    | Guilin, Guangxi Province, China     | IBSC799049   |
|                            |                 |    | Cangwu, Guangxi Province, China     | IBSC799051   |
|                            |                 |    | Luchuan, Guangxi Province, China    | IBSC799044   |
|                            |                 |    | Funing, Yunnan Province, China      | IBSC799046   |
|                            | ZJ              | 6  | Hangzhou, Zhejiang Province, China  | IBSC799050   |
|                            |                 |    | Shanghai, China                    | IBSC799059   |
|                            |                 |    | Wuhu, Hubei Province, China         | IBSC799053   |
|                            |                 |    | Xinyang, Henan Province, China      | IBSC799055   |
|                            |                 |    | Haiphong, Dak Lak Province, Vietnam | HN11357, 7111, 11946, 11950 |
| *Taxodium distichum* (L.) Rich. |                 |    | South China Botanical Garden, Guangzhou, Guangdong Province, China | (23°10′51″N, 113°21′08″E) |
|                            |                 |    | Dák Lák, Dak Lak Province, Vietnam  | IBSC799015   |

Note: N = number of individuals sampled.

\(^a\)All voucher specimens were deposited in the South China Botanical Garden Herbarium (IBSC), Guangzhou, China, or the Vietnam Academy of Science and Technology Herbarium (HN), Hanoi, Vietnam.