Development of microsatellite markers for the perennial plant *Tofieldia calyculata*

Tomáš Vlasta¹, Andrea Jarošová², Maria Šurinová¹ ², and Zuzana Münzbergová¹ ²

*Applications in Plant Sciences* 2020. Development of microsatellite markers for the perennial plant *Tofieldia calyculata*. Applications in Plant Sciences 8(3): e11327. doi:10.1002/aps3.11327

**PREMISE:** Polymorphic microsatellite markers were developed to study genetic diversity and genetic structure of populations of the locally endangered species *Tofieldia calyculata* (Tofieldiaceae).

**METHODS AND RESULTS:** Nineteen polymorphic microsatellite loci were developed using DNA-enriched libraries sequenced by illumina technology and were then used to genotype 101 individuals from five populations from Austria, Slovakia, Poland, and the Czech Republic. Of the markers tested, 68% were polymorphic in four of the five investigated populations, with one marker polymorphic in all populations. The number of alleles per locus in the populations ranged from one to 11. Levels of observed and expected heterozygosity ranged from 0.00 to 0.75 and from 0.00 to 0.84, respectively. Moreover, six of the 19 loci amplified when tested in the congeneric species *T. pusilla*.

**CONCLUSIONS:** The 19 newly developed microsatellite markers can be used to describe genetic diversity and population structure of populations of *T. calyculata*.

**KEY WORDS:** genotyping; microsatellites; *Tofieldia calyculata*; Tofieldiaceae.

*Tofieldia calyculata* (L.) Wahlenb. is a small perennial herb from the family Tofieldiaceae that is distributed mainly in mountain biotopes of the Alps and Western Carpathians. It also occurs in the lowlands of Central Europe, where it is restricted to calcareous fens and, less frequently, to wet slopes in calcareous pine woods. The species has disappeared from most of its historical localities in the lowlands as a result of drought, changes in vegetation structure, and ongoing succession. It is currently critically endangered in the Czech Republic (Grulich and Chobot, 2017), where, of more than 50 historical localities, only five localities remain (Kaplan et al., 2015). The species is also endangered or even extinct in other countries in Europe in which the species had previously been known to occur in lowland habitats. It is vulnerable in Poland and Ukraine, endangered in Lithuania, critically endangered in Croatia, and extinct in Latvia and Hungary (Gabrielová et al., 2011). Knowledge of genetic variability of the remaining lowland populations and their relationship to the alpine populations is necessary for better protection of the species.

**METHODS AND RESULTS**

**Microsatellite marker development**

A modified cetyltrimethylammonium bromide (CTAB) method (Lodhi et al., 1994) with the volume of solutions downscaled 10-fold was used to isolate the genomic DNA of 14 individuals of *T. calyculata* (two individuals in each of five populations collected across the whole distribution range, including both lowland and alpine populations; Appendix 1). The sequencing facility GenoScreen (Lille, France) prepared libraries based on previously published methodology (Malausa et al., 2011). The fragmented DNA was hybridized to eight probes (TG, TC, AAC, AAG, AGG, ACG, ACAT, and ACTC) to enrich the DNA library. The sequencing was performed on a MiSeq Nano 2 × 250 machine (Illumina), and sequencing yielded 4,998,842 reads. Of these reads, 735,657 contained microsatellites and 1235 candidate loci were identified (see Appendix S1). Of these loci, 83.24% contained dinucleotide repeats, 14.09% contained trinucleotide repeats, and 2.67% contained tetranucleotide repeats. All sequences are archived in the GenBank Sequence Read Archive (BioSample accessions SAMN11608780 and SAMN11608781). Primer3 implemented in the program QDD (Malausa et al., 2011) was used to design multiple primer pairs for every locus (for a total of 9431 primer pairs). Of these, one primer pair was selected for every locus (for a total of 1235 primer pairs). Of these selected primers, we used 50 primer pairs to identify polymorphism. Primers were synthesized (Sigma-Aldrich, St. Louis, Missouri, USA) with M13 tails preceding the 5′ end of the forward primer sequences following Schuelke (2000). Twelve individuals of *T. calyculata* were used to test amplification efficiency and polymorphism. DNA amplification
was performed in 5-μL volumes containing 2.5 μL of QIAGEN Multiplex PCR Master Mix (QIAGEN, Hilden, Germany), 0.5 μL of M13-tagged forward primer, 0.25 μL of species-specific reverse primer, 0.25 μL of fluorescently labeled (5′-FAM) M13 primer (10 μM each in initial volume for each primer), 20 ng of DNA dissolved in 0.5 μL of TE buffer, and 1 μL of H2O. The following thermocycler conditions were used following Schuelke (2000): an initial denaturation step at 95°C for 15 min; followed by 25 cycles of denaturation (95°C for 60 s), annealing (60°C for 60 s), and extension (72°C for 60 s); followed by 10 cycles of denaturation (95°C for 30 s), annealing (53°C for 45 s), and extension (72°C for 45 s); and a final extension at 72°C for 15 min. During the first 25 cycles, specific PCR products are produced, and in the following 10 cycles the fluorescent M13 tag is ligated to the M13 forward primer (Schuelke, 2000). All of the 50 primer pairs (100%) were successfully amplified, but only 44% of those loci (20) were polymorphic. We used the polymorphic primer pairs to detect variability in 20 individuals of *T. calyculata* from five populations (Appendix 1); these 20 individuals were tested in two multiplex reactions (described below). Based on this variability screening, 19 polymorphic primer pairs were chosen for further testing (Table 1); details for the remaining 31 primer pairs are provided in Appendix 2.

### Genotyping

We isolated DNA from 101 individuals of *T. calyculata* from five populations (three alpine and two lowland) and from six individuals of the congeneric species *T. pusilla* (Michx.) Pers. from two alpine populations (Appendix 1).

DNA amplification was performed in two multiplex reactions containing 2.5 μL of QIAGEN Multiplex PCR Master Mix and 20 ng of DNA dissolved in 0.5 μL of distilled water. For multiplex mix I, the PCR contained 1.94 μL of primer mix (10 μM each in initial volume) and 0.06 μL of H2O, for multiplex mix II, the PCR contained 1.875 μL of primer mix (10 μM each in initial volume) and 0.125 μL of H2O. Final volumes of primers are given in Table 1. The following thermocycler conditions were used: an initial denaturation step at 95°C for 10 min; followed by

### Table 1. Characteristics of 19 polymorphic loci designed for *Tofieldia calyculata*.

| Locus* | Primer sequences (5′–3′) | Repeat motif | Volume of forward primer (μL) | Allele size range (bp) | Multiplex | GenBank accession no. |
|--------|--------------------------|--------------|-------------------------------|------------------------|-----------|----------------------|
| Tof4   | F: GAGGGAGACGGCATACGACTC | (AGC)_5      | 0.05                          | 123–126                | I         | MN124996             |
|        | R: GGATCAAAGCACAGGCAACG   |              |                               |                        |           |                      |
| Tof7   | F: ATTCGCTGTCGCCGCGAGAG  | (AGC)_6      | 0.05                          | 124–136                | I         | MN124997             |
|        | R: GTCTCTAATGCGGCGCTG     |              |                               |                        |           |                      |
| Tof11  | F: CACAAGACTCATGACGAGC    | (AAG)_7      | 0.07                          | 131–212                | I         | MN124998             |
|        | R: GGTAAAGCTTAGGCTACCGA   |              |                               |                        |           |                      |
| Tof13  | F: TAGGCCAGGACACCCACATG   | (ACG)_4      | 0.07                          | 144–156                | I         | MN124999             |
|        | R: CCACAAACACTCTACACG     |              |                               |                        |           |                      |
| Tof19  | F: GTCGTAATTTAACCTCGCGG   | (AAG)_5      | 0.07                          | 189–198                | I         | MN125000             |
|        | R: GGGAACTCGGCTCTCAATGT   |              |                               |                        |           |                      |
| Tof22  | F: GCTTGCTGCCAGATAGAATTC | (ACG)_5      | 0.07                          | 184–193                | I         | MN125001             |
|        | R: ATACAGACCATGCGGTTCTT   |              |                               |                        |           |                      |
| Tof33  | F: TTAAGAGGAAGATGGAAGTGG  | (ACAT)_6     | 0.07                          | 233–245                | I         | MN125002             |
|        | R: CAATAAGGAGGCGAGGACT    |              |                               |                        |           |                      |
| Tof35  | F: TCTCTGTCATATAATGAGTGTTC | (AATG)_11   | 0.3                           | 222–268                | I         | MN125003             |
|        | R: CTCCTACCTCCGCTGTTGG    |              |                               |                        |           |                      |
| Tof46  | F: GTCGCTTCCTCCCTCCTGATA  | (ACT)_4      | 0.07                          | 298–301                | I         | MN125004             |
|        | R: CGTGGTGAACATAGGTGTTGA  |              |                               |                        |           |                      |
| Tof50  | F: CATGATATTTAAGTCCGTCCC  | (ACG)_4      | 0.15                          | 292–352                | I         | MN125005             |
|        | R: TTATCCCAAGATAGGCTGGCC  |              |                               |                        |           |                      |
| Tof5   | F: CCAGCTACAGCGGCCAGCATCA | (AAG)_10     | 0.0375                        | 121–169                | II        | MN125006             |
|        | R: GCAACCTCCCATGGGATCAA   |              |                               |                        |           |                      |
| Tof9   | F: AAGCCAGCCTCATCTGTG     | (AGC)_6      | 0.07                          | 135–141                | II        | MN125007             |
|        | R: ACCTGAGGCTGCTGTATGG    |              |                               |                        |           |                      |
| Tof15  | F: GACAGTATGTTGGATACCTCT  | (ACT)_5      | 0.1                           | 164–172                | II        | MN125008             |
|        | R: ACATGGCCACCAAGTACCA    |              |                               |                        |           |                      |
| Tof21  | F: GCCGCAAGTCTCACGAGAGG   | (ACT)_7      | 0.1                           | 186–195                | II        | MN125009             |
|        | R: TCGCCGCGGTGCTCAATGAG   |              |                               |                        |           |                      |
| Tof27  | F: TACGTTACCGACAACCTGTAGA | (ACAT)_5     | 0.1                           | 198–218                | II        | MN125010             |
|        | R: TTGAAAGCTTGTGCCGTTCA   |              |                               |                        |           |                      |
| Tof31  | F: TGAAAGCAGCGAGATACCT    | (AAG)_3      | 0.1                           | 224–233                | II        | MN125011             |
|        | R: CCCCACCAAGAGCAGATAGA   |              |                               |                        |           |                      |
| Tof34  | F: CGGACTGGAGAACAGAAAAAATAAA | (AGAT)_1  | 0.1                           | 197–278                | II        | MN125012             |
|        | R: AGGGCGGAGATGCCACCA     |              |                               |                        |           |                      |
| Tof42  | F: GCCGTAATGTCCGCCAGG     | (AAG)_11     | 0.1                           | 283–304                | II        | MN125013             |
|        | R: CGGGACAGCAATGATAGTA    |              |                               |                        |           |                      |
| Tof45  | F: ACACACGACAGCGACACTGA   | (AGC)_5      | 0.1                           | 290–305                | II        | MN125014             |
|        | R: GATAGATGTAAGCCTCAAGATGGATCA |       |                               |                        |           |                      |

*Optimal annealing temperature was 60°C for both multiplexes.*
35 cycles of denaturation (95°C for 30 s), annealing (60°C for 40 s), and extension (72°C for 30 s); and a final extension at 72°C for 8 min.

PCR products were diluted with ddH2O 10× (PCR product of multiplex mix I) and 20× (PCR product of multiplex mix II). Each PCR product (1 μL) was mixed with 12 μL of formamide mixed with 0.1 μL of GeneScan 500 LIZ Size Standard (Thermo Fisher Scientific, Waltham, Massachusetts, USA). Fragment lengths were determined by capillary gel electrophoresis with an ABI 3130 Genetic Analyzer using GeneMapper 4.0 (Thermo Fisher Scientific).

We used SPAGeDi (Hardy and Vekemans, 2002) to calculate levels of observed and expected heterozygosity and number of alleles in each population. GENEPOP version 4.2 (Rousset, 2008) was used to test Hardy–Weinberg equilibrium and linkage disequilibrium. The number of alleles per locus ranged from one to 11, and the mean levels of observed and expected heterozygosity ranged from 0.10 to 0.39 and 0.12 to 0.49, respectively (Table 2). The highest number of alleles and heterozygosity values were detected in two populations from the Alps, which is the center of distribution of the species. All markers were polymorphic in at least one of the studied populations, with 68% of markers polymorphic in four of five populations. Significant deviations from Hardy–Weinberg equilibrium were found in only five loci in the Monkova dolina population (Table 2), and significant linkage disequilibrium was found only in the Sosnowiec-Bory population (loci Tof22/Tof35, significant after Bonferroni correction).

We also tested cross-amplification of these loci in the closely related species *T. pusilla* using six individuals from two populations (Appendix 1). Six of 19 loci successfully amplified (Table 3). Using these microsatellites in other species in the genus *Tofieldia* may therefore be possible. More loci for testing are available in the GenBank Sequence Read Archive (BioSample accessions SAMN11608780 and SAMN11608781).

**CONCLUSIONS**

We developed, successfully amplified, and multiplexed 19 polymorphic markers in *T. calyculata*. These polymorphic loci will be used in a future study to reveal the genetic diversity of remaining lowland populations and to compare them with alpine populations.
ACKNOWLEDGMENTS

The authors thank C. Pachschwöll, T. Pachschwöll, and F. Kolář for help collecting samples in the Alps and all of those involved in obtaining collecting permits. This work was supported by the Charles University Grant Agency (grant no. 1450218).

DATA AVAILABILITY

All sequences are archived in the GenBank Sequence Read Archive (BioSample accessions SAMN11608780 and SAMN11608781). Primer sequences have been deposited to NCBI’s GenBank database; accession numbers are listed in Table 1.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

APPENDIX S1. List of all identified microsatellite loci for Tofieldia calyculata.

LITERATURE CITED

Gabrielová, J., T. Fialová, and Z. Münzbergová. 2011. Critically endangered plant species of the Czech Republic: What is the situation of their endangerment in other European countries? *Prároda* 31: 299–343.

Grulich, V., and K. Chobot. 2017. *Cřený* seznam ohrožených druhů České republiky: Cévnaté rostliny. [Red List of Threatened Species of the Czech Republic. Vascular Plants.] *Prároda* 35: 75–132.

Hardy, O. J., and X. Vekemans. 2002. SPAGeDi: A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes* 2: 618–620.

Kaplan, Z., J. Danihelka, J. Štěpáneková, P. Bureš, J. Zázvorka, Z. Hroudová, M. Ducháček, et al. 2015. Distributions of vascular plants in the Czech Republic. Part 1. *Příroda* 417–500.

Lodhi, M. A., G. Ye, N. F. Weeden, and B. I. Reisch. 1994. A simple and efficient method for DNA extraction from grapevine cultivars and *Vitis* species. *Plant Molecular Biology Reporter* 12(1): 6–13.

Malausa, T., A. Gilles, E. Méglécz, H. Blanquart, S. Duthoys, C. Costedoat, and V. Dubut. 2011. High-throughput microsatellite isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries. *Molecular Ecology Resources* 11: 638–644.

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

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.

APPENDIX 1. Accession information for Tofieldia species used in this study.

| Species name                  | Voucher specimen  | Locality         | Country | n  | Latitude        | Longitude       |
|-------------------------------|-------------------|------------------|---------|----|----------------|----------------|
| *Tofieldia calyculata* (L.) Wahlenb. | —                 | Gosausee         | Austria | 20 | 47.5243239°N | 13.5105097°E   |
| *Tofieldia calyculata*        | Tch1901 (PRA)     | Weißenbach       | Austria | 20 | 47.6702111°N  | 14.1282278°E   |
| *Tofieldia calyculata*        | —                 | Monková dolína   | Slovakia| 20 | 49.2592200°N  | 20.2309106°E   |
| *Tofieldia calyculata*        | —                 | Sosnowiec-Bory   | Poland  | 20 | 50.7670711°N  | 19.2722917°E   |
| *Tofieldia calyculata*        | —                 | Jestebské slaté | Czech Republic | 21 | 47.5131489°N  | 13.6616417°E   |
| *Tofieldia pusilla* (Michx.) Pers. | Tph1801 (PRA)    | Dachstein        | Austria | 2  | 47.0612222°N  | 12.7674444°E   |

Note: *n* = number of individuals.

APPENDIX 2. Thirty-one additional microsatellite loci developed for *Tofieldia calyculata.*

| Locus | Primer sequences (5′–3′) | Repeat motif | Allele size (bp) |
|-------|--------------------------|--------------|-----------------|
| Tof1  | F1: GGGTGAGGTTGAAGAGAGGA | (AGC)₅       | 113             |
|       | R1: AGGACTGACCAATAGGACCT |             |                 |
| Tof2  | F1: GGGAGATGACTGGCGATCGT | (AATG)₆     | —               |
|       | R1: AGAGCTCAAGGTCAACACCC |             |                 |
| Tof3  | F1: TTGCCCTCTGGCTTGAAACA | (AGC)₆     | 140             |
|       | R1: CTCCAAATTGGTGAGGGAGGT |             |                 |
| Tof6  | F1: CGCCCAACTGCAAGGCGCTG | (ACC)₅     | 141             |
|       | R1: AGAATGCTTCAAGAAGAAACA |             |                 |
| Tof8  | F1: TGCTTGCGGATTGTTGTA  | (AGAT)₅     | 157             |
|       | R1: GCATTAAAGCAGATAGGATGGA |             |                 |
| Tof10 | F1: GGTCTATGGTGCTCTCCGCA  | (AGC)₅     | 155             |
|       | R1: CGCAGGCTATCCCAACAGAGG | (ACT)₅     | 168             |
| Tof12 | F1: CCATAGAGCCAGGGGTGAT | (ACT)₅     | 173, 179        |
|       | R1: GTGGTATAGGCTCCATGCA |             |                 |
| Tof14 | F1: AGGTTCTACTCGTGCGGGTCGG | (AAG)₅     | 173, 179        |
|       | R1: CGCCTATGGAGCCGAGTAAAGAT |         |                 |

(Continues)
### APPENDIX 2. (Continued)

| Locus | Primer sequences (5′–3′) | Repeat motif | Allele size (bp)* |
|-------|-------------------------|--------------|------------------|
| Tof16 | F: GGTGCTAGCTAGCTTTAGG  | (ACAT)_6     | 192              |
|       | R: TGGTCAGGTTGACTATCAGT  |              |                  |
| Tof17 | F: CCGTGCCGAGACTACCTTTTC | (ATC)_6      | 197              |
|       | R: TTGCGGAATGGAGACTTTTTT |              |                  |
| Tof18 | F: AGAGTGACCGTGATCTTCC   | (ATTG)_6     | 191, 204         |
|       | R: TGCTCTCATCAACAAAGGAA  |              |                  |
| Tof20 | F: AGATAGAGCCATGGTGGAC   | (AAGG)_4     | —                |
|       | R: AGGAGGGAACCTTGCGGA    |              |                  |
| Tof23 | F: GGAGCAGGCTCCTGACCTTGG | (ACTG)_4     | —                |
|       | R: CCAAGATTCCCTCAATGAC   |              |                  |
| Tof24 | F: TGAAGAGCGATTAGGCTCCAGA| (AAGG)_4     | —                |
|       | R: TTGTGGCCCTCCAGTGGG    |              |                  |
| Tof25 | F: GTTTGATATTCTACAGGACT  | (AAGG)_4     | 217              |
|       | R: GTGAGAGCAGATGACCTTTA  |              |                  |
| Tof26 | F: ACCCTACAGATGCATACCTAC | (AAGG)_4     | 221              |
|       | R: ACATAGAGGCGTGAGGGGA   |              |                  |
| Tof28 | F: TTGAGGCTGCTCGTCTCAG   | (AGAT)_4     | 227              |
|       | R: GCGTTCCGTAAGGCGAGAAG  |              |                  |
| Tof29 | F: TTAGGAGGCTGGTCTGTCTC  | (AGAT)_4     | —                |
|       | R: GCTTTGGCTGCAAGGTGAC   |              |                  |
| Tof30 | F: CCCAAAGCCGCTCAAGGAA   | (AAGG)_4     | 245              |
|       | R: CCTGAAAGGCTGCTGAGT    |              |                  |
| Tof32 | F: TCTTGAGTCGCTGCACTTT   | (AAGG)_4     | 243, 250         |
|       | R: GCCTTGAAGGCTCCGACCA   |              |                  |
| Tof36 | F: GACCTCGCTCTAGTAATGAG  | (ACAT)_2     | —                |
|       | R: CCAACACTCCGCCACATGAC  |              |                  |
| Tof37 | F: ACAATGGAATGTGATATGCA  | (ACAT)_2     | 268              |
|       | R: TGAGGATGCATCCACACCA   |              |                  |
| Tof38 | F: GTGAGAGGAGACACTGGGG   | (ACAT)_2     | —                |
|       | R: GCTCTGAGCTATGACAGGTCT |              |                  |
| Tof39 | F: CCCCTTGGTCTCAGTTTGGG  | (AAGG)_4     | 280              |
|       | R: TGGAGTTGGCCTTACCAACCA |              |                  |
| Tof40 | F: TGTCAAGGGTTGAGCTTGG   | (AAGG)_4     | 288              |
|       | R: GAGCTCTGCGATATTTCCGCT |              |                  |
| Tof41 | F: CCGGCGACACCTCTATCTCT  | (AAGG)_4     | 294              |
|       | R: GACAGAAGGCTAATGCGCA   |              |                  |
| Tof43 | F: CTGCAAGACTGCTGGGCAAC  | (AAGG)_4     | 307              |
|       | R: GACCTTGAGGAAGTACGG    |              |                  |
| Tof44 | F: TGCGACTAGGCAACAGACCA  | (AAGG)_4     | —                |
|       | R: CTTGATGAGGTTGGTACGATA |              |                  |
| Tof47 | F: AGCTGACACCTAAAATCCAGG | (AAGG)_4     | 313              |
|       | R: GCAAAATGGCTGGCAGTGA   |              |                  |
| Tof48 | F: TGTTGATGGTGCCTCAGCAGA | (AAGG)_4     | —                |
|       | R: AGATGCGACGGTATATCCTTT |              |                  |
| Tof49 | F: GCCCTTCTCCACCTACCTGCG | (AAGG)_4     | 317              |
|       | R: TATGAGCGAGCTGAACTTGGC |              |                  |

---

*Note: — = unsuccessful amplification or unclear pattern.
1These loci were either monomorphic, did not amplify, or presented unclear amplification.
2Amplification was performed with M13 tail primers.