Microsatellite markers for *Anthericum ramosum*: Development, characterization, and cross-species amplification

Petr Vít1,2, Karol Krak1,2, Jan Douda1, and Bohumil Mandák1,2,3

The genus *Anthericum* L. (Agavaceae) comprises more than 50 perennial herbaceous species, distributed mainly in Africa and Central and South America. Three species, all within section *Anthericum*, are native to Europe, namely *A. baeticum* Boiss., *A. liliago* L., and *A. ramosum* L. (Hegi, 1939). The European species have different ecologies: *A. baeticum* prefers more stony slopes and rocks with predominantly south-facing exposure, these species co-occur only rarely (Rosquist, 2001). The Central European distribution of *A. ramosum* is discontinuous, with populations growing among fragmented dry steppe–like habitats. The current distribution pattern is mainly the result of past changes in agricultural practice, which have resulted in significant loss of continuous dry steppe habitats. Knowledge of genetic variation in fragmented populations of this species may serve as a stepping stone to better understand its historical origins and past migration patterns. It may also reveal the population genetic processes taking place in fragmented populations of other important steppe species, and thus offer insights into processes that may endanger these fragmented populations (e.g., loss of genetic diversity caused by genetic drift).

**METHODS AND RESULTS**

**Microsatellite development**

DNA of *A. ramosum* was extracted from silica gel–dried leaves taken from a single plant using the DNeasy Plant Mini Kit (QIAGEN, PRIMER NOTE

Applications in Plant Sciences 2020 8(2): e11323; http://www.wileyonlinelibrary.com/journal/AppsPlantSci © 2020 Vít et al. Applications in Plant Sciences is published by Wiley Periodicals, Inc. on behalf of the Botanical Society of America. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

The genus *Anthericum* L. (Agavaceae) comprises more than 50 perennial herbaceous species, distributed mainly in Africa and Central and South America. Three species, all within section *Anthericum*, are native to Europe, namely *A. baeticum* Boiss., *A. liliago* L., and *A. ramosum* L. (Hegi, 1939). The European species have different ecologies: *A. baeticum* is found in wet meadows of the Iberian Peninsula, whereas *A. liliago* and *A. ramosum* are found in dry grasslands (Hegi, 1939; Rosquist, 2001). All three European *Anthericum* species are morphologically similar, with their main discriminating characters present in their inflorescences (Hegi, 1939).

*Anthericum ramosum* is a perennial diploid (2n = 30, 32) species that is self-compatible and insect-pollinated with frequent clonal reproduction (Rosquist, 2001). This species has a wide European distribution from Spain to western Russia (Meusel et al., 1965) and is an important component of sub-Mediterranean (Pontic) dry steppe habitats (Chytrý, 1995). It occurs in open, xeric, and sunny slopes and dry open forests with elevated CaCO$_3$ contents, from lowlands to ca. 1800 m in Central Europe (Hegi, 1939). Although the ecology of the species is similar to the tetraploid *A. liliago*, which prefers more stony slopes and rocks with predominantly southern exposure, these species co-occur only rarely (Rosquist, 2001). Mixed populations with the occurrence of hybrid individuals have been documented in Scandinavia (Rosquist, 2001).

**Citation**

Vít, P., K. Krak, J. Douda, and B. Mandák. 2020. Microsatellite markers for *Anthericum ramosum*: Development, characterization, and cross-species amplification. Applications in Plant Sciences 8(2): e11323.

doi:10.1002/aps3.11323

**KEY WORDS** Agavaceae; *Anthericum ramosum*; cross-species amplification; microsatellites.
Hilden, Germany), according to the manufacturer's protocol. The DNA library was prepared with a NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, Massachusetts, USA), and it was sequenced on the Illumina MiSeq platform using the MiSeq Reagent Kit (2 × 300 bp; Illumina, San Diego, California, USA), together with another nine indexed libraries, in one run. Sequencing resulted in 1,677,883 raw reads (available at the National Center for Biotechnology Information [NCBI] Sequence Read Archive [accession no. SRR10054461]). Chloroplast reads were removed with Bowtie 2 (Langmead and Salzberg, 2012) using the reference chloroplast sequence of *A. ramosum* (GenBank accession no. KX790364.1). Sequences containing perfect microsatellite motifs (i.e., di-, tri-, and tetranucleotide repeats with minimum lengths of 14, 18, and 20 bp, respectively) were extracted using SSR_pipeline (Miller et al., 2013). Primers were designed for all extracted sequences using Primer3 (Untergasser et al., 2012), as integrated in MSATCOMMANDER version 0.8.2 (Faircloth, 2008). See Appendix 1 for settings of programs and scripts used for manipulation with sequence reads.

### Biological validation

A total of 110 randomly selected candidate primer pairs (defining repeats with 100–400 bp amplicon lengths) were tested for amplification in seven individuals of *A. ramosum* from different populations (Appendix 2). Total genomic DNA of *A. ramosum* was extracted from silica gel–dried leaves using the DNeasy 96 Plant Kit (QIAGEN). PCR was carried out in volumes of 10 μL using the QIAGEN Multiplex PCR Kit. The reaction contained 1× concentrated QIAGEN Multiplex PCR Master Mix, 0.05 μM of forward and 0.2 μM of reverse primer, and 10 ng of genomic DNA. The PCR fragments were fluorescently labeled by adding 0.2 μM M13 primer (with either FAM, NED, VIC, or PET; Thermo Fisher Scientific, Waltham, Massachusetts, USA) in the reaction as described in Schuelke (2000). The following PCR profile was used: an initial denaturation step at 95°C for 15 min; followed by 25 cycles of denaturation at 95°C for 30 s, annealing at 55°C for 30 s, and extension at 72°C for 2 min; followed by 10 cycles of denaturation at 95°C for 30 s, annealing at 50°C for 30 s,

### Table 1. Characteristics of 19 polymorphic microsatellite loci developed in *Anthericum ramosum*.

| Locus     | Primer sequences (5′–3′) | Repeat motif     | Allele size range (BP) | Fluorescent dye | GenBank accession no. |
|-----------|--------------------------|------------------|------------------------|----------------|-----------------------|
| AR-di-05  | F: CCGCAAACTCAGGAGTAGAG  | (AC)14            | 181–193                | 6-FAM<sup>m1</sup> | MN259494             |
|           | R: TGATTCTCATTGACCAGGCC  |                  |                        |                |                       |
| AR-di-12  | F: GTTGAAGTTGAGGGGATGAT | (AT)12            | 302–322                | 6-FAM<sup>m1</sup> | MN259502             |
|           | R: AGGAGAAAGGATGAAATCGC  |                  |                        |                |                       |
| AR-di-27  | F: TGATCTTAAAATGTGCGGCC  | (AG)15            | 382–420                | 6-FAM<sup>m2</sup> | MN259492             |
|           | R: GTCAATTTGTCGGCCACC    |                  |                        |                |                       |
| AR-di-33  | F: GGACACCCCTTGGGAGTTG  | (AG)15            | 204–248                | VIC<sup>m2</sup> | MN259497             |
|           | R: CAGCAAAATCTGATGGGCTC  |                  |                        |                |                       |
| AR-di-34  | F: TTTCTGGAGGAAACCTCTGC | (AG)15            | 152–186                | PET<sup>m2</sup> | MN259498             |
|           | R: GAAACAGAGGGAGGAAAG    |                  |                        |                |                       |
| AR-di-50  | F: ACCAACAATTTACAGGCAAC  | (AC)10            | 231–263                | PET<sup>m1</sup> | MN259503             |
|           | R: ACACTCTCTACTAGAACCAC  |                  |                        |                |                       |
| AR-di-96  | F: CCCCACATTTCTCTGGTAAAC | (AG)10            | 250–258                | 6-FAM<sup>m1</sup> | MN259493             |
|           | R: AAAGTCAAAGTTGAAACGC  |                  |                        |                |                       |
| AR-di-98  | F: CAGCGGATATCTCTCTCCTC  | (AG)10            | 213–257                | 6-FAM<sup>m1</sup> | MN259495             |
|           | R: GACCTGTTCTCCCGGCTAC  |                  |                        |                |                       |
| AR-di-101 | F: CACTCTGACCTCTGAGCAAG | (AG)11            | 187–251                | VIC<sup>m1</sup> | MN259496             |
|           | R: GCTGTCATCACTGACTGCTTG |                  |                        |                |                       |
| AR-di-106 | F: TCCCTAATCTCTACTACCCG | (AG)8             | 147–175                | PET<sup>m1</sup> | MN259499             |
|           | R: TCATCAATGCCGCTCTCC   |                  |                        |                |                       |
| AR-di-107 | F: CATCGAGACCGAGGAATCC  | (AC)7             | 202–238                | NED<sup>m1</sup> | MN259500             |
|           | R: TGGAGAAGCCAAATCTTCCAC |                  |                        |                |                       |
| AR-di-108 | F: TTGTATTCTCTTGGGCTCTG | (AG)7             | 134–146                | VIC<sup>m1</sup> | MN259501             |
|           | R: GGAGAGGCTGAGAAGATTCC |                  |                        |                |                       |
| AR-tet-65 | F: TACACAGGCTACGTCGCGC  | (ACGC)3           | 110–128                | VIC<sup>m1</sup> | MN259505             |
|           | R: GGTCTGTGCACAAGATGTGC |                  |                        |                |                       |
| AR-tet-91 | F: CGAGTACGTACTTTCTGTCAC| (AAAT)4           | 247–267                | NED<sup>m1</sup> | MN259504             |
|           | R: ACGAGCTCCGAGGAAGATGC |                  |                        |                |                       |
| AR-tri-16 | F: CACCCCTACATCTCCAAACAC| (AAAG)3           | 232–338                | PET<sup>m2</sup> | MN259506             |
|           | R: GCTGTTGGTATGTTGCTCTC |                  |                        |                |                       |
| AR-tri-59 | F: AAACCTCTCATATCCATCC  | (AAAG)3           | 258–285                | PET<sup>m2</sup> | MN259509             |
|           | R: TTATCCACATGTGACGGG   |                  |                        |                |                       |
| AR-tri-60 | F: TCTCTGCTACATCCGATG  | (AAAT)3           | 406–430                | VIC<sup>m1</sup> | MN259510             |
|           | R: GTTCAACCGGTAAACCTAC  |                  |                        |                |                       |
| AR-tri-86 | F: GTTGAATTTTGATTGTCGCG | (AAAT)4           | 329–386                | NED<sup>m1</sup> | MN259507             |
|           | R: GGAATTGACAAGGTTACGC  |                  |                        |                |                       |
| AR-tri-90 | F: ATGCACTCTGACGGCACGC  | (ATC)6            | 134–146                | 6-FAM<sup>m1</sup> | MN259508             |
|           | R: GCACCGAGGAGGCACTAGGC |                  |                        |                |                       |

<sup>a</sup>An M13 tag (GGAAACAGCTATGACCAT; Schuelke, 2000) was added to the 5′ end of the forward primer. Optimal annealing temperature was 55°C.

<sup>b</sup>The multiplex number is indicated as m1 or m2.
and extension at 72°C for 2 min; with a final extension at 72°C for 10 min. The success of amplification was verified by electrophoresis in a 2% agarose gel, with only 79 primer pairs showing clear bands for the seven individuals used for the polymorphism tests. One microliter of PCR product was added to a mix of 12.0 μL of Hi-Di Formamide (Applied Biosystems, Waltham, Massachusetts, USA) and 0.2 μL of GeneTrace LIZ 500 Size Standard (Carolina Biosystems, Prague, Czech Republic) for fragment analysis on the Applied Biosystems 3500 Genetic Analyzer. The electropherograms were analyzed using GeneMarker version 2.7.4 (SoftGenetics, State College, Pennsylvania, USA). Finally, 19 easily scorable, polymorphic markers were selected for genetic analysis of four A. ramosum populations and for a cross-species amplification test of A. liliago, Chlorophytum comosum (Thunb.) Jacques, Muscari comosum (L.) Mill., and M. tenuiflorum Tausch (Table 1, Appendix 2).

### Microsatellite data analysis and results

For each polymorphic microsatellite locus, the following basic genetic diversity parameters were calculated by using the package diveRsity in R (Keenan et al., 2013): the number of alleles (A), levels of observed (H<sub>o</sub>) and expected (H<sub>e</sub>) heterozygosities, and the inbreeding index f (F<sub>IS</sub>) (Weir and Cockerham, 1984) as a measure of departure from within-population random mating. The presence of null alleles was evaluated in MICRO-CHECKER (van Oosterhout et al., 2004). Deviation from Hardy–Weinberg equilibrium was calculated using Fisher’s exact test for each locus, and P values were adjusted using Bonferroni correction (Wright, 1992).

The basic genetic parameters of four populations are presented in Table 2. A total of 195 alleles were identified in the 19 polymorphic loci analyzed. Among the four populations, the number of alleles per locus ranged from two to 17. Levels of heterozygosities (H<sub>o</sub> and H<sub>e</sub>) ranged from 0.000 to 1.000 and from 0.100 to 0.900, respectively. The level of inbreeding (inbreeding coefficient F<sub>IS</sub>) ranged from –0.504 to 1.000. Two pairs of loci (AR-di-101 × AR-di-107, AR-di-108 × AR-tri-90) were significantly linked, probably because of the limited number of individuals and populations. Significant deviation from Hardy–Weinberg equilibrium was tested in 11 out of 19 loci, caused by heterozygosity deficiency in particular populations (see Table 2 for details). Four loci (AR-tri-16, AR-tri-59, AR-tri-60, and AR-tri-96) showed significant presence of null alleles (see Table 2). These loci were not further included in the cross-species amplification tests.

Cross-species amplification was tested in four related species: A. liliago, C. comosum, M. comosum, and M. tenuiflorum. Amplification of developed markers was partly successful in closely related species (i.e., A. liliago and C. comosum) and mostly unsuccessful in phylogenetically more distant species (i.e., M. comosum and M. tenuiflorum) (Table 3). Ten markers showed significant variation within A. liliago individuals; these markers may be suitable for study of populations in which both species occur, as well as for study of gene flow between both species.

### CONCLUSIONS

Nineteen novel nuclear microsatellite markers were developed for A. ramosum. It is the first set of microsatellites developed for this
These markers will be helpful for population genetic studies to reveal phylogeographic patterns of important steppe species. Several markers were successfully amplified within related species, and after optimization, they may be useful in these species for studies of phylogeography, hybridization, or gene flow.

**ACKNOWLEDGMENTS**

This study was supported by the Czech Science Foundation (18-03028S) and is part of the long-term research development project RVO 67985939. The authors thank Kateřina Machynková for her help and assistance in the laboratory and Dr. Sandy Lang for assistance with language editing.

**AUTHOR CONTRIBUTIONS**

P.V., K.K., and B.M. conceived and designed the study. P.V., K.K., J.D., and B.M. collected the plant material. P.V. designed and supervised the laboratory work. P.V. and B.M. analyzed the data and drafted the manuscript.

**DATA AVAILABILITY**

The raw sequence reads are deposited in the National Center for Biotechnology Information (NCBI; GenBank Sequence Read Archive accession no. SRR10054461). Information about plant material is stored as BioSample no. SAMN12655405 under BioProject no. PRJNA562837. Sequences of reads used for primer design were uploaded to NCBI, see Table 1 for accession numbers.

**LITERATURE CITED**

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

Chytrý, M. 1995. Are species with similar ranges confined to similar habitats in a landscape? Preslia 67: 25–40.

Faircloth, B. C. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. Molecular Ecology Resources 8: 92–94.

Hegi, G. 1939. Illustrierte Flora von Mittel-Europa. Band II, Monocotyledones (II. Teil). J. F. Lehman, Munich, Germany.

Keenan, K., P. McGinnity, T. F. Cross, W. W. Crozier, and P. A. Prodöhl. 2013. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. Methods in Ecology and Evolution 4: 782–788.

Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9: 357–359.

Meusel, H., E. J. Jäger, and E. Weinert. 1965. Vergleichende Chorologie der Zentral- und Osteuropäischen Flora, Vol. 1, G. Fischer, Jena, Germany.

Miller, M. P., B. J. Knaus, T. D. Mullins, and S. M. Haig. 2013. SSR_pipeline: A bioinformatic infrastructure for identifying microsatellites from paired-end Illumina high-throughput DNA sequencing data. Journal of Heredity 104: 881–885.

Rosquist, G. 2001. Reproductive biology in diploid Anthericum ramosum and tetraploid A. iliiago (Anthericaceae). Oikos 92: 143–152.

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

Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm, and S. G. Rozen. 2012. Primer3—New capabilities and interfaces. Nucleic Acids Research 40: e115–e115.

van Oosterhout, C., W. F. Hutchinson, D. P. Wills, and P. Shipley. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4: 535–538.

Weir, B. S., and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38: 1358–1370.

Wright, S. P. 1992. Adjusted P-values for simultaneous inference. Biometrics 48: 1005–1013.

**TABLE 3.** Results of cross-species amplification of 19 microsatellite markers developed for Anthericum ramosum and tested in four related species.

| Locus   | Anthericum liliago (N = 5) | Chlorophytum comosum (N = 1) | Muscari comosum (N = 5) | Muscari tenuiflorum (N = 5) |
|---------|---------------------------|-----------------------------|------------------------|-----------------------------|
| AR-di-05 | 183                       | —                           | —                      | —                           |
| AR-di-12 | —                         | —                           | —                      | —                           |
| AR-di-27 | 386–412                   | 276–296                     | —                      | —                           |
| AR-di-33 | 204–246                   | —                           | —                      | —                           |
| AR-di-34 | —                         | 166                         | 160–226                | —                           |
| AR-di-50 | 251–253                   | —                           | —                      | —                           |
| AR-di-98 | 219–233                   | 221                         | —                      | —                           |
| AR-di-101| —                         | —                           | —                      | —                           |
| AR-di-106| —                         | —                           | —                      | —                           |
| AR-di-107| 170–242                   | —                           | 210                    | —                           |
| AR-di-108| 118                       | —                           | —                      | —                           |
| AR-tet-65| 124–152                   | —                           | —                      | —                           |
| AR-tet-91| 251–259                   | 268                         | —                      | —                           |
| AR-tri-16| 311–314                   | —                           | —                      | —                           |
| AR-tri-59| 270                       | —                           | —                      | —                           |
| AR-tri-60| —                         | —                           | —                      | —                           |
| AR-tri-86| 329–338                   | —                           | 329                    | —                           |
| AR-tri-90| 134–146                   | 188–197                     | —                      | —                           |
| AR-tri-96| —                         | —                           | —                      | —                           |

*Note:* — indicates no amplification; N = number of individuals sampled.

*Localities for related species are given in Appendix 2. Allele size range is given if the marker was polymorphic.*
APPENDIX 1. The in silico SSR development workflow used in this study.

1. Initial sequence trimming using Trimmomatic 0.36 (Bolger et al., 2014). Paired-end reads were first trimmed with the following settings:
   ILLUMINACLIP:adapters_used.fa:2:30:10
   LEADING:20
   TRAILING:20
   SLIDINGWINDOW:4:20
   MINLEN:48

2. Removal of chloroplast reads using Bowtie 2 (Langmead and Salzberg, 2012).
   A Bowtie index was built from a set of reference DNA sequence (Anthericum ramosum, GenBank accession no. KX790364.1) using a bowtie2-build indexer. Reads of A. ramosum were mapped on the reference index using the Bowtie2 aligner with default settings.

3. Extraction of reads containing microsatellites using SSR_pipeline (Miller et al., 2013).
   Reads containing microsatellite motifs were extracted using the built-in commands quality_sort, joinseqs, and SSR_search using default settings. Only the following changes in search parameters were made to extract perfect di-, tri-, and tetranucleotide repeats:
   - dinucleotide repeats: minimum length 14 bp
   - trinucleotide repeats: minimum length 18 bp
   - tetranucleotide repeats: minimum length 20 bp

4. Design of primers for all sequences containing the above-mentioned microsatellite motifs using Primer3 (Untergasser et al., 2012), as integrated in MSATCOMMANDER version 0.8.2 (Faircloth, 2008) with default settings.

APPENDIX 2. Locality and voucher information for populations used in this study.*

| Species               | Population code | Voucher no. | N  | Collection locality | Geographic coordinates | Elevation (m) |
|-----------------------|-----------------|-------------|----|---------------------|------------------------|--------------|
| Anthericum ramosum L. | AR1<sup>b</sup> | BRNU665930  | 20 | Czech Republic, Louny| 50.410224°N, 13.807042°E | 475          |
| Anthericum ramosum    | AR3<sup>b</sup> | BRNU667304  | 20 | Poland, Raclawice    | 50.338267°N, 20.233317°E | 300          |
| Anthericum ramosum    | AR9<sup>b</sup> | BRNU667337  | 1  | Hungary, Eger        | 48.012401°N, 20.578951°E | 590          |
| Anthericum ramosum    | AR11<sup>b</sup> | BRNU667235  | 20 | Romania, Suceava     | 47.574910°N, 26.255420°E | 325          |
| Anthericum ramosum    | AR61<sup>b</sup> | BRNU667211  | 16 | Ukraine, Bovshiv     | 49.223650°N, 24.698897°E | 260          |
| Anthericum liliago L. | AL28            | BRNU667317  | 5  | Germany, Konnern     | 51.654783°N, 11.756253°E | 100          |
| Chlorophytum comosum  | CC1             |             | 1  | Commercial cultivar  | —                      | —            |
| (Thunb.) Jacques “Atlantic” |          |             |     |                      |                        |              |
| Muscari comosum (L.) Mill. | MC205 | BRNU667297  | 5  | Czech Republic, Lhánice | 49.108932°N, 16.240021°E | 375          |
| Muscari tenuiflorum Tausch | MT2      | BRNU667112  | 5  | Bulgaria, Pazardzhik | 42.125886°N, 24.387949°E | 400          |

Note: N = number of individuals sampled.
<sup>a</sup>Herbarium vouchers are deposited at the herbarium of the Department of Botany and Zoology, Masaryk University, Brno (BRNU), Czech Republic.
<sup>b</sup>Populations used for the initial PCR amplification test.