SIMPLE SEQUENCE REPEAT MARKERS FOR THE ENDANGERED SPECIES \textit{Clianthus puniceus} AND \textit{C. maximus} (\textit{Fabaceae})\textsuperscript{1}

GARY J. HOULISTON\textsuperscript{2,3}, ANA RAMÓN-LACA\textsuperscript{2}, REEMA JAIN\textsuperscript{2}, CAROLINE M. MITCHELL\textsuperscript{2}, AND DAGMAR F. GOEKE\textsuperscript{2}

\textsuperscript{2}Landcare Research, Gerald Street, Lincoln 7640, New Zealand
\textsuperscript{3}Author for correspondence: houlistong@landcareresearch.co.nz

\begin{itemize}
\item \textit{Premise of the study}: Microsatellite markers were developed for \textit{Clianthus puniceus} using a shotgun sequencing library and tested for cross amplification in the closely related \textit{C. maximus} to inform population management of these two endangered species.
\item \textit{Methods and Results}: We constructed a shotgun sequencing library using a Roche 454 sequencer and searched the resulting data set for putative microsatellite regions. We optimized 12 of these regions to produce polymorphic markers for \textit{Clianthus}. We tested these markers on four populations of \textit{C. maximus} and on four \textit{C. puniceus} individuals of known provenance. Alleles per locus ranged from two to nine, while observed and expected heterozygosities per locus ranged from 0.000 to 1.000 and 0.178 to 0.600, respectively.
\item \textit{Conclusions}: These markers will be valuable for ongoing monitoring of the genetic variation in naturally occurring populations of \textit{Clianthus} and for the selection of individuals for revegetation projects in the species’ former range.
\end{itemize}

Key words: \textit{Clianthus maximus; Clianthus puniceus; Fabaceae; simple sequence repeats (SSRs)}.

\textit{Clianthus} Sol, ex Lindl. (\textit{Fabaceae}) is a genus found only in Australia and New Zealand, where it is represented by two highly endangered species, \textit{C. puniceus} (G. Don) Banks & Sol. ex Lindl. (Heenan, 2000) and \textit{C. maximus} Colenso. \textit{Clianthus puniceus} is all but extinct in the wild (de Lange et al., 2010) and is found only as a cultivated plant; \textit{C. maximus} is represented in the wild by around 200 individuals (de Lange et al., 2010) but is also commonly cultivated as an amenity species, and was formerly cultivated by Maori (Song et al., 2008). \textit{Clianthus maximus} is found only on the east coast of the North Island of New Zealand, and is usually restricted to high-disturbance sites such as slips. These plants are known by the common name kaka beak, or ngutu-kaka, referring to the long scarlet flowers resembling the beak of the native parrot (\textit{kākā}, \textit{Nestor meridionalis} (Gmelin, 1788)). The seed pods were used as a food source by early Maori, and the flowers for ornamentation. The flowers also provide nectar for New Zealand endemic passerine birds (Anderson, 2003).

The management of \textit{Clianthus} in New Zealand is a collaborative effort between a government department (the Department of Conservation), local iwi (i.e., Maori tribes—the Lake Waikaremoana Hapu Restoration Trust), and other nongovernmental organizations. These groups manage complementary ongoing programs and frequently discover new plants in the wild. Microsatellite markers are useful to genotype these new discoveries, to identify the plants as naturally occurring or simply garden discards or escapes, and therefore of less use for restoration activities. The rapidity of next-generation discovery of microsatellite markers (Abdelkrim et al., 2009; Zalapa et al., 2012) lends itself to applications such as this.

METHODS AND RESULTS

Leaf material was sourced from a cultivated plant of \textit{C. puniceus} (Landcare Research Allan Herbarium accession no. CHR559142), and total genomic DNA was extracted using a DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany) following the manufacturer’s instructions. The resulting extraction was adjusted to a concentration of 20 ng/µL in dH2O, as determined via Spectrophotometer (Shimazu, Kyoto, Japan), and used to create a shotgun Multiplex Identifier (MID) library on a Roche 454 Jr. Genome Sequencer (Roche, Basel, Switzerland) using Roche Titanium chemistry (Margulies et al., 2005). The sequencing run resulted in 83,643 reads (average read length 461 bp) for a total yield of 34.6 Mb of sequence.

The library was searched using MSATCOMMANDER (Faircloth, 2008) for di- to hexanucleotide repeat regions with at least five repeat units, and flanked by appropriate regions for primer design. Primers were designed via the default settings of Primer3 (Rozen and Skaltsky, 2000) as implemented in MSATCOMMANDER with the following user specifications: amplification regions of 100–500 bp, an optimal oligonucleotide melting temperature range of 57–62°C, GC content range of 20–80% with an optimum rate of 50%, low levels of self- or pair-complementarity, and a maximum end-stability (AG) of 8.0 (Faircloth, 2008). We chose 48 primer pairs for screening, using an M13 tag (CAGCAGTGGTAAACGAC) on the 5′ end of the forward primer for subsequent fluorescent labeling.

The 48 primer pairs were tested on a single representative of \textit{C. puniceus} and for cross amplification in seven individuals of \textit{C. maximus} (Table 1). DNA was extracted using either a DNeasy Plant Mini Kit (QIAGEN) or iNIRON Plant DNA extraction kits (iNIRON Biotechnology, Seongnam, South Korea).
following the manufacturers’ instructions. PCR was performed in 10-μL reactions (1 μL of template DNA at 5–50 ng added to final concentrations of 0.1 μM forward primer, 0.4 μM reverse primer, 0.4 μM M13 6-FAM–labeled primer, 1× iNtRON buffer, 250 μM of dNTP mix, 40 μg/mL of bovine serum albumin [BSA; New England Biolabs, Ipswich, Massachusetts, USA], 0.08 unit iNtRON Taq DNA Polymerase [iNtRON Biotechnology], volume adjusted with filtered [0.22 μm] and autoclaved Millipore [Merck KGaA, Darmstadt, Germany] \( \text{H}_2\text{O} \). PCR conditions for M13-tagged primer testing were as follows: initial denaturation of 95°C for 4 min followed first by 30 cycles of 94°C for 30 s, 58°C for 45 s, and a final extension at 72°C for 10 min. One-microliter samples of the resulting amplified DNAs were prepared by adding 9 μL of Hi-Di formamide (Applied Biosystems, Carlsbad, California, USA) and 1 μL of LZ-labeled size standard (Applied Biosystems), before being separated on an ABI 3130xl Genetic Analyzer (Applied Biosystems) at the Landcare Research sequencing laboratory (Auckland, New Zealand). Fragments were sized and scored using GeneMapper version 3.7 (Applied Biosystems), and polymorphism and repeatability of each locus were assessed. Twelve of the 48 loci tested produced interpretable fragments and no more than two alleles per individual. A further eight loci amplified reliably but were either monomorphic or produced fixed heterozygote genotypes for all screened individuals (Table 1). The primers for the 12 polymorphic loci were then ordered labeled with different fluorescent dyes (6-FAM, NED, VIC, or PET) to allow coloading of PCR products when genotyping, and omitting the M13 tail. PCR conditions were optimized as follows: initial denaturation of 95°C for 4 min followed by 35 cycles of 94°C for 30 s, 58°C, or 63°C (see Table 1) for 30 s, and 72°C for 45 s, and a final extension at 72°C for 10 min. These primers were then tested on a total of 49 individuals from four populations of C. maximus, representing the majority of wild extant plants, and four individuals of C. puniceus (Table 2, Appendix 1). Due to the size and critically endangered status of these plants, voucher specimens were not prepared as this would have been overly destructive.

The numbers of alleles and observed (\( H_o \)) and expected (\( H_e \)) heterozygosities for each population of C. maximus were estimated using GenAlEx (Peakall and Smouse, 2006). \( H_o \) and \( H_e \) ranged from 0.000 to 1.000 and 0.178 to 0.600, respectively, while mean \( H_o \) and \( H_e \) across all loci and populations were 0.364 and 0.403. Alleles per locus ranged from two to nine, with an average of 5.5 (Table 2). We did not estimate population parameters separately for C. puniceus due to the uncertain number of extant wild-collected individuals as most material is from cultivated sources and likely not representative of the former wild distribution (Song et al., 2010).

All of the 12 loci tested produced polymorphic bands in C. maximus, and six of the 12 were polymorphic in C. puniceus. Of the four populations of C. maximus that were tested, all four had private alleles. No alleles were found in C. puniceus that were tested, all four had private alleles. No alleles were found in C. puniceus.

### Table 1. Characteristics of 12 polymorphic microsatellite loci developed in Clianthus puniceus and C. maximus (CLIANT1–12), and the other eight loci screened that were either monomorphic or produced fixed heterozygous genotypes for all individuals (CLIANT13–20).

| Locus   | Primer sequences (5′–3′)* | Repeat motif | Allele size range (bp) | \( T_a °C \) | Fluorescent dye | GenBank accession no. |
|---------|--------------------------|--------------|------------------------|-------------|----------------|----------------------|
| CLIANT1 | F: CTTAGATGAAGCCAAAGAGACAAAATG | (AGAT)_3 | 119–144 | 58 | PET | KM588656 |
|         | R: GCAGACATATGAGAACGCCGCGC | | | | | |
| CLIANT2 | F: AACCCCTTGTGTGAGACCG | (AGAT)_3 | 202–226 | 58 | 6-FAM | KM588657 |
|         | R: ACGTGTTGGAATGTTGAGT | | | | | |
| CLIANT3 | F: CTCCTGAGGATCGATTACC | (CTT)_3 | 226–229 | 58 | PET | KM588658 |
|         | R: GCCTCAGAGTTGAGAAGAGAAC | | | | | |
| CLIANT4 | F: TTTAGGTTACAAAGACCCCG | (AAC)_3 | 184–189 | 58 | NED | KM588659 |
|         | R: ATCTTTGCTTGGCTTTCG | | | | | |
| CLIANT5 | F: TGTTGTTAGAGTTGAGACCG | (CATAT)_3 | 172–208 | 58 | PET | KM588660 |
|         | R: ACCTCTAGTTGCTTTGGCG | | | | | |
| CLIANT6 | F: CCTCCACACCCACGTCC | (CTT)_10 | 145–161 | 58 | VIC | KM588661 |
|         | R: GAGATGAGGAGACGACGC | | | | | |
| CLIANT7 | F: CAAGAGAAGACGATATTGTTTG | (AAT)_3 | 174–192 | 58 | NED | KM588662 |
|         | R: GATGTGAAAGCAAGGITGAGG | | | | | |
| CLIANT8 | F: TCAGCTCTACAGCTACAG | (CTT)_5 | 154–160 | 58 | 6-FAM | KM588663 |
|         | R: GCCTCCAGGTAGATGTTG | | | | | |
| CLIANT9 | F: CAGAGGCTTGTGCGATGTT | (AAT)_3 | 201–240 | 63 | PET | KM588664 |
|         | R: TGCTCGCTACAGCTAGAAG | | | | | |
| CLIANT10 | F: GGCTGTCGCACTCAGAC | (AAG)_3 | 185–209 | 63 | VIC | KM588665 |
|         | R: GATCCGCACTCCACCTTCCT | | | | | |
| CLIANT11 | F: TCTAGTGGAGCTGAGACGC | (AAAC)_3 | 163–207 | 63 | NED | KM588666 |
|         | R: TGAGAGCCAGGGAAGCTCG | | | | | |
| CLIANT12 | F: TGAGAGGCTACCTGGCTG | (AAG)_3 | 167–173 | 63 | 6-FAM | KM588667 |
|         | R: TGTCTCCTATTTTCACGACG | | | | | |
| CLIANT13 | F: ACTCAATTGCCAAAGACTGCC | (ATGC)_4 | 174 | 58 | N/A | KP100530 |
|         | R: CCAAGAGTTCGCTTAGGTC | (CTT)_10 | 261 | 58 | N/A | KP100531 |
| CLIANT14 | F: ACTTCACTAGACGTGACG | (CTT)_10 | 261 | 58 | N/A | KP100532 |
|         | R: AGGGTAGAGCAGTGGCTG | | | | | |
| CLIANT15 | F: TCGCTCGCTTCTCCACGGC | (ATC)_3 | 246 | 58 | N/A | KP100535 |
|         | R: GTCACTATACCTGATGCGAC | | | | | |
| CLIANT16 | F: CTTCTCTTACAACAGACAC | (GGT)_3 | 239 | 58 | N/A | KP100533 |
|         | R: ATGGCCACAAATCCACTCC | | | | | |
| CLIANT17 | F: CTTCTCTCTGCTGGTCTCTGCT | (AGk)_3 | 246 | 249 | 58 | N/A | KP100534 |
|         | R: TGCACACGTGCTCTGCTTCG | | | | | |
| CLIANT18 | F: GATAAGCAACAGCAGCGCC | (ATC)_7 | 238 | 58 | N/A | KP100535 |
|         | R: TGAGTAGAAGCCTTTCTTGGG | | | | | |
| CLIANT19 | F: CCGTCCTTCGAGAGAGAG | (AGk)_3 | 174 | 58 | N/A | KP100536 |
|         | R: AGGAGAAGGAGGTGAGACTCCG | | | | | |
| CLIANT20 | F: ACTTATGATGCACTACCTGTGGCAAAC | (AGCG)_3 | 179 | 58 | N/A | KP100537 |
|         | R: GTAGAGATGCGTGGCGTGG | | | | | |

Note: \( T_a °C \) is annealing temperature used in PCR.

* M13 tail (TGTAAAACGACGGCCAGT) added to the 5′ end of each forward primer during initial screening.

http://www.bioone.org/loi/apps
Table 2. Results of primer screening in four populations of *Clianthus maximus* and one population of *C. puniceus.*

| Locus       | *C. maximus* | *C. puniceus* |
|-------------|--------------|---------------|
|             | (n = 12)     | (n = 13)      | (n = 14)      | (n = 10)     | (n = 4)       | (n = 53)      |
|             | A  | H<sub>o</sub> | H<sub>e</sub> | A  | H<sub>o</sub> | H<sub>e</sub> | A  | H<sub>o</sub> | H<sub>e</sub> | A  | H<sub>o</sub> | H<sub>e</sub> | A  | H<sub>o</sub> | H<sub>e</sub> |
| CLIANT1     | 3  | 0.083 0.156 | 4  | 0.077 0.541 | 3  | 0.000 0.255 | 5  | 0.300 0.540 | 1  | 0.000 0.000 | 6  | 0.092 0.299 |
| CLIANT2     | 4  | 0.917 0.663 | 6  | 0.692 0.683 | 3  | 1.000 0.622 | 2  | 1.000 0.500 | 3  | 0.750 0.531 | 4  | 0.872 0.600 |
| CLIANT3     | 2  | 0.083 0.080 | 2  | 0.077 0.074 | 2  | 0.214 0.191 | 3  | 0.300 0.405 | 2  | 0.250 0.219 | 4  | 0.185 0.194 |
| CLIANT4     | 1  | 0.000 0.000 | 1  | 0.000 0.000 | 2  | 0.000 0.408 | 2  | 0.000 0.480 | 1  | 0.000 0.000 | 2  | 0.000 0.178 |
| CLIANT5     | 4  | 0.500 0.538 | 5  | 0.308 0.538 | 5  | 0.143 0.661 | 2  | 0.111 0.105 | 1  | 0.000 0.000 | 8  | 0.212 0.368 |
| CLIANT6     | 6  | 0.500 0.753 | 7  | 0.231 0.683 | 2  | 0.000 0.459 | 4  | 0.500 0.655 | 2  | 0.250 0.219 | 9  | 0.376 0.554 |
| CLIANT7     | 2  | 0.083 0.497 | 3  | 0.308 0.592 | 3  | 0.000 0.292 | 4  | 0.429 0.531 | 1  | 0.000 0.000 | 7  | 0.164 0.382 |
| CLIANT8     | 2  | 1.000 0.500 | 3  | 1.000 0.536 | 2  | 1.000 0.500 | 3  | 1.000 0.545 | 3  | 1.000 0.594 | 4  | 1.000 0.535 |
| CLIANT9     | 4  | 0.083 0.601 | 3  | 0.333 0.486 | 3  | 0.071 0.309 | 2  | 0.100 0.495 | 1  | 0.000 0.000 | 6  | 0.118 0.378 |
| CLIANT10    | 3  | 0.583 0.497 | 6  | 0.077 0.707 | 4  | 0.000 0.367 | 5  | 0.300 0.600 | 2  | 0.000 0.375 | 8  | 0.192 0.509 |
| CLIANT11    | 4  | 0.917 0.740 | 2  | 0.923 0.497 | 4  | 0.714 0.594 | 2  | 0.200 0.320 | 2  | 1.000 0.500 | 4  | 0.751 0.530 |
| CLIANT12    | 3  | 0.667 0.469 | 3  | 0.308 0.269 | 4  | 0.643 0.474 | 2  | 0.400 0.320 | 1  | 0.000 0.000 | 4  | 0.403 0.306 |

Note: A = number of alleles; H<sub>e</sub> = expected heterozygosity; H<sub>o</sub> = observed heterozygosity; n = number of individuals sampled.

"a"See Appendix 1 for population location data.

CONCLUSIONS

We developed 12 microsatellite markers for the critically endangered *Clianthus* spp. in New Zealand based on 454 sequencing of total genomic DNA. While polymorphic markers were readily found, departures between observed and expected heterozygosities and low number of alleles per locus indicate that these species have possibly had severe reductions in population size. This strongly correlates with the known decrease in the wild, including one of the species becoming extinct outside of cultivation. These markers will have good utility for management of existing populations, particularly the selection of plants for revegetation plantings to ensure minimal loss of the remaining genetic variation in these species.

LITERATURE CITED

Abdelkrim, J., B. C. Robertson, J. A. L. Stanton, and N. J. Gemmell. 2009. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques* 46: 185–191.

Anderson, S. 2003. The relative importance of birds and insects as pollinators of the New Zealand flora. *New Zealand Journal of Ecology* 27: 83–94.

De Lange, P. J., P. B. Heenan, D. A. Norton, J. R. Rolfe, and J. Sawyer. 2010. Threatened plants of New Zealand. Canterbury University Press, Christchurch, New Zealand.

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

Heenan, P. B. 2000. *Clianthus* (Fabaceae) in New Zealand: A reappraisal of Colenso’s taxonomy. *New Zealand Journal of Botany* 38: 361–371.

Margulies, M., M. Egholm, W. Altman, S. Attiya, J. Bader, L. Bemben, J. Berra, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–380.

Peakall, R., and P. E. Smouse. 2006. GenAIEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

Rozen, S., and H. J. Skaletsky. 2000. 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.

Song, J., J. Murdoch, S. E. Gardner, A. Young, P. E. Jameson, and J. Clemens. 2008. Molecular markers and a sequence deletion in intron 2 of the putative partial homologue of *LEAFY* reveal geographical structure to genetic diversity in the acutely threatened legume genus *Clianthus*. *Biological Conservation* 141: 2041–2053.

Zalapa, J. E., H. Cuevas, H. Zhu, S. Steffen, D. Senalik, E. Zelden, B. McCown, et al. 2012. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *American Journal of Botany* 99: 193–208.

Appendix 1. Location data for the *Clianthus* populations included in this study.

| Species       | Location        | Collector                  | Geographic coordinates |
|---------------|-----------------|----------------------------|------------------------|
| *C. maximus*  | Waikaremoana    | Sandra Elia and Riki Winitana | 38°44.64’S, 177°10.58’E |
| *C. maximus*  | Ruakuri         | Helen Jonas                | 38°46.10’S, 177°24.53’E |
| *C. maximus*  | Mohaka River    | Pete Shaw                  | 38°59.89’S, 177°1.62’E  |
| *C. maximus*  | Tolaga Bay      | Graeme Atkins              | 38°17.13’S, 178°16.95’E |
| *C. puniceus* | Kaipara harbor  | Peter Heenan               | 36°25.34’S, 174°23.57’E |

http://www.bioone.org/loi/apps