Transcriptomic resources and marker validation for diploid and polyploid Veronica (Plantaginaceae) from New Zealand and Europe

EIKE MAYLAND-QUELHHORST1,2, HEIDI M. MEUDT2,3, AND DIRK C. ALBACH2

1 Carl-von-Ossietzky Universität Oldenburg, Carl-von-Ossietzky Straße 9–11, Oldenburg 26111, Germany; and 2 Museum of New Zealand Te Papa Tongarewa, Cable Street, P.O. Box 467, Wellington 6140, New Zealand

• Premise of the study: Polyploidy may generate novel variation, leading to adaptation and species diversification. An excellent natural system to study polyploid evolution in a comparative framework is Veronica (Plantaginaceae), which comprises several parallel, recently evolved polyploid series.

• Methods: Over 105 million Illumina paired-end sequence reads were generated from cDNA libraries of leaf tissue from eight individuals representing three European and four New Zealand species. Forty-eight simple sequence repeat (SSR) and 48 low-copy nuclear (LCN) markers were developed and validated with Fluidigm microfluidic PCR and Illumina MiSeq amplicon sequencing on 48 different individuals each.

• Results: Individual Trinity assemblies were similar regarding annotated transcripts (13,009–14,271), mean contig length (635–742 bp), N50 value (916–1133 bp), E90N50 value (1099–1308 bp), contigs with positive BLAST hits (42–63%), and gene ontology terms. Analyses of 29,738 single-nucleotide polymorphisms (8746 phylogenetically informative) mined from these transcriptomes plus two outgroups (Picrorhiza kurrooa and Plantago ovata) showed moderate to high bootstrap support for all branches and reticulation among sampled European Veronica.

• Discussion: The transcriptome sequences themselves, as well as the validated SSR (40/48) and LCN (11/48) markers derived from them, show inter- and intraspecific genetic variation. These resources will be invaluable for future population genetic, phylogenetic, and functional genetic investigations in polyploid Veronica.

Key words: low-copy nuclear (LCN) markers; polyploidy; simple sequence repeat (SSR) markers; single-nucleotide polymorphisms (SNPs); transcriptome; Veronica.

Polyploidy (whole genome duplication) is a very important process that has shaped flowering plant evolutionary history (Solitis et al., 2009). Much progress in the study of polyploid evolution has been made in the past two decades regarding both ancient paleopolyploidization (Doyle et al., 2008; Solitis et al., 2009) as well as very recent neopolyploidization (Buggs et al., 2009; Abbott et al., 2013). An important research gap (Solitis et al., 2009) is understanding polyploids of intermediate age that have diploid ancestors in the same genus, so-called mesopolyploids, which are characterized by diploid-like reproduction but whose parental subgenomes are still discernible (Mandáková et al., 2010).

Several mesoallopolyploid crop systems (e.g., cotton, soybean, tobacco, wheat) are becoming well understood and have excellent genetic resources; however, understanding natural systems is also important. Specifically, studying natural mesopolyploid species radiations may be key to understanding the importance of polyploidy in angiosperm diversification (Solitis et al., 2009). Recent plant species radiations are a significant contributor to generating plant biodiversity, and evidence suggests that polyploidy has played an important role in these radiations (Mayrose et al., 2011). Many fundamental and biologically interesting questions regarding polyploidy and diversification in plants are yet to be investigated in such systems (Doyle et al., 2008; Solitis et al., 2009; Mayrose et al., 2011).

The large, nearly cosmopolitan genus Veronica L. (Plantaginaceae) comprises approximately 450 species of annual and perennial herbs, shrubs, and small trees with centers of diversity in both Eurasia and New Zealand. The genus is an excellent example of a natural mesopolyploid (∼20 million years old) system comprising multiple lineages, including several recent species radiations, in which polyploidy and hybridization have accompanied diversification (Albach et al., 2008; Meudt et al., 2015). Northern Hemisphere Veronica species are diploids or polyploids with chromosome numbers ranging from 2n = 14–80 and base numbers of x = 6–9 and 17 (Albach et al., 2008). By contrast, Southern Hemisphere species—which evolved as a single lineage from Northern Hemisphere ancestors ∼10 million years ago (Wagstaff et al., 2002; Albach and Meudt, 2010; Meudt et al., 2015)—all have high chromosome numbers (2n = 40–124) with base chromosome numbers of x = 20 or 21 (Albach et al., 2008).
Several studies focusing on *Veronica* in both hemispheres have used standard DNA sequencing and amplified fragment length polymorphism (AFLP) fingerprinting techniques to elucidate patterns of relationship from phylogeography (Meudt and Bayly, 2008) to phylogeny of the genus as a whole (Wagstaff et al., 2002; Albach and Meudt, 2010; Meudt et al., 2015) or of particular polyploid complexes (e.g., Albach, 2007), and used these to infer the evolution of chromosome number, genome size, breeding systems, and habit (Albach and Greilhuber, 2004; Albach et al., 2008; Meudt et al., 2015). Nevertheless, a lack of variable genetic markers using standard DNA sequencing and genotyping techniques, and a lack of appropriate phylogenetic analysis methodologies that can incorporate reticulate evolution and allopolyploids, have hampered further progress in studies of *Veronica* and polyploid evolution at the population, species, and generic levels.

It has been known for some time that low-copy nuclear (LCN) markers can be extremely useful for phylogenetic reconstruction at the genus (interspecific) level, including for elucidating the evolutionary history of polyploids, for which standard uniparental DNA sequencing markers from chloroplast DNA or the internal transcribed spacer (ITS) region are not informative (e.g., Sang, 2002). Apart from LCN markers, microsatellites or simple sequence repeat (SSR) markers are useful for closely related species when traditionally genotyped and analyzed for studies at the infrageneric level, but SSRs and their flanking regions may also be useful as phylogenetic markers when high-throughput sequenced (Chatrou et al., 2009; Germain-Aubrey et al., 2016). This, however, requires new bioinformatic tools such as the workflows MarkerMiner (Chamala et al., 2015) and QDD (Meglécz et al., 2014) for the development of LCN and SSR markers, respectively, using genomic and transcriptomic resources.

High-throughput de novo transcriptome sequencing, or RNA-Seq, has proven to be an excellent source of genetic data for gene characterization and marker development in studies of natural systems with little or no additional genetic resources available (Strickler et al., 2012; Alvarez et al., 2015), as is the case for *Veronica*. The benefits of RNA-Seq are simultaneous characterization of genes and gene expression, reduced representation for large, complex genomes, and the generation of large amounts of sequence data without a reference genome. RNA-Seq also presents its challenges, particularly assembly without a reference genome, and assembly of polyploid genomes. Polyploid transcriptome assembly is an active area of research. A major issue is the differentiation of homoeologs from orthologs. Some studies have tested different pipelines, such as combining multiple k-mer assemblies in polyploid wheat (Krasileva et al., 2013), or combining assemblies from different assemblers and then using a second step to cluster redundant contigs in polyploid tobacco (Nakasugi et al., 2014). However, there are few examples to date of comparisons in natural, noncrop systems with few prior genomic resources. To date, there are no clear answers regarding which assembler, combination of assemblers, or assembly pipeline is best for polyploids and their diploid progenitors or close relatives.

The aim of our study was, therefore, twofold. First, we aimed to generate transcriptomic data for *Veronica*; second, we aimed to use these transcriptomic resources to develop and validate phylogenetically informative sequencing markers. Specifically, in this paper, we generate the first transcriptome resources for the genus *Veronica*, using short-read Illumina HiSeq 2000 (Illumina, San Diego, California, USA) sequencing of eight individuals representing seven species and five different ploidy levels. We then assemble, identify, and broadly characterize and compare a large number of expressed sequences. Single-nucleotide polymorphisms (SNPs) are mined from transcriptomes of these eight individuals plus those of two additional Plantaginaceae outgroups (*Plantago ovata* Forssk. and *Picrorhiza kurrooa* Royle ex Benth., available from public databases) and compared using phylogenetic and network analyses. Secondly, we used the transcriptomic data to discover, design, and develop two types of genetic markers (i.e., LCN and SSR markers). To test the success of our approach, we then used microfluidic PCR and Illumina MiSeq to validate 48 loci in 48 individuals for both LCN and SSR markers. We provide examples of sequence alignments and downstream phylogenetic analyses for representative loci showing their potential phylogenetic utility in *Veronica* when resequenced with high-throughput sequencing. The resource and marker development of the current study provide new, variable markers for future evolutionary studies of the genus. Furthermore, a parallel study currently underway will further examine assembly methods and analyze the transcriptomes themselves to quantify and compare underlying interspecific gene divergence and investigate the timing and mode of polyploidy in the sampled *Veronica* polyploids and their close relatives (Meudt et al., unpublished data). The current study is thus a critical first step toward ultimately understanding the role of polyploidy in generating novel genetic and morphological variation that leads to adaptation and species diversification (Doyle et al., 2008; Soltis and Soltis, 2009).

**MATERIALS AND METHODS**

**RNA extraction, cDNA library prep, and Illumina sequencing—**We sampled leaf tissue from seven greenhouse-grown individuals and from one field-collected individual representing seven species of two polyploid complexes in *Veronica* from New Zealand and Eurasia with three ploidy levels each (Appendix 1). The field-collected material was stored at −80°C on RNA Later (Life Technologies, Carlsbad, California, USA). Because we wanted to take a broad approach to analyze polyploidy in *Veronica* and develop markers for the entire genus, we sampled multiple species in two divergent lineages, rather than multiple individuals per species. Cultivated plants were grown in the same greenhouse in Oldenburg, Germany. Leaf material was harvested and placed directly into tubes with liquid nitrogen, stored at −80°C until extraction, and ground to a powder with a prechilled mortar and pestle while adding liquid nitrogen. RNA was extracted using the RNAeasy kit (QiAGEN GmbH, Hilden, Germany) following manufacturer’s instructions using 500 μL RLC buffer with 4% PVP and 1% β-mercaptoethanol. A DNase I digest and RNase inhibitor reaction was performed using 0.5 μL (20 units) RNase inhibitor, 6.0 μL 10x DNase I buffer, and 1.0 μL DNase I to the resulting 60 μL RNA extract and incubated at 37°C for 15 min. Then, 2.6 μL EDTA (0.2 M, pH = 8; final conc. 8 mM) was added. incubated for 10 min at 75°C, and the RNA was reprecipitated by adding 1:10 3 M sodium acetate, 2.5 volume 100% ethanol, centrifuging at 10,000 g for 20 min. centrifuging at full speed for 5 min, washing with 100 μL 75% ethanol, centrifuging at full speed, air-drying the resultant pellet for 10–15 min, redissolving in 25 μL RNA-free water, and storing at −80°C. Small aliquots of raw RNA extract and the reprecipitated RNA extract were run on the Teco Infiniti Pro F200 (Teco, Crawley, Germany) and Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany) to measure RNA quality and quantity. RNA from eight individuals with RNA Integrity Number (RIN) of 6.8 or greater, 260:280 ratio between 1.9–2.1, and at least 50 ng/μL (Appendix 1) were sent to BGI (BGI-Hong Kong Co. Ltd, Hong Kong, China) for Illumina TruSeq cDNA library preparation on normalized RNA and high-throughput Illumina HiSeq 2000 100- bp paired-end de novo transcriptome sequencing. The transcriptomic data generated here are publicly available in the National Center for Biotechnology Information (NCBI) Sequence Read Archive for submission SRP074674 and the Trinity assemblies in the NCBI Transcriptome Shotgun Assembly Sequence Archive (Table 1; http://www.ncbi.nlm.nih.gov/sra/SRP074674).

http://www.bioone.org/loi/apps
## Table 1: Information about Illumina sequencing reads and Trinity assemblies for the eight individuals of Veronica sampled

| Species (Phylogeny) | Geography | SRA accession | TSA accession | No. (clean reads) | No. of contigs (Trinity) | NCBI value (median contig length) | Mean contig length | N50 value (median contig length) | E90N50 value (computed with the collapseDuplicateCount.sh) | MapMan hierarchical categories (Ath_AGI_LOCUS_TAIR10_Aug2012) | MapMan store | Number of LCN markers detected in MarkerMiner | BLAT with positive BLAST hits |
|---------------------|-----------|---------------|---------------|-----------------|-----------------------|-----------------------------------|-------------------|-------------------------------|---------------------------------|------------------------------------------------|-------------|-----------------------------|----------------------------------|
| Veronica catarractae (6x) | New Zealand | SAMN00966311/GEVY00000000 | GEYS00096631/GEVY00000000 | 23,711,074 | 66,671 | 69,950 | 73,820 | 61,752 | 1118-741,642 | 916-633,493 | 1131-1708,759 | 992 (671,041) | 24,385,504 | 73,889 |
| V. planopetiolata (12x) | New Zealand | SAMN00966311/GEVY00000000 | GEYS00096631/GEVY00000000 | 24,385,098 | 64,950 | 70,255 | 73,820 | 61,752 | 1118-741,642 | 916-633,493 | 1131-1708,759 | 992 (671,041) | 24,385,504 | 73,889 |
| V. ochracea (18x) | Europe | SAMN00966311/GEVY00000000 | GEYS00096631/GEVY00000000 | 25,055,264 | 73,820 | 75,255 | 73,820 | 61,752 | 1118-741,642 | 916-633,493 | 1131-1708,759 | 992 (671,041) | 24,385,504 | 73,889 |
| V. panormitana (2x) | Europe | SAMN00966311/GEVY00000000 | GEYS00096631/GEVY00000000 | 29,050,110 | 61,451 | 68,598 | 68,598 | 46,573 | 1118-741,642 | 916-633,493 | 1131-1708,759 | 992 (671,041) | 24,385,504 | 73,889 |
| V. trichadena (18x) | Europe | SAMN00966311/GEVY00000000 | GEYS00096631/GEVY00000000 | 24,209,390 | 64,451 | 68,598 | 68,598 | 46,573 | 1118-741,642 | 916-633,493 | 1131-1708,759 | 992 (671,041) | 24,385,504 | 73,889 |
| V. cymbalaria (6x) | Europe | SAMN00966311/GEVY00000000 | GEYS00096631/GEVY00000000 | 24,490,360 | 73,820 | 75,255 | 73,820 | 61,752 | 1118-741,642 | 916-633,493 | 1131-1708,759 | 992 (671,041) | 24,385,504 | 73,889 |

Note: GO = gene ontology; LCN = low-copy nuclear; SRA = Sequence Read Archive; TSA = Transcriptome Shotgun Assembly.

*a Sequence Read Archive (SRA) accession numbers for SRA submission SPR074674 (http://www.ncbi.nlm.nih.gov/sra/SRP074674).

*b BLAT search with MapMan categories.

*c Number of LCN markers detected in MarkerMiner (Chamala et al., 2015), contigs longer than 600 bp.

** MarkerMiner 1.0 [Chamala et al., 2015].

---

**Quality control, preprocessing of reads, assembly, and Blast2GO analyses**

The following analyses were carried out on each of the eight individuals sequentially. Demultiplexed Illumina sequencing results were retrieved in FASTQ format via FTP from BGI. Between 12.5 and 13.5 million paired-end reads were generated per individual in both the forward and reverse directions (Table 1), from which single reads, adapters, and reads with a quality score (QC) cutoff of less than 20 had already been removed. After testing the effect of different QC cutoffs on the resulting sequence reads and assemblies of *V. trichadala* and *F. Fourr.,* we used QC = 40 in the bash script TrimClip.sh (De Wit et al., 2012) to remap reads with QC < 40. Reads were screened for contaminating sequences from *H. sapiens, E. coli,* mtDNA, and cpDNA using mirabait (MIRALIB version 4.0; Chevreux et al., 1999) with default settings, the respective databases downloaded from NCBI, and then removed. We used QualityStats (De Wit et al., 2012) and the Galaxy web interface (Afgan et al., 2016) to summarize quality scores, nucleotide distribution data for the forward and reverse reads, CollapseDuplicateCount.sh (De Wit et al., 2012) to calculate the fraction of duplicate reads and singletons, PECCombibiner.sh (De Wit et al., 2012) to remove orphan reads and put remaining reads in the same order in forward and reverse files, and the Velvet helper script shuffleSequences_fastQ.pl to put those two files together in one interleaved file (necessary for Velvet/Oases assembly). The resulting clean sequence reads were assembled de novo using several different assemblers including Trinity, trans-ABySS, SOAPdenovo-Trans, and Velvet/Oases. Relative to the other assemblers, Trinity produced more hits with >80% similarity to contigs >600 bp against Arabidopsis thaliana (L.) Heynh. (data not shown; comparisons done using MarkerMiner 1.0 (Chamala et al., 2015)). Therefore, we chose the de novo assemblies produced using Trinity version r20140717 (Grabherr et al., 2011, compiled for 64-bit Ubuntu) using default settings on the resulting clean sequence reads. For the purposes of marker development, a highly accurate discrimination of homoeologs in polyploids is not necessary at the transcriptome assembly stage, as the discrimination is done in the second resequencing step. Additional comparisons of different assemblers and assembly pipelines, particularly regarding polyploid transcriptomes, were outside the scope of the current study and will be addressed in a subsequent study (Meudt et al., unpublished data). Trinity assemblies of all four New Zealand, all four European, and all eight Veronica individuals were also made. Table 1 shows information about the sequence reads and statistics from the eight different individual Trinity assemblies. Functional annotation of contigs from the different assemblies was conducted using BLAT (Kent, 2002) with default settings against the TAIR database (version 10 represented gene model from 2011-01-03; Lamesch et al., 2012) and MapMan hierarchical categories (Ath_AGI_LOCUS_TAIR10_Aug2012: http://mapman.gabipd.org/web/guest/mapmanstore). Mean contig length ranged from 635–742 bp, N50 value (median contig length) from 73,720 to 73,820, E90N50 value from 1099–1307 bp (which is computed with the refN50_statistic.pl script of the Trinity package and represents the N50 of 90% of the expressed transcripts), and number (and percentage) of contigs with positive BLAST hits from 24,583–39,915 (42–63%). To demonstrate the quality and utility of the transcriptomic resources developed here, we compared the transcriptome sequences of our eight sampled individuals relative to each other and to two outgroup species, *Picrorhiza kurrooa* (http://scbb.ihbt.res.in/Picro_information/; SRR392742; Gahlan et al., 2012) and *Plantago ovata* (SRR629688; Kotwal et al., 2016). To do this, we mined the data from these 10 individuals for SNPs using Site Identification from Short Read Sequences (SISRS) version 1.0 (Schwart et al., 2015; https://github.com/rachels/SISRS/releases). SISRS identifies SNPs for phylogenetic studies directly from raw high-throughput sequence data without a reference genome and without a priori knowledge of potentially informative loci. Briefly, SISRS first assembles raw sequence reads into a “composite genome” using Velvet, maps the raw reads and individual contigs against this composite genome with Bowtie 2, and then calls SNPs with a Python script (Schwart et al., 2015). SNP discovery was performed using SISRS on four different data sets: (1) all eight Veronica individuals combined plus *P. kurrooa* and *P. ovata* as outgroups, (2) all eight Veronica individuals only, (3) the four New Zealand individuals only, and (4) the four European individuals only. The SNP data were converted to Nexus format and analyzed using NeighborNet networks (SplitsTree version 4.14.2; Huson and Bryant, 2006). In addition, GARLI version 2.01.067 (Zwickl, 2006) was used for phylogenetic tree reconstruction under maximum likelihood. We first performed a GARLI run with 10 replicates to estimate the parameter values for the model of evolution with the ModelTest version 2.1.5 (012010F; Drrrha et al., 2012) [setting ratematrix = a b c a b a a statefrequencies = estimate]; six of the 10 replicates had the same best lnL score. These estimated model parameters were then fixed for a bootstrap analysis, which was performed with 1000 replicates [parametervaluestring = M1 r 1.000000 7.301633 1.614222 1.000000 7.301633 1.000000 0.272331 0.22561 0.22541

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

3 of 24
Marker development—Two different types of markers were developed from the *Veronica* transcriptome resources generated here, LCN and SSR markers.

Low-copy nuclear markers—MarkerMiner was used with default settings to identify LCN markers from a curated set of conserved ortholog set (COS) loci (De Smet et al., 2017). MarkerMiner was developed and tested using transcriptome assemblies from 77 Lamiales species (including six from Plantaginaceae; Chamala et al., 2015), and uses a reciprocal BLAST of all transcriptomes with one another and to the reference *A. thaliana* genome. *Arabidopsis thaliana* (Brassicales) is the phylogenetically closest reference available in MarkerMiner to *Veronica* (Lamiales). Of the 1228 loci returned, 73 were classified as being “strictly” and 1155 as “mostly” single copy. MAFFT alignments of the 330 loci found in six or more individuals, of which 15 were “strictly” and 314 “mostly” single copy, were used to develop primers in Geneious (version 8.7) with Primer3 (Untergasser et al., 2012), aiming for a melting temperature of 60°C. Loci were checked manually for large introns in *Geneious* by comparing the MarkerMiner alignment to *A. thaliana*. We chose 13 “strictly” single-copy loci with a successful primer search and 35 additional “mostly” single-copy loci with successful primer searches such that all five *A. thaliana* chromosomes were equally represented in this marker set. These 48 loci were validated using Fluidigm microfluidic PCR and Illumina MiSeq amplicon sequencing of 48 individuals representing 20 Australasian species and one interspecific hybrid (Appendix 2). The combination of this technique with Illumina MiSeq amplicon sequencing of 300-bp paired-end reads has proven useful and highly efficient in recent studies for development of novel and effective nuclear sequencing markers and improving understanding of phylogenetic relationships in nonmodel genera (Gostel et al., 2015; Uribe-Convers et al., 2016).

Simple sequence repeats—Numerous SSRs were identified from Trinity assemblies of the New Zealand individuals only using QDD version 3.1 (Meglécz et al., 2014; Table 1). Settings for the search were a length of 250–350 bp of the locus and primer melting temperatures of 59–61°C. After filtering for quality (taking QDD categories A and B), repeats (removing dinucleotides for example), and length of predicted PCR product, 48 loci were chosen from the 1124 potential SSRs with primer sites found by QDD. These were validated using Fluidigm microfluidic PCR and Illumina MiSeq amplicon sequencing (see above) of 48 individuals representing 20 Australasian species and one interspecific hybrid (Appendix 3). For each SSR marker, which included the SSR repeat area and flanking regions, resulting sequences were analyzed in the same way as the LCN data (see above) and examined in Geneious and GARLI regarding SSR motif, sequence length, number of individuals successfully sequenced, number of alleles sequenced, and pairwise genetic distance. In addition, for one randomly chosen example SSR locus, the alignment was exported to GARLI and a phylogeny was reconstructed using the same settings as described above for SNPs.

RESULTS

Transcriptomes—Functional annotation of individual assemblies was similar for each of the eight individuals, with gene ontology (GO) terms assigned to 13,009–14,271 contigs (19–33%; Table 1). There was large overlap of annotated contigs of the different assemblies whether looking at assemblies of individuals of New Zealand species only (26,524 or 89.4% shared annotated contigs; Fig. 1A), European species only (25,456 or 87.8%; Fig. 1B), or all New Zealand vs. all European species (29,839 or 94.3%; Fig. 1C). On the other hand, individual species had 114–453 (0.4–1.6%) unique annotated contigs relative to other species from the same geographical area, and

![Fig. 1](http://www.bioone.org/loi/apps) Venn diagrams showing the number of annotated contigs from the *Veronica* Trinity assemblies. (A) Four New Zealand individuals. (B) Four European individuals. (C) All New Zealand vs. all European individuals.
the numbers for New Zealand and European species were comparatively very similar (compare Fig. 1A and 1B). Within the New Zealand species, V. hectorii Hook. f. and V. ochracea (Ashwin) Garn.-Jones shared the most unique annotated contigs (234 or 0.8%) relative to the other five species pairs, whereas V. catarractae G. Först. and V. ochracea shared the fewest (110 or 0.4%; Fig. 1A). Within the European species, the species pair with the most unique shared annotated contigs was V. panormitana Tineo ex Guiss. (2x) and V. cymbalaria Bodard (6x) (238 or 1.1%), whereas the two diploids V. panormitana and V. trichadena shared the fewest (53 or 0.2%) (Fig. 1B).

GO term results were also very similar; of 35 GO categories, the number of unique transcripts were largely overlapping for all species pairs, as is shown for the most divergent species pair of the eight transcriptomes sequenced (i.e., New Zealand V. hectorii vs. European V. panormitana; Fig. 2). The GO categories with the largest numbers of unique transcripts (ca. 500–3000) for these Veronica leaf transcriptomes were (from highest to lowest) “not assigned,” “proteins,” “RNA,” “signaling,” “transcription,” “mRNA,” “cell,” and “DNA” (Fig. 2A).

SNP discovery using SISRS resulted in the following number of SNPs and potential PICs: 10-individual data set (29,738 SNPs, 8746 PICs), eight Veronica individuals only (45,751 SNPs, 4021 PICs), four New Zealand individuals only (41,167 SNPs, 2302 PICs), and four European individuals only (65,278 SNPs, 1735 PICs). When the 10-individual data set was analyzed using SplitsTree (Fig. 3A–C), the NeighborNet network clearly showed a main split between all New Zealand species, V. hectorii (New Zealand hexaploid, ○) based on the larger similarity with that species compared with V. trichadena (Fig. 1; 328 vs. 132 unique annotated contigs).

**Marker development**

Low-copy nuclear markers—A range of 3–44 (average: 23.4, median: 22) of 48 individuals were successfully sequenced for each of the 48 loci, with 22 of 48 (46%) loci successfully amplifying in at least 24 (>50%) individuals (Appendix 4). For each individual, 4–40 loci were successfully amplified (average: 23.4, median: 21.5), and again less than half of the individuals (22/48, 46%) had successful amplification of at least 25 (>50%) loci (data not shown). Only one-quarter (11/48) of the loci aligned well with the corresponding transcript; these loci had mean lengths of 327–480 bp, contained large numbers of SNPs and PICs, and BLASTed to known A. thaliana genes (Appendix 4).

Figure 4 shows an alignment and GARLI tree of sequences from 22 of the 42 individuals successfully sequenced for one randomly chosen example locus, LCN-04 (two outgroups plus 10 European and 10 New Zealand Veronica individuals/species). Nearly twice as many different sequences were generated for the 10 New Zealand individuals shown here (27 sequences; 6x or 18x, “V. townsonii E6 18" to “V. melanocaulon D5 11” in the tree) relative to the 10 European individuals (14 sequences; 2x or 4x, “V. missurica E2 13” to “V. chamaedrys C3 10”). Of the 10 New Zealand individuals, five have only one sequence and are all in the same clade (V. albitflora (Pennell) Albach, V. cupressoides Hook. f., V. densifolia F. Muell., V. lavaudiana Raoul, and V. senex (Garn.-Jones) Garn.-Jones), whereas the other five have 2–8 sequences that fall into the first clade or one
of two other clades (Fig. 4). As another example highlighting the low-copy nature of the loci that were sequenced, locus LCN-38 has two orthologous copies, which is expected due to the categorization of the *A. thaliana* gene AT3G59380 as “mostly” single copy (data not shown; comparisons done using MarkerMiner 1.0 [Chamala et al., 2015]). Additional phylogenetic analyses of the other LCN loci are outside the scope of this study and will be performed elsewhere (Meudt et al., unpublished data).

**Simple sequence repeats**—Overall, 3–47 (mean: 37.7, median: 44.5) of 48 individuals were successfully sequenced for each of the 48 SSR loci, including 40 of 48 loci that were successfully sequenced for at least 26 (>50%) individuals (Appendix 3). For each individual, 0–43 loci were successfully sequenced (average: 37.7, median: 40), with all but two individuals with at least 29 (>60%) loci successfully sequenced (individuals *V. catarractae* B1 and *V. colostylis* H3 failed for all 48 and 40 loci, respectively). In general, sequences ranged from 98–851 bp in length (average: 324) and contained one or more length- and/or sequence-variable SSR motifs as well as flanking SNPs and indels within and among individuals (e.g., Fig. 5). Number of sequenced alleles (which are supported by at least 10 raw sequencing reads) per individual ranged from 1–39 (mean: 4.32, median: 3.0, n = 47), with the lower polyploids having fewer alleles than the higher polyploids (6x, mean: 3.96, n = 37; 12x, mean: 5.25, n = 5; 18x, mean: 6.09, n = 5).

As the focus of SSRs is often population genetics, we analyzed two subsets of the larger SSR data set in more detail, i.e., eight individuals of *V. chionohebe* Garn.-Jones (4), *V. trifida* Petrie (2), and their interspecific hybrid (2) (all 2n = 42) (Appendix 5), and six individuals of *V. thomsonii* Cheeseman (2n = 42), respectively (Appendix 6). For all loci in the two subsets, sequences were on average of 317–327 bp, with 1–26 alleles (mean: 4.0–4.3), 54–80 SNPs, and 41.7–52.5 PICs (see “Totals” rows in Appendix 5 and 6). Figure 5 shows an alignment of 54 different SSR sequences from one locus (SSR-08) of the eight-individual *V. chionohebe/V. trifida* subset. In loci SSR-08, the sequences ranged from 311–387 bp (average: 357 bp). The sampled individuals had on average 6.8 alleles, and individuals of *V. chionohebe* had half as many unique alleles (3–6 each) as individuals of *V. trifida* and the interspecific hybrid (8–10). The sequences of locus SSR-08 were highly variable (note the many colored bars in the alignment in Fig. 5), with 126 SNPs and 109 PICs, and 0–0.14 pairwise genetic distances (mean and median: 0.08) (Appendix 5). In the phylogenetic tree, there is support for some taxonomic clustering of sequences of *V. chionohebe* and *V. trifida*, respectively, with hybrid sequences in highly supported clades with *V. chionohebe* or *V. trifida* in three vs. four cases, respectively (see tree in Fig. 5). Additional analyses of the other SSR loci are outside the scope of this study and will be performed elsewhere (Meudt et al., unpublished data).

**DISCUSSION**

The development of transcriptomic and genomic resources and variable genetic markers in so-called natural “mesopolyploid” species radiations is key to addressing fundamental questions about polyploidy and diversification. For polyploids, functional genomic resources in particular are important to facilitate the study of gene evolution. *Veronica* is an example of a natural mesopolyploid species radiation that to date has lacked such genomic and genetic resources, and this has hindered progress in studying polyploid evolution at the population, species, and generic levels. The transcriptomic and genetic resources developed here will make further detailed studies regarding the role of polyploidy in adaptation and species diversification in *Veronica* possible.

In the current study, we sequenced and assembled leaf transcriptomes from eight individuals representing seven species of *Veronica* from polyploid species radiations in Europe and New Zealand. There was high overlap of annotated contigs (Fig. 1) and GO terms (Fig. 2) among the eight individuals, as well as good phylogenetic resolution in the network and phylogenetic analyses of SNPs generated using SISRS (Fig. 3). An outstanding challenge with de novo transcriptome assemblies of polyploids...
is differentiating homoeologs from orthologs; however, this was not an issue for developing markers in polyploid Veronica from our transcriptome assemblies, as phylogenetic relationships (Fig. 3) are consistent with hypothesized relationships and previous phylogenetic results. Such results demonstrate the utility of these transcriptomic resources for phylogenetic studies, functional analyses across the genus using reverse transcription PCR, or for further comparative transcriptomic analyses of the sampled natural allopolyploids and their diploid parental species in the two main centers of diversity for Veronica (i.e., Europe and New Zealand). The large number of transcripts unique to hexaploid V. cymbalaria (453) relative to other individuals representing species from which it likely derived (V. trichadena: 114 and V. panormitana: 256) is surprising and opens the door to studies of differential expression and functional differentiation of genes in polyploids. Common garden experiments are also planned, which will allow comparison of other individuals with the eight sequenced here.

Furthermore, the SSR and LCN genetic markers developed here from the transcriptomes, and validated using microfluidic PCR and high-throughput sequencing, are highly variable and will be extremely useful in future phylogenetic studies of Veronica, as a whole, as well as studies at the interface of intra- and interspecific levels regarding questions of population genetics, species limits, and relationships of closely related species in New Zealand Veronica. Additionally, challenges presented by genotyping SSRs in polyploids, such as determining allele dosage and unambiguously identifying alleles (Pfeiffer et al., 2011), are overcome by sequencing the SSRs and their flanking regions, which we would recommend for future sequencing markers (as opposed to being genotyped) at the interface of intra- and interspecific levels regarding questions of population genetics, species limits, and relationships of closely related species in New Zealand Veronica. In addition, of the 1124 SSRs identified in the four New Zealand Veronica individuals, we validated 48 in 48 Southern Hemisphere Veronica individuals, 40 of which were successfully sequenced for >50% of individuals. Sequenced SSRs and their flanking regions were on average 324 bp long, contained numerous SNPs and PICs, and had mean pairwise genetic distances of 0.01–0.18. The variation seen, particularly in the flanking regions of the sequenced SSRs, is equal to or much greater than that from previous studies using standard DNA sequencing techniques of the sequenced SSRs, is equal to or much greater than that from previous studies using standard DNA sequencing and genotyping markers (e.g., Wagstaff et al., 2002; Meudt and Bayly, 2008). These 40 SSRs have great potential as highly variable sequencing markers (as opposed to being genotyped) at the interface of intra- and interspecific levels regarding questions of population genetics, species limits, and relationships of closely related species in New Zealand Veronica. In addition to the potential advances for Veronica, our methodological approach may also be useful for other natural polyploid groups that lack genomic or genetic resources. Natural species that are not associated with economically important crop or other "model" species often lack genomic resources and are very limited regarding the availability of variable genetic markers. Furthermore, developing and establishing such markers using

http://www.bioone.org/loi/apps
and MiSeq amplicon sequencing of LCN and SSR markers, which were designed in MarkerMiner and QDD from transcriptomic data, is a relatively straightforward high-throughput marker validation method as well as an analysis pipeline that can be used on other natural (and polyploid) systems.

LITERATURE CITED

Abbott, R. J., D. C. Albach, S. W. Ansell, J. W. Arentzen, S. J. E. Baird, N. Bierne, J. Bougman, et al. 2013. Hybridization and speciation. Journal of Evolutionary Biology 26: 229–246.

AfGan, E., D. Baker, M. van den Beek, D. Blankenberg, D. Bouvier, M. Čech, J. Chilton, et al. 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Research 44: W3–W10.

Albach, D. C. 2007. Amplified fragment length polymorphisms and sequence data in the phylogenetic analysis of polyploids: Multiple origins of Veronica cymbalaria (Plantaginaceae). New Phytologist 176: 481–498.

Albach, D. C., and J. Greilhuber. 2004. Genome size variation and evolution in Veronica. Annals of Botany 94: 897–911.

Albach, D. C., M. M. Martínez-Ortega, L. Delgado, H. Weiss-Schneeweiss, F. Özgökcé, and M. A. Fischer. 2008. Chromosome numbers in Veronicaceae (Plantaginaceae): Review and several new counts. Annals of the Missouri Botanical Garden 95: 543–566.

Fig. 5. MAFFT alignment and GARLI phylogenetic tree (visualized in Geneious) of 54 sequences for a subset of eight New Zealand Veronica individuals of V. chionohebe, V. trifida, and their interspecific hybrid from two South Island locations from sequences of SSR locus SSR-08 mined using QDD from Trinity assemblies of leaf transcriptome data. Consensus and identity sequences are shown at the top. Base pairs that are identical to the consensus are shown in gray, whereas SNPs are shown as colors (red = A, blue = C, green = T, yellow = G, black = N). Each of the eight individuals has a unique color: three individuals of V. chionohebe (orange, red, and brown), two of V. trifida (blue, pink), and two of their hybrid (light and dark green). For each sequence in the alignment, species names are followed by location (Garvie Mountains or Pisa Range), sequencing plate location (A5, B5, C4, D4, E4, F4, G4, or H4), and number of sequence reads supporting that allele (range: 12–187). Green branches in the GARLI tree to the left of the individual names have >80% bootstrap support (see Fig. 3 for GARLI settings). Voucher information is shown in Appendix 3.

Table 1. MAFFT alignment and GARLI phylogenetic tree (visualized in Geneious) of 54 sequences for a subset of eight New Zealand Veronica individuals of V. chionohebe, V. trifida, and their interspecific hybrid from two South Island locations from sequences of SSR locus SSR-08 mined using QDD from Trinity assemblies of leaf transcriptome data. Consensus and identity sequences are shown at the top. Base pairs that are identical to the consensus are shown in gray, whereas SNPs are shown as colors (red = A, blue = C, green = T, yellow = G, black = N). Each of the eight individuals has a unique color: three individuals of V. chionohebe (orange, red, and brown), two of V. trifida (blue, pink), and two of their hybrid (light and dark green). For each sequence in the alignment, species names are followed by location (Garvie Mountains or Pisa Range), sequencing plate location (A5, B5, C4, D4, E4, F4, G4, or H4), and number of sequence reads supporting that allele (range: 12–187). Green branches in the GARLI tree to the left of the individual names have >80% bootstrap support (see Fig. 3 for GARLI settings). Voucher information is shown in Appendix 3.

| Sequence | Identity | Voucher Information |
|----------|----------|---------------------|
| V. chionohebe | | |
| V. trifida | | |
| V. chionohebe x V. trifida | | |
| V. trifida x V. chionohebe | | |
| V. trifida x V. chionohebe | | |
| V. trifida x V. chionohebe | | |
| V. trifida x V. chionohebe | | |
| V. trifida x V. chionohebe | | |
| V. trifida x V. chionohebe | | |
| V. trifida x V. chionohebe | | |
| V. trifida x V. chionohebe | | |
| V. trifida x V. chionohebe | | |
| V. trifida x V. chionohebe | | |

Fig. 6. MAFFT alignment and GARLI phylogenetic tree (visualized in Geneious) of 54 sequences for a subset of eight New Zealand Veronica individuals of V. chionohebe, V. trifida, and their interspecific hybrid from two South Island locations from sequences of SSR locus SSR-08 mined using QDD from Trinity assemblies of leaf transcriptome data. Consensus and identity sequences are shown at the top. Base pairs that are identical to the consensus are shown in gray, whereas SNPs are shown as colors (red = A, blue = C, green = T, yellow = G, black = N). Each of the eight individuals has a unique color: three individuals of V. chionohebe (orange, red, and brown), two of V. trifida (blue, pink), and two of their hybrid (light and dark green). For each sequence in the alignment, species names are followed by location (Garvie Mountains or Pisa Range), sequencing plate location (A5, B5, C4, D4, E4, F4, G4, or H4), and number of sequence reads supporting that allele (range: 12–187). Green branches in the GARLI tree to the left of the individual names have >80% bootstrap support (see Fig. 3 for GARLI settings). Voucher information is shown in Appendix 3.

Table 2. MAFFT alignment and GARLI phylogenetic tree (visualized in Geneious) of 54 sequences for a subset of eight New Zealand Veronica individuals of V. chionohebe, V. trifida, and their interspecific hybrid from two South Island locations from sequences of SSR locus SSR-08 mined using QDD from Trinity assemblies of leaf transcriptome data. Consensus and identity sequences are shown at the top. Base pairs that are identical to the consensus are shown in gray, whereas SNPs are shown as colors (red = A, blue = C, green = T, yellow = G, black = N). Each of the eight individuals has a unique color: three individuals of V. chionohebe (orange, red, and brown), two of V. trifida (blue, pink), and two of their hybrid (light and dark green). For each sequence in the alignment, species names are followed by location (Garvie Mountains or Pisa Range), sequencing plate location (A5, B5, C4, D4, E4, F4, G4, or H4), and number of sequence reads supporting that allele (range: 12–187). Green branches in the GARLI tree to the left of the individual names have >80% bootstrap support (see Fig. 3 for GARLI settings). Voucher information is shown in Appendix 3.
Applications in Plant Sciences 2016 4(10): 1600091 Mayland-Quellhorst et al.—Veronica transcriptomic resources
doi:10.3732/apps.1600091

MAYLAND-QUELLHORST et al.—Veronica transcriptomic resources...
Thiers, B. 2016 [continuously updated]. Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden’s Virtual Herbarium. Website http://sweetgum.nybg.org/science/ih [accessed 14 September 2016].

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.

Uribe-Convers, S., M. L. Settles, and D. C. Tank. 2016. A phylogenomic approach based on PCR target enrichment and high throughput sequencing: Resolving the diversity within the South American species of *Bartsia* L. (Orobanchaceae). *PLoS ONE* 11: e0148203.

Wagstaff, S. J., and P. Wardle. 1999. Whipcord hebes—Systematics, distribution, ecology and evolution. *New Zealand Journal of Botany* 37: 17–39.

Wagstaff, S. J., M. J. Bayly, P. J. Garnock-Jones, and D. C. Albach. 2002. Classification, origin and diversification of the New Zealand hebes (Scrophulariaceae). *Annals of the Missouri Botanical Garden* 89: 38–63.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin, Austin, Texas, USA.
APPENDIX 1. Information about the eight individuals of Veronica sampled for RNA-Seq.

| Speciesa | GPS coordinates | Chromosome number (Ploidy)b | 1C-value (pg)c | Collection locality and collection no. (Voucher)d | RNA 260/280 ratioe | RNA conc. (ng/µL)f | RNA RINf |
|-----------|----------------|-----------------------------|----------------|-----------------------------------------------|-------------------|-------------------|---------|
| Veronica catarractae G. Forst. | NA (cultivated plant) | 2n = 42 (6x) | 1.06 | Cult. Botanischer Garten Oldenburg (Germany), ex New Zealand, Meudt s.n. (OLD00026) | 2.11 | 1017.00 | 7.90 |
| V. hectorii Hook. f. subsp. coarctata (Cheeseman) Gam.-Jones | NA (cultivated plant) | 2n = 40 (6x) | 1.07 | Cult. Botanischer Garten Bonn 1342 (Germany), ex New Zealand, Meudt s.n. (OLD00029) | 1.94 | 121.00 | 7.10 |
| V. planopetiolata G. Simpson & J. S. Thomson | 44.52247°S, 168.6736916667°E | 2n = 84 (12x) | 2.45 | New Zealand; South Island, Otago, Meudt HMM339a (WELT SP091593) | 1.93 | 147.00 | 7.00 |
| V. ochracea (Ashwin) Gam.-Jones | NA (cultivated plant) | 2n = 124 (18x) | 2.97 | Cult. Botanischer Garten Bonn 9509 (Germany), ex New Zealand, Meudt s.n. (OLD00071) | 2.12 | 1327.00 | 6.80 |
| V. panormitana Tineo ex Guss. | 36.6672°N, 31.8989°E | 2n = 18 (2x) | 0.36 | Turkey: north of Paravallar, Albach 1114 & 5272 (OLD00214) | 2.00 | 53.00 | 8.00 |
| V. trichadena Jord. & Fourn. | 39.678536°N, 2.80062°E | 2n = 18 (2x) | 0.39 | Spain: Mallorca, Meudt HMM346L (OLD00086) | 1.98 | 302.00 | 7.50 |
| V. cymbalaria Bodard | 36.5325°N, 31.99°E | 2n = 36 (4x) | 0.76 | Turkey: Alanya Castle, Albach 1235 (OLD01171) | 2.04 | 245.00 | 6.90 |
| V. cymbalaria | 37.22778°N, 31.12972°E | 2n = 54 (6x) | 1.38 | Turkey: Analaya, Selgedos, Albach 1087 & 5300 (OLD00481) | 2.11 | 1265.00 | 7.60 |

Note: NA = not applicable.

a RNA was extracted from leaf material from greenhouse-grown material of all individuals except V. planopetiolata, which was from field-collected leaf material stored in RNAlater (Life Technologies, Carlsbad, California, USA).
b Chromosome numbers are from the literature (Albach et al., 2008).
c 1C-values (Meudt et al., 2015) were assessed for the same individual from which RNA was extracted for this study except for V. panormitana, whose 1C-value is based on the average of five other individuals from three different Turkish populations (range 0.35–0.37 pg; Meudt et al., 2015).
d Voucher specimens are lodged at herbaria at the Museum of New Zealand Te Papa Tongarewa (WELT) or Carl-von-Ossietzky Universität Oldenburg (OLD).
e RNA 260:280 ratio was calculated using the Tecan Infinite Pro F200 (Tecan, Crailsheim, Germany).
f RNA concentration and RNA Integrity Number (RIN) were calculated using the Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany); please note that the cDNA construction was made with normalized RNA.
## Appendix 2. Information about the 48 individuals of *Veronica* sampled for the LCN marker validation.

| Species                              | Subgenus       | Ploidya | Chromosome no. | Location on sequencing plate | No. of LCN markers successfully sequenced (of 48 total) |
|--------------------------------------|----------------|---------|----------------|-----------------------------|-------------------------------------------------------|
| *Lagotis integrifolia* (Willd.) Schisch. ex Vikulova | (outgroup) | 4       | 44             | Kazakhstan                 | 20                                                   |
| *Paederota lutea* L. f. (outgroup)   |                | 4       | 36             | Austria                    | 24                                                   |
| *Veronicastrum stenostachyum* (Hemsl.) T. Yamaz. | (outgroup) | 4       | 34             | China                      | 25                                                   |
| *Wulfenia carinthiaca* Jacq. (outgroup) |                | 2       | 18             | Austria                    | 25                                                   |
| *Veronica anagallis-aquatica* L. Beccabunga |            | 2       | 18             | Czech Republic            | 18                                                   |
| *V. catenata* Pennell Beccabunga |            | 2       | 18             | Czech Republic            | 18                                                   |
| *V. gentianoides* Vahl Beccabunga |            | 2       | 18             | Georgia                    | 18                                                   |
| *V. arvensis* L. Chamaedrys |            | 2       | 16             | Germany                    | 16                                                   |
| *V. chamaedrys* L. Chamaedrys |            | 4       | 32             | Norway                     | 32                                                   |
| *V. crista-galli* Steven Cochlidiosperma |            | 2       | 18             | Georgia                    | 18                                                   |
| #*V. cymbalaria* Bodard Cochlidiosperma |            | 4       | 36             | Turkey                     | 36                                                   |
| #*V. cymbalaria* Cochlidiosperma |            | 6       | 54             | Turkey                     | 54                                                   |
| *V. javanica* Blume Cochlidiosperma |            | 2       | 16             | Murata et al.              | 16                                                   |
| #*V. panormitana* Tineo ex Guss. Cochlidiosperma |          | 2       | 18             | Turkey                     | 18                                                   |
| #*V. trichadena* Jord. & Fourr. Cochlidiosperma |        | 2       | 18             | Spain                      | 18                                                   |
| *V. triloba* (Opiz) Opiz |            | 2       | 18             | Turkey                     | 18                                                   |
| *V. brownii* Roem. & Schult. Labiatoides |          | 12      | 72             | Australia                  | 72                                                   |
| #*V. triphyllos* L. Pellidosperma |            | 2       | 14             | Russia                     | 14                                                   |
| *V. cuneifolia* D. Don |            | 2       | 16             | Turkey                     | 16                                                   |
| #*V. fuhsii* Freyn & Sint. Pentasepalae |      | ?       | 32             | Turkey                     | 32                                                   |
| #*V. prostrata* L. Pentasepalae |            | 2       | 16             | Austria                    | 16                                                   |
| *V. filiformis* Sm. Pocilla |            | 2       | 14             | Germany                    | 14                                                   |
| #*V. longifolia* L. Pseudolysimachium |          | 4       | 34             | Turkey                     | 34                                                   |
| #*V. longifolia* Pseudolysimachium |            | 4       | 34             | Russia                     | 34                                                   |
| *V. schmidtiana* Regel Pseudolysimachium |           | 4       | 34             | Japan                      | 34                                                   |
| #*V. spicata* L. |            | 8       | 68             | Japan                      | 68                                                   |
| #*V. trichophylla* L. Primula |            | 4       | 24             | Spain                      | 24                                                   |
| #*V. trilobiata* L. Primula |            | 4       | 24             | Spain                      | 24                                                   |
| #*V. triloba* (Opiz) Opiz |            | 2       | 18             | Turkey                     | 18                                                   |
| #*V. brownii* Roem. & Schult. Labiatoides |          | 12      | 72             | Australia                  | 72                                                   |
| #*V. triphyllos* L. Pellidosperma |            | 2       | 14             | Germany                    | 14                                                   |
| #*V. cuneifolia* D. Don |            | 2       | 16             | Turkey                     | 16                                                   |
| #*V. fuhsii* Freyn & Sint. Pentasepalae |      | ?       | 32             | Turkey                     | 32                                                   |
| #*V. prostrata* L. Pentasepalae |            | 2       | 16             | Austria                    | 16                                                   |
| *V. filiformis* Sm. Pocilla |            | 2       | 14             | Germany                    | 14                                                   |
| #*V. longifolia* L. Pseudolysimachium |          | 4       | 34             | Turkey                     | 34                                                   |
| #*V. longifolia* Pseudolysimachium |            | 4       | 34             | Russia                     | 34                                                   |
| *V. schmidtiana* Regel Pseudolysimachium |           | 4       | 34             | Japan                      | 34                                                   |
| #*V. spicata* L. |            | 8       | 68             | Japan                      | 68                                                   |
| #*V. trichophylla* L. Primula |            | 4       | 24             | Spain                      | 24                                                   |
| #*V. trilobiata* L. Primula |            | 4       | 24             | Spain                      | 24                                                   |
| #*V. triloba* (Opiz) Opiz |            | 2       | 18             | Turkey                     | 18                                                   |
| #*V. brownii* Roem. & Schult. Labiatoides |          | 12      | 72             | Australia                  | 72                                                   |
| #*V. triphyllos* L. Pellidosperma |            | 2       | 14             | Germany                    | 14                                                   |
| #*V. cuneifolia* D. Don |            | 2       | 16             | Turkey                     | 16                                                   |
| #*V. fuhsii* Freyn & Sint. Pentasepalae |      | ?       | 32             | Turkey                     | 32                                                   |
| #*V. prostrata* L. Pentasepalae |            | 2       | 16             | Austria                    | 16                                                   |
| *V. filiformis* Sm. Pocilla |            | 2       | 14             | Germany                    | 14                                                   |
| #*V. longifolia* L. Pseudolysimachium |          | 4       | 34             | Turkey                     | 34                                                   |
| #*V. longifolia* Pseudolysimachium |            | 4       | 34             | Russia                     | 34                                                   |
| *V. schmidtiana* Regel Pseudolysimachium |           | 4       | 34             | Japan                      | 34                                                   |
| #*V. spicata* L. |            | 8       | 68             | Japan                      | 68                                                   |
| #*V. trichophylla* L. Primula |            | 4       | 24             | Spain                      | 24                                                   |
| #*V. trilobiata* L. Primula |            | 4       | 24             | Spain                      | 24                                                   |
| #*V. triloba* (Opiz) Opiz |            | 2       | 18             | Turkey                     | 18                                                   |

http://www.bioone.org/loi/apps
## Appendix 2. Continued.

| Species                        | Subgenus      | Ploidy | Chromosome no. | Country     | Voucher (Herbarium and/or Herbarium accession no.) | Location on sequencing plate | No. of LCN markers successfully sequenced (of 48 total) |
|--------------------------------|---------------|--------|----------------|-------------|---------------------------------------------------|------------------------------|-------------------------------------------------------|
| #V. planopetiolata G. Simpson & J. S. Thomson | Pseudoveronica | 12     | 84             | New Zealand | Meudt HMM39a (WELT SP01595)                      | C5                          | 22                                                    |
| V. senex (Garn.-Jones) Garn.-Jones     | Pseudoveronica | 6      | 42             | New Zealand | Garnock-Jones PGJ 2879 (OLD)                      | E5                          | 28                                                    |
| V. spectosa R. Cunn. ex A. Cunn.      | Pseudoveronica | 6      | 40             | New Zealand | Garnock-Jones PGJ 2878 (OLD)                      | F6                          | 33                                                    |
| V. tairawhiti (B. D. Clarkson & Garn.-Jones) Garn.-Jones | Pseudoveronica | 12     | 80             | New Zealand | Garnock-Jones PGJ 2888 (OLD)                      | G6                          | 9                                                     |
| V. townsonii Cheeseman              | Pseudoveronica | 6      | 40             | New Zealand | Garnock-Jones PGJ 2901 (WELT SP103482)           | E6                          | 26                                                    |

**Note:** LCN = low-copy nuclear.

*Ploidy and chromosome numbers are from the literature (Albach et al., 2008).

b Herbaria acronyms follow Thiers (2016).

#RNA-Seq sample.
## Appendix

### 3. Validation of 48 SSR markers on 48 individuals of 20 species of Southern Hemisphere Veronica subg. Pseudoveronica

| Species name                      | Section and informal group                  | Ploidy | Chromosome n<sup>a</sup> | Country | GenBank accession no. | Voucher and collection locality | Location on sequencing plate | No. SSR loci successfully sequenced (of 48 total) |
|-----------------------------------|--------------------------------------------|--------|--------------------------|---------|-----------------------|---------------------------------|-----------------------------|---------------------------------------------|
| Veronica calycina R. Br. sect. Labiatoides | 6                                           | 36     | Australia                | D3      | MB 1833, Pin Ranges (WELT) | RGC 19644, near Lithgow, NSW    | C3                          | 38                                           |
| Veronica derwentiana Andrews subsp. subglauca (B. G. Briggs & Ehrend.) sect. Hebe, snow hebe | 6                                           | 42     | Australia                | E4      | SP084029               | RGC 19649, near Lithgow, NSW    | F4                          | 38                                           |
| Veronica chionohebe Garn.-Jones sect. Hebe, snow hebe | 6                                           | 42     | New Zealand              | C4      | SP084043               | MJB 1823, Pisa Range (WELT)     | A4                          | 40                                           |
| Veronica chionohebe × Veronica trifida Petrie sect. Hebe, snow hebe | 6                                           | 42     | New Zealand              | C5      | SP084046               | MJB 1844, Garvie Mountains (WELT) | B4                          | 39                                           |
| Veronica ciliolata (Hook. f.) Garn.-Jones subsp. ciliolata | 6                                           | 42     | New Zealand              | D5      | SP084050               | MJB 1845, Garvie Mountains (WELT) | E4                          | 40                                           |
| Veronica ciliolata subsp. fiordensis (Ashwin) Meudt sect. Hebe, snow hebe | 6                                           | 42     | New Zealand              | E5      | SP084051               | MJB 1848, Garvie Mountains (WELT) | G4                          | 38                                           |
| Veronica densifolia F. Muell. sect. Hebe, snow hebe | 6                                           | 42     | New Zealand              | F5      | SP084052               | MJB 1849, Garvie Mountains (WELT) | H4                          | 40                                           |
| Veronica pulvinaris (Hook. f.) Cheseman sect. Hebe, snow hebe | 6                                           | 42     | New Zealand              | G5      | SP084053               | MJB 1850, Hauer Hills (WELT)    | A6                          | 40                                           |
| Veronica thomsonii Cheeseman sect. Hebe, snow hebe | 6                                           | 42     | New Zealand              | H5      | SP084054               | MJB 1851, Garvie Mountains (WELT) | C6                          | 38                                           |
| Veronica thomsonii sect. Hebe, snow hebe | 6                                           | 42     | New Zealand              | I5      | SP084055               | MJB 1852, Garvie Mountains (WELT) | D6                          | 40                                           |
| Veronica trigida P. McNeill sect. Hebe, snow hebe | 6                                           | 42     | New Zealand              | H6      | SP084056               | MJB 1853, Garvie Mountains (WELT) | E6                          | 38                                           |
| Veronica brachysiphon (Summerh.) Bean as Veronica vernicosa sect. Hebe, snow hebe | 18                                           | 120    | New Zealand              | I6      | SP083931               | HMM s.n., cult. Kew Gardens     | F6                          | 40                                           |
| Veronica catarractae G. Forst. sect. Hebe, snow hebe | 6                                           | 42     | New Zealand              | J6      | SP083932               | HMM 259, Mt. St. Bathans (WELT) | G6                          | 40                                           |

---

<http://www.bioone.org/loi/apps>
### Appendix 3. Continued.

| Species name                                      | Section and informal group | Ploidy | Chromosome no. | Country | Voucher and collection locality (Herbarium and/or Herbarium accession no.) | Location on sequencing plate | No. SSR loci successfully sequenced (of 48 total) |
|--------------------------------------------------|-----------------------------|--------|----------------|---------|-----------------------------------------------------------------|-----------------------------|--------------------------------------------------|
| #V. catarractae (purchased as Parahebe 'Snow')    | sect. Hebe, speedwell hebe  | 6      | 42             | New Zealand | HMM s.n., cult. Botanischer Garten Oldenburg (OLD00026) | A1                          | 41                                               |
| V. colostylis Garn.-Jones                         | sect. Hebe, speedwell hebe  | 6      | 42             | New Zealand | HMM338a, Arrowtown (WELT SP091592)                             | H3                          | 8                                                |
| V. colostylis                                     | sect. Hebe, speedwell hebe  | 6      | 42             | New Zealand | HMM341c, Moke Creek (WELT SP091595)                             | G3                          | 30                                               |
| V. hectorii Hook. f.                             | sect. Hebe, hebe            | 6      | 40             | New Zealand | PGJ 2910, cult. Otari (WELT SP103460)                            | D1                          | 38                                               |
| #V. hectorii subsp. coarctata (Cheeseman) Garn.-Jones | sect. Hebe, hebe            | 6      | 40             | New Zealand | HMM s.n., Bonn 13428 (OLD00029)                                | C1                          | 41                                               |
| V. hulkeana F. Muell. ex Hook. f. subsp. evesita (Garn.-Jones) Garn.-Jones 'Lena' | sect. Hebe, sun hebe        | 6      | 42             | New Zealand | PGJ 2874, cult. Wellington (OLD)                                | A4                          | 32                                               |
| V. lavanduliana Raoul                            | sect. Hebe, sun hebe        | 6      | 42             | New Zealand | PGJ 2881, cult. Wellington (OLD)                                | B4                          | 41                                               |
| V. macrantha Hook. f.                            | sect. Hebe, unresolved, early branching | 6      | 42             | New Zealand | HMM s.n., cult. Kew Gardens 1969-35034 (OLD)                    | D2                          | 36                                               |
| V. macrantha                                     | sect. Hebe, unresolved, early branching | 6      | 42             | New Zealand | PGJ 2924, cult. Otari (WELT SP103475)                           | C2                          | 41                                               |
| #V. ochracea (Ashwin) Garn.-Jones                 | sect. Hebe, hebe            | 18     | 124            | New Zealand | HMM s.n., Bonn 9509 (OLD00071)                                 | E1                          | 42                                               |
| V. ochracea                                      | sect. Hebe, hebe            | 18     | 124            | New Zealand | PGJ 2911, cult. Otari (WELT SP103461)                           | F1                          | 36                                               |
| V. ochracea 'James Stirling'                     | sect. Hebe, hebe            | 18     | 124            | New Zealand | HMM s.n., cult. Kew Gardens 1992-1403 (OLD)                     | G1                          | 39                                               |
| V. odora Hook. f. (as Hebe vernicosa in Botanischer Garten Bonn) | sect. Hebe, hebe            | 12     | 84             | New Zealand | HMM s.n., cult. Bonn 17475 (OLD)                               | A3                          | 29                                               |
| V. odora 'New Zealand Gold'                      | sect. Hebe, hebe            | 12     | 84             | New Zealand | HMM s.n., cult. Kew Gardens 1989-2000 (OLD)                     | B3                          | 40                                               |
| #V. planopetiolata G. Simpson & J. S. Thomson    | sect. Hebe, speedwell hebe  | 12     | 84             | New Zealand | HMM339a, Shotover Saddle (WELT SP091593)                        | H1                          | 42                                               |
| V. planopetiolata                                | sect. Hebe, speedwell hebe  | 12     | 84             | New Zealand | HMM339b, Shotover Saddle (WELT SP091593)                        | A2                          | 36                                               |
| V. planopetiolata                                | sect. Hebe, speedwell hebe  | 12     | 84             | New Zealand | HMM339c, Shotover Saddle (WELT SP091593)                        | A1                          | 43                                               |
| V. salicornoides Hook. f.                        | sect. Hebe, hebe            | 6      | 42             | New Zealand | HMM s.n., cult. Kew Gardens 1989-2004 (OLD)                     | F2                          | 38                                               |
| V. salicornoides                                 | sect. Hebe, hebe            | 6      | 42             | New Zealand | PGJ 2923, cult. Otari (WELT SP103474)                           | E2                          | 42                                               |
| V. vernicosa Hook. f.                            | sect. Hebe, hebe            | 6      | 42             | New Zealand | PGJ 2925, cult. Otari (WELT SP103476)                           | E3                          | 41                                               |
| V. vernicosa                                     | sect. Hebe, hebe            | 6      | 42             | New Zealand | PGJ 2926, cult. Otari (WELT SP103477)                           | F3                          | 39                                               |

*Ploidy and chromosome numbers are from the literature (Albach et al., 2008).

bHerbaria acronyms follow Thiers (2016). Voucher specimens are lodged at herbaria at the Museum of New Zealand Te Papa Tongarewa (WELT), Carl-von-Ossietzky Universität Oldenburg (OLD), or National Herbarium of New South Wales (NSW). Collection initials: MJB = Michael J. Bayly, HMM = Heidi M. Meudt, PGJ = Phil Garnock-Jones, RGC = R. G. Coveny.

#RNA-Seq sample.
## Appendix 4. Validation of 48 LCN markers on 48 individuals of 46 species of Veronica, representing all subgeneric lineages in the genus.

| Locus   | Primer sequences (5’–3’)                           | Sequence same as original transcript? | No. of individuals successfully sequenced | No. of different sequences in GARLI alignment | Length (bp, range) | Length (bp, mean) | No. SNPs | No. PICs | A. thaliana gene                                                                 |
|---------|----------------------------------------------------|--------------------------------------|------------------------------------------|-----------------------------------------------|-------------------|-----------------|----------|---------|----------------------------------------------------------------------------------|
| LCN-03  | F: AGCATCGCTCTAGCTGTGGT<br>R: CCGGATATGCCACTGGT    | complete                             | 18                                       | 37                                            | 158–866           | 480             | 607      | 345     | AT3G07080: EamA-like transporter family                                          |
| LCN-04  | F: AGTTTTATACATTGGAAGC<br>R: CCGCTTGACCCCTCTCAAC   | complete                             | 42                                       | 43                                            | 288–331           | 327             | 117      | 77      | AT3G07720: Galactose oxidase/keck repeat superfamily protein                     |
| LCN-08  | F: CCGCTCTGGCTCCCTCCAATAG<br>R: CCCTCCAGGGAGCGTTAAG | complete                             | 27                                       | 44                                            | 311–791           | 374             | 427      | 131     | AT4G17100: UNKNOWN                                                               |
| LCN-10  | F: GCAAGACAGGCTCCTACAGAG<br>R: AGGCTCCTCTGTCATCTAC | complete                             | 35                                       | 68                                            | 234–820           | 440             | 523      | 303     | AT4G33460: ABC transporter family protein                                         |
| LCN-13  | F: GGCACCACTGCTGCTCCAGCT<br>R: CCAAGGAATACGGGCCTTAT | partial                              | 18                                       | 19                                            | 310–912           | 422             | 354      | 112     | AT5G65760: Serine carboxypeptidase S28 family protein                            |
| LCN-20  | F: GCCATACCTGAGACCCTTG<br>R: AGCAGAAATGCACCACACC  | partial                              | 44                                       | 85                                            | 306–681           | 384             | 280      | 176     | AT1G57770: FAD/NAD(P)-binding oxidoreductase family protein                      |
| LCN-25  | F: AGGAGTGATCAGCAATAG<br>R: ACTGTTTCCCCCAATCCACC  | partial                              | 43                                       | 89                                            | 310–726           | 396             | 439      | 287     | AT2G05830: NagB/RpiA/CoA transferase-like superfamily protein                   |
| LCN-38  | F: AAGGCCCTTGGGAGAGTGG<br>R: TAGTGTCTTTGGGCCACTCC  | complete                             | 42                                       | 104                                           | 310–762           | 361             | 436      | 299     | AT3G59380: famesyltransferase A                                                  |
| LCN-43  | F: TAGATGCTGTGTCGCTGTC<br>R: AGGACAGCTCTATCTGCGCA  | complete                             | 30                                       | 50                                            | 139–750           | 396             | 329      | 185     | AT4G35850: Pentatricopeptide repeat (PPR) superfamily protein                 |
| LCN-46  | F: TGCAGACTCTTTGGGGGT<br>R: ACTCCATCATGGGGGGGCTG  | complete                             | 44                                       | 85                                            | 307–707           | 373             | 314      | 224     | AT5G13800: pheophytinase                                                        |
| LCN-48  | F: AAGCTTAOGCCGCAGATAT<br>R: TGCCTGCTTATGAGTCTGA  | complete                             | 34                                       | 42                                            | 147–734           | 379             | 445      | 280     | AT5G14520: pescadillo-related                                                  |
| LCN-01  | F: GCCATGCTTGGAGAGTGG<br>R: ATCCCTCTGTCGCGAAT    | no                                   | 17                                       |                                               |                   |                 |          |         | AT1G71810: protein kinase superfamily protein                                   |
| LCN-02  | F: AGGAGTCTGGCTGGAGAGTGG<br>R: CCGCTTGACCCCTCTCAAC | no                                   | 22                                       |                                               |                   |                 |          |         | AT2G25950: protein of unknown function (DUF1000)                               |
| LCN-05  | F: TGGCTGCTTGCTGCTTATCT<br>R: AGAACTGCAACATCTCTGGGA | partial                              | 12                                       |                                               |                   |                 |          |         | AT3G20790: NAD(P)+binding Rossmann-fold superfamily protein                    |
| LCN-06  | F: AGCCCTTGAGATGATG<br>R: AACCTCGGACTCTTTGGA    | no                                   | 33                                       |                                               |                   |                 |          |         | AT4G09730: RHE9                                                                   |
| LCN-07  | F: GCCATGCTTGGAGGGCTTC<br>R: CCACAACTCTGCCCTGCGC  | no                                   | 0                                        |                                               |                   |                 |          |         | AT4G09750: NAD(P)+binding Rossmann-fold superfamily protein                    |
| LCN-09  | F: AAGAGCTGGTGAAGAGTGG<br>R: GCCAGTCCATAGCAATCTGC | no                                   | 16                                       |                                               |                   |                 |          |         | AT4G25450: nonintrinsic ABC protein 8                                           |
| LCN-11  | F: GTGCTGCTTGCTGCTGCTG<br>R: TCCACTCTGACCGTCTCCT | no                                   | 3                                        |                                               |                   |                 |          |         | AT4G37040: methionine aminopeptidase 1D                                       |
| LCN-12  | F: GCCATGCTTGGAGAGTGG<br>R: CTCACCACTCAGCTCTGCC  | no                                   | 20                                       |                                               |                   |                 |          |         | AT5G44520: NagB/RpiA/CoA transferase-like superfamily protein                 |
| Locus | Primer sequences (5′–3′) | Sequences same as original transcript? | No. of individuals successfully sequenced | No. of different sequences in GARLI alignment | Length (bp, range) | Length (bp, mean) | No. SNPs | No. PICs | A. thaliana gene |
|-------|--------------------------|--------------------------------------|------------------------------------------|---------------------------------------------|-------------------|------------------|------------|---------|----------------|
| LCN-14 | F: CGGA TGCTTCA ATGCTA GCTG<br> R: GCAC CTGACA ACAA CTGTAG | no | 13 | | | | | | AT1G04420: NAD(P)-linked oxidoreductase superfamily protein |
| LCN-15 | F: CGGT GGTTGA AACATT TTTG<br> R: TCCA ACAAGA GTGGACC AGC | partial | 28 | | | | | | AT1G16180: Serine-domain containing serine and sphingolipid biosynthesis protein |
| LCN-16 | F: ACTC TTTPCC GCATTT CTG<br> R: CCTC ACCATCT GAACT TGG | no | 30 | | | | | | AT1G19600: pkB-like carbohydrate kinase family protein |
| LCN-17 | F: AGAC TCTACCACAGCTTCC<br> R: TGGATGAT GGGGGC C | no | 11 | | | | | | AT1G31800: cytochrome B50, family 97, subfamily A, polypeptide 3 |
| LCN-18 | F: AGTT TGGTGTGGGCA TAGG<br> R: GAA ATA CAGCT CCGGGA AGG | partial | 24 | | | | | | AT1G45280: GLU-ADT subunit B |
| LCN-19 | F: CTGTG CGCTTGGCTATAG<br> R: TTTGAC CTCCCA AAGCAC CAC | no | 31 | | | | | | AT1G453280: Class I glutamine amidotransferase-like superfamily protein |
| LCN-21 | F: TGGT GTCATGGGCTGTC<br> R: TCCA ATCTTTGCTGATGTC | no | 9 | | | | | | AT1G68010: hydroxypyruvate reductase |
| LCN-22 | F: TGGG TAAGGGTCTTTTGGT<br> R: CCAATCTCTCAATCGATGCTG | no | 21 | | | | | | AT1G68830: STT7 homolog STN7 |
| LCN-23 | F: AACG ATGGAGGAAGAACG<br> R: CAAGCATTCCATGCTGAC | no | 24 | | | | | | AT1G71240: Plant protein of unknown function (DUF639) |
| LCN-24 | F: GAAA CTCTTATGCTCAGGGTTG<br> R: TCTT ATTAGGTGTTGCCACACC | no | 13 | | | | | | AT1G75210: HAD-superfamily hydrolase, subfamily IG, 5′-nucleotidase |
| LCN-26 | F: GATA ACCTGGAAGCGACGGATT<br> R: GCTA AGACCCAA CCCCCTTTTT | no | 32 | | | | | | AT2G21280: NAD(P)-binding Rossmann-fold superfamily protein |
| LCN-27 | F: TGGG ATCGATATCATG GCC<br> R: CAGG TGATGGTTTGACGTG | partial | 18 | | | | | | AT2G23390: UNKNOWN |
| LCN-28 | F: TGCGT CCCACAGCTGAGATG<br> R: CCAA TTCCTTCC AAGAAATCA | no | 16 | | | | | | AT2G27680: NAD(P)-linked oxidoreductase superfamily protein |
| LCN-29 | F: GCCA AGGCCCA AAAAAAGCAA<br> R: TCCCTCCATATGCAACC CGG | partial | 11 | | | | | | AT2G30390: ferrochelatase 2 |
| LCN-30 | F: ATGG AAGGAGGTGGATG<br> R: TTTG CGACAGTGACCACTTC | no | 22 | | | | | | AT2G44760: Domain of unknown function (DUF3598) |
| LCN-31 | F: TCAATTTGCAGCTTGGAGGC<br> R: CAAC ACCGCCA AATGCTACAAG | partial | 19 | | | | | | AT3G06510: Glycosyl hydrolase superfamily protein |
| LCN-32 | F: AAAATG GGTGGCTGCTGTG<br> R: CAAGCC TATACCTGACAT | no | 39 | | | | | | AT3G17810: pyrimidine 1 |
| LCN-33 | F: TGGCA ATGCAATCGCTCTTGC<br> R: AGAA TGATCGTGAGGCTG | no | 28 | | | | | | AT3G17940: Galactose mutarotase-like superfamily protein |
| LCN-34 | F: CACAG AAAAGGCAAACTAGG<br> R: TGAT CCAAATCAGAGGCTG | partial | 12 | | | | | | AT3G23620: Ribosomal RNA processing Brix domain protein |
| Locus | Primer sequences (5′–3′) | Sequence same as original transcript? | No. of individuals successfully sequenced | No. of different sequences in GARLI alignment | Length (bp, range) | Length (bp, mean) | No. SNPs | No. PICs | A. thaliana gene |
|-------|--------------------------|--------------------------------------|------------------------------------------|---------------------------------------------|-------------------|-------------------|----------|---------|-----------------|
| LCN-35 | F: AAATCGCTCACGGGTTTGG  
R: TTGCAGTGGGAAATGCTCAAAA | partial | 9 | | | | | | | AT3G52190: phosphate transporter traffic facilitator1 |
| LCN-36 | F: GATCGGGGTCAATCCACCA  
R: AACGCTAATGCAATGACAC | partial | 31 | | | | | | | AT3G56460: GroES-like zinc-binding alcohol dehydrogenase family protein |
| LCN-37 | F: CAAGGAGTCTGTTAGGAAGAC  
R: GAGACAGAGAAGCGGACC | no | 7 | | | | | | | AT3G56940: dicarboxylated iron protein, putative (Crd1) |
| LCN-39 | F: CGCGTGCTTCTCTGCATGT  
R: ATTGGAGGCTCTCAGTCTT | no | 36 | | | | | | | AT3G62910: Peptide chain release factor 1 |
| LCN-40 | F: TGGGAAACTGGAATGGGTG  
R: AATTGGACGCTGGATGTGT | no | 43 | | | | | | | AT4G02790: GTP-binding family protein |
| LCN-41 | F: AG7GGGCTGAAATGGGATT  
R: CCGATGCTGCTTTGATGTGT | no | 13 | | | | | | | AT4G09020: isomylase 3 |
| LCN-42 | F: AG7GGGCTGCAATGGGATT  
R: CCGATGCTGCTTTGATGTGT | partial | 9 | | | | | | | AT4G21470: riboflavin kinase/FMN hydrolase |
| LCN-44 | F: ACAAGGATGAGCAGACGGGT  
R: TGCCCAAAGAAGCTGTAACC | partial | 14 | | | | | | | AT5G06260: TLD-domain containing nucleolar protein |
| LCN-45 | F: GSCACAGTCTGCTGACACCA  
R: CCCAGCCCCATGGTCTAAA | no | 36 | | | | | | | AT5G08710: Regulator of chromosome condensation (RCC1) family protein |
| LCN-47 | F: TCCTCAGCGAATGGAATAA  
R: AAATCTCTCTCTGCTG | no | 22 | | | | | | | AT5G14250: Proteasome component (PCI) domain protein |

*Note:* LCN = low-copy nuclear; PIC = parsimony informative character; SNP = single-nucleotide polymorphism.
### Appendix 5. Validation of 48 SSRs on a subset of the 48 New Zealand and Australian individuals of *Veronica* sequenced. Shown are eight individuals of the *Veronica chionohebe*/*V. trifida* subset (A5, B5, C4, D4, E4, F4, G4, and H4; see Appendix 3).

| SSR locus | Primer sequences (5′–3′) | SSR motif, Main (additional) | Sequence length (bp) | No. of individuals successfully sequenced | No. of alleles | Pairwise genetic distance | SNP/PCs/Introns | Notes |
|-----------|--------------------------|----------------------------|---------------------|------------------------------------------|---------------|---------------------------|----------------|-------|
| SSR-01    | F: TGGAACAGCCATTGCATCAAA  |
| R: TCGTGACCTCACGGTGTCAG | ACA (ATG) | 310–692 | 353 | 8 | 3 2 2 2 3 3 3 2 4 | 2.9 | 0.03 0.01 0.01 | 23 14 | two large introns, motif in central exon, sequences partially not covering complete locus |
| SSR-02    | F: GATTGTTTCAGCCAAGATATCTCA  |
| R: CTTGTCGGACGCGCGACCAT | GAT | 208–476 | 328 | 7 | 0 2 2 2 1 1 1 1 | 1.5 | 0.08 0.04 0.02 | 28 27 | incomplete; several genes amplified by primers; same locus as SSR-42 |
| SSR-03    | F: TTGAGACGCAAGATTTCTGCAA  |
| R: GCCCTACGGCCTCATCAGTT | ACT | | | | | | | |
| SSR-04    | F: TGTCGAAACGTTGCATCAAA  |
| R: TCGTGACCTCACGGTGTCAG | AGT | 127–239 | 216 | 8 | 3 4 5 4 4 5 5 3 | | 0.06 0.02 0.01 | 23 15 | |
| SSR-05    | F: GGTGACGGATTTACTAGCTAAGT  |
| R: AGTCGGGAAAGAGATTGGGC | GATA | 293–329 | 328 | 8 | 4 1 1 4 5 6 1 6 | | 0.08 0.04 0.02 | 28 27 | |
| SSR-06    | F: AATGAACAACGCGACGCCAG | CAGT | 3 | | | | | | |
| R: ACTGTGACCTCATGCTTACG | TGA | 3 | | | | | | |
| SSR-07    | F: AGCTGGAACGCTCAGAATCC | GACA | 358–424 | 386 | 8 | 7 10 10 9 9 9 10 11 | | 0.33 0.18 0.14 | 190 174 | three orthologues? |
| SSR-08    | F: CCAATGACGGCTCCAGTGCTG | CACA | 311–387 | 357 | 8 | 9 10 4 3 6 6 8 | | 0.14 0.08 0.08 | 126 109 | |
| SSR-09    | F: TTGGACTCTTGTGTGTGA  |
| R: CCTATAATTGTGCTCTCCA | ACT | 310–401 | 325 | 2 | 1 3 1 2 0 0 0 | | 1 0 | 3 | at least three introns, sequence not covering complete locus |
| SSR-10    | F: CTTAAATTGATACGATGC | TGCAT | 266–280 | 272 | 8 | 3 4 1 1 4 3 4 1 | | 0.04 0.02 0.02 | 20 14 | |
| SSR-11    | F: AAAGACGCAAACGCTCAGAC  |
| R: GGCACACAGGGATGCAGTAC | AGCAG (TIG, ATG) | 264–293 | 283 | 8 | 2 8 9 4 5 4 4 2 | | 0.08 0.04 0.04 | 42 31 | two orthologues? |
| SSR-12    | F: TTGGCACTGCGTTTAAGAATG | TATA | | | | | | | |
| R: ATACCGACGCGCTACAATCC | ACTG | | | | | | | |
| SSR-13    | F: TCTTCTTCAGCTGCAAACCTTCC | TACA | 301–620 | 377 | 7 | 3 4 2 3 4 1 0 3 | | 0.51 0.06 0.02 | 91 43 | |
| SSR-14    | F: TGTTGACTCAATCCGTCTCCG | TTTAG | 290–306 | 293 | 8 | 2 1 1 1 4 4 2 2 1 4 | | 0.02 0.01 0.01 | 9 6 | one unambiguous locus |
| SSR-15    | F: GCGCAAGAACGCTGCAAG | CATC | 310–355 | 355 | 2 | 2 2 0 0 2 0 0 0 | | 1 0 | |
| SSR-16    | F: GACACGAACTCTGCGAAGCTC | ATG | 301–620 | 377 | 7 | 3 4 2 3 4 1 0 3 | | 0.51 0.06 0.02 | 91 43 | |
| SSR-17    | F: AGATCGCTGTCTTTCAGCTC  |
| R: GCGATGACTCAAAAACCACAA | TGCAT | 203–257 | 238 | 8 | 1 1 7 2 5 6 5 6 | | 0.09 0.04 0.03 | 39 27 | two orthologues? |
| SSR-18    | F: TCTGTGACTCAACTGCAGAAGAG | TACG | 268–303 | 282 | 8 | 6 6 2 1 4 1 4 3 | | 0.09 0.04 0.03 | 35 33 | two orthologues? |
| SSR-19    | F: TTGGCACTGCGTTTAAGAATG | TATA | | | | | | | |
| R: AGCAGAACGCAAATCGTGTGGG | GACA | | | | | | | |
| SSR-20    | F: CATGCTGTCACTCATTGAAAATGTTGGCC | ACTG | 186–253 | 228 | 8 | 2 2 8 3 4 8 8 5 2 8 5 | | 0.1 0.04 0.05 | 33 31 | two orthologues? |
| SSR locus | Primer sequences (5′–3′) | SSR motif, Main (additional) | Sequence length (bp) | No. of individuals successfully sequenced | No. of alleles | Pairwise genetic distance | No. | Notes |
|-----------|--------------------------|-----------------------------|---------------------|------------------------------------------|-------------|--------------------------|-----|-------|
| SSR-21    | F: AGGAGGTGACGAGAAAGAG   | GAT                         | 238–256             | 8                                        | 5           | 1.0                      | 38  | two or three orthologues? |
|           | R: CCGCCGTCCTCTCCCTGTT   |                             |                     |                                          |             |                          |     |                    |
| SSR-22    | F: AGGCTGTCTGTTGAAAACCG  | GAT                         | 286–348             | 2                                         | 1           | 0.36                     | 111 | two orthologues?         |
|           | R: GACTGACACTGCTGCA       |                             |                     |                                          |             |                          |     |                    |
| SSR-23    | F: CATCCACAAAAGTGACAGCCT  | GAT                         | 268–315             | 8                                        | 3           | 0.19                     | 88  | three orthologues?       |
|           | R: TGGAATGCTGATTAAAGGGA   |                             |                     |                                          |             |                          |     |                    |
| SSR-24    | F: GATGCTGTGACGAGAAAGAG  | GAT                         | 275–315             | 8                                        | 3           | 0.22                     | 25  | two orthologues?         |
|           | R: CCGCCGTCCTCTCCCTGTT   |                             |                     |                                          |             |                          |     |                    |
| SSR-25    | F: AGGCTGTCTGTTGAAAACCG  | GAT                         | 286–348             | 8                                        | 3           | 0.05                     | 25  | two orthologues?         |
|           | R: GACTGACACTGCTGCA       |                             |                     |                                          |             |                          |     |                    |
| SSR-26    | F: GTGGGCGAACAAATTTGTTT  | ATC                         | 288–315             | 8                                        | 3           | 0.19                     | 88  | three orthologues?       |
|           | R: TCACTAAGCGATCTGCTGCT   |                             |                     |                                          |             |                          |     |                    |
| SSR-27    | F: CCGAGGCTGCTAATGACTAAGT | ACT                         | 259–262             | 8                                        | 2           | 0.02                     | 8   | one unambiguous locus    |
|           | R: GGCAGACGCTTTGAGAGGGA   |                             |                     |                                          |             |                          |     |                    |
| SSR-28    | F: CCGAAATGCAGATCTGAGCTG | ACT                         | 266–284             | 8                                        | 3           | 0.05                     | 25  | two orthologues?         |
|           | R: GGGAGCAAGAGCTCTGCA     |                             |                     |                                          |             |                          |     |                    |
| SSR-29    | F: GACACCAACTTCTCTTCACAT  | ACT                         | 300–355             | 8                                        | 6           | 0.14                     | 65  | two orthologues?         |
|           | R: AAGAGTGTGTCAGCTTGAGTT  |                             |                     |                                          |             |                          |     |                    |
| SSR-30    | F: TTCTTGCCCTCTGTGGTGGTCC | GAT                         | 278–290             | 8                                        | 2           | 0.04                     | 16  | three orthologues?       |
|           | R: TTCTTGCCCTCTGTGGTGGTCC |                             |                     |                                          |             |                          |     |                    |
| SSR-31    | F: CGATAGAGGTGAGGAGGAG    | GAT                         | 2                           | 2                                        |             |                          |     | sequence different to transcript |
|           | R: CATTGGTGACCTGATGCTCT   |                             |                     |                                          |             |                          |     |                    |
| SSR-32    | F: GTGCTACTGATGACTGAGAGA  | GAT                         | 158–248             | 8                                        | 1           | 0.07                     | 17  | two orthologues?         |
|           | R: TCCCTGCACTGCACTGCA     |                             |                     |                                          |             |                          |     |                    |
| SSR-33    | F: GCTGCTGTGGTGAGGAGGAG   | GAT                         | 2                           | 5                                        |             |                          |     | three orthologues?       |
|           | R: ACTGCTGATGCTGCTGCTT    |                             |                     |                                          |             |                          |     |                    |
| SSR-34    | F: ATGGCTCACAGATCATGGCTCTC | ATC (CAA)                 | 317–501              | 8                                        | 5           | 0.34                     | 141 | three orthologues?       |
|           | R: TGACATGTTTGCGAGATTTG   |                             |                     |                                          |             |                          |     |                    |
| SSR-35    | F: TGGACATGCTGGGAAAACAC   | ATC (CAA)                 | 317–501              | 8                                        | 5           | 0.34                     | 141 | three orthologues?       |
|           | R: ACATGTCAGTCTGCTGGTCT   |                             |                     |                                          |             |                          |     |                    |
| SSR-36    | F: AAGGACATGACTGAGGAGGAGA | ATC (CAA)                 | 317–501              | 8                                        | 5           | 0.34                     | 141 | three orthologues?       |
|           | R: ACATGTCAGTCTGCTGGTCT   |                             |                     |                                          |             |                          |     |                    |
| SSR-37    | F: ATGGACAGCTGGTGGTGTCC   | GAT                         | 112–280              | 8                                        | 13          | 0.09                     | 45  | two or three orthologues? |
|           | R: CGAGAAGACAGAGATGTCCT   |                             |                     |                                          |             |                          |     |                    |
| SSR-38    | F: CCGCAAGGTCAGCAGCTAGAC  | TCA                         | 310–569              | 8                                        | 7           | 0.22                     | 174 | three orthologues?       |
|           | R: GTGCACTGCTCAACAGCTTC   |                             |                     |                                          |             |                          |     |                    |
| SSR-39    | F: AGGCTTCAGTCTGATGCTC    | TCA                         | 450–480              | 8                                        | 4           | 0.05                     | 55  | two orthologues?         |
|           | R: TGAGTAAAGGAGAAAGAAGAC  |                             |                     |                                          |             |                          |     |                    |
| SSR-40    | F: GCCCTGGTGAGGACATTTG    | GAT                         | 1                           | 1                                        |             |                          |     | sequences different to transcript |
|           | R: ATGCTATGGTAAAGGATGCT   |                             |                     |                                          |             |                          |     |                    |
| SSR-41    | F: GTAGACGATGACTGAGGCTCT  | TCA                         | 375–380              | 5                                        | 1           | 0.05                     | 23  | three orthologues?       |
|           | R: GCAGTCTCTCTGGAGGTCTTT  |                             |                     |                                          |             |                          |     |                    |
| SSR-42    | F: AGTCTACACTAACTGAAAGGAC | GAT                         | 7                           | 7                                        |             |                          |     | three orthologues?       |
|           | R: AGCTTATCCTCCAGTCATTTAGC |                             |                     |                                          |             |                          |     |                    |
| SSR locus | Primer sequences (5′–3′) | SSR motif, Main (additional) | Sequence length (bp) | No. of individuals successfully sequenced | No. of alleles | Pairwise genetic distance | No. SNPs | PICs | Introns | Notes |
|-----------|--------------------------|-----------------------------|---------------------|------------------------------------------|----------------|--------------------------|---------|------|---------|-------|
| SSR-43    | F: ACCATCA; AAGGGCTCCGTAG | ATG                         | 192–416             | 8                                        | 1              | 1                        | A5      | B5   | C4      | 1.6   |
|           | R: TTGGGATGGGGCCCCCT       | 382                         |                     |                                          |                |                          |         |      |         |       |
| SSR-44    | F: GTTATACGCACTGCCACCCCTGC | ATC (TGG, CACC)            | 283–310             | 8                                        | 2              | 4                        | 0.06    | 0.03 | 0.03    | 31    |
|           | R: AGTGAGGCGCATGCTCTCTGG  | 297                         |                     |                                          |                |                          |         |      |         | 21    |
| SSR-45    | F: GTTCGTCTTGAAGATGGACATGA | GG (GGAATT, TGT)           | 147–632             | 8                                        | 2              | 4                        | 0.02    | 0.01 | 0.02    | 19    |
|           | R: ACAATTGTTCCATAGTTTTGAA | 316                         |                     |                                          |                |                          |         |      |         | 15    |
| SSR-46    | F: TGCTGTAGCCAGAAGGCCC    | GTATGCTCAGAAGAAGAAA         | 3                   |                                          |                |                          |         |      |         |       |
|           | R: GGTGGGTCAGAAGAGAAA     |                             |                     |                                          |                |                          |         |      |         |       |
| SSR-47    | F: CAGGAGCCAGATGCTGAAA    | TGAGAT (GGAATT, TGT)       | 264–288             | 8                                        | 1              | 2                        | 0.04    | 0.01 | 0.02    | 15    |
|           | R: ACCACTTTGCTTTAAGAGAAA  | (GGAATT, TGT)              | 272                 |                                          |                |                          |         |      |         | 19    |
| SSR-48    | F: CTTCCTCAGATCCGATGAGG   |                             | 0                   |                                          |                |                          |         |      |         |       |
|           | R: CAAATCTTTGCTTTAATAGC   |                             |                     |                                          |                |                          |         |      |         |       |
| Totals    |                          |                             | 112–692             | 316.8                                    | 6.7            | 3.7                      | 0.51    | 0.04 | 0.04    | 54.7  |

Note: PICs = parsimony informative characters; SNPs = single-nucleotide polymorphisms; SSRs = simple sequence repeats.
APPENDIX 6. Validation of 48 SSRs on a subset of the 48 New Zealand and Australian individuals of *Veronica* sequenced. Shown are six individuals of the *V. thomsonii* subset (C5, D5, E5, F5, G5, and H5; see Appendix 3).

| SSR locus | Primer sequences (5′–3′) | SSR motif, Main (additional) | Sequence length (bp) | No. of individuals successfully sequenced | No. of alleles | Pairwise genetic distance | No. SNPs | No. PCs | Introns | Notes |
|-----------|--------------------------|-------------------------------|----------------------|------------------------------------------|---------------|---------------------------|----------|--------|--------|-------|
| SSR-01    | F: TGAGACGACATCCATGACAA   | ACA (ATG)                     | 310–694              | 6                                        | 4             | 4                         | 3        | 2      | 2      | 4     | two large introns, motif in central exon, sequences partially not covering complete locus |
|           | R: TGCTGACTAACTACACCTCCAG |                              |                      |                                          |               |                           |          |        |        |       | incomplete; several genes amplified by primers; same locus as SSR-42 |
| SSR-02    | F: GATTGTTCAGCAAGAATTCTCA | GAT                           | 127–242              | 6                                        | 2             | 3                         | 9        | 6      | 2      | 9     | several genes amplified by primers |
|           | R: CTGGTCAGCAAGAATTCTCCAG |                              |                      |                                          |               |                           |          |        |        |       | at least 1 |
| SSR-03    | F: TGAGACGACATCCATGACAA   | ACT                           | 127–242              | 6                                        | 6             | 6                         | 4        | 2      | 2      | 4     | several genes amplified by primers |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | not transcript sequence amplified |
| SSR-04    | F: TTGGTCAACCGCTGCGACTG   | GAT                           | 243–314              | 6                                        | 6             | 6                         | 4        | 2      | 4      | 6     | several genes amplified by primers |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | at least 1 |
| SSR-05    | F: TTGGTCAACCGCTGCGACTG   | AGA                           | 310–371              | 6                                        | 6             | 6                         | 4        | 2      | 4      | 6     | several genes amplified by primers |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | not transcript sequence amplified |
| SSR-06    | F: TTGGTCAACCGCTGCGACTG   | GAT                           | 300–371              | 6                                        | 6             | 6                         | 4        | 2      | 4      | 6     | several genes amplified by primers |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | at least 1 |
| SSR-07    | F: TTGGTCAACCGCTGCGACTG   | AGA                           | 310–371              | 6                                        | 6             | 6                         | 4        | 2      | 4      | 6     | several genes amplified by primers |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | not transcript sequence amplified |
| SSR-08    | F: TTGGTCAACCGCTGCGACTG   | GAT                           | 342–418              | 6                                        | 8             | 8                         | 4        | 4      | 4      | 8     | several genes amplified by primers |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | at least 1 |
| SSR-09    | F: TTGGTCAACCGCTGCGACTG   | GAT                           | 342–418              | 6                                        | 8             | 8                         | 4        | 4      | 4      | 8     | several genes amplified by primers |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | at least 1 |
| SSR-10    | F: TTGGTCAACCGCTGCGACTG   | AGA                           | 342–418              | 6                                        | 8             | 8                         | 4        | 4      | 4      | 8     | several genes amplified by primers |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | at least 1 |
| SSR-11    | F: AAACGACGACTGACG       | ACGA (TTG, AT, AG)           | 265–294              | 6                                        | 7             | 7                         | 5        | 5      | 5      | 12    | two orthologues? |
|           | R: GGCTGACTGCTGCTGCTGCTG |                              |                      |                                          |               |                           |          |        |        |       | two orthologues? |
| SSR-12    | F: TTGGTCAACCGCTGCGACTG   | AATC                          | 199–263              | 6                                        | 5             | 5                         | 9        | 6      | 6      | 9     | two orthologues? |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | two orthologues? |
| SSR-13    | F: TTGGTCAACCGCTGCGACTG   | GAT                           | 342–418              | 6                                        | 8             | 8                         | 4        | 4      | 4      | 8     | two orthologues? |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | two orthologues? |
| SSR-14    | F: TTGGTCAACCGCTGCGACTG   | GAT                           | 286–311              | 6                                        | 2             | 2                         | 2        | 1      | 4      | 1     | two orthologues? |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | two orthologues? |
| SSR-15    | F: TTGGTCAACCGCTGCGACTG   | GAT                           | 300–371              | 6                                        | 6             | 6                         | 4        | 4      | 4      | 6     | two orthologues? |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | two orthologues? |
| SSR-16    | F: TTGGTCAACCGCTGCGACTG   | GAT                           | 310–371              | 6                                        | 6             | 6                         | 4        | 4      | 4      | 6     | two orthologues? |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | two orthologues? |
| SSR-17    | F: TTGGTCAACCGCTGCGACTG   | GAT                           | 342–418              | 6                                        | 8             | 8                         | 4        | 4      | 4      | 8     | two orthologues? |
|           | R: TCGGACTCTACACCTCCAG    |                              |                      |                                          |               |                           |          |        |        |       | two orthologues? |
| SSR locus | Primer sequences (5′–3′) | SSR motif, Main (additional) | Sequence length (bp) | No. of individuals successfully sequenced | No. of alleles | Pairwise genetic distance | SNP | PICs | Introns | Notes |
|-----------|-------------------------|-----------------------------|---------------------|------------------------------------------|----------------|------------------------|-----|------|--------|-------|
| SSR-18    | F: TCTGGCTGACACCTGACATCAAGGAG | 260–295 | 4 | 4 | 0.08 | 0.02 | 0.01 | 31 | 24 | sequences not covering locus |
| SSR-19    | F: TGGCAACATGCACACTGTTT | TATC (ATA, TAC) | 190–253 | 232 | 6 | 5 | 10 | 5.7 | 0.09 | 0.04 | 0.04 | 26 | 24 | two orthologues? |
| SSR-20    | F: CATTGTATATCTGTAATAGTGTTG | GTTA (ACA, GTGA) | 238–256 | 245 | 6 | 5 | 4 | 10 | 4.8 | 0.12 | 0.05 | 0.02 | 39 | 32 | two orthologues? |
| SSR-21    | F: ATGGATAAGGCACGTATGAGG | GAT | 285–345 | 336 | 6 | 3 | 6 | 1 | 3 | 6 | 3.5 | 0.41 | 0.15 | 0.01 | 133 | 126 | 1 |
| SSR-22    | F: AGGGCGTTATGGAAACCGGG | GAT | 238–256 | 245 | 6 | 5 | 4 | 6 | 5 | 6 | 10 | 5.7 | 0.09 | 0.04 | 0.04 | 26 | 24 | two orthologues? |
| SSR-23    | F: CACACACAAAACGTAAGACT | TGTGATCTTGATGAAAGGAA | 190–253 | 232 | 6 | 5 | 4 | 6 | 5 | 6 | 10 | 5.7 | 0.09 | 0.04 | 0.04 | 26 | 24 | two orthologues? |
| SSR-24    | F: GTCACACTCCCTCGTCGCA | GAT | 275–287 | 289 | 6 | 1 | 2 | 3 | 5 | 2 | 5 | 2 | 0.04 | 0.01 | 0.01 | 13 | 5 | 0 |
| SSR-25    | F: GGTGTATAAGGCACGTGGA | TGTGATCTTGATGAAAGGAA | 190–253 | 232 | 6 | 5 | 4 | 6 | 5 | 6 | 10 | 5.7 | 0.09 | 0.04 | 0.04 | 26 | 24 | two orthologues? |
| SSR-26    | F: GTGGCCGAGCCAAGGTTGTTT | ATC | 302–306 | 302 | 6 | 3 | 4 | 2 | 2 | 2 | 4 | 2.8 | 0.2 | 0.12 | 0.17 | 95 | 81 | 0 |
| SSR-27    | F: TCACTAATACACTCTACGTCGTC | ACT | 259–263 | 259 | 6 | 2 | 3 | 1 | 1 | 1 | 3 | 1.7 | 0.04 | 0.01 | 0.01 | 14 | 4 | 0 |
| SSR-28    | F: GCAGAACGAGCCACATGAGGGG | ACT | 257–290 | 273 | 6 | 4 | 3 | 6 | 4 | 3 | 6 | 4 | 0.05 | 0.02 | 0.03 | 25 | 18 | 0 |
| SSR-29    | F: GACACACAAATCTCTCTTGAGTC | ACT | 300–351 | 320 | 6 | 7 | 4 | 5 | 7 | 4 | 2 | 2 | 7 | 4.8 | 0.13 | 0.06 | 0.06 | 61 | 58 | 1 |
| SSR-30    | F: TCTGGCTCTGGTGTTGCTTG | TGA | 275–279 | 289 | 6 | 1 | 2 | 3 | 5 | 2 | 5 | 2 | 0.04 | 0.01 | 0.01 | 13 | 5 | 0 |
| SSR-31    | F: GCTGAGGAGTGGAGAGAAGC | no sequences |
| SSR-32    | F: GTGCTGATTGCTACATCAAGAGAG | GAT | 236–251 | 241 | 6 | 6 | 3 | 6 | 3 | 3 | 9 | 5.3 | 0.07 | 0.03 | 0.03 | 26 | 23 | 0 |
| SSR-33    | F: GCCAGGATGGTTGAGAACCAGCA | no sequences |
| SSR-34    | F: ATGGCAAGATTTGCTGGTCTGCT | ACT | 455–455 | 455 | 5 | 7 | 7 | 6 | 6 | 0 | 1 | 7 | 5 | 0.47 | 0.16 | 0.06 | 192 | 141 | 1 |
| SSR-35    | F: GCCACGCTCTGGGTAATAGGGG | ATC (CA) A | 463–501 | 484 | 5 | 7 | 4 | 7 | 6 | 6 | 0 | 1 | 7 | 5 | 0.47 | 0.16 | 0.06 | 192 | 141 | 1 |
| SSR-36    | F: ATTGCCCTGGGGAAACAATTCAAGG | TCA | 240–253 | 245 | 6 | 2 | 4 | 4 | 2 | 4 | 2 | 4 | 3 | 0.02 | 0.01 | 0.01 | 8 | 6 | 0 |
| SSR-37    | F: AGTGCACCTCTCTCTGCTGTTCC | GAT | 184–280 | 275 | 6 | 10 | 26 | 16 | 14 | 10 | 2 | 2 | 26 | 13 | 0.08 | 0.03 | 0.03 | 32 | 29 | 0 |
| SSR-38    | F: CACCAATCTGATGATGACATCAAGC | TCA | 310–568 | 536 | 6 | 10 | 8 | 6 | 8 | 6 | 8 | 6 | 10 | 7.7 | 0.23 | 0.08 | 0.06 | 171 | 98 | 2 |
| SSR-39    | F: ATCTCTGCATCTGATCAGTCAACA | TCA | 406–480 | 456 | 6 | 6 | 4 | 3 | 5 | 4 | 5 | 6 | 4.5 | 0.06 | 0.03 | 0.03 | 52 | 36 | 2 |
| SSR-40    | F: GTGGCTGTTGAGACTGTTG | no sequences |

**Notes:**
- SSR-18: sequences not covering locus
- SSR-20: two orthologues?
- SSR-21: sequences not covering locus, two loci?
- SSR-22: no sequences
- SSR-23: sequences not covering locus, two loci?
- SSR-24: sequences not covering locus
- SSR-25: no sequences
- SSR-26: three orthologues?
- SSR-27: one unambiguous locus
- SSR-28: two orthologues?
- SSR-29: three orthologues?
- SSR-30: one unambiguous locus
- SSR-31: no sequences
- SSR-32: two or three orthologues?
- SSR-33: different orthologous locus sequenced?
- SSR-34: sequences different to transcript
- SSR-35: two orthologues?
- SSR-36: one unambiguous locus
- SSR-37: two or three orthologues?
- SSR-38: two or three orthologues?
- SSR-39: two or three orthologues?
- SSR-40: sequences different to transcript, but consistent
### APPENDIX 6. Continued.

| SSR locus | Primer sequences (5’–3’) | SSR motif, Main (additional) | Sequence length (bp) | No. of individuals successfully sequenced | No. of alleles | Pairwise genetic distance | SNPs | PICs | Introns | Notes |
|-----------|--------------------------|-----------------------------|----------------------|------------------------------------------|----------------|--------------------------|------|------|---------|-------|
| SSR-41    | F: GTAAGCAAAGTAGATTGGTTCACCT<br>R: GGCGTGCCTCTGGTGGTTTT | GAT | 376–380 379 | 5 | 2 | 0 | 1 | 1 | 1 | 1 | 2 | 1 | 0.06 | 0.02 | 0.04 | 24 | 4 | 1 |
| SSR-42    | F: AGCTAACTCAAATACGATGCAA<br>R: AGCTCATTCCACGTATTACC | GAT | 125–362 313 | 6 | 3 | 2 | 3 | 2 | 1 | 1 | 3 | 2 | 0.39 | 0.18 | 0.15 | 180 | 68 | 0 | two orthologues? |
| SSR-43    | F: ACCCTGAAACCTTGAGCT<br>R: TTTGGGTGGCCCT GCCTCAC | ATG | 157–416 366 | 4 | 2 | 0 | 2 | 0 | 0 | 1 | 1 | 2 | 0.8 | | | | | | |
| SSR-44    | F: GTTTAAAGCATCCACGCGTG<br>R: AGCTAGGACATGCTGCGT | ATC (TCG, CACC) | 283–316 296 | 6 | 3 | 4 | 4 | 3 | 5 | 2 | 2 | 5 | 3.5 | 0.05 | 0.03 | 0.03 | 21 | 16 | 0 | two orthologues? |
| SSR-45    | F: GTGGCTTTGAAAGATGCCATGA<br>R: ACAATGTCTCCACGGCCCTG | | 148–359 303 | 6 | 6 | 1 | 10 | 6 | 1 | 1 | 1 | 10 | 4.2 | 0.6 | 0.43 | 0.47 | 405 | 311 | ? | three orthologues? second transcript different to transcript |
| SSR-46    | F: TCGCTGATGACGCAAGG<br>R: GGCTGCTGACGAAAGGCA | | | | | | | | | | | | | | | | | |
| SSR-47    | F: CAGAGACACCAGCTGCTGACA<br>R: ACCATGGATCAAGCAAGAC<br> | TGAGAT (GGAATT, TGT) | 264–294 271 | 6 | 3 | 7 | 4 | 6 | 2 | 2 | 2 | 1 | 0.03 | 0.01 | 0.01 | 17 | 10 | 0 | two orthologues? |
| SSR-48    | F: CTTCCACTCCTGAAATCAGTGAG<br>R: CAACTCTCGCCCTTTATCGA | | | | | | | | | | | | | | | | | |
| Totals    | | | 125–1051 327.3 | 5 | 4.3 | 4.6 | 5.4 | 4.1 | 3.9 | 3.2 | 1 | 26 | 4.3 | 0.6 | 0.1 | 0.1 | 80.3 | 52.5 | 0–2 |

**Note:** PICs = parsimony informative characters; SNPs = single-nucleotide polymorphisms; SSRs = simple sequence repeats.