



Primers for Castilleja and their Utility Across Orobanchaceae: II. Single-Copy Nuclear Loci

Authors: Latvís, Maribeth, Jacobs, Sarah J., Mortimer, Sebastian M. E., Richards, Melissa, Blischak, Paul D., et al.

Source: Applications in Plant Sciences, 5(9)

Published By: Botanical Society of America

URL: <https://doi.org/10.3732/apps.1700038>

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

PRIMERS FOR *CASTILLEJA* AND THEIR UTILITY ACROSS OROBANCHACEAE: II. SINGLE-COPY NUCLEAR LOCI¹

MARIBETH LATVIS^{2,8}, SARAH J. JACOBS^{3,4,5}, SEBASTIAN M. E. MORTIMER^{3,4},
MELISSA RICHARDS³, PAUL D. BLISCHAK⁶, SARAH MATHEWS⁷, AND DAVID C. TANK^{3,4,5}

²Department of Natural Resource Management, South Dakota State University, 1390 College Avenue, Brookings, South Dakota 57007 USA; ³Department of Biological Sciences, University of Idaho, 875 Perimeter Drive, MS 3051, Moscow, Idaho 83844-3051 USA; ⁴Stillinger Herbarium, University of Idaho, 875 Perimeter Drive, MS 1133, Moscow, Idaho 83844-1133 USA; ⁵Institute for Bioinformatics and Evolutionary Studies (IBEST), University of Idaho, 875 Perimeter Drive, MS 3051, Moscow, Idaho 83844-3051 USA; ⁶The Ohio State University, Department of Evolution, Ecology, and Organismal Biology, 318 W. 12th Avenue, Columbus, Ohio 43210 USA; and ⁷Australian National Herbarium, CSIRO National Research Collections, Canberra, Australia

- *Premise of the study:* We developed primers targeting nuclear loci in *Castilleja* with the goal of reconstructing the evolutionary history of this challenging clade. These primers were tested across other major clades in Orobanchaceae to assess their broader utility.
- *Methods and Results:* We assembled low-coverage genomes for three taxa in *Castilleja* and developed primer combinations for the single-copy conserved ortholog set (COSII) and the pentatricopeptide repeat (PPR) gene family. These primer combinations were designed to take advantage of the Fluidigm microfluidic PCR platform and are well suited for high-throughput sequencing applications. Eighty-seven primers were designed for *Castilleja*, and 27 were found to have broader utility in Orobanchaceae.
- *Conclusions:* These results demonstrate the utility of these primers, not only across *Castilleja*, but for other lineages within Orobanchaceae as well. This expanded molecular toolkit will be an asset to future phylogenetic studies in *Castilleja* and throughout Orobanchaceae.

Key words: *Castilleja*; high-throughput sequencing; microfluidic PCR; Orobanchaceae; single-copy nuclear genes.

Although the plastome has long been considered the workhorse of phylogenetic inference in plants, reliance on chloroplast data alone may limit the ability to identify inheritance patterns of polyploids, as well as introgression and hybridization events (Godden et al., 2012; Twyford and Ennos, 2012). The reliable estimation of an underlying species tree also depends upon the acquisition of multiple, unlinked loci—especially for recent and rapid radiations—shifting the focus toward the development of single- or low-copy nuclear gene regions for phylogenetic analyses. The process of identifying and developing these nuclear gene regions using traditional methods can be time consuming and costly, but the increasing availability of high-throughput sequencing data, as well as new bioinformatic approaches, allows for the efficient and cost-effective exploration of the nuclear genome.

Here we focus on developing a suite of putatively single-copy nuclear gene regions in *Castilleja* L. (Orobanchaceae;

“the paintbrushes”), a clade rich with polyploid and hybrid taxa, and the product of an ongoing rapid radiation (Tank and Olmstead, 2008, 2009). Previous studies using the nuclear ribosomal ITS and ETS regions, the low-copy nuclear gene *waxy*, and the plastid *trnL-F* and *rps16* intron regions hinted at cytonuclear discordance in some taxa, and most relationships among closely related taxa were unresolved (Tank and Olmstead, 2008, 2009). We recently developed primer combinations targeting the most variable regions of the plastome in *Castilleja* (Latvis et al., 2017a), and now present a companion set of nuclear primers with the goal of obtaining a resolved species tree for this challenging clade, as well as to aid in the detection of introgression and hybrid speciation. Primers for microsatellite markers have also been developed by Fant et al. (2013) for population-level investigations. In part, we follow the approach outlined by Blischak et al. (2014) to develop nuclear primer combinations from genome-skimming data, while following specifications for the Fluidigm Access Array microfluidic PCR system (Fluidigm, South San Francisco, California, USA) (see Latvis et al., 2017a). Thus, all primer combinations use the same annealing temperature of 60°C and may be amplified in parallel prior to high-throughput sequencing or traditional Sanger sequencing. We specifically target putatively single-copy genes from the conserved ortholog set (COSII) and pentatricopeptide repeat (PPR) domains, both of which have been highlighted for their phylogenetic utility in plants (COSII: Wu et al., 2006; PPR domain:

¹Manuscript received 15 April 2017; revision accepted 20 July 2017.

This research was supported by resources at the Institute for Bioinformatics and Evolutionary Studies (IBEST; NIH/NCRR P20RR16448 and P20RR016454) and by the following awards from the National Science Foundation: DEB-1253463 (awarded to D.C.T.), DEB-1502061 (awarded to D.C.T. for S.J.J.), and DEB-1455399 (support for P.D.B.).

⁸Author for correspondence: Maribeth.Latvis@sdstate.edu

doi:10.3732/apps.1700038

Applications in Plant Sciences 2017 5(9): 1700038; <http://www.bioone.org/loi/apps> © 2017 Latvis et al. Published by the Botanical Society of America. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC-BY-NC-SA 4.0), which permits unrestricted noncommercial use and redistribution provided that the original author and source are credited and the new work is distributed under the same license as the original.

TABLE 1. Nuclear primer pairs designed for *Castilleja* (locus and region amplified), amplicon lengths, and validation results for Orobanchaceae and outgroup taxon *Paulownia*. All pairs were designed for an annealing temperature of 60°C (±1°C). Boldfaced rows correspond to core Orobanchaceae primers, defined by successful amplification in two or more major clades in Orobanchaceae (see Fig. 1).

Locus (Region)	Primer sequences (5'–3') ^a		Amplicon length (bp) ^b	Clade I:	Clade II:	Clade III:	Clade IV:	Clade V:	Clade VI:	Clade VI:	Paulowniaceae: <i>Paulownia fortunei</i> (outgroup)
	<i>Lindenbergia</i> sp. ^c	<i>Schwalbea americana</i> ^c		<i>Orobancha californica</i> ^c	<i>Castilleja lineariloba</i> , <i>C. pumila</i> , <i>C. lemmonii</i> ^a	<i>Lamoureauxia virgata</i> ^c	<i>Pedicularis</i> sp. ^c	<i>Neobartisia pervinacae</i> ^c	<i>Rhinanthus alectonophilus</i> ^c		
A12g34560_688F		F: CTGTTTGTGGCAGCAAGTACC	498	X	X	X	X	X	X	X	X
A12g34560_1186R (COSII)		R: AAAGGCTGTGTCACAACTGAAGC									
A15g46630_851F		F: TGACAAACAGATTGTCAAATGCA	499	X	X	X	X	X	X	X	X
A15g46630_1350R (COSII)		R: TTGGCAGGACGGGATTTAAGT									
A13g04260_165F		F: GGAGTGGGAAGTTGAATTGGC	500	X	X	X	X	X	X	X	X
A13g04260_665R (COSII)		R: AAGGCTGAAGGCATAGGAGC									
A15g12370_3513F		F: ATGCACTGCTGGAAGCCCTTT	511	X	X	X	X	X	X	X	X
A15g12370_4024R (COSII)		R: GAGATGGCTGACTGAAGCT									
A12g28390_1096F		F: TGCAATGCAAGTTATCAGTCTTG	499	X	X	X	X	X	X	X	X
A12g28390_1595R (COSII)		R: GACATGGGATTTGATCTTGGCAG									
A12g26430_2136F		F: TTGCAAGAATTCACAGCTTCTTGTG	496	X	X	X	X	X	X	X	X
A12g26430_2632R (COSII)		R: GCGCCAGTTGTGTTTGG									
A14g24190_1278F		F: CTGAAGTCCAGATTTAGCAATAGTAC	499	X	X	X	X	X	X	X	
A14g24190_1777R (COSII)		R: TCAGGCATTTGGACAGATTAGGT									
A15g27620_538F		F: CCGCATATCTTTCAACTTCGCG	799	X	X	X	X	X	X	X	
A15g27620_1337R (COSII)		R: CTTGGGAAAGGATTTGAACAGG									
A13g62010_1226F		F: CAACATCTGCTTTGCTCTCAGG	499	X	X	X	X	X	X	X	
A13g62010_1725R (COSII)		R: GCTTCAGATTTGTTATTCATCCA									
A15g48300_1121F		F: GCTCCAGGAATAACTTCACTTCC	500	X	X	X	X	X	X	X	X
A15g48300_1621R (COSII)		R: GCAGTAAGGCATAATTTGTGGTTG									
A13g62010_1425F		F: GATAGCCATAGCATAAGCAACCC	501	X	X	X	X	X	X	X	
A13g62010_1926R (COSII)		R: AGGCAATCTGTGTTAGTAGATG									
A14g24830_948F		F: AAACAACAACCTCGCAGCC	499	X	X	X	X	X	X	X	
A14g24830_1447R (COSII)		R: AAGTGTCTCTGGCTCTACAGT									
A13g04260_147F		F: TCCTAGAAGATGGTCACAGGAGT	499	X	X	X	X	X	X	X	
A13g04260_646R (COSII)		R: CTCGTATACGAGCCGCGAAGT									
A15g27620_1180F		F: CGAATCCTAACTTGCAATGAAGA	500	X	X	X	X	X	X	X	
A15g27620_1680R (COSII)		R: CGATGCTTCAAAATTTCCAGT									
A15g26360_1312F		F: GGCAATGGTTTGGTTCTCACC	499	X	X	X	X	X	X	X	
A15g26360_1811R (COSII)		R: TGTGTGGACTTGTGAAGAGTTGT									
A12g29210_514F		F: TGAGACCATGGATTGCAAAAGC	500	X	X	X	X	X	X	X	
A12g29210_1014R (COSII)		R: GCTGTGCACTCAGAGCAG									
A13g62010_2_1674F		F: AGCCAGACCTTACTCATCTCT	500	X	X	X	X	X	X	X	
A13g62010_2_2174R (COSII)		R: CAGATCTTAAGGAAAGAGGCTGTAG									
A13g62010_1674F		F: AGCCAGCACTTACTCATCTCT	616	X	X	X	X	X	X	X	
A13g62010_2290R (COSII)		R: ATATCTGTGTATAGCATGGCAACT									
A12g38020_1797F		F: GTGCTATTATATGCACTGACTTGT	500	X	X	X	X	X	X	X	
A12g38020_2297R (COSII)		R: CAGCATGACGAGCAACATCTG									
A12g38770_387F		F: TTCTTAACACGCCCAAGCCA	499	X	X	X	X	X	X	X	
A12g38770_886R (COSII)		R: CCTAGCTCTTCCCTGGGCA									
A14g11120_1226F		F: TCCGCAATACGACCTATGCC	490	X	X	X	X	X	X	X	
A14g11120_1716R (COSII)		R: AATTTATGCTCAATTTGCTTGGC									
A11g80150_1016F		F: GCATGTAGGCGCAGCAGCTTA	512	X	X	X	X	X	X	X	
A11g80150_1528R (PPR)		R: GCGTTCGGGAAGGGTTTATA									
A11g04200_1908F		F: GAGAGTACTTGCCTGAAGC	487	X	X	X	X	X	X	X	
A11g04200_2395R (COSII)		R: GACTCCTCTTTCATGTCAGATG									
A12g07050_1067F		F: TGCAAACTGTAAAGAAGGCCCA	467	X	X	X	X	X	X	X	
A12g07050_1534R (COSII)		R: TGTAGAGAAAGATAACATTACAGCA									
A14g02720_1652F		F: CGAAATCTGTTTCCGAGAGCG	500	X	X	X	X	X	X	X	
A14g02720_2152R (COSII)		R: CATCTCCATCGACTCTGAGG									
A13g09920_917F		F: TGCAACGAGTTAAAGAATTCATCG	499	X	X	X	X	X	X	X	
A13g09920_1416R (COSII)		R: ACCTCCCAATGAAAGACCTTT									
A13g09920_1316F		F: ACCTCCCAATGAAAGACCTTT	543	X	X	X	X	X	X	X	
A13g09920_1859R (COSII)		R: GTCGATATGCGAAATGATGCC									

TABLE I. Continued.

Locus (Region)	Primer sequences (5'-3') ^a	Amplicon length (bp) ^b	Clade I: <i>Lindenbergia</i> sp. ^c	Clade II: <i>Schwalbea</i> <i>americana</i> ^c	Clade III: <i>Orobanchae</i> <i>californica</i> ^c	Clade IV: <i>Castilleja lineariloba</i> , <i>C. pumila</i> , <i>C. lemmonii</i> ^d	Clade IV: <i>Lamoureauxia virgata</i> ^c	Clade IV: <i>Pedicularis</i> sp. ^c	Clade V: <i>Neobanksia peruviana</i> ^c	Clade V: <i>Rhinanthus alectonophilus</i> ^c	Clade VI: <i>Harveya purpurea</i> ^c	Clade VI: <i>Physocalyx major</i> ^c	Paulowniaceae: <i>Paulownia fortunei</i> ^c (outgroup)
A15g26360_322F	F: TTACCGGTAATTCCTATATGC	496				X						X	
A15g26360_818R (COSII)	R: ATGCTATGTCTGTGCGACGA						X						
A15g49970_503F	F: AGCAGCTACAAGACATCGC	495				X							
A15g49970_998R (COSII)	R: AGCTGGGAAATGATGAATCC						X						
A15g52210_456F	F: ATCCAGCAGCAACCTCAAC	499				X							
A15g52210_955R (COSII)	R: GGATTTGCAAGGAGCTCCAC						X						
A1g09620_1187F	F: GGCTACACGCGAGCTGAAA	501				X							
A1g09620_1490F	R: ACCCGACACCTTCAGATCAA						X						
A1g09620_1688R (COSII)	F: AGCTCATATGTTGAGGAGG	499				X							
A1g09620_1989R (COSII)	R: CCTCTCCATCCATGCAATCT						X						
A1g28190_483F	F: GTTCCAGGTCACCCGC	500				X							
A1g28190_983R (COSII)	R: TGCAACCTGTTTCTGTAATGG						X						
A2g38020_892F	F: TTTGACATTTCTGTTGCTTACCA	499				X							
A2g38020_1391R (COSII)	R: CGTACAGCATTTAAGACGCG						X						
A4g22670_93F	F: GAATTTGTTGTTGGCATACAG	496				X							
A4g22670_589R (COSII)	R: AGTCTAGCTTTGAGCCACA						X						
A15g26360_1322F	F: TGTTCTCACCTTTCTTGTACTCA	499				X							
A15g26360_1821R (COSII)	R: GGTGCAATATGTTGGACTTG						X						
A2g29210_902F	F: TTGAATGCAAGCAGTTAATGAA	499				X							
A2g29210_1401R (COSII)	R: ACCGCCATACAGCAGTTTG						X						
A1g01940_828F	F: TCATAACTATGTAAGCCAAACCATGAT	399				X							
A1g01940_1227R (COSII)	R: TATAAAGGTTTTCATGATACAGGTGG						X						
A4g24620_1780F	F: CTGTTTCCCTGTATACCAAGATTG	443				X							
A4g24620_2223R (COSII)	R: CAGTTATATCAACAGCTCAGATGG						X						
A15g62530_439F	F: TGGAACAAAATGCTTCGATTACTCC	499				X							
A15g62530_938R (COSII)	R: TGCGGCTGAGAGAGTCATG						X						
A15g62010_1468F	F: ACAAATAGATATGACCTGCGCA	501						X					
A15g62010_1969R (COSII)	R: GCACATGGAGCAGGTTTG							X					
A4g11120_1399F	F: CCCATGTGAAGTGGGATA	460				X							
A4g11120_1859R (COSII)	R: GCTTGTGACGCAATCCAG						X						
A2g41680_402F	F: GTACACAAACAAGTCTCGGC	500				X							
A2g41680_902R (COSII)	R: GGATCCGGATGTTTGCTGT						X						
A15g22830_933F	F: AAAGACCTAATGCACTATTGTGAAC	499				X							
A15g22830_1432R (COSII)	R: ACCAGTCTCTGTTTCTAATACGA						X						
A15g26730_1284F	F: AGATACATCTTATCTGATCATGGA	486				X							
A15g26730_1770R (COSII)	R: TGTTTCAGGACAAGTGGCTCAG						X						
A2g17670_501F	F: ATCCAATAGCCGCGCTTAC	497				X							
A2g17670_998R (PPR)	R: AAGGAGTTCGGTATATAAGCCTGAT						X						
A4g37170_962F	F: CAACCTATACCGGGTTTCTTCAC	504				X							
A4g37170_1466R (PPR)	R: ATGGGTATGCTCAGAAATGGTCAG						X						
A4g30825_1603F	F: GGTGACATGGTGTATGATCGGT	499				X							
A4g30825_2102R (PPR)	R: TTCAAGACCGGACTTGACA						X						
A13g14730_878F	F: ACGGGTGTACCTAACAGG	500				X							
A13g14730_1378R (PPR)	R: ACGGGCTTCTCTCATGCTC						X						
A15g18475_417F	F: ACGGATGAAGCTATCGAGATGG	498				X							
A15g18475_915R (PPR)	R: AACATGACACACATAATATGCT						X						
A1g13040_1457F	F: TGTAGCTCACCATGTCAGGC	501				X							
A1g13040_1958R (PPR)	R: AGTGGAGCTCAAAATGATAGGT						X						
A1g16890_816F	F: AAACCATCGATCGTCTATTACTTTG	499				X							
A1g16890_1315R (PPR)	R: TCCACCTCCATTTGATGAAACTTG						X						
A1g20300_782F	F: GTGATCGAGCTTTGTGCC	504				X							
A1g20300_1286R (PPR)	R: TCGGCTCCACCTCATCTCA						X						
A1g07230_362F	F: GGACCTCATCAACGGTTC	508				X							
A1g07230_870R (COSII)	R: GTCAATCCACGGCGAAATCAA						X						
A1g17760_1370F	F: CTTGCTCATTCCTCACTTCGG	497				X							
A1g17760_1867R (COSII)	R: AACCATTCACGTATTGAAACCCA						X						
A1g26640_1171F	F: TGCAAGAGCCCTGACAATTC	499				X							
A1g26640_1670R (COSII)	R: TGGGTCCTCAGAAATTCATCTG						X						

TABLE I. Continued.

Locus (Region)	Primer sequences (5'-3') ^a	Amplicon length (bp) ^b	Clade I: <i>Lindenbergia</i> sp. ^c	Clade II: <i>Schwalbea americana</i> ^c	Clade III: <i>Orobanchaceae</i>	Clade IV: <i>Castilleja lineariloba</i> , <i>C. pumila</i> , <i>C. lemmingii</i> ^d	Clade IV: <i>Lamoureauxia virgata</i> ^e	Clade IV: <i>Pedicularis</i> sp. ^e	Clade V: <i>Neobanksia peruviana</i> ^e	Clade V: <i>Rhinanthus alectonolophus</i> ^e	Clade VI: <i>Harveya purpurea</i> ^e	Clade VI: <i>Physocalyx major</i> ^e	Paulowniaceae: <i>Paulownia fortunei</i> ^e (outgroup)
At1g30580_2218F	F: TGGAAGGCTGCTGAAGTTGA	500				X							
At1g30580_2718R (COSII)	R: GGCAATGATCTTTGGGAGAAAC												
At1g43860_2011F	F: AGCATACAATTCAGGAGTGCATC	504				X							
At1g43860_2515R (COSII)	R: CTCGGCGACCATCACCAC												
At1g59600_576F	F: GGAGACCTGAAACTAGGTCC	499				X							
At1g59600_1075R (COSII)	R: AAGGTGTTGCTGAGGCTC												
At1g69220_607F	F: AGCCTCGTCGTAACATCA	467				X							
At1g69220_1075R (COSII)	R: GGCACTGTATAAAGCTCGAGA												
At2g03120_425F	F: TGGGATCTGTTGTAGTTTATCCAGA	500				X							
At2g03120_925R (COSII)	R: ACGAAGACGACATCCCT												
At2g27170_2888F	F: CCATCTTCATCCCTCCTGC	475				X							
At2g27170_3366R (COSII)	R: TTGACAGTCAATGGCTCAGCC												
At2g42490_4066F	F: ACTTAGCCTTTGGACCAACC	475				X							
At2g42490_4564R (COSII)	R: GCTGGAGAAACGACAACTC												
At3g03600_524F	F: AGGCTCGAACTGTGTTGTC	499				X							
At3g03600_1023R (COSII)	R: GCCTTAAACCCCTCAACTC												
At3g04870_957F	F: TCTTCCCATCTTCCCAAGGTC	455				X							
At3g04870_1412R (COSII)	R: ACTCAAAACAGGCTCAACTCA												
At3g05000_1412F	F: ACCAACTAAACAGAGCCCTCAATCA	500				X							
At3g09090_1427F	R: TTGCACTTTTACTTTCCCTGTGG												
At3g09090_1912R (COSII)	F: TTTCTGTTTGTGATAGTTTACAGT	485				X							
At3g09090_1912R (COSII)	R: ATGCCGCGACAAATGATCCA												
At3g13235_2063F	F: GGCAACATGATCCGGCCTA	482				X							
At3g13235_2545R (COSII)	R: TTGGGAAACTAAACCAACTCTGC												
At3g16150_956F	F: ACACAATTAACAGATCAATGG	404				X							
At3g16150_956F	R: CCTTATTTGCTCGAATGCTCG												
At3g16150_1360R (COSII)	F: TCAAGTCGAAACCAAGCATCTGA	499				X							
At3g17205_926F	R: GGACTGCTGGTAATGCATCTG												
At3g17205_1425R (COSII)	F: CAGAACTCGGAGACTTCTCAAC	499				X							
At3g55560_121F	R: GCCACCGTCCTTCAATFACA												
At3g55360_620R (COSII)	F: CATAACTGCTTGGGATTAACCG	499				X							
At3g62580_1318F	R: GCAATCTGTTATCTTTACGCC												
At3g62580_1817R (COSII)	F: AAGCTGCCATGCTTTAAATGTC	494				X							
At4g00560_2034F	R: GGAACTGCTCATCTCTCTCTTG												
At4g12230_1167F	F: TGTAGATTGGATCAAGCAATGGC	417				X							
At4g12230_1167F	R: TCAATGCTTGAATCCACTCATGG												
At4g12230_1584R (COSII)	F: ATCACTCAAGAGAGAGGCGG	499				X							
At4g23100_808F	R: CCTTCATATATCGTATTCTACACGG												
At4g23100_1307R (COSII)	F: CCAACCAACTGCTATGTCAGC	499				X							
At4g29490_2406F	R: CCTAGCTCACCTCATGGCAG												
At4g32280_1754F	F: CCGATGGCTAATCAACGCC	500				X							
At4g32280_2254R (COSII)	R: ATACGTTGCACAGACTCGACA												
At4g33625_852F	F: AGTAGGCGTCATTGTAGAACC	499				X							
At4g33625_1351R (COSII)	R: CATCTCTGCTCAAGAAATTGG												
At4g38630_507F	F: GTTACCTGTTCTAGGACGTGC	498				X							
At4g38630_1005R (COSII)	R: ATCAAGTTGGGATCCACACC												
At5g06830_1899F	F: GGCCCAACCCATTGCCATTA	499				X							
At5g06830_2398R (COSII)	R: ACTAGACTCTGGACTTTGCAAG												
At5g07270_1779F	F: CACATCCTTGCAACCTCAAGC	499				X							
At5g07270_2278R (COSII)	R: ACTGCTCTTCATATGGCTGCT												
At5g12200_438F	F: TCGGTTCTTACTAGTGTCAAC	499				X							
At5g12200_937R (COSII)	R: ATCAGAGTTAGGGATGCTGTGT												
At5g13640_387F	F: GGGAACCTGATTACAGGGAATC	498				X							
At5g13640_885R (COSII)	R: CTAGCTACGAGCAATGCGC												
At5g18580_933F	F: GCTGCTTTTATGACGGTGT	499				X							
At5g18580_1432R (COSII)	R: ACAGTTACCTCGGTGAAGAG												

TABLE 1. Continued.

Locus (Region)	Primer sequences (5'-3') ^a	Amplicon length (bp) ^b	Clade I: <i>Lindenbergia</i> sp. ^c	Clade II: <i>Schwalbea americana</i> ^c	Clade III: <i>Orobanchaceae</i>	Clade IV: <i>Castilleja lineariloba</i> , <i>C. pumila</i> , <i>C. lemmonii</i> ^d	Clade IV: <i>Lamoureauxia virgata</i> ^e	Clade IV: <i>Pedicularis</i> sp. ^e	Clade V: <i>Neobartsia peruviana</i> ^e	Clade V: <i>Rhinanthus alectonophilus</i> ^e	Clade VI: <i>Harveya purpurea</i> ^e	Clade VI: <i>Physocalyx major</i> ^e	Paulowniaceae: <i>Paulownia fortunei</i> ^e (outgroup)
A15g20080_925F	F: ATACGAGATAGCATTTGGTACTCGA	499				X							
A15g20080_L1424R (COSII)	R: AGATGCCATCCTTAAGAACCCCTG												
A15g23540_L1611F	F: AATCCTTCCATGAGTGAACCCCA	491				X							
A15g23540_L2102R (COSII)	R: ATTACTCTCAACATGGATGATAGTTGG												
A15g58490_254F	F: GCTATCAATCCAGCCCTAATTG	410				X							
A15g58490_664R (COSII)	R: TTGCAATCTGTAAATTATGAGCCCA												

^aPrimer sequence for the “*Castilleja*-specific primer.” To make the target-specific primer for subsequent microfluidic PCR, conserved sequence tags CS1 (5'-ACACTGACGACATGGTTCTTACA) and CS2 (5'-TACGGTAGCAGAGACTTGGTCT) were added to each forward and reverse primer, respectively.

^bAmplicon length (bp) estimated from BLAST hit alignments.

^cPCR validations using DNAs from Bennett and Mathews (2006).

^dPCR validations were considered successful for *Castilleja* when amplification occurred for all three taxa, representing one annual lineage (*C. lineariloba*) and two perennial lineages (*C. pumila* and *C. lemmonii*).

^eTaxa that were both PCR validated and had primer combinations evaluated in silico against their respective assemblies (raw read files available in the NCBI Sequence Read Archive submission SRP100222).

Li et al., 2008; Yuan et al., 2009, 2010). We also test these primers for their broader applicability across Orobanchaceae following the approach outlined in Latvis et al. (2017a) with the goal of finding a subset of nuclear gene regions that would amplify across this more inclusive clade. Previous phylogenetic studies within Orobanchaceae employed the nuclear phytochrome genes *PHYA* and *PHYB*, the nuclear ribosomal ITS region, and the plastid *matK* and *rps2* genes. Orobanchaceae is the largest clade of parasitic angiosperms, and plastome reduction and accelerated rates of molecular evolution in retained plastid genes have been documented (see discussion in Bennett and Mathews, 2006). Additionally, phytochrome genes regulate responses to light and can be significantly modified in parasites (Bennett and Mathews, 2006). Therefore, the development of additional single-copy nuclear regions would provide a much-needed source of phylogenetic information in Orobanchaceae, and may provide a more reliable estimate of branch lengths for further studies of diversification and character evolution.

METHODS AND RESULTS

We assembled contigs from raw reads of three low-coverage *Castilleja* genomes, *C. cusickii* Greenm., *C. foliolosa* Hook. & Arn., and *C. tenuis* (A. Heller) T. I. Chuang & Heckard (Latvis et al., 2017a; National Center for Biotechnology Information [NCBI] Sequence Read Archive [SRA] accession SRP100222) using CAP3 (Huang and Madan, 1999) with the default settings. The accessions were sequenced as 100-bp single-end reads on an Illumina HiSeq 2000 (Illumina, San Diego, California, USA), yielding ~12.5 million reads per taxon (Uribe-Convers et al., 2014) and an average depth of coverage of ~0.8x. These taxa include both annual and perennial lineages of *Castilleja* and span the phylogenetic breadth of the clade (Tank and Olmstead, 2008, 2009). These assemblies were then culled to include only contigs of 1 Kb or larger using a custom R script. The culled assemblies and script are available from the Dryad Digital Repository (<http://doi.org/10.5061/dryad.52v62>; Latvis et al., 2017b).

To search for hits among our contigs, available COS sequences were obtained from Sol genomics (<https://solgenomics.net>), and PPR loci were mined from the *Mimulus* L. genome on Phytozome (Hellsten et al., 2013; <https://phytozome.jgi.doe.gov>) using the 127 PPR orthologs identified in *Arabidopsis* Heynh. by Yuan et al. (2009) as references. Both gene sets may be found in Uribe-Convers et al. (2016) and were used to construct local BLAST databases for the search (-makeblastdb). We used TBLASTX to search each *Castilleja* CAP3 assembly (with contigs of 1 KB or greater) against both the COS and PPR databases, indicating tab-delimited output (-outfmt 6). Output files were filtered for alignment length >200 and a maximum E-value of 1e-10, and were culled to include only unique hits.

Hits shared between *C. cusickii*, *C. foliolosa*, and *C. tenuis* were placed together into individual FASTA files (data available from the Dryad repository: <http://doi.org/10.5061/dryad.52v62>; Latvis et al., 2017b), imported into Geneious R7 version 7.0.6 (Kearse et al., 2012), and aligned with MAFFT version 7.017b under the default settings (Katoh and Standley, 2013). We designed primer pairs using Primer3 (Untergasser et al., 2012) using the same specifications for the Fluidigm Access Array system as Latvis et al. (2017a), but with a size range between 400–525 bp and an optimal size of 500 bp. We designed 10–30 primer pairs for each identified locus and prioritized them based on desired size and the presence of multiple G or C bases at the 3' end of the primers (GC clamp). This also allowed us to design overlapping sets of primers with the potential to combine them after sequencing to produce longer contigs for downstream phylogenetic analyses. Suitable primer pairs were validated for *Castilleja* with PCR following the same amplification protocol and using the same *Castilleja* accessions as in Latvis et al. (2017a) and visualized on an agarose gel. We present 87 nuclear primer combinations specifically designed and validated for *Castilleja* (Table 1).

To investigate whether any of these primer combinations would amplify successfully across Orobanchaceae, we searched for our selected *Castilleja* primers against an assembled low-coverage genome for *Lamoureauxia multifida* Kunth using BLAST (Altschul et al., 1990). *Lamoureauxia multifida* was sequenced on an Illumina HiSeq 2000 at the University of Oregon as 100-bp paired-end reads, and contigs were assembled using SPAdes (Bankevich et al., 2012) under the

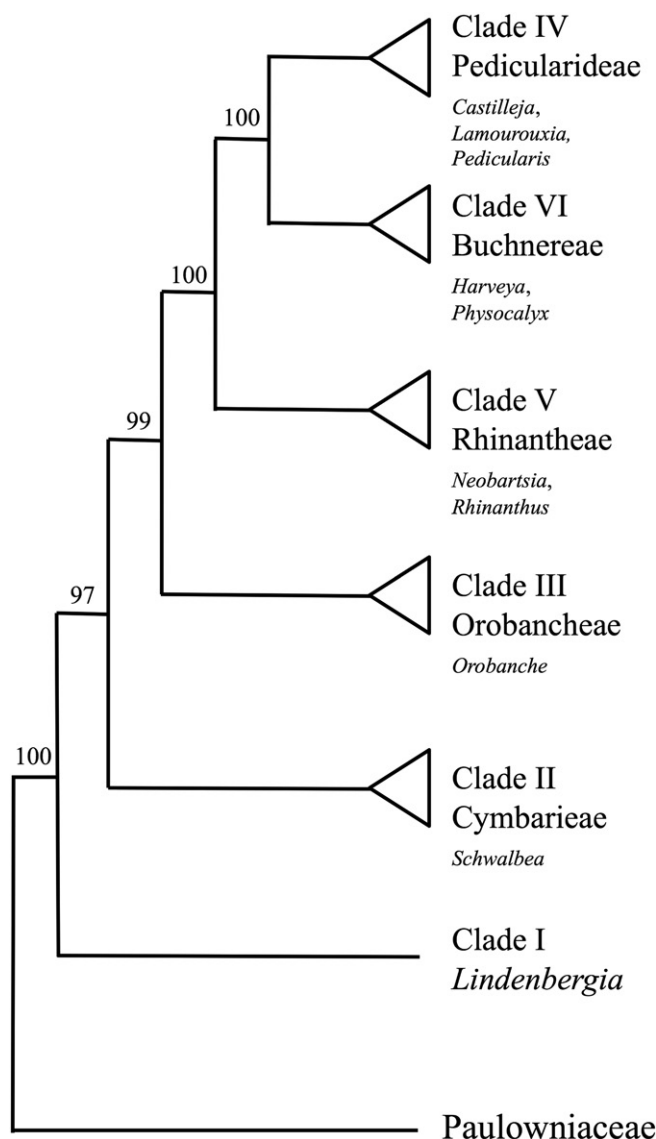


Fig. 1. Relationships among major clades within Orobanchaceae modified from McNeal et al. (2013) with taxa used for primer validation indicated (see text). Bootstrap support values for clades are indicated along the branches and follow McNeal et al. (2013). Reproduced from Latvis et al. (2017a).

default settings. BLAST search parameters, assessment of suitable hits, and subsequent PCR validation with *Lamourouxia virgata* Kunth, *Physocalyx major* Mart., and *Neobartsia filiformis* (Wedd.) Uribe-Convers & Tank are described in Latvis et al. (2017a). Primer combinations with successful amplification in *Lamourouxia* Kunth and at least one other taxon were selected for further PCR testing with other major lineages in Orobanchaceae (sensu McNeal et al., 2013; Fig. 1). This second round of PCR validation follows Latvis et al. (2017a), except that two of the accessions used for testing were changed. As in Latvis et al. (2017a), we also included a negative control and conserved sequence-tagged “universal” primers for the *trnL-F* region as a positive control for all primer pairs. We used *Neobartsia peruviana* (Walp.) Uribe-Convers & Tank instead of *N. filiformis*, and *Paulownia fortunei* (Seem.) Hemsl. instead of *P. elongata* Siebold & Zucc. (Appendix 1). Of the 87 nuclear primer combinations specifically designed for *Castilleja*, we identified 27 with broader applicability in Orobanchaceae, chosen if they successfully amplified in Pedicularideae (Clade IV; including *Castilleja*, *Lamourouxia*, and *Pedicularis* L.) and at least one of the other major clades highlighted in Fig. 1. Validation results are presented in Table 1 with these “core Orobanchaceae” combinations boldfaced.

CONCLUSIONS

We present 87 nuclear primer pairs specifically designed for *Castilleja* that target COSII and PPR loci. Although we target the same putative single-copy nuclear domains as previous studies (Wu et al., 2006; Li et al., 2008; Yuan et al., 2010; Blischak et al., 2014; Uribe-Convers et al., 2016), we developed primers for different loci and present unique primer combinations in this study. As with our chloroplast primers (Latvis et al., 2017a), all combinations were designed with the Fluidigm microfluidic PCR system in mind, allowing for parallelization of amplification for downstream high-throughput sequencing platforms. Of these, we identify a set of 27 primer combinations with broader utility across Orobanchaceae. The development of primers for putatively single-copy nuclear loci will greatly enhance efforts to understand evolutionary history at multiple taxonomic scales, both for *Castilleja* and across Orobanchaceae.

LITERATURE CITED

- ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MEYER, AND D. J. LIPMAN. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.
- BANKEVICH, A., S. NURK, D. ANTIPOV, A. GUREVICH, M. DVORKIN, A. S. KULIKOV, V. LESIN, ET AL. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19: 455–477.
- BENNETT, J. R., AND S. MATHEWS. 2006. Phylogeny of the parasitic plant family Orobanchaceae inferred from phytochrome A. *American Journal of Botany* 93: 1039–1051.
- BLISCHAK, P. D., A. J. WENZEL, AND A. D. WOLFE. 2014. Gene prediction and annotation in *Penstemon* (Plantaginaceae): A workflow for marker development from extremely low-coverage genome sequencing. *Applications in Plant Sciences* 2: 1400044.
- FANT, J. B., H. WOLF-WEINBERG, D. C. TANK, AND K. A. SKEGEN. 2013. Characterization of microsatellite loci in *Castilleja sessiliflora* and transferability to 24 *Castilleja* species (Orobanchaceae). *Applications in Plant Sciences* 1: 1200564.
- GODDEN, G. T., I. E. JORDAN-THADEN, S. CHAMALA, A. A. CROWL, N. GARCÍA, C. C. GERMAIN-AUBREY, J. M. HEANEY, ET AL. 2012. Making next-generation sequencing work for you: Approaches and practical considerations for marker development and phylogenetics. *Plant Ecology and Diversity* 5: 427–450.
- HELLSTEN, U., K. M. WRIGHT, J. JENKINS, S. SHU, Y. YUAN, S. R. WESSLER, J. SCHMUTZ, ET AL. 2013. Fine-scale variation in meiotic recombination in *Mimulus* inferred from population shotgun sequencing. *Proceedings of the National Academy of Sciences USA*.
- HUANG, X., AND A. MADAN. 1999. CAP3: A DNA sequence assembly program. *Genome Research* 9: 868–877.
- KATOH, K., AND D. M. STANDLEY. 2013. MAFFT Multiple Sequence Alignment Software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.
- KEARSE, M., R. MOIR, A. WILSON, S. STONES-HAVAS, M. CHEUNG, S. STURROCK, S. BUXTON, ET AL. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics (Oxford, England)* 28: 1647–1649.
- LATVIS, M., S. M. E. MORTIMER, D. F. MORALES-BRIONES, S. TORPEY, S. URIBE-CONVERS, S. J. JACOBS, S. MATHEWS, AND D. C. TANK. 2017a. Primers for *Castilleja* and their utility across Orobanchaceae: I. Chloroplast primers. *Applications in Plant Sciences* 5: 1700020.
- LATVIS, M., S. J. JACOBS, S. M. E. MORTIMER, M. RICHARDS, P. D. BLISCHAK, S. MATHEWS, AND D. C. TANK. 2017b. Data from: Primers for *Castilleja* and their utility across Orobanchaceae. II. Single-copy nuclear loci. Dryad Digital Repository. <http://doi.org/10.5061/dryad.52v62>.
- LI, M., J. WUNDER, G. BISSOLI, E. SCARFONI, S. GAZZANI, E. BARBARO, H. SAEDLER, AND C. VAROTTO. 2008. Development of COS genes as universally applicable markers for phylogenetic reconstructions of closely related plant species. *Cladistics* 24: 727–745.

- MCNEAL, J. R., J. R. BENNETT, A. D. WOLFE, AND S. MATHEWS. 2013. Phylogeny and origins of holoparasitism in Orobanchaceae. *American Journal of Botany* 100: 971–983.
- TANK, D. C., AND R. G. OLMSTEAD. 2008. From annuals to perennials: Phylogeny of subtribe *Castillejinae* (Orobanchaceae). *American Journal of Botany* 95: 608–625.
- TANK, D. C., AND R. G. OLMSTEAD. 2009. The evolutionary origin of a second radiation of annual *Castilleja* (Orobanchaceae) species in South America: The role of long distance dispersal and allopolyploidy. *American Journal of Botany* 96: 1907–1921.
- TWYFORD, A. D., AND R. A. ENNOS. 2012. Next-generation hybridization and introgression. *Heredity* 108: 179–189.
- INTERGASSER, 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., J. R. DUKE, M. J. MOORE, AND D. C. TANK. 2014. A long-PCR based method for chloroplast genome enrichment and phylogenomics in angiosperms. *Applications in Plant Sciences* 2: 1300063.
- 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.
- WU, F., L. A. MUELLER, D. CROUZILLAT, V. PÉTIARD, AND S. D. TANKSLEY. 2006. Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: A test case in the Euasterid plant clade. *Genetics* 174: 1407–1420.
- YUAN, Y.-W., C. LIU, H. E. MARX, AND R. G. OLMSTEAD. 2009. The pentatricopeptide repeat (PPR) gene family, a tremendous resource for plant phylogenetic studies. *New Phytologist* 182: 272–283.
- YUAN, Y., C. LIU, H. E. MARX, AND R. G. OLMSTEAD. 2010. An empirical demonstration of using pentatricopeptide repeat (PPR) genes as plant phylogenetic tools: Phylogeny of Verbenaceae and the *Verbena* complex. *Molecular Phylogenetics and Evolution* 54: 23–35.

APPENDIX 1. Voucher information for species used in this study.

Species	Voucher accession no. (Herbarium) ^a	Collection locality	Geographic coordinates
<i>Castilleja cusickii</i> Greenm.	Tank 2009-01 (ID)	Idaho, USA	45.884241°N, 116.230195°W
<i>Castilleja foliolosa</i> Hook. & Arn.	A. Colwell 03-09 (YM)	California, USA	35.3926°N, 120.3522°W
<i>Castilleja lemmonii</i> A. Gray	Jacobs 2015-088 (ID)	California, USA	37.907982°N, 119.258583°W
<i>Castilleja lineariloba</i> (Benth.) T. I. Chuang & Heckard	Tank 2002-04 (WTU)	California, USA	37.41387°N, 120.10833°W
<i>Castilleja pumila</i> Wedd.	Uribe-Convers 2011-120 (ID)	La Libertad, Peru	7.99506°S, 78.44197°W
<i>Castilleja tenuis</i> (A. Heller) T. I. Chuang & Heckard	Tank 2001-13 (WTU)	Washington, USA	46.118133°N, 121.5158°W
<i>Harveya purpurea</i> Harv.	Randle 79 (OS)	NA	NA
<i>Lamouroxia multifida</i> Kunth	Mejia 695 (CAS)	Chiapas, Mexico	NA
<i>Lamouroxia virgata</i> Kunth	Mejia 581 (CAS)	Chiapas, Mexico	16.713611°N, 92.614722°W
<i>Lindenbergia</i> sp. Kunth	Armstrong 1163 (ISU)	NA	NA
<i>Neobartsia peruviana</i> (Walp.) Uribe-Convers & Tank	Uribe-Convers 13-011 (ID)	NA	NA
<i>Orobanche californica</i> Cham. & Schltdl.	Bennett 72 (A)	Cultivated	Cultivated
<i>Paulownia fortunei</i> (Seem.) Hemsl.	s.n. (A)	Cultivated	Cultivated (https://sheffields.com)
<i>Pedicularis</i> sp. L.	Krasek and Bennett s.n. (A)	Slovenia	NA
<i>Physocalyx major</i> Mart.	G. O. Romão 2528 (ESA)	Minas Gerais, Brazil	19.2635°S, 43.5508°W
<i>Rhinanthus alectorolophus</i> (Scop.) Pollich	Bennett 85 (A)	NA	NA
<i>Schwalbea americana</i> L.	Kirkman s.n. (PAC)	NA	NA

Note: NA = not available.

^aHerbarium acronyms are per Index Herbariorum (<http://sweetgum.nybg.org/science/ih/>).