

PRIMERS FOR *CASTILLEJA* AND THEIR UTILITY ACROSS OROBANCHACEAE: I. CHLOROPLAST PRIMERS¹

MARIBETH LATVIS^{2,8}, SEBASTIAN M. E. MORTIMER^{3,4}, DIEGO F. MORALES-BRIONES^{3,4,5}, SAMUEL TORPEY^{3†}, SIMON URIBE-CONVERS⁶, SARAH J. JACOBS^{3,4,5}, 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; ⁶Department of Ecology and Evolutionary Biology, University of Michigan, 830 North University Avenue, Ann Arbor, Michigan 48109 USA; and ⁷Australian National Herbarium, CSIRO National Research Collections, Canberra, Australia

- **Premise of the study:** Chloroplast primers were developed from genomic data for the taxonomically challenging genus *Castilleja*. We further tested the broader utility of these primers across Orobanchaceae, identifying a core set of chloroplast primers amplifying across the clade.
- **Methods and Results:** Using a combination of three low-coverage *Castilleja* genomes and sequence data from 12 *Castilleja* plastomes, 76 primer combinations were specifically designed and tested for *Castilleja*. The primers targeted the most variable portions of the plastome and were validated for their applicability across the clade. Of these, 38 primer combinations were subsequently evaluated in silico and then validated across other major clades in Orobanchaceae.
- **Conclusions:** These results demonstrate the utility of these primers, not only across *Castilleja*, but for other clades in Orobanchaceae—particularly hemiparasitic lineages—and will contribute to future phylogenetic studies of this important clade of parasitic plants.

Key words: *Castilleja*; chloroplast; hemiparasite; high-throughput sequencing; microfluidic PCR; Orobanchaceae.

The plastome is heavily relied upon in plant systematics, owing to its conserved nature and orthology, particularly for the study of deeper evolutionary divergences. Moreover, discordance between the uniparentally inherited plastome and the biparentally inherited nuclear genome may provide insights into introgression events and their direction (Twyford and Ennos, 2012). However, the low rate of molecular evolution in the plastome can become a hindrance when reconstructing relationships between closely related taxa, requiring large amounts of data to resolve these relationships (Uribe-Convers et al., 2016). In an attempt to alleviate this problem, several recent studies have leveraged available high-throughput sequencing data for the development of variable taxon-specific plastid (and nuclear) regions (e.g., Uribe-Convers et al., 2016).

Castilleja L. (Orobanchaceae; “the paintbrushes”) is a taxonomically challenging clade that includes ~200 hemiparasitic species, many of which have a complicated history of polyploidy and/or hybridization (Heckard and Chuang, 1977). Microsatellite markers have been developed in *Castilleja* for population genetic studies (Fant et al., 2013), and broader, genus-wide phylogenetic reconstructions within *Castilleja* used two chloroplast regions (*trnL-F* and the *rps16* intron), nuclear ribosomal spacers (ITS and ETS), and a low-copy nuclear gene (*waxy*) (Tank and Olmstead, 2008, 2009). However, species-level relationships lacked resolution in Tank and Olmstead (2008, 2009), limiting conclusions regarding diversification and hybridization. Here, we follow Uribe-Convers et al. (2016) for primer design and validation of the most highly variable chloroplast regions in *Castilleja*. Because these primers were designed for the Fluidigm Access Array microfluidic PCR system (Fluidigm, South San Francisco, California, USA), annealing temperature specifications are consistent across all primer combinations; this allows for parallelization of PCR and is ideal for high-throughput sequencing platforms (see Uribe-Convers et al., 2016 for application of this approach). Although our initial focus was the development of *Castilleja*-specific primers, we evaluated their utility in silico in three other lineages of Orobanchaceae to obtain a subset of “core” chloroplast primers with the potential to amplify across the clade. Once identified, we surveyed this set of core primers to assess their

¹Manuscript received 27 February 2017; revision accepted 10 April 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-1210895 (awarded to D.C.T. for S.U.C.).

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

[†]Deceased 21 May 2017.

doi:10.3732/apps.1700020

TABLE 1. All primer pair sequences designed for *Castilleja* (names 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). Combinations are listed from most variable to least variable, according to our prioritization scheme (see text). 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: <i>Lindenbergia</i> sp. ^c	Clade II: <i>Schwalbea</i> <i>americana</i> ^c	Clade III: Orobanchae <i>californica</i> ^c	Clade IV: <i>Lamourovaxia</i> <i>virgate</i> ^c	Clade IV: <i>Necobartsia</i> <i>filiformis</i> ^c	Clade V: <i>Rhinanthus</i> <i>electrolophus</i> ^c	Clade VI: <i>Harveya</i> <i>purpurea</i> ^c	Clade VI: <i>Physocodyx</i> <i>major</i> ^c	Paulowniaceae: <i>Paulownia</i> <i>elongata</i> ^c (outgroup)
Cas_120561_F_	F: GTCRAAAGCATCCCATACCA	810									
Cas_121371_R	R: TTTTAGCTCGGTTACCGGTGT		X								
Cas_111970_F_	F: GGTGAAAGTGAGGAGAAAGA	819									
Cas_112789_R	R: TCAAGAAAGACAGGTTTGA		X								
Cas_129331_F_	F: TGAGTTTAAATCAACCCGGAGA	795				X					
Cas_130126_R	R: GACCCCTTCTGAACAAATCA										
Cas_112854_F_	F: ACATAGTATGTCGATTCATAAGGA	892									
Cas_113746_R	R: GGAGGAGCCCACTCCTATT		X								
Cas_126859_F_	F: GAACGGATCCAAAGATCTCCTC	854									
Cas_127713_R	R: GCGAAATGCGCTTCATA		X								
Cas_59866_F_	F: TTGCCGTCAAAGACATTCG	758			X						
Cas_60624_R	R: GCGTGTGAAACAGCCCTCAG						X		X		X
Cas_127891_F_	F: GGATTCCTTGATAGTGAAGACAGA	529									
Cas_128420_R	R: GAAGGATCTGGACGATCGAA		X								
Cas_130168_F_	F: ACAACCGAGTCTTGTTCAAA	592									
Cas_130760_R	R: GGTGAAAGTGAGGAGAAAGA		X								
Cas_126110_F_	F: TCTAATCGATAATTAGGCCAAAAGA	758									
Cas_126868_R	R: GGATCGTCTTATCAACAACA		X								
Cas_32159_F_	F: AATCGGATCAATATCAATATACAA	586			X						
Cas_32745_R	R: ATTCGCCAATATACAGCAG						X		X		X
Cas_77140_F_	F: TGTCCGAATGGCTCTATTG	894									
Cas_78034_R	R: TCATCTCGTACAGCTCAAGCA		X								
Cas_10778_F_	F: TCAGTTTGTGATGATCCTTTGATGA	747									
Cas_46472_F_	F: GGAAGTATCCGATTTTCATTG	690									
Cas_47162_R	R: CTPAATGTTGGAAATAAGGCTCCT		X								
Cas_47758_F_	F: CGCGATACGGACGATTTTA	880									
Cas_48638_R	R: GAATCTGTATCAAGAAATCGAAATTA		X								
Cas_17609_F_	F: AACTCTCCAGAGCCCGTAT	803									
Cas_18412_R	R: CCACGATAGACAGAAACAATCA		X								
Cas_33546_F_	F: GACTCGTTTGGAAATTAATCAA	953									
Cas_34499_R	R: CTTCAACCATTTCCGAGCAC		X								
Cas_67504_F_	F: TTGTACCGAGGGGCATCTTTAG	839									
Cas_68343_R	R: ACCGGAATTAATCGTTATAGTAAGCA		X								
Cas_72399_F_	F: GGCCTGGTTAGATTGATCCT	846									
Cas_73245_R	R: TTTCAITGGATTATGATCGAGAG		X								
Cas_62840_F_	F: ATTCGTTGACCCGCTACTGA	932			X						
Cas_63772_R	R: AAGAGAGAAATCCACCAAGGTAAA						X		X		X

TABLE I. Continued.

Locus (Region)	Primer sequences (5'-3') ^a	Amplicon length (bp) ^b	Clade I: <i>Lindenberglia</i> sp. ^c	Clade II: <i>Schwalbea</i> <i>ameritcanae</i>	Clade III: <i>Orobanche</i> <i>californiceae</i>	Clade IV: <i>Castilleja</i> <i>lineariloba</i> , <i>C. paumotu</i> , <i>C. lemmonii</i> ^d	Clade IV: <i>Lamoureauxia</i> <i>virgate</i> ^e	Clade IV: <i>Pedicularis</i> sp. ^c	Clade V: <i>Neobartisia</i> <i>filiformis</i> ^e	Clade V: <i>Rhinanthus</i> <i>electrolophilus</i> ^e	Clade VI: <i>Harveya</i> <i>purpurea</i> ^e	Clade VI: <i>Physoclyx</i> <i>major</i> ^e	Paulowniaceae: <i>Paulownia</i> <i>elongata</i> ^e (outgroup)
Cas_65707_F_	F: CTGGGAATCCCTTGTACC	927				X							
Cas_66634_R	R: TCCGGATTAGTTATCCCTTA						X						
(<i>petA-psbJ</i>)													
Cas_69456_F	F: CAACCTAAGCGACCCCTAAATACA	718				X							
Cas_70174_R	R: TACTCGCGATCTTTCTCT												
(<i>trnP-trp33</i>)													
Cas_4537_F	F: GFTTCCTTGACCAACACACAG	782				X							
Cas_5319_R	R: ATCCCAACAACACAGACTTC												
(<i>trnK</i>)													
Cas_48611_F	F: TGTAAATTCGATTTCTTGTACAAAT	909				X			X				
Cas_49520_R	R: CAGATACAGATTTGGCCATC												
(<i>trnF-trnL</i>)													
Cas_29425_F	F: TTGAAGCGAAGTAGGATAAATTGA	866				X							
Cas_30291_R	R: TCTACCAGAGGCTACAACTGA												
(<i>trnB-psbN</i>)													
Cas_125001_F	F: AATAATCCCAACGCCCTTACA	858				X							
Cas_125859_R	R: AATGTTTCAATTAGCTCTCGAAATG												
(<i>psbI-5-ndhH</i>)													
Cas_21290_F	F: TGTTCGATTCACAVATGATCGTTT	746	X	X		X			X	X	X		X
Cas_22036_R	R: CGTGAAGGCTTTCTTTTACA												
(<i>trnE</i>)													
Cas_20851_F	F: CTGGTAGAGAGTGGTGGATCT	456				X							
Cas_21307_R	R: GFTTGAATTTGGGAGAAGCTG												
(<i>trnG-atpA</i>)													
Cas_11589_F	F: AGCTTCCCAAGCTAACGATG	872				X			X				
Cas_12461_R	R: CTGGAATCAGACCCGCTATT												
(<i>trnG-atpA</i>)													
Cas_47139_F	F: GGAGACCTTTTATCCACCAGTTAG	550				X							
Cas_47689_R	R: TTTTCGATTTGGGTATGGCTTC												
(<i>trnS-trpA</i>)													
Cas_14073_F	F: TTCGATTCATTTGGCTCTCA	651	X	X		X			X	X	X		X
Cas_14724_R	R: TGGAAAGGAGTGTGTGTGA												
(<i>atpF</i> intron)													
Cas_122476_F	F: ACGGCTCCATAGGTCACA	797				X							
Cas_123331_R	R: TCGTGTAAAAGAAATTCATCTCA												
(<i>ndhA</i> intron)													
Cas_73947_F	F: TCTTGTTCCTGAAATGGGCTC	551				X							
Cas_74498_R	R: GTTACGTTTCCACATCAAAAGTA												
(<i>clpP</i>)													
Cas_123306_F	F: AATGAGATGAATTCCTTTAACAGC	798	X			X							X
Cas_124104_R	R: TGAATTTGGCTGATATTTATGAGC												
(<i>ndhH-ndhA</i>)													
Cas_24256_F	F: ATAAACCGGTAATCGGAAG	781				X							
Cas_25037_R	R: TETCATCCCAAGTCAATCCAA												
(<i>trnO1</i>)													
Cas_85769_F	F: CATCAGGATATACCATAGTTGGCTTT	648				X							
Cas_86417_R	R: CCATAGATTTGCCCTTCATA												
(<i>trnL2</i>)													
Cas_36699_F	F: GCGGTCGCGAGAATATATGA	745				X							
Cas_37444_R	R: TTAATTCACAATGGGAATCCCTG												
(<i>psbC-psbZ</i>)													
Cas_61880_F	F: GCAATGGCTCTTTTATTTCTTCA	951	X	X		X							X
Cas_62831_R	R: GGCCTCGGATGCTCCATATA												
(<i>psal-ycf4</i>)													
Cas_71554_F	F: TCCAAATGGCTTCGGCTACTA	877				X							
Cas_72431_R	R: AATCATCCGGTTAGGATCAATCT												
(<i>trnD-trnI2</i>)													
Cas_5508_F	F: CCCATTCATTTCCCTTTAAATCCG	722				X							
Cas_6230_R	R: TTAGCTCAACAGTTTGTATAGCTTG												
(<i>trnK-trnS</i> (b))													

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</i> <i>amertcane</i>	Clade III: <i>Obobanche</i> <i>californice</i>	Clade IV: <i>Lamourosaxia</i> <i>virgate</i>	Clade IV: <i>Pedicularis</i> sp. ^c	Clade V: <i>Neobarbista</i> <i>filiformis</i> ^c	Clade V: <i>Rhinanthus</i> <i>electrolophilus</i> ^c	Clade VI: <i>Harveya</i> <i>purpurea</i> ^c	Clade VI: <i>Physocadyx</i> <i>major</i> ^c	Paulowniaceae: <i>Paulownia</i> <i>elongata</i> ^c (outgroup)
Cas_13394_F Cas_14062_R (<i>atpA-atpF</i>)	F: CGAGCAATACCATCGCCTAC R: TTGGTTCGGGAGGGGATTAI	668	X	X	X	X	X	X	X	X	X	X
Cas_19198_F Cas_19976_R (<i>rpoC2</i>)	F: TCCCTGGAGTGGCCAAATAAG R: CCTTTCGTGAAATTAAGGCCAAA	778	X	X	X	X	X	X	X	X	X	X
Cas_124082_F Cas_124968_R (<i>ndhH</i>)	F: CFTCAATAATACGCCAAATTTCA R: ATGGACCCCGAACGACTAGG	886	X	X	X	X	X	X	X	X	X	X
Cas_52327_F Cas_52920_R (<i>ndhC-trnV</i>)	F: TCGATAAATACAGATACACCCCAATACA R: GCAAGAAATCCTAGGCCAAGA	593	X	X	X	X	X	X	X	X	X	X
Cas_50548_F Cas_51414_R (<i>rnf-ndhI</i>)	F: TCGGTTTCAGATACAAAATAATCCA R: AGGGTCAATTTGTCTGCTTGG	866	X	X	X	X	X	X	X	X	X	X
Cas_27800_F Cas_28312_R (<i>rpoB</i>)	F: CCGTACAGAACGAATACGC R: GATCCGGGGGATTAATTTG	712	X	X	X	X	X	X	X	X	X	X
Cas_20009_F Cas_20813_R (<i>rpoC2</i>)	F: TTGTCTTGGTCCCAATTCATATAC R: TCAATCATCCACTCCATCG	804	X	X	X	X	X	X	X	X	X	X
Cas_14705_F Cas_15624_R (<i>atpH</i>)	F: TCACACACACTCCCTTTTCCA R: GATATCGAAGTAGTTCCGGATTAAGTCA	919	X	X	X	X	X	X	X	X	X	X
Cas_25017_F Cas_25720_R (<i>rpoC1-rpoB</i>)	F: TTGGATTTGACTGGGANGACA R: TAATTAGACGGCCCAAGAG	703	X	X	X	X	X	X	X	X	X	X
Cas_40118_F Cas_40881_R (<i>psaB</i>)	F: TCCTGGATATATTTGGTGATGA R: CAATTGGTTTACGCCACTAATGAA	763	X	X	X	X	X	X	X	X	X	X
Cas_44810_F Cas_45699_R (<i>ycf3</i>)	F: ACACCGCTCCTCAAGACTTT R: CCAATCGAAGTTGTTGAAGTG	889	X	X	X	X	X	X	X	X	X	X
Cas_121734_F Cas_122486_R (<i>ndhA-ndhI</i>)	F: CTGGCAGCTCGTATTTGTTT R: TGAGGCGCGTATGAGGTAAA	752	X	X	X	X	X	X	X	X	X	X
Cas_93851_F Cas_94660_R (<i>ycf2</i>)	F: CGGAGCTGGAACTGCTAACT R: GTCGGGTAGAGACCAAGA	749	X	X	X	X	X	X	X	X	X	X
Cas_54595_F Cas_55457_R (<i>rnlM-atpB</i>)	F: GGGAGTCAFTGGTTCAAATCC R: GCCTTTCITATCAACAACCCCTTT	862	X	X	X	X	X	X	X	X	X	X
Cas_94709_F Cas_95300_R (<i>ycf2</i>)	F: CGGATCTAGTTCAATGGCTATT R: TCTGCAATAAATCTCGATGTGA	591	X	X	X	X	X	X	X	X	X	X
Cas_70574_F Cas_71412_R (<i>rps18-rpL20</i>)	F: GGNFCGAATTTGATTAAGAACATGA R: AGCTCGGAGCGTAGAACAAA	838	X	X	X	X	X	X	X	X	X	X
Cas_78484_F Cas_79253_R (<i>petB-petD</i>)	F: CCTTACTCGGGACCAATC R: CAATGCAAGGAATGAATGC	769	X	X	X	X	X	X	X	X	X	X
Cas_80388_F Cas_81242_R (<i>rpoA</i>)	F: TTTCTAGACTGCCCAATATCTGTTTT R: AAGCCGACACAATAGGCATT	854	X	X	X	X	X	X	X	X	X	X
Cas_81995_F Cas_82887_R (<i>rpl36-rps8</i>)	F: CCGCTACAGAACGAATACGC R: GTATCCGGGGATTAATTTG	892	X	X	X	X	X	X	X	X	X	X
Cas_85146_F Cas_85791_R (<i>rps3-rp12</i>)	F: TCGGAACGTATFAGGAACAATAATCA R: GGCAACTATGGTATATCTCTGATGTG	645	X	X	X	X	X	X	X	X	X	X

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</i> <i>americana</i> ^c	Clade III: <i>Orobancha</i> <i>californica</i> ^c	Clade IV: <i>Lamoureauxia</i> <i>virgata</i> ^c	Clade IV: <i>Pedicularis</i> sp. ^c	Clade IV: <i>Neobartisia</i> <i>filiformis</i> ^c	Clade V: <i>Rhinanthus</i> <i>alectorolophus</i> ^c	Clade VI: <i>Harveya</i> <i>purpurea</i> ^c	Clade VI: <i>Physocadyx</i> <i>major</i> ^c	Paulowniaceae: <i>Paulownia</i> <i>elongata</i> ^c (outgroup)
Cas_96241_F_	F: TCCGAGATCTCTATTGGAATTC	789	X	X	X	X	X	X	X	X	X	X
Cas_97030_R	R: TTCCATCGAATTGAGTATGATTTGT											
(<i>vcf/indlfb</i>)		769						X		X	X	
Cas_38180_F_	F: CGCCCAAGATCAAGATAAA											
Cas_38949_R	R: ACTGAAACATATATGCCAAGA	803	X	X	X	X	X	X	X	X	X	X
(<i>trng-rpsL4</i>)												
Cas_21932_F_	F: CGCGTCAGATATCCAGCAT	832	X	X	X	X	X	X	X	X	X	X
Cas_22735_R	R: TTCAGGCCCTTTCATATGGT											
(<i>rpoC2-rpoC1</i>)		802	X	X	X	X	X	X	X	X	X	X
Cas_12567_F_	F: AGCGGTCACTATTCCTTCA											
Cas_13399_R	R: TGGTCGTATTCACGGCTTG	939										
(<i>atpA</i>)												
Cas_25855_F_	F: GATATATGCCACTGGAACG	778	X	X	X	X	X	X	X	X	X	X
Cas_26657_R	R: AAGGCCCTGAAAGGATCACTA											
(<i>rpoB</i>)		902										
Cas_64793_F_	F: TAAGCCCTGGATATTTGAGG											
Cas_65732_R	R: AAATGGGTACAGGGATTTCC	801	X	X	X	X	X	X	X	X	X	X
(<i>petA-psbJ</i>)												
Cas_23417_F_	F: TTCCTGAAGTATTTCCATCAATC	792	X	X	X	X	X	X	X	X	X	X
Cas_24195_R	R: CGATACATTTCCAAATCGAG											
(<i>rpoC1</i>)		835	X	X	X	X	X	X	X	X	X	X
Cas_66623_F_	F: ACCTAATCCGGAATATGAACCA											
Cas_67525_R	R: TCTAAGATGCCCCCTCGGTACA	840	X	X	X	X	X	X	X	X	X	X
(<i>psbJ-petL</i>)												
Cas_90084_F_	F: AGAATCAGACCTATTCGCCAAA	923	X	X	X	X	X	X	X	X	X	X
Cas_90885_R	R: TGCCCTCAATATGTGTTC											
(<i>vcf2</i>)		757	X	X	X	X	X	X	X	X	X	X
Cas_18394_F_	F: TTGTTTGGTCTATCFGTGAAA											
Cas_19186_R	R: TGGCCATATGGAAATCC	839	X	X	X	X	X	X	X	X	X	X
(<i>rpoC2</i>)												
Cas_42897_R	F: GGCATAAGCGTGGGTATTT	834	X	X	X	X	X	X	X	X	X	X
(<i>psbA</i>)	R: TCAGGTGCATGTATCTTTACCG											
Cas_92095_F_	F: CMAFTCCAGATCCATTC	815	X	X	X	X	X	X	X	X	X	X
Cas_92935_R	R: TCGGTGCAGATGTAGATACC											
(<i>vcf2</i>)		839	X	X	X	X	X	X	X	X	X	X
Cas_87589_F_	F: TTGCTGCGTACTCTTCAG											
Cas_88512_R	R: ACGAATCGGTGTGTATATTTCA	839	X	X	X	X	X	X	X	X	X	X
(<i>rpl2-rpl23</i>)												
Cas_26951_F_	F: CTCATTCCTCGAGACAAGG	839	X	X	X	X	X	X	X	X	X	X
Cas_27708_R	R: CCGACTCCAGAAATTTGGT											
(<i>rpoB</i>)		839	X	X	X	X	X	X	X	X	X	X
Cas_91056_F_	F: AATGAAATACGATCAACCAACTT											
Cas_91895_R	R: TCATAATATTGATACGGGCCITTT	834	X	X	X	X	X	X	X	X	X	X
(<i>vcf2</i>)												
Cas_104111_F_	F: TTGGFTTGCACCTCTTCACA	815	X	X	X	X	X	X	X	X	X	X
Cas_104945_R	R: ATTTCCAGCTCTTCCCTTTCG											
(<i>rnl-rnl16</i>)		815	X	X	X	X	X	X	X	X	X	X
Cas_34914_F_	F: GAGCTTGTCTCATCTGTTC											
Cas_35729_R	R: ATTGCTCCAGCCAGATACC											
(<i>psbD-psbC</i>)												

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

^bAmplicon length (bp) estimated from *Castilleja* plastome 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 both were PCR validated and had primer combinations evaluated in silico against their respective plastome assemblies (raw read files available in the NCBI Sequence Read Archive submission SRP100222).

performance using additional sampling across Orobanchaceae. Orobanchaceae represents the largest parasitic clade of angiosperms and has well-documented modifications to the plastome, such as reduction and accelerated rates of molecular evolution; however, the most comprehensive phylogenetic investigation to date was based on only five gene regions (McNeal et al., 2013). Thus, an expanded molecular toolkit will be of great benefit for future investigations in the clade.

METHODS AND RESULTS

Three species of *Castilleja* were selected for genome skimming (*C. cusickii* Greenm., *C. foliolosa* Hook. & Arn., *C. tenuis* (A. Heller) T. I. Chuang & Heckard; Appendix 1), with taxa chosen to include both annual and perennial lineages (National Center for Biotechnology Information [NCBI] Sequence Read Archive [SRA] accession SRP100222). DNA extraction, purification, Illumina library construction, and subsequent cleaning of reads followed Uribe-Convers et al. (2016). Samples were sequenced as 100-bp single-end reads on an Illumina HiSeq 2000 (Illumina, San Diego, California, USA) at the University of Oregon, and cleaned reads were assembled against a reference genome (*Sesamum indicum* L. JN637766) using the Alignreads pipeline version 2.25 (Straub et al., 2011). In addition to these three low-coverage genomes, we also used existing data for 12 *Castilleja* plastomes generated by Uribe-Convers et al. (2014) using a long-PCR approach. Fifteen plastomes in total were aligned using MAFFT version 7.017b under the default settings (Katoh and Standley, 2013).

We used a custom R script (Uribe-Convers et al., 2016) to identify the most variable regions of the alignment spanning 400–1000 bp that were flanked by conserved regions, enabling prioritization based on predicted amplicon size and variability. Regions containing ambiguous bases were discarded, and those missing from one or more taxa in the alignment, particularly in the plastomes generated through the long-PCR method, were given lesser priority. We used Primer3 (Untergasser et al., 2012) to design primer pairs for the selected regions with an annealing temperature of 60°C (±1°C), and allowing no more than three continuous nucleotides of the same base, following the specifications of the Fluidigm Access Array System protocol.

We validated each primer combination using PCR with three high-quality *Castilleja* DNA isolations chosen to represent major lineages, sensu Tank and Olmstead (2008) (*C. lineariloba* (Benth.) T. I. Chuang & Heckard, *C. lemmonii* A. Gray, and *C. pumila* Wedd.; Appendix 1), but different than those selected for genome skimming and primer design, and a negative control. Because we followed the approach of Uribe-Convers et al. (2016), it was necessary for our validation conditions to simulate the four-primer reaction of the Fluidigm microfluidic PCR using a standard thermocycler. Therefore, our target-specific primers include a 5' conserved sequence (CS) tag, obtained from the Fluidigm Access Array System protocol, which provides an annealing site for Illumina sequencing adapters and sample-specific barcodes. PCR amplification followed Uribe-Convers et al. (2016), and amplicons were visualized on a standard agarose gel. In total, 76 primer combinations were successfully designed and validated (Table 1).

To test the broader utility of our *Castilleja*-specific primers, we searched for matches in two published plastome assemblies for *Lamourouxia virgata* Kunth (Pedicularideae, Clade IV; Fig. 1) and *Neobartsia stricta* (Kunth) Uribe-Convers & Tank (Rhinantheae, Clade V) (NCBI SRA accessions SRR1023133 and SRR1023130, respectively; Uribe-Convers et al., 2014). We assembled the plastome for a third taxon, *Physocalyx major* Mart. (Buchnereae, Clade VI; NCBI SRA accession SRP100222), to include in our comparison. *Physocalyx major* was sequenced on an Illumina HiSeq 2000 at the University of Oregon as 100-bp paired-end reads. Cleaned reads for *P. major* were mapped to three reference plastomes with one copy of the inverted repeat region removed (*Sesamum indicum* JN637766, *Neobartsia inaequalis* (Benth.) Uribe-Convers & Tank KF922718, *Castilleja paramensis* F. González & Pabón-Mora KT959111) using Bowtie2 (Langmead and Salzberg, 2012). Consensus sequences of the resultant contigs were obtained and used as final references. Contigs were then imported into Geneious R7 version 7.0.6 (Kearse et al., 2012), and a consensus sequence was obtained by calling regions with less than 5× coverage as “N” and using the “Highest Quality” as a threshold.

Separate BLAST databases were created for *Lamourouxia* Kunth, *Neobartsia* Uribe-Convers & Tank, and *Physocalyx* Pohl assemblies (-makeblastdb), and blastn_short was used to search for matching hits with the list of *Castilleja* chloroplast primers. Hits were further considered if both primer pairs (1) occurred on the same contig and (2) had predicted amplicon sizes between 350–1000 bp. Once we obtained a set of primer hits for the three taxa, they were validated with

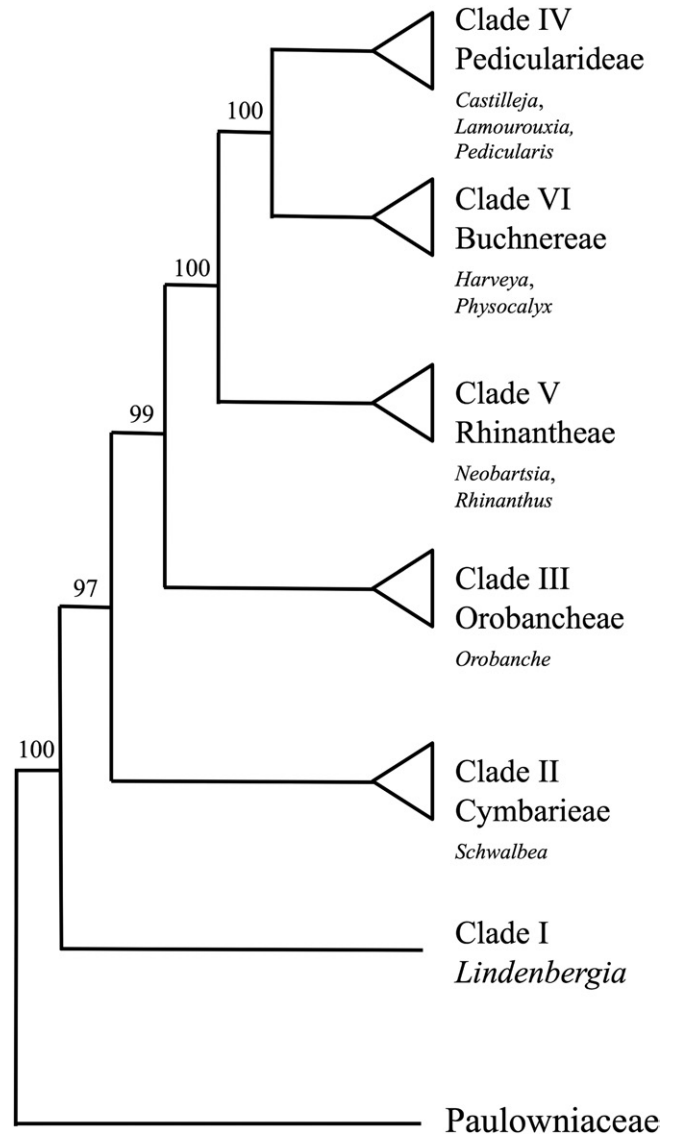


Fig. 1. Relationships among major clades within Orobanchaceae modified from McNeal et al. (2013), with taxa used for primer validation indicated (see text). Bootstrap support values for clades are indicated along the branches and follow McNeal et al. (2013).

PCR using *L. virgata*, *P. major*, and *Neobartsia filiformis* (Wedd.) Uribe-Convers & Tank (Appendix 1), as described above. Primer pairs with amplification in at least two out of three taxa above were chosen for another round of PCR validation with expanded taxon sampling that represented all major lineages of Orobanchaceae (sensu McNeal et al., 2013; Appendix 1): *Lindenbergia* sp. Lehm. (Clade I), *Schwalbea americana* L. (Cymbarieae, Clade II), *Orobanche californica* Cham. & Schldl. (Orobancheae, Clade III), *Pedicularis* sp. L. (Pedicularideae, Clade IV), *Rhinanthus alectorolophus* (Scop.) Pollich (Rhinantheae, Clade V), *Harveya purpurea* Harv. (Buchnereae, Clade VI), and *Paulownia* Siebold & Zucc. (Paulowniaceae; outgroup). As a positive control, we included CS-tagged “universal” primers for the *trnL-F* region (“trn-c” and “trn-f” of Taberlet et al., 1991, in Tank and Olmstead, 2008).

Out of the 76 primer pairs designed and validated for *Castilleja*, we identified 36 pairs with applicability across Orobanchaceae (referred to as core Orobanchaceae primers; these are boldfaced in Table 1). These were chosen based on amplification across a large phylogenetic breadth of the clade, but allowing for some failures. For example, *Orobanche*, a holoparasite, failed for most primer combinations, a result that is likely due to the reduction and modification of the plastome in this lineage (see Bennett and Mathews, 2006). Higher success rates were noted for hemiparasites.

CONCLUSIONS

We report 76 primer pairs designed to target the most variable regions of the chloroplast genome in *Castilleja*. We further demonstrate their utility across other major clades in Orobanchaceae, particularly with hemiparasitic taxa, and present a subset of 38 core Orobanchaceae primers. Although these primer combinations target similar highly variable plastid regions as in other angiosperm-wide studies (e.g., Ebert and Peakall, 2009), few of the primers reported here overlap directly with them. Two exceptions are Cas_11589 F (*trnG*) and Cas_61880 F (*psaI*) (Table 1), which were also developed by Ebert and Peakall (2009). Notably, our primer combinations were designed with the same annealing temperature to take advantage of the Fluidigm microfluidic PCR system and high-throughput sequencing platforms, but will also be useful for traditional PCR and Sanger sequencing.

LITERATURE CITED

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.

EBERT, D., AND R. PEAKALL. 2009. A new set of universal *de novo* sequencing primers for extensive coverage of noncoding chloroplast DNA: New opportunities for phylogenetic studies and cpSSR discovery. *Molecular Ecology Resources* 9: 777–783.

FANT, J. B., H. WOLF-WEINBERG, D. C. TANK, AND K. A. SKOGEN. 2013. Characterization of microsatellite loci in *Castilleja sessiliflora* and transferability to 24 *Castilleja* species (Orobanchaceae). *Applications in Plant Sciences* 1: 1200564.

HECKARD, L. H., AND T.-I. CHUANG. 1977. Chromosome numbers, polyploidy, and hybridization in *Castilleja* (Scrophulariaceae) of the Great Basin and Rocky Mountains. *Brittonia* 29: 159–172.

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.

LANGMEAD, B., AND S. SALZBERG. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9: 357–359.

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.

STRAUB, S. C. K., M. FISHBEIN, T. LIVSHULTZ, Z. FOSTER, M. PARKS, K. WEITEMIER, R. C. CRONN, AND A. LISTON. 2011. Building a model: Developing genomic resources for common milkweed (*Asclepias syriaca*) with low-coverage genome sequencing. *BMC Genomics* 12: 211.

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.

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., 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.

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>Lamourouxia 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 filiformis</i> (Wedd.) Uribe-Convers & Tank	Uribe-Convers 13-027 (ID)	La Paz, Bolivia	16.32796°S, 67.9457°W
<i>Orobanche californica</i> Cham. & Schltld.	Bennett 72 (A)	Cultivated	Cultivated
<i>Paulownia elongata</i> Siebold & Zucc.	s.n. (A)	Cultivated	Cultivated (https://sheffields.com)
<i>Pedicularis</i> sp. L.	Krajsek and Bennett s.n. (A)	NA	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/>).