PLASTID PRIMERS FOR ANGIOSPERM PHYLOGENETICS AND PHYLOGEOGRAPHY

LINDA M. PRINCE

Department of Botany, The Field Museum, 1400 S. Lake Shore Drive, Chicago, Illinois 60605 USA

Application Article

1 Manuscript received 8 September 2014; revision accepted 24 March 2015.

Whole genome sequencing is more available and less expensive than ever before, yet most scientists continue to rely on targeted, comparative sequencing for phylogenetics and phylogeography. Identifying the most appropriate markers to employ has been challenging. Information for model organisms abounds (e.g., grasses; Saski et al., 2007; Bortiri et al., 2008; Leseberg and Duvall, 2009), and a few studies have specifically screened the same set of markers across a diversity of plant groups, ranking the utility of these markers either explicitly or implicitly (Shaw et al., 2005, 2007, 2014). These studies are exceedingly valuable, demonstrating there is no one-size-fits-all answer to the question “which markers?” The second critical question to “which markers” is “which primers?” Hundreds of primer sequences have been published, many designed for specific taxonomic groups. “Universal” primers may have extremely limited utility. There was little consistency in likelihood of amplification success for any given publication across lineages or within lineage across publications.

Key words: comparative sequencing; complete chloroplast genome; cpDNA.

Discussion:
The Tortoise and the Hare II paper was revolutionary in assessing sequence variability for all regions studied across a broad diversity of flowering plants, and providing a ranking of that variability. In the mid-2000s, a small number of complete chloroplast genome sequences were available for land plants and some of those were not annotated (e.g., Medicago truncatula Gaertn. [GenBank NC_003119]; Saski et al., 2005). Grivet et al. (2001) were visionary when they moved beyond analyzing regions commonly being used to design primers for less-known and potentially faster-evolving regions of the chloroplast genome. They were the first to take advantage of the new genomic data boom, providing a set of 20 universal chloroplast primers designed around the complete chloroplast data from seven flowering plant species. Around the same time, I developed nondegenerate primers for 36 noncoding regions in the large and small single-copy regions of the chloroplast genome (published here). These near-universal primers were designed based on the complete chloroplast genome sequences of 16 flowering plant species (see Appendix 1).

Grivet et al. (2001) and I designed primers, but Shaw et al. (2007) took an even more applied approach when they examined sequences for three different taxon pairs (Atropa/Nicotiana, Lotus/Medicago, and Saccharum/Oryza), specifically searching for faster-evolving regions. Shaw et al. (2014) go one step further, comparing complete chloroplast genome sequences for 25 (primarily congeneric) sister species pairs. They examined sequence diversity for 107 single-copy noncoding regions, providing the most comprehensive analysis to date.

There are now at least 150 primer pairs available to amplify almost every intergenic, intron, and exon region of the chloroplast genome, including portions of the inverted repeats, thanks to the efforts of Shaw et al. (2005, 2007, 2014) and others (Ebert and Peakall, 2009; Scarcelli et al., 2011; Dong et al., 2012, 2013, 2015; Prince, 2015).
flowering plants, but does vary and can be highly informative, for example, as in the 22-kb inversion in almost all Asteraceae (Jansen and Palmer, 1987a, 1987b) and the 78-kb inversion in Fabaceae subtribe Phaseolinae (Bruneau et al., 1990). Some primer combinations are not useful in particular groups of plants due to structural rearrangements. In some cases, the downloaded genomes differ in the identification of specific genes.

**Primer utility**—The chloroplast genomes for species of eight genera (Acors L., Amborella Baill., Canna L., Ceratophyllum L., Cymbidium Sw., Helianthus L., Magnolia L., and Nelumbo Adans.) and for subspecies of *F. vesca*, *G. herbaceaem*, *O. europaea*, and *O. sativa* were compared to 130 primer pairs published by Shaw et al. (2005, 2007), Scarcelli et al. (2011), Dong et al. (2012), and those designed here. Complete chloroplast genome sequences were downloaded from GenBank (accession numbers, taxonomic identity, and original publication information provided in Appendix 3) and aligned manually in Sequencer (Gene Codes Corporation, Ann Arbor, Michigan, USA). A separate file containing the primer sequences was imported and automatically assembled using the settings “dirty data” and 100% sequence similarity with a minimum overlap of 16 bp. An alignment included with successively lower levels of sequence similarity. Primers that failed to align automatically, or that aligned incorrectly, were realigned manually whenever possible (guided by the GenBank annotations). Alignment of the two *Gossypium* sequences required inversion of a large region of one taxon (arbitrarily selected as *G. herbaceaem* subsp. *africanum* (G. Watt) Vollesen) approximately corresponding to bases 115,132–135,355 in the final alignment. The *Oryza* alignment includes O. *nivara* Sharma & Shastry because it is a potential progenitor of *O. sativa* (Li et al., 2006; but see Huang et al., 2012 for an alternative view point).

As mentioned above, degenerate primers provide broader utility, but reduced amplification efficiency. If a mismatch was detected in the last five bases at the 3′ end of the primer, the mismatch was inferred to be fatal (IBT, 2009). If more than three mismatches were detected within any given primer, amplification was inferred to be unsuccessful. These criteria are arbitrary but have worked for me personally and are probably more strict than necessary.

**Sequence variability within species**—The sequences of *F. vesca*, *G. herbaceaem*, *O. europaea*, and *O. sativa* were examined manually to assess the variation of the 130 regions. Length of the inferred ampiclon was noted along with the number of mismatched bases (aka inferred substitutions; excluding primer regions), the number of insertions/deletions (indel) events, and the number of inversions. These data provided an estimate of the utility of the regions for inferring phylogeny among closely related subspecies, and potential for application to phylogeographic studies. Shaw et al. (2014) specifically avoided these types of comparisons due to the very small number of parsimony informative characters. Sequence diversity was estimated using three criteria calculated as: (1) {number of substitutions*2+number of indels+number of inversions}/amplicon length, (2) number of substitutions+indels+inversions, and (3) sequence diversity (number of substitutions/sequence length). The first criterion (criterion 1) is a weighted rank, and includes information on the number of inferred substitutions (weighted twice as heavily as the other two components), indels, and inversions. Substitutions were weighted more heavily because chloroplast indels may be more homoplous (Kelchner and Clark, 1997), especially among closely related taxa. Inversions are often low in homoplasy (Graham et al., 2000) and thus could be weighted more heavily, but are relatively rare so weighting was not employed. The 10 most variable regions for each species were identified, as measured under each criterion. Frequency of any specific “top 10” primer pair was summed across the four species.

RESULTS

**Primers designed here**—The 72 primers targeted noncoding regions of the chloroplast genome with amplicon sizes of 500–1800 bp. Degenerate primers were avoided because they were assumed to decrease priming efficiency, as were mismatches within the last five bases at the 3′ end of the primer. Only two primers required degenerate bases: one primer with two degenerate bases and another primer with one degenerate base. None of these degeneracies were located within the last five bases. In
### Table 1. Region, primer name, primer sequence, amplicon position, and amplicon length for plastid noncoding regions relative to the *Nicotiana tabacum* L. (GenBank Z000441.1) genome.

| Region                        | Primer name                     | Tm (°C) | Primer sequence                  | Amplicon position | Amplicon length (bp) |
|-------------------------------|---------------------------------|---------|----------------------------------|-------------------|----------------------|
| trnG(UUG)–psbK IGS           | trnG-IGSR                       | 62.7    | ACCCGTGGCTCTACGGCTTGG            | 7457–8018         | 562                  |
| psbK–trnS(GCU) IGS           | psbK-IGSR                       | 50.9    | ATCGAAACTGTGAGCGATGTTGG          |                   |                      |
| trnS(GUC)–atpA IGS           | trnS_GUC-IGSR                   | 47.9    | CCAATGCTAGTGGTAGTTGAGG           | 7937–8719         | 783                  |
| atpF intron                   | atpF-E2R                        | 47.3    | CTTGGATTGCTTAATATTTAAAT          | 12,582–13,372     | 794                  |
| atpF–atpH IGS                | atpF-E1R                        | 48.1    | ACGAAACTGTGAGCGATGTTGG           |                   |                      |
| atpH–atpI IGS                | atpH-IGSR                       | 48.5    | GGTGTTTATGTTGCTTGG              | 13,352–13,927     | 757                  |
| rpoC1 intron                  | rpoC1-intR                      | 49.9    | AAGTGGAACTTGCTTATG              | 23,004–23,976     | 973                  |
| trnS(UAG)–psbZ IGS           | trnS_UAG-IGSR                   | 49.2    | ATCAAGAATCAGCGCATC              | 37,209–37,620     | 412                  |
| psaA–ycf3 IGS                | psaA-IGSR                       | 50.2    | CCGGAACAGATAAACAT              | 43,460–44,295     | 827                  |
| ycf3 intron 2                 | ycf3-E3F                        | 48.4    | CCGGGAATATATTGAGG               |                   |                      |
| ycf3 intron 1                 | ycf3-E1R                        | 50.0    | TCCGGAATATATTGAGG               | 45,370–46,163     | 794                  |
| ycf3–trnS(GGA) IGS           | ycf3-trnS(GGA)-IGSR             | 45.5    | ACAATTGAGAAGGTTATTTG             | 46,214–47,147     | 961                  |
| rps4–trnT(UGU)               | rps4-trnT(UGU)                  | 56.2    | TCCGGAATATATTGAGG               | 48,065–48,570     | 506                  |
| tmF(GAA)–ndhJ IGS            | tmF-GAA-IGSR                    | 56.4    | ATCCTGGTGACACGGATTTGA            | 50,277–51,024     | 747                  |
| ndhC–trnV(UAC) IGS           | ndhC-UAC-IGSR                   | 52.9    | ATCATATTGGAAGCAGACAAAT          | 52,644–53,776     | 1132                 |
| trnV(UAC) intron             | trnV-UAC-E2R                    | 58.3    | GGTCAATTGATGGTAGGG              | 53,757–54,380     | 624                  |
| trnV(UAC)–atpE IGS           | trnV-UAC-IGSR                   | 52.8    | GGTGTTTATGTTGCTTGG              | 54,361–55,032     | 672                  |
| atpB–rbcL IGS                | atpB-IGSR                       | 48.4    | AAGTGGAACTTGCTTATG              | 56,756–57,615     | 859                  |
| rbcL–accD IGS                | rbcL-IGSR                       | 53.9    | ATCTGGTGATTTGTTGG               |                   |                      |
| accD–psaI IGS                | accD-IGSR                       | 58.5    | GGTCAATTGATGGTAGGG              | 58,960–59,865     | 905                  |
| psal–ycf4 IGS                | psal-IGSR                       | 51.1    | ATCTGGTGATTTGTTGG               | 61,143–62,161     | 1018                 |
| ycf4–ycf10 IGS               | ycf4-IGSR                       | 51.8    | ATCTGGTGATTTGTTGG               | 62,127–62,682     | 556                  |
| petA–psbJ IGS                | petA-IGSR                       | 51.9    | ATCTGGTGATTTGTTGG               | 63,153–63,541     | 389                  |
| petL–psaJ IGS                | petL-IGSR                       | 58.8    | ATCTGGTGATTTGTTGG               | 65,255–66,388     | 1133                 |
| psaA–rpsL20 IGS              | psaA–rpsL20-IGSR                | 52.4    | ATCTGGTGATTTGTTGG               | 68,322–69,671     | 1350                 |
| rpsL20–rps12 IGS             | rpsL20-rps12-IGSR               | 46.8    | ATCTGGTGATTTGTTGG               | 69,565–71,404     | 1840                 |
| clpP intron 2                 | clpP-E2R                        | 51.6    | ATCTGGTGATTTGTTGG               | 71,372–72,319     | 947                  |
| clpP intron 1                 | clpP-E1R                        | 51.2    | ATCTGGTGATTTGTTGG               | 72,573–73,466     | 893                  |
| clpP–psbB IGS                | clpP-IGSR                       | 54.9    | ATCTGGTGATTTGTTGG               | 73,446–74,451     | 1005                 |
| psbB–petD IGS                | psbB-IGSR                       | 52.2    | ATCTGGTGATTTGTTGG               | 74,481–74,970     | 490                  |
| petB–petD IGS                | petB-IGSR                       | 51.5    | ATCTGGTGATTTGTTGG               | 77,214–78,377     | 1163                 |
| rps3–rps19 IGS               | rps3-IGSR                       | 50.8    | ATCTGGTGATTTGTTGG               | 78,805–79,760     | 956                  |
| ccsA–ndhD IGS                | ccsA-IGSR                       | 45.9    | ATCTGGTGATTTGTTGG               | 85,863–86,504     | 642                  |
| ndhD–IGSR                    | ndhD-IGSR                       | 43.6    | ATCTGGTGATTTGTTGG               | 116,344–117,614   | 1271                 |
contrast, 17 of the Scarcelli et al. (2011) primers have at least one degenerate base in the last five bases at the 3' end of the primer, and so are assumed to fail for at least some taxa.

**Primer evaluation**—Three of the four sets of primers examined here were equally likely to amplify target chloroplast regions (81–85% should work; see Table 2). The Dong et al. (2012) primers were least likely to work based on the 12 species examined here (65% on average) and were particularly poorly matched to the *Oryza* genome (29% amplification success predicted), and only moderately suited for *Amborella* (52%), *Cymbidium* (52%), and *Helianthus* (57%). However, the Dong et al. (2012) primer pair *trnH-psbA* was not expected to work on any of the target species, possibly due, in part, to an extra “A” near the 3’ end of the published sequence for the *trnH* primer. The primers designed here were poorly matched to three of the four monocots (*Cymbidium, Oryza, and Canna*; 61%, 64%, and 67%, respectively), despite being a good match for *Acorus* (81%). Scarcelli et al. (2011) primers were designed with monocots in mind and did an exceptional job matching the monocot genomes examined here, with amplification success ranging from 82–97%. They were almost equally good for the monocot genomes examined here, with amplification success of 72–93%.

On average, the Shaw et al. (2005, 2007) and Scarcelli et al. (2011) primers are more degenerate, yet they were only slightly more likely to amplify the target sequences than the nondegenerate primers designed here, at least for nonmonocot taxa. With so many different primers available, most regions could be amplified in almost all target taxa provided an appropriate primer pair was selected. Indeed, many primer pairs should work in all 12 species examined here. Details of the inferred priming success are provided in Appendix S1, and species-specific notes on primer/sequence mismatches are provided in Appendix S2.

**Primer utility × sequence variability**—Shaw et al. (2014) conveniently summarized sequence variability across the chloroplast genome including the identification of the 13 fastest-evolving regions for six taxonomic groups (magnoliids, monocots, eudicots I, eudicots II, euasterids I, and euasterids II). Summing across these major groups, 28 different regions were identified as the most variable. Primers to amplify those 28 regions are detailed in Table 3, along with the Shaw et al. (2014) rank for each region (in bold typeface above each primer region), for each taxon examined here. Multiple primer pairs are available for each of the 28 regions except the *trnF-trnl* (Shaw et al., 2005 only), *ycf4-ycf10* (or *cena*; current study only), and *ndhD-psaC* (none of the publications examined). The *ndhD-psaC* region was ranked 10th fastest for eudicots I, but as there are no primers to be evaluated this region will not be discussed further.

Among the basal dicot grade (*Amborella* and *Magnolia*), successful primers are available for all 27 regions. Primer selection is more challenging for *Amborella* than for *Magnolia*. The top ranked region was the *rpl32-trnL* intergenic spacer (IGS). Shaw et al. (2007) primers will work for both taxa; Dong et al. (2012) primers will not. In contrast, *rps16-trnQ*, the second highest ranked region, has three sets of primers available (Shaw et al., 2007; Scarcelli et al., 2011; and Dong et al., 2012), all of which should work.

Among the monocots sampled (*Acorus, Cymbidium, Oryza, and Canna*), *Acorus* was the least difficult sequence to match and *Oryza* the most difficult. Structural rearrangements are the primary reason for failure to amplify across all available primers (e.g., *rbcL-accD* in *Oryza* and *petA-psbJ in Cymbidium*). One region cannot be amplified in *Acorus*—the *accD-psaD* IGS, despite the availability of four different primer pairs. In all, four regions cannot be amplified in *Cymbidium* with the primers studied here: *petN-psbM, psbM-trnD, atpB-rbcL, and petA-psbJ*. The *ndhA* region can be amplified in only some species of *Cymbidium* due to fatal substitutions in some species for all three primer pairs evaluated here. In *Oryza*, the *trnS[GU]-trnG[GCC], trnF-trnD, rbcL-accD, accD-psaD*, and *rps15-ycf1* cannot be amplified using any primer pair. In *Canna*, *ndhF-rpl32* will not amplify with either of the available primer pairs.

Unfortunately, according to Shaw et al. (2014), *ndhF-rpl32* is the most variable and *psbM-trnD* is the third most variable region for monocots.

Basal eudicots were not evaluated by Shaw et al. (2014) in detail, so direct comparisons cannot be made here. Fortunately, at least one primer pair was successful for each of the 27 fastest-evolving regions, with the exception of the *ycf4-ycf10* region. The only available primers for this region were designed here, and they will not work for *Ceratophyllum*. In general, *Ceratophyllum* was more difficult to match than was *Nelumbo*.

Shaw et al. (2014) detailed variability of higher eudicots for four major groups: eudicots I, eudicots II, euasterids I, and euasterids II. Only a single species representing each group was included here: *Fragaria* (eudicots I) could not be amplified for a single region, the *ycf4-ycf10* IGS. According to Shaw et al. (2014), the fastest region for this clade was the *ndhA* intron. Both the Shaw et al. (2007) and Scarcelli et al. (2011) primers should work, but the Dong et al. (2012) primers will not. The second fastest region was the *trnS[GU]-trnG[GCC], which should amplify with any of the primer pairs (Shaw et al., 2005; Scarcelli et al., 2011; or Dong et al., 2012).

The sole representative of eudicots II and euasterids I (*Gossypium* and *Olea*, respectively) could successfully be amplified by at least one pair of primers studied here. The fastest region

**TABLE 1. Continued.**

| Region | Primer name | $T_m$ (°C)* | Primer sequence | Amplicon position | Amplicon length (bp) |
|--------|-------------|-------------|----------------|-------------------|---------------------|
| psaC–ndhE IGS | psaC-IGSR | 44.9 | TCTATACACGGTACATCATAAA | 119,351–119,713 | 363 |
| ndhE–ndhI IGS | ndhE-IGSF | 42.4 | TTCATCAATTATGTGAAAC | 119,912–121,251 | 1340 |
| rps15–ycf1 IGS | rps15-IGSR | 45.6 | GAAATAATAGGCGACTCAA | 125,374–126,001 | 628 |
| ycf1-IGSR | 45.6 | ATTATCGATTAAGAGATTACG | 125,374–126,001 | 628 |

*Melting temperature ($T_m$) based on 50 mM NaCl solution.
for eurosids II was the ndhF-rpl32 IGS. The Shaw et al. (2007) primer pair should work, but the Scarcelli et al. (2011) primer pair likely will not. The second most variable region was the psbZ-trnG IGS. For this region, both the Scarcelli et al. (2011) and Dong et al. (2012) primers should work, but the Shaw et al. (2005; as trnM-trnS) primers will not. In euasterids I, the fastest region was the rps16-trnQ IGS. For *Olea*, the Shaw et al. (2007) and Scarcelli et al. (2011) primers should work, but not so the Dong et al. (2012) primers. The next-fastest region was the rpl32-trnL IGS. Both the Shaw et al. (2007) and Dong et al. (2012) primers should work.

Primer failure in *Helianthus* (euasterids II) was primarily due to structural rearrangements (e.g., *trnS*[GCU]-*trnG*[GCC], *rpsB-trnc*, *trnE-trnT*, *rbcL-accD*). rpl32-trnL IGS was the fastest region according to Shaw et al. (2014), and either the Shaw et al. (2007) or Dong et al. (2012) primers should successfully amplify this region. The adjacent ndhF-rpl32 IGS was the second most variable region. Both the Shaw et al. (2007) or the Scarcelli et al. (2011) primers should work.

Subspecific sequence variability—Intraspecific sequence variation was evaluated in four species: *F. vesca, G. herbaceum, O. europaea*, and *O. sativa*. This represents a tiny fraction of angiosperm diversity, but is the first analysis of subspecific diversity across the entire chloroplast genome for multiple species, in the context of available primer resources. Appendix S3 identifies the fastest-evolving regions among the four species, under three different criteria. On average, only five inversions per chloroplast genome were detected here and the distribution across species was very different. *Gossypium* and *Oryza* each had 10 inversions, *Fragaria* none, and *Olea* only one. Details of subspecific comparisons for all regions are provided in Appendix S2.

No single genic region was identified as the top 10 fastest for all four species. Pooling data across all three criteria, the most frequently identified genic region was the psbZ-trnM IGS with eight occurrences out of a maximum of 12 possible, followed by the *trnS*[GCU]-*trnG*[GCC] IGS, with six occurrences, *rps16-trnQ* IGS and *trnT*[GGU]-*psbB* IGS each with five, and *rps12-psbB* IGS and *rps4-trnT*[UGU] IGS each with four occurrences. Data for individual species have limited general application, but are provided below.

*Oryza sativa*, the only monocot in this comparison, showed highest variation, based on rank, for *clpP-psbB* (0.0195, 924 bp), *atpB-rbcL* (0.0168, 1070 bp), and *psbM-trnD*[GUC] (0.0150, 523 bp). Two of the same regions were identified as fastest under criterion 2, *atpB-rbcL* (12 characters, 1070 bp) and *clpP-psbB* (11 characters, 924 bp), plus *rbcL-accD* (13 characters, 1824 bp). Sequence divergence was highest in and around the *clpP* region including what would be the *clpP* intron 2 (1.9455%, 257 bp), *clpP* intron 1 (1.0050%, 199 bp), and *clpP-psbB* (0.7576%, 924 bp). In contrast, the three fastest regions per Shaw et al. (2014) for monocots were *ndhF-rpl32* (rank 1), *ndhC-trnV* (rank 2), and *psbM-trnD* (rank 3).

The highest variation for *Fragaria* under criterion 1 was for *trnW*[CCA]-*psaJ* (0.0101, 789 bp), *trnT*[GGU]-*psbD* (0.0098, 1527 bp), and *trnP*[UGG]-*rpl18* (0.0090, 1563 bp). Under criterion 2: *trnT*[GGU]-*psbD* (eight characters; 1527 bp), *trnP*[UGG]-*rpl18* (eight characters, 1563 bp), and *petN-trnD* (seven characters, 2504 bp). Under criterion 3, the top three regions were *trnT*[GGU]-*psbD* (0.4584%, 1527 bp), *psbB-psbH* (0.4451%, 674 bp), and *rps4-trnT*[UGU] (0.4435%, 451 bp). Shaw et al. (2014) eurosids I top three regions were *ndhA* intron 924 bp). In contrast, the three fastest regions per Shaw et al. (2014) for monocots were *ndhF-rpl32* (rank 1), *ndhC-trnV* (rank 2), and *psbM-trnD* (rank 3).

The highest variation for *Fragaria* under criterion 1 was for *trnW*[CCA]-*psaJ* (0.0101, 789 bp), *trnT*[GGU]-*psbD* (0.0098, 1527 bp), and *trnP*[UGG]-*rpl18* (0.0090, 1563 bp). Under criterion 2: *trnT*[GGU]-*psbD* (eight characters; 1527 bp), *trnP*[UGG]-*rpl18* (eight characters, 1563 bp), and *petN-trnD* (seven characters, 2504 bp). Under criterion 3, the top three regions were *trnT*[GGU]-*psbD* (0.4584%, 1527 bp), *psbB-psbH* (0.4451%, 674 bp), and *rps4-trnT*[UGU] (0.4435%, 451 bp). Shaw et al. (2014) eurosids I top three regions were *ndhA* intron
| Approx. Nicotiana order | Genomic region | Publication | Basal dicot grade/Magnoliids | Monocots | Basal eudicot grade | Eurosids I | Eurosids II | Euasterids I | Euasterids II | Average |
|------------------------|----------------|-------------|-----------------------------|----------|---------------------|------------|------------|-------------|-------------|---------|
|                        |                |             | Amborella | Magnolia | Acorus | Cymbidium | Oryza | Cannia | Centrophilyum | Nelumbo | Fragaria | Gossypium | Olea | Helianthus |
| 1                      | trnH-psbA IGS  | Dong et al. | NO**     | NO**     | NO     | NO**     | NO   | NO     | NO          | NO**     | 8°      |
| 2                      | trnH-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | NO | YES | YES | YES | YES | YES | YES | 92% |
| 3                      | trnH-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 92% |
| 4                      | matK exon (including matK) | Dong et al. | YES | YES | YES | NO | YES | YES | YES | YES | YES | YES | YES | YES | 92% |
| 5                      | trnK IGS       | Scarcelli et al. | YES | YES | YES | YES* | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 6                      | trnK IGS       | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 7                      | trnK IGS       | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 8                      | trnK IGS       | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 9                      | trnK IGS       | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 10                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 11                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 12                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 13                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 14                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 15                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 16                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 17                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 18                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 19                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 20                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 21                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 22                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 23                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 24                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 25                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 26                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 27                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 28                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 29                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 30                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 31                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 32                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 33                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 34                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 35                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 36                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 37                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 38                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 39                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 40                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 41                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 42                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 43                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 44                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 45                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 46                     | trnL-psbA IGS  | Dong et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 47                     | trnL-psbA IGS  | Scarcelli et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
| 48                     | trnL-psbA IGS  | Shaw et al. | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | YES | 100% |
## Table 3. Continued.

| Approx. Nicotiana order | Genomic region | Publication\(^b\) | Basal dicot grade/Magnoliids | Monocots | Basal eudicot grade | Eurosids I | Eurosids II | Euasterids I | Euasterids II | Average |
|-------------------------|----------------|-------------------|-------------------------------|----------|---------------------|------------|------------|-------------|-------------|---------|
|                         |                |                   | Amborella | Magnolia | Acorus | Cymbidium | Oryza | Canna | Centrophylum | Nelumbo | Fragaria | Gossypium | Olea | Helianthus |
|                         |                |                   | YES      | YES     | YES     | YES     | YES     | YES     | YES     | YES     | 100%    |
|                         |                |                   | 50       | 55      | 70     | 71      | 72      | 73      | 74      | 75      | 100%    |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | YES     | YES     | 100%    |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | YES     | YES     | 100%    |
|                         |                |Dong et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 100%    |
|                         |                |Prince (here)      | YES      | YES     | YES     | YES     | YES     | YES     | YES     | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | YES     | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Shaw et al.        | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |
|                         |                |Scarcelli et al.   | YES      | YES     | YES     | YES     | YES     | YES     | NO      | YES     | 92%     |

\(a\) YES* = will not work for at least one species in the genus; NO** = will work if psbA primer is synthesized with one fewer A at the 3' end.

\(b\) Shaw et al., 2005, 2007; Scarcelli et al., 2011; Dong et al., 2012.

\(c\) Shaw et al. (2014) rank for the region within the specified taxonomic group.
trnS(GCU)-trnG(GCC) (rank 2), and rps16 intron (rank 3).

In *Gossypium*, the most informative regions under criterion 1 were *psbZ-trnfM* (CAU) (0.0534, 1179 bp), *trnH*(GUG)-*psbA* (0.0444, 496 bp), and *rps4-trnF* (UGU) (0.0425, 635 bp). Criterion 2 fastest regions were *trnS*(UGA)-*trnfM* (GCC) with 39 variable characters over 1673 bp, followed by *psbZ-trnfM* (CAU) with 37 characters for 1179 bp, and *trnT*(UGU)-*trnL* (UAA) with 33 characters over 1470 bp. Sequence divergence (criterion 3) was highest for *psbZ-trnfM* (CAU) (2.0253%, 1179 bp), then *trnS*(UGA)-*trnfM* (GCC) (1.6736%, 1673 bp), and finally the *rps16* intron (1.6181%, 927 bp). Eurosid II top three regions for Shaw et al. (2014) were *ndhF-rpl32* (rank 1), *psbZ-trnG* (rank 2), and *trnT-trnL* (rank 3).

For *Olea*, the most informative regions under criterion 1 were *psbC-psbZ* (0.0411, 1045 bp), *trnS*(UGA)-*trnfM* (0.0333, 1203 bp), and *clpP* intron 2 (0.0313, 702 bp). The highest number of variable characters (criterion 2) were found in *rps16-trnQ* (29 characters, 2739 bp), *psbC-psbZ* (22 characters, 1045 bp), and *trnS*(UGA)-*trnfM* (21 characters, 1203 bp). Criterion 3 (percent sequence divergence) was highest in the same three regions as under criterion 1: *psbC-psbZ* (2.0096%, 1045 bp), *trnS*(UGA)-*trnfM* (1.5794%, 1203 bp), and *clpP* intron 2 (1.4245%, 702 bp). Shaw et al. (2014) eurosterids I top three included *rps16-trnQ* (rank 1), *rpl32-trnL* (rank 2), and *ndhC-trnV* (rank 3).

DISCUSSION

A large number of “universal” primers have been published for amplification of various chloroplast regions. Some are more degenerate than others, presumably to be more widely applicable. Degeneracy is not required, however, and may not lead to greater success in the laboratory. On the other hand, nondegenerate primers with poor fit are likely to fail, and some primers published as “universal” are not necessarily so. The universal barcoding primers of Dong et al. (2012) were the least likely to be useful across the 12 taxa examined here, with an average success rate of 65%, and a very poor 29% success rate in *Oryza*. In contrast, the primers designed by Scarcelli et al. (2011) specifically for monocots were exceedingly well-matched to the monocots sampled (97% in *Acorus*, 93% in *Cymbidium*, 92% in *Oryza*, and 88% in *Canna*), and a good match across all angiosperms.

Unlike previous analyses, this study used published genomes and primer sequences to infer the likelihood of amplification success. Only a small number of published primers were evaluated, and additional primers will be added to future analyses. Indeed, as mentioned in the introduction, Ebert and Peakall (2009) and Dong et al. (2013) have primers that could be evaluated as well as those of Doorduin et al. (2011) designed for species of Asteraeae. The evaluation conducted here shows parallels to prior studies in that general conclusions or recommendations are difficult to distill. For each region, there may be a number of primer pair options. Which primer pair is best is highly variable and depends upon the taxon being investigated. Scarcelli et al. (2011) primers are the best option for monocots in general, but will fail in specific combinations (e.g., *trnH*-psbA for *Canna*, atpF intron/exon for *Cymbidium*, and *trnD-trnT* for *Oryza*). Dong et al. (2012) primers are generally less successful, but they are the only primers that will work for *psbM-trnD* in Amborella and Magnolia. In several instances, a primer will work for some, but not all species in a genus, like the Scarcelli et al. (2011) *matK* primers in *Cymbidium* or the *trnK-rps16* primers in *Helianthus*. Table 3 provides a quick summary of primer match for the top regions according to Shaw et al. (2014).

Prior studies have done an excellent job assessing variability of various noncoding regions across a diversity of angiosperms, particularly the recent work of Shaw et al. (2014). Those studies focused on infrageneric or even intergeneric comparisons. Here I compare sequence variability within species to see if the same markers are identified as the most variable, under slightly different criteria. This comparison was specifically avoided by Shaw et al. (2014) due to the small number of variable characters. The fastest regions identified here for *Oryza* were (depending upon criterion) *clpP-psbB*, atpB-rbcL, *psbM-trnD*, and *rbcL-accD*. In contrast, Shaw identified *ndhF-rpl32*, *ndhC-trnV*, and *psbM-trnD* as the fastest regions for monocots, with only one region of overlap between the two. For *Fragaria* (eurosids I), the list has no overlap at all. *Olea* (eurosids II) and *Gossypium* (eurosids I) each only overlap for a single region between the two studies. The lack of consensus over which region is the most variable at lower taxonomic levels has been pointed out by a number of papers including Särkinen and George (2013) for *Solanium*, and for 19 species pairs as demonstrated by Shaw et al. (2014). The comparison made here only adds to the argument that there is an acute need for additional comparative information.

Shaw et al. (2014) provided a solid foundation for which markers evolve the most quickly in major angiosperm clades, yet the fastest regions identified here for subspecies comparisons share little overlap with Shaw’s regions. This finding suggests the need for a thorough exploration of markers prior to undertaking a large comparative sequencing project. The methods employed here to examine expected primer utility can easily be applied to any taxon, provided complete chloroplast genomic data are available. When complete genome data are lacking, the results presented here can provide a rough estimate of the “best primers,” but this remains a work in progress.

LITERATURE CITED

ASANO, T., T. TSUEZUKI, S. TAKAMASHI, H. SHIMADA, AND K. KADOWAKI. 2004. Complete nucleotide sequence of the sugarcane (*Sacccharum officinarum*) chloroplast genome: A comparative analysis of four monocot chloroplast genomes. *DNA Research* 11: 93–99.

ASMUSSEN, C. B., AND M. W. CHASE. 2001. Coding and noncoding plastid DNA in palm systematics. *American Journal of Botany* 88: 1103–1117.

BARRETT, C. F., C. D. SPECHT, J. LEEBRENS-MACK, D. W. STEVENSON, W. B. ZIMLEFFER, AND J. J. DAVIS. 2014. Resolving ancient radiations: Can complete plastid gene sets elucidate deep relationships among the tropical gingers (Zingiberales)? *Annals of Botany* 113: 119–133.

BERNARD, G., P. HERNANDEZ, B. KHADARI, G. DORADO, AND V. SAVOLAINEN. 2011. Genomic profiling of plastid DNA variation in the Mediterranean olive tree. *BMC Plant Biology* 11: 80.

BOCK, D. G., N. C. KANE, D. P. EBERT, AND L. H. RIESEBERG. 2004. Complete nucleotide sequence of the sugarcane (*Sacccharum officinarum*) chloroplast genome: A comparative analysis of four monocot chloroplast genomes. *DNA Research* 11: 93–99.

BRACKETT, C. F., C. D. SPECHT, J. LEEBRENS-MACK, D. W. STEVENSON, W. B. ZIMLEFFER, AND J. J. DAVIS. 2014. Resolving ancient radiations: Can complete plastid gene sets elucidate deep relationships among the tropical gingers (Zingiberales)? *Annals of Botany* 113: 119–133.

BERNARD, G., P. HERNANDEZ, B. KHADARI, G. DORADO, AND V. SAVOLAINEN. 2011. Genomic profiling of plastid DNA variation in the Mediterranean olive tree. *BMC Plant Biology* 11: 80.

BOCK, D. G., N. C. KANE, D. P. EBERT, AND L. H. RIESEBERG. 2004. Complete nucleotide sequence of the sugarcane (*Sacccharum officinarum*) chloroplast genome: A comparative analysis of four monocot chloroplast genomes. *DNA Research* 11: 93–99.

BRACKETT, C. F., C. D. SPECHT, J. LEEBRENS-MACK, D. W. STEVENSON, W. B. ZIMLEFFER, AND J. J. DAVIS. 2014. Resolving ancient radiations: Can complete plastid gene sets elucidate deep relationships among the tropical gingers (Zingiberales)? *Annals of Botany* 113: 119–133.
Applications in Plant Sciences 2015 3(6): 1400085
doi:10.3732/apps.1400085

Prince—Plastid primers for phylogenetics

DONG, W., J. LIU, J. YU, L. WANG, and S. ZHOU. 2012. Highly variable plastid markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. PLoS ONE 7: e53071. 10.1371/journal.pone.0053071.

DONG, W., C. XI, T. CHENG, K. LIN, and S. ZHOU. 2013. Sequencing angiosperm plastid genomes made easy: A complete set of universal primers and a case study on the phylogeny of Saxifragales. Genome Biology and Evolution 5: 989–997.

DOORDUIIN, L., B. GRAVENDEL, Y. LAMMERS, Y. ARIYUREK, T. CHIN-AWONG, and K. VRIELING. 2011. The complete plastid genome of 17 individuals of pest species Jacobaea vulgaris: SNPs, microsatellites and barcoding markers for population phylogenetic studies. DNA Research 18: 93–105.

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

GAUT, B. S., S. V. MUSE, W. D. CLARK, and M. T. CLEGG. 1992. Relative rates of nucleotide substitution at the rbcL locus of monocotyledonous plants. Journal of Molecular Evolution 35: 292–303.

GOREMYKIN, V. V., K. I. HIRSCH-ERNST, S. WOLFEL, and F. H. HELLLWIG. 2003. Analysis of the Amborella trichopoda plastid genome sequence suggests that Amborella is not a basal angiosperm. Molecular Biology and Evolution 20: 1499–1505.

GOREMYKIN, V. V., K. I. HIRSCH-ERNST, S. WOLFEL, and F. H. HELLLWIG. 2004. The plastid genome of Nymphaea alba: Whole-genome analyses and the problem of identifying the most basally angiosperm. Molecular Biology and Evolution 21: 1445–1454.

GOREMYKIN, V. V., B. HOLLAND, K. I. HIRSCH-ERNST, and F. H. HELLLWIG. 2005. Analysis of Acorus calamus plastid genome and its phylogenetic implications. Molecular Biology and Evolution 22: 1813–1822.

GRAHAM, S. W., P. A. REEVES, A. C. E. BURNS, and R. G. OLMSTEAD. 2000. Microstructural changes in noncoding plastid DNA: Interpretation, evolution, and utility of indels and inversions in basal angiosperm plastid genome. International Journal of Plant Sciences 161: S83–S96.

GREVET, D., B. HEINZE, G. G. VENDRIMIN, and R. J. PETIT. 2001. Genome walking with consensus primers: Application to the large single copy region of plastid DNA. Molecular Ecology Notes 1: 345–349.

HEINZE, B. 2007. A database of PCR primers for the plastid genomes of higher plants. Plant Methods 3: 4.

HIRATSUKA, J., H. SHIMADA, R. WHITTIER, T. ISHIBASHI, M. SAKAMOTO, M. MOORE, M. J. BELL, P. S. SOLTIS, and D. E. SOLTIS. 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. Proceedings of the National Academy of Sciences, USA 104: 19363–19368.

HIROYA, Y., K. ISINO, T. KOHMA, A. ENDO, M. HANAKOA, T. SHINA, T. TERACHI, et al. 2002. Structural features of a wheat plastome as revealed by complete sequences of chloroplast DNA. DNA Research 9: 263: 581 – 585.

HOBLEY, W. L., B. SEARS. 2000. Complete nucleotide sequence of the plastid chromosome, representing plastome I of the five distinguishable Euonothera plastomes. Molecular & General Genetics 263: 581 – 585.

IDT (INTEGRATED DNA TECHNOLOGIES). 2009. Degenerate sequences and non-standard bases: A quick look. Technical publication downloaded from http://www.idtdna.com/pages/support/technical-vault/reading-room/technical-reports [accessed 3 December 2014].

JENSEN, R. K., and J. D. PALMER. 1987a. A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). Proceedings of the National Academy of Sciences, USA 84: 5818–5822.

JENSEN, R. K., and J. D. PALMER. 1987b. Chloroplast DNA from lettuce and Barnadesia (Asteraceae): Structure, gene localization, and characterization of a large inversion. Current Genetics 11: 553–564.

Katoh, T., T. Kaneko, S. Sato, Y. Nakamura, and A. Tabata. 2000. Complete structure of the plastid genome of a legume, Lotus japonicus. DNA Research 7: 323–330.

Kelchner, S. A., and L. G. Clark. 1997. Molecular evolution and phylogenetic utility of the chloroplast rpl16 intron in Chusquea and the Bambusoidaeae (Poaceae). Molecular Phylogenetics and Evolution 8: 385–397.

Leslie, C. H., and M. R. Duval. 2009. The complete plastid genome of Coix lacryma-jobi and a comparative molecular evolutionary analysis of plastomes in cereals. Journal of Molecular Evolution 69: 311–318.

Li, C., A. Zhou, and T. Sang. 2006. Genetic analysis of rice domestication syndrome with the wild annual species, Oryza nivara. New Phytologist 170: 185–194.

Maer, R. M., K. Neckermann, G. L. Ieder, and H. Kossel. 1995. Complete sequence of the maize plastid genome: Gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. Journal of Molecular Biology 251: 614–628.

Moore, M. J., C. D. Bell, P. S. Soltis, and D. E. Soltis. 2007. Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. Proceedings of the National Academy of Sciences, USA 104: 19363–19368.

Niisuna, W., A. Liston, R. Cronn, T. L. Ashman, and N. Bassil. 2013. Insights into phylogeny, sex function and age of Fragaria based on whole plastome genome sequencing. Molecular Phylogenetics and Evolution 66: 17–29.

Ochiara, Y., K. Isino, T. Kohma, A. Endo, M. Hanakoa, T. Shina, T. Terachi, et al. 2002. Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. Molecular Genetics and Genomics 266: 740–746.

Rambaut, A. 1996. Se-Al (v2.0a11) Sequence Alignment Editor. Available at http://evolve.zoo.ox.ac.uk/ University of Oxford, Oxford, United Kingdom.

Roychle, W. 2002. Oligo Primer Analysis Software v. 6. Molecular Biology Insights, Cascade, Colorado, USA.

Särkinen, T., and M. George. 2013. Predicting plastid marker variation: Can complete plastid genomes from closely related species help? PLoS ONE 8: e82266. 10.1371/journal.pone.0082266.

Sasaki, C., S.-B. Lee, H. Daniell, T. C. Wood, J. Tomkins, H.-G. Kim, and R. K. Jensen. 2005. Complete plastome genome sequence of Glycine max and comparative analyses with other legume genomes. Plant Molecular Biology 59: 309 – 322.

Sasaki, C.-S. B.-Lee, S. Fiehlheim, C. Guda, R. K. Jensen, H. Joo, J. Tomkins, et al. 2007. Complete plastome genome sequences of Hordeum vulgare, Sorghum bicolor and Agrostis stolonifera, and comparative analyses with other grass genomes. Theoretical and Applied Genetics 115: 571–590.

Sato, S., Y. Nakamura, T. Kaneko, E. Asami, and S. Tabata. 1999. Complete structure of the plastid genome of Arabidopsis thaliana. DNA Research 6: 283–290.

Scarcelli, N., A. Barnaud, W. Eisnerhardt, U. A. Treher, M. Seveno, A. d’Anfayr, Y. Vignous, and J.-C. Pentaud. 2011. A set of 100 plastid DNA primer pairs to study population genetics and phylogeny in monocotyledons. PLoS ONE 6: e19594. 10.1371/journal.pone.0019594.

Schmitz-Linneweber, C., R. M. Maier, J. P. Alcaraz, A. Cottet, R. G. Herrmann, and R. Mach. 2001. The plastid chromosome of spinach (Spinacia oleracea): Complete nucleotide sequence and gene organization. Plant Molecular Biology 45: 307–315.

Schmitz-Linneweber, C., R. Riegel, T. G. Du, H. Hupper, R. G. Herrmann, and R. M. Maier. 2002. The plastid chromosome of Atropa belladonna and its comparison with that of Nicotiana tabacum: The role of RNA editing in generating divergence in the process of plant speciation. Molecular Biology and Evolution 19: 1602–1612.

Shahid Masood, M., T. Nishikawa, S. Fukukawa, P. K. Nigewa, T. Tsuzuki, and K. Kadwani. 2004. The complete nucleotide sequence of wild rice (Oryza nivara) plastome chromosome: First genome wide comparative sequence analysis of wild and cultivated rice. Gene 340: 133–139.
Applications in Plant Sciences 2015 3(6): 1400085
doi:10.3732/apps.1400085

Prince—Plastid primers for phylogenetics

---

# Appendix 1. Complete chloroplast genome sequences used to design universal flowering plant primers for 36 plastid noncoding regions.

| Family       | Species                        | GenBank number and version | Reference                                                                 |
|--------------|--------------------------------|----------------------------|---------------------------------------------------------------------------|
| **Basal Dicot Grade:** |                                |                            |                                                                           |
|              | *Amborella trichopoda* Baill.  | NC_005086.1; Goremykin et al., 2003. |
| **Monocots:** |                                |                            |                                                                           |
| 2. Oryza *nivara* Sharma & Shastry; | NC_005973.1; Shahid Masood et al., 2004. |
| 3. *Oryza sativa* L.; | NC_001320.1; Hiratsuka et al., 1989. |
| 4. *Saccharum* hybrid; | NC_005878.2; Calsa et al., 2004. |
| 5. *Saccharum officinarum* L.; | NC_006084.1; Asano et al., 2004. |
| 6. *Triticum aestivum* L.; | NC_002762.1; Ogihara et al., 2002. |
| 7. *Zea mays* L.; | NC_001666.2; Maier et al., 1995. |
| **Eudicots:** |                                |                            |                                                                           |
| 8. *Arabidopsis thaliana* (L.) Heynh.; | NC_000932.1; Sato et al., 1999. |
| 9. *Atropa belladonna* L.; | NC_004561.1; Schmitz-Linneweber et al., 2001. |
| 10. *Calycaentus floridus* L. var. *glaucus* (Wild.) Torr. & A. Gray; | NC_004993.1; Goremykin et al., unpublished (Goremykin, V., K. Hirsch-Ernst, S. Wolfl, and F. Hellwig. Complete structure of the chloroplast genome of *Calycaentus fertils*. Direct GenBank submission 9 July 2003). |
| 11. *Lota japonicus* (Regel) K. Larsen; | AP002983.1; Kato et al., 2000. |
| 12. *Medicago truncatula* Gaertn.; | NC_003119.6; Lin et al., unpublished (Lin, S., H. Wu, H. Jia, P. Zhang, R. Dixon, G. May, R. Gonzalez, and B. A. Roe. *Medicago truncatula* variety Jema Long A-17 chloroplast, complete sequence. Direct GenBank submission 31 August 2001). |
| 13. *Nicotiana tabacum* L.; | Z00044.1; Shinozaki et al., 1986. Note: this sequence has been updated since this article was published. |
| 14. *Nymphaea alba* L.; | NC_006050.1; Goremykin et al., 2004. |
| 15. *Oenothera elata* Kunth subsp. *hoekeri* (Torr. & A. Gray) W. Dietr. & W. L. Wagner; | NC_002693.1; 10.1371/journal.pone.0037128. |
| 16. *Spinacia oleracea* L.; | NC_002202.1; Schmitz-Linneweber et al., 2001. |

---

References:

1. SHAW, J. E. B. LICKEY, E. E. SCHILLING, and R. L. SMALL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *American Journal of Botany* 104: 1987–2004.

2. SHAW, J., H. L. SHAFER, O. R. LEONARD, M. J. KOVACH, M. SCHORR, and A. B. MORRIS. 2014. Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: The tortoise and the hare IV. *American Journal of Botany* 101: 2012 – 2014.

3. SHINOZAKI, K., M. ORIME, M. TANAKA, T. WAKASUGI, N. HAYASHIDA, T. MITSUBAYASHI, N. ZAITA, et al. 1986. The complete nucleotide sequence of tobacco chloroplast genome: Its gene organization and expression. *EMBO Journal* 5: 2043–2049.

4. SHULAEV, V., D. J. SARGENT, R. N. CROWHURST, T. C. MOCKLER, O. FOLKERTS, A. L. DELCHER, P. JAIWAL, et al. 2011. The genome of woodland strawberry (*Fragaria vesca*). *Nature Genetics* 43: 109–116.

5. TANG, J., H. XIA, M. CAO, X. ZHANG, W. ZENG, S. HE, W. TONG, et al. 2004. A comparison of rice chloroplast genomes. *Plant Physiology* 135: 412–420.

6. TIMME, R. E., J. V. KUEHL, J. L. BOORE, and R. K. JANSEN. 2007. A comparative analysis of the *Lactuca* and *Helianthus* (Asteraceae) plastid genomes: Identification of divergent regions and categorization of shared repeats. *American Journal of Botany* 94: 302–312.
### APPENDIX 2. Comparison of chloroplast regions with published primer pairs.

| Approx. Nicotiana | Primary type | Location | Genomic region | Shaw et al., 2005, 2007 | Ebert and Peakall, 2009 | Scarcelli et al., 2011 | Dong et al., 2012 | Dong et al., 2013 | Current study |
|-------------------|--------------|----------|----------------|--------------------------|--------------------------|------------------------|-----------------|-----------------|---------------|
| IGS LSC           |              | trnH(GUG)-psbA | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| exon              |              | psbA exon    | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| IGS LSC           |              | trnK(UUU)-matK | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| exon              |              | matK exon    | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| IGS LSC           |              | matK-trnK5` | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| IGS LSC           |              | trnK(UUU)-psr16 | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| intron            |              |                         |               |                         |                         |                        |                 |                 |               |
| IGS LSC           |              | psbK-trnS(GCU) | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| IGS LSC           |              | trnS(GCU)-trn(G(UCC)) and intron | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| intron            |              |                         |               |                         |                         |                        |                 |                 |               |
| IGS LSC           |              | trnU(UCC)-trn(A) | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| exon              |              | atpA exon    | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| IGS LSC           |              | rps2 exon    | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| intron            |              |                         |               |                         |                         |                        |                 |                 |               |
| IGS LSC           |              | rps2-trnB | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| exon              |              |                         |               |                         |                         |                        |                 |                 |               |
| IGS LSC           |              | rpoC1 exon    | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| intron            |              |                         |               |                         |                         |                        |                 |                 |               |
| IGS LSC           |              | rpoC1-trn(C)  | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| exon              |              |                         |               |                         |                         |                        |                 |                 |               |
| IGS LSC           |              | rps2-exon    | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| intron            |              |                         |               |                         |                         |                        |                 |                 |               |
| IGS LSC           |              | rps2-trnB | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| exon              |              |                         |               |                         |                         |                        |                 |                 |               |
| IGS LSC           |              | rpoC2-exon    | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| intron            |              |                         |               |                         |                         |                        |                 |                 |               |
| IGS LSC           |              | rps2-trnB | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| exon              |              |                         |               |                         |                         |                        |                 |                 |               |
| IGS LSC           |              | rps2-trnB | ✓              | ✓                        | ✓                        | ✓                      | ✓               | ✓               |               |
| exon              |              |                         |               |                         |                         |                        |                 |                 |               |

---

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

11 of 14
| Approx. Nicotiana order | Primary type | Location | Genomic region | Shaw et al., 2005, 2007 | Ebert and Peakall, 2009 | Scarcelli et al., 2011 | Dong et al., 2012 | Dong et al., 2013 | Current study |
|-------------------------|--------------|----------|----------------|-------------------------|-----------------------|-----------------------|-----------------|-----------------|---------------|
| 69                      | Exon         | LSC      | accD exon      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 70                      | IGS          | LSC      | accD-psaJ      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 71                      | IGS          | LSC      | psaJ-ycf4      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 72                      | Exon         | LSC      | ycf4 exon      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 73                      | IGS          | LSC      | ycf4-ycf10(cemA) | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 74                      | Exon         | LSC      | cemA           | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 75                      | IGS          | LSC      | ycf4-petA      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 76                      | Exon         | LSC      | petA exon      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 77                      | IGS          | LSC      | petA-psbJ      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 78                      | IGS          | LSC      | psbJ-psbE      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 79                      | IGS          | LSC      | petA-psbL      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 80                      | IGS          | LSC      | psbE-petL      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 81                      | IGS          | LSC      | petL-psaJ      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 82                      | IGS          | LSC      | petL-trnP(UGG) | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 83                      | IGS          | LSC      | tmrW(CCA)-psaJ | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 84                      | IGS          | LSC      | tmrP(UGG)-rps18 | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 85                      | IGS          | LSC      | psaJ-rpl20     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 86                      | IGS          | LSC      | rps18-rps12    | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 87                      | IGS          | LSC      | rpl20-rps12    | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 88                      | IGS          | LSC      | rps12-psbB     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 89                      | IGS          | LSC      | rps12-clpP     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 90                      | Intron       | LSC      | clpP intron 2  | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 91                      | Intron       | LSC      | clpP intron 1  | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 92                      | IGS          | LSC      | clpP-psbB      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 93                      | Exon         | LSC      | psbB exon      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 94                      | IGS          | LSC      | psbB-psbH      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 95                      | IGS          | LSC      | psbH-petBE2    | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 96                      | Intron       | LSC      | petB intron/exon 2 | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 97                      | IGS          | LSC      | petBE2-petDE2  | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 98                      | Intron       | LSC      | petD intron/exon 2 | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 99                      | IGS          | LSC      | petD-rpoA      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 100                     | Exon         | LSC      | rpoA exon      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 101                     | IGS          | LSC      | rpoA-rps11     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 102                     | IGS          | LSC      | rps11-rps8     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 103                     | Exon         | LSC      | rps8 exon      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 104                     | IGS          | LSC      | rpl36-rpl14    | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 105                     | IGS          | LSC      | rps8-rpl16     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 106                     | Intron       | LSC      | rpl16 intron   | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 107                     | IGS          | LSC      | rpl16-rps3     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 108                     | Exon         | LSC      | rps3 exon      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 109                     | IGS          | LSC      | rps3-rps19     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 110                     | IGS          | LSC      | rpl22-rpl2     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 111                     | Intron       | IRb      | rpl2 intron/exon 1-2 | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 112                     | IRb          | IRb      | rpl23-ycf2     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 113                     | Exon         | IRb      | ycf2 exon      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 114                     | IRb          | IRb      | ycf2-ndhB      | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 115                     | Exon         | IRb      | ndhB exon 2    | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 116                     | Intron       | IRb      | ndhB intron/exon 1 | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 117                     | IRb          | IRb      | ndhB-rps7     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 118                     | IRb          | IRb      | rps7-rps12    | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 119                     | Intron       | IRb      | rps12 intron/exon | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 120                     | IGS          | IRb      | rps12-trnP(GAC) | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 121                     | IGS          | IRb      | trnP(GAC)-rnn16 | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 122                     | Exon         | IRb      | rnn16 exon    | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 123                     | IRb          | IRb      | rnn16-trnP(GAU) | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 124                     | Intron       | IRb      | trnP(GAU) intron | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 125                     | Intron       | IRb      | trnA(UGC) intron | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 126                     | IGS          | IRb      | trnA(UGC)-rnn23 | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 127                     | Exon         | IRb      | rnn23 exon    | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 128                     | IGS          | IRb      | rnr4.5-trnN(GU) | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 129                     | IRb          | IRb      | trnN(GU)-ycf1  | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 130                     | IRb          | SSC      | ndhF exon     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 131                     | Exon         | SSC      | ndhF-rps32    | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 132                     | IGS          | SSC      | ndhF-rps32    | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 133                     | IGS          | SSC      | rps32-ccsA    | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 134                     | IGS          | SSC      | rps32-trnP(UAG) | ✓                     | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 135                     | Exon         | SSC      | ccsA exon     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
| 136                     | IGS          | SSC      | ccsA-ndhD     | ✓                       | ✓                     | ✓                     | ✓               | ✓               | ✓             |
## APPENDIX 2. Continued.

| Approx, Nicotiana order | Primary type | Location | Genomic region | Shaw et al., 2005, 2007 | Ebert and Peakall, 2009 | Scarcelli et al., 2011 | Dong et al., 2012 | Dong et al., 2013 | Current study |
|-------------------------|--------------|----------|----------------|-------------------------|------------------------|----------------------|------------------|------------------|---------------|
| 137 Exon                | SSC          | ndhD exon|                 | ✓                       | ✓                      |                      |                   |                  |               |
| 138 IGS                 | SSC          | ndhD-ndhE|                 | ✓                       | ✓                      |                      |                   |                  |               |
| 139 IGS                 | SSC          | psaC-ndhE|                 |                         |                        |                      |                   |                  |               |
| 140 IGS                 | SSC          | psaC-ndhG|                 | ✓                       |                        |                      |                   |                  |               |
| 141 IGS                 | SSC          | ndhE-ndhF|                 | ✓                       |                        |                      |                   |                  |               |
| 142 Exon                | SSC          | ndhG exon|                 | ✓                       |                        | *                    |                   |                  |               |
| 143 IGS                 | SSC          | ndhG-ndhI|                 | ✓                       | ✓                      |                      |                   |                  |               |
| 144 Intron              | SSC          | ndhA intron|               | ✓                       | ✓                      | ✓                    | ✓                |                  |               |
| 145 IGS                 | SSC          | ndhA-ndhH|                 | ✓                       |                        |                      |                   |                  |               |
| 146 Exon                | SSC          | ndhH exon|                 | ✓                       |                        |                      |                   |                  |               |
| 147 IGS                 | SSC          | ndhH-rps15|               | ✓                       |                        |                      |                   |                  |               |
| 148 IGS                 | SSC/IRa      | rps15-ycf1 |               | ✓                       |                        |                      |                   |                  |               |
| 149 IGS                 | IRa          | ycf1-rrn5 |               | ✓                       |                        |                      |                   |                  |               |
| Bonus                   | LSC          | rbcL-psaI |                 | ✓                       |                        |                      |                   |                  |               |
| Bonus                   | LSC          | trnS-psbD |                 | ✓                       |                        |                      |                   |                  |               |

*a Several regions overlap.

*b IR = inverted repeat; LSC = large single-copy region; SSC = small single-copy region.

*Slightly different region from that listed.
APPENDIX 3. Complete chloroplast genome sequences used to assess primer utility. Format: Organism; GenBank number and version; publication.

| Basal Dicot Grade: |  |
|-------------------|--|
| 1. Amborella trichopoda Baill.; NC_005086.1; Goremykin et al., 2003. |  |
| 2. Magnolia grandiflora L.; NC_020318.1; Li et al., unpublished (direct GenBank submission dated 22 February 2013). |  |

| Monocots: |  |
|-----------|--|
| 3. Acorus calamus L.; AJ879453.1; Goremykin et al., 2005. |  |
| 4. Cymbidium aloifolium (L.) Sw.; NC_021429.1; Yang et al., 2013. |  |
| 5. Cymbidium maunii Rchb. f.; NC_021433.1; Yang et al., 2013. |  |
| 6. Cymbidium sinense (Jacks. ex Andrews) Willd.; NC_021430.1; Yang et al., 2013. |  |
| 7. Cymbidium tortisepalum Fukuy.; NC_021431.1; Yang et al., 2013. |  |
| 8. Cymbidium tracyanum Rolfe; NC_021432.1; Yang et al., 2013. |  |
| 9. Oryza nivara Sharma & Shastry; NC_005973.1; Shahid Masood et al., 2004. |  |
| 10. Oryza sativa L. Indica group; NC_008155.1; Tang et al., 2004. |  |
| 11. Oryza sativa L. Japonica group; NC_001320.1; Hiratsuka et al., 1989. |  |
| 12. Canna indica L.; KF601570.1; Barrett et al., 2014. |  |

| Basal Eudicot Grade: |  |
|---------------------|--|
| 13. Ceratophyllum demersum L.; NC_009962.1; Moore et al., 2007. |  |
| 14. Nelumbo lutea Willd.; NC_015605.1; Quan and Ding, unpublished (direct GenBank submission dated 16 February 2009). |  |
| 15. Nelumbo nucifera Gaertn.; NC_015610; Quan and Ding, unpublished (direct GenBank submission dated 16 February 2009). |  |

| Eurosids I: |  |
|------------|--|
| 16. Fragaria vesca L. subsp. bracteata (A. Heller) Staudt; NC_018766.1; Njuguna et al., 2013. |  |
| 17. Fragaria vesca L. subsp. vesca; NC_015206.1; Shulaev et al., 2011. |  |

| Eurosids II: |  |
|-------------|--|
| 18. Gossypium herbaceum L.; NC_023215.1; Shang et al., unpublished (Shang, M., K. Wang, J. Hua, F. Liu, C. Wang, X. Zhang, Y. Wang, and S. Li. Gossypium herbaceum chloroplast, complete genome. Direct GenBank submission 11 February 2011). |  |
| 19. Gossypium herbaceum L. subsp. africanaum (G. Watt) Vollesen; NC_016692.1; Xu et al., 2012. |  |

| Euasterids I: |  |
|--------------|--|
| 20. Olea europaea L.; NC_013707.2; Messina, unpublished (Messina, R. Olea europaea chloroplast, complete genome. Direct GenBank submission 3 March 2007). |  |
| 21. Olea europaea L. subsp. cuspidata (Wall. ex G. Don) Cif.; NC_015604.1; Besnard et al., 2011. |  |
| 22. Olea europaea L. subsp. europea; NC_015401.1; Besnard et al., 2011. |  |
| 23. Olea europaea L. subsp. maroccana (Greuter & Burdet) P. Vargas; NC_015623.1; Besnard et al., 2011. |  |

| Euasterids II: |  |
|---------------|--|
| 24. Helianthus annuus L.; NC_007977.1; Timme et al., 2007. |  |
| 25. Helianthus decapetalus L.; NC_023110.1; Bock et al., 2014. |  |
| 26. Helianthus decapetalus L.; NC_023109.1; Bock et al., 2014. |  |
| 27. Helianthus giganteus L.; NC_023107.1; Bock et al., 2014. |  |
| 28. Helianthus grosseserratus M. Martens; NC_023108.1; Bock et al., 2014. |  |
| 29. Helianthus hirsutus Raf.; NC_023111.1; Bock et al., 2014. |  |
| 30. Helianthus maximiliani Schrad.; NC_023114.1; Bock et al., 2014. |  |
| 31. Helianthus strumosus L.; NC_023113.1; Bock et al., 2014. |  |
| 32. Helianthus tuberosus L.; NC_023112.1; Bock et al., 2014. |  |