Data Article

Data characterizing the chloroplast genomes of extinct and endangered Hawaiian endemic mints (Lamiaceae) and their close relatives

Andreanna J. Welch a,*, Katherine Collins a, Aakrosh Ratan b, Daniela I. Drautz-Moses c, Stephan C. Schuster c, Charlotte Lindqvist a,***

a Department of Biological Sciences, University at Buffalo (SUNY), Buffalo, NY 14260, USA
b Department of Public Health Sciences and Center for Public Health Genomics, University of Virginia, Charlottesville 22908, VA, USA
c Singapore Centre on Environmental Life Sciences Engineering, Nanyang Technological University, Singapore 637551, Singapore

Article info

Article history:
Received 3 March 2016
Received in revised form 5 March 2016
Accepted 7 March 2016
Available online 14 March 2016

Keywords:
Hawaii
Lamiaceae
Plastid genomes
Genome structure

Abstract

These data are presented in support of a plastid phylogenomic analysis of the recent radiation of the Hawaiian endemic mints (Lamiaceae), and their close relatives in the genus Stachys. “The quest to resolve recent radiations: Plastid phylogenomics of extinct and endangered Hawaiian endemic mints (Lamiaceae)” [1]. Here we describe the chloroplast genome sequences for 12 mint taxa. Data presented include summaries of gene content and length for these taxa, structural comparison of the mint chloroplast genomes with published sequences from other species in the order Lamiales, and comparisons of variability among three Hawaiian taxa vs. three outgroup taxa. Finally, we provide a list of 108 primer pairs targeting the most variable regions within this group and designed specifically for amplification of DNA extracted from degraded herbarium material.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Specifications Table

| Subject area                      | Biology, genetics, genomics                  |
|----------------------------------|---------------------------------------------|
| More specific subject area       | Molecular phylogenetics and evolution       |
| Type of data                     | Tables and figures                           |
| How data was acquired            | High-throughput sequencing of contemporary and herbarium samples was conducted on the Illumina HiSeq 2500 and MiSeq platforms, followed by both de novo and reference-guided assemblies, mapping, and functional annotation |
| Data format                      | Raw, and analyzed                            |
| Experimental factors             | De novo assemblies were created using SOAPdenovo and reference-guided assemblies were created using YASRA. Sequences were functionally annotated using DOGMA. SNPs were called and filtered using SAMtools and BCFtools |
| Experimental features            | Data include chloroplast genome gene content, structure, and comparisons of variable loci in a suite of recently diverged species and outgroups |
| Data source location             | Hawaii, North America, South America, Europe, Africa, and Asia |
| Data accessibility               | Data are published with this article         |

Value of the data

- These data provide a summary of the characteristics and structure of the chloroplast genomes of several taxa within Lamiaceae, which can be used to increase our understanding of molecular evolution of the chloroplast genome, as well as the evolution of its structure and function.
- A comparison of variable regions in mints can be used to identify rapidly evolving regions in other taxa.
- Primer sequences described here can be used to target highly variable regions in closely related taxa.

1. Data

Raw, demultiplexed sequence reads have been deposited in the NCBI sequence read archive (SRP070171) and full chloroplast genomes for 12 mint taxa have been deposited in GenBank (KU724130-KU724141). Data presented in the text include tables and figures giving information on gene content and variability in these 12 species, as well as comparison of genome structure with other members of the order Lamiales.

2. Experimental design, materials and methods

2.1. Samples, library construction, and shotgun sequencing

We selected 12 Hawaiian mint taxa for shotgun sequencing (five contemporary and seven from herbarium collections ranging up to ~100 years old), of which two extinct species were represented by two accessions each (see Tables 1 and 2 in [1]). We also sequenced four Stachys species, representing both close and more distantly related relatives.

DNA extraction, library construction and shotgun sequencing followed the methods described in [1]. Briefly, approximately 100 mg dried leaf tissue was homogenized using the TissueLyser system (Qiagen), and DNA was extracted using the DNeasy plant mini kit (Qiagen). DNA isolated from herbarium samples was processed separately from contemporary samples using stringent protocols and controls to prevent and detect any potential contamination. For contemporary samples, DNA extracts were sheared to 200–600 bp via sonication in a Covaris S220; DNA from herbarium samples is...
naturally degraded and therefore was not sheared further. Genomic shotgun sequencing libraries were constructed following the standard Illumina Tru-seq protocol for contemporary samples, or the NEBNext Library Prep Mastermix kit (New England Biolabs) for herbarium samples. Libraries were quantified using the PicoGreen High Sensitivity assay and then pooled and sequenced on the Illumina

### Table 1

Gene content of the chloroplast genome of *Stenogyne haliakalae* and 11 additional mint species.

| Gene Products          | Genes                                      |
|------------------------|--------------------------------------------|
| Photosystem I          | psaA, psaB, psaC, psaI, psaJ, ycf3 [20], ycf4 [20] |
| Photosystem II         | psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ/lhbA |
| Cytochrome b6/f        | petA, petB, petD, petG, petL, petN          |
| ATP synthase           | atpA, atpB, atpE, atpF, atpH, atpL          |
| Rubisco                | rbcL                                       |
| NADH oxidoreductase    | ndhA, ndhB, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK |
| RNA polymerase         | rpoA, rpoB, rpoC1, rpoC2                   |
| Large subunit ribosomal proteins | rpl2 [a], rpl14, rpl16, rpl20, rpl22, rpl23 [b], rpl32, rpl33, rpl36 |
| Small subunit ribosomal proteins | rps2, rps3, rps4, rps7 [a], rps8, rps11, rps12 [b], rps14, rps15, rps16, rps18, rps19 |
| Other functions        | accD, ccsA, cemA, clpP, matK, infA         |
| Unknown functions      | ycf1 [c], ycf2 [d], ycf15 [e]              |
| Ribosomal RNAs         | rrn23 [a], rrn16 [b], rrn5 [c], rrn4.5 [d] |
| Transfer RNAs          | trnA(UGC) [a], trnC(GCA), trnD(GUC), trnE(UUC), trnF(GAA), trnG(GCC), trnG(UCC), trnH(GUG), trnI(CAU) [e], trnI(GAU) [f], trnK(UUU), trnL(UAA), trnL(UAG), trnM(CAU), trnN(GUU) [f], trnP(UGC), trnQ(UUG), trnR(ACG) [g], trnR(UCU), trnS(GCU), trnS(GGA), trnS(UGA), trnT(GGU), trnT(UGU), trnV(UAC), trnV(GAC) [h], trnW(CCA), trnY(GUA) |

*a* Gene fully located within the inverted repeats.

*b* Gene partially located within the inverted repeats.

### Table 2

Genes containing introns in the chloroplast genomes of *Stenogyne haliakalae* and 11 additional mint species. Numbers represent the lengths (bp) of exons and introns in *S. haliakalae*.

| Gene     | Location | # Introns | Exon I | Intron I | Exon II | Intron II | Exon III |
|----------|----------|-----------|--------|----------|---------|-----------|----------|
| atpF     | LSC      | 1         | 143    | 656      | 410     | 291       | 616      | 227      |
| clpP     | LSC      | 2         | 70     | 658      | 291     | 616       | 227      |
| ndhA     | SSC      | 1         | 552    | 1020     | 538     | 776       |          |
| ndhB     | IR       | 1         | 755    | 680      | 776     |           |          |
| petB     | LSC      | 1         | 5      | 718      | 651     | 474       |          |
| petD     | LSC      | 1         | 7      | 728      | 474     |           |          |
| rpl16    | LSC      | 1         | 8      | 907      | 392     |           |          |
| rpl2     | IR       | 1         | 390    | 658      | 433     |           |          |
| rpoC1    | LSC      | 1         | 434    | 736      | 1634    |           |          |
| rps12    | LSC/IR[a] | 2       | 113    | a        | 231     | 537       | 25       |
| rps16    | LSC      | 1         | 39     | 875      | 226     |           |          |
| trnA-UGC | IR       | 1         | 37     | 807      | 34      |           |          |
| trnG-UCC | LSC      | 1         | 22     | 690      | 47      |           |          |
| trnL-GAU | IR       | 1         | 34     | 938      | 36      |           |          |
| trnK-UUU | LSC      | 1         | 36     | 2509     | 34      |           |          |
| trnL-UAA | LSC      | 1         | 36     | 488      | 49      |           |          |
| trnV-UAC | LSC      | 1         | 37     | 579      | 36      |           |          |
| ycf3     | LSC      | 2         | 123    | 713      | 229     | 725       | 152      |

*a* Trans-spliced
Table 3
Lengths (bp) of the long single copy region (LSC), short single copy region (SSC), and inverted repeats regions (IR) for the chloroplast genomes of *Stenogyne haliakalae* and 11 additional mint species.

| Species                  | LSC (bp) | IR (bp) | SSC (bp) | Total (bp) |
|--------------------------|----------|---------|----------|------------|
| *Haplostachys haplostachya* | 81,755   | 25,441  | 17,495   | 150,132    |
| *Haplostachys linearifolia* | 81,752   | 25,441  | 17,495   | 150,129    |
| *Phyllostegia velutina*    | 81,765   | 25,440  | 17,496   | 150,141    |
| *Phyllostegia waimeae*     | 81,744   | 25,448  | 17,517   | 149,722    |
| *Stachys byzantina*       | 81,774   | 25,464  | 17,552   | 150,254    |
| *Stachys coccinea*        | 82,171   | 25,470  | 17,563   | 150,674    |
| *Stachys sylvatica*       | 81,807   | 25,414  | 17,560   | 150,195    |
| *Stenogyne bifida*        | 81,752   | 25,441  | 17,495   | 150,129    |
| *Stenogyne haliakalae*    | 81,364   | 25,436  | 17,499   | 149,736    |
| *Stenogyne kanehoana*     | 81,739   | 25,441  | 17,495   | 150,116    |
| *Stenogyne sessilis*      | 81,743   | 25,441  | 17,495   | 150,120    |

Fig. 1. Comparison of structure and similarity among 15 complete chloroplast genomes from the order Lamiales. *Cistanche deserticola* (102,657 bp) and *Epifagus virginiana* (70,028 bp), both members of the Orobanchaceae, are not considered here because they are parasitic and lack chlorophyll, thus demonstrating largely reduced chloroplast genomes. Blocks with the same color represent homologous regions free of internal structural changes for that subset of taxa, and those above the centerline for each taxon are in the same orientation as in *Stenogyne haliakalae*, whereas those below the line are in the reverse direction. Within each block a similarity profile for the region is plotted. Areas outside of blocks are presumed to represent lineage-specific regions of the chloroplast genome. One copy of the inverted repeat has been trimmed so that homology of the remaining repeat (area shaded in light gray) can be shown.
HiSeq and MiSeq platforms. Adapter sequences were trimmed from the reads using the Adapter-Removal software [2]. Assessment of DNA damage in old herbarium specimens was conducted using mapDamage 2.0 [3]. The presence of misincorporations characteristic of damaged DNA molecules typically found in old and degraded samples suggests that the data from herbarium samples are authentic, however, the overall levels of damage were low and within the range expected based on the age of the specimens (see Supplementary Figs. 2 and 3 in [1]).

2.2. Assembly of the Hawaiian mint reference chloroplast genome

Because no chloroplast genome sequence from a closely related taxon was available at the time this study was conducted, we implemented a combined reference-guided and de novo assembly approach [4] to determine the first complete chloroplast genome sequence for a Hawaiian mint. We assembled the sequence for Stenogyne haliakalae, an extinct species, as it had the largest number of reads. Briefly, the approach involved conducting both reference-guided assembly in YASRA 2.32 [5] with olive (Olea europaea, NC_013707; [6]) as the reference, as well as de novo assembly in

![Fig. 2. Conservation among 11 complete mint chloroplast genomes. The sequence for Haplostachys linearifolia was excluded due to missing data. A physical map is given at the top to show gene content and organization (see Fig. 2 in [1] for gene names and products). In the lower panels, regions of the genome are represented by bars, and those that are conserved among all 11 species are colored mauve, whereas those that are conserved among subsets of the taxa have different colors. The height of the bar shows the degree of similarity.](attachment:image.png)
Table 4
Sliding window analysis of variability of complete chloroplast genome sequences of three Hawaiian and three Stachys taxa (\textit{Stenogyne haliakalae}, \textit{Stenogyne bifida}, \textit{Haplostachys haplostachya}, \textit{Stachys chamissonis}, \textit{Stachys coccinea}, and \textit{Stachys sylvatica}). PICs – Potentially informative characters.

| Begin | End  | # PICs Stachys | # PICs Hawaiian |
|-------|------|----------------|-----------------|
| 1     | 999  | 8              | 1               |
| 1000  | 1999 | 7              | 2               |
| 2000  | 2999 | 6              | 4               |
| 3000  | 3999 | 11             | 1               |
| 4000  | 4999 | 7              | 0               |
| 5000  | 5999 | 5              | 1               |
| 6000  | 6999 | 10             | 1               |
| 7000  | 7999 | 8              | 0               |
| 8000  | 8999 | 6              | 2               |
| 9000  | 9999 | 9              | 0               |
| 10,000| 10,999| 4            | 1               |
| 11,000| 11,999| 3             | 0               |
| 12,000| 12,999| 4             | 0               |
| 13,000| 13,999| 9             | 4               |
| 14,000| 14,999| 4             | 2               |
| 15,000| 15,999| 1             | 0               |
| 16,000| 16,999| 3             | 0               |
| 17,000| 17,999| 6             | 0               |
| 18,000| 18,999| 5             | 3               |
| 19,000| 19,999| 2             | 0               |
| 20,000| 20,999| 3             | 0               |
| 21,000| 21,999| 0             | 0               |
| 22,000| 22,999| 5             | 1               |
| 23,000| 23,999| 1             | 0               |
| 24,000| 24,999| 0             | 0               |
| 25,000| 25,999| 1             | 0               |
| 26,000| 26,999| 4             | 1               |
| 27,000| 27,999| 5             | 3               |
| 28,000| 28,999| 9             | 2               |
| 29,000| 29,999| 4             | 1               |
| 30,000| 30,999| 11            | 1               |
| 31,000| 31,999| 5             | 1               |
| 32,000| 32,999| 3             | 0               |
| 33,000| 33,999| 3             | 1               |
| 34,000| 34,999| 8             | 1               |
| 35,000| 35,999| 1             | 0               |
| 36,000| 36,999| 1             | 0               |
| 37,000| 37,999| 5             | 0               |
| 38,000| 38,999| 0             | 0               |
| 39,000| 39,999| 1             | 0               |
| 40,000| 40,999| 5             | 0               |
| 41,000| 41,999| 3             | 1               |
| 42,000| 42,999| 1             | 0               |
| 43,000| 43,999| 5             | 2               |
| 44,000| 44,999| 6             | 1               |
| 45,000| 45,999| 10            | 0               |
| 46,000| 46,999| 3             | 1               |
| 47,000| 47,999| 7             | 0               |
| 48,000| 48,999| 0             | 0               |
| 49,000| 49,999| 6             | 1               |
| 50,000| 50,999| 3             | 0               |
| 51,000| 51,999| 2             | 0               |
| 52,000| 52,999| 3             | 1               |
| 53,000| 53,999| 5             | 3               |
| 54,000| 54,999| 8             | 3               |
| 55,000| 55,999| 8             | 2               |
| 56,000| 56,999| 2             | 0               |
| 57,000| 57,999| 6             | 1               |
| Begin | End   | # PICs Stachys | # PICs Hawaiian |
|-------|-------|---------------|----------------|----------------|
| 58,000| 58,999| 9             | 1              |
| 59,000| 59,999| 4             | 2              |
| 60,000| 60,999| 3             | 3              |
| 61,000| 61,999| 7             | 1              |
| 62,000| 62,999| 2             | 0              |
| 63,000| 63,999| 6             | 1              |
| 64,000| 64,999| 5             | 1              |
| 65,000| 65,999| 6             | 1              |
| 66,000| 66,999| 8             | 0              |
| 67,000| 67,999| 12            | 0              |
| 68,000| 68,999| 6             | 0              |
| 69,000| 69,999| 0             | 0              |
| 70,000| 70,999| 7             | 0              |
| 71,000| 71,999| 5             | 3              |
| 72,000| 72,999| 3             | 0              |
| 73,000| 73,999| 3             | 1              |
| 74,000| 74,999| 4             | 2              |
| 75,000| 75,999| 2             | 1              |
| 76,000| 76,999| 5             | 0              |
| 77,000| 77,999| 5             | 0              |
| 78,000| 78,999| 7             | 0              |
| 79,000| 79,999| 7             | 1              |
| 80,000| 80,999| 4             | 1              |
| 81,000| 81,999| 10            | 0              |
| 82,000| 82,999| 2             | 2              |
| 83,000| 83,999| 0             | 1              |
| 84,000| 84,999| 1             | 0              |
| 85,000| 85,999| 1             | 0              |
| 86,000| 86,999| 2             | 0              |
| 87,000| 87,999| 0             | 0              |
| 88,000| 88,999| 1             | 1              |
| 89,000| 89,999| 0             | 0              |
| 90,000| 90,999| 1             | 0              |
| 91,000| 91,999| 0             | 0              |
| 92,000| 92,999| 1             | 0              |
| 93,000| 93,999| 0             | 0              |
| 94,000| 94,999| 2             | 1              |
| 95,000| 95,999| 1             | 0              |
| 96,000| 96,999| 2             | 0              |
| 97,000| 97,999| 0             | 0              |
| 98,000| 98,999| 0             | 0              |
| 99,000| 99,999| 1             | 0              |
| 100,000| 100,999| 1             | 0              |
| 101,000| 101,999| 1             | 0              |
| 102,000| 102,999| 1             | 0              |
| 103,000| 103,999| 0             | 0              |
| 104,000| 104,999| 5             | 2              |
| 105,000| 105,999| 1             | 0              |
| 106,000| 106,999| 4             | 0              |
| 107,000| 107,999| 9             | 0              |
| 108,000| 108,999| 4             | 0              |
| 109,000| 109,999| 16            | 3              |
| 110,000| 110,999| 6             | 2              |
| 111,000| 111,999| 10            | 1              |
| 112,000| 112,999| 7             | 0              |
| 113,000| 113,999| 8             | 3              |
| 114,000| 114,999| 4             | 0              |
| 115,000| 115,999| 5             | 0              |
| 116,000| 116,999| 6             | 4              |
| 117,000| 117,999| 9             | 4              |
| 118,000| 118,999| 3             | 0              |
SOAPdenovo v1.05 [7]. Assembly methods are described in more detail in [1]. The resulting contigs from both approaches were split into overlapping sequences, and then used as input for a further reference guided-assembly step in YASRA. Gaps between the final contigs were closed using PCR (see [1] for PCR reaction conditions and Supplementary Table 1 of this paper for primer information) and Sanger sequencing in both directions from high-quality DNA extracted from a contemporary sample of *Stengyne bifida*. This ensured that amplification could be carried out over potentially large gaps, which would not be possible with degraded DNA from the extinct *Stenogyne haliakalae*. Contigs and Sanger sequences were aligned in Sequencher 4.7 (Gene Codes) to create a pseudo-reference sequence [4]. Reads from *Stenogyne haliakalae* were then mapped to the pseudo-reference using BWA v. 0.6.2 [8]. The reference sequence was further refined through Sanger sequencing of areas with low coverage or poor mapping quality (e.g., the border between the inverted repeat and single copy region). Reads were mapped to the final sequence, PCR duplicates were flagged and removed with the MarkDuplicates tool of the Picard command line toolset (http://picard.sourceforge.net/index.shtml), and a consensus sequence was called using SAMtools [9].

### 2.3. Assembly of additional mint chloroplast genomes

Complete or nearly complete chloroplast genomes were assembled using similar methods for 11 additional taxa: seven from the endemic Hawaiian mints (two of which were from herbarium samples) and four *Stachys* outgroups (see Tables 1 and 2 in [11]). Since the Hawaiian mints have diverged recently, we used the new chloroplast genome sequence from *Stenogyne haliakalae* as the reference during reference-guided assembly for the remaining Hawaiian taxa. The resulting contigs were aligned to create an interim sequence, and then the reads were mapped to this using BWA and a final consensus sequence called using SAMtools.

Chloroplast genome sequences for the *Stachys* outgroup taxa were assembled in a similar manner. We first assembled the chloroplast genome sequence for *Stachys chamissonis*, as this species is most closely related to the Hawaiian lineage. We conducted independent YASRA runs using *Olea europaea* as the reference, in addition to newly available sequences from *Stenogyne haliakalae, Sesamum indicum* (NC_016433) [10], *Origanum vulgare* (JX880022) [11], and *Salvia miltiorrhiza* (NC_020431) [12]. The contigs from all five runs were aligned to create an interim sequence. The reads were then mapped to the interim sequence using BWA and a consensus called using SAMtools. Once the *Stachys chamissonis* sequence was assembled we used this as the reference in YASRA for reference guided assembly of both *Stachys coccinea* and *Stachys sylvatica*. For *Stachys byzantina*, the most distantly related outgroup, we performed the initial reference-guided assembly using the sequence from *Stachys chamissonis*, as well as *Olea europaea* and *Sesamum indicum*. The rest of the assembly proceeded as described for the other *Stachys* species.

**Table 4 (continued)**

| Begin   | End    | # PICs Stachys | # PICs Hawaiian |
|---------|--------|----------------|-----------------|
| 119,000 | 119,999| 5              | 0               |
| 120,000 | 120,999| 11             | 2               |
| 121,000 | 121,999| 10             | 1               |
| 122,000 | 122,999| 7              | 5               |
| 123,000 | 123,999| 11             | 2               |
| 124,000 | 124,999| 4              | 1               |
| **Min** | 0      | 0              | 0               |
| **Max** | 16     | 5              | 104             |
| **Total** | 565   | 5              | 104             |
| **Mean** | 4.52  | 0.83           |                 |
| **Median** | 4.00  | 0.00           |                 |
Table 5
Comparison of variability of coding genes by exon in complete chloroplast genome sequences of three Hawaiian and three *Stachys* taxa (*Stenogyne hallakalae*, *Stenogyne bifida*, *Haplostachys haplostachya*, *Stachys chamissonis*, *Stachys coccinea*, and *Stachys sylvatica*). Note: Exon numbers are defined by position in overall chloroplast genome sequence (not direction of gene) and exons that are completely conserved among these taxa and/or shorter than 100 bp have been excluded. PICs – Potentially informative characters.

| Region          | Length | #PICs per locus Stachys | #PICs per locus Hawaiian | % PICs per locus Stachys | % PICs per locus Hawaiian |
|-----------------|--------|-------------------------|--------------------------|--------------------------|--------------------------|
| psbA            | 1056   | 0                       | 2                        | 0.00                     | 0.19                     |
| matK            | 1530   | 14                      | 5                        | 0.92                     | 0.33                     |
| psbK            | 186    | 1                       | 0                        | 0.54                     | 0.00                     |
| atpA            | 1524   | 5                       | 1                        | 0.33                     | 0.07                     |
| atpF Exon1      | 411    | 1                       | 0                        | 0.24                     | 0.00                     |
| atpI            | 744    | 1                       | 0                        | 0.13                     | 0.00                     |
| rps2            | 711    | 1                       | 0                        | 0.14                     | 0.00                     |
| rpoC2           | 4083   | 16                      | 3                        | 0.39                     | 0.07                     |
| rpoC1 Exon1     | 1635   | 1                       | 0                        | 0.06                     | 0.00                     |
| rpoC1 Exon2     | 435    | 1                       | 0                        | 0.23                     | 0.00                     |
| rpoB            | 3213   | 4                       | 0                        | 0.12                     | 0.00                     |
| psbD            | 1062   | 4                       | 0                        | 0.38                     | 0.00                     |
| psbC            | 1422   | 3                       | 1                        | 0.21                     | 0.07                     |
| psaB            | 2205   | 6                       | 0                        | 0.27                     | 0.00                     |
| psaA            | 2253   | 2                       | 0                        | 0.09                     | 0.00                     |
| rps4            | 606    | 1                       | 0                        | 0.17                     | 0.00                     |
| ndhJ            | 477    | 3                       | 0                        | 0.63                     | 0.00                     |
| ndhK            | 678    | 3                       | 0                        | 0.44                     | 0.00                     |
| atpE            | 402    | 2                       | 0                        | 0.50                     | 0.00                     |
| atpB            | 1497   | 4                       | 0                        | 0.27                     | 0.00                     |
| rbcL            | 1446   | 12                      | 6                        | 0.83                     | 0.41                     |
| accD            | 1467   | 3                       | 0                        | 0.20                     | 0.00                     |
| ycf4            | 555    | 2                       | 0                        | 0.36                     | 0.00                     |
| cemA            | 690    | 2                       | 1                        | 0.29                     | 0.14                     |
| petA            | 963    | 2                       | 1                        | 0.21                     | 0.10                     |
| rpl23           | 201    | 1                       | 0                        | 0.50                     | 0.00                     |
| rpl20           | 387    | 2                       | 0                        | 0.52                     | 0.00                     |
| psbB            | 1527   | 9                       | 0                        | 0.59                     | 0.00                     |
| petB Exon2      | 651    | 2                       | 1                        | 0.31                     | 0.15                     |
| rpoA            | 1014   | 3                       | 1                        | 0.30                     | 0.10                     |
| rpl36           | 114    | 2                       | 0                        | 1.75                     | 0.00                     |
| rps8            | 414    | 1                       | 0                        | 0.24                     | 0.00                     |
| rpl14           | 369    | 2                       | 0                        | 0.54                     | 0.00                     |
| rpl16 Exon1     | 393    | 3                       | 0                        | 0.76                     | 0.00                     |
| rps3            | 663    | 3                       | 0                        | 0.45                     | 0.00                     |
| rpl22           | 465    | 1                       | 1                        | 0.22                     | 0.22                     |
| rps19           | 279    | 6                       | 0                        | 2.15                     | 0.00                     |
| rpl2 Exon1      | 434    | 1                       | 0                        | 0.23                     | 0.00                     |
| ycf2            | 6849   | 5                       | 2                        | 0.07                     | 0.03                     |
| ndhB Exon1      | 756    | 1                       | 0                        | 0.13                     | 0.00                     |
| rps7            | 468    | 2                       | 1                        | 0.43                     | 0.21                     |
| rnr23           | 2811   | 1                       | 0                        | 0.04                     | 0.00                     |
| ndhF            | 2229   | 16                      | 0                        | 0.72                     | 0.00                     |
| rpl32           | 177    | 2                       | 0                        | 1.13                     | 0.00                     |
| ccsA            | 970    | 3                       | 2                        | 0.31                     | 0.21                     |
| ndhD            | 1521   | 11                      | 1                        | 0.72                     | 0.07                     |
| psaC            | 244    | 1                       | 0                        | 0.41                     | 0.00                     |
| ndhE            | 306    | 1                       | 2                        | 0.33                     | 0.65                     |
| ndhG            | 531    | 1                       | 0                        | 0.19                     | 0.00                     |
| ndhI            | 507    | 2                       | 0                        | 0.39                     | 0.00                     |
| ndhA Exon1      | 539    | 3                       | 0                        | 0.56                     | 0.00                     |
| ndhA Exon2      | 553    | 3                       | 1                        | 0.54                     | 0.18                     |
| ycf1 SSC        | 4487   | 43                      | 11                       | 0.96                     | 0.25                     |
| Min             | 114    | 0                       | 0                        | 0.00                     | 0.00                     |
2.4. Gene content and structure of mint chloroplast genomes

The *Stenogyne haliakalae* reference sequence and sequences from additional species were annotated using a combination of DOGMA [13], tRNAscan-SE [14], and additional manual BLAST searches. The borders of the inverted repeats were identified with the program Inverted Repeats Finder [15]. Overall the chloroplast genome sequences assembled here are very similar to other Lamiales. Table 1 shows the gene content of the *Stenogyne haliakalae* chloroplast genome, which mirrors the gene content of the chloroplast genomes for the 11 additional mint taxa, including the *Stachys* outgroups. Table 2 lists the genes that contain introns (and the number of introns present) in the mint taxa investigated here. Table 3 gives the lengths of the full chloroplast genome for each sequence assembled, as well as the lengths of the inverted repeats and single copy regions.

To compare the genome structure of the Hawaiian and *Stachys* taxa to other taxa in the order Lamiales, we conducted analyses in Mauve 2.3.1 [16] (Fig. 1). In the analysis we included *Stenogyne haliakalae*, *Stachys byzantina*, *Ajuga reptans* (NC_023102), *Andrographis paniculata* (NC_022451), *Boea hygrometrica* (NC_016468), *Jasminum nudiflorum* (NC_008407), *Lindenber gia philippensis* (NC_022859), *Olea europaea* (NC_013707), *Origanum vulgare* (JX880022), *Pinguicula ehlersiae* (NC_023463), *Salvia miltiorrhiza* (NC_020431), *Schwalbea americana* (NC_023115), *Sesamum indicum* (NC_016433), *Tectona grandis* (NC_020098), and *Utricularia gibba* (NC_021449). The chloroplast genome sequences of some taxa in this analysis demonstrated rearrangements and inversions. Therefore, one of the inverted repeats was trimmed off at the coordinates suggested by Inverted Repeats Finder so that homology could be determined with the remaining region. Seed weight was set to 19, the gap opening penalty was set to $200/C_0$, and the gap extension penalty to $30/C_0$.

To compare the genome structure of all of the Hawaiian and *Stachys* mints to each other, we conducted a separate analysis in Mauve (Fig. 2). The sequences were assumed to be collinear. The seed weight was set to 7, the gap opening penalty was set to $200/C_0$, and the gap extension penalty to $30/C_0$.

### Table 5 (continued)

| Region | Length | #PICs per locus | % PICs per locus | \% PICs per locus |
|--------|--------|----------------|-----------------|-----------------|
| Max    | 6849   | 43             | 2.15            | 0.65            |
| Total  | 61110  | 225            | 23.43           | 3.45            |
| Mean   | 1153   | 4.25           | 0.44            | 0.07            |
| Median | 678    | 0.00           | 0.33            | 0.00            |

2.5. Variability in the chloroplast genome sequences of Hawaiian mints and *Stachys* outgroups

We investigated variability among the three highest quality chloroplast genome sequences of the Hawaiian mints (*Stenogyne haliakalae*, *Stenogyne bifida*, *Haplostachys haplostachya*), and the three highest quality sequences from the *Stachys* outgroups (*Stachys chamissonis*, *Stachys coccinea*, and *Stachys sylvatica*). These species represent all of the main lineages within our samples, and were sequenced at $>10 \times$ depth. We used BWA to map the reads for each of these species onto the *Stenogyne haliakalae* reference genome and used SAMtools and BCFtools to call SNPs with a SNP quality score $>30$. Annotations of the *S. haliakalae* reference genome were transferred to the locations of the SNPs. We compared the levels of chloroplast genome diversity among the six genomes by identifying unique, variable positions, which we refer to as potentially informative characters (PICs) [17,18]. We did not include indels and inversions in this definition, which have been included in other analyses of chloroplast genome variability.

To analyze diversity among the chloroplast genomes, we compared the number of PICs present in 1000 bp non-overlapping sliding windows across the entire chloroplast genome sequences (Table 4, see also Fig. 5 in [1]). We also compared the number of PICs per locus for coding (Table 5, Fig. 3a), intron (Table 6, Fig. 3b), and intergenic spacer and pseudogene regions (Table 7, Fig. 3c). Because this
approach does not take into account the length of the locus, very long loci appear to have more PICs than shorter loci. Therefore, we also divided the number of PICs by the total length of the region to give the percent PICs per locus (Tables 5–7, Fig. 4). However, very short regions may still appear to
Table 6
Comparison of variability of introns in complete chloroplast genome sequences of three Hawaiian and three Stachys taxa (Stenogyne haliakalae, Stenogyne bifida, Haplostachys haplostachya, Stachys chamissonis, Stachys coccinea, and Stachys sylvatica). Note: Exon and intron numbers are defined by position in the overall chloroplast genome sequence (not direction of gene). The trnK-UUU intron contains the gene matK. Here trnK-UUU-1 refers to the region between exon 1 and matK, and trnK-UUU-2 refers to the region between matK and exon2. PICs – Potentially informative characters.

| Region             | Length | #PICs per locus Stachys | #PICs per locus Hawaiian | % PICS per locus Stachys | % PICS per locus Hawaiian |
|--------------------|--------|--------------------------|--------------------------|--------------------------|--------------------------|
| trnK-UUU - 1      | 264    | 5                        | 0                        | 1.89                     | 0.00                     |
| trnK-UUU - 2      | 714    | 5                        | 0                        | 0.70                     | 0.00                     |
| rps16              | 874    | 5                        | 1                        | 0.57                     | 0.11                     |
| trnG-UCC          | 689    | 7                        | 0                        | 1.02                     | 0.00                     |
| atpF               | 655    | 3                        | 0                        | 0.46                     | 0.00                     |
| rpoC1              | 762    | 4                        | 1                        | 0.52                     | 0.13                     |
| ycf3-1            | 724    | 0                        | 1                        | 0.00                     | 0.14                     |
| ycf3-2            | 712    | 2                        | 0                        | 0.28                     | 0.00                     |
| trnL-UAA          | 487    | 4                        | 0                        | 0.82                     | 0.00                     |
| trnV-UAC          | 478    | 1                        | 1                        | 0.21                     | 0.21                     |
| clpP - 1          | 615    | 6                        | 0                        | 0.98                     | 0.00                     |
| clpP - 2          | 657    | 4                        | 0                        | 0.61                     | 0.00                     |
| petB               | 717    | 2                        | 0                        | 0.28                     | 0.00                     |
| petD               | 727    | 4                        | 2                        | 0.55                     | 0.28                     |
| rpl16              | 906    | 6                        | 0                        | 0.66                     | 0.00                     |
| rpl2               | 657    | 2                        | 2                        | 0.30                     | 0.30                     |
| trnL-GAU          | 937    | 1                        | 0                        | 0.11                     | 0.00                     |
| ndhA               | 1019   | 10                       | 7                        | 0.98                     | 0.69                     |
| Min                | 264    | 0                        | 0                        | 0.00                     | 0.00                     |
| Max                | 1019   | 10                       | 7                        | 1.89                     | 0.69                     |
| Sum                | 12594  | 71                       | 15                       | 10.95                    | 1.86                     |
| Mean               | 699.7  | 3.94                     | 0.83                     | 0.61                     | 0.10                     |
| Median             | 713    | 4.00                     | 0.00                     | 0.56                     | 0.00                     |

Table 7
Comparison of variability of intergenic spacers and pseudogenes in complete chloroplast genome sequences of three Hawaiian and three Stachys taxa (Stenogyne haliakalae, Stenogyne bifida, Haplostachys haplostachya, Stachys chamissonis, Stachys coccinea, and Stachys sylvatica). Note: Regions that are completely conserved among these taxa and/or shorter than 100 bp have been excluded. PICs – Potentially informative characters.

| Region             | Length | #PICs per locus Stachys | #PICs per locus Hawaiian | % PICS per locus Stachys | % PICS per locus Hawaiian |
|--------------------|--------|--------------------------|--------------------------|--------------------------|--------------------------|
| trnH-GUG-psbA      | 323    | 8                        | 1                        | 2.48                     | 0.31                     |
| psbA-trnK-UUU      | 244    | 2                        | 0                        | 0.82                     | 0.00                     |
| trnK-UUU-rps16     | 643    | 5                        | 0                        | 0.78                     | 0.00                     |
| rps16-trnQ-UUG     | 828    | 10                       | 1                        | 1.21                     | 0.12                     |
| psbK-psbl          | 374    | 7                        | 0                        | 1.87                     | 0.00                     |
| psbl-trns-GCU      | 121    | 1                        | 1                        | 0.83                     | 0.83                     |
| trnS-GCU-trnG-UCC  | 694    | 5                        | 1                        | 0.72                     | 0.14                     |
| trnG-UCC-trnR-UCU  | 160    | 2                        | 0                        | 1.25                     | 0.00                     |
| atpF-atpH          | 365    | 5                        | 1                        | 1.37                     | 0.27                     |
| atpH-atpl          | 975    | 8                        | 5                        | 0.82                     | 0.51                     |
| rps2-rpoC2         | 337    | 2                        | 0                        | 0.59                     | 0.00                     |
| rpoC2-rpoC1        | 154    | 1                        | 0                        | 0.65                     | 0.00                     |
| rpoB-trnC-GCA      | 1098   | 6                        | 4                        | 0.55                     | 0.36                     |
| trnC-GCA-petN      | 436    | 4                        | 0                        | 0.92                     | 0.00                     |
| petN-psbM          | 509    | 5                        | 2                        | 0.98                     | 0.39                     |
| psbM-trnD-GUC      | 518    | 3                        | 0                        | 0.58                     | 0.00                     |
| trnE-UUC-trnT-GGU  | 570    | 7                        | 1                        | 1.23                     | 0.18                     |
| Region                          | Length | #PICs per locus | % PICs per locus | #PICs per locus | % PICs per locus |
|--------------------------------|--------|----------------|-----------------|----------------|-----------------|
| trnF-GGU-psbD                  | 1075   | 8              | 0.74            | 1              | 0.09            |
| psbC-trnS-UGA                  | 232    | 3              | 1.29            | 1              | 0.43            |
| trnS-UGA-psbZ                  | 328    | 3              | 0.91            | 0              | 0.00            |
| psbZ-trnG-GCC                  | 224    | 2              | 0.89            | 0              | 0.00            |
| trnG-GCC-trnM-CAU              | 149    | 1              | 0.67            | 0              | 0.00            |
| psaA-ycf3                      | 730    | 7              | 0.96            | 0              | 0.00            |
| ycf3-trnS-GGA                  | 316    | 2              | 0.63            | 2              | 0.63            |
| trnS-GGA-rps4                  | 162    | 1              | 0.62            | 0              | 0.00            |
| rps4-trnT-UGU                  | 385    | 3              | 0.78            | 1              | 0.26            |
| trnT-UGU-trnL-UAA              | 690    | 9              | 1.30            | 0              | 0.00            |
| trnL-UAA-trnF-GAA              | 295    | 3              | 1.02            | 0              | 0.00            |
| trnF-GAA-ndhj                  | 658    | 0              | 0.00            | 1              | 0.15            |
| ndhC-trnV-UAC                  | 944    | 5              | 0.53            | 0              | 0.00            |
| trnV-UAC-trnM-CAU              | 188    | 1              | 0.53            | 0              | 0.00            |
| trnM-CAU-atpE                  | 213    | 2              | 0.94            | 0              | 0.00            |
| atpB-rbcL                      | 801    | 1              | 0.12            | 1              | 0.12            |
| rbcL-accD                      | 707    | 7              | 0.99            | 2              | 0.28            |
| accD-psaI                      | 397    | 1              | 0.25            | 1              | 0.25            |
| psaI-ycf4                      | 434    | 7              | 1.61            | 0              | 0.00            |
| ycf4-ermA                      | 306    | 5              | 1.63            | 1              | 0.33            |
| cemA-petA                      | 217    | 1              | 0.46            | 1              | 0.46            |
| petA-psbJ                      | 1059   | 9              | 0.85            | 3              | 0.28            |
| psbE-petL                      | 915    | 6              | 0.66            | 1              | 0.11            |
| petL-petG                      | 174    | 2              | 1.15            | 0              | 0.00            |
| petG-trnW-CCA                  | 125    | 1              | 0.80            | 0              | 0.00            |
| trnP-UGG-psaj                  | 270    | 2              | 0.74            | 1              | 0.37            |
| psaj-rpl33                     | 478    | 5              | 1.05            | 0              | 0.00            |
| rpl33-rps18                    | 146    | 0              | 0.00            | 1              | 0.68            |
| rps18-rpl20                    | 216    | 5              | 2.31            | 0              | 0.00            |
| rpl20-rps12                    | 739    | 10             | 1.35            | 0              | 0.00            |
| psbB-psbT                      | 184    | 1              | 0.54            | 1              | 0.54            |
| psbH-petB                      | 124    | 1              | 0.81            | 0              | 0.00            |
| petB-petD                      | 194    | 1              | 0.52            | 0              | 0.00            |
| infA-rps8                      | 123    | 1              | 0.81            | 0              | 0.00            |
| rps8-rpl14                     | 173    | 3              | 1.73            | 0              | 0.00            |
| rpl14-rpl16                    | 126    | 1              | 0.79            | 0              | 0.00            |
| rpl16-rps3                     | 141    | 1              | 0.71            | 1              | 0.71            |
| rps12-3’end-trnV-GAC           | 1448   | 3              | 0.21            | 0              | 0.00            |
| trnA-UGC-rmn23                 | 199    | 1              | 0.50            | 0              | 0.00            |
| rnm4.5-rmn5                    | 224    | 2              | 0.89            | 1              | 0.45            |
| rnm5-trnR-AGC                  | 234    | 2              | 0.85            | 1              | 0.43            |
| trnR-AGC-trnN-GUU              | 569    | 2              | 0.35            | 0              | 0.00            |
| ycf1Truncated                  | 1059   | 1              | 0.09            | 0              | 0.00            |
| ndhF-rpl32                     | 435    | 5              | 1.15            | 2              | 0.46            |
| rpl32-trnL-UAG                 | 737    | 13             | 1.76            | 2              | 0.27            |
| ccsA-ndhD                      | 180    | 4              | 2.22            | 0              | 0.00            |
| ndhD-psaC                      | 105    | 1              | 0.95            | 0              | 0.00            |
| psaC-ndhE                      | 251    | 5              | 1.99            | 0              | 0.00            |
| ndhG-ndhI                      | 375    | 6              | 1.60            | 0              | 0.00            |
| rps15-ycf1 SSC                 | 389    | 3              | 0.77            | 0              | 0.00            |
| Min                             | 105    | 0              | 0.00            | 0              | 0.00            |
| Max                             | 1448   | 13             | 2.48            | 5              | 0.83            |
| Sum                             | 29192  | 250            | 62.72           | 43             | 10.44           |
| Mean                            | 435.7  | 3.73           | 0.94            | 0.64           | 0.16            |
| Median                          | 328.0  | 3.00           | 0.82            | 0              | 0.00            |
have a high percentage of variable sites, when in fact only a small number of the sites were variable. To minimize this, we have excluded regions less than 100 bp in length, and for clarity we have also excluded regions that were conserved among all six taxa.

To identify the most variable regions of the mint chloroplast genome for targeted re-sequencing and high resolution phylogenetic analyzes, reads from all 15 taxa subjected to shotgun sequencing...
Table 8
Tailed primers used for multiplex amplification and targeted re-sequencing in mints. Sequences complimentary to the Illumina sequencing adapters were appended to the end of each primer (Forward: 5’ TCGTGGCAGCGTCAGATGTGATGTAAGAAGACAG-[locus specific sequence] and Reverse: 5’ GTCTCGTGGGCTCGGAGATGTGATGTAAGAAGACAG-[locus specific sequence]). The loci within each region are also indicated. When the names of two genes are given with an underscore between them, the intergenic spacer between these genes is included in the amplified region.

| Primer   | Sequence 5’ – 3’ | Tm  | Loci in region (excluding priming sites) |
|----------|------------------|-----|-----------------------------------------|
| Mint204F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCG CCC TCT ACT ATA ATG GTA GA | 69.3 | trnH-GUG_psbA |
| Mint204R | GTC TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCG CCC TCT ACT ATA ATG GTA GA | 68.1 | |
| Mint771F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCA TGA GCG GCT AGC ACT TCT | 70.3 | psbA |
| Mint771R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA GGC TGA GCA CAA CAT TCT | 70.2 | |
| Mint867F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCA TGA GCG GCT AGC ACT TCT | 70.4 | psbA |
| Mint867R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA GGA ACC ATG CAT AAC ACT TCT | 69.1 | |
| Mint1170F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCA TGA GCG GCT AGC ACT TCT | 70.1 | |
| Mint1170R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA GGA ACC ATG CAT AAC ACT TCT | 69.5 | |
| Mint1939F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCA TGA GCG GCT AGC ACT TCT | 67.6 | trnK-UUU Intron1, matK |
| Mint1939R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA GGA GTG TCG GTA TTT GAC TAT GA | 68.3 | |
| Mint3998F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCA TGA GCG GCT AGC ACT TCT | 71.2 | trnK-UUU Intron2 |
| Mint3998R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA GGA TTC GAA CCG AAC | 68.6 | |
| Mint4546F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCA TGA GCG GCT AGC ACT TCT | 68.2 | trnK-UUU Exon2_rps16 Exon1 |
| Mint4546R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA TCG TGA AAT GTA TAG AGC A | 69.2 | |
| Mint4714F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCA GAT CCA TTA TCA CG | 70.1 | trnK-UUU Exon2_rps16 Exon1 |
| Mint4714R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA GGA GTG TCG GTA TTT GAC TAT GA | 67.8 | |
| Mint5533F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TCC GAT CCA GTC ATT GTG GAC TCT | 69.0 | rps16 Intron |
| Mint5533R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA GGA GCA GAT CAC TTA TCA CG | 69.7 | |
| Mint5985F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCG CCC GAG AAA TGA ATG A | 69.9 | rps16 Intron |
| Mint5985R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA TCG TGA AAT GTA TAG AGC A | 69.3 | |
| Mint6787F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TCA GCA ATT TTA TCA CG | 68.5 | rps16 Exon2_trnQ-UUG |
| Mint6787R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA GGA GCT GAG TGG GTA TCA CCA | 69.7 | |
| Mint7334F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TCA AAA ACG CAA CCA AAA TG | 68.4 | trnQ-UUG_psbK, psbK |
| Mint7334R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GTC GCA TAA CAT CCA GGA GTG TCG TTT GCT CG | 68.8 | |
| Mint8052F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG GGT TTC TCT | 69.9 | psbl-trnS-GCU, trnS-GCU |
| Mint8052R | GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GCA GGA GAT GCA GGA GTG TCG TTT GCT GA | 70.3 | |
| Mint8691F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCG AAC TCA AAA ATA AAC TGT CG | 68.7 | trnS-GCA_trnG-UCC Exon1 |
| Mint8691R | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCG AAC TCA AAA ATA AAC TGT CG | 69.7 | |
| Primer         | Sequence 5’–3’ | Tm  | Loci in region (excluding priming sites)                                                                 |
|---------------|---------------|-----|--------------------------------------------------------------------------------------------------------|
| Mint8791F     | GTC TCG GCT CGG GTA GAT TAA GAG ACA ATT AAG C   | 69.8| trnS-GCU_trnG-UCC Exon1, trnG-UCC Exon1                                                              |
| Mint8791R     | TGC TCG GCA GCG TCA GAT TAA GAG ACA ATT AAG CATTTA G  | 68.9|                                                                                                       |
| Mint10381F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG AATTATT A AAATTT TCG AAATTT TTCT TGC GAG GAT | 69.0|                                                                                                       |
| Mint10381R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.7|                                                                                                       |
| Mint10531F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 69.9|                                                                                                       |
| Mint10531R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.9|                                                                                                       |
| Mint11293F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.9|                                                                                                       |
| Mint11293R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.9|                                                                                                       |
| Mint13441F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 70.3|                                                                                                       |
| Mint13441R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.8|                                                                                                       |
| Mint13731F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 69.4|                                                                                                       |
| Mint13731R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.7|                                                                                                       |
| Mint14054F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 67.8|                                                                                                       |
| Mint14054R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 67.9|                                                                                                       |
| Mint16824F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 67.7|                                                                                                       |
| Mint16824R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.8|                                                                                                       |
| Mint17255F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 70.8|                                                                                                       |
| Mint17255R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.3|                                                                                                       |
| Mint17928F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.9|                                                                                                       |
| Mint17928R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.8|                                                                                                       |
| Mint20423F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 70.6|                                                                                                       |
| Mint20423R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 69.3|                                                                                                       |
| Mint20572F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 69.4|                                                                                                       |
| Mint20572R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 69.4|                                                                                                       |
| Mint21910F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 69.0|                                                                                                       |
| Mint21910R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 69.0|                                                                                                       |
| Mint22486F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.8|                                                                                                       |
| Mint22486R    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 68.8|                                                                                                       |
| Mint23294F    | TGT TGC GGT GCA GCG TCA GTA GAT TAA GAG ACA ATT AAG ATTTAT ATAG ATG TTT GAC ATG GAA ATG | 69.2|                                                                                                       |

Table 8 (continued)
| Primer   | Sequence 5’ – 3’                | Tm  | Loci in region (excluding priming sites) |
|----------|---------------------------------|-----|----------------------------------------|
| Mint23294R | TGC TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG AAA TTC GAC TCC GCA TGG TT | 69.7 |                                  |
| Mint23536F | TCG TGC GCA GCG TCA GAT GTG TAT AAG AGA CAG AAA TTC GAC TCC GCA TGG TT | 69.1 | rpoB                                  |
| Mint23536R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG CAC TAA ATA CAA ATT | 67.2 |                                  |
| Mint24088F | TCG TCA GCG GCA TGT GTA TAT TAA GAG AAG GAC GCT ATT CGA TAA TGG CTG | 69.3 | rpoB                                  |
| Mint24088R | GTC TCG TGG CTC GCA AGA TGT GTA TAA GAG ACA GGT GTG CAC TAA ATA CAA ATT | 69.3 |                                  |
| Mint24204R | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG AAA TTC GAC TCC GCA TGG TT | 69.9 | rpoB                                  |
| Mint24204R | TCG TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG TTA AGT TGA GGT ATT TTA | 70.0 |                                  |
| Mint27582F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CTC ATT GCT TCA GCA TGG TT | 69.1 | rpoB_trnC-GCA                         |
| Mint27582R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG CAC TAA ATA CAA ATT | 68.4 |                                  |
| Mint27884F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CTC ATT GCT TCA GCA TGG TT | 70.4 | trnC-GCA, trnC-GCA_petN, psbM         |
| Mint27884R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG CAC TAA ATA CAA ATT | 69.4 |                                  |
| Mint28672F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CTC ATT GCT TCA GCA TGG TT | 68.4 | petN_psbM, psbD                      |
| Mint28672R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG CAC TAA ATA CAA ATT | 69.9 |                                  |
| Mint28881F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CTC ATT GCT TCA GCA TGG TT | 70.0 | petN_psbM, psbD                      |
| Mint28881R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG CAC TAA ATA CAA ATT | 69.7 |                                  |
| Mint30267F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CTC ATT GCT TCA GCA TGG TT | 69.9 | trnE-UUC_trnT-GGU, psbD             |
| Mint30267R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG TTA AGT TGA GGT ATT TTA | 69.3 |                                  |
| Mint31916F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CTC ATT GCT TCA GCA TGG TT | 71.0 | psbD                                  |
| Mint31916R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG CAC TAA ATA CAA ATT | 69.5 |                                  |
| Mint32003F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CTC ATT GCT TCA GCA TGG TT | 70.9 | psbD                                  |
| Mint32003R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG CAC TAA ATA CAA ATT | 69.6 |                                  |
| Mint32348F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CTC ATT GCT TCA GCA TGG TT | 68.9 | psbD                                  |
| Mint32348R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG CAC TCA GCA TGG TT | 70.0 |                                  |
| Mint33303F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CTC ATT GCA TCA GCA TGG TT | 70.6 | psbC                                  |
| Mint33303R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG CAC TCA GCA TGG TT | 69.1 |                                  |
| Mint34416F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CTC ATT GCA TCA GCA TGG TT | 70.1 | trnS-UGA_trnS-UGA_psbZ              |
| Mint34416R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG CAC TCA GCA TGG TT | 69.6 |                                  |
| Mint35190F | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CTC ATT GCA TCA GCA TGG TT | 68.9 | trnS-GCC_trnS-GCC_psbM              |
| Mint35190R | GTC TCG GCT CGG AGA TGT GTA TAA GAG ACA GGT GTG CAC TCA GCA TGG TT | 70.1 |                                  |
| Primer      | Sequence 5′ – 3′                      | Tm   | Loci in region (excluding priming sites)          |
|------------|--------------------------------------|------|--------------------------------------------------|
| Mint35459F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG AGG ATT TGA ACC CGT GAC CT | 70.2 | trnM-CAU, trnM-CAU_rps14                         |
| Mint35459R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GAC CAA CTC TGG GGA TCA TAA AGG | 68.9 |                                                  |
| Mint36179F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG CCA AAT TGG CTT GCA AAA TG | 69.7 | psaB                                             |
| Mint36179R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GAT GGG GTG GGA TGT TCT CTT TAT | 70.0 |                                                  |
| Mint36649F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG CCA TGC CGA AGA GAA GTT AGG CTA CCT | 69.7 | psaB                                             |
| Mint36649R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GTC CCA TGC CAA AAT AGG CTT TTT CTG ACT | 70.2 |                                                  |
| Mint36907F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG CAT TGC CAT CAT | 69.9 |                                                  |
| Mint36907R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GGA GAT TAC AAC CGG GAA GAATT AGG | 70.0 |                                                  |
| Mint38366F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG CCA TGC TGA TCA ATC TTG GCC TTT AAA TTT | 69.4 |                                                  |
| Mint38366R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GAT TGA GAG GAA GGG ACT | 68.9 |                                                  |
| Mint41292F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG TCG AAA CTC TTT GGG GGA TCT ATC ATC | 69.5 | ycF3 Exon1, ycF intron1                           |
| Mint41292R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GGA GGT TTA CTA CTC TCT CCA AAG AAAT GAA | 69.9 |                                                  |
| Mint43421F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG TCG GTG GGA TTA CTA TCT CTA CTA | 69.5 | ycF3 Exon3, rnt-S-GGA                             |
| Mint43421R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GAG CCA GCC TGA ACC AAT CTA AAT CCG GAC CAC CT | 70.0 |                                                  |
| Mint47802F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG GGT GGT TGA TAC AGG AAT ATT CTA TAC AAT GAT | 69.1 | ndhK                                             |
| Mint47802R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GGT GTC CAC TTA AAC CCG GAA | 69.3 |                                                  |
| Mint50255F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG TGG GTA CCT AAA CGG GCA CT | 70.3 | trnV-UAC Intron, trnV-UAC Exon2                  |
| Mint50255R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA AGT TGG TAG AGC ACC TCG T | 70.0 |                                                  |
| Mint51642F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG TGG GGA TTA CTA TCC TCC TTA | 68.9 | atpB                                             |
| Mint51642R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GTC CAG GAA GTC TTT TGA AAG TTT GGG GGC GGA GTG | 69.2 |                                                  |
| Mint51809F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG CAA AGG ATC GGT TCA TAC AAT CTA AAT TTG GCC TTT GGT | 70.3 | rbcl                                             |
| Mint51809R | GTC TCG TGG CCG TCA GAT GTG TAT AAG AGA CAG GAT CTC CCT | 68.2 |                                                  |
| Mint52064F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG GGA AAT ATT CCG CCA TTG TTG | 69.4 | atpB                                             |
| Mint52064R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GTA TCA TCA TCA CCA CCA ATC | 69.1 |                                                  |
| Mint53793F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG GAC AAC TGG TGG GAC CCA TG | 70.3 | rbcl                                             |
| Mint53793R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GAC CTA GCC TTA | 69.3 |                                                  |
| Mint53955F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG AAA GCC CTA CGT GCT CTA TG | 70.0 | rbcl                                             |
| Mint53955R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GAG AAT TCT CTT CCA ATT CAA CAA ACC ATT | 68.7 |                                                  |
| Mint54300F | TCG TCG GCA GGG TCA GAT GTG TAT AAG AGA CAG TGT TTG AAT GGT TGG GAG CAA AGA ATC | 68.8 | rbcl                                             |
| Mint54300R | GTC TCG TGG CCG AGA TGG GTA AAG GAG ACA GAT TGG CAG TGA ATC CTC CTA | 69.3 |                                                  |
| Primer          | Sequence 5’ – 3’ | Tm  | Loci in region (excluding priming sites) |
|-----------------|------------------|-----|-----------------------------------------|
| Mint54974F      | TCG GCA CGG TCA  | 70.3| rbcL, rbcL_accD                         |
|                 | TGT AAG AGA CAG  |     |                                         |
|                 | TGG AGT CAG GTG  |     |                                         |
|                 | TAT CGT TCT GAG  |     |                                         |
| Mint54974R      | GTC TCG TGG CGG  | 68.9|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CAA TTA |     |                                         |
|                 | AAA AAT GCT GCA |     |                                         |
| Mint55144F      | TCG GCA CGG TCA  | 69.4| rbcL, accD                              |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG CGG TAA TTA |     |                                         |
|                 | AAC GAA ACC AAG |     |                                         |
| Mint55144R      | GTC TCG TGG CGG  | 69.4|                                         |
|                 | AGA TGT TGA TAA |     |                                         |
|                 | GAG ACA GTT GTG  |     |                                         |
|                 | GAT CCA AGA CAA  |     |                                         |
| Mint55263F      | TCG GCA CGG TCA  | 70.5| rbcL_accD                              |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG AGA CAG TCT  |     |                                         |
|                 | CCA TTT TTA TTC  |     |                                         |
| Mint55263R      | GTC TCG TGG CGG  | 69.5|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CAA TTA |     |                                         |
|                 | ACA GAG CTA GAA  |     |                                         |
| Mint56316F      | TCG GCA CGG TCA  | 70.3| accD                                   |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG CGG CAA CTA  |     |                                         |
|                 | CCA GAA GAA ATT |     |                                         |
|                 | GAA CCC ACA ACT  |     |                                         |
| Mint56316R      | GTC TCG TGG CGG  | 70.4|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CTA GAA  |     |                                         |
| Mint57465F      | TCG GCA CGG TCA  | 68.2| accD_psal                              |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG TTT TTA TGT |     |                                         |
|                 | AGT CAC TTA GCA  |     |                                         |
| Mint57465R      | GTC TCG TGG CGG  | 68.2|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CTA GAA  |     |                                         |
|                 | GGT GCT ACC AAG |     |                                         |
| Mint58710F      | TCG TCG GCA CGG  | 69.0| ycf4, ycf4_cemA                        |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG TGA TGA  |     |                                         |
|                 | GAA TTA ACC TAC |     |                                         |
| Mint58710R      | GTC TCG TGG CGG  | 67.7|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CTA GAA  |     |                                         |
| Mint59661F      | TCG GCA CGG TCA  | 67.6| cemA, cemA_petA                       |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG TTT TTA TTC  |     |                                         |
| Mint59661R      | GTC TCG TGG CGG  | 69.3|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CTA GAA  |     |                                         |
| Mint60529F      | TCG GCA CGG TCA  | 69.4| petA                                   |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG TTT TTA ATC  |     |                                         |
| Mint60529R      | GTC TCG TGG CGG  | 69.1|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CTA GAA  |     |                                         |
| Mint60945F      | TCG GCA CGG TCA  | 69.5| petA_psbj                              |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG TTT TTA ATT  |     |                                         |
| Mint60945R      | GTC TCG TGG CGG  | 67.2|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CTA GAA  |     |                                         |
| Mint61133F      | TCG GCA CGG TCA  | 71.0| petA_psbj                              |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG TTT TTA ACC  |     |                                         |
| Mint61133R      | GTC TCG TGG CGG  | 68.8|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CTA GAA  |     |                                         |
| Mint62265F      | TCG GCA CGG TCA  | 70.0|                                         |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG TTT TTA ACC  |     |                                         |
| Mint62265R      | GTC TCG TGG CGG  | 70.0|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
| Mint64488F      | TCG GCA CGG TCA  | 69.4| trnP-UGG, trnP-UGG_psaJ                |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG AAA ACA AAC  |     |                                         |
|                 | CCA GCA TTA CCA |     |                                         |
| Mint64488R      | GTC TCG TGG CGG  | 68.6|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CTA GCG  |     |                                         |
|                 | TGT AAT TCA GCA  |     |                                         |
| Mint65276F      | TCG GCA CGG TCA  | 68.0| psal_rpl33, rpl33                       |
|                 | TAT AAG AGA CAG |     |                                         |
|                 | CAG TTT ACT ATG  |     |                                         |
|                 | GCT TGT TTA TTA  |     |                                         |
| Mint65276R      | GTC TCG TGG CGG  | 68.8|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CCA GGG  |     |                                         |
|                 | GGG GTG GCT GAT  |     |                                         |
| Mint65764F      | TCG GCA CGG TCA  | 69.6| rps18                                   |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG CCA TTA CCC  |     |                                         |
| Mint65764R      | GTC TCG TGG CGG  | 68.8|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CCA GGG  |     |                                         |
|                 | TCG ACT CCA TCA  |     |                                         |
| Mint68295F      | TCG GCA CGG TCA  | 67.5| clpP_Intron1, clpP_Exon2               |
|                 | TGT TAT AAG AGA |     |                                         |
|                 | CAG TTT CCA AAT  |     |                                         |
| Mint68295R      | GTC TCG TGG CGG  | 69.0|                                         |
|                 | TCA GAT TTA GAA |     |                                         |
|                 | AGA CAG CCA GGA  |     |                                         |

Table 8 (continued)
## Table 8 (continued)

| Primer         | Sequence 5' – 3' | Tm  | Loci in region (excluding priming sites) |
|----------------|------------------|-----|----------------------------------------|
| Mint68800F     | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCT TTT GGT GCA TAC GGT TC | 70.1 | clpP Intron2                           |
| Mint68800R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GCC ATC GTG ATT TGG ATT GAA | 68.9 |                            |
| Mint71208F     | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG ATC GTG CGA CTT TGA AAT CC | 69.1 | psbB                                   |
| Mint71208R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GCC CAA GTT TTT GGA ATG CTC | 69.2 |                            |
| Mint71701F     | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CGA GAA CCA CTT AAA GTG GTA CAA C | 70.0 | psbT, psbT_psbN, psbN                   |
| Mint71701R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GCC GGG TAC GCC TGA TAT ACC | 69.6 |                            |
| Mint72221F     | TCG TCG GCC GCA TCA GAT GTG TAT AAG AGA CAG GCC TCC TTT TGT GCG TGC TCG GCC TGG GCC | 70.2 | psbH, psbH_petB Exon1                  |
| Mint72221R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GCC GGG CAA ATT GTA GTT CTA | 69.5 |                            |
| Mint73489F     | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CGA GAA CCA TTT GGA ATT GGT GTT | 70.3 | petB Exon2                             |
| Mint73489R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GCC AGT CAA GGT GGA TGG TCC | 70.0 |                            |
| Mint74004F     | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TTA TGG GAG TGT GCC ACT TG | 69.6 | petD Intron                           |
| Mint74004R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GCC GAG GAG CTT ACT CAT GTA CAA C | 69.5 |                            |
| Mint74125F     | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG AAG ATG GCC TGG GCC TGG AC | 70.6 | petD Intron                           |
| Mint74125R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GAA TGG GCC GCC TAA GGA AAT TCC | 68.3 |                            |
| Mint75956F     | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CAT AAG GCC GTA TAC CAA C | 69.9 | rpoA                                   |
| Mint75956R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GTG AGA ATG TCC GCC ATG AAT | 69.4 |                            |
| Mint76974F     | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TCT GCG GAT TAG TCG ACA TTT | 69.3 | rpl36                                  |
| Mint76974R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GCC GAA ACA GGT GAG ATG CAA C | 68.9 |                            |
| Mint79374F     | TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG ATT GCT TCT CGG TTC ATT TC | 68.6 | rpl16 Intron                          |
| Mint79374R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GCC ACG ACA TGG CTC GCG TCC | 69.4 |                            |
| Mint79823F     | TCG TCG GCC GCA GCC TCA GAT GTG TAT AAG AGA CAG AAT AGT AAC CGA TGG GGT GG | 69.0 | rpl16 Intron, rpl16 Exon2, rpl16 Exon2_rps3 |
| Mint79823R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GTC CGG AAC AAT CAA C | 68.6 |                            |
| Mint80282F     | TCG TCG GCC GGC TCA GAT GTG TAT AAG AGA CAG TGC CTG TCC CAA ACA ATT TCC | 68.6 | rps3                                  |
| Mint80282R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GCC CGA GGA ATC GAA GTA TAA | 69.1 |                            |
| Mint80833F     | TCG TCG GCC GCA GCC TCA GAT GTG TAT AAG AGA CAG TAG CCC GGG GTT TTA ATT TC | 69.2 | rpl22                                  |
| Mint80833R     | GTC TCG TGG GCT CGG AGA TGT TAA GAG ACA GCC GCT CTT ACT AGG ACG AAA CAC | 69.9 |                            |
| Mint10608BF    | TCG TCG GCC GCA GCC TCA GAT GTG TAT AAG AGA CAG CGG ACC CGT TTC TGA AGA GT | 70.1 | ycf1Truncated, ycf1-Truncated_ndhF, ndhF |
| Mint10608BR    | GTC TCG TGG GCC TCA GAT GTG TAA GAG ACA GAT TTT GGA CAA GGG ACT AAA GAG G | 68.1 |                            |
| Mint107075F    | TCG TCG GCC GCA GCC TCA GAT GTG TAT AAG AGA CAG CCC ATC GTG TTC TTT TAG | 69.4 | ndhF                                   |
| Primer     | Sequence 5’ – 3’                                                                 | Tm  | Loci in region (excluding priming sites) |
|------------|----------------------------------------------------------------------------------|-----|------------------------------------------|
| Mint107075R| GTC TCG GCT CGG AGA TGT GTA TAA GAC ACA GGG ATT CGG CAA GGT GAT ATG               | 69.4|                                           |
| Mint107435F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG TAT TGG GCG ATT CAT AA               | 68.8| ndhF                                     |
| Mint107435R| GTC TCG GCT CGG AGA TGT GTA TAA GAC ACA GGG GGT AAA GGG TAT TCC AAA A           | 69.1|                                           |
| Mint107714F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG TGA ATG TTT AAA TGC TTT TGT ACT    | 69.0| ndhF                                     |
| Mint107714R| GTC TCG GCT CGG AGA TGT GTA TAA GAC ACA GGT GTA CAT CAT TCT CGG TTT GTG GCA    | 69.1|                                           |
| Mint108575F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG CGC TTT TTG ACA AGC GGA ACT        | 69.1| ndhF                                     |
| Mint108575R| GTC TCG GCT CGG AGA TGT GTA TAA GAC ACA GGT GCT CAC GAC ATT AGT                 | 68.8| rpl32, rpl32_trnL-UAG                   |
| Mint109649F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG GTA TCG GCC AGC GGT AAA AGT        | 68.8|                                           |
| Mint109649R| GTC TCG GCT CGG AGA TGT GTA TAA GAC ACA GAT TCC CGG TTT AAG AAG ATG             | 70.4| rpl32_trnL-UAG                          |
| Mint110066F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG TCT CGC TAT CAA TCC AAA              | 69.2|                                           |
| Mint110066R| GTC TCG TGG GCT CGG AGA TGT GTA TAA GAC GTA GAA CCC TCC CTC CCC AAA             | 70.4| trnL-UAG, trnL-UAG_ccsA, ccsA            |
| Mint110450F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG GGT AGA CAC GCT GCT CTT AGG        | 70.4|                                           |
| Mint110450R| GTC TCG TGG GCT CGG AGA TGT GTA TAA GAC GTC GAA ACG TCG GCA AAA GGA             | 67.9|                                           |
| Mint111053F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG GTA TCG GCC ATC GAA AAA GAA         | 70.0| ccsA                                     |
| Mint111053R| GTC TCG TGG GCT CGG AGA TGT GTA TAA GAC GTA GGT GAC ATT CAC GGA TCA TCT TTG    | 68.3|                                           |
| Mint113090F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG GGA TCA TCC GAT TGA AAA TGA        | 68.8| ndhD                                     |
| Mint113090R| GTC TCG TGG GCT CGG AGA TGT GTA TAA GAC ACA GAC CCA TTT TCG TTG GCT             | 68.9|                                           |
| Mint113960F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG TTT AAC AAA A                    | 69.7| ndhE                                     |
| Mint113960R| GTC TCG TGG GCT CGG AGA TGT GTA TAA GAC ACA CCT ATT TAT TTT CCA TTA GTG        | 67.9|                                           |
| Mint116500F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG CCG TTA CGG TCG CTA TAA CAG       | 69.9| ndhA Intron                              |
| Mint116500R| GTC TCG TGG GCT CGG AGA TGT GTA TAA GAC ACA GGC AGA CAG AAT TCC ATT GCT C      | 69.2|                                           |
| Mint116861F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG TTT CAA TCT CCG TTT TTG GA         | 68.7| ndhA Intron                              |
| Mint116861R| GTC TCG TGG GCT CGG AGA TGT GTA TAA GAC ACA GGA ATT GGG GCT TAA AGT GTG TTT    | 68.4|                                           |
| Mint116914F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG TGG ATA GGA ACA TAC TCT GG        | 69.3| ndhA Intron                              |
| Mint116914R| GTC TCG TGG GCT CGG AGA TGT GTA TAA GAC ACA GGT GAT GGT TAG GAA GAC CAA A     | 69.0|                                           |
| Mint117269F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG CAA ATT TTT AAG CGG TTT T        | 68.9| ndhA Intron                              |
| Mint117269R| GTC TCG TGG GCT CGG AGA TGT GTA TAA GAC ACA GCG TGT CAT GTC GCA AGA           | 69.9|                                           |
| Mint117433F| TCG TCG GCA CGG TCA GAT GTG TAT AAG AGA CAG TGT CTC ACC GAA TCA CAC GTA       | 69.7| ndhA Exon2                                |
| Mint117433R| GTC TCG TGG GCT CGG AGA TGT GTA TAA GAC ACA GGG ATC TTA GTA TTG GTG T          | 69.6|                                           |
The partial genomes from all of the historical samples, except *Phyllostegia variabilis*, were mapped to the *Stenogyne haliakalae* reference sequence using BWA. SNPs were called with SAMtools and BCFtools, filtering out those with SNP quality < 30. We selected a total of 108 variable loci (see Fig. 2 in [1] for a diagram of locations) identified from single copy regions, including (1) all the regions that had a variant position among the Hawaiian mints (except where every individual had an alternative allele as compared to the reference sequence) and (2) additional regions that had variant positions among at least two of the *Stachys* species. 100 bp of flanking sequence on either side of the SNPs was retrieved from the reference genome, and PCR primers were designed using BatchPrimer3 [19], with further manual examination for quality control (e.g. to ensure that primer sequences did not fall into a gap for one of the other taxa). Sequences complementary to the Illumina sequencing adapters were appended to the end of each primer (Table 8) so that sequencing libraries could be prepared directly from the cleaned multiplex PCR products. Overall, these regions represent roughly 20,000 bp of sequence from the chloroplast genome and contain additional variable sites beyond the initial targeted SNP.

**Acknowledgments**

We thank the Bernice P. Bishop Museum, The National Tropical Botanic Garden, and the Volcano Rare Plant Facility for access to samples. Webb Miller provided helpful advice on data analysis. We also thank the Center for Computational Resources at the University of Buffalo for access to computational resources. This work was supported by the College of Arts and Sciences, University at Buffalo.

### Table 8 (continued)

| Primer          | Sequence 5' – 3' | Tm  | Loci in region (excluding priming sites) |
|-----------------|------------------|-----|----------------------------------------|
| Mint119376F     | TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TTT TGA CAA ATA GGC CAG CA | 69.3 | rps15                                  |
| Mint119376R     | GTC TCG TGG CTT CAG AAG TAA GAG ACA GTG TTA CAT GTA AAC TTT TGT TTT C | 68.0 |                                       |
| Mint122192F     | TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TTT TTC GTG TTT TTT TT | 67.9 | ycf1 single copy                       |
| Mint122192R     | GTC TCG TGG CTT CAG AAG TAA GAG ACA GTG TTA CAT GTA AAC TTT TGT TTT C | 70.0 |                                       |
| Mint122492F     | TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG AAA AAC ACT TTT GAG AAG AAG GA | 68.2 | ycf1 single copy                       |
| Mint122492R     | GTC TCG TGG CTT CAG AAG TAA GAG ACA GTG TTA CAT GTA AAC TTT TGT TTT C | 69.4 |                                       |
| Mint122638F     | TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TTT TTC GTC TCA AAC AAA AG | 68.8 | ycf1 single copy                       |
| Mint122638R     | GTC TCG TGG CTT CAG AAG TAA GAG ACA GTG TTA CAT GTA AAC TTT TGT TTT C | 69.1 |                                       |
| Mint123159F     | TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TTT GTT GAA ATT GAT GAA ATG | 69.2 | ycf1 single copy                       |
| Mint123159R     | GTC TCG TGG CTT CAG AAG TAA GAG ACA GTG TTA CAT GTA AAC TTT TGT TTT C | 69.4 |                                       |
| Mint124145F     | TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TTT TGC TCA GAT GGA ATT GGA TTA CTT C | 69.1 | ycf1 single copy                       |
| Mint124145R     | GTC TCG TGG CTT CAG AAG TAA GAG ACA GTG TTA CAT GTA AAC TTT TGT TTT C | 69.3 |                                       |
Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2016.03.037.

References

[1] A.J. Welch, K. Collins, A. Ratan, D.I. Drautz-Moses, S.C. Schuster, C. Lindqvist, The quest to resolve recent radiations: Plastid phylogenomics of extinct and endangered Hawaiian endemic mints (Lamiaceae), Mol. Phylogen. Evol. 99 (2016) 16–33.
[2] S. Lindgreen, AdapterRemoval: easy cleaning of next-generation sequencing reads, BMC Res. Notes 5 (2012) 337.
[3] H. Jónsson, A. Ginolhac, M. Schubert, P.L.F. Johnson, L. Orlando, mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters, Bioinformatics 13 (2013) 1682–1684.
[4] R. Cronn, A. Liston, M. Parks, D.S. Gernandt, R. Shen, T. Mockler, Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology, Nucleic Acids Res. 36 (2008) e122.
[5] A. Ratan, Pennsylvania State University, University Park, PA, 2009.
[6] R. Mariotti, N.G.M. Culterra, C.M. Díez, L. Baldoni, A. Rubini, Identification of new polymorphic regions and differentiation of cultivated olives (*Olea europaea L.*) through plastome sequence comparison, BMC Plant Biol. 10 (2010) 211.
[7] R. Li, H. Zhu, J. Ruan, W. Qian, X. Fang, Z. Shi, Y. Li, S. Li, G. Shan, K. Kristiansen, S. Li, H. Yang, J. Wang, J. Wang, De novo assembly of human genomes with massively parallel short read sequencing, Genome Res. 20 (2010) 265–272.
[8] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics 25 (2009) 1754–1760.
[9] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, & 1000 Genome Project Data Processing Subgroup, The sequence alignment/map (SAM) format and SAMtools, Bioinformatics 25 (2009) 2078–2079.
[10] D.-K. Yi, K.-J. Kim, Complete chloroplast genome sequences of important oilseed crop *Sesamum indicum L.*, PLoS One 7 (2012) e35872.
[11] B. Lukas, J. Novak, The complete chloroplast genome of *Origanum vulgare L.* (Lamiaceae), Gene 528 (2013) 163–169.
[12] J. Qian, J. Song, H. Gao, Y. Zhu, J. Xu, X. Pang, H. Yao, C. Sun, X. Li, C. Li, J. Liu, H. Xu, S. Chen, The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*, PLoS One 8 (2013) e57607.
[13] S.K. Wyman, R.K. Jansen, J.L. Boore, Automatic annotation of organellar genomes with DOGMA, Bioinformatics 20 (2004) 3252–3255.
[14] P. Schattner, A.N. Brooks, T.M. Lowe, The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snRNAs, Nucleic Acids Res. 33 (2005) W686–W689.
[15] P.E. Warburton, J. Giordano, F. Cheung, Y. Gelfand, G. Benson, Inverted repeat structure of the human genome: The X-chromosome contains a preponderance of large, highly homologous inverted repeats that contain testes genes, Genome Res. 14 (2004) 1861–1869.
[16] A.E. Darling, B. Mau, N.T. Perna, progressiveMauve: multiple genome alignment with gene gain, loss, and rearrangement, PLoS One 5 (2010) e11147.
[17] J. Shaw, E.B. Lickey, J.T. Beck, S.B. Farmer, W.B. Liu, J. Miller, K.C. Siripun, C.T. Winder, E.E. Schilling, R.L. Small, The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis, Am. J. Bot 92 (2005) 142–166.
[18] J. Shaw, E.B. Lickey, E.E. Schilling, R.L. Small, Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III, Am. J. Bot. 94 (2007) 275–288.
[19] F.M. You, N. Huo, Y.Q, Gu, M.-C. Luo, Y. Ma, D. Hane, G.R. Lazo, J. Dvorak, O.D. Anderson, BatchPrimer3: a high throughput web application for PCR and sequencing primer design, BMC Bioinf. 9 (2008) 253.
[20] E. Boudreau, Y. Takahashi, C. Lemieux, M. Turmel, J.-D. Rochais, The plastome ycf3 and ycf4 open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem 1 complex, EMBO J. 16 (1997) 6095–6104.