rpl16

81189　　　　　　　　　　　　　　　　　81582644　82652

A. The distance between exons is too short

petD

766888　　　　　　　　　　　　　　　77198　　77672
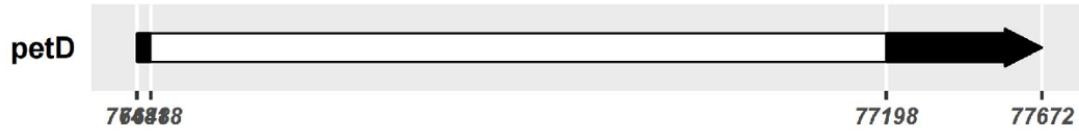
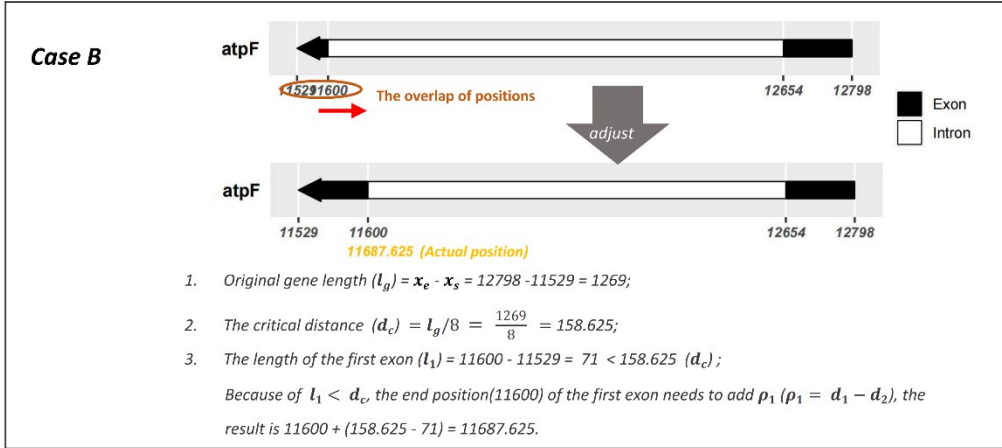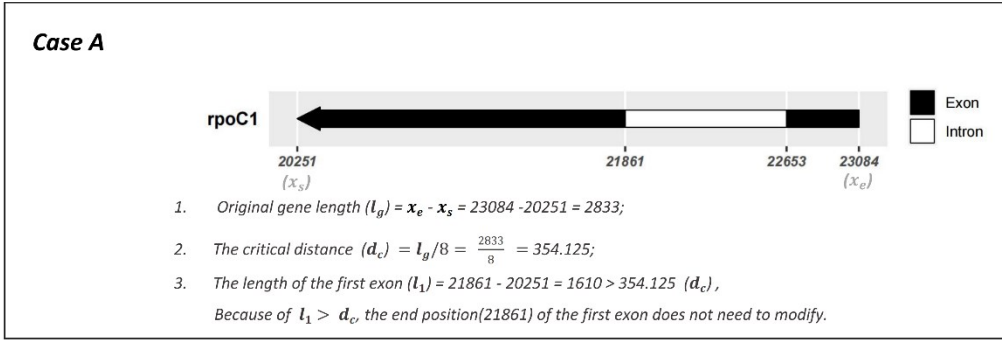**B. The distance between the start and end position of exon is too short**

1

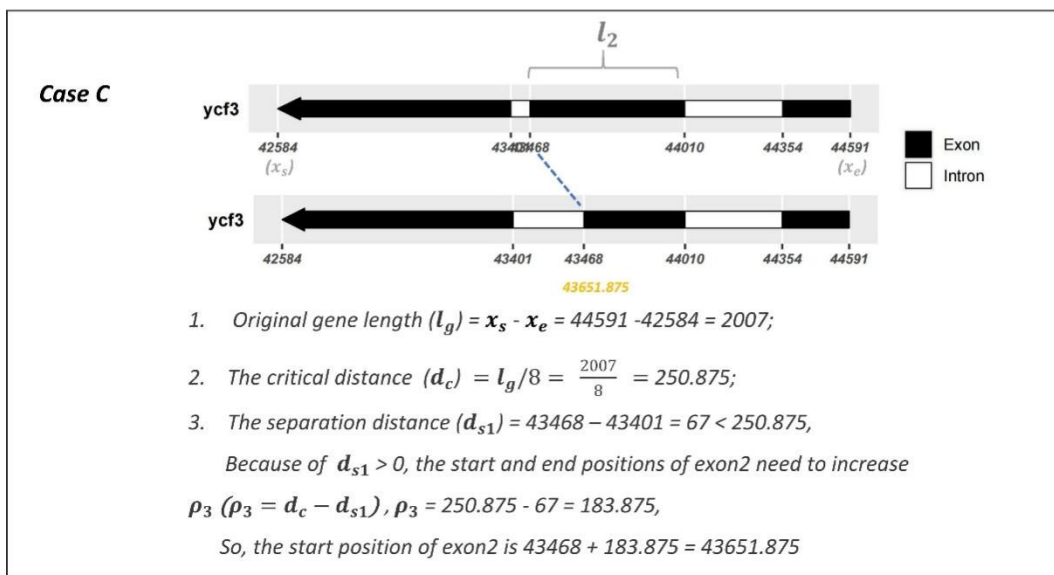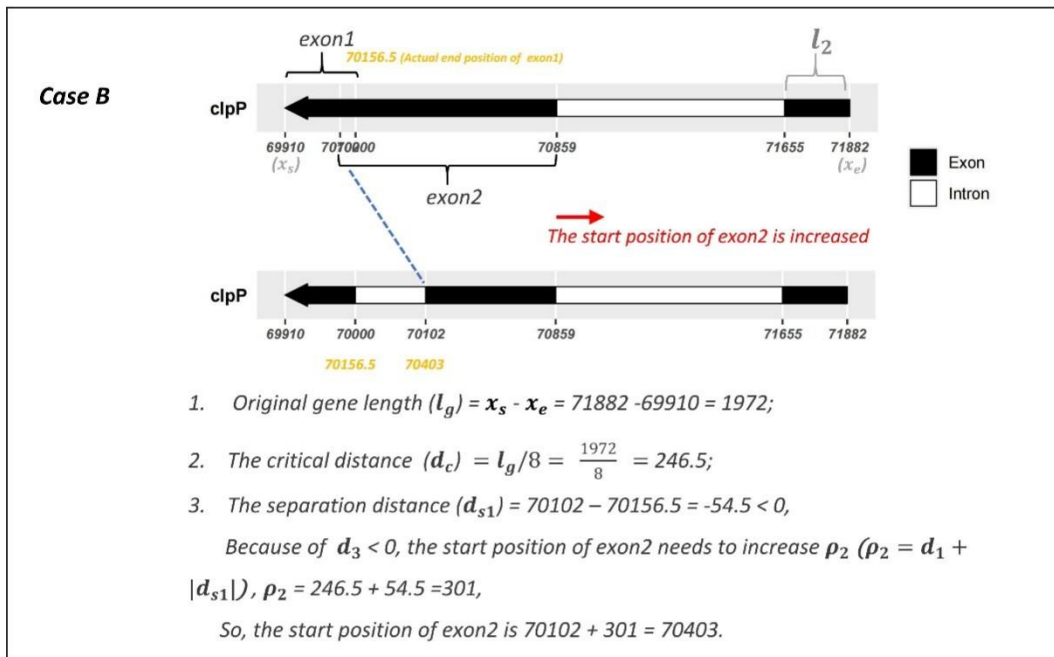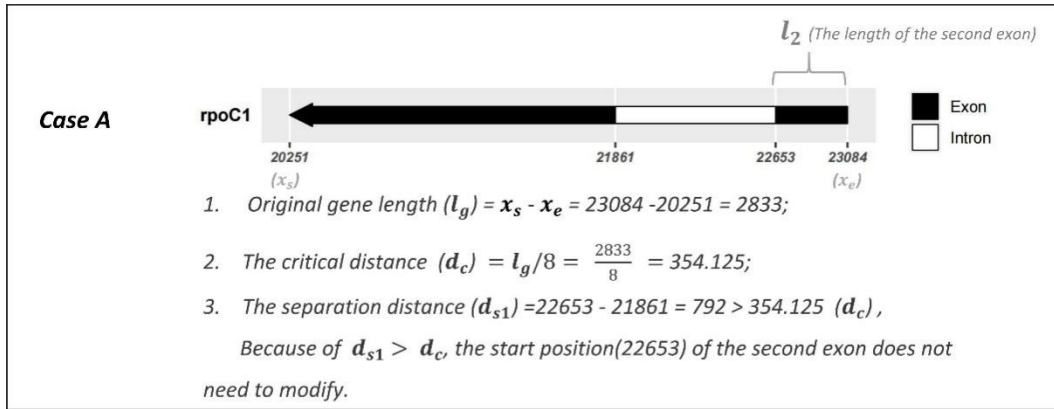2　**Supplementary Figure S1. Typical problems in the display of the structures of**
3　**the cis-splicing genes.** (A) The size of the intron is too small. The end position of the
4　upstream exon overlaps with that of the start position of the downstream exon. (B)
5　The size of the exon is too small. The positions of the start and end positions of the
6　exons overlap. Where they are too many exons, the problem further exacerbates.
7

Case A

rpoC1

Exon
Intron

20251 (x_s)   21861   22653   23084 (x_e)

1. Original gene length (l_g) = x_e - x_s = 23084 - 20251 = 2833;

2. The critical distance (d_c) = $l_g/8 = \frac{2833}{8}$ = 354.125;

3. The length of the first exon (l_1) = 21861 - 20251 = 1610 > 354.125 (d_c),

   Because of $l_1 > d_c$, the end position(21861) of the first exon does not need to modify.

Case B

atpF

11529 11600   The overlap of positions   12654   12798

adjust

Exon
Intron

atpF

11529   11600   12654   12798
11687.625 *(Actual position)*

1. Original gene length (l_g) = x_e - x_s = 12798 - 11529 = 1269;

2. The critical distance (d_c) = $l_g/8 = \frac{1269}{8}$ = 158.625;

3. The length of the first exon (l_1) = 11600 - 11529 = 71 < 158.625 (d_c) ;

   Because of $l_1 < d_c$, the end position(11600) of the first exon needs to add ρ_1 ($\rho_1 = d_1 - d_2$), the result is 11600 + (158.625 - 71) = 11687.625.

Supplementary Figure S2. The schematic depiction of the steps 1 and 2 in the CSA algorithm to modify the end position of the exon 1.

(A) the length of the exon 1 is > d_c and no change is needed. (B) the length of the exon 1 is < d_c. The size of the exon is scaled.

Case A

l_2 *(The length of the second exon)*

rpoC1

20251 (x_s) 21861 22653 23084 (x_e)

Exon / Intron

1. Original gene length $(l_g) = x_s - x_e = 23084 - 20251 = 2833$;

2. The critical distance $(d_c) = l_g/8 = \frac{2833}{8} = 354.125$;

3. The separation distance $(d_{s1}) = 22653 - 21861 = 792 > 354.125\ (d_c)$,

   Because of $d_{s1} > d_c$, the start position(22653) of the second exon does not need to modify.

Case B

exon1 70156.5 *(Actual end position of exon1)* l_2

clpP

69910 (x_s) 70000 70100 70859 71655 71882 (x_e)

exon2

The start position of exon2 is increased

clpP

69910 70000 70102 70859 71655 71882

70156.5 70403

Exon / Intron

1. Original gene length $(l_g) = x_s - x_e = 71882 - 69910 = 1972$;

2. The critical distance $(d_c) = l_g/8 = \frac{1972}{8} = 246.5$;

3. The separation distance $(d_{s1}) = 70102 - 70156.5 = -54.5 < 0$,

   Because of $d_3 < 0$, the start position of exon2 needs to increase ρ_2 $(\rho_2 = d_1 + |d_{s1}|)$, $\rho_2 = 246.5 + 54.5 = 301$,

   So, the start position of exon2 is $70102 + 301 = 70403$.

Case C

l_2

ycf3

42584 (x_s) 43401 43468 44010 44354 44591 (x_e)

Exon / Intron

ycf3

42584 43401 43468 44010 44354 44591

43651.875

1. Original gene length $(l_g) = x_s - x_e = 44591 - 42584 = 2007$;

2. The critical distance $(d_c) = l_g/8 = \frac{2007}{8} = 250.875$;

3. The separation distance $(d_{s1}) = 43468 - 43401 = 67 < 250.875$,

   Because of $d_{s1} > 0$, the start and end positions of exon2 need to increase ρ_3 $(\rho_3 = d_c - d_{s1})$, $\rho_3 = 250.875 - 67 = 183.875$,

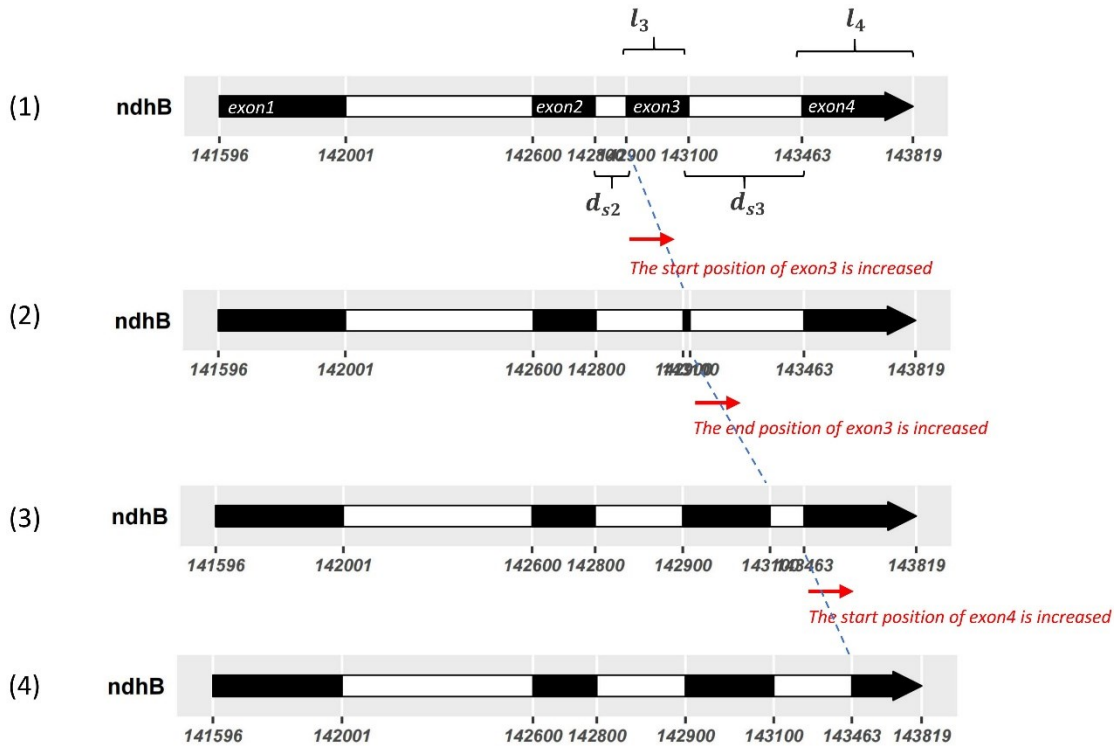   So, the start position of exon2 is $43468 + 183.875 = 43651.875$

14

15 **Supplementary Figure S3. The schematic depiction of the steps 4 in the CSA**

16 **algorithm to adjust the start and end positions of exon 2.**

17 In case A, no adjustment of the start and end positions of exon 2 is needed. In case B

18 and C, only the start position of exon 2 is needed.

19

20

Supplementary Figure S4. The schematic depiction of steps 5 in the CSA algorithm to adjust the start and end positions of the exon 3 and those after for the *ndh*B gene. (1) After the scaling of exon 2, the positions of exon 2 and exon 3 overlap. (2) Adjust the start position of exon 3, resulting the overlap of the start and end positions of exon 3. (3) Adjust the end position of exon 3, resulting in the overlap of positions for exons 3 and 4. (4) Adjust the start position of exon 4, resulting in well-separated positions.

(A) Bird's-eye view of exon1/intron1 boundary

(B) Read mapping result of exon1/intron1 boundary

(C) Bird's-eye view of intron1/exon2 boundary

(D) Read mapping result of intron1/exon2 boundary

(E) Bird's-eye view of exon2/intron2 boundary
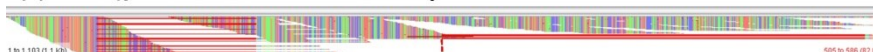
(F) Read mapping result of exon2/intron2 boundary

(G) Bird's-eye view of intron2/exon3 boundary

(H) Read mapping result of intron2/exon3 boundary

30

6
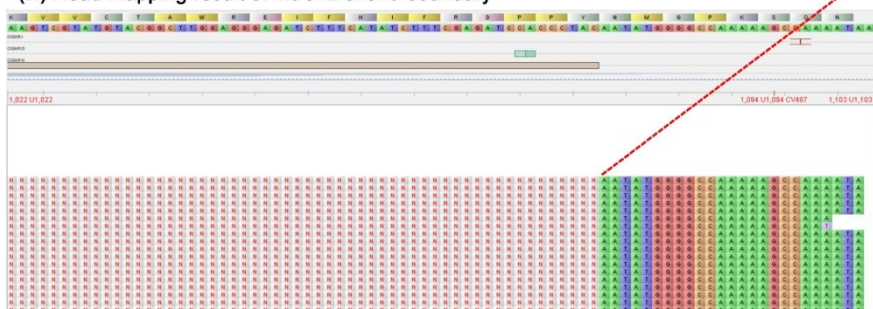
Supplementary Figure S5. The mapping results of RNA-seq reads to the exon/intron/exon boundaries of the *rps*12 gene on the *Glycine max* chloroplast genome. These results exemplify a three-exon model of the *rps*12 gene. (A) A bird's-eye view of the read mapping result at the exon1/intron1 boundary. (B) A base-level view of the read mapping result at the exon1/intron1 boundary. (C) A bird's-eye view of the read mapping result at the intron1/exon2 boundary. (D) A base-level view of the read mapping result at the intron1/exon2 boundary. (E) A bird's-eye view of the read mapping result at the exon2/intron2 boundary. (F) A base-level view of the read mapping result at the exon2/intron2 boundary. (G) A bird's-eye view of the read mapping result at the intron2/exon3 boundary. (H) A base-level view of the read mapping result at the intron2/exon3 boundary. The exon/intron/exon junctions are indicated with read lines.

(A) Bird's-eye view of exon1/intron1 boundary

(B) Read mapping result of exon1/intron1 boundary

(C) Bird's-eye view of intron1/exon2 boundary

(D) Read mapping result of intron1/exon2 boundary
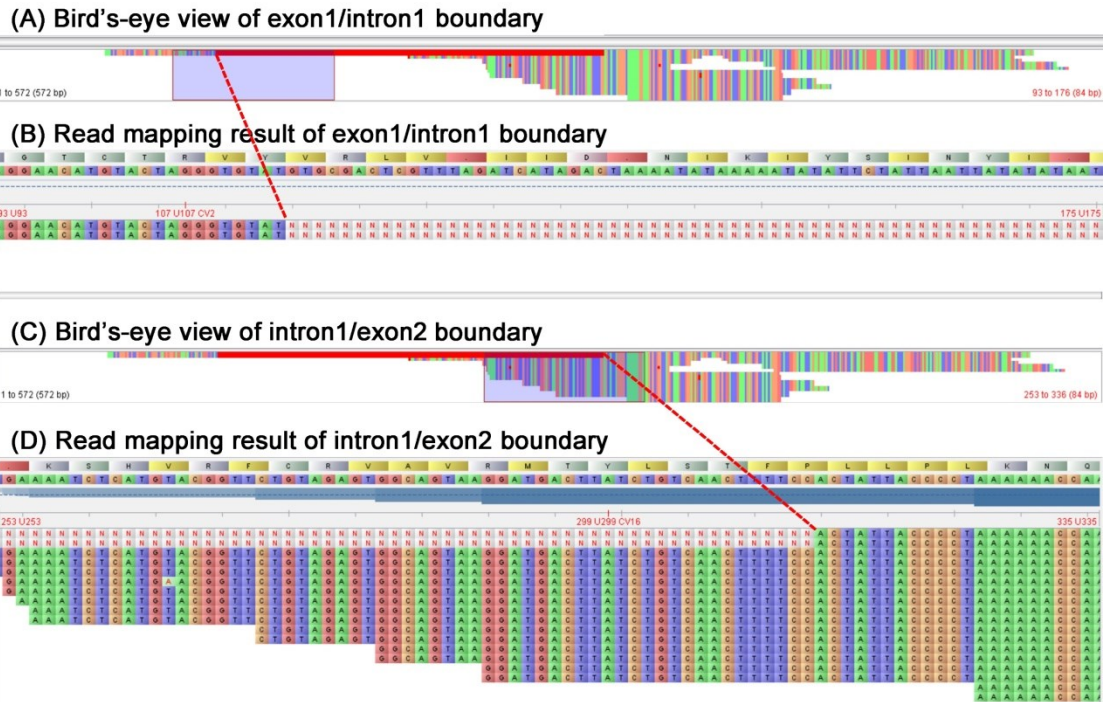
Supplementary Figure S6. The mapping results of RNA-seq reads to the exon/intron and intron/exon boundaries of the *rps*12 gene on the *Cicer arietinum* chloroplast genome. These results exemplify a two-exon model of the *rps*12 gene. (A) A bird's-eye view of the read mapping result at the exon1/intron1 boundary. (B) A base-level view of the read mapping result at the exon1/intron1 boundary. (C) A bird's-eye view of the read mapping result at the intron1/exon2 boundary. (D) A base-level view of the read mapping result at the intron1/exon2 boundary. The exon/intron/exon junctions are indicated with read lines.

Note:

Due to the large size of the supplementary figure S7-S37 (~300 MB), we could not upload them to the journal's server after trials for days. As a result, we uploaded the supplementary figures to our server. They can be accessed from the following website: http://www.1kmpg.cn/cpgview/sfigures/index.html.

We apologize for any inconvenience this might cause, and we greatly appreciate your understanding.

Supplementary Figure S7. The gene maps for the chloroplast genome of *Agrimonia pilosa* generated by CPGView.
(A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4, 5, and 6.

Supplementary Figure S8. The gene maps for the chloroplast genome of *Fagopyrum dibotrys* generated by CPGView.
(A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4, 5, and 6.

Supplementary Figure S9. The gene maps for the chloroplast genome of *Menispermum dauricu* generated by CPGView.
(A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4, 5, and 6.

Supplementary Figure S10. The gene maps for the chloroplast genome of *Magnolia biondii* generated by CPGView.
(A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4, 5, and 6.

Supplementary Figure S11. The gene maps for the chloroplast genome of *Sanguisorba officinalis* generated by CPGView.
(A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4, 5, and 6.

Supplementary Figure S12. The gene maps for the chloroplast genome of *Apocynum venetum* generated by CPGView.
(A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4, 5, and 6.

93 **Supplementary Figure S13. The gene maps for the chloroplast genome of**
94 *Ligusticum sinense* **generated by CPGView.**
95 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
96 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
97 5, and 6.

98 **Supplementary Figure S14. The gene maps for the chloroplast genome of** *Rosa*
99 *chinensis* **generated by CPGView.**
100 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
101 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
102 5, and 6.

103 **Supplementary Figure S15. The gene maps for the chloroplast genome of** *Aster*
104 *tataricus* **generated by CPGView.**
105 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
106 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
107 5, and 6.

108 **Supplementary Figure S16. The gene maps for the chloroplast genome of**
109 *Dianthus chinensis* **generated by CPGView.**
110 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
111 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
112 5, and 6.

113 **Supplementary Figure S17. The gene maps for the chloroplast genome of**
114 *Prunus japonica* **generated by CPGView.**
115 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
116 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
117 5, and 6.

118 **Supplementary Figure S18. The gene maps for the chloroplast genome of**
119 *Prunella vulgaris* **generated by CPGView.**
120 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
121 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
122 5, and 6.

123 **Supplementary Figure S19. The gene maps for the chloroplast genome of**
124 *Saururus chinensis* **generated by CPGView.**
125 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
126 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
127 5, and 6.

128 **Supplementary Figure S20. The gene maps for the chloroplast genome of**
129 *Achyranthes bidentata* **generated by CPGView.**

130   (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing

131   gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,

132   5, and 6.

133   **Supplementary Figure S21. The gene maps for the chloroplast genome of *Salvia***

134   ***officinalis* generated by CPGView.**

135   (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing

136   gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,

137   5, and 6.

138   **Supplementary Figure S22. The gene maps for the chloroplast genome of**

139   ***Celosia argentea* generated by CPGView.**

140   (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing

141   gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,

142   5, and 6.

143   **Supplementary Figure S23. The gene maps for the chloroplast genome of**

144   ***Euphorbia lathyris* generated by CPGView.**

145   (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing

146   gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,

147   5, and 6.

148   **Supplementary Figure S24. The gene maps for the chloroplast genome of**

149   ***Paeonia lactiflora* generated by CPGView.**

150   (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing

151   gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,

152   5, and 6.

153   **Supplementary Figure S25. The gene maps for the chloroplast genome of**

154   ***Schizonepeta tenuifolia* generated by CPGView.**

155   (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing

156   gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,

157   5, and 6.

158   **Supplementary Figure S26. The gene maps for the chloroplast genome of**

159   ***Gleditsia sinensis* generated by CPGView.**

160   (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing

161   gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,

162   5, and 6.

163   **Supplementary Figure S27. The gene maps for the chloroplast genome of**

164   ***Houttuynia cordata* generated by CPGView.**

165   (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing

166   gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,

167 5, and 6.

168 **Supplementary Figure S28. The gene maps for the chloroplast genome of**
169 *Xanthium spinosum* **generated by CPGView.**
170 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
171 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
172 5, and 6.

173 **Supplementary Figure S29. The gene maps for the chloroplast genome of**
174 *Abelmoschus manihot* **generated by CPGView.**
175 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
176 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
177 5, and 6.

178 **Supplementary Figure S30. The gene maps for the chloroplast genome of**
179 *Euryale ferox* **generated by CPGView.**
180 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
181 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
182 5, and 6.

183 **Supplementary Figure S31. The gene maps for the chloroplast genome of**
184 *Trichosanthes kirilowii* **generated by CPGView.**
185 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
186 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
187 5, and 6.

188 **Supplementary Figure S32. The gene maps for the chloroplast genome of**
189 *Aristolochia tubiflora* **generated by CPGView.**
190 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
191 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
192 5, and 6.

193 **Supplementary Figure S33. The gene maps for the chloroplast genome of**
194 *Diospyros rhombifolia* **generated by CPGView.**
195 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
196 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
197 5, and 6.

198 **Supplementary Figure S34. The gene maps for the chloroplast genome of**
199 *Pachysandra terminalis* **generated by CPGView.**
200 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
201 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
202 5, and 6.

203 **Supplementary Figure S35. The gene maps for the chloroplast genome of**
204 ***Sassafras tzumu* generated by CPGView.**
205 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
206 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
207 5, and 6.

208 **Supplementary Figure S36. The gene maps for the chloroplast genome of**
209 ***Rumex acetosa* generated by CPGView.**
210 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
211 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
212 5, and 6.
213
214 **Supplementary Figure S37. The gene maps for the chloroplast genome of**
215 ***Chimonanthus praecox* generated by CPGView.**
216 (A) The cpgenome_r gene map; (B) The cis-splicing gene map; (C) The trans-splicing
217 gene map. Detailed descriptions of these maps can be found in the legends for Fig. 4,
218 5, and 6.

219

220