1   Supplementary Figures
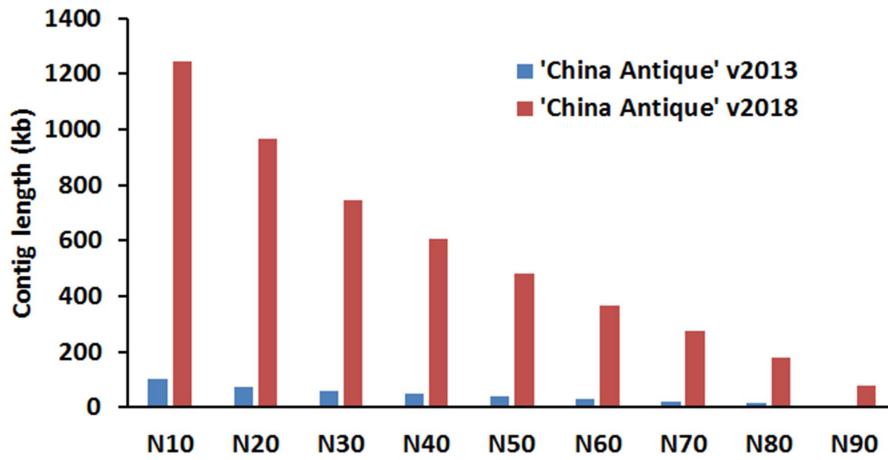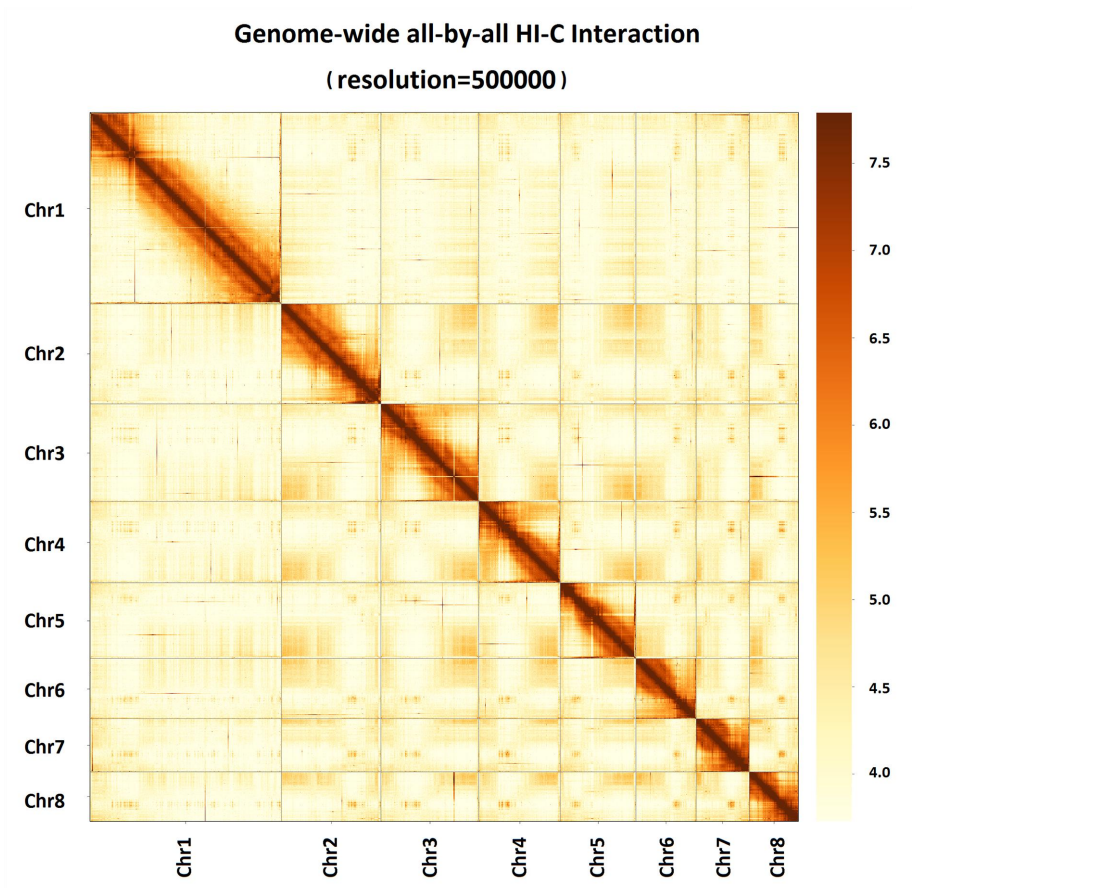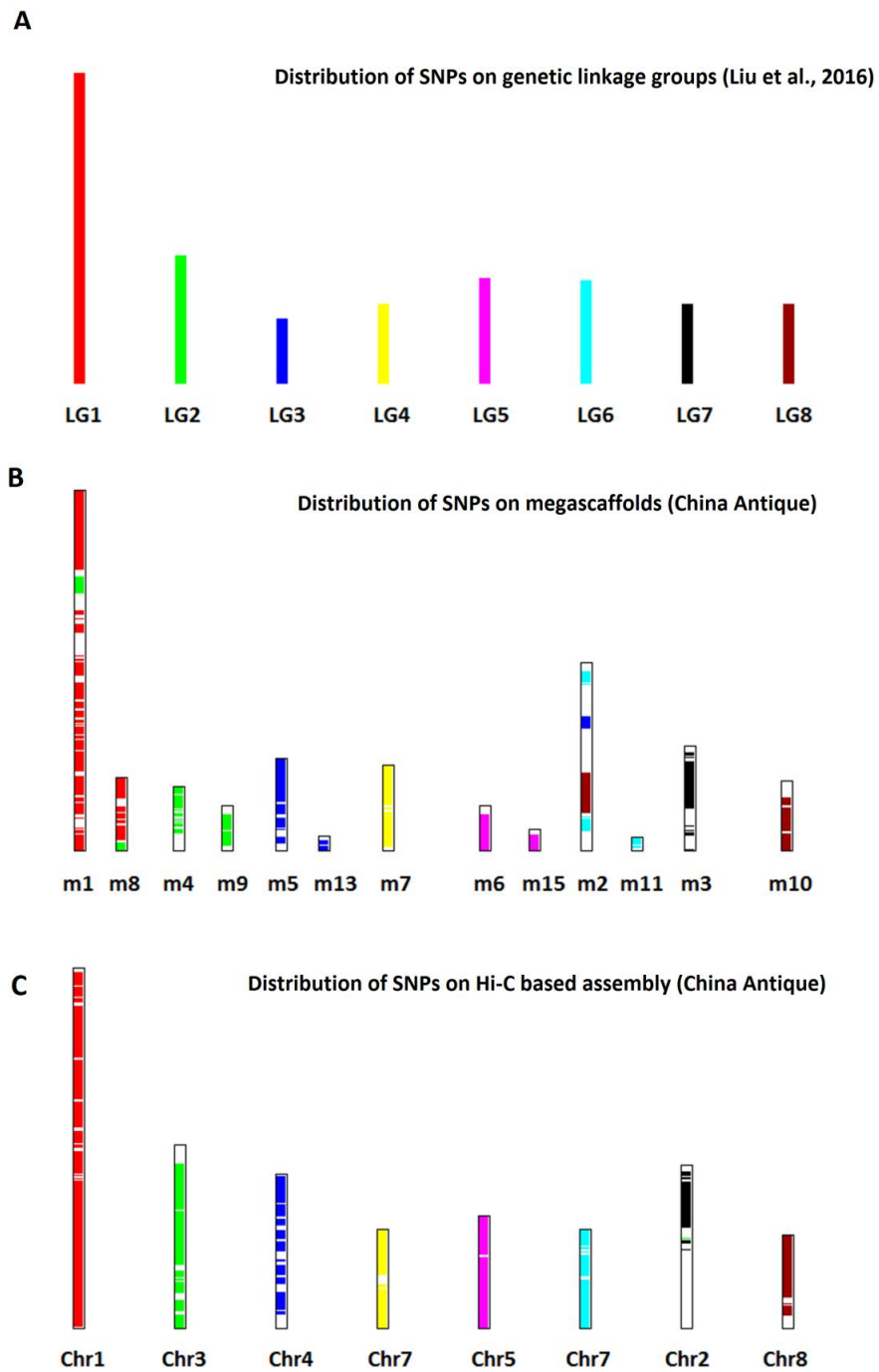

2

3   **Supplementary Figure S1. Comparison of contig N10-N90 between the old and current**

4   **version of the *Nelumbo nucifera* var. 'China Antique' assembly.**
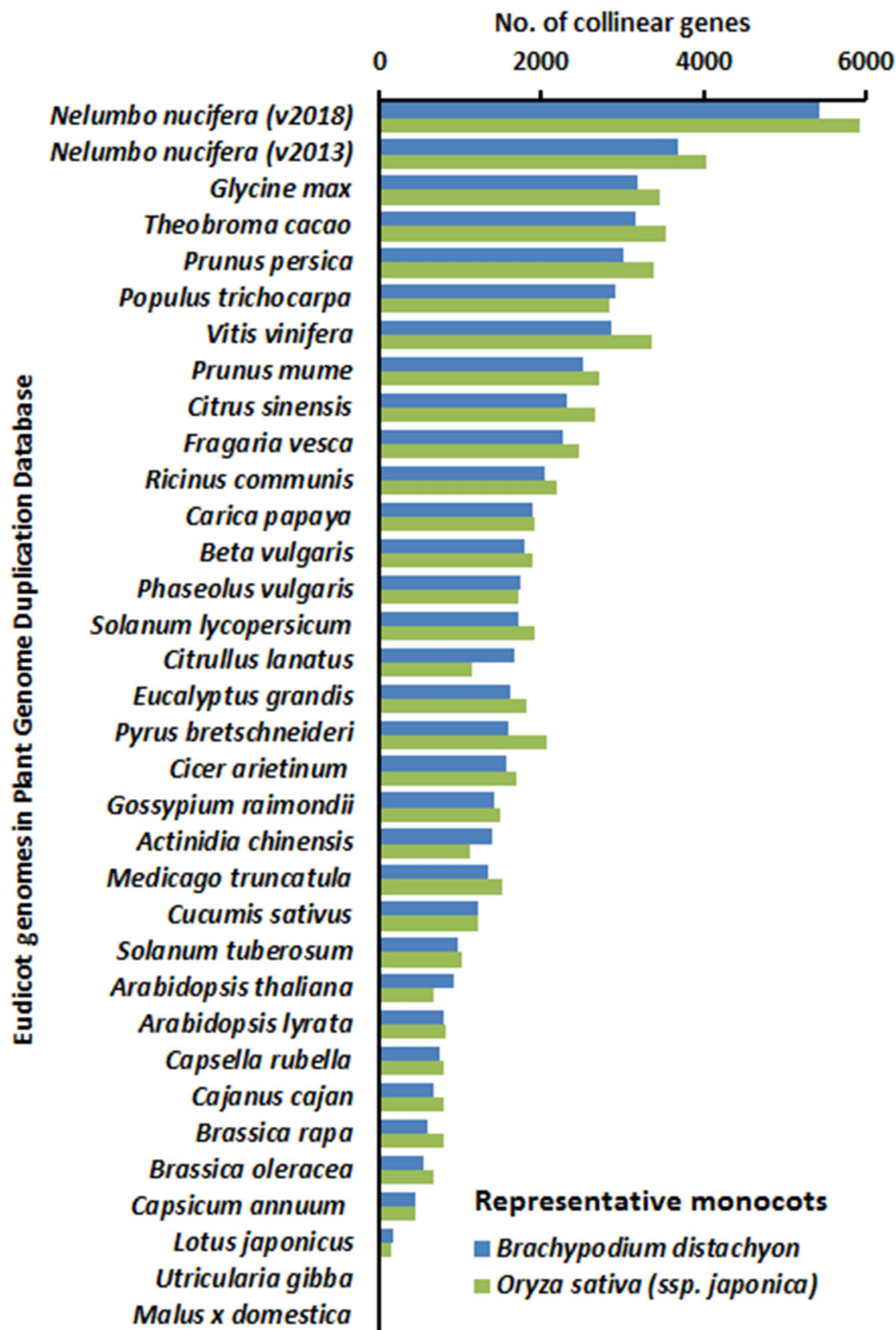

5

6   **Supplementary Figure S2. Chromatin interaction matrix for eight chromosomes from**

7   **HI-C experiment. Darker color indicates stronger chromatin interaction.**

Supplementary Figure S3. Distribution of SNP markers on the eight linkage groups (A), on the 'China Antique' assembly v2013 (B) and our current version (C).

Supplementary Figure S4. Comparison of the numbers of syntenic ortholog genes in *Brachyopodium distachyon* and *Oryza sativa* in relationship to eudicot assemblies from Plant Genome Duplication Database including two lotus assemblies (v2013 and the current version).
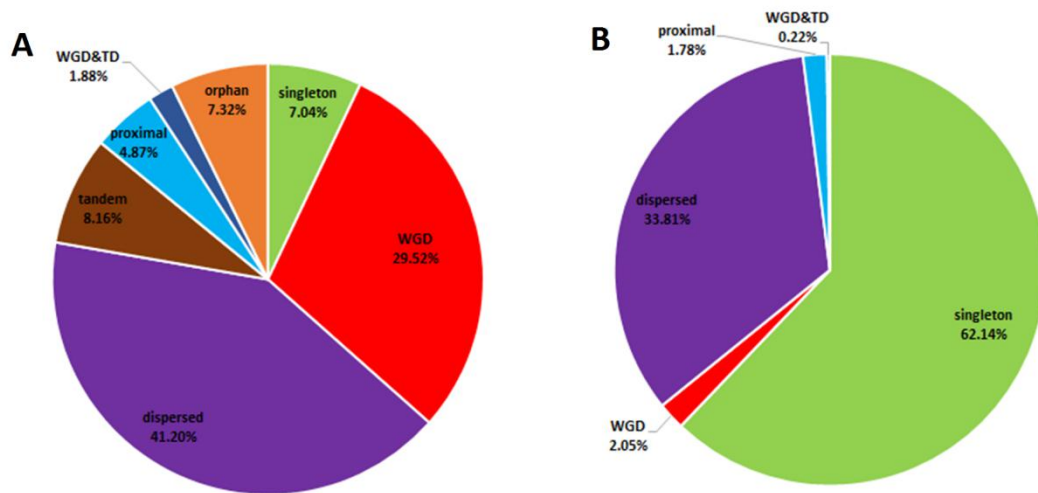
17
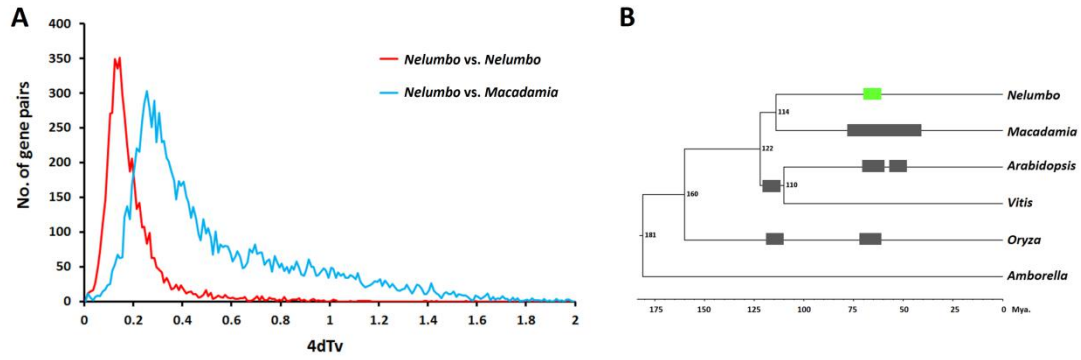
Supplementary Figure S5. Detection of lotus whole genome duplication and its relation to species divergence. A. 4dTv distributions of intra-specific (*Nelumbo*) syntenies and predicted orthologous gene pairs between lotus and *Macadamia*. B. Chronogram showing divergence time and WGDs in lotus and other key taxa.

22

23
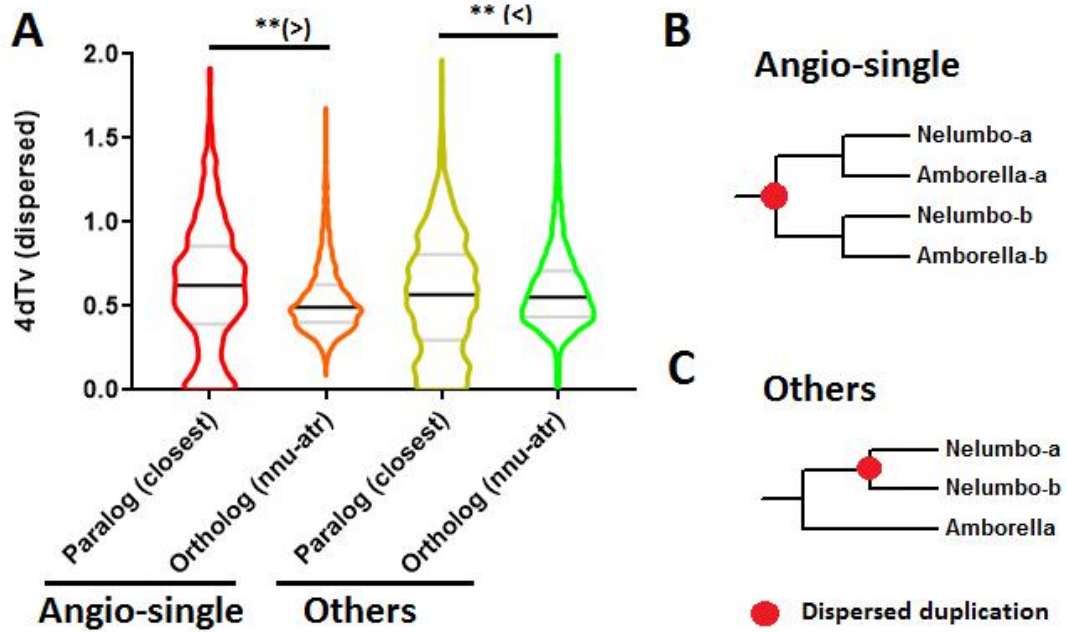

24

Supplementary Figure S6. Pie chart displaying the distribution of genes of different groups based on origins and duplication status. A. All lotus genes. B. Lotus orphan genes.

27

Supplementary Figure S7. 4dTv divergences between the closest paralogs and between orthologs (*Nelumbo-Amborella*) for dispersed duplicates as 'angiopserm conserved single copy genes' (Angio-single) and the other dispersed duplicates. A. Violin plot of 4dTv divergence for orthologs and paralogs. B. Presumed duplication event before split of *Nelumbo* and *Amborella*. C. Presumed duplication event after split of *Nelumbo* and *Amborella*.
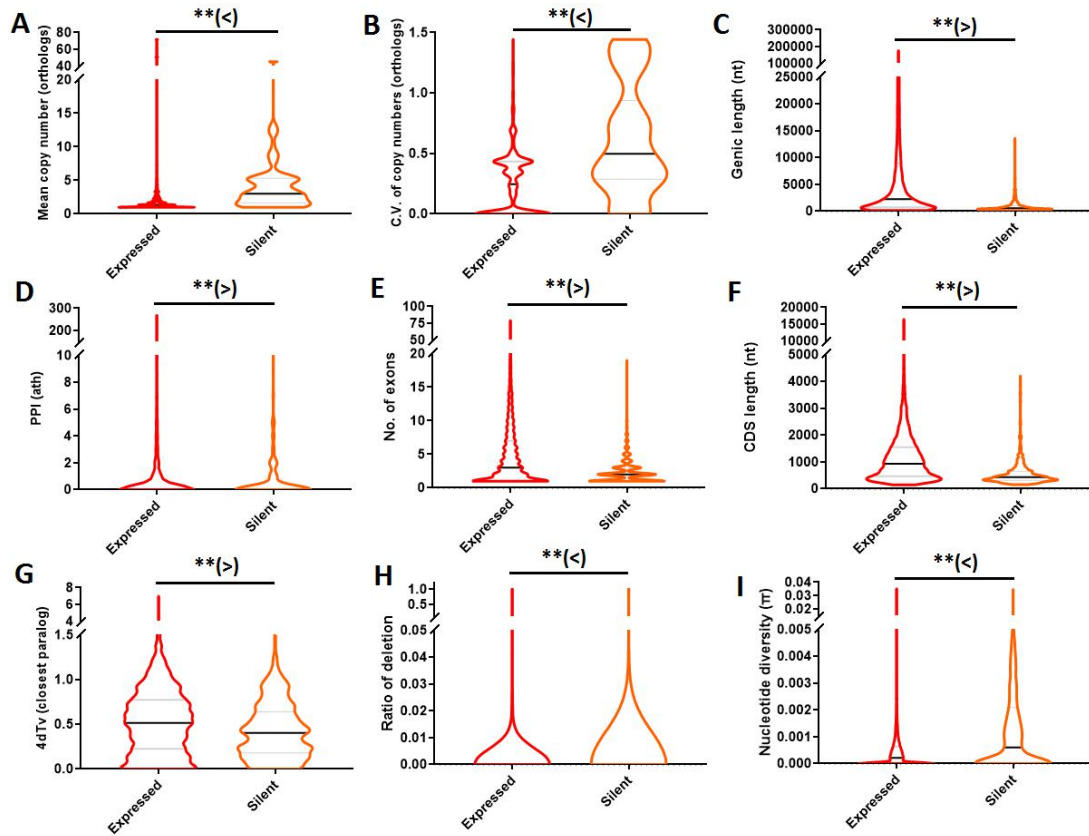
28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

Supplementary Figure S8. Differences between expressed dispersed duplicates and silent dispersed duplicates. A. Mean copy number of orthologs in three related taxa. B. C.V. of copy number of orthologs in three related taxa. C. Gene length. D. Protein-protein interactions inferred from Arabidopsis. E. The number of exons. F. CDS length. G. 4dTv between the closest paralogs in lotus. H. Ratio of deleted sequence in CDS. I. Nucleotide diversity in CDS.
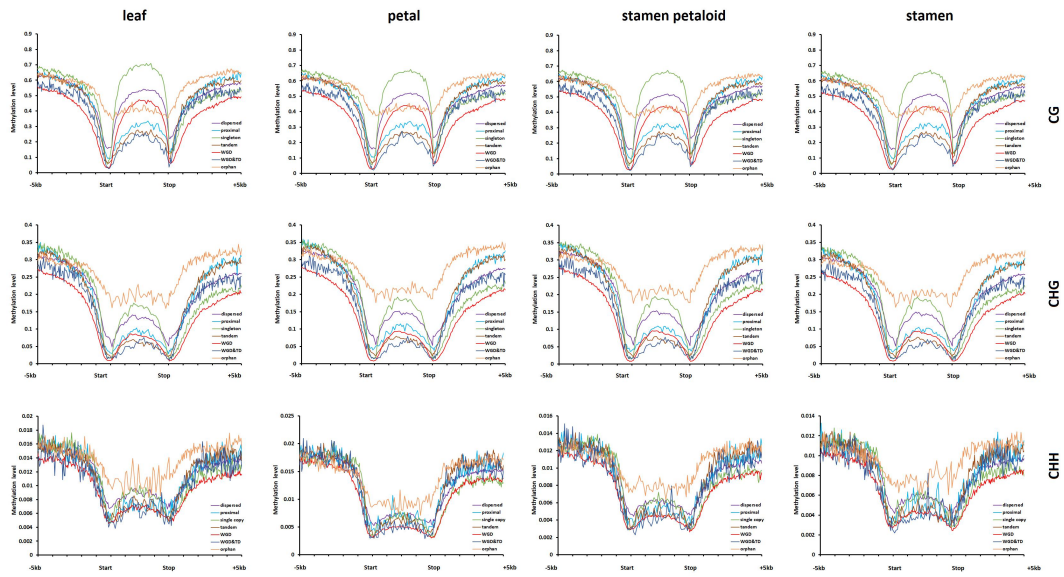
50

51

52

53

54

55

56

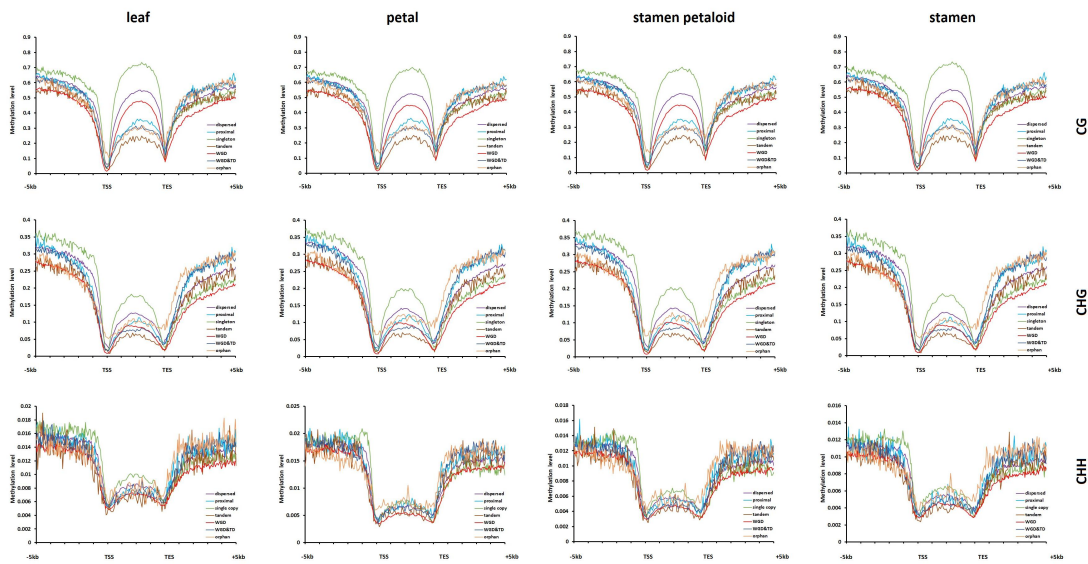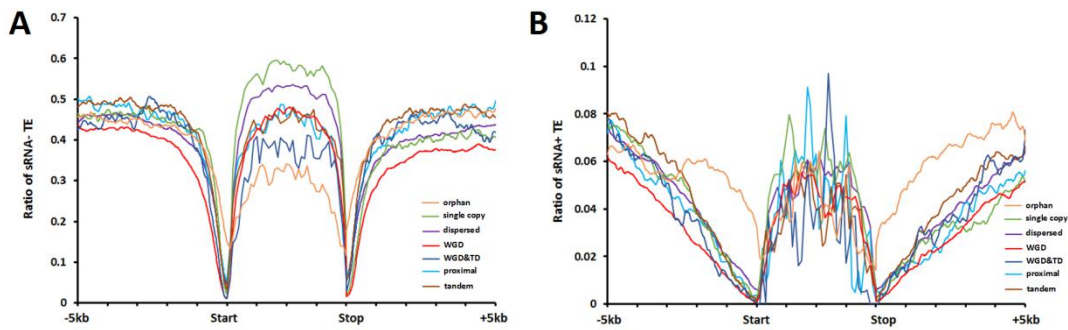57

58

59

Supplementary Figure S9. Differences in average CG, CHG and CHH methylation level (ML) along the gene (on genic and flanking regions) among different gene groups based on duplication status in lotus leaf, petal, stamen petaloid, and stamen.

63

Supplementary Figure S10. Differences in average CG, CHG and CHH methylation level (ML) along the representative transcript (on transcript and flanking regions) among different gene groups based on duplication status in lotus leaf, petal, stamen petaloid, and stamen.

69

Supplementary Figure S11. Differences in TE distribution of genic and flanking regions among different gene groups (based on duplication status) in lotus. A. sRNA- TEs. B. sRNA+ TEs.

73


74
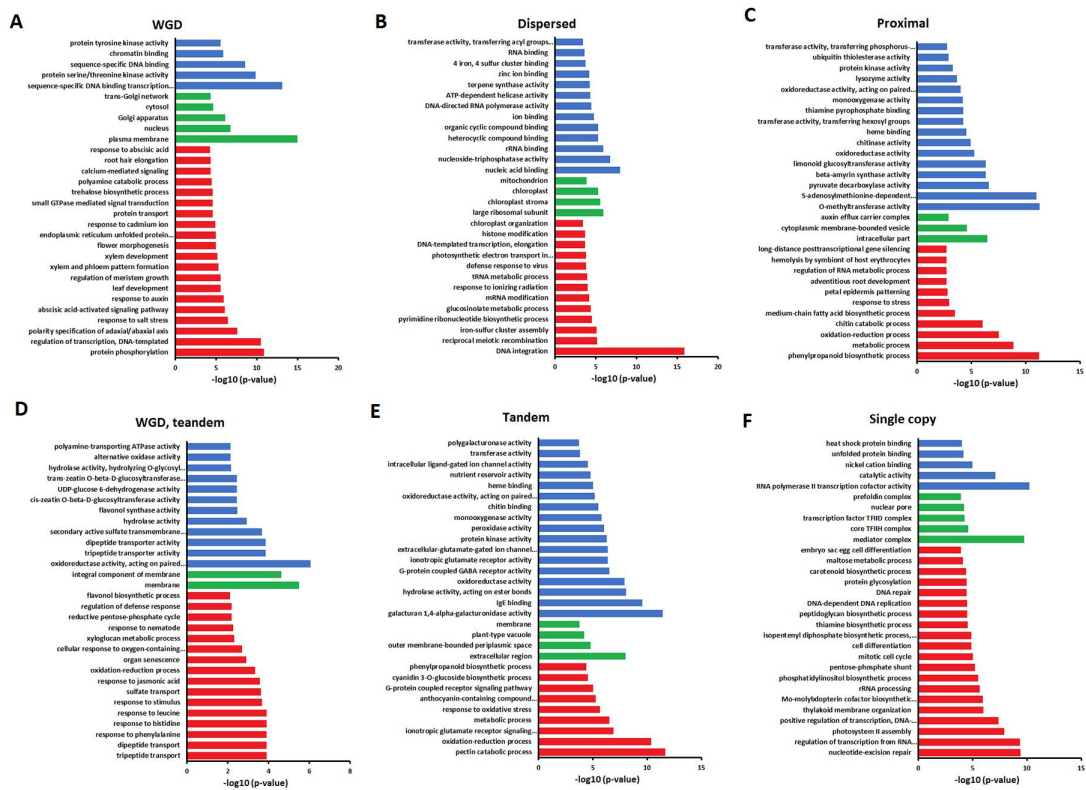
Supplementary Figure S12. Top 30 enriched Gene Ontology terms of genes from different gene groups (based on duplication status) in lotus. A. Genes retained from WGD. B. Dispersed duplicate genes. C. Proximal duplicate genes. D. Genes underwent both WGD and tandem duplication. E. Tandemly duplicated genes. F. Single copy genes.

79
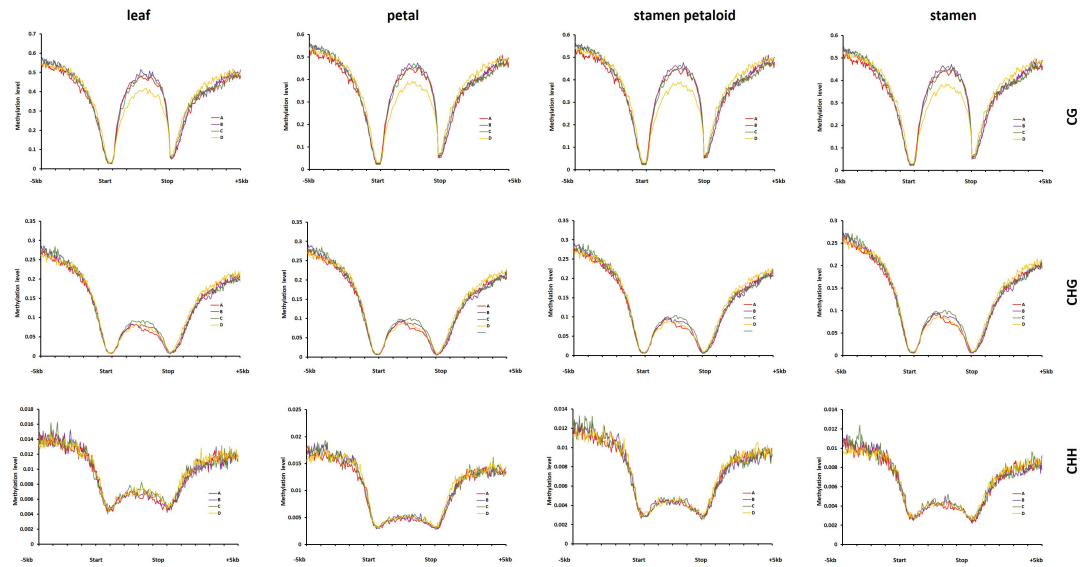
Supplementary Figure S13. Average CG, CHG and CHH methylation level (ML) in lotus leaf, petal, stamen petaloid, and stamen along the gene among duplicate genes that display different levels of expression divergence (Group A, Group B, Group C and Group D).

84
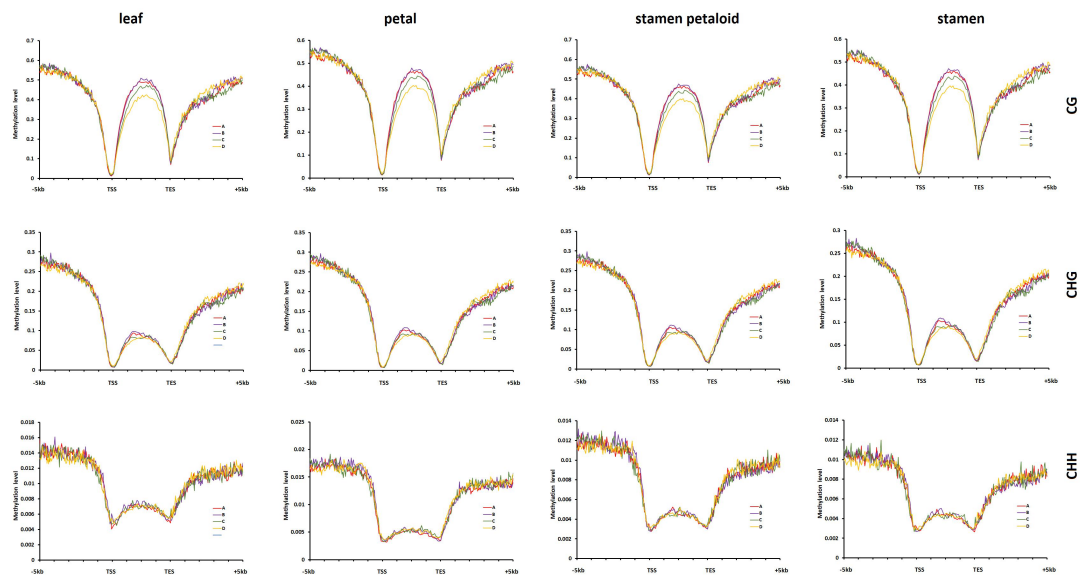
Supplementary Figure S14. Average CG, CHG and CHH methylation level (ML) in lotus leaf, petal, stamen petaloid, and stamen along the representative transcript among duplicate genes that display different levels of expression divergence (Group A, Group B, Group C and Group D).

89
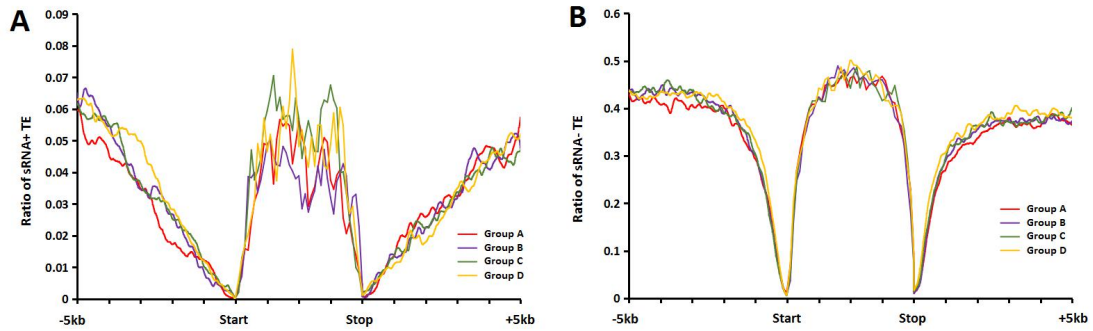
Supplementary Figure S15. Differences in TE distribution of genic and flanking regions among WGD duplicates displaying different levels of expression divergence (Group A, Group B, Group C and Group D). A. sRNA- TEs. B. sRNA+ TEs.
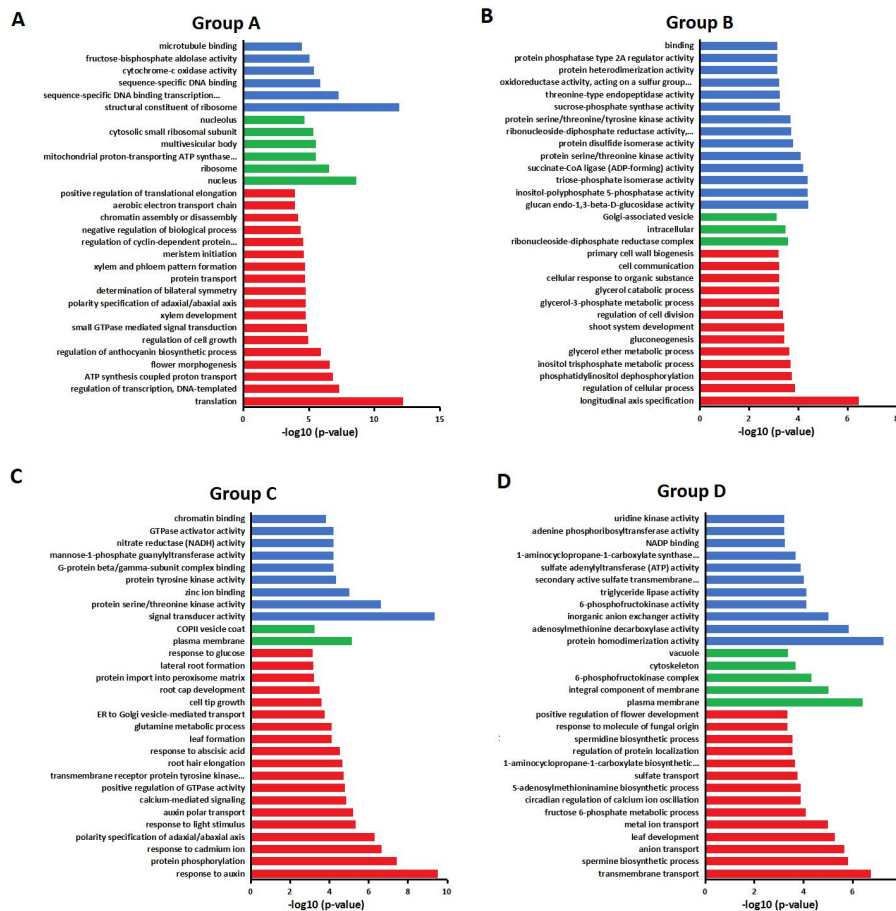
93

94
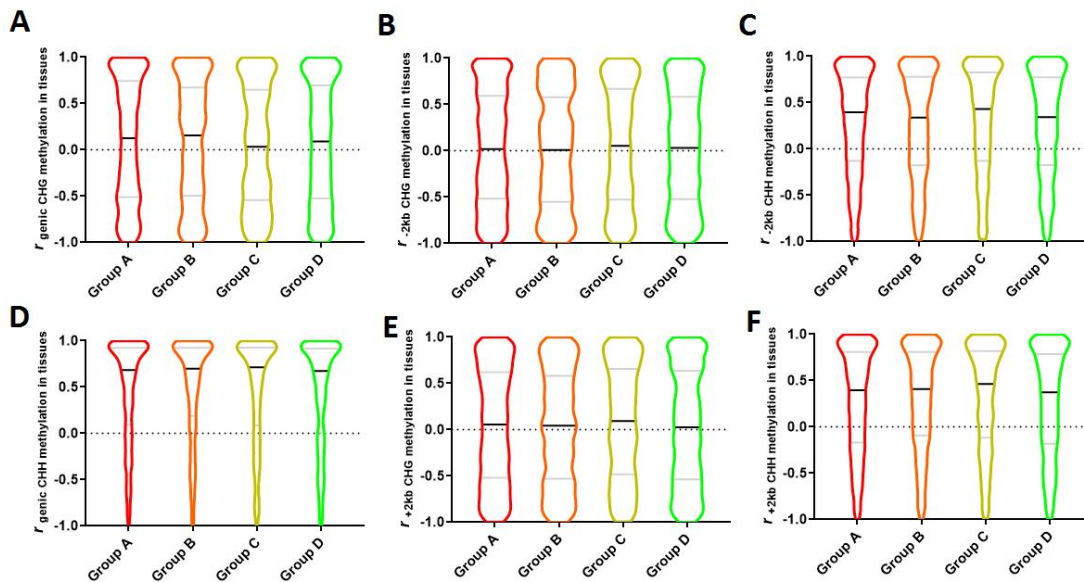
Supplementary Figure S16. Violin plots of *r* (correlation coefficient) for CHG and CHH methylation on gene body, gene upstream (-2kb) and downstream (+2kb) in different tissues between WGD-derived duplicate genes with different level of expression divergence (Group A, Group B, Group C and Group D). A-C: CHG. D-F: CHH. Black line: median; grey line: quantile.
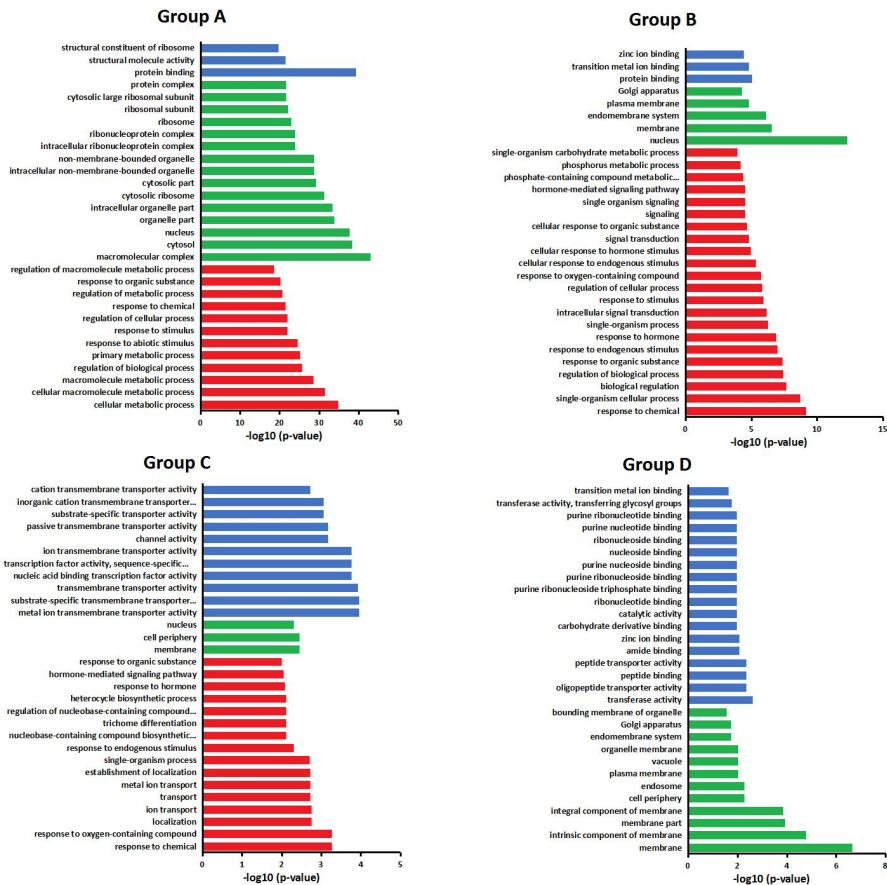
100

101

Supplementary Figure S17. Top 30 Gene Ontology terms enriched in lotus WGD

duplicates that display different levels of expression divergence (Group A, Group B,

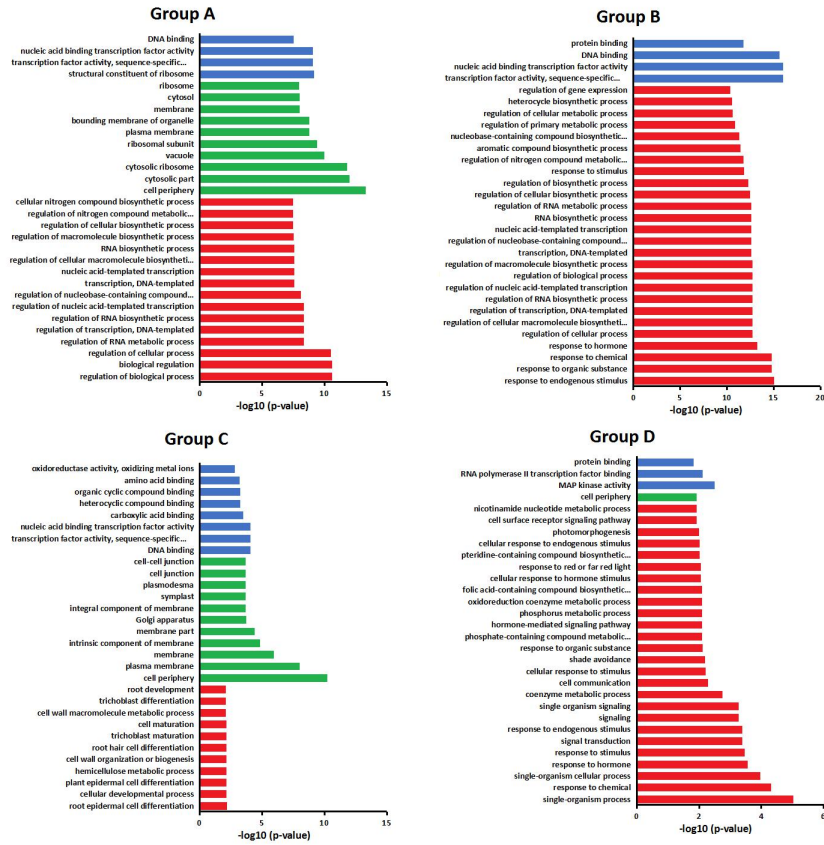**Group C and Group D).** A. Group A. B. Group B. C. Group C. D. Group D.

105
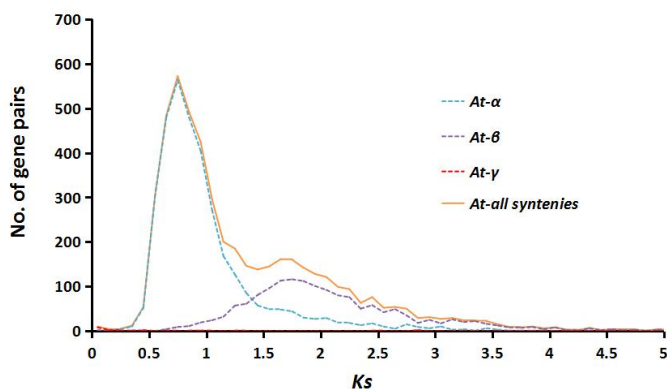
Supplementary Figure S18. Grouping of *Arabidopsis* syntenies into three presumed

WGDs, At-α, At-β and At-γ by pairwise *t*-test based on *Ks* values.

108

Supplementary Figure S19. Top 30 Gene Ontology terms enriched in At-α WGD duplicates that display different levels of expression divergence (Group A, Group B, Group C and Group D). Note that GO terms with more than 10,000 annotated genes, such as cellular process, are not shown. A. Group A. B. Group B. C. Group C. D. Group D.
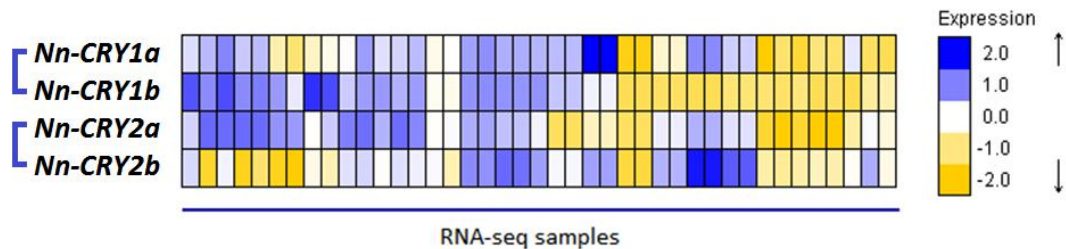
113

Supplementary Figure S20. Top 30 Gene Ontology terms enriched in At-β WGD duplicates that display different levels of expression divergence (Group A, Group B, Group C and Group D). Note that GO terms with more than 10,000 annotated genes, such as cellular process, are not shown. A. Group A. B. Group B. C. Group C. D. Group D.
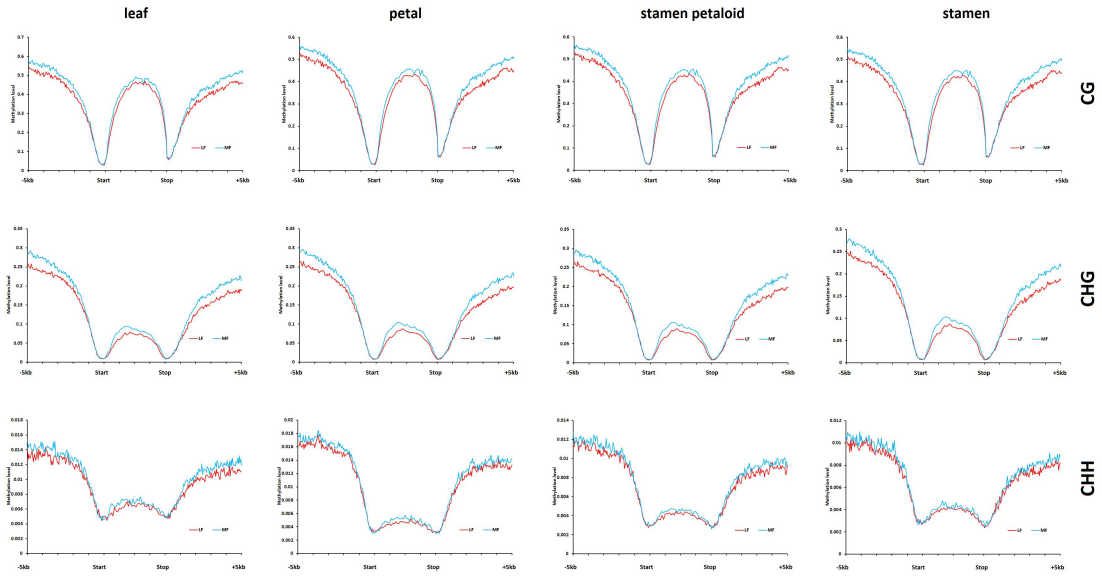
118

119

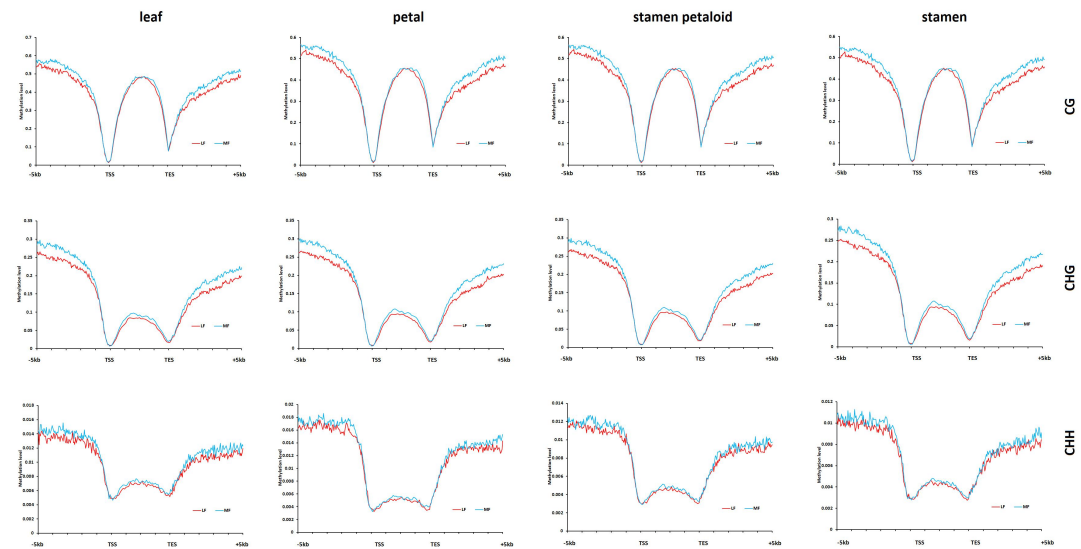**Supplementary Figure S21. Distinct expression patterns of lotus *CRY1a,b* and *CRY2a,b*.**

121

Supplementary Figure S22. Differences in average CG, CHG and CHH methylation level (ML) in lotus leaf, petal, stamen petaloid, and stamen along gene (on genic and flanking regions) between duplicates that belong to less fractionated blocks (LFs) and more fractionated blocks (MFs).
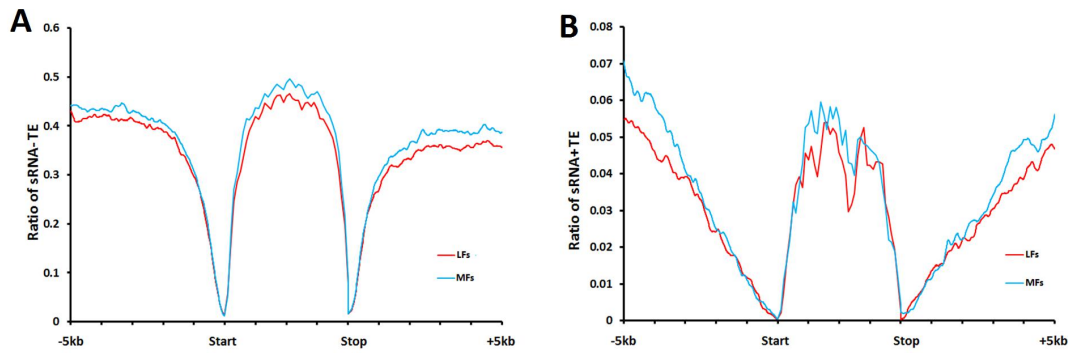
126

127

128

Supplementary Figure S23. Differences in average CG, CHG and CHH methylation level (ML) in lotus leaf, petal, stamen petaloid, and stamen along the representative

14

131   **transcript (on transcript and flanking regions) between duplicates that belong to less**

132   **fractionated blocks (LFs) and more fractionated blocks (MFs) .**
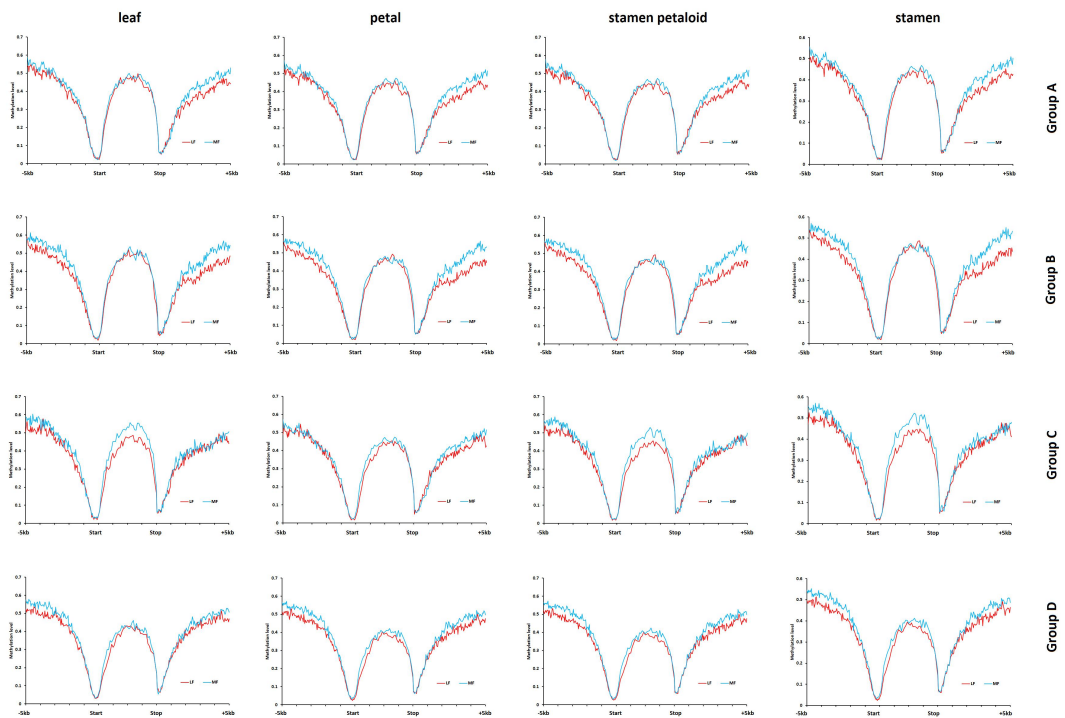
133

134

135   **Supplementary Figure S24. Differences in average TE distribution along gene (on genic**

136   **and flanking regions) between duplicates that belong to less fractionated blocks (LFs)**

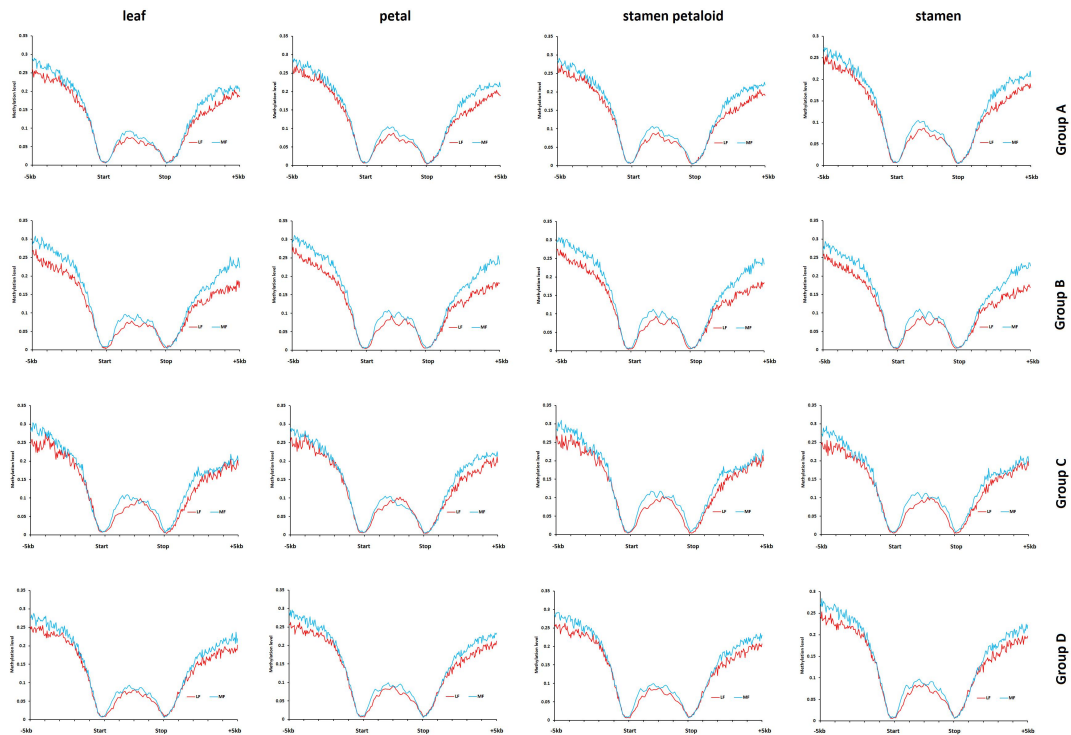137   **and more fractionated blocks (MFs).** A. sRNA- TEs. B. sRNA+ TEs.

138

139

140   **Supplementary Figure S25. Differences in average CG methylation level (ML) in lotus**

141 **leaf, petal, stamen petaloid, and stamen along gene (on genic and flanking regions)**

142 **between duplicates that belong to less fractionated blocks (LFs) and more fractionated**

143 **blocks (MFs) in Group A, Group B, Group C and Group D.**
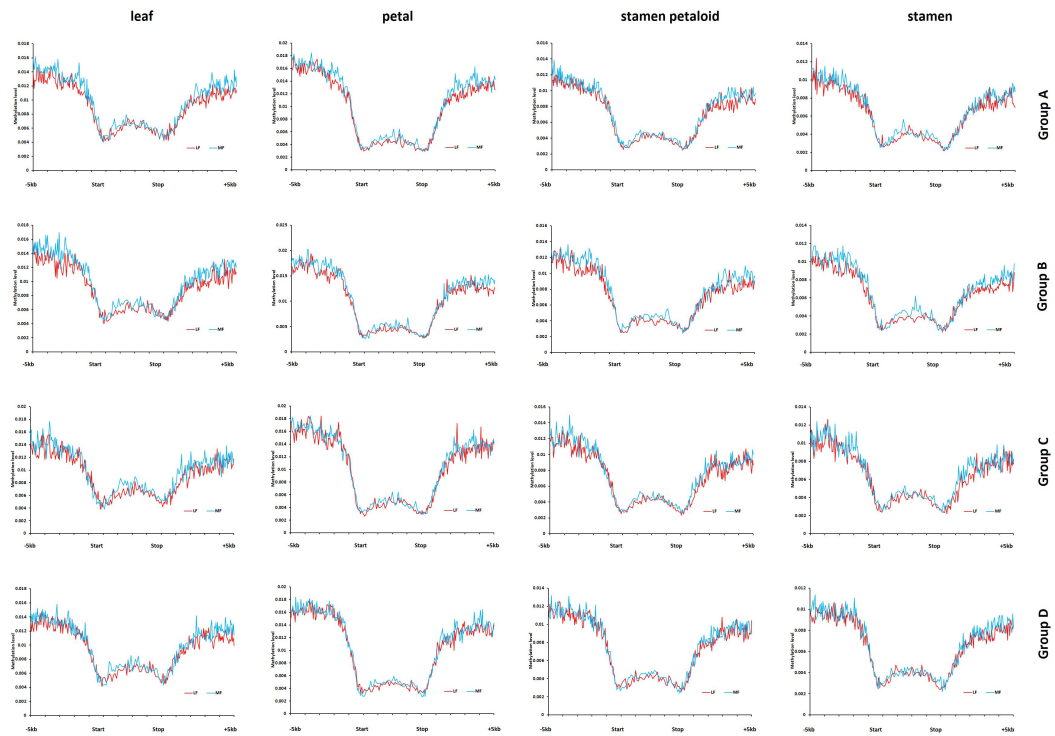
144

145

146

147 **Supplementary Figure S26. Differences in average CHG methylation level (ML) in lotus**

148 **leaf, petal, stamen petaloid, and stamen along gene (on genic and flanking regions)**

149 **between duplicates that belong to less fractionated blocks (LFs) and more fractionated**

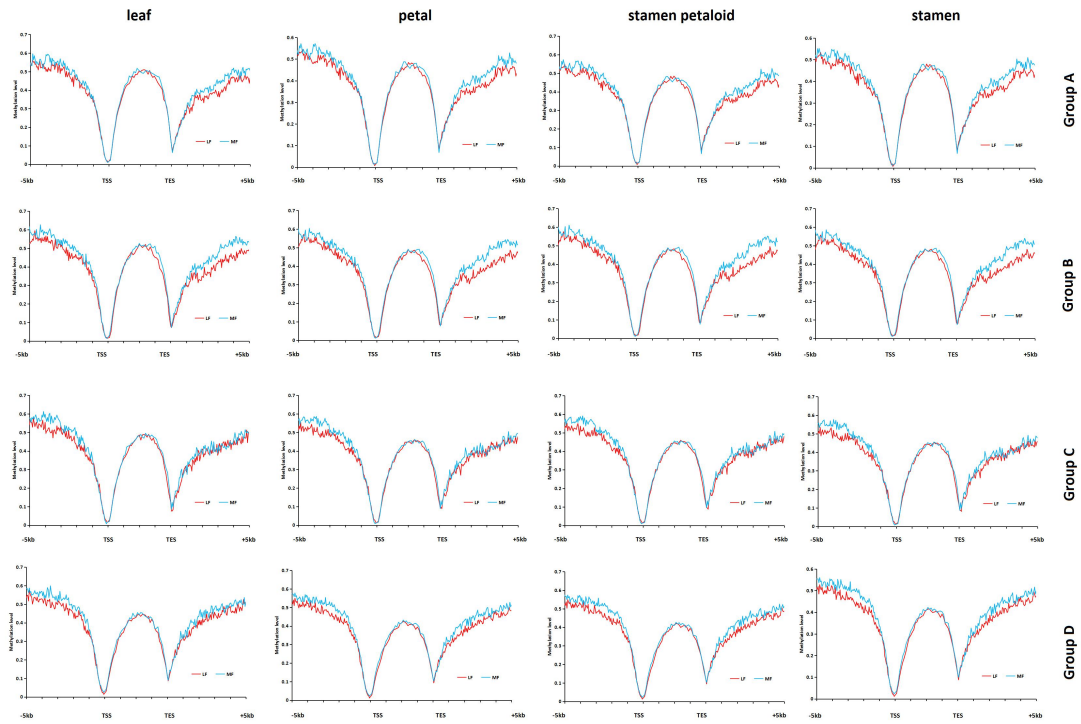150 **blocks (MFs) in Group A, Group B, Group C and Group D.**

151

152

Supplementary Figure S27. Differences in average CHH methylation level (ML) in lotus leaf, petal, stamen petaloid, and stamen along gene (on genic and flanking regions) between duplicates that belong to less fractionated blocks (LFs) and more fractionated blocks (MFs) in Group A, Group B, Group C and Group D.

153
154
155
156

157

158

159

Supplementary Figure S28. Differences in average CG methylation level (ML) in lotus leaf, petal, stamen petaloid, and stamen along the representative transcript (on transcript and flanking regions) between duplicates that belong to less fractionated blocks (LFs) and more fractionated blocks (MFs) in Group A, Group B, Group C and Group D.
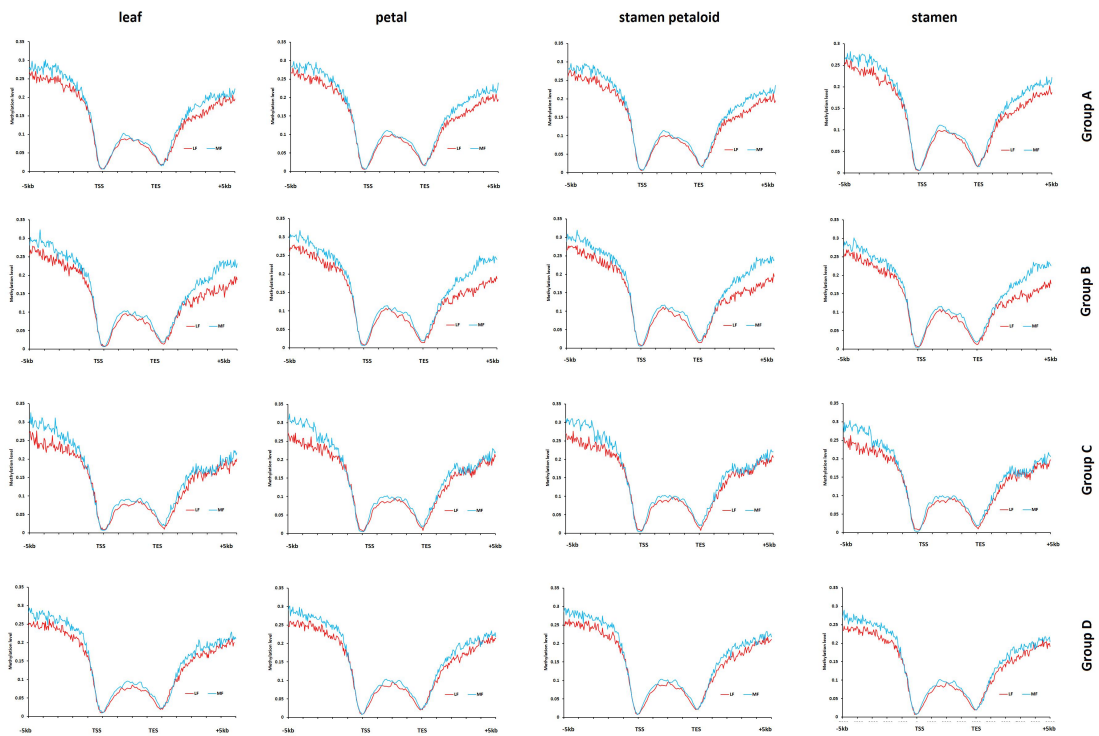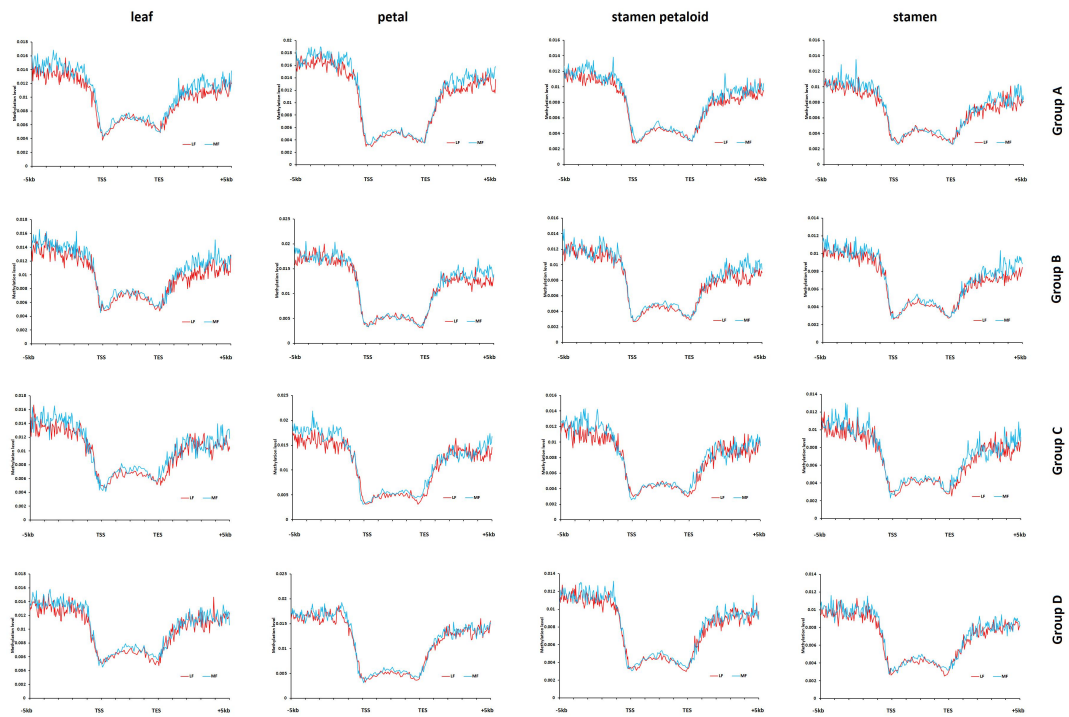
160
161
162
163
164

165

166

167

Supplementary Figure S29. Differences in average CHG methylation level (ML) in lotus leaf, petal, stamen petaloid, and stamen along the representative transcript (on transcript and flanking regions) between duplicates that belong to less fractionated blocks (LFs) and more fractionated blocks (MFs) in Group A, Group B, Group C and Group D.

168
169
170
171
172

173

174

Supplementary Figure S30. Differences in average CHH methylation level (ML) in lotus leaf, petal, stamen petaloid, and stamen along the representative transcript (on transcript and flanking regions) between duplicates that belong to less fractionated blocks (LFs) and more fractionated blocks (MFs) in Group A, Group B, Group C and Group D.