Figure S1. Density distribution of consensus full-length transcripts obtained by Nanopore sequencing. The average length of reads is shown in a red dashed line.

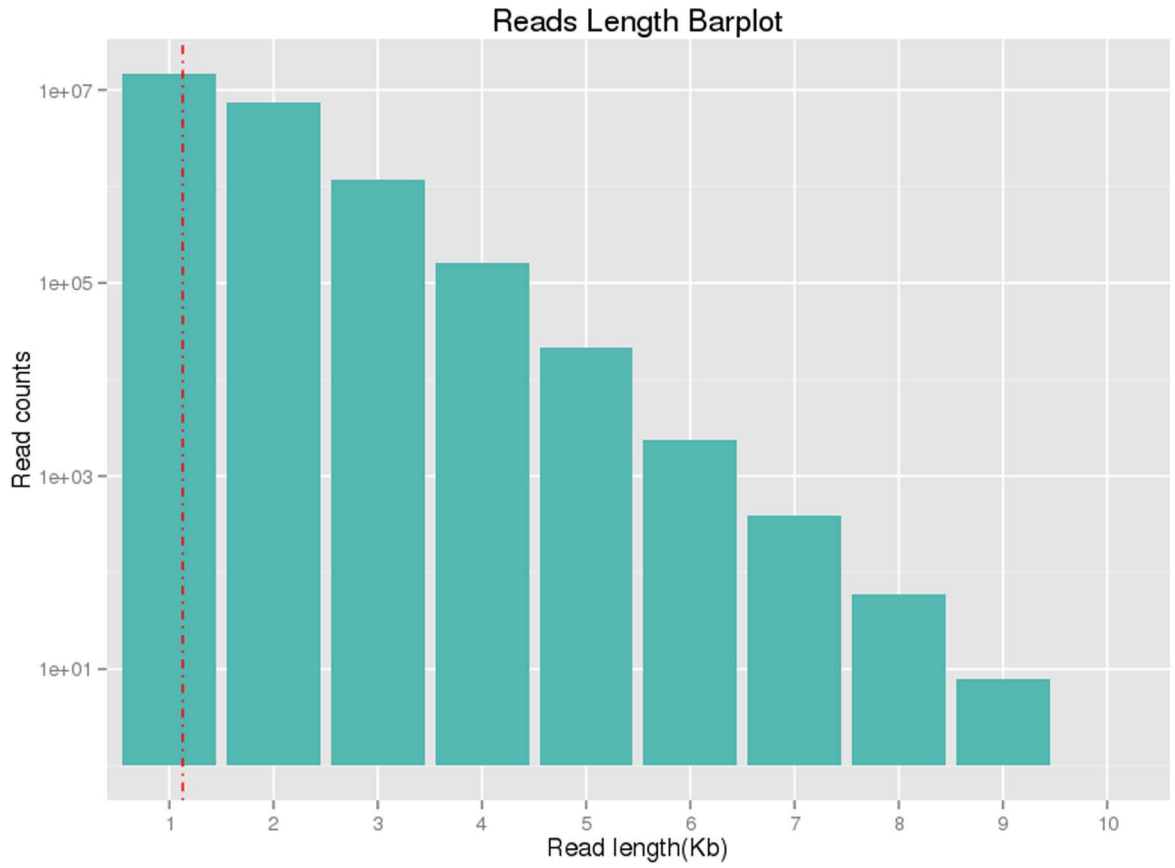


Figure S2. Distribution of the number of isoforms per gene in *N. lutea* (a), *N. nucifera* (b), and *Arabidopsis* (c).

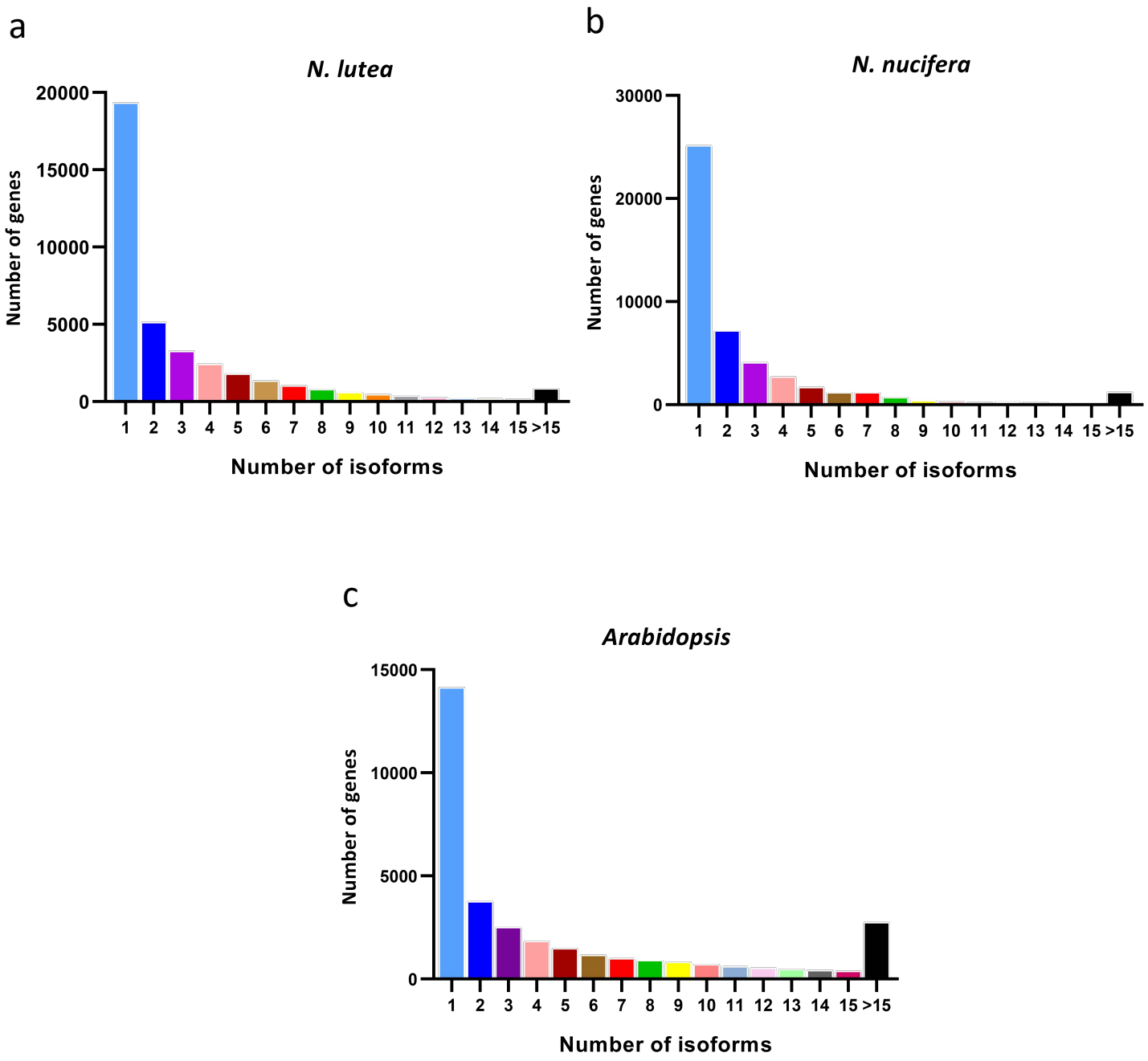


Figure S3. Pipeline to identify interspecies conserved AS events in the same ortholog group. For the different orthologous genes in the same group, up to 45 bp on either side of the splice junctions were extracted (yellow arrows). For each gene, extracted sequences were grouped and concatenated per splice type event. This allows comparing junction-specific sequences from a specific gene between different species. This comparison was performed between junction-specific sequences that belong to the same AS type. Hereto we used BLASTN, retaining matches with p-value < 10-5 and a proportion of perfectly matching bases in the target > 80%.

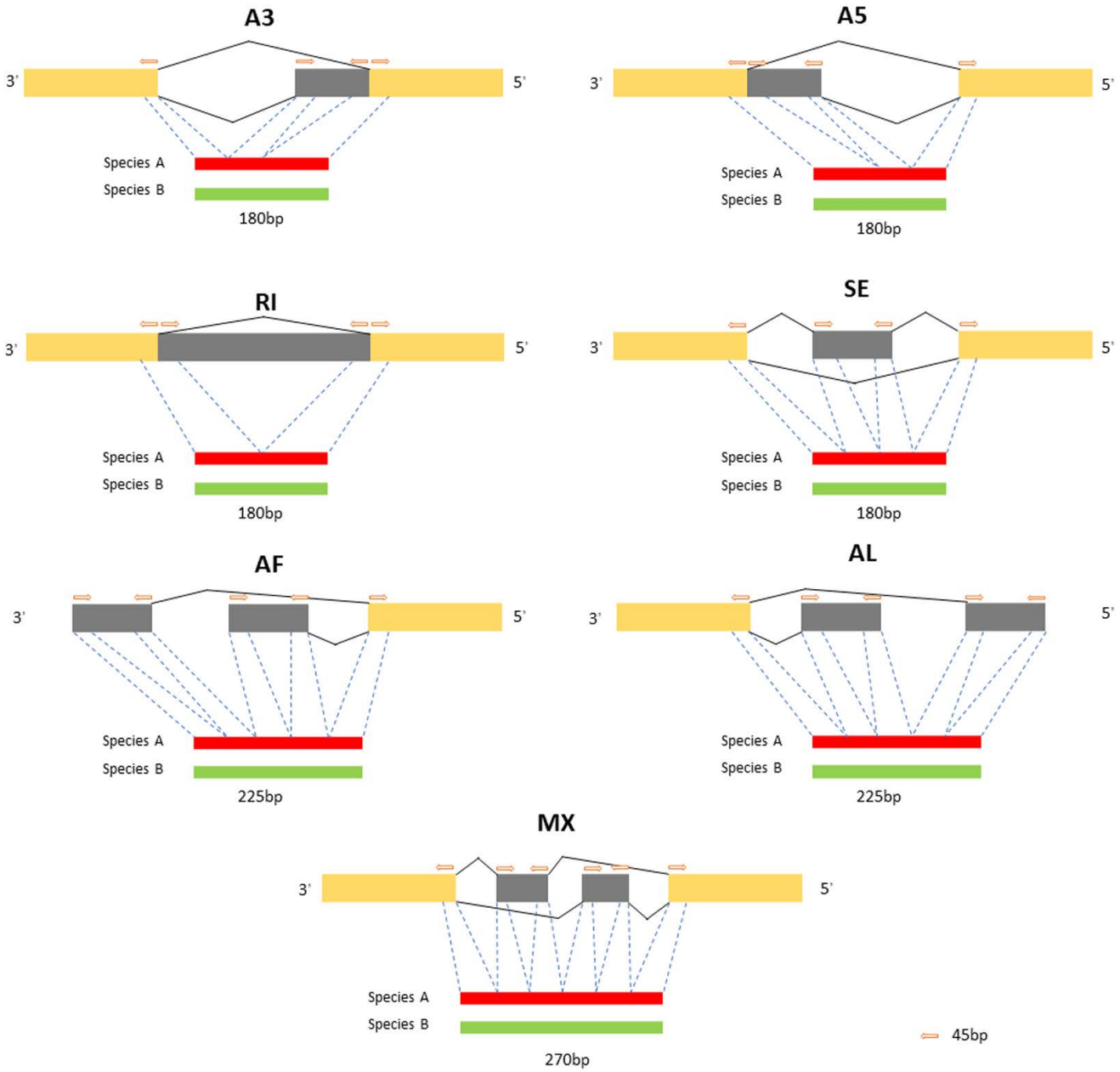
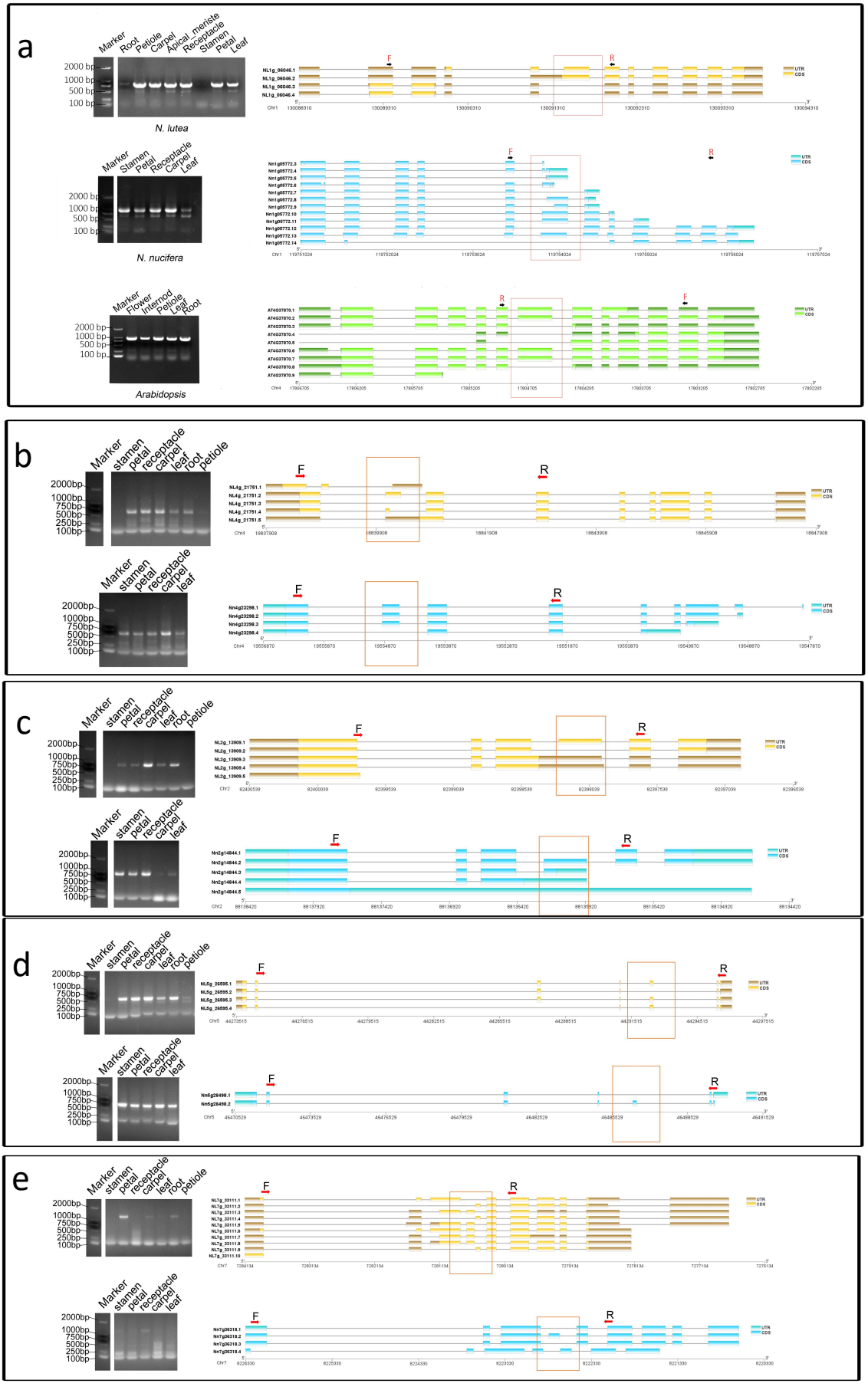
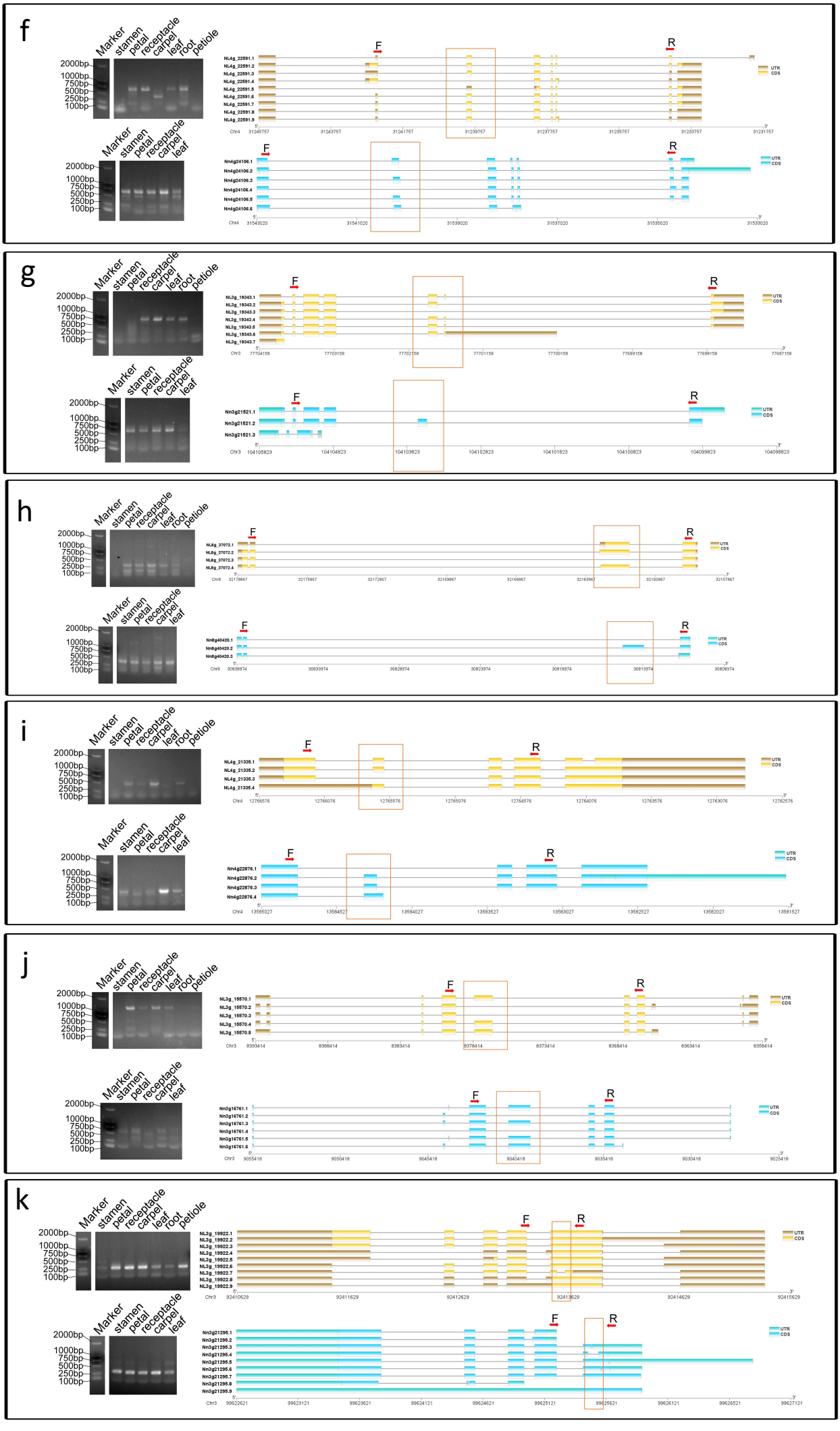
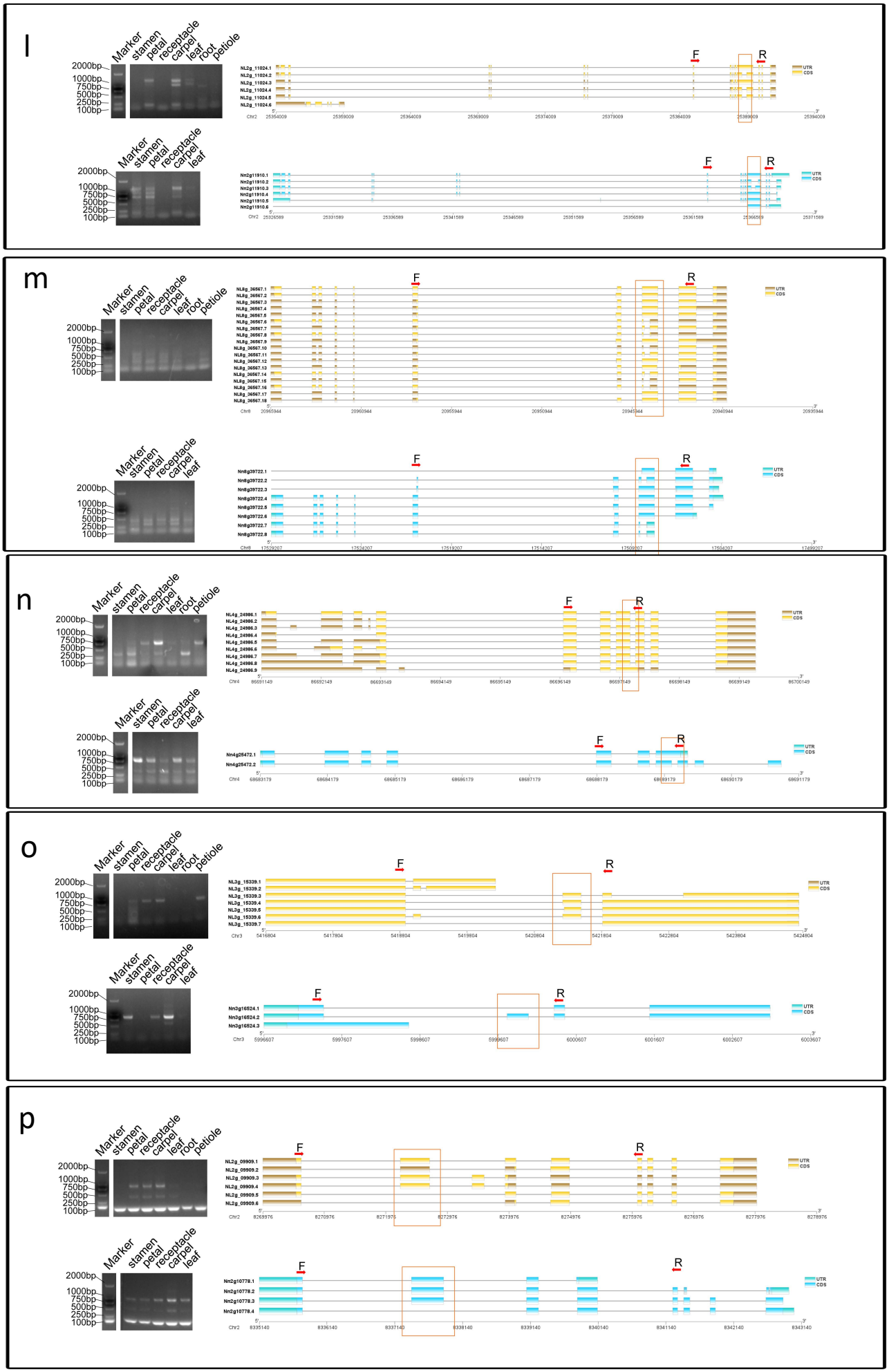


Figure S4. (a-s): RT-PCR validation of interspecies conserved AS events that are shared between orthologous genes of *N. lutea*, *N. nucifera*, and *Arabidopsis*. Each black rectangular frame was the RT-PCR validation of an interspecies conserved AS event. The amplified fragments of the RT-PCR experiment for orthologous genes in three species tissues and the 2k DNA marker were tested in agarose gel and shown on the panel at the left. Non-contiguous gel lanes (2k DNA marker vs. RT-PCR products) are demarcated by white spaces. The predicted structure of each isoform is shown on the panel at the right. Yellow boxes and arrowed lines show respectively exons and introns in *N. lutea*, blue boxes and arrowed lines and green boxes and arrowed lines show the same, but in respectively *N. nucifera*, and *Arabidopsis*. The arrows show the loci of the PCR primers (F, forward and R, reverse) on the first isoform of each gene. The red rectangles show the interspecies conserved AS events. The genomic locus of this interspecies conserved AS event in each of the three species can be found in Table S3.




f
Marker
stamen
petal
receptacle
carpel
leaf
root
petiole
2000bp
1000bp
750bp
500bp
250bp
100bp
F
R
UTR
CDS
g
h
i
j
k


l
Marker
stamen
petal
receptacle
carpel
leaf
root
petiole
2000bp
1000bp
750bp
500bp
250bp
100bp
F
R
UTR
CDS
m
n
o
p

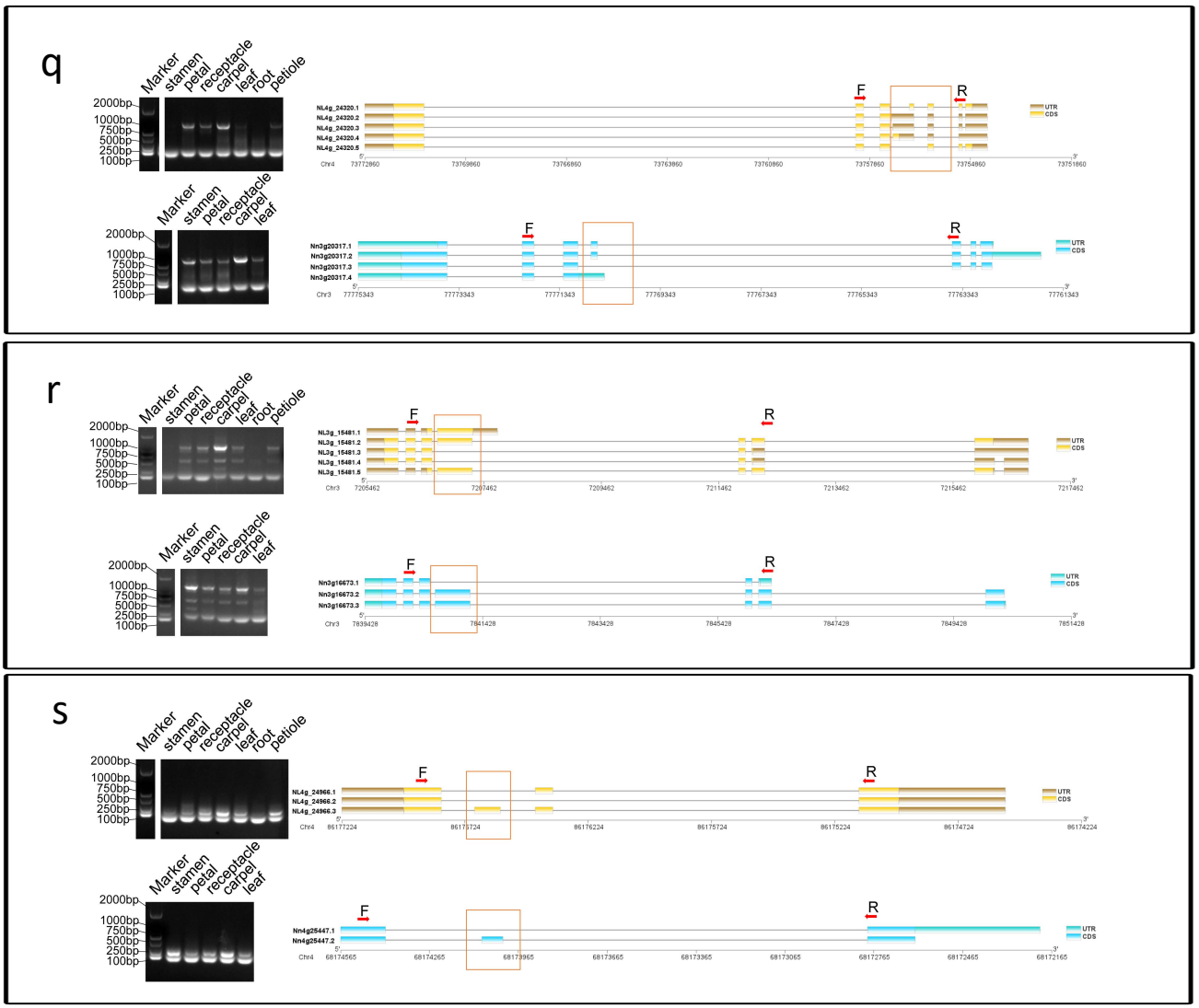

q
Marker
stamen
petal
receptacle
carpel
leaf
root
petiole
2000bp
1000bp
750bp
500bp
250bp
100bp
F
R
NL4g_24320.1
NL4g_24320.2
NL4g_24320.3
NL4g_24320.4
NL4g_24320.5
UTR
CDS
Chr4
73772860
73769860
73766860
73763860
73760860
73757860
73754860
73751860
Nn3g20317.1
Nn3g20317.2
Nn3g20317.3
Nn3g20317.4
Chr3
77775343
77773343
77771343
77769343
77767343
77765343
77763343
77761343
r
NL3g_15481.1
NL3g_15481.2
NL3g_15481.3
NL3g_15481.4
NL3g_15481.5
Chr3
7205462
7207462
7209462
7211462
7213462
7215462
7217462
Nn3g16673.1
Nn3g16673.2
Nn3g16673.3
7839428
7841428
7843428
7845428
7847428
7849428
7851428
s
NL4g_24966.1
NL4g_24966.2
NL4g_24966.3
Chr4
86177224
86176224
86175724
86175224
86174724
86174224
Nn4g25447.1
Nn4g25447.2
68174565
68174265
68173965
68173665
68173365
68173065
68172765
68172465
68172165

Figure S5. Box plots showing the distribution of the number of AS events per gene in genes with different copies in *N. lutea* (a), *N. nucifera* (b), and *Arabidopsis* (c). Pairwise comparisons of the number of AS events between gene groups were performed using a Mann-Whitney *U* test, ** means $p < 0.01$.

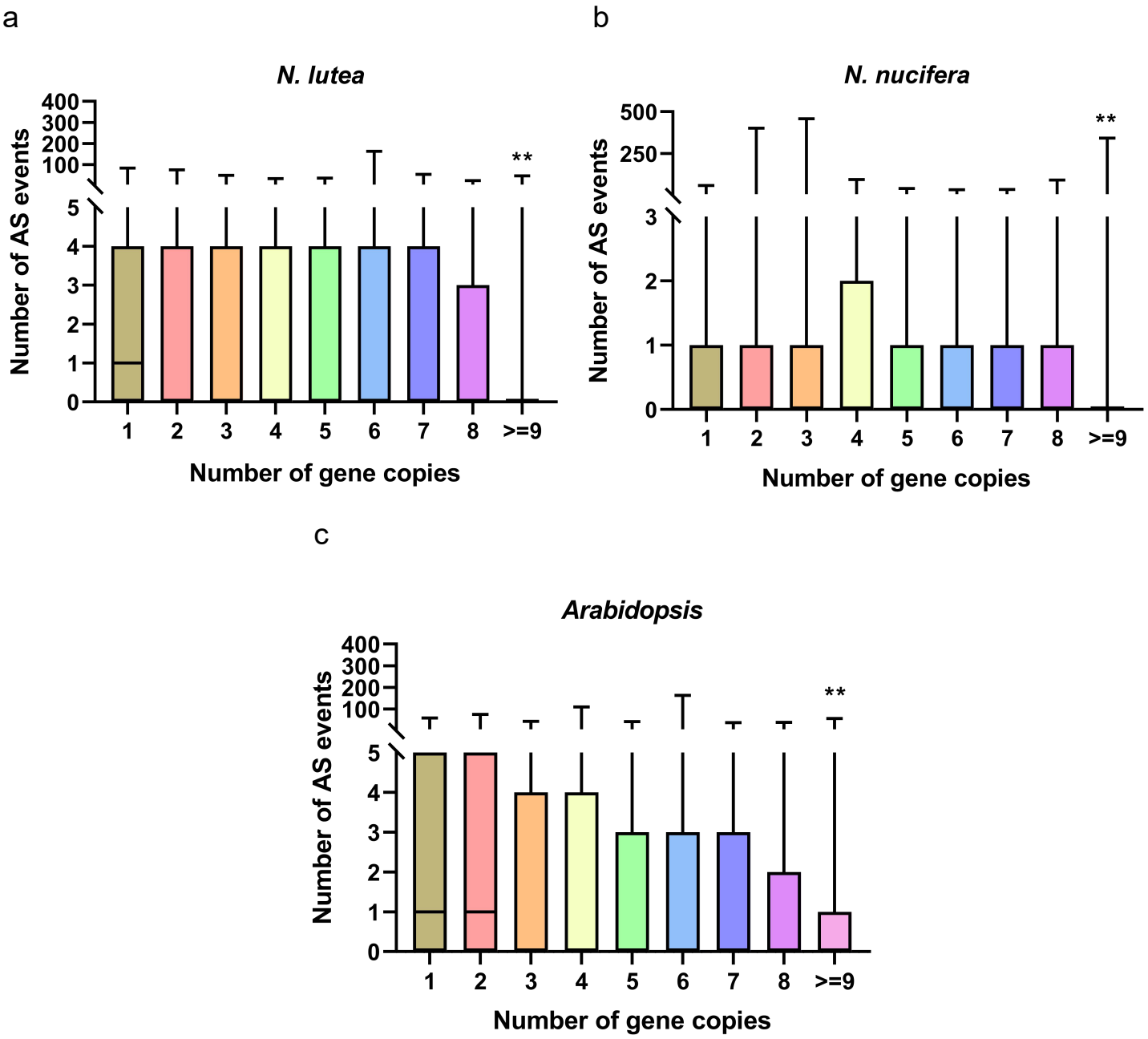


Figure S6. Percentage of respectively single-copy or duplicated genes that undergo AS events. The percentage represents the ratio of the number of either the single-copy or duplicated genes that undergo AS events versus the total number of singe-copy or duplicated genes in *N. lutea*, *N. nucifera*, and *Arabidopsis*. The significance of the difference between single-copy and duplicated genes is tested by the chi-square test, * means p-value < 0.01.

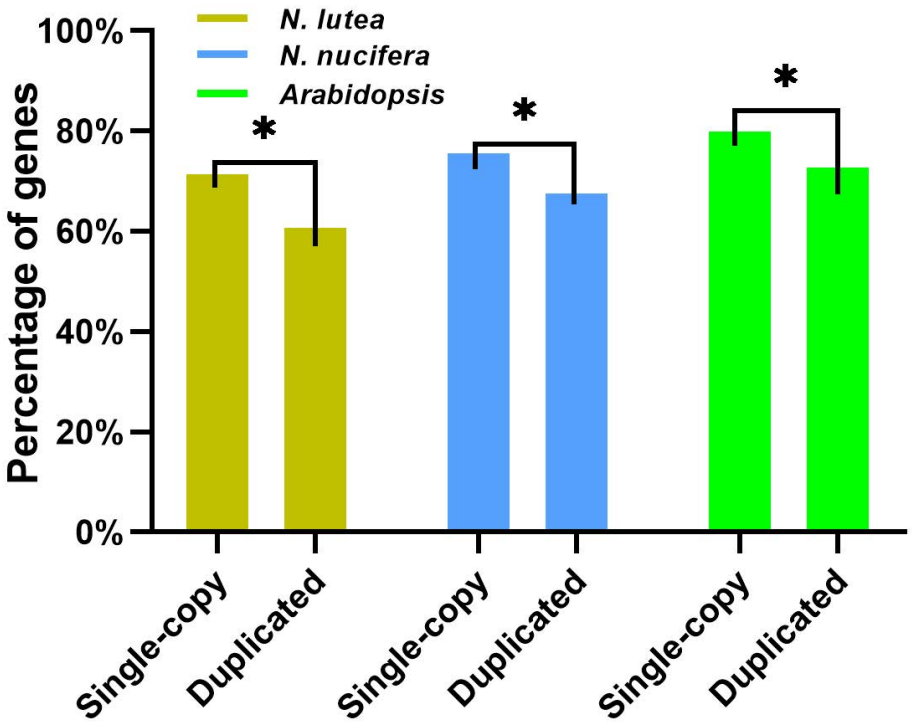


Figure S7. The bar graph shows the percentage of genes with at least one interspecies conserved AS event in single-copy genes or duplicated genes in *N. lutea* (yellow) and *N. nucifera* (blue). Because an interspecies conserved AS event might involve multiple orthologs within one OG, we consider for the duplicated genes the number of genes that undergo at least one interspecies conserved AS event versus the total number of duplicated genes in each species. (* means p-value < 0.01, *t*-test).

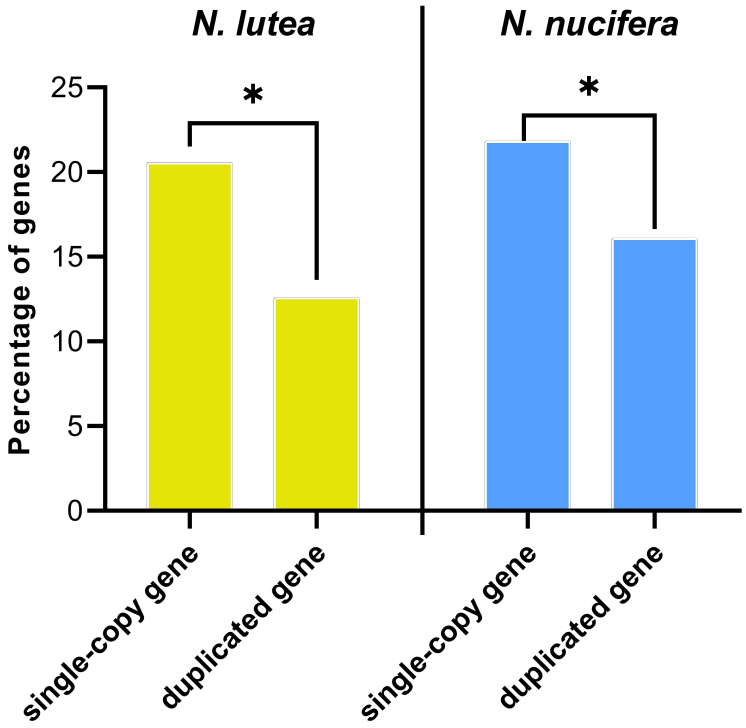


Figure S8. Bar chart showing the average number of AS events per gene. Genes are grouped according to their origin of duplication and according to the species in which they occur.

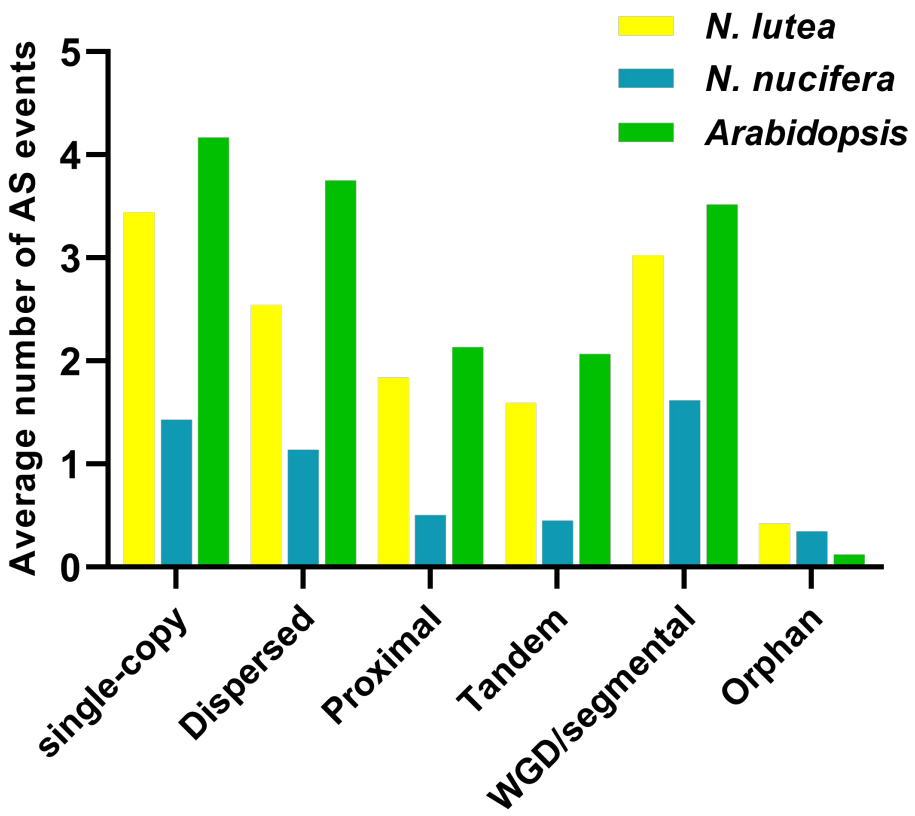
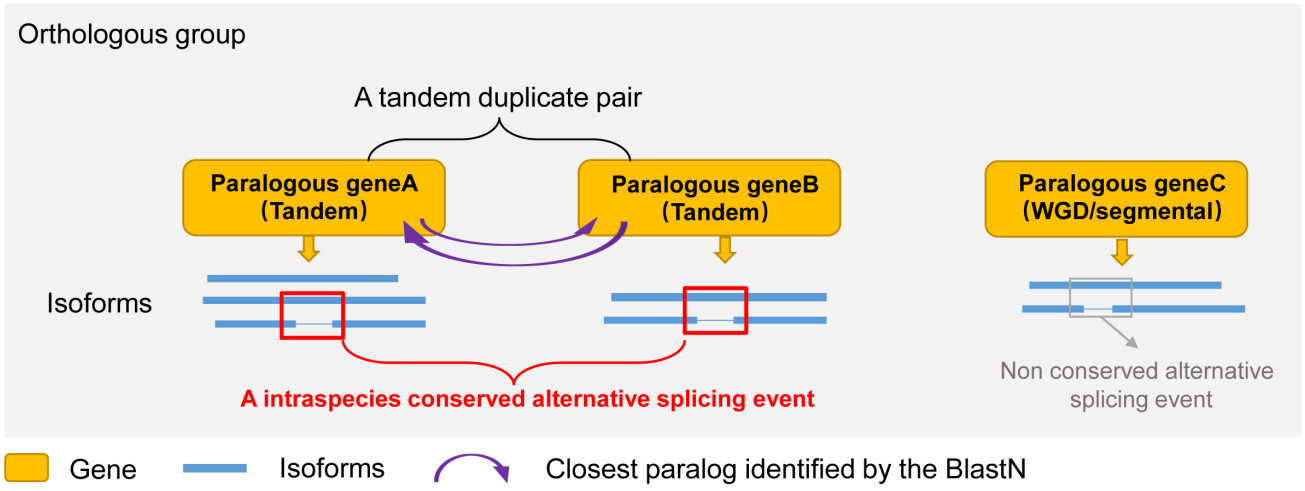


Figure S9. Intraspecies conserved AS events in paralogous gene pairs. (a) Schematic diagram of the identification process of intraspecies conserved AS events. Herein we identified an intraspecies conserved AS event (the red frame) in a tandem duplicate pair of paralogous genes using the 'splice junction-based approach' (see experimental procedure). Because an orthologous group might involve multiple paralogs, we only consider paralog pairs (i.e. the query gene and its closest paralog identified by BlastN) with the same origin of duplication. If two paralogous genes were each others mutual best hits, we regard them as a single paralogous pair to avoid counting them twice. (b) Percentages of the number of paralogous gene pairs with at least one conserved AS event versus the total number of paralogous gene pairs of a certain origin (duplication types) were calculated in N. lutea, N. nucifera, and Arabidopsis. Pairwise comparisons between different duplicated gene groups were tested by the chi-square test, * means p-value < 0.01.

a



b

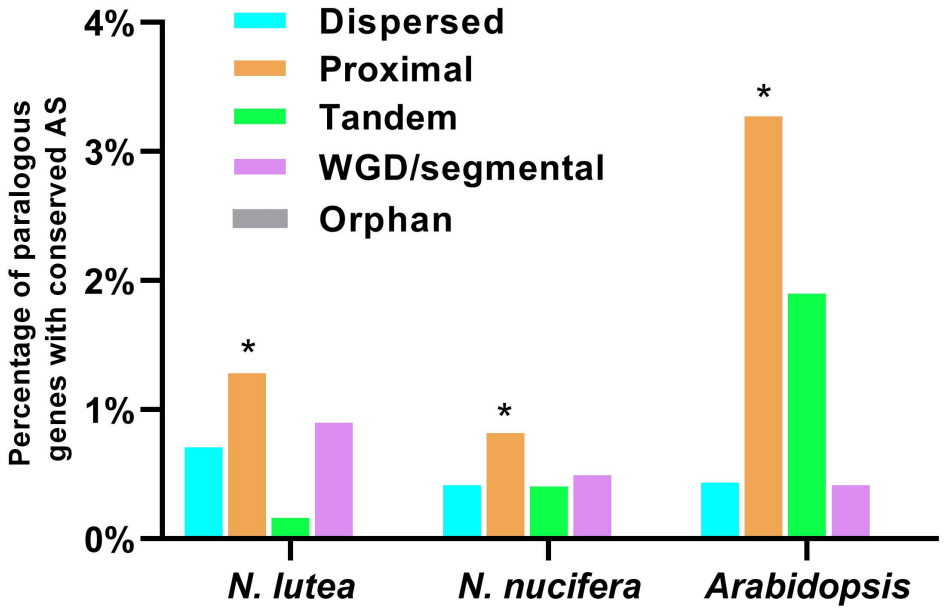


Figure S10. Heatmap showing the presence/absence of the candidate *N. nucifera* lineage-specific isoforms in other *N. nucifera* cultivars as obtained from transcriptome analysis. Green: observed isoforms; Gray: isoforms not detected.

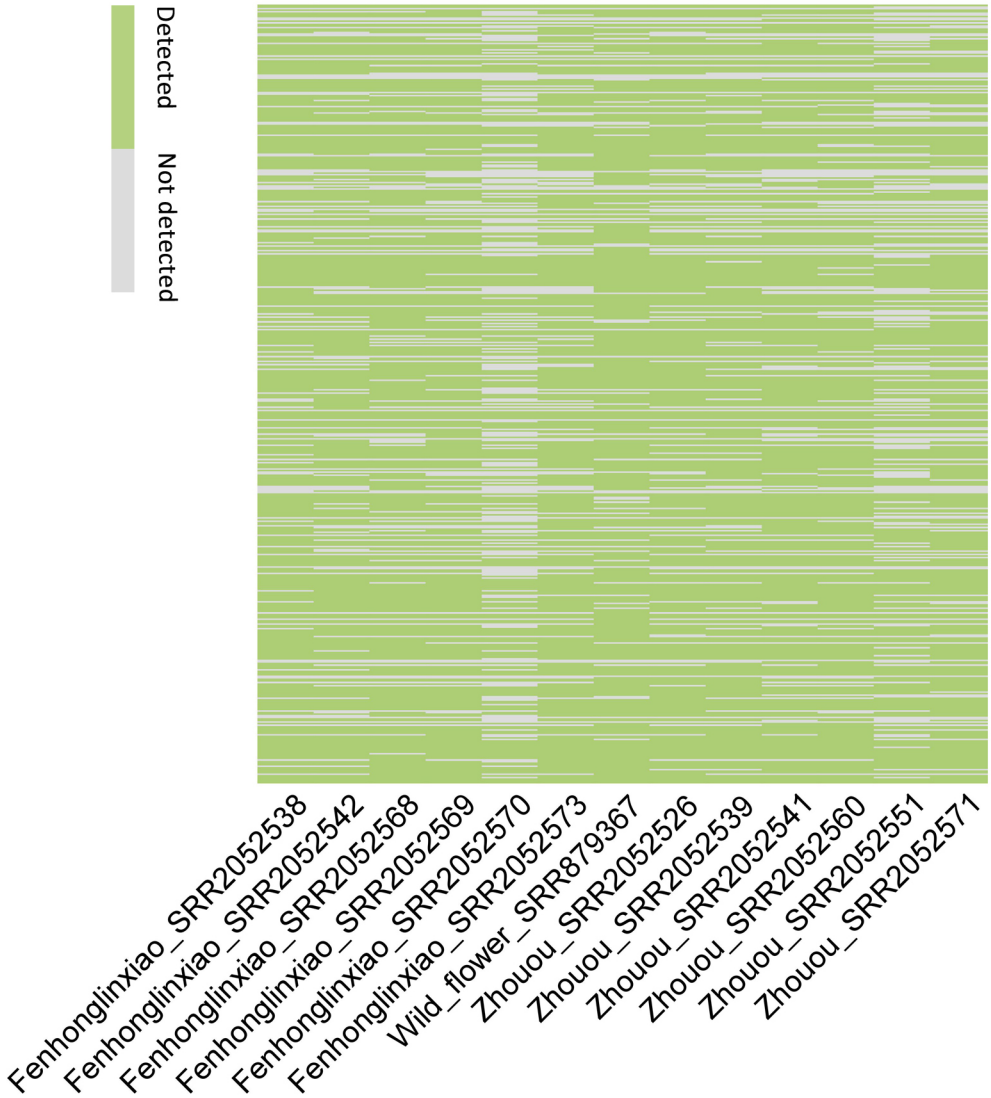
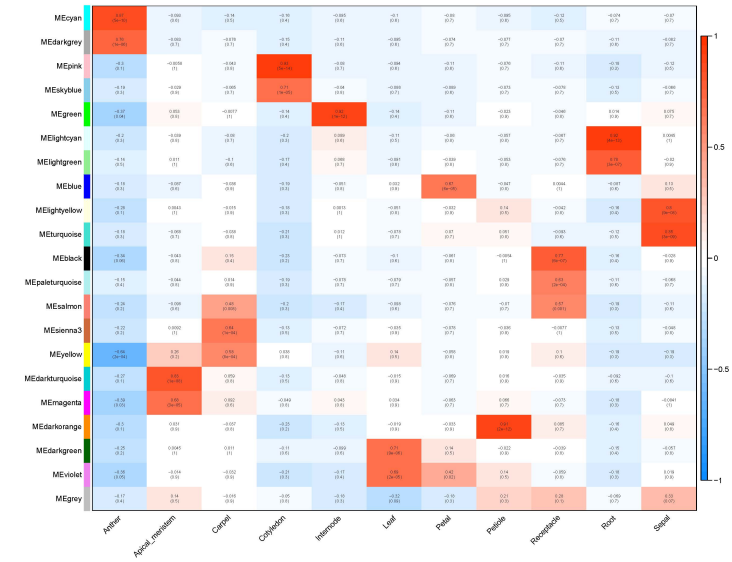
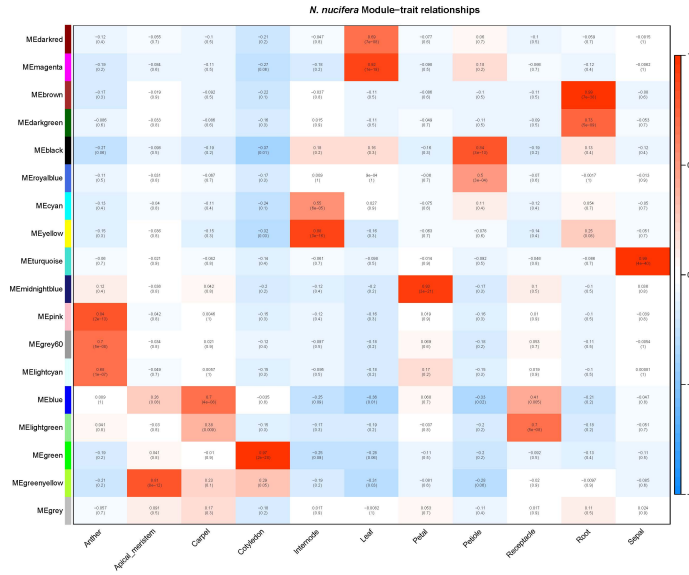


Figure S11. Figure S11. The heatmap of WGCNA module-tissue association in *N. lutea* (a), *N. nucifera* (b), and *Arabidopsis* (c). Each column corresponds to a tissue. Each row represents a module.Each cell indicates the association between the expression profile of isoforms in the corresponding module and the tissue type. This association is expressed as the correlation coefficient and p-value (both indicated in each of the matrix entries). The color corresponds to the value of the correlation coefficient.

a



b



c

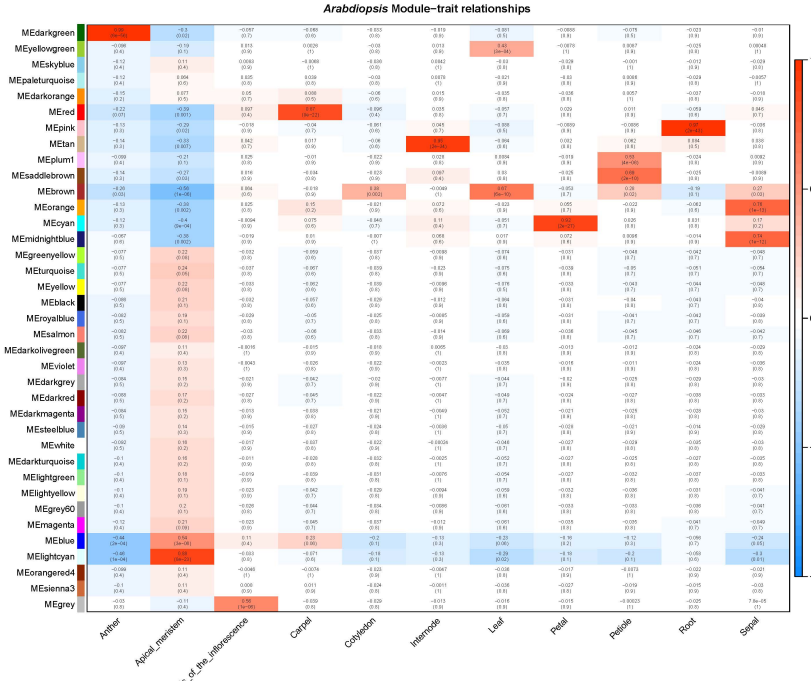


Figure S12. Distribution of the polymorphism value (PV) for the genes in *N. lutea*, *N. nucifera* and *Arabidopsis*.

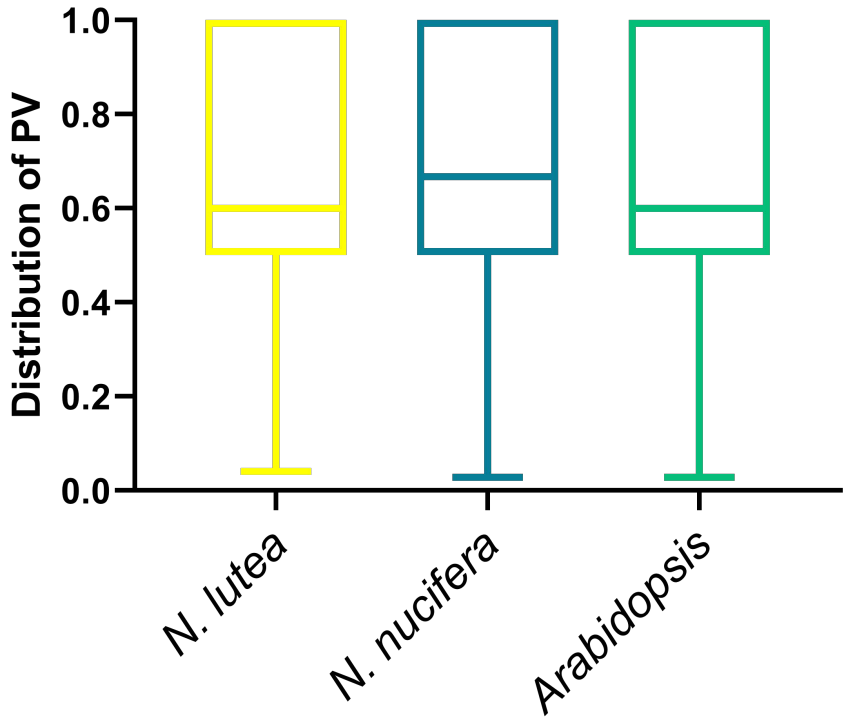


Figure S13. Distribution of the number of isoformsof duplicated genes displaying respectively single- or multiple- tissue bias expression patterns plots are shown for *N. lutea*, *N. nucifera*, and *Arabidopsis*. The significance of the difference in number of isoforms between duplicated genes with single- or multiple- tissue bias expression in three species were tested by the Mann-Whitney *U* test, ** means $p < 0.01$.

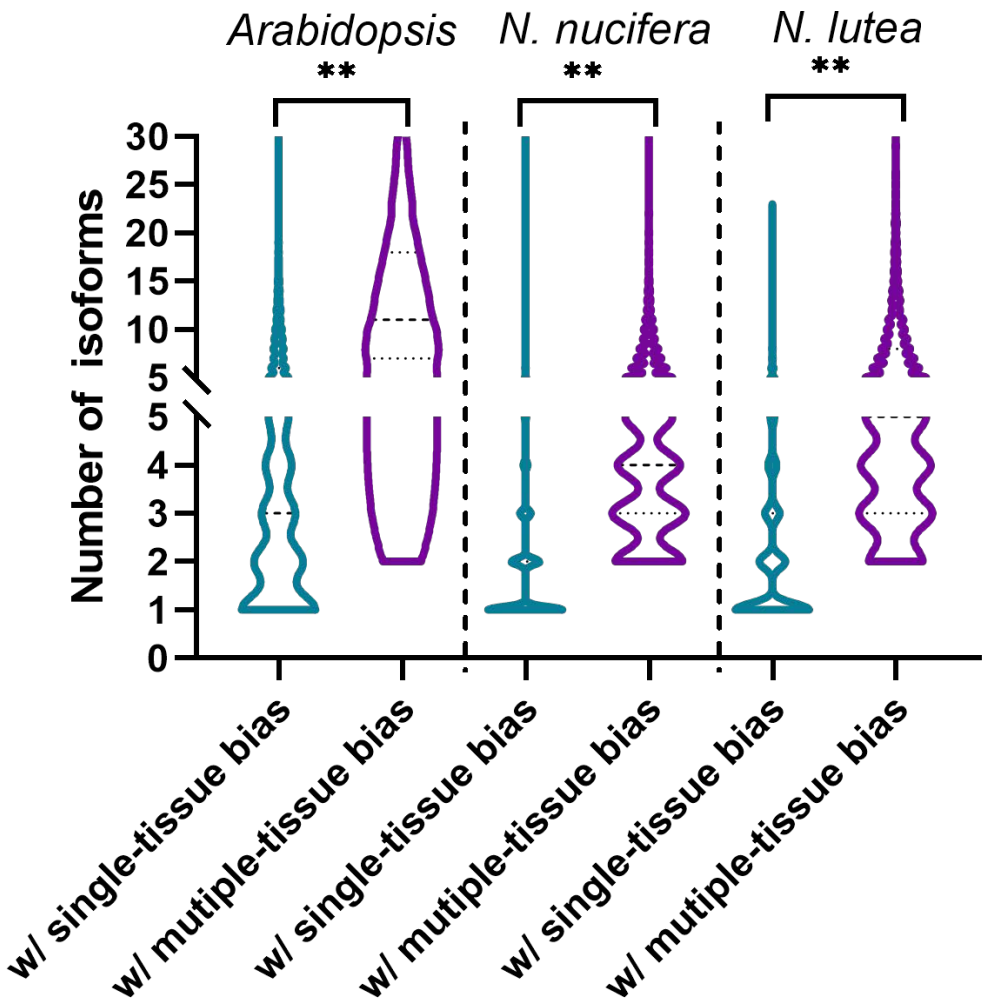


Figure S14. Verification of matching tissues between *N. lutea*, *N. nucifera*, and *Arabidopsis*. The heatmap shows the degree to which the tissue-specific expression vectors of orthologs correlate in respectively matching and non-matching tissues in the pairwise comparison between (a) *N. lutea* vs *Arabidopsis*, (b) *N. nucifera* vs *Arabidopsis*, and (c) *N. lutea* vs *N. nucifera*. The tissue-specific expression vectors contained the expression values of orthologous genes only. The color of each cell expresses the correlation in expression between matchingtissues of each of the two species. (d). The violin plot shows the distribution of the correlation coefficients obtained between matching tissues versus the distribution of the correlations observed when comparing tissue-specific expression vectors of non-matching tissues. Results of the pairwise comparisons of *N. lutea* vs *Arabidopsis*, *N. nucifera* vs *Arabidopsis*, and *N. lutea* vs *N. nucifera* are displayed. The significance was tested with the Mann-Whitney *U* test, * means $p < 0.05$.

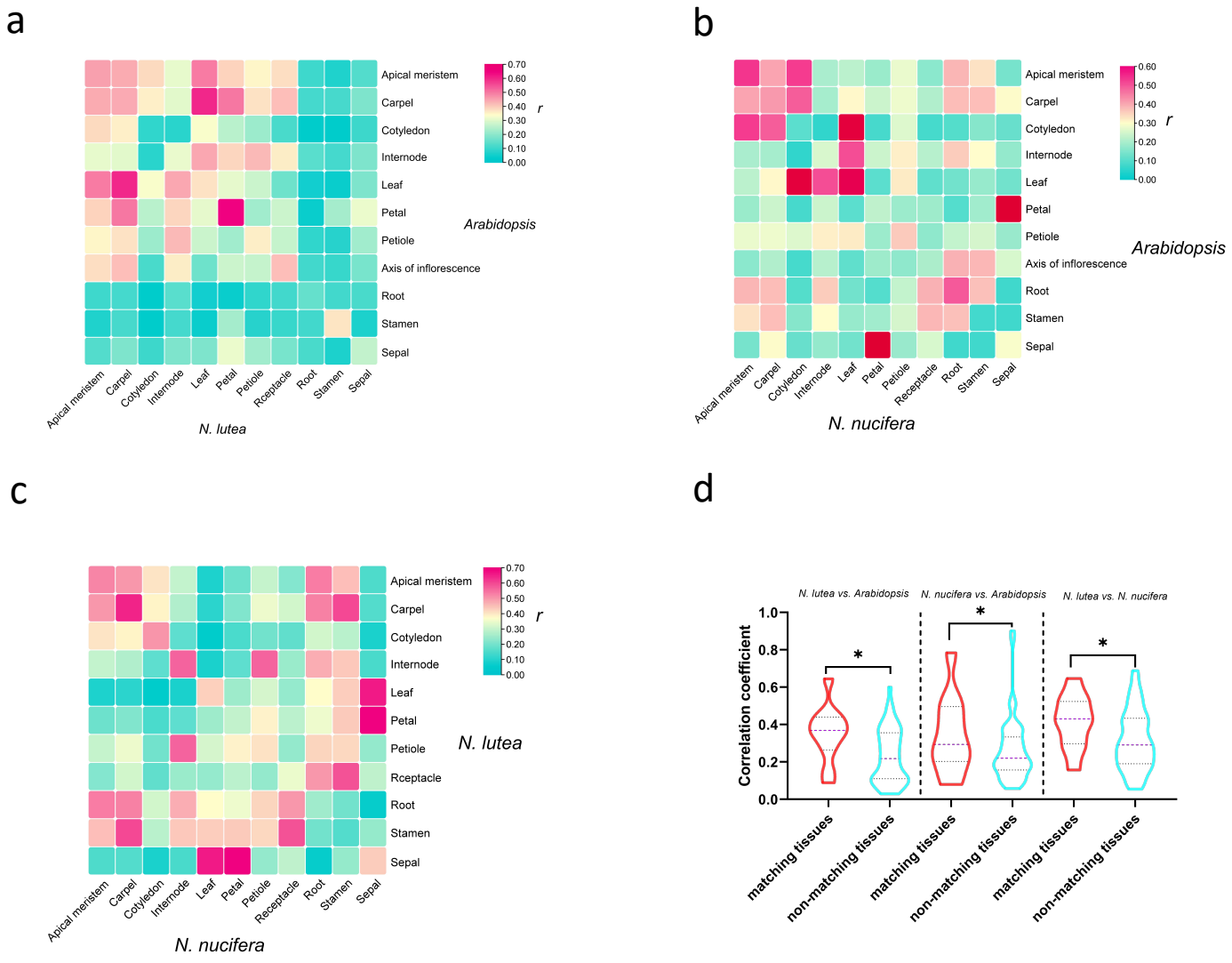


Figure S15. Conceptual illustration of genes with single- and multiple-tissue bias and the conserved tissue bias between orthologous genes. (a) If all isoforms of a gene were assigned to a module set representative of a single tissue, the gene was said to have a single tissue bias, otherwise if the isoforms of the gene were assigned to module sets representative of multiple tissues, this gene was sais to have a multiple-tissue bias. The blue hexagon and circles indicate respectively the gene and its corresponding isoforms. The parallelogram represents the module set representative of the tissue (i.e. tissue-specific module set). (b-c) Schematic representation of the orthologs (genes from the same orthologous group) of which the expression pattern shows a conserved single- (b) and multiple- (c) tissue bias in pairwise comparisons between *N. lutea*, *N. nucifera*, and *Arabidopsis*. The circles with different colors are isoforms from the equally colored orthologous genes. The parallelograms with the same color represent matching tissues in the three species. The expression of the orthologs in the red rectangle shows a conserved tissue bias.

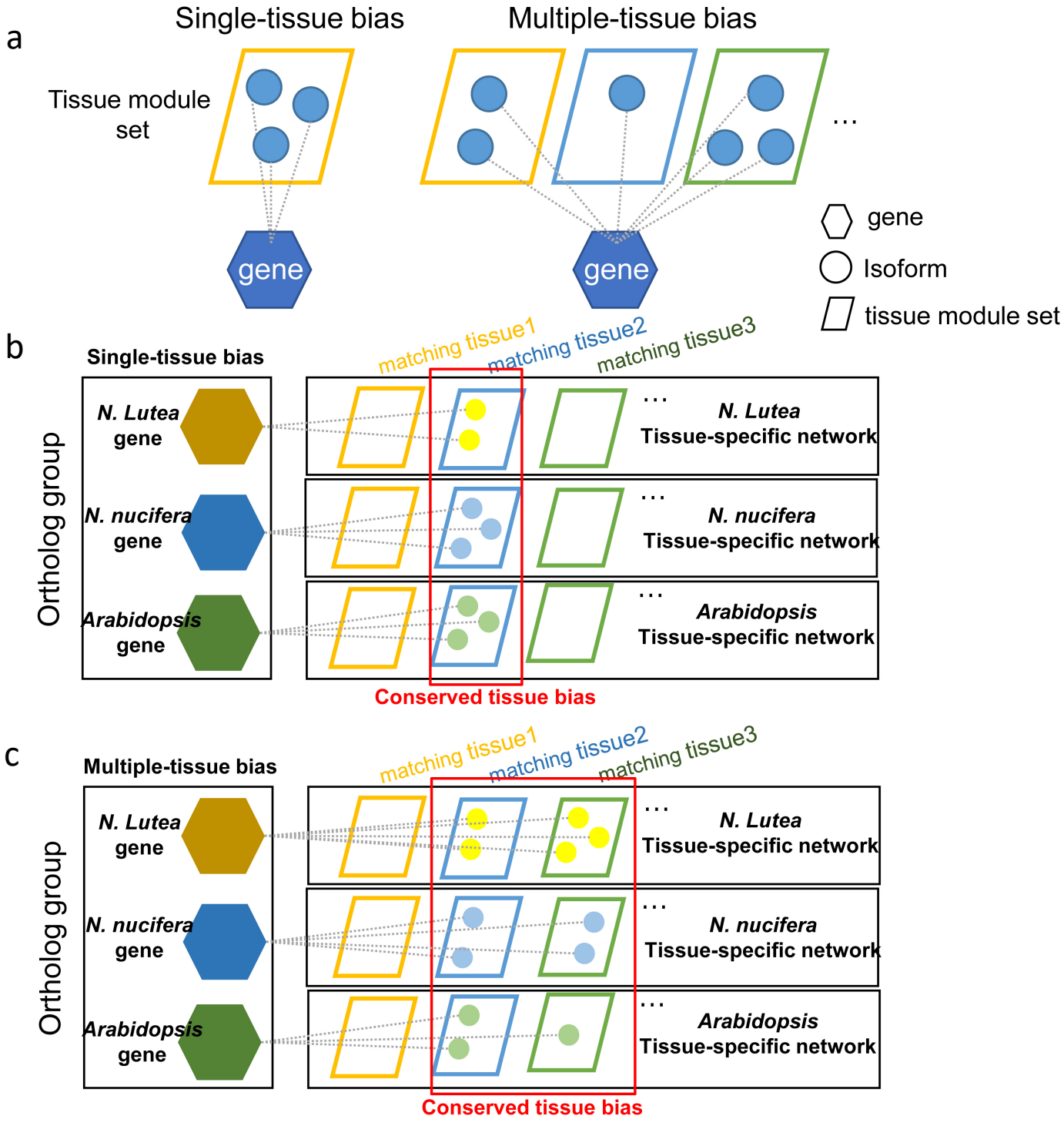
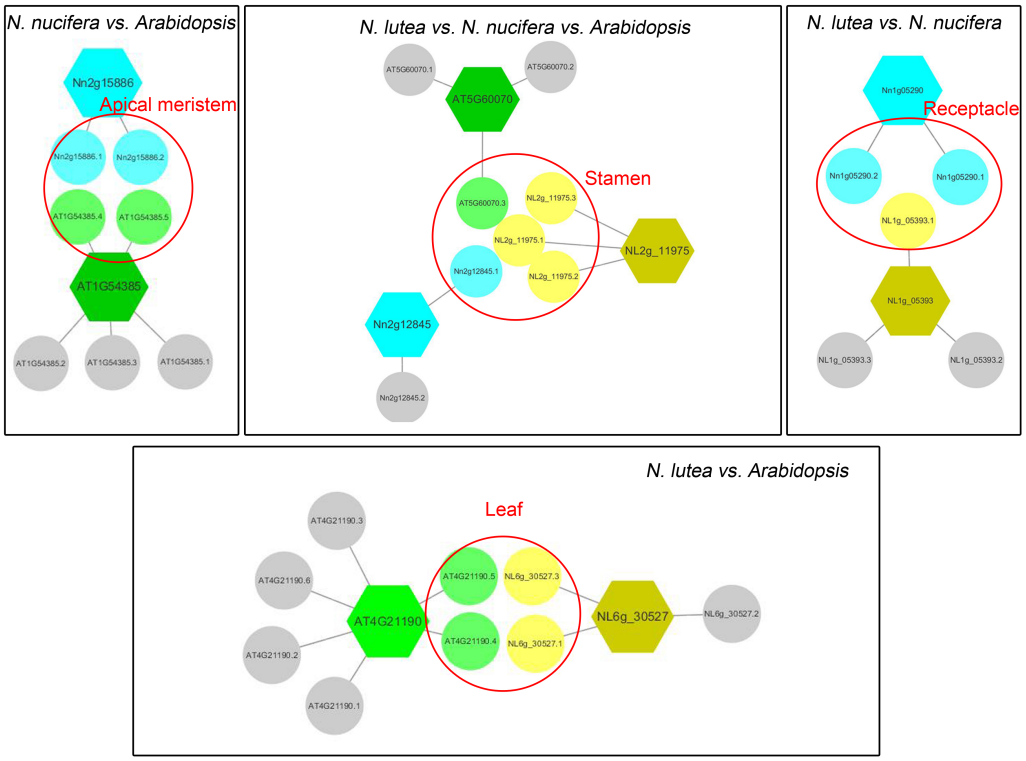


Figure S16. Examples of the conservation of tissue-specific expression for orthologous gene pairs that either show single-tissue (a) or multiple-tissue (b) biased expression. The hexagons represent the orthologous genes in *N. lutea* (yellow), *N. nucifera* (cyan), and *Arabidopsis* (green) and the circles represent the isoforms of the orthologs. Isoforms that display a bias in expression towards the same tissue (represented by the circle in red) are colored whereas isoforms that do not show this bias are in grey. Displaying a bias in expression to a tissue is assessed by checking whether the isoform was assigned to module set representative for that tissue.

a



b

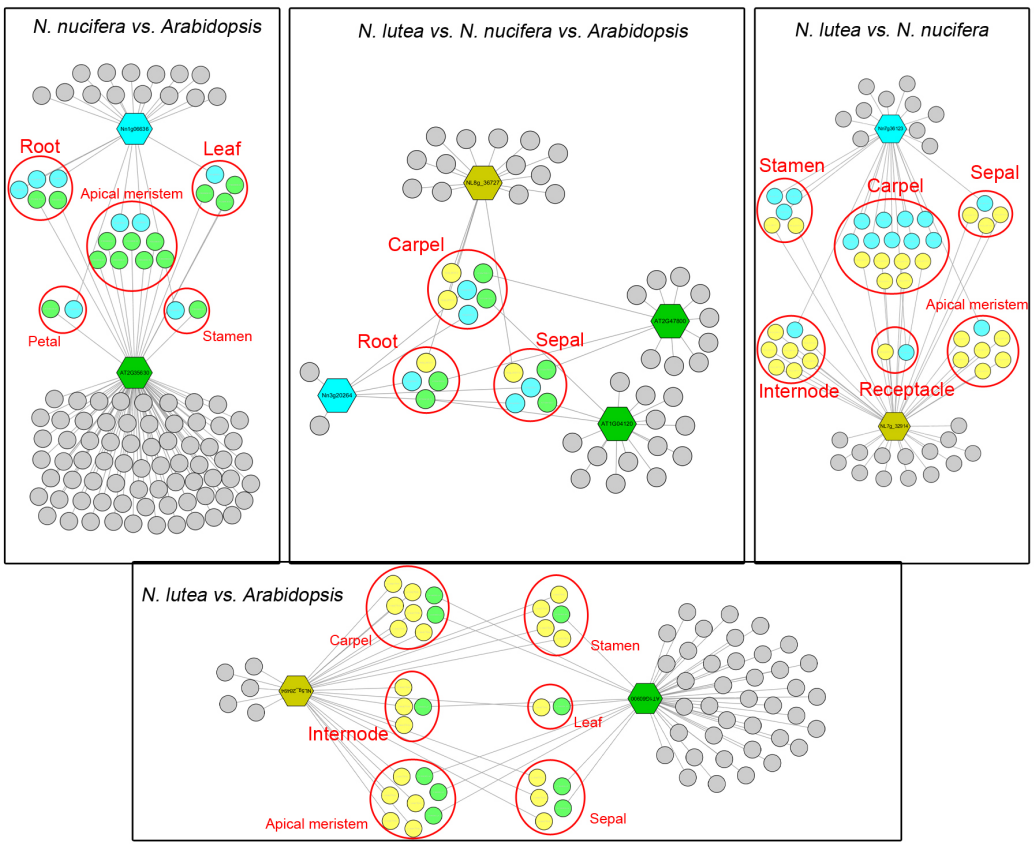


Figure S17. Phylogenetic analysis of MADS-box genes from the two Nelumbo species and *Arabidopsis*. The genes indicated with red, yellow, blue, and green rectangles are respectively A-class, B-class, C-class, and E-class genes in the 'ABCE' model.

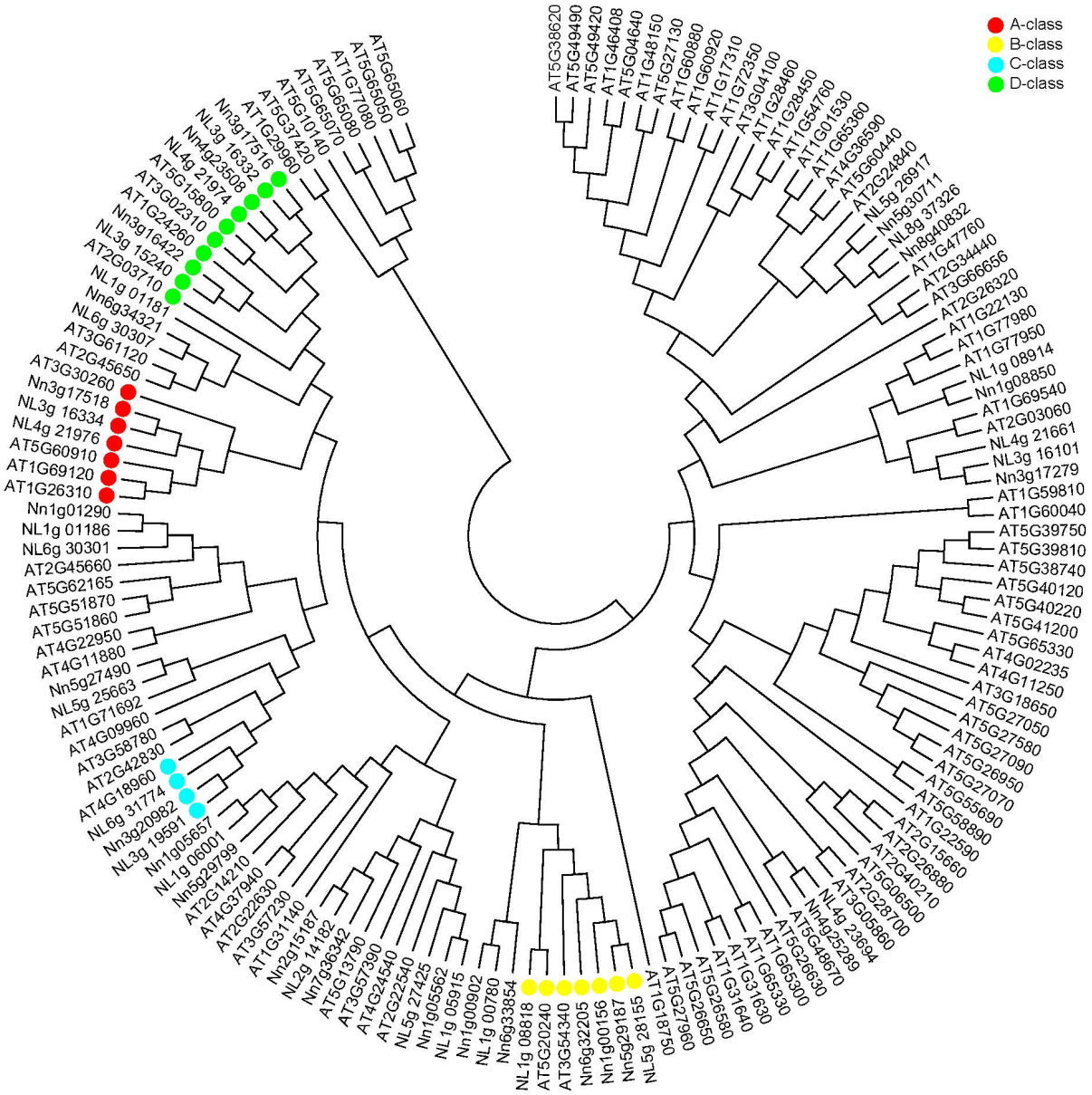


Figure S18. Tissue-specific module networks of isoforms for the 'ABCE' module in *Arabidopsis thaliana*. The filled squares in red (A), yellow (B), blue (C), and green (E) represent the 'ABCE' model genes. The colored circles (red, green, yellow, and blue) represent the isoforms of the genes belonging to distinct classes of the 'ABCE' model, and the grey circles represent isoforms not present in the tissue-specific expression modules. The circle's colors other than the ones indicated above are isoforms that are specifically expressed in non-floral tissue modules. The lines link genes and their corresponding isoforms. Different combinations of isoforms from the 'ABCE' model genes participate in the formation of different floral tissues.