

Figure S1. Seven families of repeats (R1-R7) of *Rhazya* mitochondrial genome involved in recombination. **A.** Illustration of corrected PacBio reads mapped against the assembled *Rhazya* mitochondrial genome. Blue and red lines indicate large and small repeat regions, respectively. Grey lines indicate regions of PacBio reads not aligned to the mitochondrial genome. **B.** Seven repeat families of *Rhazya* plotted on the mitochondrial genome. Blue lines within the circle indicate the positions of the pairs of repeats, twisted connecting lines denote repeats in reverse orientation. The letters A through F on the map indicated seven single-copy regions separated by repeats.

Figure S2. PCR strategy for identifying intramolecular recombination across mitochondrial repeats. Red lines on the outside of the master circle indicate the position and orientation of PCR primers to identify mitochondrial recombination, relative to master circle. Blue (direct repeats) or yellow (inverted repeats) lines on the inside of the master circle indicate the position and orientation of PCR primers to detect mitochondrial recombination; matched color (blue or yellow) lines on the outside of the eight alternative circles indicate alternative orientations of primers/products. Each repeat (R2; orange, R6; purple, and R7; green) or repeat cluster (R2-R3-R4; black) is indicated by an arrowhead on the circles and double-headed arrows show six possible alternative arrangements. Individual primer sequences are indicated by the numbers one through 12 and correspond to Table S12. The letters A through F on the maps indicate seven single-copy regions separated by repeats.

Figure S3. The two mitochondrial ORFs representing the nuclear (R)-mandelonitrile lyase sequence. **A.** Schematic diagram of the (R)-mandelonitrile lyase sequence in the *Rhazya* nucleus (nDNA; red line), the mitochondrial fragments of 247 and 1,894 bp, with the two ORFs (blue block arrows) in the *Rhazya* mitochondria (mtDNA; blue line), and the transcript (orange). A TE (transposable element) insertion is shown in red between the two fragments in the *Rhazya* mitochondrial DNA. The model (malate:quinone oxidoreductase; *Mqo* and

glucose-methanol-choline (GMC) oxidoreductase; *GMC_oxred_C*) of a putative nuclear protein-coding gene was predicted using the Conserved Domain Architecture Retrieval Tool (CDART) [82]. **B.** The broken outlines represent the full-length nuclear sequence, the red shaded arrow within indicates the proportion of the sequence present in six asterid mitochondrial genomes with the length of each indicated above. **C.** Nucleotide sequence alignment of the nuclear and mitochondrial (ORF185 and ORF430) copies and the mitochondrial transcript from *Rhazya stricta*. Red box shows 5 bp deletion that caused a frame shift to generate split ORFs. **D.** Amino acid sequence alignment of nuclear and two mitochondrial ORFs. Each ORF (blue; ORF185, red; ORF430) contains a conserved domain (*Mqo*; blue shaded box and *GMC_oxred_C*; red shaded box).

Figure S4. Potential chimeric ORFs. Schematic diagram of the genomic regions surrounding the ORFs containing repeats. **A.** ORF56b overlaps with repeat region 6 (R6) associated with *rpl2* and contains small fragments of three mitochondrial genes (*rpl2*, *matR*, and *ccmFn*). **B.** ORF318 and ORF324 were identified in repeat cluster regions associated with *atp6*. Repeats are numbered and indicated by solid dark blue arrows and are diagrammed on the mitochondrial genome map. The maps are broken (-//-) to focus on regions of interest. ORFs are indicated by light blue block arrows. Red bands on the ORFs indicate the transmembrane helices regions.

Figure S5. Conserved mitochondrial gene blocks among eight asterids. All pairwise comparisons were made relative to *R. stricta*. All protein-coding genes and rRNAs are arranged from top to bottom for *Rhazya*, and from left to right for the other asterids based on the order in published genome maps. Red genes names indicate losses of each mitochondrial genome. Blue squares represent the presence of a given gene in each genome. Red boxes indicate colinear gene blocks between two genomes. Shaded red boxes indicate shared colinear gene blocks among the eight asterid genomes.

Figure S1

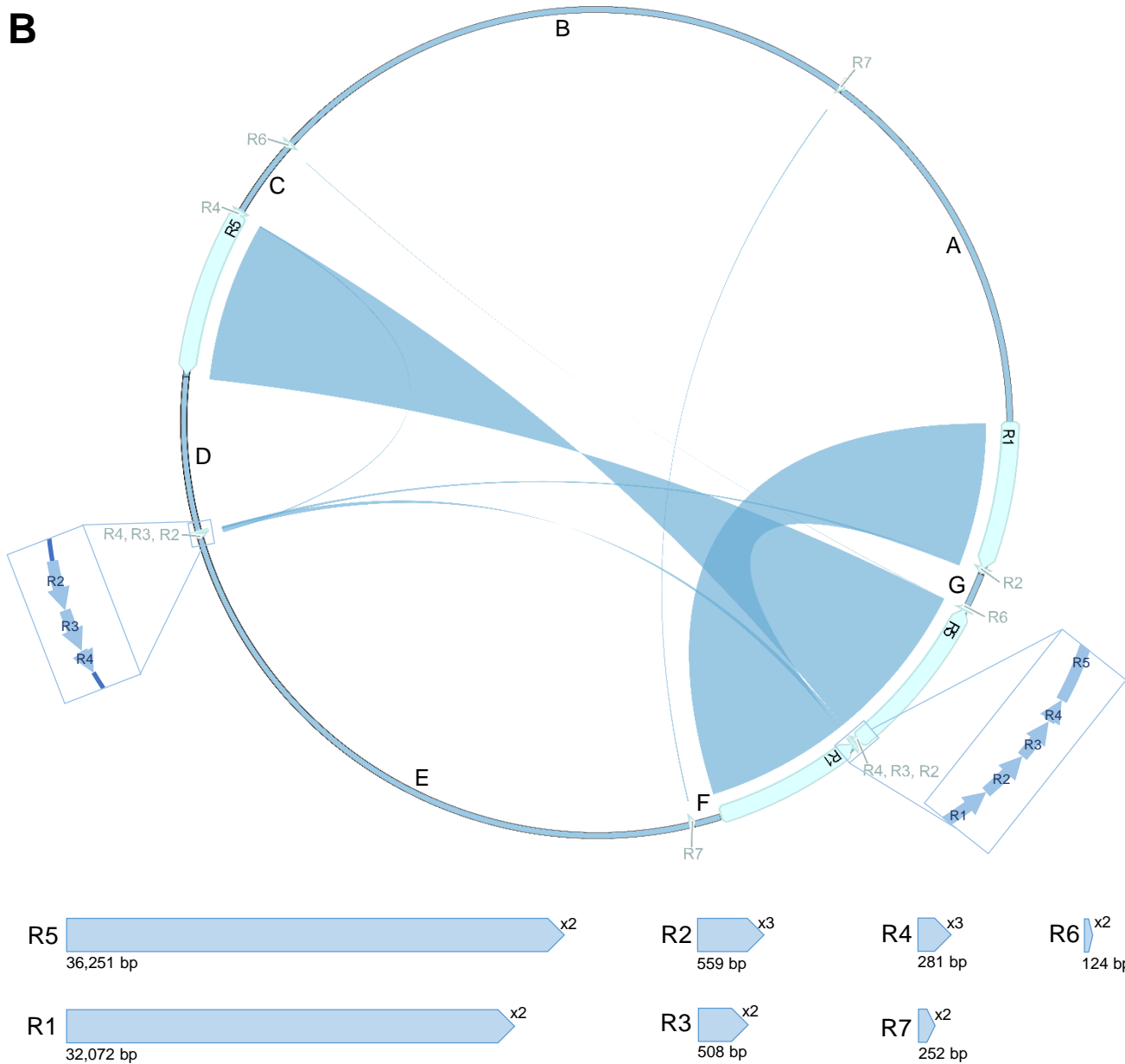
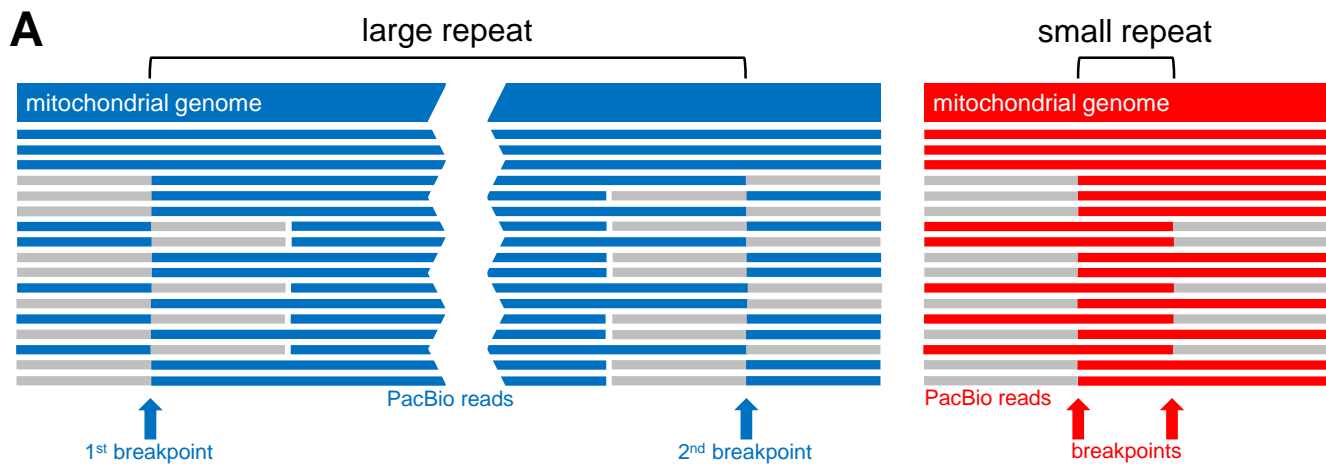


Figure S2

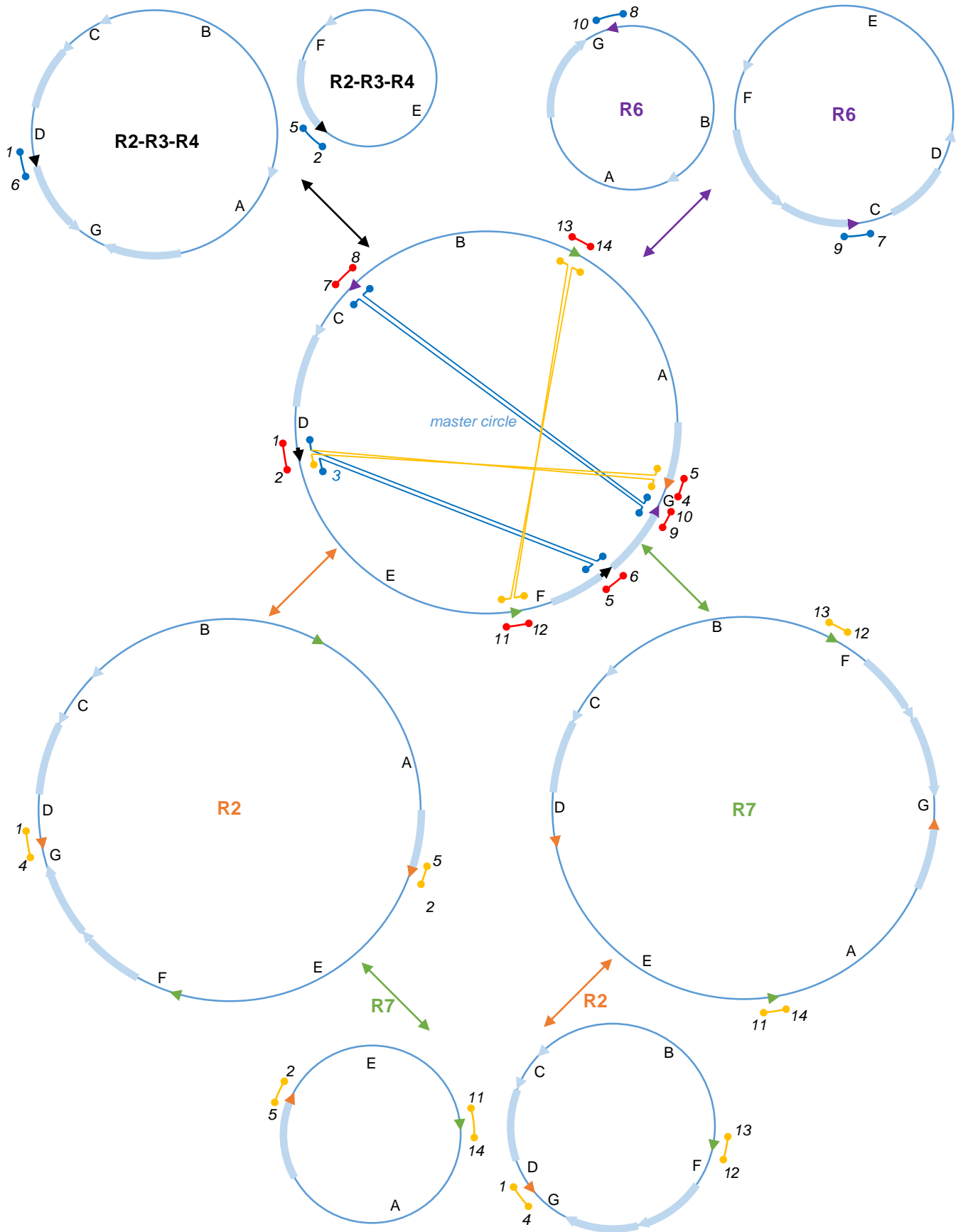


Figure S4

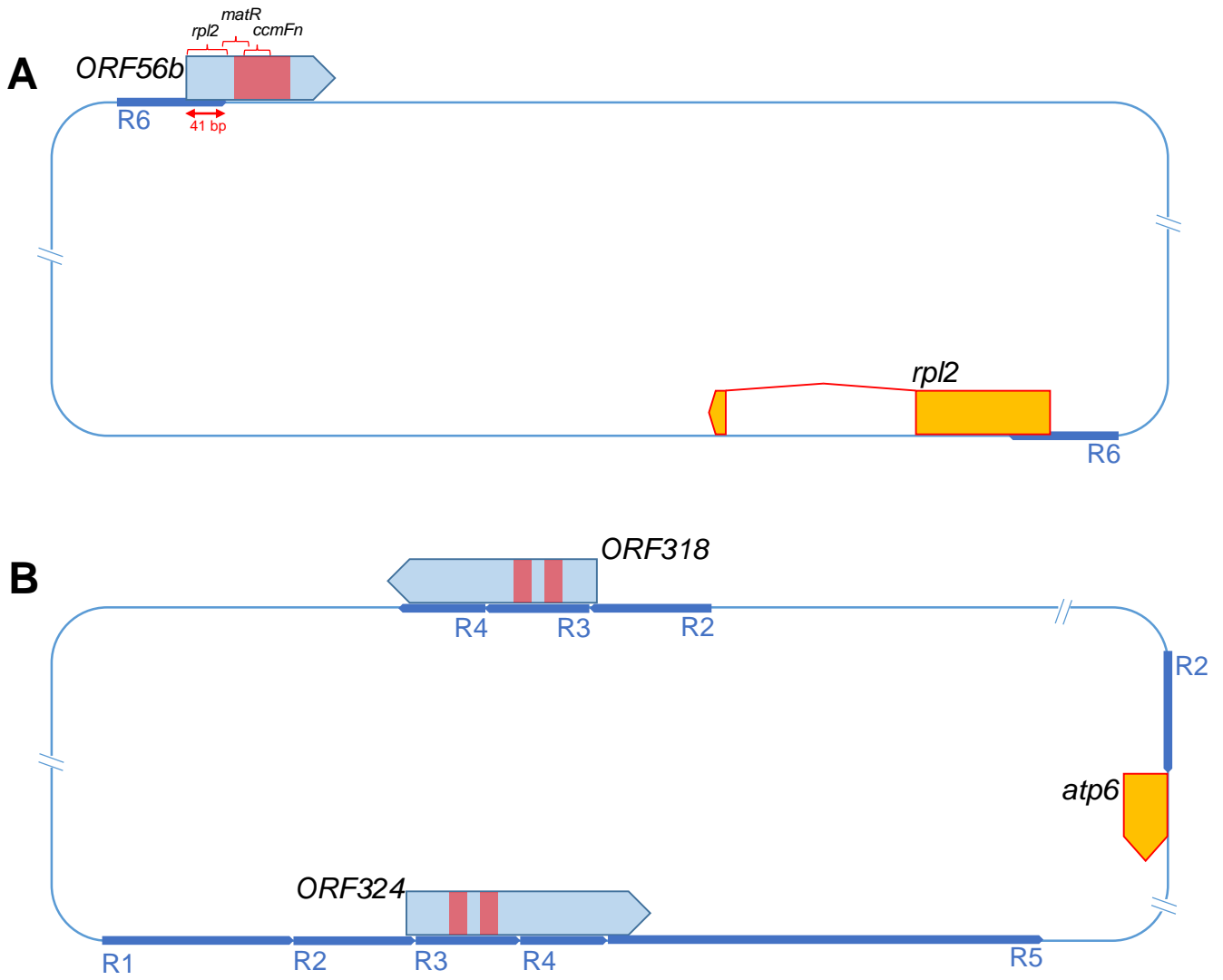
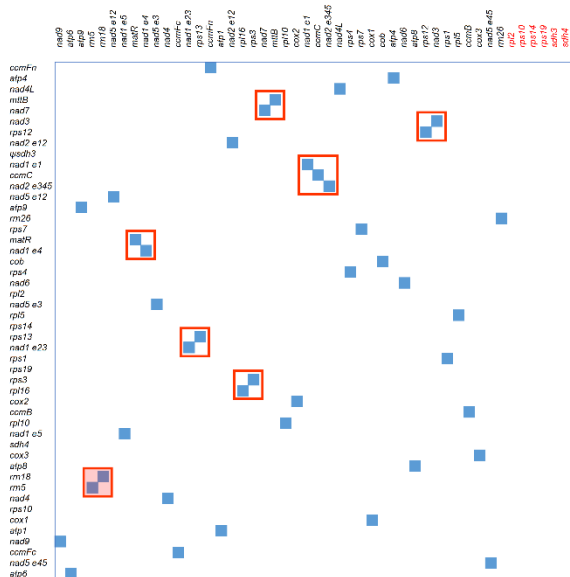


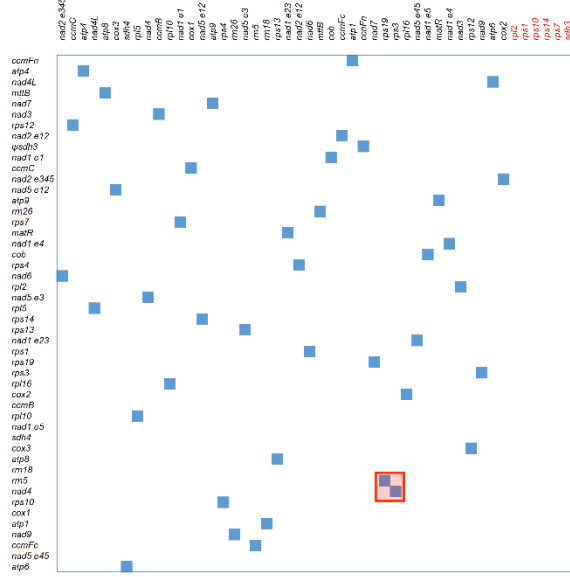
Figure S5

Daucus

Rhazya



Helianthus



Vaccinium

Rhazya

