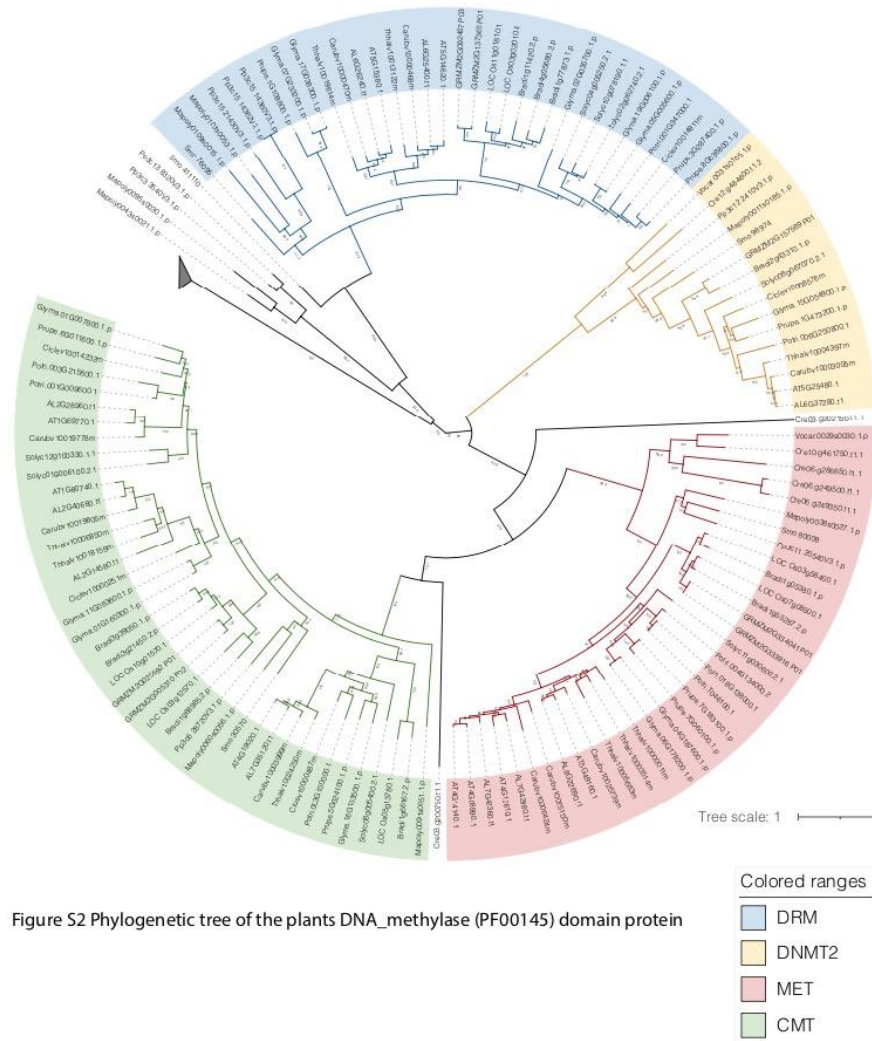


Figure S1 Phylogenetic tree of the plants SAD\_SRA (PF02182) domain protein

**Figure S1.** Circular phylogram of plant SAD\_SRA (PF02182) domain proteins. The different clades are colored as indicated in the associated legend. The bootstrap number is represented on each branch. SAD\_SRA (PF02182) domain protein sequences were downloaded from Phytozome [71] with a Pfam domain filter (SAD\_SRA, PF02182). We restricted our analysis to the following representative species: *A. lyrata*, *A. thaliana*, *B. distachyon*, *C. rubella*, *C. reinhardtii*, *C. clementina*, *E. salsugineum*, *G. max*, *M. polymorpha*, *O. sativa*, *P. patens*, *P. trichocarpa*, *P. persica*, *S. moellendorffii*, *S. lycopersicum*, *V. carteri*, *Z. mays*. The sequences were cleaned using CD-HIT suite with a sequence identity cut-off of 1 and length of the sequence to skip 300 [100] and further manually curated to remove protein variants (i.e., only one representative protein per locus). Alignment and phylogenetic reconstructions were performed using the function “build” of ETE3 v3.0.0b32 [101] as implemented on the GenomeNet (<https://www.genome.jp/tools/ete/>, accessed on 18 July 2021) using default parameters. The mid-point rooted tree was exported in a Newick format. Visualization and illustration were performed using iTOL v6.1.2 [102].



**Figure S2.** Circular phylogram of plant DNA methylase (PF00145) domain proteins. The different clades are colored as indicated in the associated legend. The bootstrap number is represented on each branch. DNA methylase (PF00145) domain protein sequences were downloaded from Phytozome [71] with a Pfam domain filter (DNA methylase, PF00145). We restricted our analysis to the following representative species: *A. lyrata*, *A. thaliana*, *B. distachyon*, *C. rubella*, *C. reinhardtii*, *C. clementine*, *E. salsgineum*, *G. max*, *M. polymorpha*, *O. sativa*, *P. patens*, *P. trichocarpa*, *P. persica*, *S. moellendorfii*, *S. lycopersicum*, *V. carteri*, *Z. mays*. The sequences were cleaned using CD-HIT with a sequence identity cut-off of 1 and length of the sequence to skip 300 [100] and further manually curated to remove protein variants (i.e., only one representative protein per locus). Alignment and phylogenetic reconstructions were performed using the function “build” of ETE3 v3.0.0b32 [101] as implemented on the GenomeNet (<https://www.genome.jp/tools/ete/>, accessed on 18 July 2021) using default parameters. The mid-point rooted tree was exported in a Newick format. Visualization and illustration were performed using iTOL v6.1.2 [102]