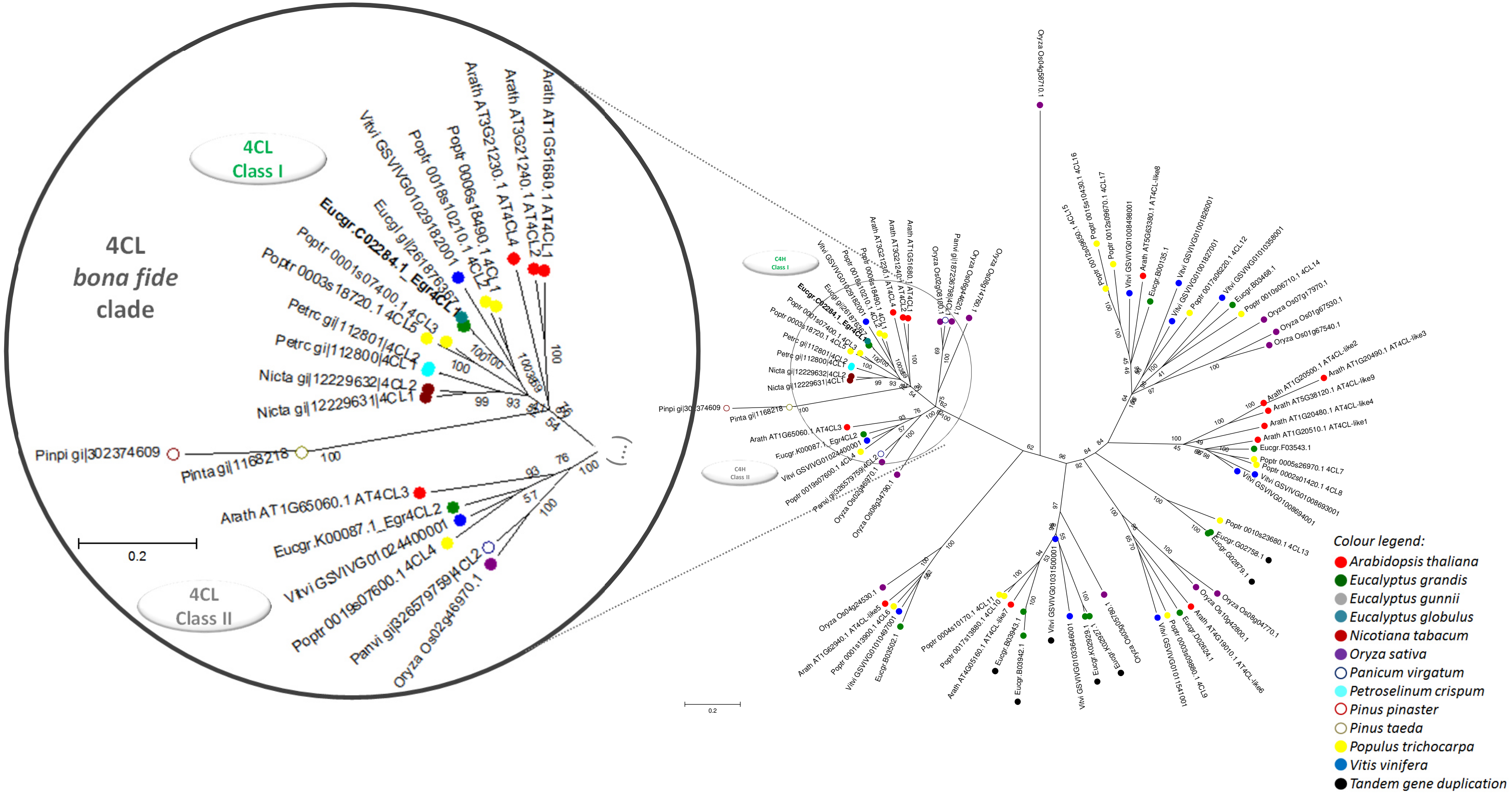
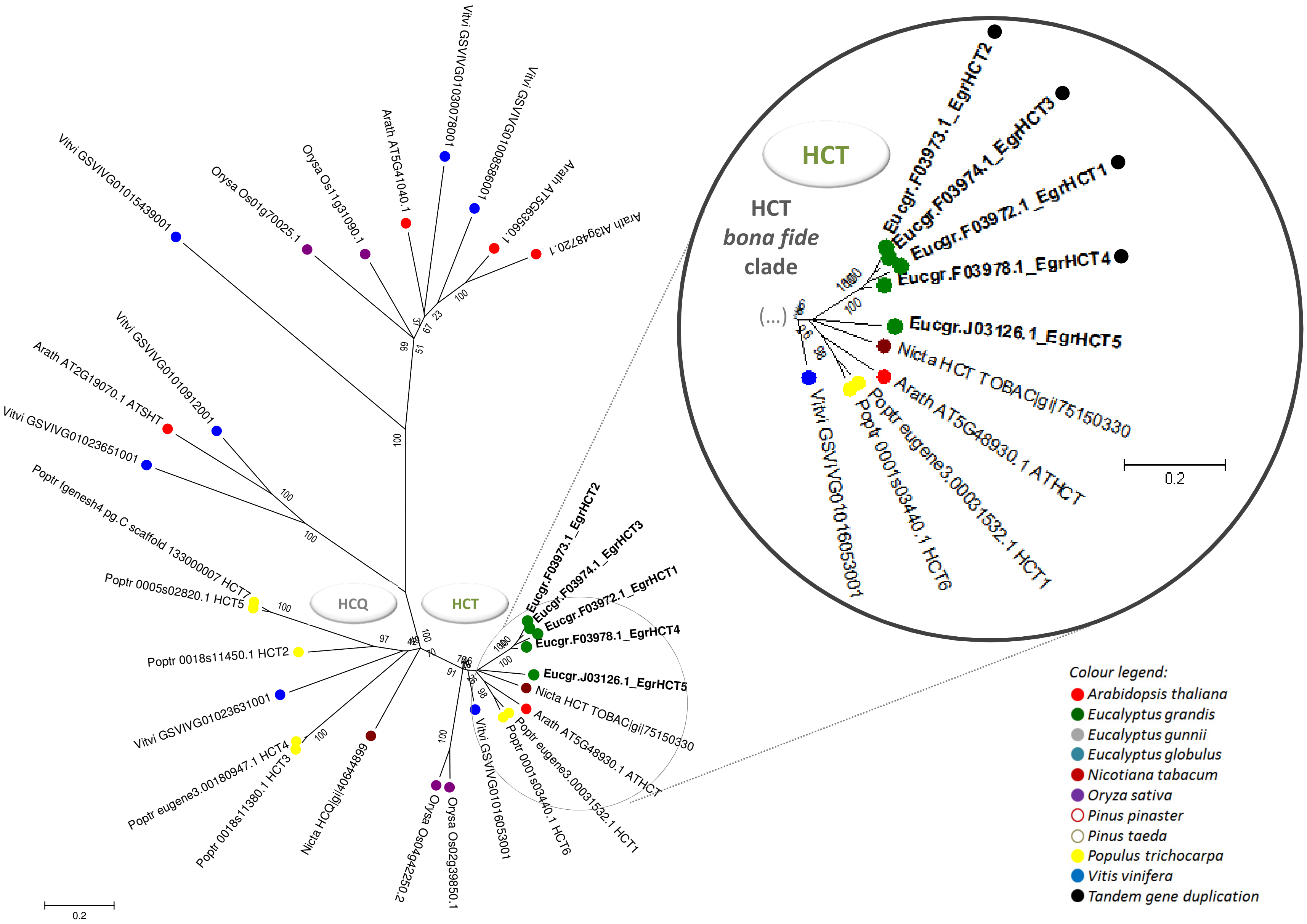
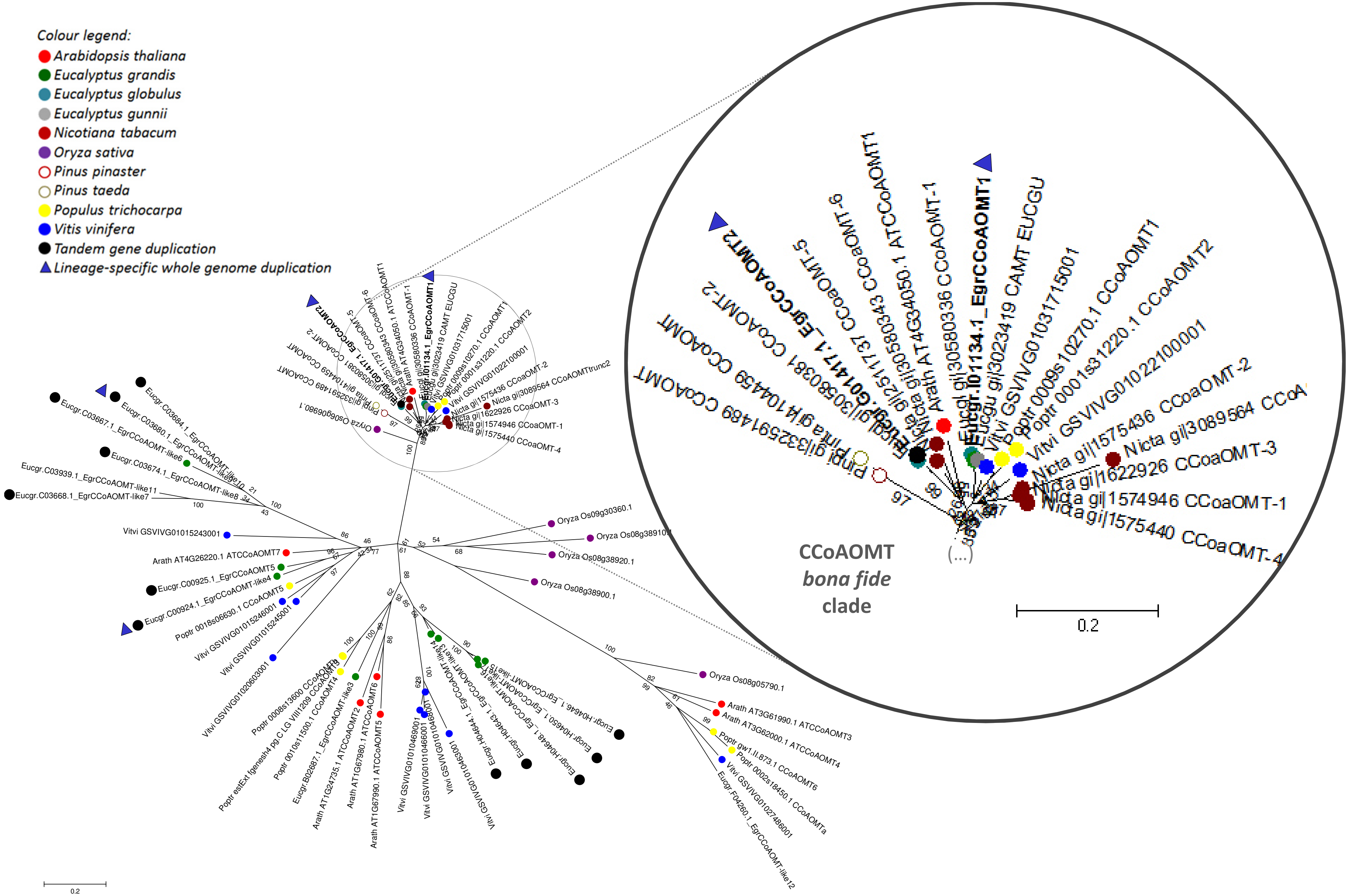
Supporting Information Figs S1–S7



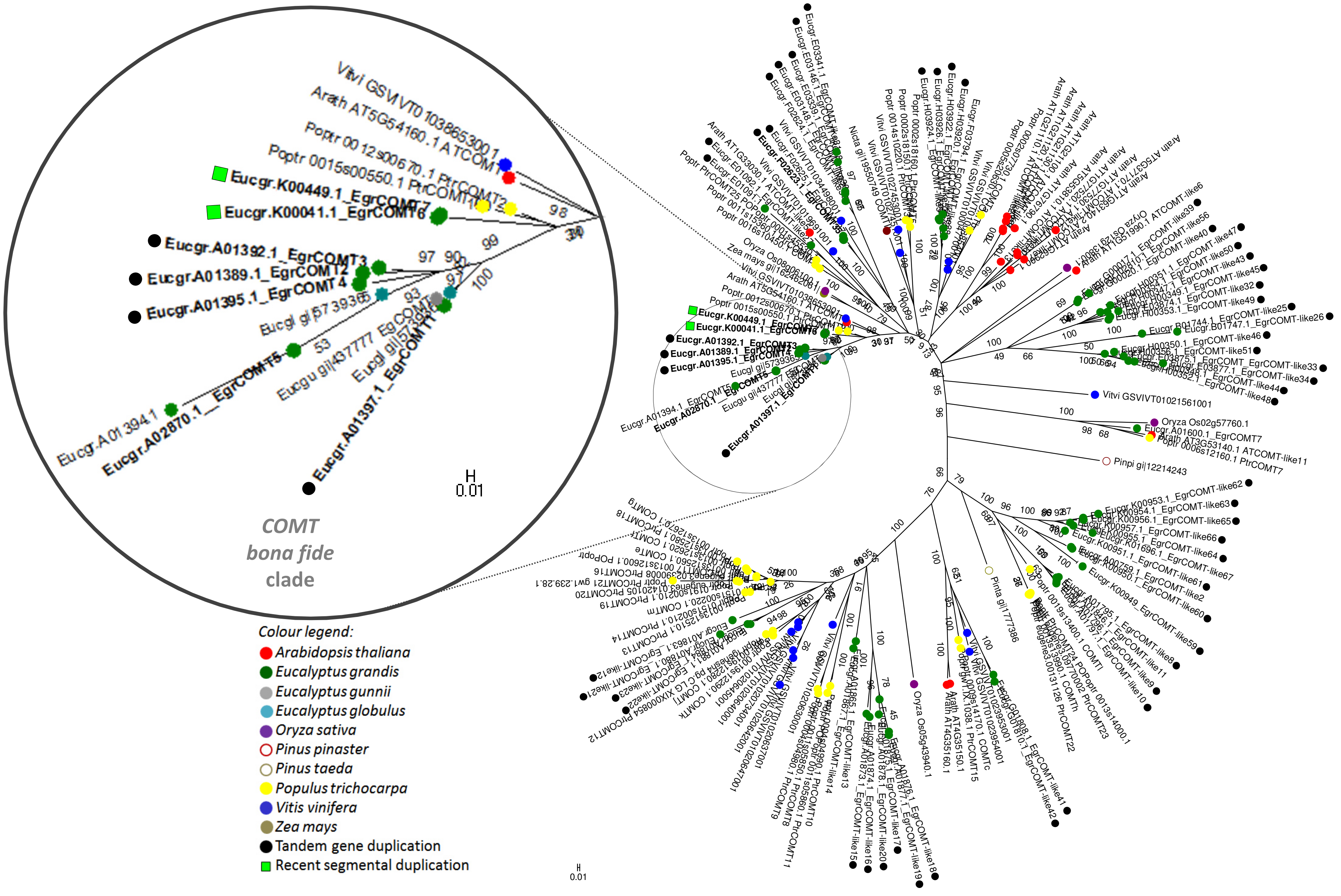
Supplementary Figure S1. Unrooted protein phylogenetic tree of the 4CL (4-coumarate:CoA ligase) multigene family. Large comparative phylogenetic analysis generated for the two Egr4CL proteins, eleven Egr4CL-like proteins and 64 related proteins from other plant species. A total of 983 non-ambiguous amino acid positions were considered in final dataset. Subclade containing the *bona fide* genes is highlighted.



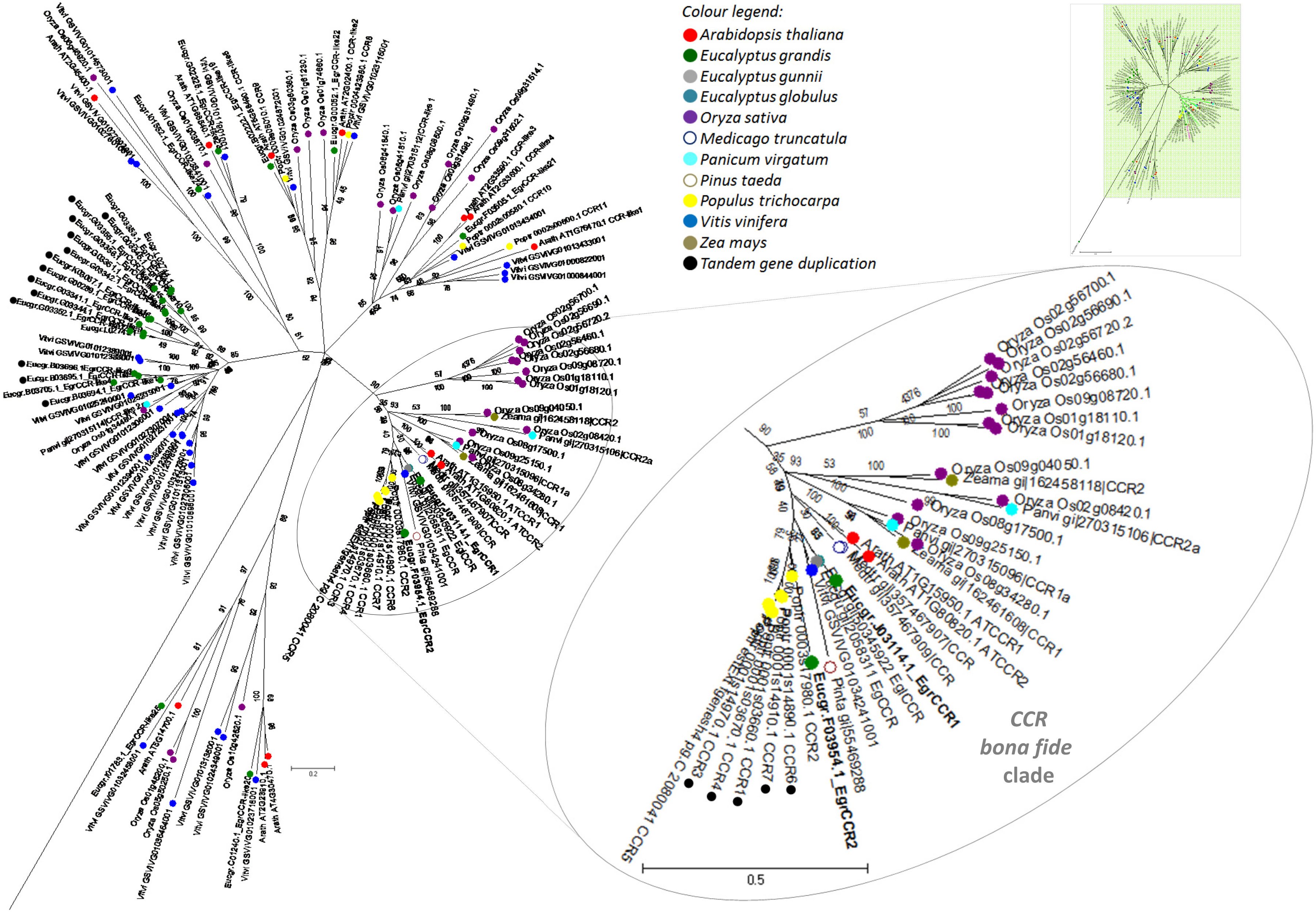
Supplementary Figure S2. Unrooted protein phylogenetic tree of the HCT/HCQ (hydroxycinnamoyl-CoA:shikimate / quinate hydroxycinnamoyltransferase) multigene family. Large comparative phylogenetic analysis generated for the five EgrHCT proteins and 25 related proteins from other plant species. A total of 121 non-ambiguous amino acid positions were considered in the final dataset. Subclade containing the *bona fide* genes is highlighted.



Supplementary Figure S3. Unrooted protein phylogenetic tree of the CCoAOMT (caffeoyl CoA 3-O-methyltransferase) multigene family. Large comparative phylogenetic analysis generated for the two EgrCCoAOMT proteins, fifteen EgrCCoAOMT-like and 44 related proteins from other plant species. A total of 70 non-ambiguous amino acid positions were considered in the final dataset. Subclade containing the *bona fide* genes is highlighted.

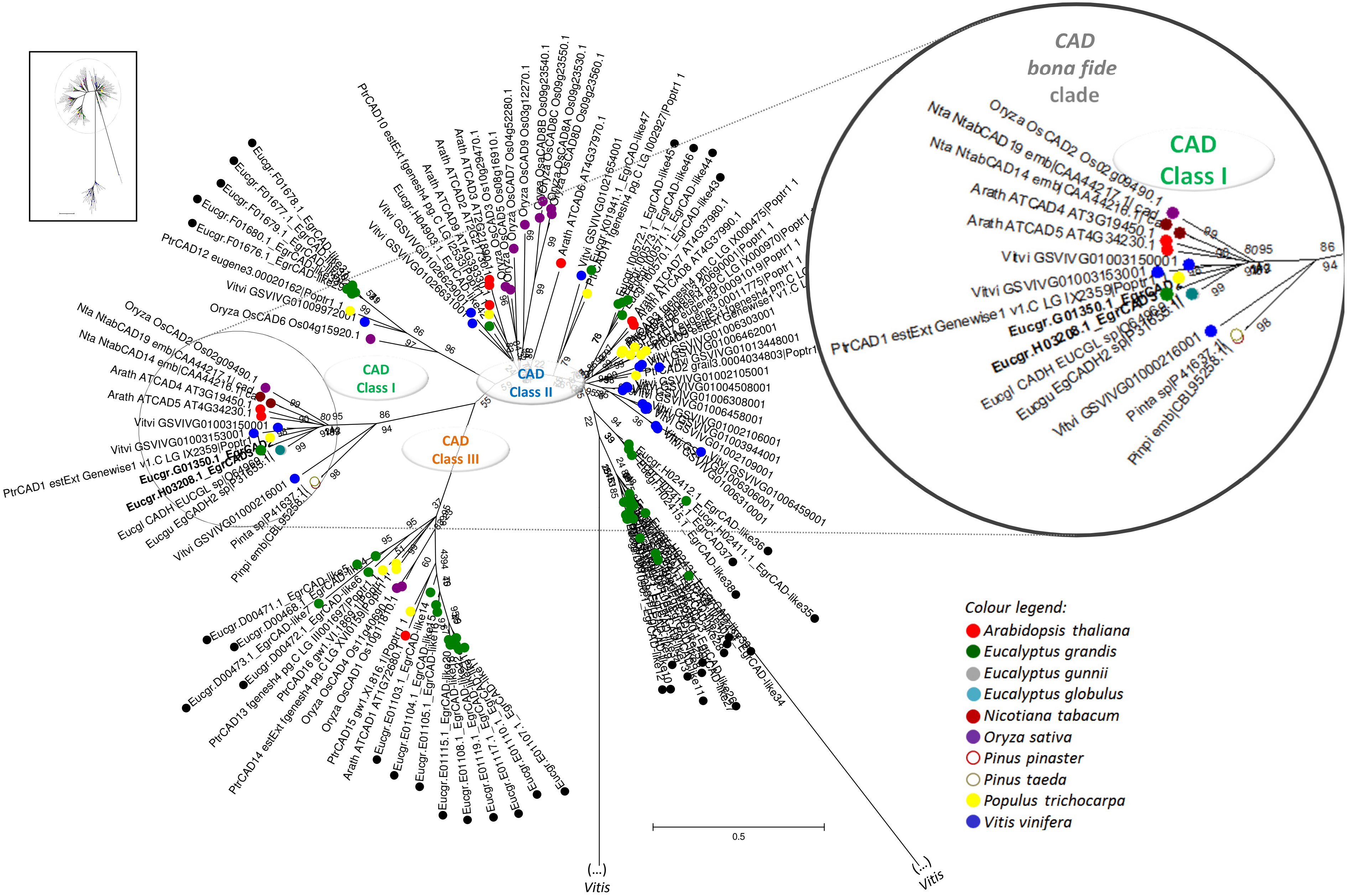


Supplementary Figure S4. Unrooted protein phylogenetic tree of the COMT (caffeic acid/5-hydroxyferulic acid O-methyltransferase) multigene family: Large comparative phylogenetic analysis generated for the seven EgrCOMT proteins, 60 EgrCOMT-like and 84 related proteins from other plant species. A total of 901 non-ambiguous amino acid positions were considered in the final dataset. Subclade containing the *bona fide* genes is highlighted.

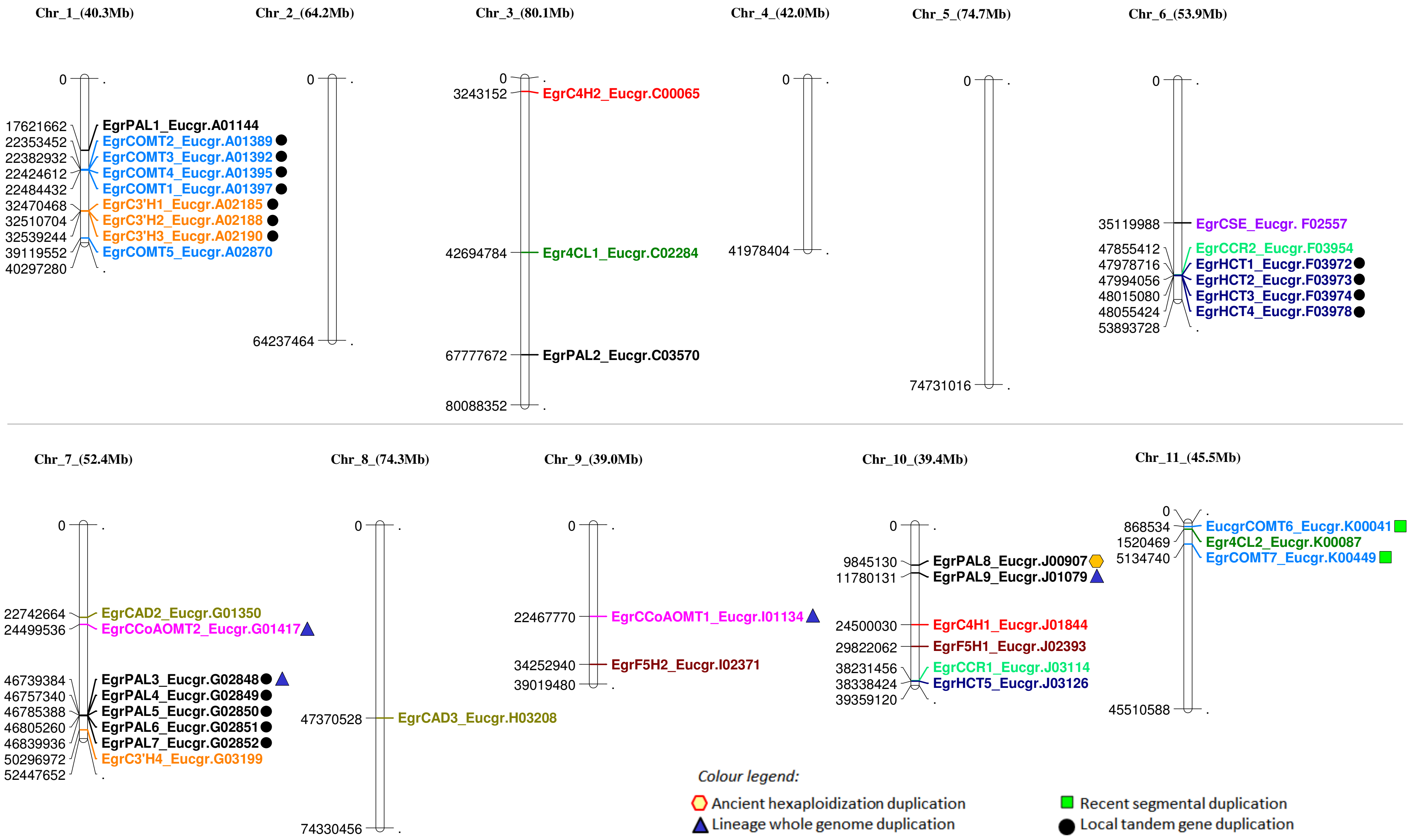


Supplementary Figure S5. Unrooted protein phylogenetic tree of the CCR (cinnamoyl CoA reductase) multigene family

Large comparative phylogenetic analysis generated for the two EgrCCR proteins, 25 EgrCCR-like and 95 related proteins from other plant species. A total of 122 non-ambiguous amino acid positions were considered in the final dataset. Subclade containing the *bona fide* genes is highlighted.



Supplementary Figure S6. Unrooted protein phylogenetic tree of the CAD (cinnamyl alcohol dehydrogenase) multigene family Large comparative phylogenetic analysis generated for the two EgrCAD proteins, 44 EgrCAD-like and 95 related proteins from other plant species. A total of 881 non-ambiguous amino acid positions were considered in the final dataset. Subclade containing the *bona fide* genes is highlighted.



Supplementary Figure S7. Physical position of the 38 *Eucalyptus grandis* genes involved in the *Eucalyptus grandis* phenylpropanoid and lignin branch biosynthetic pathway. The physical position of the putative *bona fide* lignification genes was verified. For this, the physical distances were retrieved from BioMart for each gene and graphical display was performed using MapChart 2.2 (Voorips 2002).