

Supporting Information

Zhao et al. 10.1073/pnas.1009170107

SI Materials and Methods

Constructs for Transfection Studies. SND1 (At1g32770), MYB46 (At5g12870), and MYB58 (At1g16490) were cloned into the vector p2GW7 (1). The fragment (1-3414) of the vector pPGWL7 (1), which contains the firefly luciferase ORF was transferred to p2GW7 to generate a small high-copy vector with the firefly luciferase sequence. The promoters (\approx 1-kb regions) of F5H, C4H, COMT, and CCoAOMT were cloned into pENTR vector and then switched into the modified vector. The primers used for promoter cloning were: ProF5H-Forward 5'-CACCGATTCTCACATTTGTTTTATATATATG-3, ProF5H-Reverse 5'-GTTAACCTTGTTGTTTCTATTTTTGGTTTC-3; ProC4H-Forward 5'-CACCTCTAATTGAATAGCATGTTACCCGGTCCAACC-3, ProC4H-Reverse 5'-TGTTGAGTGTGTTTTTGTGTTTTGGGTAATGAAA-3; ProCOMT-Forward 5'-CACCGGTCACGTACATTTTGGTTGTGTGACCATATTTTTC-3, ProCOMT-Reverse 5'-GGT-

TGTGATTGATGAGGTTTTTGAGATTGTAGTGGTG-3; and ProCCoAOMT-Forward 5'-CACCGCATCAAAAACATGTCAAAACACTAGCAATTCATC-3, ProCCoAOMT-Reverse 5'-TGTTAAATCTTGGGAATGTTTTGCTTCTTTTTCTG-3.

F5H Promoter Fragments for Competiton EMSA Assays. P1 to P6 are 50-bp fragments from -349 to -100 of the F5H promoter region, with 10-bp overlaps between each other. Their sequences are:

- P1, 5'-atcaagcatatcaaacgtgtgtaaatatctattctatctctcaaatctca-3';
- P2, 5'-ctcaaatctcaatgttacgTTTTTctcctcaaatgctttattagttgcgt-3';
- P3, 5'-ttagttgcgtccacttgcaatataataatctattgatacatatTTTT-3';
- P4, (5'-catatTTTTcttaaccaagtttacactctcaagcaactattatata-3';
- P5, 5'-tattatataagaccataatataatctcacaattcagcaacccttagagt-3';
- P6, 5'-acccttagagttgcaatctcacatcaactctcatccaacaacaaca-3').

1. Karimi M, Inzé D, Depicker A (2002) GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci* 7:193–195.

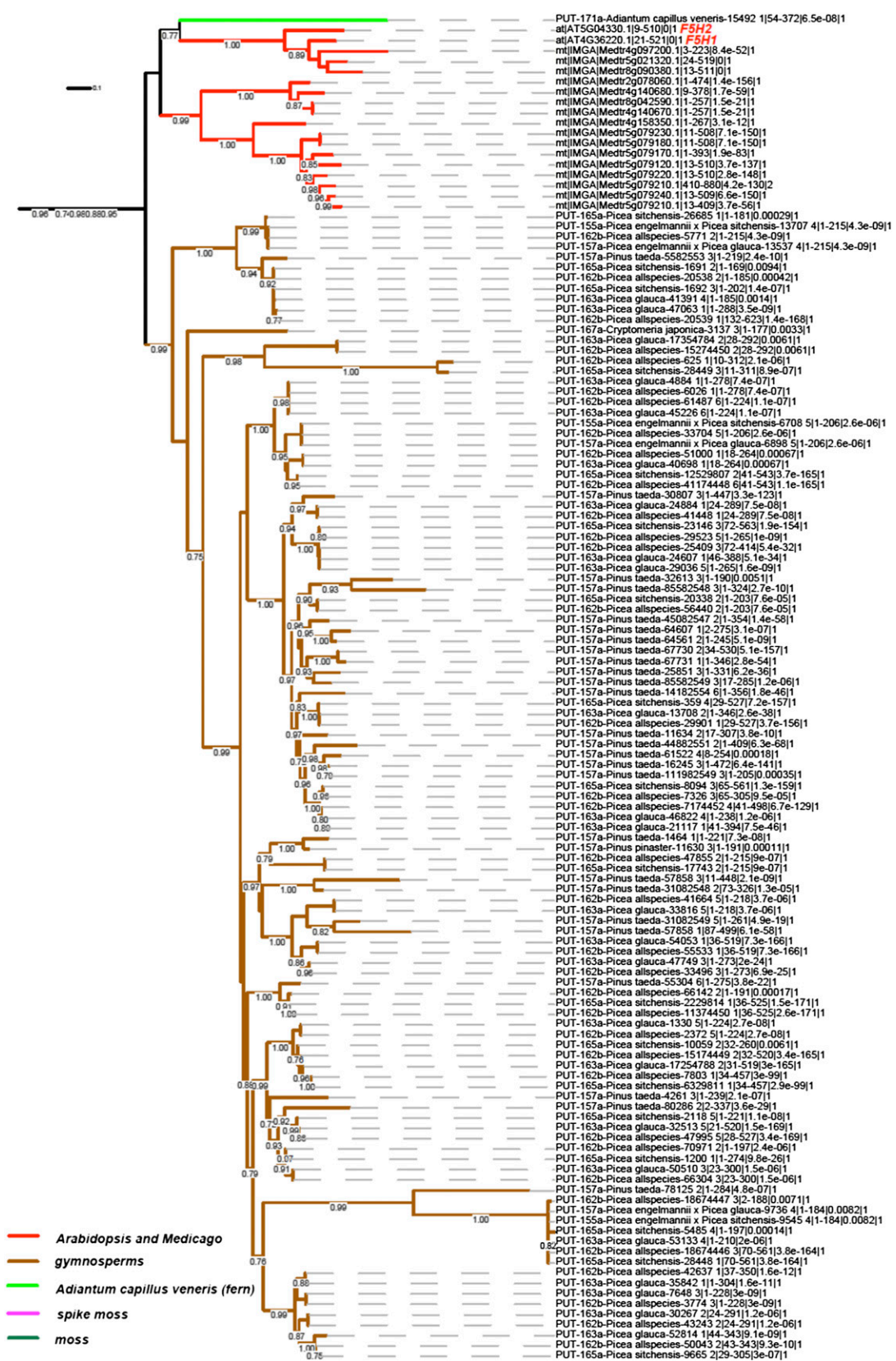


Fig. S1. Subclade containing *Arabidopsis* F5H proteins. This subclade is taken from Fig. 6A and shows the evolutionary orthologs of *Arabidopsis* F5H proteins in other organisms.

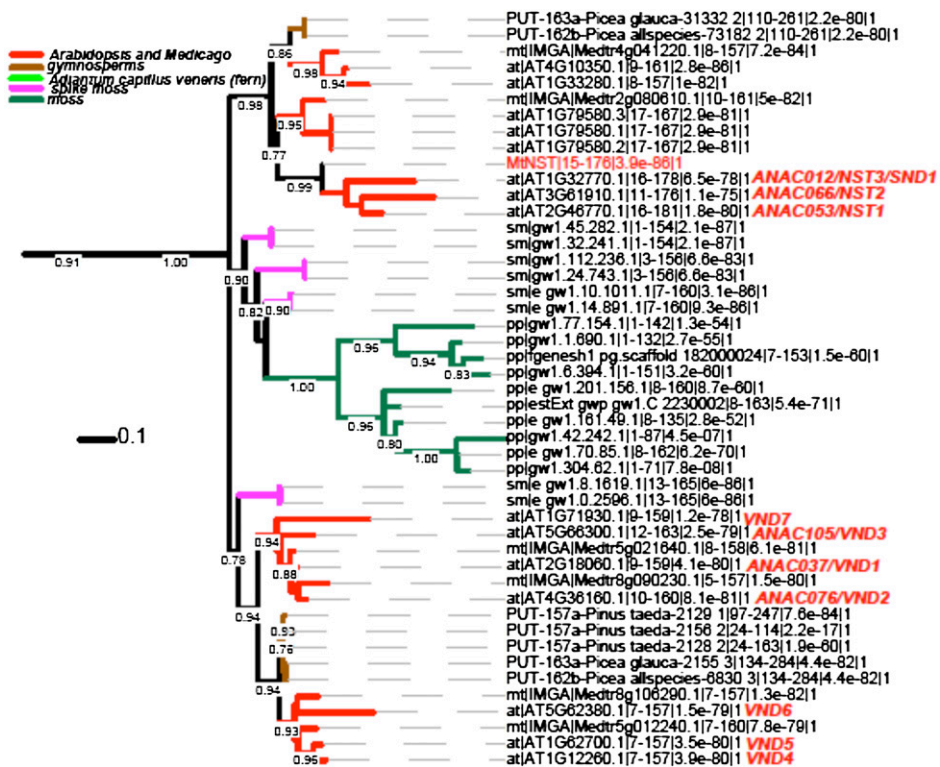


Fig. S2. Subclade containing *Arabidopsis* NAC proteins. This subclade is taken from Fig. 6B and shows the evolutionary orthologs of *Arabidopsis* NST proteins in other organisms.

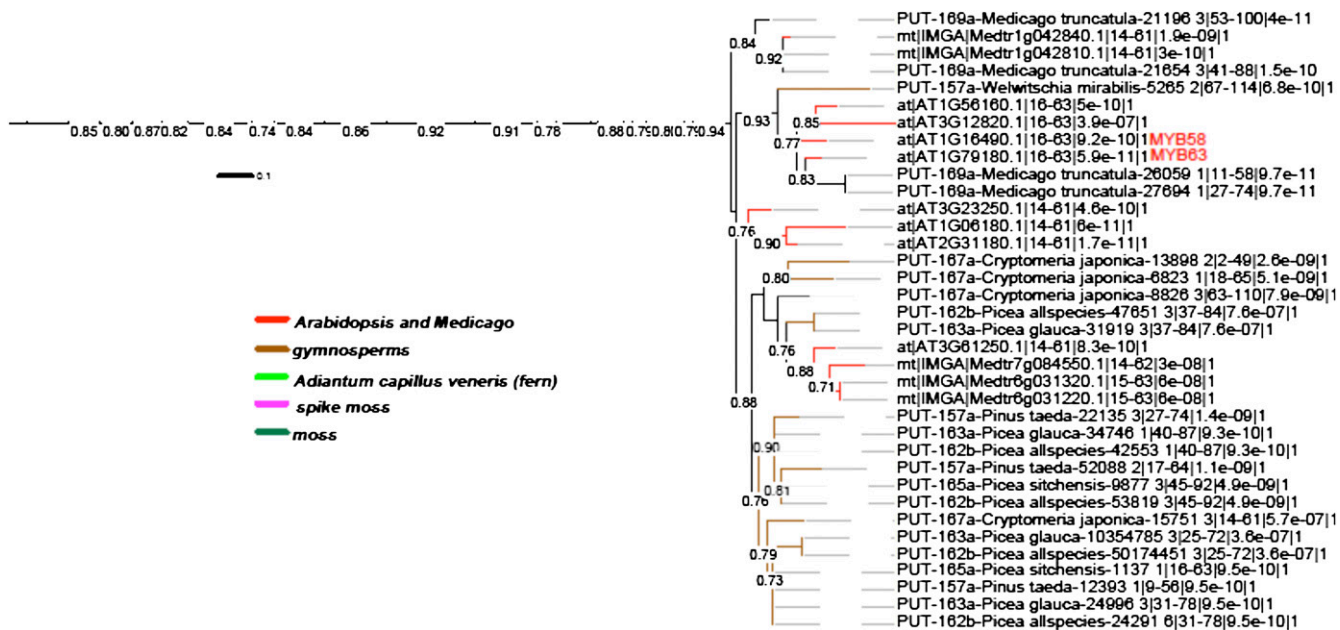


Fig. S3. Subclade containing the N-terminal domain of *Arabidopsis* MYB58 protein. This subclade is taken from Fig. 6C and shows the evolutionary orthologs of *Arabidopsis* MYB58 proteins in other organisms.

Table S1. The 21 land-plant and algal species used in the phylogenetic analyses

Species	Clade	No. of UniGenes/ proteins
<u>Cyanidioschyzon merolae 10D</u>	Unicellular red algae	<u>5014</u>
<u>Chlamydomonas reinhardtii</u>	Chlorophytic green algae	<u>14598</u>
<i>Mesostigma viride</i>	Green algae	7371
<i>Marchantia polymorpha</i>	Liverwort	10959
<u>Physcomitrella patens ssp patens</u>	Moss	<u>35938</u>
<u>Selaginella moellendorffii</u>	Spike moss	<u>34697</u>
<i>Adiantum capillus veneris</i>	Fern	16944
<i>Cycas rumphii</i>	Gymnosperms	10901
<i>Cryptomeria japonica</i>	Gymnosperms	24299
<i>Ginkgo biloba</i>	Gymnosperms	10210
<i>Picea abies</i>	Gymnosperms	5184
<i>Picea allspecies</i>	Gymnosperms	72011
<i>Picea engelmannii x Picea glauca</i>	Gymnosperms	13880
<i>Picea engelmannii x Picea sitchensis</i>	Gymnosperms	13880
<i>Picea glauca</i>	Gymnosperms	53255
<i>Picea sitchensis</i>	Gymnosperms	29178
<i>Pinus pinaster</i>	Gymnosperms	12901
<i>Pinus taeda</i>	Gymnosperms	77540
<i>Welwitschia mirabilis</i>	Gymnosperms	6606
<u>Arabidopsis thaliana</u>	Dicot	<u>33410</u>
<u>Medicago truncatula</u>	Dicot	<u>53423</u>

Six sequenced plant genomes are underlined; the remainder are EST assemblies taken from plant GDB. Species are ordered from lower plants to higher plants.