

Supplementary Material

1 Supplementary Tables

Supplementary Table S1. NCBI genome assembly accession numbers and sources of the peptide annotations used for other Fagales species.

| Species | Abbreviation | NCBI assembly | Source of peptide annotations |
|-------------------------|--------------|-----------------|---|
| <i>Alnus glutinosa</i> | Alngl | GCA_003254965.1 | http://gigadb.org/dataset/101042 |
| <i>Betula pendula</i> | Bpev | GCA_900184695.1 | https://www.hardwoodgenomics.org/Genome-assembly/2986212 |
| <i>Casuarina glauca</i> | Casgl | GCA_003255045.1 | http://gigadb.org/dataset/101051 |
| <i>Corylus avellana</i> | Corav | GCA_901000735.1 | https://www.hardwoodgenomics.org/Genome-assembly/3472010 |
| <i>Fagus sylvatica</i> | FSB | GCA_003347535.1 | https://www.hardwoodgenomics.org/Genome-assembly/2365974 |
| <i>Juglans nigra</i> | Juni | GCA_002916485.1 | https://www.hardwoodgenomics.org/Genome-assembly/3641871 |
| <i>Quercus robur</i> | Qrob | GCA_900291515.1 | http://www.oakgenome.fr/?page_id=587 |

Supplementary Table S2. Metabolites identified and/or provisionally annotated in red alder tissues. P, Phenolic acid; O, Organic acid; G, Gallic acid derivative; E, Ellagitannin; F, Flavonoid; Pa, Proanthocyanidin; D, Diarylheptanoid; Po, Polyamine. Pl, Pollen, Ca, Catkins.

| Feature | RT (min) | Category | Identification/Putative annotation | Molecular formulae | Calculated mass [M-H] ⁻ | ppm error | Nodules | | | Roots | | | Stems | | | Buds | | | Leaves | | Pl | Ca |
|----------|----------|----------|---|---|------------------------------------|-----------|---------|----|----|-------|----|----|-------|----|----|------|----|----|--------|----|----|----|
| | | | | | | | Sp | Su | Fa | Sp | Su | Fa | Sp | Su | Fa | Sp | Su | Fa | Su | Fa | Su | Su |
| M191T128 | 2.13 | O | Quinic acid ^a | C ₇ H ₁₂ O ₆ | 191.0556 | 0.13 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M173T157 | 2.62 | O | Shikimic acid ^a | C ₇ H ₁₀ O ₅ | 173.0450 | 1.10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M191T299 | 4.98 | O | Citric acid ^b | C ₆ H ₈ O ₇ | 191.0192 | 0.16 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M331T342 | 5.7 | G | β-Glucogallin ^a | C ₁₃ H ₁₆ O ₁₀ | 331.0665 | 0.51 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M783T489 | 8.15 | E | Pedunculagin (α or β anomer) ^a | C ₃₄ H ₂₄ O ₂₂ | 783.0681 | 1.00 | ND | ND | ND | √ | √ | √ | ND | √ | √ | ND | ND | ND | √ | √ | ND | √ |
| M783T581 | 9.68 | E | Pedunculagin (α or β anomer) ^a | C ₃₄ H ₂₄ O ₂₂ | 783.0681 | 1.04 | ND | ND | ND | √ | √ | √ | ND | √ | √ | ND | ND | ND | √ | √ | ND | √ |
| M483T594 | 9.90 | G | Digalloyl glucose ^{b,c} | C ₂₀ H ₂₀ O ₁₄ | 483.0775 | 4.14 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| M785T626 | 10.43 | E | Tellimagrandin I (α or β anomer) ^{a,d} | C ₃₄ H ₂₆ O ₂₂ | 785.0838 | 1.27 | ND | ND | ND | √ | √ | √ | √ | √ | √ | ND | ND | ND | √ | √ | ND | √ |
| M353T634 | 10.57 | P | <i>trans</i> -Chlorogenic acid ^a | C ₁₆ H ₁₈ O ₉ | 353.0873 | 0.29 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M633T645 | 10.75 | E | Strictinin ^a | C ₂₇ H ₂₂ O ₁₈ | 633.0728 | 0.76 | ND | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M289T647 | 10.78 | F | (+)-Catechin ^a | C ₁₅ H ₁₄ O ₆ | 289.0712 | 0.83 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M325T650 | 10.83 | P | <i>cis-p</i> -Coumaric acid glucoside ^a | C ₁₅ H ₁₈ O ₈ | 325.0923 | 1.54 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| M353T659 | 10.98 | P | Cryptochlorogenic acid (4- <i>O</i> -Caffeoyl quinic acid) ^a | C ₁₆ H ₁₈ O ₉ | 353.0873 | 0.43 | √ | √ | √ | √ | ND | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M577T688 | 11.47 | Pa | Procyanidin B2 ^a | C ₃₀ H ₂₆ O ₁₂ | 577.1346 | 0.07 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M785T701 | 11.68 | E | Tellimagrandin I (α or β anomer) ^a | C ₃₄ H ₂₆ O ₂₂ | 785.0838 | 0.55 | ND | √ | ND | √ | √ | √ | √ | √ | √ | √ | ND | ND | ND | √ | ND | √ |
| M337T720 | 12.00 | P | <i>cis-4-O-p</i> -Coumaroyl quinic acid ^c | C ₁₆ H ₁₈ O ₈ | 337.0923 | 1.57 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M635T729 | 12.15 | G | Trigalloyl glucose ^{b,d} | C ₂₇ H ₂₄ O ₁₈ | 635.0884 | 4.41 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| M289T734 | 12.23 | F | (-)-Epicatechin ^a | C ₁₅ H ₁₄ O ₆ | 289.0712 | 0.35 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M337T739 | 12.31 | P | <i>trans-3-O-p</i> -Coumaroyl quinic acid ^{a,b,c,d} | C ₁₆ H ₁₈ O ₈ | 337.0923 | 0.59 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND |
| M337T748 | 12.47 | P | <i>trans-4-O-p</i> -Coumaroyl quinic acid ^c | C ₁₆ H ₁₈ O ₈ | 337.0923 | 0.09 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M865T759 | 12.65 | Pa | Pa trimer ^c | C ₄₅ H ₃₈ O ₁₈ | 865.1980 | 0.31 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M625T776 | 12.93 | F | Dihexosylquercetin ^c | C ₂₇ H ₃₀ O ₁₇ | 625.1405 | 0.87 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | √ | √ |
| M367T782 | 13.03 | P | 4- <i>O</i> -Feruloyl quinic acid ^c | C ₁₇ H ₂₀ O ₉ | 367.1029 | 0.01 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |

| Feature | RT (min) | Category | Identification/Putative annotation | Molecular formulae | Calculated mass [M-H] ⁻ | ppm error | Nodules | | | Roots | | | Stems | | | Buds | | | Leaves | | Pl | Ca |
|-----------|----------|----------|--|---|------------------------------------|-----------|---------|----|----|-------|----|----|-------|----|----|------|----|----|--------|----|----|----|
| | | | | | | | Sp | Su | Fa | Sp | Su | Fa | Sp | Su | Fa | Sp | Su | Fa | Su | Fa | Su | Su |
| M935T791 | 13.18 | E | Casuarinin ^a | C ₄₁ H ₂₈ O ₂₆ | 935.0791 | 0.88 | ND | ND | √ | √ | √ | √ | √ | √ | ND | ND | ND | √ | √ | ND | √ | |
| M337T818 | 13.63 | P | <i>cis</i> -3- <i>O</i> - <i>p</i> -Coumaroyl quinic acid ^c | C ₁₆ H ₁₈ O ₈ | 337.0923 | 0.41 | √ | √ | ND | √ | √ | ND | √ | √ | √ | √ | √ | √ | √ | √ | √ | |
| M615T827 | 13.78 | F | Quercetin- <i>O</i> -(<i>O</i> -galloyl)-hexoside ^c | C ₂₈ H ₂₄ O ₁₆ | 615.0986 | 0.90 | ND | ND | ND | ND | √ | ND | √ | √ | √ | √ | √ | √ | √ | √ | √ | |
| M937T835 | 13.92 | E | Tellimagrandin II ^{a,b,d} | C ₄₁ H ₃₀ O ₂₆ | 937.0947 | 1.71 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | |
| M436T839 | 13.98 | Po | <i>N</i> ¹ , <i>N</i> ¹⁰ -bis(<i>p</i> -Coumaroyl) spermidine ^a | C ₂₅ H ₃₁ N ₃ O ₄ | 436.2236 | 0.03 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | √ | √ | |
| M615T845 | 14.08 | F | Quercetin- <i>O</i> -(<i>O</i> -galloyl)-hexoside ^c | C ₂₈ H ₂₄ O ₁₆ | 615.0986 | 1.51 | ND | ND | ND | √ | √ | ND | √ | √ | √ | √ | √ | √ | √ | √ | √ | |
| M436T872 | 14.53 | Po | <i>N</i> ¹ , <i>N</i> ¹⁰ -bis(<i>p</i> -Coumaroyl) spermidine ^c (isomer) | C ₂₅ H ₃₁ N ₃ O ₄ | 436.2236 | 1.01 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | √ | √ | |
| M507T884 | 14.73 | D | Hirsutanonol 5- <i>O</i> -glucoside ^c | C ₂₅ H ₃₂ O ₁₁ | 507.1866 | 0.88 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | |
| M463T892 | 14.87 | F | Quercetin-3- <i>O</i> -glucoside (isoquercitrin) ^a | C ₂₁ H ₂₀ O ₁₂ | 463.0877 | 0.35 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ | |
| M477T895 | 14.92 | F | Quercetin glucuronide ^c | C ₂₁ H ₁₈ O ₁₃ | 477.0669 | 0.23 | √ | √ | √ | ND | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | |
| M939T920 | 15.33 | G | Pentagalloyl glucose ^{a,b} | C ₄₁ H ₃₂ O ₂₆ | 939.1104 | 1.49 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | |
| M477T932 | 15.53 | D | Oregonin ^a | C ₂₄ H ₃₀ O ₁₀ | 477.1761 | 0.53 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | |
| M491T959 | 15.98 | D | 1-(3,4-Dihydroxyphenyl)-7-hydroxyphenyl-heptane-3-one 5- <i>O</i> -β-D-glucopyranoside ^c | C ₂₅ H ₃₂ O ₁₀ | 491.1917 | 0.37 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | |
| M447T965 | 16.08 | F | Kaempferol-3- <i>O</i> -glucoside ^{a,b} | C ₂₁ H ₂₀ O ₁₁ | 447.0927 | 1.25 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | √ | √ | |
| M491T968 | 16.13 | D | 1-Hydroxyphenyl-7-(3,4-dihydroxyphenyl)-heptane-3-one 5- <i>O</i> -β-D-glucopyranoside ^c | C ₂₅ H ₃₂ O ₁₀ | 491.1917 | 0.25 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | |
| M447T974 | 16.23 | F | Quercetin-rhamnoside (quercitrin) ^{a,c} | C ₂₁ H ₂₀ O ₁₁ | 447.0927 | 0.24 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | |
| M629T977 | 16.28 | D | Galloyl oregonin like ^c | C ₃₁ H ₃₂ O ₁₄ | 629.1870 | 0.82 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | |
| M345T1006 | 16.77 | D | Hirsutanonol ^a | C ₁₉ H ₂₂ O ₆ | 345.1338 | 0.46 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | |
| M461T1018 | 16.96 | D | Alnuside A ^{a,d} | C ₂₄ H ₃₀ O ₉ | 461.1812 | 0.09 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | |
| M625T1019 | 16.98 | D | Rubranoside D ^c | C ₃₀ H ₄₂ O ₁₄ | 625.2496 | 1.16 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | |
| M629T1022 | 17.03 | D | Galloyl oregonin like ^c | C ₃₁ H ₃₄ O ₁₄ | 629.1870 | 0.38 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | |
| M461T1028 | 17.13 | D | Alnuside B ^a | C ₂₄ H ₃₀ O ₉ | 461.1812 | 0.08 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | |

| Feature | RT (min) | Category | Identification/Putative annotation | Molecular formulae | Calculated mass [M-H] ⁻ | ppm error | Nodules | | | Roots | | | Stems | | | Buds | | | Leaves | | Pl | Ca |
|-----------|----------|----------|--|---|------------------------------------|-----------|---------|----|----|-------|----|----|-------|----|----|------|----|----|--------|----|----|----|
| | | | | | | | Sp | Su | Fa | Sp | Su | Fa | Sp | Su | Fa | Sp | Su | Fa | Su | Fa | Su | Fa |
| M475T1045 | 17.42 | D | Platyphylloside ^a | C ₂₅ H ₃₂ O ₉ | 475.1968 | 0.63 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M431T1060 | 17.67 | F | Kaempferol-3-rhamnoside (afzelin) ^c | C ₂₁ H ₂₀ O ₁₀ | 431.0978 | 0.28 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M493T1066 | 17.77 | D | Rubranoside A ^a | C ₂₅ H ₃₄ O ₁₀ | 493.2074 | 0.50 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M629T1073 | 17.88 | D | Galloyl oregonin like ^c | C ₃₁ H ₃₄ O ₁₄ | 629.1870 | 0.84 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M625T1099 | 18.32 | D | Rubranoside C ^c | C ₃₀ H ₄₂ O ₁₄ | 625.2496 | 0.64 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M463T1117 | 18.62 | D | Rubranoside B ^a | C ₂₄ H ₃₂ O ₉ | 463.1968 | 0.14 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M445T1120 | 18.67 | D | Platyphyllonol 5- <i>O</i> -D-xylopyranoside ^c | C ₂₄ H ₃₀ O ₈ | 445.1862 | 0.45 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M577T1180 | 19.67 | D | Butanoyl hirsutanonol 5- <i>O</i> -β-D-glucopyranoside ^c | C ₂₉ H ₃₈ O ₁₂ | 577.2285 | 0.57 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M623T1192 | 19.87 | D | 2'''- <i>O</i> - <i>p</i> -Coumaroyl oregonin ^a | C ₃₃ H ₃₆ O ₁₂ | 623.2129 | 0.26 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M285T1214 | 20.23 | F | Luteolin ^a | C ₁₅ H ₁₀ O ₆ | 285.0399 | 0.10 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M547T1233 | 20.55 | D | Butanoyl oregonin ^c | C ₂₈ H ₃₆ O ₁₁ | 547.2179 | 0.39 | √ | √ | √ | ND | √ | ND | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M461T1240 | 20.67 | D | Aceroside VII ^c | C ₂₅ H ₃₄ O ₈ | 461.2175 | 0.47 | √ | √ | √ | ND | √ | ND | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M591T1264 | 21.07 | D | Methylbutanol hirsutanonol 5- <i>O</i> -β-D-glucopyranoside ^c | C ₃₀ H ₄₀ O ₁₂ | 591.2442 | 1.19 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M582T1274 | 21.23 | Po | Tri- <i>p</i> -coumaroyl spermidine ^c | C ₃₄ H ₃₇ N ₃ O ₆ | 582.2604 | 1.03 | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | ND | √ | √ |
| M593T1284 | 21.40 | D | 1,7-bis(4-Hydroxyphenyl)-3-heptanol 3- <i>O</i> -β-D-glucopyranosyl-(1→3)-β-D-xylopyranoside ^c | C ₃₀ H ₄₂ O ₁₂ | 593.2598 | 0.32 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M561T1321 | 22.02 | D | Alnuside C ^c | C ₂₉ H ₃₈ O ₁₁ | 561.2336 | 0.43 | √ | √ | √ | ND | ND | ND | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M431T1323 | 22.05 | D | 1,7-bis(4-Hydroxyphenyl)-3-heptanol 3- <i>O</i> -β-D-xylopyranoside ^c | C ₂₄ H ₃₂ O ₇ | 431.2070 | 0.73 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M269T1353 | 22.55 | F | Apigenin ^a | C ₁₅ H ₁₀ O ₅ | 269.0450 | 0.11 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ | √ |
| M575T1366 | 22.77 | D | 1-(3,4-Dihydroxyphenyl)-7-hydroxyphenyl-heptane-3-one methylbutanoyl 5- <i>O</i> -β-D-glucopyranoside ^c | C ₃₀ H ₄₀ O ₁₁ | 575.2492 | 0.61 | √ | √ | √ | ND | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| M637T1394 | 23.23 | D | Cinnamoyl hirsutanonol 5- <i>O</i> -β-D-glucopyranoside ^c | C ₃₄ H ₃₈ O ₁₂ | 637.2285 | 0.86 | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | ND | √ |
| M345T1416 | 23.60 | F | Viscidulin III ^c | C ₁₇ H ₁₄ O ₈ | 345.0610 | 0.60 | ND | ND | ND | ND | √ | ND | √ | √ | ND | √ | √ | √ | √ | √ | √ | √ |
| M283T1673 | 27.88 | F | 7- <i>O</i> -Methylapigenin (genkwanin) ^a | C ₁₆ H ₁₂ O ₅ | 283.0607 | 0.15 | ND | ND | ND | ND | ND | ND | √ | √ | ND | √ | √ | √ | √ | √ | √ | √ |

| Feature | RT (min) | Category | Identification/Putative annotation | Molecular formulae | Calculated mass [M-H] ⁻ | ppm error | Nodules | | | Roots | | | Stems | | | Buds | | | Leaves | | Pl | Ca |
|-----------|----------|----------|------------------------------------|--|------------------------------------|-----------|---------|----|----|-------|----|----|-------|----|----|------|----|----|--------|----|----|----|
| | | | | | | | Sp | Su | Fa | Sp | Su | Fa | Sp | Su | Fa | Sp | Su | Fa | Su | Fa | Su | Su |
| M313T1679 | 27.98 | F | Dimethoxyluteolin ^c | C ₁₇ H ₁₄ O ₆ | 313.0712 | 0.02 | ND | ND | ND | ND | ND | ND | √ | √ | ND | √ | √ | √ | √ | √ | √ | √ |
| M329T1686 | 28.10 | F | Unidentified flavonoid | C ₁₇ H ₁₄ O ₇ | 329.0661 | 0.60 | ND | ND | ND | ND | ND | ND | √ | √ | ND | √ | √ | √ | ND | √ | √ | √ |

a = confirmed using authentic standards, including proanthocyanidins (procyanidin B2), diarylheptanoids (alnusides A and B, hirsutanonol, oregonin, 2'''-O-p-coumaroyl oregonin, platyphylloside and rubranosides A and B), gallic acid derivatives and ellagitannin related [β -glucogallin, pentagalloyl glucose, pedunculagin (α or β anomers), tellimagrandin I (α or β anomers), strictinin, casuarinin, tellimagrandin II], polyamine [N^1, N^{10} -bis(*p*-coumaroyl) spermidine], flavonoids [(+)-catechin, (-)-epicatechin, quercetin-3-*O*-glucoside, quercitrin, genkwanin, kaempferol-3-*O*-glucoside, luteolin and apigenin], phenolic acids/derivatives (*trans*-chlorogenic, 4-*O*-caffeoyl quinic and *trans*-3-*O*-*p*-coumaroyl quinic acids, as well as *cis*-4-*O*-*p*-coumaric acid glucoside), organic acids (citric, shikimic, and quinic acids).

b = metabolite with lower detection limit (trace)

c = provisionally annotated by analysis of mass spectrometric fragmentation patterns and comparison to literature reports such as by Sati et al. (1), Jackrel (2), and Jackrel et al. (3).

d = several known alder metabolites were not detected in these Clone 639 tissues.

Supplementary Table S3. Red alder homologs to *bona fide* shikimate, chorismate, phenylalanine, and β -glucogallin pathway genes.

| Enzyme | Plant species | NCBI accession number | <i>A. rubra</i> homologs | Name | Amino acid identity (%)* | Predicted localization** |
|--|-------------------------------|--------------------------|--------------------------|-------------|--------------------------|--------------------------|
| 3-Deoxy-D-arabinoheptulosonate 7-phosphate synthase (DAHPS) | <i>Arabidopsis thaliana</i> | AAA32784 (4) | ALNRU00118289 | ArDAHPS1 | 80.5 | Plastid |
| | | | ALNRU00091501 | ArDAHPS2 | 80.4 | Plastid |
| | | | ALNRU00046501 | ArDAHPS3 | 77.6 | Plastid |
| | | | ALNRU00044315 | ArDAHPS4 | 74.2 | Plastid |
| | | | ALNRU00084407 | ArDAHPS5 | 72.8 | Plastid |
| 3-Dehydroquinate synthase (DHQS) | <i>Solanum lycopersicum</i> | AAL77575 (5) | ALNRU00027821 | ArDHQS | 76.8 | Plastid |
| Bifunctional 3-dehydroquinate dehydratase/shikimate dehydrogenase (DHQD-SDH) | <i>A. thaliana</i> | Q9SQT8 (6) | ALNRU00075273 | ArDHQD-SDH1 | 73.2 | Cytoplasm |
| | | | ALNRU00016551 | ArDHQD-SDH2 | 50.6 | Cytoplasm |
| | | | ALNRU00016553 | ArDHQD-SDH3 | 64.1 | Cytoplasm |
| | | | ALNRU00016554 | ArDHQD-SDH4 | 63.6 | Cytoplasm |
| | | | ALNRU00118494 | ArDHQD-SDH5 | 47.8 | Cytoplasm |
| | | | ALNRU00050250 | ArDHQD-SDH6 | 53.3 | Cytoplasm |
| Shikimate kinase (SK) | <i>S. lycopersicum</i> | NP_001234112 (7) | ALNRU00033270 | ArSK | 50.4 | Plastid |
| 5-Enolpyruvylshikimate-3-phosphate synthase (EPSPS) | <i>A. thaliana</i> | P05466 (8) | ALNRU00062892 | ArEPSPS | 73.3 | Plastid |
| Chorismate synthase (CS) | <i>Capnoides sempervirens</i> | CAA43034 (9) | ALNRU00062533 | ArCS | 76.1 | Plastid |
| Chorismate mutase (CM) | <i>A. thaliana</i> | CAB54518 (10) | ALNRU00054560 | ArCM1 | 70.8 | Plastid |
| | | | ALNRU00061145 | ArCM2 | 49.1 | Cytoplasm |
| | | | ALNRU00010637 | ArCM3 | 61.4 | Plastid |
| Prephenate amino-transferase (PPA-AT) | <i>A. thaliana</i> | ADM67558 (11) | ALNRU00090523 | ArPPA-AT | 71.9 | Plastid |
| Arogenate dehydratase (ADT) | <i>A. thaliana</i> | At5g22630 (AtADT5) (12) | ALNRU00051159 | ArADT1 | 73.3 | Plastid |
| | | | ALNRU00103384 | ArADT2 | 53.4 | Plastid |
| | | | ALNRU00048921 | ArADT3 | 52.5 | Plastid |
| β -glucogallin forming UDP glucosyltransferase (UGT) | <i>Quercus robur</i> | AHA54051 (UGT84A13) (13) | ALNRU00005258 | ArUGT1 | 87.1 | Cytoplasm |
| | | | ALNRU00139344 | ArUGT2 | 87.2 | Cytoplasm |
| | | | ALNRU00005266 | ArUGT3 | 71.7 | Cytoplasm |
| | | | ALNRU00126768 | ArUGT4 | 71.1 | Cytoplasm |
| | | | ALNRU00005272 | ArUGT5 | 70.6 | Cytoplasm |

*used Clustal Omega Multi Sequence Alignment Tool (14).

**used LOCALIZER 1.0.4 (15) and DeepLoc 2.0 (16).

Supplementary Table S4. Red alder homologs to *bona fide* phenylpropanoid pathway genes.

| Enzyme | Plant species | NCBI accession number | <i>A. rubra</i> homologs | Name | Amino acid identity (%)* |
|--|-----------------------------|------------------------------|--------------------------|----------|--------------------------|
| Phenylalanine ammonia lyase (PAL) | <i>Arabidopsis thaliana</i> | At2g37040 (AtPAL1) (17) | ALNRU00113829 | ArPAL1 | 83.9 |
| | | | ALNRU00062591 | ArPAL2 | 82.2 |
| | | | ALNRU00103226 | ArPAL3 | 81.8 |
| | | | ALNRU00062587 | ArPAL4 | 80.8 |
| Cinnamate 4-hydroxylase (C4H) | <i>Helianthus tuberosus</i> | CYP73A1, CAA78982 (18) | ALNRU00092227 | ArC4H1 | 89.1 |
| | | | ALNRU00122270 | ArC4H2 | 62.0 |
| | | | ALNRU00028961 | ArC4H3 | 62.2 |
| Hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase (HCT) | <i>Nicotiana tabacum</i> | CAD47830 (NtHCT) (19) | ALNRU00070014 | ArHCT1 | 79.5 |
| | | | ALNRU00070002 | ArHCT2 | 76.0 |
| | | | ALNRU00069972 | ArHCT3 | 75.3 |
| <i>p</i> -Coumarate 3-hydroxylase (C3H) | <i>A. thaliana</i> | CYP98A3, At2g40890 (20) | ALNRU00003369 | ArC3H1 | 82.3 |
| | | | ALNRU00003318 | ArC3H2 | 77.4 |
| | | | ALNRU00003289 | ArC3H3 | 74.6 |
| | | | ALNRU00138548 | ArC3H4 | 74.6 |
| 4-Coumarate CoA ligase (4CL) | <i>A. thaliana</i> | At1g51680 (At4CL1) (21) | ALNRU00130727 | Ar4CL1 | 69.8 |
| | | | ALNRU00011055 | Ar4CL2 | 69.6 |
| | | | ALNRU00087439 | Ar4CL3 | 70.3 |
| | | | ALNRU00090869 | Ar4CL4 | 60.3 |
| Ferulate 5-hydroxylase (F5H) | <i>A. thaliana</i> | CYP84A1, At4g36220 (22) | ALNRU00027975 | ArF5H | 62.3 |
| Caffeic acid <i>O</i> -methyltransferase (COMT) | <i>A. thaliana</i> | At5g54160 (AtCOMT1) (23, 24) | ALNRU00041884 | ArCOMT1 | 78.5 |
| | | | ALNRU00090858 | ArCOMT2 | 62.3 |
| Caffeoyl CoA <i>O</i> -methyltransferase (CCOMT) | <i>Petroselinum crispum</i> | AAA33851 (25) | ALNRU00005083 | ArCCOMT1 | 78.6 |
| | | | ALNRU00005087 | ArCCOMT2 | 71.8 |
| | | | ALNRU00025885 | ArCCOMT3 | 58.0 |
| Cinnamoyl CoA reductase (CCR) | <i>A. thaliana</i> | At1g15950 (AtCCR1) (26) | ALNRU00069900 | ArCCR1 | 78.9 |
| | | | ALNRU00065491 | ArCCR2 | 50.3 |
| Cinnamyl alcohol dehydrogenase (CAD) | <i>A. thaliana</i> | At4g34230 (AtCAD5) (27) | ALNRU00005245 | ArCAD1 | 79.3 |
| | | | ALNRU00011316 | ArCAD2 | 62.9 |
| | | | ALNRU00085651 | ArCAD3 | 50.1 |
| | | | ALNRU00028036 | ArCAD4 | 47.6 |
| | | | ALNRU00048433 | ArCAD5 | 49.2 |
| | | | ALNRU00061461 | ArCAD6 | 48.3 |
| | | | ALNRU00085657 | ArCAD7 | 48.5 |
| | | | ALNRU00085652 | ArCAD8 | 48.7 |
| | | | ALNRU00061417 | ArCAD9 | 47.5 |
| | | | ALNRU00061363 | ArCAD10 | 47.9 |

*used Clustal Omega Multi Sequence Alignment Tool (14).

Supplementary Table S5. Red alder homologs to *bona fide* flavonoid pathway genes.

| Enzyme | Plant species | NCBI accession number | <i>A. rubra</i> homologs | Name | Amino acid identity (%)* |
|--|--------------------------------|-------------------------|--------------------------|-----------|--------------------------|
| Chalcone synthase (CHS) | <i>Antirrhinum majus</i> | CAA27338 (28) | ALNRU00052590 | ArCHS1 | 86.1 |
| | | | ALNRU00080687 | ArCHS2 | 73.9 |
| | | | ALNRU00080685 | ArCHS3 | 86.2 |
| | | | ALNRU00080658 | ArCHS4 | 72.7 |
| Chalcone isomerase (CHI) | <i>Petunia × hybrida</i> | CAA32730 (29) | ALNRU00071279 | ArCHI | 70.5 |
| Flavanone 3-hydroxylase (F3H) | <i>Arabidopsis thaliana</i> | AAC49176 (30) | ALNRU00027879 | ArF3H | 84.6 |
| Flavonoid 3'-hydroxylase (F3'H) | <i>A. thaliana</i> | At5g07790 (31) | ALNRU00072988 | ArF3'H1 | 71.3 |
| | | | ALNRU00026572 | ArF3'H2 | 71.2 |
| Flavonoid 3',5'-hydroxylase (F3',5'H) | <i>P. hybrida</i> | CAA80266 (32) | ALNRU00009602 | ArF3',5'H | 76.8 |
| Flavonol synthase (FLS) | <i>P. hybrida</i> | CAA80264 (33) | ALNRU00076204 | ArFLS1 | 73.9 |
| | | | ALNRU00037056 | ArFLS2 | 67.5 |
| Dihydroflavonol 4-reductase (DFR) | <i>A. majus</i> | P14721 (34) | ALNRU00000326 | ArDFR1 | 69.4 |
| | | | ALNRU00000329 | ArDFR2 | 62.2 |
| Leucoanthocyanidin reductase (LAR) | <i>Desmodium uncinatum</i> | CAD79341 (35) | ALNRU00034769 | ArLAR1 | 66.2 |
| | | | ALNRU00056862 | ArLAR2 | 58.1 |
| Leucoanthocyanidin dioxygenase (LDOX) also called anthocyanidin synthase (ANS) | <i>A. thaliana</i> | Q96323 (36) | ALNRU00013023 | ArLDOX | 79.9 |
| Anthocyanidin reductase (ANR) | <i>Medicago truncatula</i> | AAN77735 (37) | ALNRU00124301 | ArANR1 | 76.6 |
| | | | ALNRU00012650 | ArANR2 | 74.9 |
| Flavone synthase II (FNS II) | <i>Gerbera</i> hybrid cultivar | AAD39549 (CYP93B2) (38) | ALNRU00042823 | ArFNSII-1 | 57.1 |
| | | | ALNRU00004682 | ArFNSII-2 | 51.3 |
| | | | ALNRU00004681 | ArFNSII-3 | 51.4 |
| | | | ALNRU00004684 | ArFNSII-4 | 48.5 |

*used Clustal Omega Multi Sequence Alignment Tool (14).

Supplementary Table S6. Red alder homologs to *bona fide* dirigent protein genes

| Enzyme | Plant species | NCBI accession number | <i>A. rubra</i> homologs | Name | Amino acid identity (%)* |
|------------------|-----------------------------|-----------------------------------|--------------------------|---------------------------|--------------------------|
| Dir-a sub-family | <i>Pisum sativum</i> | P13240 (DRR206) (39) | ALNRU00062277 | ArDIR1 | 73.2 |
| | | | ALNRU00062278 | ArDIR2 | 62.0 |
| | | | ALNRU00062271 | ArDIR3 | 51.2 |
| Dir-b sub-family | <i>P. sativum</i> | PsCam039127 6OOD (PsPTS1) (40) | ALNRU00004390 | ArDIR4 | 55.7 |
| | | | ALNRU00004385 | ArDIR5 | 53.3 |
| | | | ALNRU00004380 | ArDIR6 | 53.8 |
| | | | ALNRU00004378 | ArDIR7 | 54.4 |
| | | | ALNRU00091352 | ArDIR8 | 49.7 |
| | | | ALNRU00091353 | ArDIR9 | 48.4 |
| | | | ALNRU00091336 | ArDIR10 | 46.7 |
| | | | ALNRU00135575 | ArDIR11 | 46.7 |
| | | | ALNRU00091354 | ArDIR12 | 45.4 |
| | | | ALNRU00091345 | ArDIR13 | 43.2 |
| | | | ALNRU00113764 | ArDIR15 | 46.0 |
| | | | ALNRU00116248 | ArDIR21 | 49.2 |
| | | | Dir-d sub-family | <i>Gossypium hirsutum</i> | GhDIR4 (41) |
| ALNRU00093034 | ArDIR16 | 44.6 | | | |
| ALNRU00065374 | ArDIR17 | 46.3 | | | |
| ALNRU00093041 | ArDIR20 | 46.0 | | | |
| ALNRU00065375 | ArDIR22 | 41.1 | | | |
| ALNRU00065377 | ArDIR25 | 44.6 | | | |
| ALNRU00123362 | ArDIR27 | 38.1 | | | |
| Dir-e sub-family | <i>Arabidopsis thaliana</i> | At2g28670 (AtDIR10) (42) | ALNRU00042827 | ArDIR18 | 60.2 |
| | | | ALNRU00071249 | ArDIR19 | 61.1 |
| | | | ALNRU00071250 | ArDIR23 | 41.8 |
| | | | ALNRU00052656 | ArDIR24 | 40.9 |
| | | | ALNRU00071251 | ArDIR26 | 36.4 |

*used Clustal Omega Multi Sequence Alignment Tool (14).

Supplementary Table S7. Sequence comparisons (percent identity) of chalcone synthases (CHS), curcumin synthase (CURS) and diketide CoA synthase (DCS).

| | AmCHS* | CIDCS | CICURS | ArCHS1 | ArCHS2 | ArCHS3 |
|---------|--------|-------|--------|--------|--------|--------|
| CIDCS | 65.8 | | | | | |
| CICURS1 | 60.2 | 62.6 | | | | |
| ArCHS1 | 86.1 | 65.5 | 60.7 | | | |
| ArCHS2 | 73.9 | 62.5 | 55.3 | 74.8 | | |
| ArCHS3 | 86.2 | 65.8 | 59.4 | 89.5 | 76.5 | |
| ArCHS4 | 72.7 | 61.2 | 48.8 | 72.7 | 91.9 | 74.6 |

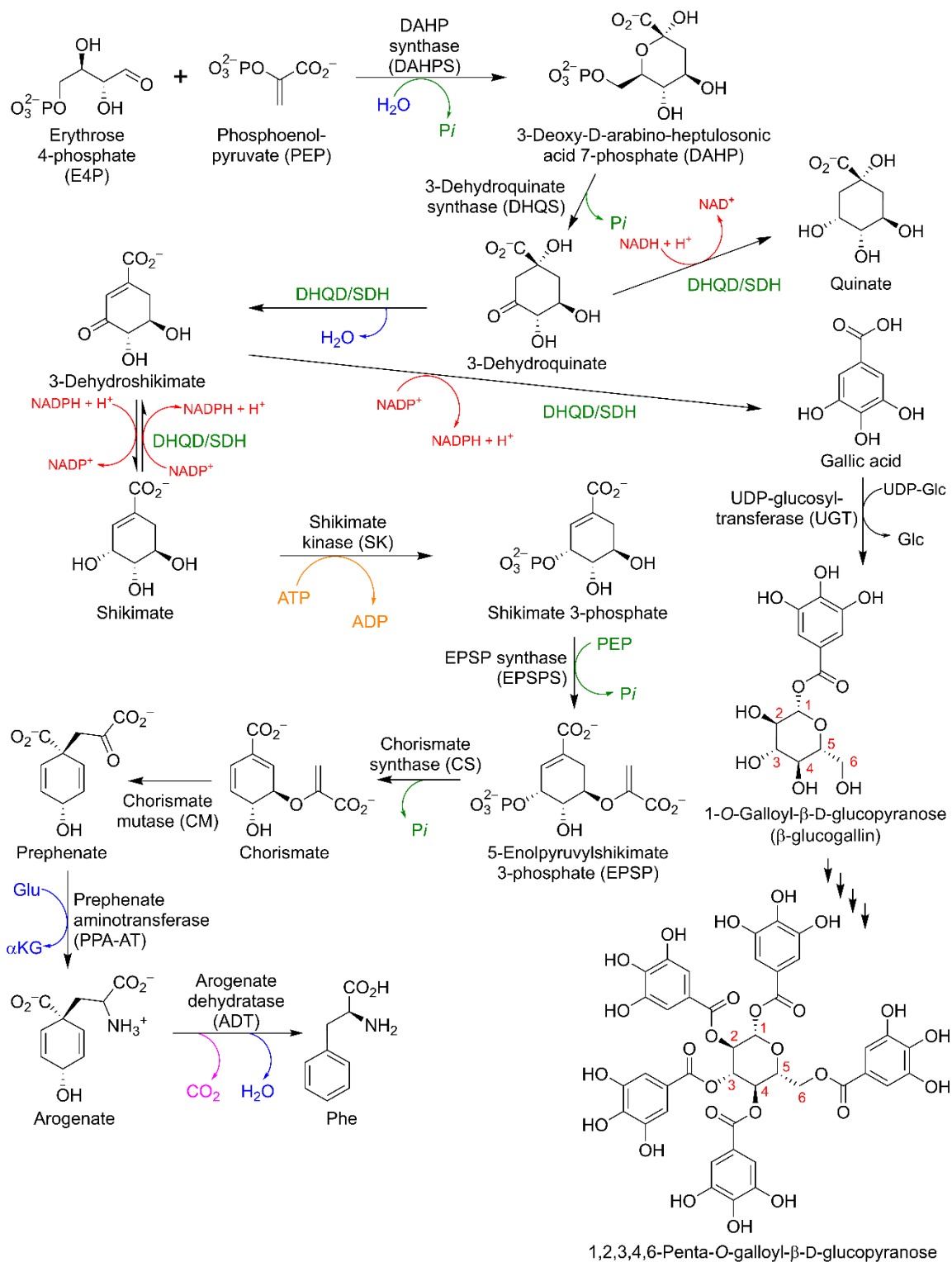
*AmCHS, *Antirrhinum majus* CHS (CAA27338); CICURS1, *Curcuma longa* CURS (BAH56226); CIDCS, *C. longa* DCS (BAH56225); ArCHS1 – ArCH3, putative *Alnus rubra* CHS (ALNRU00052590, ALNRU00080687, ALNRU00080685)

Supplementary Table S8. Red alder homologs *bona fide* pinoresinol-lariciresinol reductase, allylphenol synthase, propenylphenol synthase, and secoisolariciresinol dehydrogenase genes.

| Enzyme | Plant species | NCBI accession number | <i>A. rubra</i> homologs | Name | Amino acid identity (%)* |
|---|-----------------------------|------------------------|--------------------------|--------|--------------------------|
| Pinoresinol-lariciresinol reductase (PLR) | <i>Forsythia intermedia</i> | AAC49608 (FiPLR1) (43) | ALNRU00108755 | ArPLR1 | 77.5 |
| Allylphenol synthase (APS) | <i>Larrea tridendata</i> | AHA90804 (LtAPS1) (44) | ALNRU00041001 | ArAPS1 | 81.8 |
| | | | ALNRU00117167 | ArAPS2 | 81.8 |
| | | | ALNRU00040994 | ArAPS3 | 77.0 |
| | | | ALNRU00041004 | ArAPS4 | 77.0 |
| | | | ALNRU00117169 | ArAPS5 | 77.0 |
| Propenylphenol synthase (PPS) | <i>L. tridendata</i> | AHA90806 (LtPPS1) (44) | ALNRU00072637 | ArPPS1 | 64.4 |
| | | | ALNRU00072634 | ArPPS2 | 63.1 |
| Secoisolariciresinol dehydrogenase (SDH) | <i>Podophyllum peltatum</i> | AF352734 (PpSDH) (45) | ALNRU00043749 | ArSDH1 | 56.2 |
| | | | ALNRU00043779 | ArSDH2 | 56.2 |
| | | | ALNRU00043753 | ArSDH3 | 55.8 |
| | | | ALNRU00043507 | ArSDH4 | 56.0 |
| | | | ALNRU00043783 | ArSDH5 | 53.5 |
| | | | ALNRU00049317 | ArSDH6 | 53.4 |
| | | | ALNRU00043813 | ArSDH7 | 50.2 |
| | | | ALNRU00049319 | ArSDH8 | 54.1 |

*used Clustal Omega Multi Sequence Alignment Tool (14).

2 Supplementary Figures



Supplementary Figure S1. Current simplified shikimate-chorismate pathway and offshoots to phenylalanine, quinate, β-glucogallin, and 1,2,3,4,6-penta-O-galloyl-β-D-glucopyranose. DHQD-SDH: bifunctional 3-dehydroquinone dehydratase/shikimate dehydrogenase.

```

AtDAHPS      1  MALNASTLSTPSYGGDLHRPSNRQSEFTFPAVNTKPKSVNLV---TAVHAAEPAANAVSVKE-----SVAS
ArDAHPS1    1  MALNNTATLSSKSLYSNTQT---LSSHQPFFSVVPTTTTKTNSGKLVTRISAVHAAEPTK-----PS
ArDAHPS2    1  MALNNTATLSSKSLYSNTQT---LSSHQPFFSVVPTTTTKTNSGKLVTRISAVHAAEPTK-----PS
ArDAHPS3    1  MALNASSVPTKSFITQTSLLPST-----KPHQPSIPSRRLSPLSISAVHAAEPAANPVVSDKPPKQVTPPPTKT
ArDAHPS4    1  MALNIVTSTLSSRPFLQPHQPAKPLH-----HHHLSLPDRK---KPTISAVHAAEPTK-----SKSTPT
ArDAHPS5    1  MALNLLKGSVN-----LTPPTITKALPLPHFL-----TP---HFTTTRIPSVSS-----SSSVTDNAN

AtDAHPS      68  SSSGALKWTFESWKLKKALQLPYPNAAELESVLKTEAFPPVIFAGEARNLEERLAAMVGAFLFLQGGDCAESFKEFN
ArDAHPS1    61  AAAFPSKWTVESWKSPTLQLPEYPNQEELETVLKTIEAFPPVIFAGEARHLEELAEAAIGNAFLFLQGGDCAESFKEFN
ArDAHPS2    61  AAAFPSKWTVESWKSPTLQLPEYPNQEELETVLKTIEAFPPVIFAGEARHLEELAEAAIGNAFLFLQGGDCAESFKEFN
ArDAHPS3    74  HNAFVFTWVESWKKKALQLPEYPNQEELESVLKTEAFPPVIFAGEARNLEERLAEAAIGNAFLFLQGGDCAESFKEFN
ArDAHPS4    57  TTPPTRWSLDSWKSKEVQLPEYDQALDFVINTLESFPPVIFAGEARSLERLAEAAQKAFLLQGGDCAESFKEFN
ArDAHPS5    54  TSTGFSNMAPYSWKSKKARQLPYPDPDLRSVQLTLESFPPVIFAGEARKLEERLAAMVGAFLFLQGGDCAESFKEFN

AtDAHPS     148  ATNIRDTRFVLLQMSIVLTFGGQVPVIKVGMRMAGQFAKPRSDAFEEKDGVKLPYSYKGDNINGDTEFEKSRIPDPERMIRA
ArDAHPS1    141  ANNIRDTRFVLLQMGAVLMFGQVPI---VGRMAGQFAKPRSDSLEEKDGVKLPYSYKGDNINGDTEFEKSRIPDPERMIRA
ArDAHPS2    141  ANNIRDTRFVLLQMGAVLMFGQVPIVGRMAGQFAKPRSDSLEEKDGVKLPYSYKGDNINGDTEFEKSRIPDPERMIRA
ArDAHPS3    154  ANNIRDTRFVLLQMGAVLMFGQVPIVGRMAGQFAKPRSDPFEEKDGVKLPYSYKGDNINGDTEFEKSRIPDPERMIRA
ArDAHPS4    137  ANNIRDTRFVLLQMGVLMFGQVPIVGRMAGQFAKPRSDPFEEKDGVKLPYSYKGDNINGDTEFEKSRIPDPERMIRA
ArDAHPS5    134  GNNIRDTRFVLLQMGIALTVGAQVPIVGRMAGQFAKPRSDPFEEKDGVKLPYSYKGDNINGDTEFEKSRIPDPERMIRA

AtDAHPS     228  YTQSAATLNLLRAFATGGYAAIQRVTVQWNLDFVHQSEQLDRYQELANRVDEALGFMSACGLGTDHPIMTTTDFWTSHECL
ArDAHPS1    218  YCQSAATLNLLRAFATGGYAAIQRISEWNLDFAEHSEQDQRYQELANRVDEALGFMAAAGLTVDHPVMTTEFEWTSHECL
ArDAHPS2    221  YCQSAATLNLLRAFATGGYAAIQRISEWNLDFAEHSEQDQRYQELANRVDEALGFMAAAGLTVDHPVMTTEFEWTSHECL
ArDAHPS3    234  YCQSAATLNLLRAFATGGYAAIQRVTVQWNLDFAEHSEQDQRYRELARRVDEALGFMTAAGLTVDHPIMTTTDFWTSHECL
ArDAHPS4    217  YGQSVATLNLLRAFATGGYAAIQRVTVQWNLDFAEHSEQDQRYRELARRVDEALGFMSAAGLTVDHPIMTTTDFWTSHECL
ArDAHPS5    214  YLQSVGTLNLLRAFATGGYAAIQRVTVQWNLDFVHSEQDQRYMELARRVDEALGFMAAAGLTVNHPIMTTTEFEWTSHECL

AtDAHPS     308  LLPYEQSLTRLDSTSGLYDCAHMLVWVGERTRLDGAHVEFLRGIANPLGIKVSNKMDPNELVKLVEILNPNKPKGRIT
ArDAHPS1    298  HLPYEQSLTRKLDSTSGLYDCAHMLVWVGERTRLDGAHVEFLRGIANPLGIKVSNKMDPNELVKLVEILNPNKPKGRIT
ArDAHPS2    301  HLPYEQSLTRKLDSTSGLYDCAHMLVWVGERTRLDGAHVEFLRGIANPLGIKVSNKMDPNELVKLVEILNPNKPKGRIT
ArDAHPS3    314  LLPYEQSLTRLDSTSGLYDCAHMLVWVGERTRLDGAHVEFLRGIANPLGIKVSNKMDPNELVKLVEILNPNKPKGRIT
ArDAHPS4    297  LLPYEQALTRDSTSGLYDCAHMLVWVGERTRLDGAHVEFLRGIANPLGIKVSNKMDPNELVKLVEILNPNKPKGRIT
ArDAHPS5    294  HLPYEQALTRDSTSGCYDCAHMLVWVGERTRLDGAHVEFLRGVSNPLGIKVSNKMDPEKELVKLVEILNPNKPKGRIT

AtDAHPS     388  VIVRMGAENMRVKLPHLIRAVRRSGQIVTWVCDPMHGNTIKSTCGLKTRAFDSTLAEVRAFFDVHEQEGSHPGGVHLEMT
ArDAHPS1    378  IIVRMGAENMRVKLPHLIRAVRRAGQIVTWVCDPMHGNTIKAPCGLKTRPFDAITLAEVRAFFDVHEQEGSHPGGVHLEMT
ArDAHPS2    381  IIVRMGAENMRVKLPHLIRAVRRAGQIVTWVCDPMHGNTIKAPCGLKTRPFDAITLAEVRAFFDVHEQEGSHPGGVHLEMT
ArDAHPS3    394  IIVRMGAENMRVKLPHLIRAVRRAGHIVTWVSDPMHGNTIKAPCGLKTRPFDSITLAEVRAFFDVHEQEGSHPGGVHLEMT
ArDAHPS4    377  VIVRMGAENMRVKLPHLIRAVRAGQIVTWVSDPMHGNTIKAPCGLKTRSFDAITLAEVRAFFDVHEQEGSYPGGVHLEMT
ArDAHPS5    374  IIVRMGANMRVKLPHLIRAVRQAGLIVTWVSDPMHGNTIKAPSGLKTRPFDAITRAEIRAFFDVHEQEGSYPGGVHLEMT

AtDAHPS     468  GQNVTECIGGSRTVTFDDLSSRYHTHCDPRLNASQSLELAFIAERLRKRRIGSORV-S----
ArDAHPS1    458  GQNVTECIGGSRTVTFDDLSSRYHTHCDPRLNASQSLELAFIAERLRKRRIGQRI-ASLSL
ArDAHPS2    461  GQNVTECIGGSRTVTFDDLSSRYHTHCDPRLNASQSLELAFIAERLRKRRIGQRI-ASLSL
ArDAHPS3    474  GQNVTECIGGSRTVTFDDLSSRYHTHCDPRLNASQALELAFIAERLRKRRIRSQPLSSLGF
ArDAHPS4    457  GQNVTECIGGSRTVTFDDLSSRYHTHCDPRLNASQSLELAFIAERLRKRRIGSRRLSVS--
ArDAHPS5    454  GQNVTECIGGSRTVTFDDLSSRYHTHCDPRLNASQSLELAFIAERLRKRRIGSRVNDVQQRG

```

Supplementary Figure S2. Protein sequence alignment of *bona fide* 3-deoxy-D-arabino-heptulosonic acid 7-phosphate synthase (DAHPS) from *Arabidopsis thaliana* (AtDAHPS (4)) and that of putative homologs (ArDAHPS1 – ArDAHPS5) in red alder.

```

SlDHQS 1 MASSFCPEKQALSFITNSTHQLHQSRAIPRDIHVRFPAEVSPSSRCGLKS-----KATTRLKVIATSAT
ArDHQS 1 MATITNPFSLSL-----KNSADIFSRVPEFNPNSVSLRSEFVSPGSSDLSLIRGQNSVVSASRVRSVRICASSA

SlDHQS 64 KVMDSSSSKASSQAPTVEVDLGTRSYPIYIGAGLLDQPDLLQRHGHGKRVLVVTNTTVAPLYLDRITISALTDGNPNVIV
ArDHQS 70 QVMDQSVAKTDSKVEVIVDVNLGDRSYPIYIGSGLLDQPHLLQRHGHGKRVLVVTNNTTVAPLYLDRVVEALTKGNPNVSV

SlDHQS 144 ESVILPDGEQPKNMTLMKFDKAIESRLDRRCTFVALGGGVIGDMCGFAAASYLRGVNFIQIPTTVMAQVDSSVGGKTG
ArDHQS 150 ESVILPDGEKPKMDTLMKFDKAIESRLDRRCTFVALGGGVIGDMCGFAAASFLRGVNFIQIPTTVMAQVDSSVGGKTG

SlDHQS 224 INHPLGKNMIGAFYQPQCVLIDTDTLNTLPDRELASGLAEVIKYGLIRDAEFFEWQEQNPLLRARDPTAFTYAIKRSC
ArDHQS 230 INHRLGKNLIGAFYQPQCVLIDTDTLNTLPDRELASGLAEVIKYGLIRDAEFFEWQERNIQALMSRDPSEALAYAIKRSC

SlDHQS 304 NKADVVSQDEKESGRATLNLGHTFGHAEETGVGYCQWLHGEAVAAGTVMVAVDMSRRLGWIDDSIVQRVQKILQQAQLPT
ArDHQS 310 NKADEVVSLDEKEGGIRATLNLGHTFGHAEETGVGYCQWLHGEAVAAGMVMVAVDLSRRLGWIDDSIVKRVHNIILTQAKLPT

SlDHQS 384 SPPEIMTVEMFKSIMAVDKKVADGLRLRILLLKGLGNCVFTGDYDOKAIDETLRAFCKS
ArDHQS 390 AFAKTMTVEMFKSIMAVDKKVADGLRLRILLLKGPLGNCVFTGDYDRKAIDETLRAFCKS

```

Supplementary Figure S3. Protein sequence alignment of *bona fide* 3-dehydroquinase synthase (DHQS) from *Solanum lycopersicum* (SlDHQS (5)) and that of a putative homolog (ArDHQS) in red alder.

```

AtDHQD-SDH          68 -----MEPSNVYVASNSTEMEIGSHDI VKNPSLI CAPVMADSIDKVIETSKAHEL GADIVE IRLDWLKD FNPTE
FvDHQD-SDH3         1 -----MGSSSTLV CAPIMAEVHKMVRDMSKARL GADIVE IRLDYLKV FNSNQ
FvDHQD-SDH4         1 -----MDSPTLTVASAVQVGGG GMRKSSSTLVCAPIMAESVAKMVMEMGRAKAV GADIVE IRLDHLKV FDSSE
NtDHQD-SDH1         1 -----MELV---VDSGVRKMEGEAMTRNETLVCAPIMADTVQMLINLMQKAKIS GADIVE IRLDSLKS FNPQS
SlDHQD-SDH1         1 -----MELV---VDSGVKKMEGEAMRKNQTLVCAPIMADSVQMLITLMQKAKIS GADIVE IRLVDSLKS FNPFP
CasDHQD-SDH4        1 -----MELVAVPVVAGCGVKMSE GMRKNSTLVCVPLMADSVQMLITVMNKAQLGADIVE IRLDSLKS FNPPL
VvDHQD-SDH1         1 -----MESGGMSKRNSTLVCVPTMGETIEKVMVDMSKAKIS GADIVE IRLDTLKR FNPPO
FSB010265701        1 -----MDSVVLLASSSAGLQMGGGEIRKNPTLVCAPIMAESVDKVINMDKAKQGGADIVE IRLDSLKS FNPNE
FSB012889301        1 -----MDSNVLLASSSAGLQIGGGEIRKNPTLVCAPIMAESVDKVINMDKAKQGGADIVE IRLDSLKS FNPNE
Qrob_P0275560       1 -----MDSNVLLASSSAGLEMGIREFIRKNQTLVCAPIMAEVVDKVINMDKAKQSDADIVE IRLDSLKS FNPYE
Juni_01730          1 -----MASGGGRNGPTLVCAPIMAESVDKVMANMEKAKQGGADIVE IRLDSLKS FNPHE
JrDHQD-SDH         1 -----TLV CAPIMAESVDKVINMDKAKQGGADIVE IRLDSLKS FNPEN
Juni_19979          1 -----MESSNVLLASSAALQMGSGG-RNNPTLVCAPIMAESVDKVINMDKAKQGGADIVE IRLDSLKS FNPED
ArDHQ-SDH1          1 -----MESGNVLLASSAALQMRSGGARNNSTLVCAPVMADSVDKMVTNMDKAKREGADIVE IRLDSLKS FDPYE
Alng188491S37275   1 -----MESGNVLLASSAALQMRSGGARNNSTLVCAPVMADSVDKMVTNMDKAKREGADIVE IRLDSLKS FDPYE
Bpev01_c0420-g0007 1 -----MESRNVLLASSAALQMGSGGVRNNPTLVCAPVMAESVDKVMVTNMDKAKQDGADEIVE IRLDSLKS FDPYE
Corav_882           1 -----
Casg115S00444      1 -----MLASSAPQMATEGVRRNPSTLVCAPIMAESIDKVINMDKAKQDGADEIVE IRLDSLKS FHPDE
EgDHQD-SDH5         1 -----MASTGNILLATSSAPHEMESGAVRKNSTLVCVPMADSVDEMVIQLDKAKSSGADIVE IRLVDGLKNLSPHE
EcDHQD-SDH1         1 -----MASTGNVLLATSSAPHEMESGAVRKNSTLVCVPMADSVDEMVIQLDKAKSSGADIVE IRLVDSLKNLSPHE
CsDHQD-SDH3         1 -----MESPPLLVASGSKLVSGGMRKNPTLVCVPTMGESVDKVMVDMGKAKASGADIVE IRLDGLKN FNPPE
PoptrDHQD-SDH1     1 -----MDSASNVLASSPSAAAAGVGGSGVRRNPSTLVCPTMADSVDKMAIIMAEAKSVGADIVE IRLDSLKD FNPNS
PoptrDHQD-SDH3     1 -----LINSVMV CAPLMAQSVEQMVIMDMSAKAQQADIVE IRLDCIK FQPRQ
PoptrDHQD-SDH2     1 -----LAINSTMV CAPLMSRSVEQMVIMDMSAKAQQADIVE IRLDYINS FQPSQ
EgDHQD-SDH4         1 -----MGSLSLSSVGLTMV CAPVMGESVDQVVEEMHKAKAQQADIVE IRLDCIK FQAHQ
EcDHQD-SDH4a        1 -----MGSLSLSSVGLTMV CAPVMGESVDQVVEEMHKAKAQQADIVE IRLDCIK FQAHQ
EcDHQD-SDH4b        1 -----MGSLSLSSVGLTMV CAPVMGESVDQVVEEMHKAKAQQADIVE IRLDCIK FQAHQ
CasDHQD-SDH1        1 -----MGSVGV-LTNSTTICAPLMSQSVEQMVSDMNQAKAQQADIVE IRLDCIK FQPNR
SlDHQD-SDH2         1 -----MGSVGL-LKNSAMV CAPLMATSVDQLIDEMVEAKSQGADIVE IRLDAIHN FQPHK
NtDHQD-SDH2         1 -----MGSVGL-LKNSAMV CAPLMAQSVQVLSNMVQAKAQQADIVE IRLDCINN FQPK
VvDHQD-SDH1         1 -----TKNTTMI CAPLMAQSVEQVLSNMVQAKAQQADIVE IRLDCINN FQPK
VvDHQD-SDH2         1 -----MDDVGV-LKKE TMICTPLMGQSVEQMVSDMNMHAKVEGADIVE IRLDYINN FHPQ
Qrob_P0208710      1 -----MGSIGMVMVSKTTLVCAPLMAETVEQIMTDMYKAKVQGADIVE IRLDYIKN FQPRQ
FSB015730401       1 -----MGSLSI-VSKSTLVCAPLMSQVTEQIVTDMYKAKVQGADIVE IRLDYIKN FKPFO
Juni_23842         1 -----MGSIGM-LNNSTLVCAPLMAQSVEQMVSDMJEAKAQQADIVE IRLDCIKN FQPRQ
Bpev01_c1356-g0002 1 -----MGSIGI-VDNSTLVCAPIVAQSVEQMVTDKMQAKAQQADIVE IRLDCLKN FQPRQ
Corav_4151          1 -----MGTVGI-FDRSTLVCAPLVAQSVEQMVIMDKQAKAQQADIVE IRLDYIKN FQPRQ
Alng120613S19962   1 -----MGSVGI-LENSTLVCAPLVAQSVEQMVIMDKQAKAQQADIVE IRLDYIKN FQPRQ
ArDHQD-SDH2         1 -----MGSVGI-LENSTLVCAPLVAQSVEQMVIMDKQAKAQQADIVE IRLDYIKN FQPRQ
CasDHQD-SDH2        1 -----MASGSFSFATSDVQTSSTSGVRSPTLCTPLIGTTVDQMLTDMRKAKEIGADIVE IRLDCLRF FNPFP
VvDHQD-SDH3         1 -----MGSLLPFTVSDLQTSVSGVRSNPTLCTPLMGTTVBQMLTEMRKAKEIGADIVE IRLDCLRFN FSPAQ
FvDHQD-SDH1         1 -----MGSLLPFTTSDLHTSTGGFLSSPTLCTPLMGTTVQDMLIEMHKAKEIGSDVVE IRLDCLRFN FNPSS
EgDHQD-SDH3         1 -----MGSVPFTTSDLQTSSTGGFRSSPTLCTPLMGTTVQDMLIEMRKAKEIGADIVE IRLDCLRFN FNPFO
EcDHQD-SDH3         1 -----MGSVPFTTSDLQTSSTGGFRSSPTLCTPLMGTTVQDMLIEMRKAKEIGADIVE IRLDCLRFN FNPFO
Qrob_P0208720      1 -----MGSLLPFTVSDLQASINCARSSPTLCTPLMGTTVQDMLIEMGKAKEIGADIVE IRLDCLRT FNPFP
FSB015730501       1 -----MGSLLPFTVSGLQTSIDGARSPTLCTPLMGTTVQDMLIEMGKAKEIGADIVE IRLDCLRF FNPFP
Corav_14055         1 -----MASLPFTVSGLQTSSTGGRRSSPTLCTPLMGTTVQDMLIEMGKAKEIGSDVVE IRLDCLRFN FNPFP
Corav_11016         1 -----MASLPFTVSGLQTSSTGGRRSSPTLCTPLMGTTVQDMLIEMGKAKEIGSDVVE IRLDCLRFN FNPFP
Bpev01_c1356-g0003 1 -----MASLPFTVSGLQTSSTGGRRSSPTLCTPLMGTTVQDMLIEMGKAKEIGSDVVE IRLDCLRFN FNPFP
ArDHQD-SDH3         1 -----MGSLLPFTVSGLQTSSTGGRRSSPTLCTPLMGTTVQDMLIEMGKAKEIGSDVVE IRLDCLRFN FNPFP
CasDHQD-SDH3        1 -----MTCAPVMAENVDQMLLMRKAKELGADIVE IRLDYLKN FSPHH
DkDHQD-SDH         1 -----MTCAPVMAETAEQMLGQMRKAKELGADIVE IRLDYLKN FSPQO
VvDHQD-SDH4         1 -----MTLSSVPLATSDI QIPE-GARRNSTLVCVPTMADSVQMLGQIRKAKEVGGADIVE IRLDYLKN FSPRO
EgDHQD-SDH2         1 -----MTLSSIPLTAADLQIPA-GGRRNSTLVCAPVMGESVDQMLGQIRAAKEQGGADIVE IRLDFLKS FSPRO
EcDHQD-SDH2         1 -----MTLSSIPLTAADLQIPA-GGRRNSTLVCAPVMGESVDQMLGQIRAAKEQGGADIVE IRLDFLKS FSPRO
PoptrDHQD-SDH5     1 -----MDLQSDAD-GERRNSTLVCAPIMAESVDQMLVQMKRAKELGADVAE IRLDFLKN FSPRN
FvDHQD-SDH2         1 -----MTLSSIPLVASDLQISY-GTGRNSTLVCAPVMGESVDQMLRQLQAKELGADIVE IRLDYIKN FSPRO
Casg1267S13432     1 -----MTLSSVPLAATDLKMGD-AVRRNSTLVCAPVMAETVDQMLIQTRNAKELGADIVE IRLDSLKN FSPRO
Qrob_P0208730      1 -----MTLSSIPLATS DLQISG-GIRRSSTLVCAPVMAETVDQMLLMRKAKELGADIVE IRLDFLKN FNPFO
FSB015730601       1 -----MTLSSIPLATS DLQIAD-GTRRSSTLVCAPVMAETVDQMLLMRKAKELGADIVE IRLDFLKN FSPRO
Juni_23843         1 -----MTLDSIPLATMDLQIAD-GVRRNGTQI CAPVMAETVDQMLTQARKAKELGADIVE IRLDFLKN FNPFO
ArDHQD-SDH4         1 -----MTLIGSVPLATSDLQIVD-GARRNSTLVCAPVMAETVDQMLIQTRKAKELGADIVE IRLDFLKN FSPRO
Alng16638S34349   1 -----
Bpev01_c1356-g0004 1 -----MTLGSVPLATTDLQIVG-GARRNSTLVCAPVMAETVDQMLIQARKAKELGADIVE IRLDFLKN FSPFO
Corav_941           1 -----MTLGSVPLATPDLQIVG-GARRNSTLVCAPVMAETVDQMLIQARKAKELGADIVE IRLDFLKN FSPRO
Corav_7742         1 -----MTLGSVPLATPDLQIVG-GARRNSTLVCAPVMAETVDQMLIQARKAKELGADIVE IRLDFLKN FSPRO
SlDHQD-SDH3        1 MQVLTWFLVLYMFPPLYKRERWSVKEERDIIMGLKNDLVVYTRLECESCEEMTCCIEKAKEEGADIVE IRLCDF-IFSDI
EgDHQD-SDH1        40 VAEGVTAQTLHSSSRIFFLHCPSSRRERPREMDRNGVLCAPLECEETLCEMLSSMDKAKAHGADIVE IRLVDMAM-SFGRVS
PoptrDHQD-SDH4     1 -----MAFKNNLLVCTPLECETAGEMLSSMKRAEHEGADITEL IRLDSL-SFHSNS
CsDHQD-SDH2        1 -----MEVAAKNSLLVCTOLECETTEEMQASIEQAKVEGADIVE IRLCIDS-MEFSHIS

```

```

Qrob_P0339740 1 -----MAFKNNLLVCTQLQECETMEEMIASMEKAKAEGADIVELCLDMS-SFSHTS
Juni_24457 1 -----MAFKNNILVCTPLECETEMEEMVGSMEKAKAEGADIVELCVDSM-SFSHTS
Casg1106806827 1 -----MAFKNNLLVCTSLCETEMEEMIASMDKAKAEGADIVELCVDSL-SFSHTS
Bpev01_c0171-g0015 1 -----MAFKNNLLICTPLECETMEEMOASMEKAKAEGADIVELCLDMS-QFSHTS
ArDHQD-SDH5 1 -----MAFKNNLLICTPLECETMEEMOASMEKAKAEGADIVELCLDMS-SFSHTS
ArDHQD-SDH6 1 -----

```

```

AtDHQD-SDH 138 DLKTLIKKSELPPTLFTYR-----PWEGGQYEGDE--NRRDVLRLAMELGADYIDVELQVASEF
FvDHQD-SDH3 50 DLKTLIKESFMPPTLFTYR-----PWEGGQYDGDDE--KHRODLRLAMELGADYIDVELQVAHEF
FvDHQD-SDH4 67 DVKTLIDQSFPLPTLFTYR-----PWEGGQYDGDDE--KSRLDALRLAMELGADYIDVELQVAQEF
NtDHQD-SDH1 65 DIDTLIKQSFPLPTLFTYR-----PWEGGQYAGDE--VSRDLALRLAMELGADYIDVELKVAIDEF
SLDHQD-SDH1 66 DIDTLIKQCFPLPTLFTYSVVLGVGGQILLIRYYKIGIPWEGGQYAGDE--KSRLDALRLAMELGADYIDVELKVAIGEF
CasDHQD-SDH4 69 DIETLIKRCPLPTLFTYR-----PWEGGQYDGDDE--SSRLEALRLAMELGADYIDVELQVIHEF
VvDHQD-SDH1 55 DLVLRIRKCFPLPTLFTYR-----PWEGGQYEGDE--NSRRDALRLAMELGADYIDVELKVAHEF
FSB010265701 71 DLKTLIKECFPLPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELQVAQEF
FSB012889301 71 DLKTLIKECFPLPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELQVAQEF
Qrob_P0275560 71 DLKTLIKECFPLPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELQVAQEF
Juni_01730 55 DLKTLIKECFPLPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELQVAHEF
JrDHQD-SDH 45 DLKTLIKASPLPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELQVACEF
Juni_19979 69 DLKTLIKASPLPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELQVACEF
ArDHQ-SDH1 70 DLKTLIKECFPLPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELQVAHEF
Alng188491S37275 70 DLKTLIKECFPLPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELQVAHEF
Bpev01_c0420-g0007 70 DLKTLIKECFPLPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELQVAHEF
Corav_882 -----
Casg115800444 64 DLKTLIKECFPLPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELQVAHEF
EgDHQD-SDH5 72 DLKTLIKASALPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELKVAHEF
EcDHQD-SDH1 72 DLKTLIKASALPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELKVAHEF
CsDHQD-SDH3 68 NIKTLIKESFVPTLFTYR-----PWEGGQYDGDDE--NRRDVLRLAMELGADYIDVELQVAHEF
PoptrDHQD-SDH1 75 DIKTLILHSELPPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELKVAHEF
PoptrDHQD-SDH3 55 DLETLIRNKPLPVIIVYR-----PWEGGQYEGDE--HRRLEALRLANLGLGADYIDVELKVADEL
PoptrDHQD-SDH2 55 DLETLIRNKPLPVIIVYR-----PWEGGQYEGDE--HRRLEALRLANLGLGADYIDVELKVASDL
EgDHQD-SDH4 56 DLEIILKSKPLPVIIVYR-----KCEGGQYEGDE--TARLEALRSALKLGADYIDVELKVADEL
EcDHQD-SDH4a 56 DLEIILKSKPLPVIIVYR-----KCEGGQYEGDE--TARLEALRSALKLGADYIDVELKVADEL
EcDHQD-SDH4b 56 DLEIILKSKPLPVIIVYR-----KCEGGQYEGDE--TARLEALRSALKLGADYIDVELKVADEL
CasDHQD-SDH1 55 DLQILLNKELPVLIVYR-----QWEGGQYGGDE--NRRDVLRLAMELGADYIDVELKVAHEF
SLDHQD-SDH2 55 HLQLLFKNKPLPVLIVYR-----IWERNDFEADA--HKQLEALRLAKELGADYIDVELKTIASEF
NtDHQD-SDH2 55 DLQVLLNNNLPVLIVYR-----IWEGNDFEEDDHI--HKQLEALRLAKELGADYIDVELKTIASDF
CsDHQD-SDH1 55 DLEIILKSKPLPVLIVYR-----KWEGGQYEGDE--HRRLEALRLANLGLGADYIDVELKVASNI
VvDHQD-SDH2 55 DLEIILRNKPLPVLIVYR-----PWEGGQYEGDE--HSRLEALRLAKELGADYIDVELKVASDF
Qrob_P0208710 56 DLEIILKSKPLPVLIVYR-----KWEGGQYEGDE--HTRLEALRLAKELGADYIDVELKVASNL
FSB015730401 55 DLEIILKSKPLPVLIVYR-----KWEGGQYEGDE--HSRLEALRLAKELGADYIDVELKVASNH
Juni_23842 55 DLQTIIRKPLPVLIVYR-----RWEGGQYEGDE--CTRVEALRLAKELGADYIDVELKVASNL
Bpev01_c1356-g0002 55 DLEIILRNKPLPVLIVYR-----KWEGGQYEGDE--HMRVETLRLAKELGADYIDVELKVASDL
Corav_4151 55 DLEIILRNKPLPVLIVYR-----KWEGGQYEGDE--HMRVETLRLAKELGADYIDVELKVASDL
Alng120613S19962 55 DLEIILRNKPLPVLIVYR-----KWEGGQYEGDE--HMRVETLRLAKELGADYIDVELKVASDL
ArDHQD-SDH2 55 DLEIILRNKPLPVIIVYR-----KWEGGQYEGDE--HTRVETLRLAKELGADYIDVELKVASDL
CasDHQD-SDH2 69 DLQILLIKQSFPLPTLFTYR-----PWEGGQYEGDE--NRRDVLRLAMELGADYIDVELQVAQEF
VvDHQD-SDH3 67 DLQILLIKQSFPLPTLFTYR-----PWEGGQYEGDE--NKRODALRLAMELGADYIDVELQVAHEF
FvDHQD-SDH1 67 DLQILLIKQSFPLPTLFTYR-----PWEGGQYEGDE--TKRODALRLAMELGADYIDVELQVAHEF
EgDHQD-SDH3 67 DLQILLIKQSFPLPTLFTYR-----PWEGGQYEGDE--SKRODALRLAMELGADYIDVELQVAHDF
EcDHQD-SDH3 67 DLQILLIKQSFPLPTLFTYR-----PWEGGQYEGDE--SKRODALRLAMELGADYIDVELQVAHDF
Qrob_P0208720 67 DLEILIKRSSLPPTLFTYR-----PWEGGQYEGDE--NKRODALRLAMELGADYIDVELQVAHDF
FSB015730501 67 DLEILIKCSSLPPTLFTYR-----PWEGGQYEGDE--NKRODALRLAMELGADYIDVELQVAHDF
Corav_14055 67 DLEFLVKQSFPLPTLFTYR-----PWEGGQYEGDE--TKRODALRLAMELGADYIDVELQVAHDF
Corav_11016 67 DLEFLVKQSFPLPTLFTYR-----PWEGGQYEGDE--TKRODALRLAMELGADYIDVELQVAHDF
Bpev01_c1356-g0003 67 DLEFLVKQSFPLPTLFTYR-----PWEGGQYEGDE--TKRODALRLAMELGADYIDVELQVAHDF
ArDHQD-SDH3 67 DLEFLVKQSFPLPTLFTYR-----PWEGGQYEGDE--TKRODALRLAMELGADYIDVELQVAHDF
CasDHQD-SDH3 44 DLQILLIKQCFPLPTLFTYR-----PWEGGQYDGDDE--TKRODALRLAMELGADYIDVELQVAHEF
VvDHQD-SDH 44 HLEVLIKQSFPLPTLFTYR-----PWEGGQYDGDDE--SRQATLRLAMELGADYIDVELQVAHEF
DkDHQD-SDH4 68 DLQFLVKQSFPLPTLFTYR-----PWEGGQYDGDDE--KRRLDALRLAMELGADYIDVELQVAHEF
EgDHQD-SDH2 68 DLEIILKQSFPLPTLFTYR-----PWEGGQYEGDE--SRRLDALRLAMELGADYIDVELQVAQEF
EcDHQD-SDH2 68 DLEIILKQSFPLPTLFTYR-----PWEGGQYEGDE--SRRLDALRLAMELGADYIDVELQVAQEF
PoptrDHQD-SDH5 58 DLEALIKQCFPLPTLFTYR-----PWEGGQYDGDDE--NKRODALRLAMELGADYIDVELQVAHEF
FvDHQD-SDH2 68 DLETLIKRSPLPTLFTYR-----PWEGGQYEGDE--KRRLDALRLAMELGADYIDVELQVAHDF
Casg1267S13432 68 DLEILINKSSLPPTLFTYR-----PWEGGQYDGDDE--SKRONALRLAMELGADYIDVELQVAHDF
Qrob_P0208730 68 DLEILIKQSFPLPTLFTYR-----PWEGGQYDGDDE--SKRONALRLAMELGADYIDVELQVAHDF
FSB015730601 68 DLEILIKQSFPLPTLFTYR-----PWEGGQYDGDDE--SKRONALRLAMELGADYIDVELQVAHDF
Juni_23843 69 DLEILIKQCFPLPTLFTYR-----PWEGGQYDGDDE--SKRONALRLAMELGADYIDVELQVAHDF
ArDHQD-SDH4 69 DLEILIKQCFPLPTLFTYR-----PWEGGQYDGDDE--SKRONALRLAMELGADYIDVELQVAHDF
Alng16638S34349 -----

```


Bpev01_c1356-g0004 68 DLEILLIKQSELPPLTITYR-----PKWEGGOYDGDG---SRKROTALRQAMELGADYIDVELQVAHEF
 Corav_941 68 DLEILLIKQSELPPLTITYR-----PKWEGGOYDGDG---GRKROTALRLAMELGADYIDVELKVAHDF
 Corav_7742 68 DLEILLIKQSELPPLTITYR-----PKWEGGOYDGDG---GRKROTALRLAMELGADYIDVELKVAHDF
 SlDHQD-SDH3 80 QLEELLKQKRSPLSIVSFR-----FKSPINS---EGKKTICIQVLRKLAVELDVEFEVVDTOVVC HQ
 EgDHQD-SDH1 219 EVEELIRRRTLPAIVSFR-----LNSARASRRQNDKTTCLQVLRRLAELDVEFEVMEHEVVS HF
 PoptrDHQD-SDH4 50 EVERLIKQRTLPSIVSFR-----LEPSRISSNKDRKNTCLQVLRRLAELDVEFEVEMDYEVASB D
 SlDHQD-SDH2 52 EVDKLIQHETLPAIVSYR-----LKSSRKSSDACKNTCLQVLRRLAELDVEFEVEMDYEVASD P
 Csdh_Q0339740 50 EVEKLIKQRLPAIVSYR-----LKSSRTSGKGECKITCMQVLRRLAELDVEFEVEMDFEVASD I
 Juni_24457 50 EVKKLIQHRTLPAIVSYR-----LKSSRTSGKGD SKFLCLOVLRRLAELDVEFEVEMDYEVASD I
 Casg1106S06827 50 EIEELIKQRTLPTIIVCFR-----LQSSGTSKGKCKFTCSQVLRRLAELDVEFEVEMDYEVV SNI
 Bpev01_c0171-g0015 50 EVEELIKQRTLPSIVSFR-----LKSSRTSGKGDCKFTCLQVLRRLAELDVEFEVEMDYEVASD I
 ArDHQD-SDH5 50 EVQKLIKQRTLPSIVSFR-----LKSSGSSGKGDCKFTCLQVLRRLAELDVEFEVEMDYEVASD V
 ArDHQD-SDH6 1 -----MDYEVASDV

AtDHQD-SDH IK-SIDGKKPKGKFKVIVSSHNYQN-TPSVEDLGNLVARIQATGADIVKIATTAVDITDVARVFMHITSKAQV---PTIGLV
 FvDHQD-SDH3 ID-SINGNKPEKFKVIVSSHNYQD-TPSWKDLGDLVARIQATGADIVRIITTTALDITDVARVFIQISVSHV---PVI GLA
 FvDHQD-SDH4 VD-FIRDKKPEKFKVIVSSHNYQE-TPSVEALGNLVARIQATGADIVKIATTALDITDVARVFIQITVHSQV---PTI GLV
 NtDHQD-SDH1 NT-ALHGKNSAKCKVIVSSHNYDN-TPSSEELGNLVARIQASGADIVKFAATTALDITDVARVFOITVHSQV---PII AMV
 SlDHQD-SDH1 NN-ALHGKNSAKCKVIVSSHNYES-TPSAEDLGNLVARIQASGADIVKFAATTAQDITDVARVFOITVHSQV---PII AMV
 CasDHQD-SDH4 NS-SMHGKKPAKCKVIVSSHNYEN-TPSVEDLGNLVARIQATGADIVKIATTALDITDVARVFOITVHSQV RGVPII AMV
 VvDHQD-SDH1 IN-SIHGRKPEKFKVIVSSHNYQN-TPSVEDLGNLVSIQATGADIVKIATTALDITDVARVFOITVHSQV---PVI GLV
 FSB010265701 NE-SIYGKKPEKFKVIVSSHNYQN-TPSVEDLGNLVARIQATGADIVKFAATTALDITDVARVFOITVHSQV---PII AIV
 FSB012889301 NE-SIYGKKPEKFKVIVSSHNYQN-TPSVEDLGNLVARIQATGADIVKFAATTALDITDVARVFOITVHSQV---PII AIV
 Qrob_P0275560 NE-SIFGKKPEKFKVIVSSHNYHD-TPSVEDLSNLVARIQEAAGADIVKIATTALDITDVAHFQITVHSQV---PII AIV
 Juni_01730 ND-SIYGKKPEKFKVIVSSHNYLD-TPSVEDLGNLVARIQATGADIVKIATTALDITDVARVFOITVHSQV RSVPII GIV
 JrDHQD-SDH ND-SIYGRKPEKFKVIVSSHNYQD-TPSAEDLGNLVARIQATGADIVKIATTALDITDVARVFOITVHSQV---PII GIV
 Juni_19979 ND-SIYGRKPEKFKVIVSSHNYQD-TPSAEDLGNLVARIQATGADIVKIATTALDITDVARVFOITVHSQV---PII GIV
 ArDHQ-SDH1 ND-SIYGKKPEKFKVIVSSHNYLD-TPSVEDLGNLVARIQATGADIVKIATTALDITDVARVFOITVHSQV---PII AMV
 Alng188491S37275 ND-SIYGKKPEKFKVIVSSHNYLD-TPSVEDLGNLVARIQATGADIVKIATTALDITDVARVFOITVHSQV---PII AMV
 Bpev01_c0420-g0007 ND-SIYGKKPEKFKVIVSSHNYLD-TPSVEDLGNLVARIQATGADIVKIATTALDITDVARVFOITVHSQV---PII AMV
 Corav_882 -----MV
 Casg115S00444 ND-SICGKKPEKFKVIVSSHNYLD-TPSVEDLGNLVARIQATGADIVKIATTALDITDVARVFOISVHSQV---PFI GMV
 EgDHQD-SDH5 NA-SIQGRKPEKFKVIVSSHNYEN-TPSAEDLSNLVARIQAAAGADIVKIATTALDITDVARVFMHITVHSQV SSVPII AMV
 EcDHQD-SDH1 NA-SIQGRKPEKFKVIVSSHNYEN-TPSAEDLSNLVARIQAAAGADIVKIATTALDITDVARVFMHITVHSQV SSVPII GMV
 CsdhQD-SDH3 ND-SIRGKKPEKFKVIVSSHNYQY-TPSVEDLSNLVARIQASGADIVKFAATTALDITDVARVFOITVHSQV---PII GLV
 PoptrDHQD-SDH1 NE-LLRGNKPKGKFKVIVSSHNYEN-TPSVEELGNLVARIQAAAGADIVKIATTALDITDVARVFOITVHSQV RSVPII GLV
 PoptrDHQD-SDH3 IW-ELKNKHONGKFKVIVSSYLNGA-TPSKENLSHLVARIQATGADIVKIVSNADDITEMERIFHLLSHQEV---PAVAYS
 PoptrDHQD-SDH2 VR-EVKNKHQTGKFKVIVSSYLSGA-TPSKEDLSHLVASMQATGADIVKIVSNANDITELDRIFHLLSHSEV---PAVAYS
 EgDHQD-SDH4 KQ-KQNKLHCGGTKVIVSCFLDGV-TPSKEELSNLATHMKATGADIVKIIVTSASNITELAGLFHLLSYSQM---PVVAYA
 EcDHQD-SDH4a MS-KQNKLHCGGTKVIVSCFLDGV-TPSKEELSNLATHMKATGADIVKIIVTSASNITELAGLFHLLSYSQM---PVVAYA
 EcDHQD-SDH4b MS-KQNKLHCGGTKVIVSCFLDGV-TPSKEELSNLATHMKATGADIVKIIVTSASNITELARIFHLLSYSQM---PVVAYA
 CasDHQD-SDH1 ME-EHNSNKHRSSKIIIVSRYIDGT-TPSEEDLSQLVAHMOSTGADIVKLVSKSSSITELPRIFHLLSHQOI---PLIAYS
 SlDHQD-SDH2 AK-NEKSSWSSGOKLITSCFVDNV-TPSKEDLSQVVAHMOSTGADIVKIVINANDITBLEKMFHLLSHQOQ---PLIAYS
 NtDHQD-SDH2 TK-KEKPRWSSGOKVIVSCFVDNV-TPSKEDLSQVVAHMOSTGADIVKIVTNANDITELKMFHLLSHQOQ---PLIAYS
 CsdhQD-SDH1 LG-KQYSSHQSGTRFIVSONLDE-TPSEEDLGLVSRMQATGADIVKLVFSVNDITBIARIFOLLSHQOQ---PLIAYS
 VvDHQD-SDH2 LG-KQKMDQHSSTRTIVSCYVDGV-TPPTEDIICRVALLQSTGADMIKLVINATNITETIKIFHLLSHQOQ---PLIAYS
 Qrob_P0208710 ME-ELKMNCHSGSKVIVSCYVNGM-TPSEEYLSYLVANMQATGADIVKLVINATNITETITRIFHLLSHQOQ---PLIAYS
 FSB015730401 ME-ELKMNCHSGSKVIVSCYVNGM-TPSEEYLSYLVANMQATGADIVKLVINATNITETITRIFHLLSHQOQ---PLIAYS
 Juni_23842 RC-ELKMDHHDGSKVIVSCYVNGM-TPSLEDLSYLVASMQATGADIVKLVTKAIDITETISRIIFHLLSHQOQ---PLIAYS
 Bpev01_c1356-g0002 MG-ELRMDHHSKSKIIIVSRYVNSM-DSLEEDLSYLVASMQATGADIVKLVNTAITSITETISRIIFHLLSHQOQ---PLIAYS
 Corav_4151 KG-ELRMNHHSKSKIIIVSRYVNSM-ASLEEDLSYLVASMQATGADIVKLVNTVITITETIPRIIFHLLSHQOQ---PLIAYS
 Alng120613S19962 MG-ELRKNRHSGSKIIIVSRYVNSM-ASLEEDLSYLVASITQTGANIKLVANATSITETISRIIFHLLSHQOQ---PLIAYS
 ArDHQD-SDH2 MG-ELRKNRHSGSKIIIVSRYVNSM-ASLEEDLSYLVASITQTGANIKLVANATSITETISRIIFHLLSHQOQ---PLIAYS
 CasDHQD-SDH2 NN-SISAKKPEKFKVIVSSQNFHS-TPSAEAIAGNLVARIQATGADIVKIATTALDITDVARVFOITVHSQI---PMI GIA
 VvDHQD-SDH3 NN-SIYGKKPQNFKIVSSHNFHN-TPSTEAIAGNLVARIQASGADIVKIATTALDITDVARVFOITVHSQV---PTI AIV
 FvDHQD-SDH1 NN-SIYEKKPDNFKIVSSHNFHN-TPSSEAIAGNLVARIQATGADIVKIATTALDITDCARIFQITVHSQV---PTI GIV
 EgDHQD-SDH3 NS-SILGKKPDNFKIVSSHNFHN-TPSSEAIAGNLVARIQATGADIVKIATTALDITDCARIFQITVHSQI---PTI GIV
 EcDHQD-SDH3 NS-SILGKKPDNFKIVSSHNFHN-TPSSEAIAGNLVARIQATGADIVKIATTALDITDCARIFQITVHSQI---PTI GIV
 Qrob_P0208720 TS-SINGKPPDNFKVIVSSHNFHN-TPSAEAIAGNLVARIQATGADIVKIATTALDITDCARIFRIMVHSQV---PTI GIV
 FSB015730501 NS-SIYGSKPENEKIVSSHNFHN-TPSAEAIAGNLVARIQATGADIVKIATTALDITDCARIFQIMVHSQV---PTI GIV
 Corav_14055 NS-SIYGKKPENEKIVSSHNFHN-TPSAEAIAGNLVARIQATGADIVKIATTALDITDCARIFQIMVHSQV---PTI GIV
 Corav_11016 NS-SIYGKKPENEKIVSSHNFHN-TPSAEAIAGNLVARIQATGADIVKIATTALDITDCARIFQIMVHSQV---PTI GIV
 Bpev01_c1356-g0003 NS-SIYGKKPENEKIVSSHNFHN-TPSAEAIAGNLVARIQATGADIVKIATTALDITDCARIFQITVHSQV---PTI GIV
 ArDHQD-SDH3 NS-SIYGKKPENEKIVSSHNFHN-TPSAEAIAGNLVARIQATGADIVKIATTALDITDCARIFQITVHSQV---PTI GIV
 CasDHQD-SDH3 YK-LIQGKKPEKVKIIVSSHNYEN-TPSAEEIGNLAARIQATGADIVKIATTALDITDCARIFOLLAHSQV---PTI GIV
 DkDHQD-SDH FS-SIQEESPKRPFKIVSSHNYEN-TPSAEEIGNLAARIQATGADIVKIATTALDITDCARIFOLLAHSQV---PTI GIV
 VvDHQD-SDH4 IN-SIQGKTSKVKIIVSSHNYQN-TPSAEELGNLVARIQATGADIVKIATTALDITDCARIFQVLAHSQV---PTI GIV
 EgDHQD-SDH2 FS-SIQGKPEKAKIIVSSHNYQN-TPSSEELGNLVAKIQATGADIVKIATTALDISDCPRIFEVLAHSQV---PTI AIA
 EcDHQD-SDH2 FS-SIQGKPEKAKIIVSSHNYQN-TPSSEELGNLVAKIQATGADIVKIATTALDISDCPRIFEVLAHSQV---PTI AIA
 PoptrDHQD-SDH5 YN-FIQGKKPEKVKIIVSSHNYEC-TPSIEEIGDLVARIQATGADIVKIATTALDITDCARIFHIVNLQV---PMI GLV

FvDHQD-SDH2 YS-SIQGKKPERVKIIIVSSHNYES-TPSLEEIGNLVARIQATGADIVKATTALDITDNASTFOVLARSQV---PMIGLV
 Casg1267S13432 YN-TIQGKKPQKVKIIIVSSHNYQN-TPSLEEIGNLVARIQATGADIVKATTGLDITDSARIFQILVHSQV---PMIGIV
 Qrob_P0208730 YN-AIQGKKPEKVKIIIVSSHNYQN-TPSVEEIGDLVAKILSTGADIVKATTALDITDSARVFOVLVHSQV---PMIGIV
 FSB015730601 YN-TIQGKKPEKVKIIIVSSHNYQN-TPSVEEIGNLVARILATGADIVKATTGLDITDSARIFQVLVHSQV---PMIGLV
 Juni_23843 YN-AIQGKKPAKVKIIIVSSHNYQN-TPSLEEIGNLVARIQATGADIVKATTAVDITDSARIFQILVHSQV---PMIGIV
 ArDHQD-SDH4 YN-TIQGKKPEKVKIIIVSSHNYQN-TPSLEEIGNLVARILATGADIVKATTAVDITDSARIFQILVHSQV---PMIGIV
 Alng16638S34349 --GLPGKKPEKVKIIIVSSHNYQN-TPSLEEIGNLVARILATGADIVKATTAVDITDSARVFOVLVHSQV---PMIGIV
 Bpev01_c1356-g0004 YK-TIQGKKPEKVKIIIVSSHNYQN-TPSLEEIGNLVARIQATGADIVKATTAVDITDSARIFQVLVHSQV---PMIGLV
 Corav_941 YN-TIQGKKPEKVKIIIVSSHNYQN-TPSLEEIGNLVARIQATGADIVKATTAVDITDSARVFOILVHSQV---PMIGIV
 Corav_7742 YN-TIQGKKPEKVKIIIVSSHNYQN-TPSLEEIGNLVARIQATGADIVKATTAVDITDSARVFOILVHSQV---PMIGIV
 SLDHQD-SDH3 YV-AELMKSRNSKIIIVSYVNGG-NPFTKDTLNSIINLQSTGADIVKLVIDVAYITDPAVFHMLTHQV---PLIVRA
 EgDHQD-SDH1 NI-DELMKRSRNSKIIIVSRHLNGD-KPKCKEKLNLIAIMQSSGCDVTKLVLDVDYITDLAPLFQLLTHQV---PLIATT
 PoptrDHQD-SDH4 VM-AEYVYNSRNTKIIIVSSYLNGGKPTTEKLGDIACMQSTGADVTKLVLDVEKITDLAPVFTMLTHQV---PLIALA
 CsDHQD-SDH2 LM-SEIITYSRNTKIIIVSSYLNGGKPTTEKLGDIACMQSTGADVTKLEIIVDSITDLAPVFTMLTHQV---PLIALA
 Qrob_P0339740 GM-AECMVYRANSKIIIVSSYVNGG-KPSTNERLGNLIACMQSTGADVTKLVLDVDYITDLAPVQMLTHQV---PLIATA
 Juni_24457 IV-RAEHLVGRVNSRIIVSSYVNGG-KPSTEKLGDIACMQSTGADVTKLVLDVDYITDLAPVFTMLTHQV---PLIATA
 Casg1106S06827 GM-AEYVYHRAKSKIIIVSYVNGG-KPSTEKLGDIACMQSTGADVTKLVINVDYITDLAPVQMLTHQV---PLIAVA
 Bpev01_c0171-g0015 AM-AEHVYKRADSKIIIVSSYVNGG-KPSTEKLGDIACMQSTGADVTKLEINVDYITDLAPVQMLTHQV---PLIAVA
 ArDHQD-SDH5 GM-AEHVYKRADSKIIIVSSYVNGG-KPSTERLGHLIACMQSTGADVTKLEINVDYITDLAPVQMLTHQV---PLIAVA
 ArDHQD-SDH6 GM-AEHVYKRADSKIIIVSSYVNGG-KPSTERLGHLIACMQSTGADVTKLEINVDYITDLAPVQMLTHQV---PLIAVA

AtDHQD-SDH MGERGLMSRILCSKFGGILTFGTIDSSKVSAPGQPTIKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPIVHNAQAFKSV
 FvDHQD-SDH3 MGERGLISRILCAKFGGILTFGTIDSGILVSVPGQPAIKDILDVYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEGFKSVG
 FvDHQD-SDH4 MGERGLISRILCAKFGGILTFGTIDSGAVSAPGQPTIKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEGFKSVG
 NtDHQD-SDH1 MGERGLMSRILCSKFGGILTFGTILEVGVKVSAPGQPTIKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFRSVG
 SLDHQD-SDH1 MGERGLMSRILCPKFGGILTFGTILEVGVKVSAPGQPTVEDLLNLYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFRSVG
 CasDHQD-SDH4 MGERGLMARILCPKFGGILTFGTILEVGVKVSAPGQPTMDLVLNLYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKVVG
 VvDHQD-SDH1 MGERGLISRILCPKFGGILTFGTILEVGVKVSAPGQPTIKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSVG
 FSB010265701 MGERGLISRILCPKFGGILTFGTILEVGVKVSAPGQPTMKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSVG
 FSB012889301 MGERGLISRILCPKFGGILTFGTILEVGVKVSAPGQPTMKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSVG
 Qrob_P0275560 MGERGLISRILCPKFGGILTFGTILEVGVKVSAPGQPTMKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSVG
 Juni_01730 MGERGLISRILCPKFGGILTFGTILEVGVKVSAPGQPTMKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSVG
 JrDHQD-SDH MGERGLISRILCPKFGGILTFGTILEVGVKVSAPGQPTMKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSVG
 Juni_19979 MGERGLISRILCPKFGGILTFGTILEVGVKVSAPGQPTMKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSVG
 ArDHQ-SDH1 MGERGLISRILSPKFGGILTFGTILEVGVKVSAPGQPSMKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSAG
 Alng188491S37275 MGERGLISRILSPKFGGILTFGTILEVGVKVSAPGQPSMKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSAG
 Bpev01_c0420-g0007 MGERGLISRILCPKFGGILTFGTILEVGVKVSAPGQPTMKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSAG
 Corav_882 MGERGLISRILCPKFGGILTFGTILEVGVKVSAPGQPTMKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSAG
 Casg115300444 MGERGLISRILCPKFGGILTFGTILEVGVKVSAPGQPSMKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSAG
 EgDHQD-SDH5 MGERGLISRVLCAKFGGILTFGTILEVGVKVSAPGQPTIKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSAG
 EcDHQD-SDH1 MGERGLISRVLCAKFGGILTFGTILEVGVKVSAPGQPTIKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSAG
 CsDHQD-SDH3 MGERGLISRILCAKFGGILTFGTILEVGVKVSAPGQPTIKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSAG
 PoptrDHQD-SDH1 MGERGLISRILCAKFGGILTFGTILEVGVKVSAPGQPTIKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKSAG
 PoptrDHQD-SDH3 VGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLDSLRQAYKVDINSNTKVFGLISKPVSHSKSPILLNNEAFRHN
 PoptrDHQD-SDH2 IGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLDSLRQAYKVENINSNTKVFGLISKPVSHSKSPILLNNEAFRHN
 EgDHQD-SDH4 VGERGLISQLLSPKFGGILVYGSMEGN--AVPGLPTLES LRQAYKVEHINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 EcDHQD-SDH4a VGERGLISQLLSPKFGGILVYGSMEGN--AVPGLPTLES LRQAYKVEHINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 EcDHQD-SDH4b VGERGLISQLLSPKFGGILVYGSMEGN--AVPGLPTLES LRQAYKVEHINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 CasDHQD-SDH1 IGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLDSLRQAYKVDINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 SLDHQD-SDH2 VGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLDSLRQAYKVDINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 NtDHQD-SDH2 IGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLDSLRQAYKVDINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 CsDHQD-SDH1 VGERGLISQLLSPKFGGILVYGSMEGN--SIPGLPTLES LRQAYKVEHINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 VvDHQD-SDH2 IGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLES LRQAYKVEHINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 Qrob_P0208710 VGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLES LRQAYKVEHINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 FSB015730401 VGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLES LRQAYKVDINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 Juni_23842 VGERGLISQLLSPKFGGILVYGSMEGN--SIPGLPTLDSLRQAYKVEHINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 Bpev01_c1356-g0002 AGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLDSLRQAYKVDINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 Corav_4151 EGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLDSLRQAYKVDINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 Alng120613S19962 VGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLDSLRQAYKVEHINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 ArDHQD-SDH2 VGERGLISQLLCPKFGGILVYGSMEGN--SIPGLPTLES LRQAYKVEHINSNTKVFGLISKPVSHSKSPILLNNEAFRHM
 CasDHQD-SDH2 MGERGLISRLLSPKFGGILTYGALEAGAIAPGQPTVKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKVVG
 VvDHQD-SDH3 MGERGLISRLLSPKFGGILTYGALEAGAIAPGQPTAKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKVVG
 FvDHQD-SDH1 MGERGLISRLLSPKFGGILTYGALEAGAIAPGQPTAKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKVVG
 EgDHQD-SDH3 MGERGLISRLLSPKFGGILTYGALEAGAIAPGQPTAKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKVVG
 EcDHQD-SDH3 MGERGLISRLLSPKFGGILTYGALEAGAIAPGQPTAKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKVVG
 Qrob_P0208720 MGERGLISRLLSPKFGGILTYGALEAGAIAPGQPTAKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKVVG
 FSB015730501 MGERGLISRLLSPKFGGILTYGALEAGAIAPGQPTAKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKVVG
 Corav_14055 MGERGLISRLLSPKFGGILTYGALEAGAIAPGQPTAKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKVVG
 Corav_11016 MGERGLISRLLSPKFGGILTYGALEAGAIAPGQPTAKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKVVG
 Bpev01_c1356-g0003 MGERGLISRLLSPKFGGILTYGALEAGAIAPGQPTAKDLDLIYNFRIRGPDTKVYGIIGKPVSHSKSPILLNNEAFKVVG

ArDHQD-SDH3 MGGRGLISRILSPKFGGVLTYGALGAGASAPGQPTAKDLLDIYNFRLLRPDTPVYGIIGKPVGHSKSPLLFNAAFKSVG
 CasDHQD-SDH3 MGGRGLMSRILCAKFGGFLTFGALGALGAGASAPGQSTLKLDDLEIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFRSAG
 DkDHQD-SDH MGGRGLISRILCAKFGGFLTFGALGALGAGASAPGQPTLRDLDDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKAA
 VvDHQD-SDH4 MAERGLISRILSAKFGSVLTFGSLGALGAGASAPGQPTVKDLLDIYNFRLLGPDTKVHGVIGKPIGHSKSPHLYNNAFAFKSVG
 EgDHQD-SDH2 MGGRGLISRILAAKFGGVLTFGALGAGASAPGQPSIKDLLDIYNFRLLGPDTKVHGVIGKPIGHSKSPHLYNNAFAFRSVN
 EcDHQD-SDH2 MGGRGLISRILAAKFGGVLTFGALGAGASAPGQPSIKDLLDIYNFRLLGPDTKVHGVIGKPIGHSKSPHLYNNAFAFKSVN
 PoptrDHQD-SDH5 MGGRGLISRILAAKFGGVLTFGALGAGASAPGQPTVKDLLDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKSVG
 FvDHQD-SDH2 MGGRGLISRILAAKFGGVLTFGALGAGASAPGQPTVKDLLDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKANN
 Casg1267S13432 MGDKGLISRILSAKFGGFLTFGSLGALGAGASAPGQPSIKDLLDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKSVN
 Qrob_P0208730 MGBKGLMSRILSAKFGGFLTFGSLGALGAGASAPGQPTVKDLLDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKSVN
 FSB015730601 MGBKGFISRILSAKFGGFLTFGSLGALGAGASAPGQPTVKDLLDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKSVN
 Juni_23843 MGBKGFISRILSAKFGGFLTFGSLGALGAGASAPGQPTVKDLLDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKSVN
 ArDHQD-SDH4 MGRGRFISRILSAKFGGFLTFGSLGALGAGASAPGQPSIKDLLDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKSVN
 Alng16638S34349 MGRGRFISRILSAKFGGFLTFGSLGALGAGASAPGQPSIKDLLDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKSVN
 Bpev01_c1356-g0004 MGRGRFISRILSAKFGGFLTFGSLGALGAGASAPGQPSIKDLLDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKSVN
 Corav_941 MGRGRFISRILSAKFGGFLTFGSLGALGAGASAPGQPSIKDLLDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKSVN
 Corav_7742 MGRGRFISRILSAKFGGFLTFGSLGALGAGASAPGQPSIKDLLDIYNFRLLGPDTKVHGVIGNPIGHSKSPHLYNNAFAFKSVN
 SlDHQD-SDH3 AGDRGLISQLLGPFGGFLVYGSLEGG--YTPGLPSTLTIKQVYKLYVNPDTRIFGVISNVPVGHSGKPLLHNAFAFRHTG
 EgDHQD-SDH1 VGRGLISQLLGPFGGFLVYGSLEGG--AIPGLPTLTLVSLKQVYKLYVNPDTRIFGVISNVPVGHSGKPLLHNAFAFRHTG
 PoptrDHQD-SDH4 VGRGLISQLLGPFGGFLVYGSLEGG--AVPGMPTLTLVSLKQVYKLYVNPDTRIFGVISNVPVGHSGKPLLHNAFAFRHTG
 CsDHQD-SDH2 VGRGLISQLLGPFGGFLVYGSLEGG--SVPGMPTLTLVSLKQVYKLYVNPDTRIFGVISNVPVGHSGKPLLHNAFAFRHTG
 Qrob_P0339740 VGRGLISQLLGPFGGFLVYGSLEGG--SVPGMPTLTLVSLKQVYKLYVNPDTRIFGVISNVPVGHSGKPLLHNAFAFRHTG
 Juni_24457 AGSRGLISQLLGPFGGFLVYGSLEGG--SIPGLPTLTLVSLKQVYKLYVNPDTRIFGVISNVPVGHSGKPLLHNAFAFRHTG
 Casg1106S06827 VGRGLISQLLGPFGGFLVYGSLEGG--SVPGMPTLTLVSLKQVYKLYVNPDTRIFGVISNVPVGHSGKPLLHNAFAFRHTG
 Bpev01_c0171-g0015 VGRGLISQLLGPFGGFLVYGSLEGG--SVPGMPTLTLVSLKQVYKLYVNPDTRIFGVISNVPVGHSGKPLLHNAFAFRHTG
 ArDHQD-SDH5 VGRGLISQLLGPFGGFLVYGSLEGG--SVPGMPTLTLVSLKQVYKLYVNPDTRIFGVISNVPVGHSGKPLLHNAFAFRHTG
 ArDHQD-SDH6 VGRGLISQLLGPFGGFLVYGSLEGG--SVPGMPTLTLVSLKQVYKLYVNPDTRIFGVISNVPVGHSGKPLLHNAFAFRHTG

381

AtDHQD-SDH FNGYVYVHLLVDDNLSVSLFQAYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNTILRRKSDGKLL
 FvDHQD-SDH3 FDCGYVYVHLLVDDIANFLHTYSSSDFSGF-----SVCIPHKEAALQCCDEVDPLAKSIGAVNCIVRRPTDGKLF
 FvDHQD-SDH4 FNGYVYVHLLVDDIANFLKTYSSSDFAGF-----SVTIIPHKEAALQCCDEVDPLAKSIGAVNCIIRRENDGKLY
 NtDHQD-SDH1 FNGYVYVHLLVDDIANFLFRTYSSSDFAGS-----AVTIIPHKEAALQCCDEVDPLAKSIGAVNCVVSRLDGGKLF
 SlDHQD-SDH1 FNGYVYVHLLVDDIANFLFRTYSSSDFAGS-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRRE-DGKLF
 CasDHQD-SDH4 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIVRRKSDGKLF
 VvDHQD-SDH1 FNGYVYVHLLVDDIANFLFHTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRRES-DGKLF
 FSB010265701 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----SITIPHKEAALQCCDEVDPLAKSIGAVNCIVRRST-DGKLF
 FSB012889301 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----SITIPHKEAALQCCDEVDPLAKSIGAVNCIIRRS-DGKLF
 Qrob_P0275560 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIVRRPTDGKLF
 Juni_01730 FNGYVYVHLLVDDIANFLFKTYSSSDFSGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRREPTDGKLI
 JrDHQD-SDH FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRREPTDGKLV
 Juni_19979 FNGYVYVHLLVDDIANFLFRTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRREPTDGKLV
 ArDHQD-SDH1 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIVRRPTDGKLF
 Alng188491S37275 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIVRRPTDGKLF
 Bpev01_c0420-g0007 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRREPTDGKLF
 Corav_882 FNGYVYVHLLVDDIANFLFRTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRREPTDGKLF
 Casg115S00444 FDCGYVYVHLLVDDIANFLFRTYSSSDFAGFRFIFV-EHYES SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRREPTDGKLF
 EgDHQD-SDH5 FDCGYVYVHLLVDDIANFLQTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRROSDGKLF
 EcDHQD-SDH1 FDCGYVYVHLLVDDIANFLQTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRROSDGKLF
 CsDHQD-SDH3 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRROSDGKLF
 PoptrDHQD-SDH1 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----SCTIPHKEAALQCCDEVDPLAKSIGAVNCIIRROSDGKLF
 PoptrDHQD-SDH3 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGY SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 PoptrDHQD-SDH2 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGY SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 EgDHQD-SDH4 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 EcDHQD-SDH4a FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 EcDHQD-SDH4b FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 CasDHQD-SDH1 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 SlDHQD-SDH2 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 NtDHQD-SDH2 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 CsDHQD-SDH1 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 VvDHQD-SDH2 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 Qrob_P0208710 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----EFAGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 FSB015730401 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 Juni_23842 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 Bpev01_c1356-g0002 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 Corav_4151 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 Alng120613S19962 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 ArDHQD-SDH2 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI
 CasDHQD-SDH2 FNGYVYVHLLVDDIANFLQTYSSSDFAGF-----AGF SVGFIPYKEAVVQFCDEVDPLAKSIGAVNTIIRRPDGGKLI

VvDHQD-SDH3 LNNAVYVHLLVDDVEKFFNTYSAPDFISGC-----SCTIPHKEWAIKOMDIDPIAKKIGAINNIVRRP-DGKLT
 FvDHQD-SDH1 LNNAVYVHFLVDDVEKFFNTYSVDFSGC-----SCTIPHKEAALKOMDEIDPIAKKIGAINNIVRRP-DGRLV
 EgDHQD-SDH3 LNNAVYVHFLVDDVEKFFNTYSVDFASGC-----SCTIPHKEAALKOMDEIDPIAKKIGAINNIVRRP-DGTLT
 EcDHQD-SDH3 LNNAVYVHFLVDDVEKFFNTYSVDFASGC-----SCTIPHKEAALKOMDEIDPIAKKIGAINNIVRRP-DGTLT
 Qrob_P0208720 LDNAVYVHYLVDDVEKFFNTYSVDFASGC-----SCTIPHKEAALKOMDEIDPIAKKIGAINNIVRRP-DGKLI
 FSB015730501 LDNAVYVHYLVDDVEKFFNTYSVDFWCGC-----SCTIPHKEAALKOMDIDPIAKKIGAINNIVRRP-DGKLT
 Corav_14055 LNNAVYVHFLVDDVKKFFDFTYSIIDWA-GC-----SCTIPHKEAALROMNEIDPIAKKIGAINNIVRRP-DGTLT
 Corav_11016 LNNAVYVHFLVDDVKKFFDFTYSIIDWA-GC-----SCTIPHKEAALROMNEIDPIAKKIGAINNIVRRP-DGTLT
 Bpev01_c1356-g0003 LNNAVYVHYLVDDVKKFFDFTYSVDFWAGC-----SCTIPHKEAALKOMNEIDPIAKKIGAINNIVRRP-DGTLT
 ArDHQD-SDH3 LNNAVYVHYLVDDVKKFFDFTYSVDFWAGC-----SCTIPHKEAALKOMNEIDPIAKKIGAINNIVRRP-DGTLT
 CasDHQD-SDH3 FNGIYLPPLLVDSVANFNNTYSVDFV-GY-----SYTIPHKEAGPKCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 DkDHQD-SDH FNGIYLPPLLVDNVANFNNTYSVDFV-GY-----SYTIPHKEAGLECCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 VvDHQD-SDH4 FNGIYLPPLLVDSVKNFLNTYSVDFV-GY-----SYTIPHKEAGLECCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 EgDHQD-SDH2 FNGIYLPPLLVDNVANFNINAYSSVDFV-GY-----SYTIPHKEAGLECCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 EcDHQD-SDH2 FNGIYLPPLLVDNVANFNINAYSSVDFV-GY-----SYTIPHKEAGLECCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 PoptrDHQD-SDH5 FNGIYLPPLLVDSVANFYISTYSVDFV-GY-----SYTIPHKEAGLECCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 FvDHQD-SDH2 FNGIYLPPLLVDSVANFNINTYNSVDFV-GY-----SYTIPHKEAGFKCCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 Casg1267S13432 FNGIYLPPLLVDNVANFLKAYSSVDFV-GY-----SYTIPHKEAGFNCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 Qrob_P0208730 FNGIYLPPLLVDNVANFNNTYSVDFV-GY-----SYTIPHKEAGFNCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 FSB015730601 FNGIYLPPLLVDNVANFNNTYSVDFV-GY-----SYTIPHKEAGFNCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 Juni_23843 FNGIYLPPLLVDNVANFLKTYSSVDFV-GY-----SYTIPHKEAGFNCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 ArDHQD-SDH4 FNGIYLPPLLVDNVANFLKTYSSVDFV-GY-----SYTIPHKEAGFNCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 Alng16638S34349 FNGIYLPPLLVDNVANFLKTYSSVDFV-GY-----SYTIPHKEAGFNCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 Bpev01_c1356-g0004 FNGIYLPPLLVDNVANFLKTYSSVDFV-GY-----SYTIPHKEAGFNCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 Corav_941 FNGIYLPPLLVDNVANFLKTYSSVDFV-GY-----SYTIPHKEAGFNCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 Corav_7742 FNGIYLPPLLVDNVANFLKTYSSVDFV-GY-----SYTIPHKEAGFNCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 SlDHQD-SDH3 YNGIYVPLLVDDNVKEFFRVTYSVDFV-GY-----AGF SVGIPHKEAAVRCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 EgDHQD-SDH1 YNGIYVPLLVDDNVKEFFRVTYSVDFV-GY-----AGF SVGIPHKEAAVRCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 PoptrDHQD-SDH4 YNGIYVPLLVDDNVKEFFRVTYSVDFV-GY-----AGF SVGIPHKEAAVRCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 CsDHQD-SDH2 FNGIYVPLLVDDNVKEFFRVTYSVDFV-GY-----AGF SVGIPHKEAAVRCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 Qrob_P0339740 YNGIYVPLLVDDNVKEFFRVTYSVDFV-GY-----AGF SVGIPHKEAAVRCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 Juni_24457 YNGIYVPLLVDDNVKEFFRVTYSVDFV-GY-----AGF SVGIPHKEAAVRCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 Casg1106S06827 YNGIYVPLLVDDNVKEFFRVTYSVDFV-GY-----AGF SVGIPHKEAAVRCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 Bpev01_c0171-g0015 YNGIYVPLLVDDNVKEFFRVTYSVDFV-GY-----AGF SVGIPHKEAAVRCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 ArDHQD-SDH5 YNGIYVPLLVDDNVKEFFRVTYSVDFV-GY-----AGF SVGIPHKEAAVRCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI
 ArDHQD-SDH6 YNGIYVPLLVDDNVKEFFRVTYSVDFV-GY-----AGF SVGIPHKEAAVRCDEVDPIDPIAKKIGAINNIVRRP-PSDGKLI

AtDHQD-SDH GYNTDYVGAISAIEDGLRSG-----DPSVSPSSSSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 FvDHQD-SDH3 GYNTDYVGAISAIEDGLRSG-----D--KSNLIGSPLAGRLEFVIM GAGGAGRALAYGAKQKGAR
 FvDHQD-SDH4 GYNTDYVGAISAIEDGLRSG-----N--GSHVTGSPLAGRLEFVVI GAGGAGRALAYGAKQKGAR
 NtDHQD-SDH1 GYNTDYVGAISAIEDGLRSG-----P--SMSGSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 SlDHQD-SDH1 GYNTDYVGAISAIEDGLRSG-----P--SISGSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 CasDHQD-SDH4 GYNTDYVGAISAIEDGLRSG-----H--VNGMVVSPSGLKLFVVI GAGGAGRALAYGAKQKGAR
 VvDHQD-SDH1 GYNTDYVGAISAIEDGLRDLH-----K--ISSTS GSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 FSB010265701 GYNTDYVGAISAIEDGLRSG-----N--ISYAADSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 FSB012889301 GYNTDYVGAISAIEDGLRSG-----N--ISYAADSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 Qrob_P0275560 GYNTDYVGAISAIEDGLRSG-----N--ISATADSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 Juni_01730 GYNTDYVGAISAIEDGLRSG-----N--SNNTGDSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 JrDHQD-SDH GYNTDYVGAISAIEDGLRSG-----N--SSNTADSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 Juni_19979 GYNTDYVGAISAIEDGLRSG-----N--SSNTADSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 ArDHQ-SDH1 GYNTDYVGAISAIEDGLRSG-----N--ISNTADSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 Alng188491S37275 GYNTDYVGAISAIEDGLRSG-----N--ISNTANSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 Bpev01_c0420-g0007 GYNTDYVGAISAIEDGLRSG-----N--ISNTADSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 Corav_882 GYNTDYVGAISAIEDGLRSG-----N--ISNTADSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 Casg115S00444 GYNTDYVGAISAIEDGLRSG-----NN--ISNTADSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 EgDHQD-SDH5 GYNTDYVGAISAIEDGLRSG-----N--GNSAGASPLNGKLFVVI GAGGAGRALAYGAKERGAK
 EcDHQD-SDH1 GYNTDYVGAISAIEDGLRSG-----N--GNSAGASPLNGKLFVVI GAGGAGRALAYGAKERGAK
 CsDHQD-SDH3 GYNTDYVGAISAIEDGLRSG-----N--VSGGVSSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 PoptrDHQD-SDH1 GYNTDYVGAISAIEDGLRSG-----N--VSNVTGSPLAGKLFVVI GAGGAGRALAYGAKERGAK
 PoptrDHQD-SDH3 GYNTDCEGSIATIEDALRD-----QKYVNG-RSLNSPLAGKLFVVI GAGGAGRALAYGAKSRGAR
 PoptrDHQD-SDH2 GYNTDCEGSIATIEDALRD-----QRYING-ASLNSPLAGKLFVVI GAGGAGRALAYGAKSRGAR
 EgDHQD-SDH4 GYNTDCEASVTIATEDALRD-----CRCINGEKSLVSPLAGKLFVVI GAGGAGRALAYGAKTRGAR
 EcDHQD-SDH4a GYNTDCEASVTIATEDALRD-----CRCINGEKSLVSPLAGKLFVVI GAGGAGRALAYGAKTRGAR
 EcDHQD-SDH4b GYNTDCEASVTIATEDALRD-----CRCINGEKSLVSPLAGKLFVVI GAGGAGRALAYGAKTRGAR
 CasDHQD-SDH1 GYNTDCEASITATEDALKV-----WGCTNGEVSPLSPLNGKLFVVI GAGGAGRALAYGAKSRGAR
 SlDHQD-SDH2 GYNTDCEASITATEDALKV-----ANGELVPCSLARKLFVVI GAGGAGRALAYGAKSRGAR
 NtDHQD-SDH2 GYNTDCEASITATEDALKV-----NGLTNGAAFLPSPLAGKLFVVI GAGGAGRALAYGAKSRGAR
 CsDHQD-SDH1 GYNTDCEASITATEDALKE-----RGYKNGTASFSPLAGRLEFVVI GAGGAGRALAYGAKSRGAR
 VvDHQD-SDH2 GYNTDCEASITATEDALRE-----RGLPNGEAPLNSPLNGKLFVVI GAGGAGRALAYGAKSRGAR
 Qrob_P0208710 GYNTDREASITATEDALKH-----QGYTNGETSFSPLAGKLFVVI GAGGAGRALAYGAKSRGAR

FSB015730401 GVNNDCEASITATEDALKQ-----RGCTNGEATFNSPLAGKQFVLVLAGGAGRALAFGAKSRGAQ
 Juni_23842 GVNNDCEASICATEDALKE-----QGCTNNGASLSSPLAGKQFVLVLAGGAGRALAFGAKSRGAW
 Bpev01_c1356-g0002 GVNNDCEASITATEDSLKE-----QGCTNCEASFSSPLAGKLFVVLVLAGGAGRALAFGAKSRGAR
 Corav_4151 GVNNDCEASITATEDALKE-----QGCTNCEASFSSPLAEKLFVVLVLAGGAGRALAFGAKSRGAR
 Alngl20613S19962 GVNNDCEASITATEDALKE-----QGCTNCEAPFSSPLAGKQFVLVLAGGAGRALAFGAKSRGAR
 ArDHQD-SDH2 GVNNDCEASITATEDALKE-----QGCTNCEAPFSSPLAGKQFVLVLAGGAGRALAFGAKSRGAR
 CasDHQD-SDH2 SVNNDYIGAI SAIEDGLRES-----NGSSPATSSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 VvDHQD-SDH3 AFNNDYIGAI SAIEDGLRES-----NGSSPAVGSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 FvDHQD-SDH1 AFNNDYIGAI SAIEDGLRGM-----NGAIPAGKSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 EgDHQD-SDH3 AFNNDYIGAI SAIEDGLRGL-----NVVSPGA-SPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 EcDHQD-SDH3 AFNNDYIGAI SAIEDGLRGL-----NGVSSGA-SPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 Qrob_P0208720 AFNNDYIGAI SAIEDGLRGL-----NGTIPAGSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 FSB015730501 AFNNDYIGAI SAIEDGLRGL-----NGTNPAAGSPLAGKLFVVMGAGGAGRALAFGAAQK GAR
 Corav_14055 AFNNDYIGAI SAIEDGLRGL-----NGTHPAVGSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 Corav_11016 AFNNDYIGAI SAIEDGLRGL-----NGTHPAVGSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 Bpev01_c1356-g0003 AFNNDYIGAI SAIEDGLRGL-----NGKHPAAGSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 ArDHQD-SDH3 AFNNDYIGAI SAIEDGLRGL-----NGKHPAAGSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 CasDHQD-SDH3 GVNVDYLGAIAAIEEALGGS-----NGSSSNAVSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 DkDHQD-SDH GVNVDYLGAIAAIEEALGGS-----SSANNGSVSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 VvDHQD-SDH4 GVNVDYLGAIAAIEEGLRAS-----NGTTSVGSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 EgDHQD-SDH2 GVNVDYLGAIAAIEEALRAS-----NGASSTTSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 EcDHQD-SDH2 GVNVDYLGAIAAIEEALRAS-----NGASPTTSPLAGKLFVVLVLAGGAGRALAFGAAQK GAR
 PoptrDHQD-SDH5 GVNVDYLGAIAAIEEALGAS-----NGAPASVSPLAGKLFVVMGAGGAGRALAFGAAQK GAR
 FvDHQD-SDH2 GVNVDYLGAIAAIEEGLRGLG-----LNGSNNGSGSPLAGKLFVVMGAGGAGRALAFGAAQK GAR
 Casg1267S13432 GVNVDYLGAIAAIEEGLRAS-----NGASNGSGSPLAGKLFVVMGAGGAGRALAFGAAQK GAR
 Qrob_P0208730 GVNVDYLGAIAAIEEGLRAS-----NGASNGSGSPLAGKLFVVMGAGGAGRALAFGAAQK GAR
 FSB015730601 GVNVDYLGAIAAIEEGLRAS-----NGASNGSGSPLAGKLFVVMGAGGAGRALAFGAAQK GAR
 Juni_23843 GVNVDYLGAIAAIEEGLRAS-----NGASQASGSPLAGKLFVVMGAGGAGRALAFGAAQK GAR
 ArDHQD-SDH4 GVNVDYLGAIAAIEEGLRA-----NGANGASGSPLAGKLFVVMGAGGAGRALAFGAAQK GAR
 Alngl6638S34349 GVNVDYLGAIAAIEEGLRA-----NGANGASGSPLAGKLFVVMGAGGAGRALAFGAAQK GAR
 Bpev01_c1356-g0004 GVNVDYLGAIAAIEEGLRAS-----NGASPVSGSPLAGKLFVVMGAGGAGRALAFGAAQK GAR
 Corav_941 GVNVDYLGAIAAIEEGLRAS-----NGASRVSGSPLAGKLFVVMGAGGAGRALAFGAAQK GAR
 Corav_7742 -----AVRHR-----NGSNEARQGD-----
 SlDHQD-SDH3 GVNNDCEACVITATEDALRE-----RQKTNGHAS--NVSPLAGKLFVVLVLAGGAGRALAFGAKSRGAR
 EgDHQD-SDH1 GVNNDCEASITATEDALRE-----RHAANGEARAMDASPLAGKLFVVLVLAGGAGRALAFGAKSRGAR
 PoptrDHQD-SDH4 GVNNDCEASISATEDALRE-----RRITQKGVLL--EASPLSGKLFVVLVLAGGAGRALAFGAKSRGAR
 CsDHQD-SDH2 GVNNDCEASISATEDALRE-----RQGINGVAS--HTSPLAGKLFVVLVLAGGAGRALAFGAKSRGAR
 Qrob_P0339740 GVNNDCEASITATEDALRE-----RRVANGAAL--RTSPLAGKLFVVLVLAGGAGRALAFGAKSRGAR
 Juni_24457 GVNNDCEASITATEDALRGVVERAYWLKVVLFSGEKYRLLKGEAS--DTSPLAGKLFVVLVLAGGAGRALAFGAKSRGAR
 Casg1106S06827 GVNNDCEASITATEDALGCG-----RKVDHGEAS--YASPLSGKLFVVLVLAGGAGRALAFGAKSRGAR
 Bpev01_c0171-g0015 GVNNDCEASITATEDALRD-----RRVKNGDAS--HASPLAGKLFVVLVLAGGAGRALAFGAKSRGAR
 ArDHQD-SDH5 GVNNDCEACITATEDALKD-----KRVENDAS--HASPLAGKLFVVLVLAGGAGRALAFGAKSRGAR
 ArDHQD-SDH6 GVNNDCEACITATEDALKD-----KRVENDAS--HASPLAGKLFVVLVLAGGAGRALAFGAKSRGAR

483

AtDHQD-SDH VVIANRNTYERARELAEATIGGKALSLLDLDNYHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 FvDHQD-SDH3 IIVIANRNTYDRARLADAVGGDALPFADLANFHPEDGAILANSTSVGMOPNIDETPIPKHALRSYSLVFDVAVYTPKTRLL
 FvDHQD-SDH4 IIVIANRNTYDRARELADTIGGAEALSISDLDFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 NtDHQD-SDH1 VVIANRNTYERARELADVGGQALSLEDELSNFPENDMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 SlDHQD-SDH1 VVIANRNTYERARELADVGGAEALSLEDELSNFPENDMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 CasDHQD-SDH4 VVIANRNTYERARELADTIGGDAISLADLSNFPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 VvDHQD-SDH1 VVIANRNTYERARELADAVGGDALSLADLNNFHPENGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 FSB010265701 VVIANRNTYERARELADTVGGQALSADLLENFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 FSB012889301 VVIANRNTYVRS-----
 Qrob_P0275560 VVIANRNTYDRARELANTIGGDVAVTLADLENFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 Juni_01730 VVIANRNTYDRARELADTIGGDALSADLDFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 JrDHQD-SDH VVIANRNTYDRARELADTIGGDALSADLDFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 Juni_19979 VVIANRNTYDRARELADTIGGDALSADLDFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 ArDHQ-SDH1 VVIANRNTYDRARELSDIVGGDALSLADLENFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 Alngl8849S137275 VVIANRNTYDRARELSDIVGGDALSLADLENFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 Bpev01_c0420-g0007 VVIANRNTYDRARELADTIGGDALSVDLDLNFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 Corav_882 VVIANRNTYDRARELADTIGGDALSADLDFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 Casg115S00444 VVIANRNTYDRARELADTIGGDALSADLDFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 EgDHQD-SDH5 VVIANRNTYDRARELAETIGGDALSADLDFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 EcDHQD-SDH1 VVIANRNTYDRARELAETIGGDALSADLDFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 CsDHQD-SDH3 VVIANRNTYDRARELAETIGGDALSADLDFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 PoptrDHQD-SDH1 VVIANRNTYERAKVLADIIIGGDAITLADLENFHPEDGMILANITSIGMOPKVDETPIPKHALRSYSLVFDVAVYTPKTRLL
 PoptrDHQD-SDH3 LIIIFDIDFERAKSLARAVSGEAOHFESLAHFQPEKGAITLANATPIGMFPSTDRIPVAEATLGNVQLVFDVAVYTPKTRLL
 PoptrDHQD-SDH2 LIIIFDIDDRAKSLAQVVSGEAOHFESLAHFQPEKGAITLANATPIGMFPSTDRIPVAEATLGNVQLVFDVAVYTPKTRLL
 EgDHQD-SDH4 LIIIFDIDFERAKMLAHAVSGEARPFQDIPYFOPEKGSITLANATPIGMFPNKDRIPVSEVPPYK-----

EcDHQD-SDH4a I I I F D I D F E R A K M L A H A V S G E A R P F G D I P Y F O P E K G S I L A N A T P I G M F P N K D R I P V S E A S L G V F O L V F D A V Y T P R K T I L L
 EcDHQD-SDH4b I I I F D I D F E R A K M L A H A V S G E A R P F G D I P Y F O P E K G S I L A N A T P I G M F P N K D R I P V S E A S L G V F O L V F D A V Y T P R K T I L L
 CasDHQD-SDH1 V V I F D I D F D R A K S L A L A V S G E A R P F E N L V S F O P E K G A I L A N A T P L G M F P N T D R I P V A E T L S Y Q L V F D A V Y T P R K T R L L
 SlDHQD-SDH2 I V I F D I D F D R A K A L A A A V S G E A L P F F E K L A S F O P E K G A I L A N A T P I G M F P N K D R I P V P E G S L K D Y V V V F D A V Y T P R R T I L L
 NtDHQD-SDH2 I V I F D I D F D R A K A L A A A V S G E A L P F F E N L A S F O P E K G A I L A N A T P I G M F P N K D R I P V S E A S L K D Y V V V F D A V Y T P R K T I L L
 CsDHQD-SDH1 V V I F D I D F E R A K S L A S D V M G A A R P F F E D I L N F O P E K G A I L A N A T P L G M F P N T D R V P V S E E T L R D Y Q L V F D A V Y T P R K T R L L
 VvDHQD-SDH2 I V I F D I D F D R A N S L A H A V S G E A R L Y E D V A N F O P E K G A I L A N A T P V G M F P N T D R I P V A E E T L S Y Q L V F D S V Y T P R K T R L L
 Qrob_P0208710 V V I F D I D F E R A K I L A H E V S G E A R P F F E E V N F H P E K G A I L A N A T P L G M F P N T D R I P V A E V T L Q D Y Q L V F D S V Y T P R K T R L L
 FSB015730401 I V I F D I D F D R A N S L A C A V S G E A R P F F E L V N F H P E K G A I L S N A T P L G M F P N T D R I P V A E A T L Q D Y Q L V F D S V Y T P R K T R L L
 Juni_23842 V I I F D I D F D R A K S L A C A V S G E A R P Y N D I V N F O P E K G A I L A N A T P L G M F P N T D R I P V A E A T L G D Y Q L V F D S V Y T P R K T R L L
 Bpev01_c1356-g0002 V I I F D I D F D R A K S L A D A V S G E A R P F E D I V N F O P N K G S I L A N A T P L G M F P N T D R I P V S E A T L A B Y Q L V F D S V Y T P R K T R L L
 Corav_4151 V I I F D I D F D R A K S L A C A V S G E A R P F E D I V N F O P S K G A I L A N A T P L G M F P N T D R I P V S E A T L A B Y Q L V F D S V Y T P R K T R L L
 Alngl20613S19962 V I I F D I D F D R A K S L A C A V S G E A R P F E D I V N F O P N K G A I L A N A T P L G M F P N T D R I P V S E A T L A B Y Q L V F D S V Y T P R K T R L L
 ArDHQD-SDH2 V I I F D I D F D R A K S L A C A V S G E A R P F E D I V N F O P N K G A I L A N A T P L G M F P N T D R I P V S E A T L A B Y Q L V F D S V Y T P R K T R L L
 CasDHQD-SDH2 V V V A N R T F E R A K E L A E K V G G K A L T L E E V N D F H P E E G M I L A N T S V G M F P N I D E T P I S K E A L K H Y D L V F D A I Y T P K D T R L L
 VvDHQD-SDH3 V V V A N R T F E R A K L A D K V G G E A M T L A E I N F H P E E G M I L A N T S V G M F P N I D T P I P K H A L K H Y S L V F D A I Y T P K D T R L L
 FvDHQD-SDH1 V V V A N R T F E R A K E L A D K V G G Q A M T L E E V N F H P E E G M I L A N T S V G M F P N V D D T P I S K Q A L K H Y C L V F D A I Y T P K E T R L L
 EgDHQD-SDH3 V V V A N R T F E R A K E L A D K V G G Q A M T L A E V N F H P E E G M I L A N T S V G M F P K I D E T P L A K H A L K N Y C L V F D A I Y T P K D T R L L
 EcDHQD-SDH3 V V V A N R T F E R A K E L A D K V G G Q A M T L A E V N F H P E E G M I L A N T S V G M F P K I D E T P L A K H A L K H Y C L V F D A I Y T P K D T R L L
 Qrob_P0208720 V V V A N R T F E R A K E L A D K V G G E A M T L A E I N F H P E E G M V L A N T S V G M F P K I D E T P I S K E A L K H Y C V F D A I Y T P K D T R L L
 FSB015730501 V V V A N R T F E R A K E L A D K V G G Q A I T L A E V D N F H P E E G M V L A N T S L G M F P K I D E T P I S K E S L K N Y C L V F D A I Y T P K E T R L L
 Corav_14055 V V V A N R T F E R A K E L A D K V G G Q A M T L A E V D S F H P E E G M V L A N T S V G M F P K I D E S P I S K Q S L K H Y C L V F D A I Y T P K E T R L L
 Corav_11016 V V V A N R T F E R A K E L A D K V G G Q A M T L A E V D S F H P E E G M V L A N T S V G M F P K I D E S P I S K Q S L K H Y C L V F D A I Y T P K E T R L L
 Bpev01_c1356-g0003 V V V A N R T F E R A K E L A D K V G G E A M T L A E V N F H P E E G M V L A N T S L G M F P K I D E A P I S K Q S L K H Y C L V F D A I Y T P K E T R L L
 ArDHQD-SDH3 V V V A N R T F E R A K E L A D K V G G Q A M T L A E V D N F H P E E G M V L A N T S V G M F P K I D E A P I S K Q S L K H Y C L V F D A I Y T P K D T R L L
 CasDHQD-SDH3 V V V A N R T Y E K A K E L A S K V G G E A I T L A E L D D F H P E D G M I L A N T S V G M F P K T D A T P I S K K A L N R Y S L V F D A I Y T P K W T R L L
 DkDHQD-SDH V V V A N R T Y E K A K D L A R K T G G E S M P L T E L N D F H P E D G M I L A N T S V G M F P N T D A T P I S K E A L S R Y S L V F D A I Y T P K W T R L L
 FvDHQD-SDH4 V V V A N R T Y E K A K E L A S K V G G E A M T L A E L E N F H P E D G M I L A N T S V G M F P N I D N T P L S K A L S R Y S L V F D A I Y T P K L T R L L
 EgDHQD-SDH2 V V V A N R T Y E K A K E L A S K V G G Q A I T L A E L E N F H P E D G M V L A N T S V G M F P N V D T P L P K N A L S R Y C L V F D A I Y T P K L T R L L
 EcDHQD-SDH2 V V V A N R T Y E K A K E L A S K V G G Q A I T L A E L E N F H P E D G M V L A N T S V G M F P N V D T P L P K N A L S R Y C L V F D A I Y T P K L T R L L
 PoptrDHQD-SDH5 V V V A N R T Y G K A K E L A S K V G G Q A I A L A K L K D F H P E E G M I L A N T S V G M F P R I E D T P L A K E A L K H Y A L V F D A I Y T P K L T I L L
 FvDHQD-SDH2 V V V A N R S F D K A K T L A D K V G G E A I T L A E L E N F H P E D G M V L A N T S V G M F P K T D Q T P I P K E A L K H Y S L V F D A I Y T P K W T R L L
 Casg1267S13432 V V V A N R T Y D K A K Q L A S K V G G E A I T L A E L E N F H P E D G M I L A N T S V G M F P K I E N T P L P K K A L K H Y S L V F D A I Y T P K M T R L L
 Qrob_P0208730 V V V A N R S Y D K A K Q L A S K V G G E A I S L A E L E N F H P E E G M I L A N T S V G M F P K I E N T P L P K H A L K H Y S L V F D A I Y T P K L T R L L
 FSB015730601 V V V A N R S Y D K A K Q L A S K V G G Q A I T L A E L E N F H P E D G M I L A N T S V G M F P K I E N T P L P K H A L K H Y S L V F D A I Y T P K Q T R L L
 Juni_23843 V V V A N R T Y D K A K Q L A S K V G G E A I T L A E L E N F H P E D G M I L A N T S V G M F P K I E N T P L S K E A L K H Y S L V F D A I Y T P K L T R L L
 ArDHQD-SDH4 V V V A N R T Y D K A K L A S K V G G E A I T L A D L E N F H P E E G M I L A N T S V G M F P K I E D T P L A K K A L K H Y S L V F D A I Y T P K L T R L L
 Alngl16638S34349 V V V A N R T Y D K A K Q L A S K V G G E A I T L A D L E N F H P E E G M I L A N T S V G M F P K I E D T P L A K K A L K H Y S L V F D A I Y T P K L T R L L
 Bpev01_c1356-g0004 V V V A N R T Y D K A K Q L A N K V G G E A I T L A E L E N F H P E D G M I L A N T S T S ----- N A L K H Y S L V F D A I Y T P K L T R L L
 Corav_941 V V V A N R S Y D K A K Q L A S K V G G Q A I T L A E L E N F H P E D G M I L A N T S V G M F P K I E D T P L A K Q A L K H Y S L V F D A I Y T P K L T R L L

AtDHQD-SDH REAEEESGATIVSGSEMFEVROAYEQFE--IFTGLFAP--KELYWQ-----IMSKY-----
 FvDHQD-SDH3 SEAAEESGVRVVGVEMLIGQAYEQFE--RFTGLFAP--KELFRK-----IMDNCL-----
 FvDHQD-SDH4 REAEEESGATVSGSEMFIGQAYEQFE--RFTGLFAP--KELFRK-----VVEST-----
 NtDHQD-SDH1 REAHESGVKIVTGVEMFIGQAYEQYE--RFTGLASS--KGTFOENYGWILRRASLSLFNAALLVTFPPKSLHSCVIAMVLD
 SlDHQD-SDH1 REAQESGAKIVTGVEMFIGQAYEQYE--RFTGLFAP--KELFRK-----IMSTY-----
 CasDHQD-SDH4 REAEEESGAKIVTGVEMFIGQAYEQFE--RFTGLFAP--KELFGK-----IMAKY-----
 VvDHQD-SDH1 REAQESGATIVTGLEMFICQAYEQFE--RFTGLFAP--PKELFKQFIS-----NIQ-----
 FSB010265701 RESEESGATIVTGLEMFICQAYEQFE--RFTGLFAP--KELFRK-----IMAKY-----
 FSB012889301 -----
 Qrob_P0275560 REAEEESGATIVSGSEMFIGQAYEQFE--RFTGLFAP--KQLFRK-----IMANS-----
 Juni_01730 ISLRGSLGCPRRNRLGKLWQIRGA--HLTGNLSR--GFFCGK-----PVQ-----
 JrDHQD-SDH REAEEESGAKIVTGLEMFICQAYEQFE--RFTGLFAP--KELFRK-----VMANN-----
 Juni_19979 REAEEESGAKIVTGLEMFICQAYEQFE--RFTGLFAP--KELFRK-----VMANN-----
 ArDHQ-SDH1 REAEEESGATIVTGLEMFICQAYEQFE--RFTGLFAP--KELFRK-----IMAKY-----
 Alngl188491S37275 REAEEESGATIVTGLEMFICQAYEQFE--RFTGLFAP--KELFRK-----IMAKY-----
 Bpev01_c0420-g0007 REAEEESGATIVTGLEMFICQAYEQFE--RFTGLFAP--KELFRK-----IMANY-----
 Corav_882 REAEEESGATIVTGLEMFICQAYEQFE--RFTGLFAP--KELFRK-----IMAKY-----
 Casg115S00444 REAEEESGATIVSGSEMFIGQAYEQFE--RFTGLFAP--KELFRE-----VMGIH-----

```

EgDHQD-SDH5 KEAECEGATIVSGEMFICQAYCQYESTKRTVQENY-VEVLKR-----LVDLDSIHIIYIFSHL----
EcDHQD-SDH1 KEAECEGATIVSGEMFICQAYCQYESTRYTGLFAP-KELFRK-----IMSKY-----
CsDHQD-SDH3 REAESGATIVSGEMFICQAYCQYESTRFTGLPGK-MNAPHLYKFFVL-----LLYSFNKHFIFTYFLFSFGN
PoptrDHQD-SDH1 REAESGAKIVTIGEMFICQAYCQYESTRFTGLFAP-KELFRK-----IMS-----
PoptrDHQD-SDH3 KDADAAGATIVSGEMFLRQALCQFN--LFTGREAP-KDFMREIVLAKF-----
PoptrDHQD-SDH2 EDADAAGATIVSGEMFLRQALCQFS--LFTGREAP-KDFMREIVLAKF-----
-----
EgDHQD-SDH4 KEAFAAGATIVSGEMFLRQALCQFN--LFTGGEAP-KEFMREIVLAKF-----
EcDHQD-SDH4a KEAFAAGATIVSGEMFLRQALCQFN--LFTGGEAP-KEFMREIVLAKF-----
EcDHQD-SDH4b KEAFAAGATIVSGEMFLRQALCQFN--LFTGGEAP-KEFMREIVLAKF-----
CasDHQD-SDH1 KEAFAAGATIVSGEMFLRQALCQFN--LFTGGEAP-KEFMREIVLAKF-----
SlDHQD-SDH2 EDAAAGALIVSGEMFLRQALCQFN--LFTGSKAP-EEFMRDIVMSKF-----
NtDHQD-SDH2 KDAFAAGATIVSGEMFLRQALCQFH--LFTRLKAP-EEFMRDIVMAKF-----
CsDHQD-SDH1 KDAFAAGATIVSGEMFLRQALCQFN--LFTGGEAP-KEFMREIVLAKF-----
VvDHQD-SDH2 KEAFAAGATIVSGEMFLRQALCQFN--LFTGGEAP-EEFMRDIIILSKF-----
Qrob_P0208710 KDAFAAGATIVSGEMFLRQALCQFK--LFTGGEAP-EEFMRDIIILAKF-----
FSB015730401 KEAFAAGATIVSGEMFLRQALCQFN--LFTGGEAP-EEFMRDIIILAKF-----
Juni_23842 KEAAVAGATIVSGEMFLRQALCQFN--LFTGGEAP-EEFMRDIIILAKF-----
Bpev01_c1356-g0002 KEAFAAGATIVSGEMFLRQALCQFN--LFTGGEAP-EEFMRDIIYAKF-----
Corav_4151 KDAFAAGATIVSGEMFLRQALCQFN--LFTGGEAP-EEFMRDIIYAKF-----
Alng120613S19962 KEAFAAGATIVSGEMFLRQALCQFN--LFTGREGT-ENFLDLLLKNVNIHA-----
ArDHQD-SDH2 KEAFAAGATIVSGEMFLRQALCQFN--LFTSREAP-EEFMRDIIYAKF-----
CasDHQD-SDH2 REARECGATIVYGTMLIRQGFEOYK--NFTGLQAP-EEFMRDIIYAKF-----
VvDHQD-SDH3 REAKESGATIVYGTMLIRQGFEOYK--NFTGLFAP-EEFMRDIIYAKF-----
FvDHQD-SDH1 REAKETGAIVYGTMLIRQGFEOYK--NFTGLFAP-EEFMRDIIYAKF-----
EgDHQD-SDH3 REARETGAIVYGTMLIRQGFEOYK--NFTGLFAP-EEFMRDIIYAKF-----
EcDHQD-SDH3 REARETGAIVYGTMLIRQGFEOYK--NFTGLFAP-EEFMRDIIYAKF-----
Qrob_P0208720 REARETGAIVYGTMLIRQGFEOYK--NFTGLFAP-EEFMRDIIYAKF-----
FSB015730501 REARETGAIVYGTMLIRQGFEOYK--NFTGLFAP-EEFMRDIIYAKF-----
Corav_14055 REARDTGAIVYGTMLIRQGFEOYK--NFTGLFAP-EEFMRDIIYAKF-----
Corav_11016 REARDTGAIVYGTMLIRQGFEOYK--NFTGLFAP-EEFMRDIIYAKF-----
Bpev01_c1356-g0003 REARETGAIVYGTMLIRQGFEOYK--NFTGLFAP-EEFMRDIIYAKF-----
ArDHQD-SDH3 REARETGAIVYGTMLIRQGFEOYK--NFTGLFAP-EEFMRDIIYAKF-----
CasDHQD-SDH3 QEAQDSGAKVVLGTEMFINQAFVQFE--RFTGMFAP-KELIRETLAKNT-----
DkDHQD-SDH REAKETGAKVVLGTEMFLNQAFVQFE--KFTGLFAP-KDLIRETLAKNT-----
VvDHQD-SDH4 REAQESGATIVYGTMLIRQAFVQFE--RFTGMFAP-KELIRETLAKNT-----
EgDHQD-SDH2 REAQEVGATPVYGTMLIRQAFVQFE--RFTGMFAP-KELIRETLAKNT-----
PoptrDHQD-SDH5 REAQEAGATIVYGTMLIRQAFVQFE--RFTGMFAP-KELIRETLAKNT-----
FvDHQD-SDH2 TEAQESGAAVVGTEMFLNQAFVQFE--KFSGLFAP-KELIRETLAKNT-----
Casg1267S13432 TEAESGATIVYGTMLIRQAFVQFE--KFTGLPGK-KEQGNLNPGR-----
Qrob_P0208730 REAESGATIVYGTMLIRQAFVQFE--KFTGLFAP-KELIRETLAKNT-----
FSB015730601 TEAESGATIVYGTMLIRQAFVQFE--KFTGLFAP-KELIRETLAKNT-----
Juni_23843 REAQESGATIVYGTMLIRQAFVQFE--KFTGLFAP-KELIRETLAKNT-----
ArDHQD-SDH4 REAQESGATIVYGTMLIRQAFVQFE--KFSGLFAP-KELIRETLAKNT-----
Alng16638S34349 REAQESGATIVYGTMLIRQAFVQFE--KFSGLFAP-KELIRETLAKNT-----
Bpev01_c1356-g0004 REAQESGATIVYGTMLIRQAFVQFE--KFSGLFAP-KELIRETLAKNT-----
Corav_941 REAQESGATIVYGTMLIRQAFVQFE--KFSGLFAP-KELIRETLAKNT-----
Corav_7742 REAQESGATIVYGTMLIRQAFVQFE--KFSGLFAP-KELIRETLAKNT-----
SlDHQD-SDH3 QEAIEVGATVVSQEMFVRQALCQFK--LFTNGLAP-VDFMRRIVYEQF-----
EgDHQD-SDH1 QEAIEQAGATVVSQEMFVRQALCQFK--LFTSGLAP-EEFMRRLVLEQF-----
PoptrDHQD-SDH4 REAEVGAIVVVSQEMFVRQALCQFR--LFTGGLAP-EEFMRRLVLEQF-----
CsDHQD-SDH2 REAEVGAIVVVSQEMFVRQALCQFR--LFTGGLAP-EEFMRRLVLEQF-----
Qrob_P0339740 QEAIEVGATVVSQEMFVRQALCQFR--LFTGGLAP-EEFMRRLVLEQF-----
Juni_24457 QEAIEIGATVVSQEMFVRQALCQFR--LFTGGLAP-EEFMRRLVLEQF-----
Casg1106S06827 QEAIEVGAIVVVSQEMFVRQALCQFR--LFTGGLAP-EEFMRRLVLEQF-----
Bpev01_c0171-g0015 QEAIEVGATVVSQEMFVRQALCQFR--LFTDGLAP-EEFMRRLVLEQF-----
ArDHQD-SDH5 QEAIEVGATVVSQEMFVRQALCQFR--LFTGGLAP-EEFMRRLVLEQF-----
ArDHQD-SDH6 QEAIEVGATVVSQEMFVRQALCQFR--LFTGGLAP-EEFMRRLVLEQF-----

```

Supplementary Figure S4. Protein sequence alignments of *bona fide* bifunctional 3-dehydroquinate dehydratases/shikimate dehydrogenases (DHQD-SDH) from *Arabidopsis thaliana* (At) (6), *Camellia sinensis* (Cas) (46), *Citrus sinensis* (Cs), *Diospyros kaki* (Dk) (47), *Eucalyptus camaldulensis* (Ec) (48), *E. grandis* (Eg), *Fragaria vesca* subsp. *vesca* (Fv), *Juglans regia* (Jr) (49), *Nicotiana tabacum* (Nt) (50), *Populus trichocarpa* (Poptr) (51)), *Solanum lycopersicum* (Sl) (52), *Vitis vinifera* (Vv) (53) and that of putative homologs (ArDHQD-SDH1 – ArDHQD-SDH6) in red alder. The N-terminal signal peptide of AtDHQD-SDH was not included. The DHQD domain is shown as a red-range bar whereas the SDH

domain is in *light blue*. The consensus motif SX[TG] is highlighted in *blue, green, purple* and *Fuchsia*. The NRT sequence motif typically associated with NADP⁺ preference is in *red*, whereas the DID motif associated with NAD⁺ preference is *green* (54). Names of Group I DHQD-SDHs are depicted in *blue*, Group II in *green*, Group III in *dark red*, Group IV in *purple* and Group V in *Fuchsia*.

AtDHQD-SDH (AAF08579), CasDHQD-SDHa (AYP64306), CasDHQD-SDHb (AYP64307), CasDHQD-SDHc (AYP64308), CasDHQD-SDHd (AYP643079), CsDHQD-SDH1 (orange1.1g010050m), CsDHQD-SDH2 (orange1.1g010101m), CsDHQD-SDH3 (orange1.1g007151m), DkDHQD-SDH (BAI40147), EcDHQD-SDH2 (BBL52471), EcDHQD-SDH3 (BBL52472), EcDHQD-SDH4a (BBL52473), EcDHQD-SDH4b (BBL52474), EgDHQD-SDH1 (Eucgr.H01214.1), EgDHQD-SDH2 (Eucgr.H04428.1), EgDHQD-SDH3 (Eucgr.H04427.1), EgDHQD-SDH4 (Eucgr.B01770.1), EgDHQD-SDH5 (Eucgr.J00263.1), FvDHQD-SDH1 (XP004302480), FvDHQD-SDH2 (XP004302479), FvDHQD-SDH3 (XP004289250), FvDHQD-SDH4 (XP004288087), JrDHQD-SDH (AAW65140), NtDHQD-SDH1 (AS90325), NtDHQD-SDH2 (AAS90324), PoptrDHQD-SDH1 (Potri.010G019000), PoptrDHQD-SDH2 (Potri.013G029900), PoptrDHQD-SDH3 (Potri.005G043400), PoptrDHQD-SDH4 (Potri.014G135500v), PoptrDHQD-SDH5 (Potri.013G029800), SIDHQD-SDH1 (AAC17991), SIDHQD-SDH2 (XP010327280), SIDHQD-SDH3 (XP004242317), VvDHQD-SDH1 (ANC67814), VvDHQD-SDH2 (ANC67815), VvDHQD-SDH3 (ANC67816), VvDHQD-SDH4 (ANC67817). Accession numbers correspond to the NCBI or the Phytozome (v13) databases. Finally, sequences from other Fagales species were retrieved (using BLASTP) from the corresponding databases as indicated in Supplementary Table S1.

S1SK 1 MEARVQSLSQLSSTWNSDKVVRKPSGLIRFSEKWNKPEHRVWVSCHLQPEKAAHSDRRVQL---KVSCSP-----QNVQ
 ArSK 1 MEATCGAALQLSITLITGSKRNV---ALVPLNQR---VREICIMKICDFRRKSVCTRQRRFDYGLLHSHSCNHSHGLVFEVE

S1SK 73 ASVLESQCFSASIDEIETLKNKAEVEVEYLDGRCVYLVGMGCGKTTVGRILAEITLGYSEFDCDRLTEQAVGGITVAEIF
 ArSK 75 IRILESGNSYASFDLWLLKNKAEVETSYLNGRSLFLVGMGSGKTTVGRILSEALGYSEFSDSDRYVELAMGGTVAQIF

S1SK 153 ELRGESEFFRDNETEVLHKLSLMHRLELVSTGGGAVVRPINWVHMKGISVLDVPLALALRITTEGTKSRPLLHEESGDV
 ArSK 155 EQRGEGFFRDHESEALRKLKSLVPRQILVSTGGGAVVRPINWVYMRQGITVLDVPLALALRIAAVGTDSRPLLNEESGDA

S1SK 233 YDTTLKRLITLIMETREGENYANASARVLENLAL-KREKDVCHITFAEITTEVLTIQIENFTKTQKSVVVL---
 ArSK 235 YTKAFVGLFTLTKKRSEAYADADVTISLLDIAANLRLEDVSDITPTVIAIEVLTIQIEKFLGGKNGRSRLRYP

Supplementary Figure S5. Protein sequence alignment of *bona fide* shikimate kinase from *Solanum lycopersicum* (S1SK (7)) and that of a putative homolog (ArSK) in red alder.

AtEPSPS 1 MAQVSRICNGVQNPISLIS-NLKSQSQRKSPISVSLKTOQHPRAYPISSSWGLKSGMTIIGSELRLPKVMSVSTAE---
 ArEPSPS 1 MAQVSRICSGAQSIFHNNPKPHKPKSLNSVNEP-----SQFLVSSNWTLLKKNVCSSRRAGVVRVSASVAIPEKPS

AtEPSPS 77 KASEIVLQPIREISGLIKLPGSKLSNRILLALALSEGTTVVDNLLNSDDNYMLDALRRLGLNVEITSENNRAIVEGCG
 ArEPSPS 77 TVPEIVLQPIREISGAITLPGSKLSNRILLALALSEGTTVVDNLLNSDDNYMLDALRRLGLNVEITSENNRAIVEGCG

AtEPSPS 157 GIFFASIDSSTDIELMLGNAGTAMRPLTAAVTAAGGNASYLDGVPRMRERPIGLVVGGLKQLGADVCEITLGTNCPVVRV
 ArEPSPS 157 GIFFVKGESDEIQLELMLGNAGTAMRPLTAAVTAAGGNSSYLDGVPRMRERPIGLVVGGLKQLGADVCEITLGTNCPVVRV

AtEPSPS 237 NANGGLPGGKVKLSGSISSQYLTAALLMSAPLALGDVEITLIDKLISSVPYVEMTLKLMERFGVSVVEHSDSWDRFVVRGGQK
 ArEPSPS 237 ICKGGLPGGKVKLSGSISSQYLTAALLMSAPLALGDVEITLIDKLISSVPYVEMTLKLMERFGVSVVEHSDSWDRFVVRGGQK

AtEPSPS 317 YKSPGNAYVEGDASSASYFLAGAAITGETVTVEGCGTISLQGDVKFAEVLEKMGCKVSWTENSVTVTGPPRDAFGMRHLR
 ArEPSPS 317 YKSPGNAYVEGDASSASYFLAGAAITGETVTVEGCGTISLQGDVKFAEVLEKMGCKVSWTENSVTVTGPPRDSKRRHLR

AtEPSPS 397 AIDVNMNMKMPDVAMTLAVVALFADGPTTIRDVASWRVK-----ETERMIAICTE---LRKLGATVEEGSDY
 ArEPSPS 397 AIDVNMNMKMPDVAMTLAVVALFADGPTTIRDGSGKLESERDRKDDCHLHRTQEGKFHVAICTSEIHPSPFLGATVEEGPDY

AtEPSPS 460 CVITPEKKIKTAEIDTYDDHRMAMAFSLAACADVPIITINDPGCTRKTFFPDYFOVLERITKH
 ArEPSPS 477 CVITPEENINVTAEIDTYDDHRMAMAFSLAACADVPIITINDESCCTRKTFFPDYFEVLRQRFKH

Supplementary Figure S6. Protein sequence alignment of *bona fide* 5-enolpyruvylshikimate-3-phosphate synthase from *Arabidopsis thaliana* (AtEPSPS (8)) and that of a putative homolog (ArEPSPS) in red alder.

CsCS 1 -MASSLSTKPFISGSRRRSTTDGSGWSYFQTSDLRQISNOQSVQISRRQTAFLKLVQASGSSFGKVQVTTMGESHGGG
 ArCS 1 MASSLSTKPFISGSRRTDGF-----SGLRHLF-SHILPSSVHISIRTRSPKKLQQAAGNIFGNVFRVTTMGESHGGG

CsCS 80 VGCVIDGCPPRFETSEADIQSDLDRRRPGQSRIITPRKETDTCKIYSGVAEGITTGPIHISVPNTDQRGNDYSEMAKAY
 ArCS 73 VGCVIDGCPPRFPLSEADMQFDLDRRRPGQSRIITPRKETDTCKIYSGVAEGITTGPIHISVPNTDQRGNDYSEMAKAY

CsCS 160 RPSHADATYDFKYGVRSVQGGGRSSARETIGRVAAGAAKKILKAYAGTEVLAYVSOAHKVVLPPEGLVDHEIISLQIES
 ArCS 153 RPSHADATYDMKYGTRAVQGGGRSSARETIGRVAAGAAKKILKSLAGIEVLAYVSOVHKVVLPPEGLVDHYSLALQIES

CsCS 240 NIVRCFDEEYAKMIAAIDAVRVKGSVGGVVTCIMRNVPRGLGSPVFDKLEAELAKACMSLPATKGFDFGSGFSTFLT
 ArCS 233 NIVRCFNPEYAKMIAAIDAVRVKGSVGGVVTCIARD CPRGLGSPVFDKLEAELAKAVMSLPATKGFDFGSGFAGTLLT

CsCS 320 GSEHNDFYTDENGRIRTRTNRSGGIQGGISNGEIIINMRIAFAKPTSTIGKKONTVTREREELIARGRHDPVVPRAVP
 ArCS 313 GSEHNDFYTDENGRIRTRTNRSGGIQGGISNGEIIINMRIAFAKPTATIAKKQLTVARDKKEELIARGRHDPVVPRAVP

CsCS 400 MVEAMVALVLDQLMLQHAQGNLFSINPALQEPLETVSSAASLQGV-----
 ArCS 393 MVEAMVALVLDQLMLQHAQGNLFPINPDLOEPLFLPRLDDVAEHEHEHDDLIVG

Supplementary Figure S7. Protein sequence alignment of *bona fide* chorismate synthase from *Capnoides sempervirens* (CsCS (9)) and that of a putative homolog (ArCS) in red alder.

| | | |
|-------|-----|--|
| AtCM1 | 1 | MRSSCCSSIGCFDHRRELSTSTPISTLPLPSYKSFVSRCS----LPQPSKPRSCTS-----SVHAVMTLAGSLTK |
| AtCM2 | 1 | -----MARVFESD |
| ArCM1 | 1 | MEAKLLGPTPAITAH-HASRFSRPISPLVQL--SRQSFSPS-----QASSLAKRGTQ--SVQATASAGSLTIK |
| ArCM2 | 1 | -----MVVRL--LILFLFS--HGIV-----RNRDNSMAEAL |
| ArCM3 | 1 | MEAKLLETGKPTPAL-----TFKSRPILHFASQTTLKFLQSHSGRNGRPLLVSSASSTSSPFRYARK |
| AtCM1 | 72 | KRVDESESLTLEGIIRNSLIRQEDSIIFGLLERAICYNADTYDPAFDMDGFNGSLVEMVKKCTEKLHAKVGRFKSPDEH |
| AtCM2 | 9 | SGSGCSNVLSDLIRESLIRQEDIIIFSLIERAKFPLNSPAFESRCLDSGSFSSLTETEFVRETEIIQAKVGRMEYPIEN |
| ArCM1 | 67 | KRVDESENLTLEGIIRSLIQEDSIIFSLLERAICYNADTYDPAFNSMDGFHGSLSVEMVKEEKLHAKVGRFKSPDEH |
| ArCM2 | 28 | SNAGSANALTLEROSLIRQEDIIIFGLIERARFPIINSFTYINVSYANIPGFSGSLAQETVRETEALQAKAGRETNPIEH |
| ArCM3 | 68 | KRVDESEITLTDGIRHSLIRQEDSIIFSLLERAICYNADTYDHTFMSMDGFHGSLSVEFMVRETEKLHAKVDRKSPDEH |
| AtCM1 | 152 | PPFFDLPEFMLPPLQYPRVLFHAADSININKKIWNMYFRDLVPRLVKRGDDGNYXSTAVCDATCLQCLSKRIHYGKFA |
| AtCM2 | 89 | PPFLENTPHSVFTHKYPSALHPKALSININKQLWDIYFELLELFVKEGDDGNYPSTAASDLACLQALSFRHYGKFA |
| ArCM1 | 147 | AFPFDLPEFVLPPLQYPOVLHPFADSININKKWDMYFRDLVPRLVEEGDDGNCSTAVCDATCLQALSFKRIHYGKFA |
| ArCM2 | 108 | PPFFKDLPEPLVPSQNHTEDLHPSAASININKSLWDIYFDKLLALFAAPGDDGNYEATAASDLVCLQALSFRHYGKFA |
| ArCM3 | 148 | SFFPADLPEFMLPPLQYPOVLHPCADSININKKIWNMYFRDLLPRLVKTGDDGNYGSAVCDTCLQALSFKRIHYGKFA |
| AtCM1 | 232 | EAKFQASPEAYESAIAKQDKDALMDMLTPTVEAIKKRVEMKATVGOEVKVGMEEKEEEEEGNESHVYKISFILVVD |
| AtCM2 | 169 | EVKFRDAPQDYEPAIRAQDREALMKLLTEKVEEMVKKRVQKKAETFGQEVKFNISGYG-----ESKKKYKIDPLLASR |
| ArCM1 | 227 | ETKFAASPEAYEAAIKQDRQRLMDMLTTPKVEEAKKRVEMKATVGOEVTDLGED-----GAEPVYKIVPSLVVD |
| ArCM2 | 188 | EVKFAASPEYEPAIRAKDRDALMKLLTVEVVEEMVKKRVAKKAMVFGQEVTKNKTVD-----TGKYKIDPSVTSR |
| ArCM3 | 228 | EAKFLSSPTSYEAAIRVQDRACLVELLLETVEEAVKKRVETKATIFGOVVRMNOEED-----EAGPAYKIKERLVAN |
| AtCM1 | 312 | LYGHWIMPLTKEQVEYLLRRLD |
| AtCM2 | 243 | IYGEWITPLTKLVQVEYLLRRLD |
| ArCM1 | 300 | LYGHWIMPLTKEQVEYLLRRLY |
| ArCM2 | 260 | LYGEWIMPLTKLVQVEYLLRRLD |
| ArCM3 | 301 | LYGEWIMPLTKQVQVEYLLRRLD |

Supplementary Figure S8. Protein sequence alignment of *bona fide* chorismate mutases from *Arabidopsis thaliana* (AtCM1 and AtCM2 (10)) and that of putative homologs (ArCM1 – ArCM3) in red alder.

| | | |
|----------|-----|---|
| AtPPA-AT | 1 | -----MA-SQSVAVIVSSAAARGE |
| ArPPA-AT | 1 | MGADGSCDSPHPTPTTPPKLFSSHQPALARTLKS PCFQNSVAFLLLVNPSRIQHLPVLFMANSLHSSAIPRLSLRQQ |
| AtPPA-AT | 19 | SFPD---SKKPIGSRVFQQ-PLRLSFSYCK--SGNMSRITCALAKPNIAETLSSSVDSLSPRVQSLKPSKTVITDLAA |
| ArPPA-AT | 81 | SLGLHSTSEFSFRSWSFASHPRTEPLRSLSEGTQVKSPPRWSALV-RA-SRFDAMEVDLSLSPRVNSVKPSKTAITDQAT |
| AtPPA-AT | 93 | TLVQSGVVPVIRLAAGEPDFDTPKVAEAGINAIREGFTRYTLNAGITEIRAICRKLKEENGISYAPDQIVVNSGAKCSI |
| ArPPA-AT | 160 | ALVQAGVVPVIRLAAGEPDFDTPAPVAEAGINAIREGFTRYTPNAGTVEIRAICRKLKEENGISYTPDQIVVNSGAKQCI |
| AtPPA-AT | 173 | LQAVLAVCSPGDEVIIPAPYVWSYTEQARLADATPVVIPTKISNNFLLDPKDLLESKITEKSRLLILCSPSNPTGSVYPKS |
| ArPPA-AT | 240 | LQAVLAVCSPGDEVIIPAPYVWSYFEMARLSDATPVIIPTSIENNFLLDPKVLLESKITEKSRLLILCSPSNPTGSVYPKR |
| AtPPA-AT | 253 | LLEEIARIYAKHPRLVLVLSDEIYEHIIYAPATHTSFASLFGMERTLTVNGFSKAFAMTGWRGLGYLAGPKHIAACSKIQ |
| ArPPA-AT | 320 | LLEEIAQVYAKHPRLVLVLSDEIYEHIIYAPATHTSFASLFGMERTLTVNGFSKAFAMTGWRGLGYLAGPKHFAACGKIQ |
| AtPPA-AT | 333 | GQVSGASSIAQKAGVAALGLGKAGGETVAEMVKAYRERRDFLVKSLGDIKGVVISEPQGFYLFIDFSAYYGSSEAEFGF |
| ArPPA-AT | 400 | SQSSGASSISQKAGVAALGLGYAGGEAWSTMVKARERRDFLVKSFQGLEGVVISEPQGFYLFIDFSSYYGSEAEFGF |
| AtPPA-AT | 413 | LINDSSSIALYFLDKFQVAVVPGDAFGDDSCIRISYATSLDVLQAAVEIRKALEPLRATVSV |
| ArPPA-AT | 480 | KIENSSESLCRYLLDKGQVAVVPGDAFGDDSCIRISYAASLITTLQAAVEIRKALVSLRPAVVP |

Supplementary Figure S9. Protein sequence alignment of *bona fide* prephenate aminotransferase (AtPPA-AT (11)) from *Arabidopsis thaliana* and that of a putative homolog (ArPPA-AT) in red alder.

```

AtADT5 1  -----MQTISPAFSCDLKSVIQPNLTAKKARYSHVNGKRVSVRCSYRSESESF PNGVGS SRADWQSSCAILASKVVS AE
ArADT1 1  -----MQAIAPSAFNSLKS LTG-----TIRPALSRRGPSRLVWQSVYRSDSVNFPNGVGLSRSDWQSSCAILSSKVV SQE
ArADT2 1  MALKGAAIWVCARTSCSDLGASDLGSRNSGSAVSLRFDLEKFRKWECCSLGALAQRAITPVEDEKPHVPGVSSGAVEST
ArADT3 1  ----MAAIVRSPTNPLSRHVS SSKPSPSEHRSKASVTIHVPYFKRRRYFPVVLASLQGDHDNDNNDNRNAQALQLQRILDDSS

AtADT5 74  NSSS-----VAVVNGHSNGSVDL SLMPSKQHNKPG---LIQPLTITDLS PAPS HGSITRVAYQGVPG
ArADT1 70  OPTTEKPGAAGGGGAAGGADHVA AVNGH-KASIDLDLVPISTDSNNKPPSQLAHRALTIITDLS PAMP HGSITRVAYQGVPG
ArADT2 81  QGNES-----RGFHKDLNLLP-----KPLSANDLSSCP SNGSKVRVAYQGLPG
ArADT3 77  QYDV-----VSKDPNMLP-----RELSSQFSDTISHGS RITRVAYQGVGG

AtADT5 135  AYSEAAAGKAYPNSEAI PCDFVAFQAVE LWIADRAVLPVENS IGGSIHRNYDLLRRHLHIVGEVQIPVHHCLLALPG
ArADT1 149  AYSEAAAGKAYPNCEAI PCDFEVAFAQAVE LWIADRAVLPVENS IGGSIHRNYDLLRRHLHIVGEVQLPVHHCLLALPG
ArADT2 124  AYSEAAALKAYPKSEIVPCDFEEAFQAVE LWLVDRAVLPVENS VGGSIHRNYDLLRRHLHIVGEVQMOVVHHCLLGLPG
ArADT3 117  AYSEAAEKAYPNCEAVPCDFDTAFQAVE LWLVDRAVLPVENS IGGSIHRNYDLLRRHLHIVGEVVKFAVHHCLMANHG

AtADT5 215  VRTDCITRVISHPQALAQTEGSLNKLTPKAAIEAFHDTAAA EYIANNLEDTAAVASARA AEIYGLQILADGIQDDAGN
ArADT1 229  VRKEYLTRVISHPQALAQCELT LTKLGLNVAREAVDDTAGAAEFVAANNLEDTAAVASARA ADLYGMNILADGIQDDSSN
ArADT2 204  VRKEDLKS VLSHPMALLQCDMMLSNLGVVRINADDTAGAAQMVASTGLRDTAAVAS SRAAEIYGLDILA EKIQDDDDN
ArADT3 197  VKLEELKRVLSHPQALAQCEHTLTRLG--LVREAVDDTAGAAKHVAE H KLEDTGAVASSTAAMIYGLNILAQDIQDDCDN

AtADT5 295  VTRFLMLARDPIIPRTDRP EKTSIVFAAQEHKGTSVLFKVL SFAFAFRNISLTKIESRPHQNC PVRVVGDENVGTSKHF EY
ArADT1 309  VTRFVMLAREPIIPRTDRP EKTSIVFA--HDKGTSVLFKVL SFAFAFRNISLTKIESRPHRNRPI RLVDDANVGTAKHF EY
ArADT2 282  ITRFLILAREPIIPGTDRP YKTSIVFT--LEEGPGILFKALSVFALRGINLSKIESR PQQRPLRVVDDSN DGS AKYFDY
ArADT3 275  VTRFLILAREPIIPGIDRP EKTSIVFS--LEEGPGVLFKALAVFALRQINLTKIESR PLRKQLLGASDDKNGV PKYFDY

AtADT5 375  IFYVDFEASMAEVARAQNALAEVQEYTSFLRVLGSY PDMTPWSTLPSE DV
ArADT1 387  MFYVDFEASMAEVR AQNALAEVQEFTSFLRVLGSY PDMTPWCPSGGD--
ArADT2 362  LFFIDFEASMAEPRAQALCHLQEF SFLRVLGSYPTDTVL-----
ArADT3 353  IFYVDFEASMAEQNAQNALRHLKEFATFLRVLGSY PVDTSMI-----

```

Supplementary Figure S10. Protein sequence alignment of *bona fide* arogenate dehydratase from *Arabidopsis thaliana* (AtADT5 (12)) and that of putative homologs (ArADT1 – ArADT3) in red alder.

```

UGT84A13 1 MGSEA-LVHVFLVSFPGQGHVNPLLRGKRLASKGLLVTFSTPESIGKMRKASNITD-EPAPVGEFGFIRFEFFEDGWDE
UGT84A22 1 MGSES-LVHVFLVSFPGQGHVNPLLRGKRLASKGLLVTFSTPESIGKMRKASNCTD-EPTPVGDGFIIRFEFFEDGWDE
UGT84A23 1 MGSESSLVHVFLVSFPGQGHVNPLLRGKRLASKGLLVTFSTPESIGKMRKASNISD-QPAPVGDGFIIRFEFFEDGWDE
UGT84A24 1 MGSES-LVHVFLVSFPGQGHVNPLLRGKRLASKGLLVTFSTPESIGKMRKASNIGE-EPSPIGDFIRFEFFEDGWDE
ArUGT1 1 MGSEA-LVHVLLVSFPGQGHVNPLLRGKRLASKGLLVTFSTPESIGKMRKASNITD-EPAPVGEFGFIRFEFFEDGWDE
ArUGT2 1 MGSEA-LVHVLLVSFPGQGHVNPLLRGKRLASKGLLVTFSTPESIGKMRKASNITD-EPAPVGEFGFIRFEFFEDGWDE
ArUGT3 1 MDSGA-PTHVFLVSFPAQGHINPLLRGKRLAAKGLLVTFSTPESIGKMRKANNTDD-QAIPVGEFGFIRFEFFEDGWDE
ArUGT4 1 ---MC-SPHVLLVSFPGQGHINPLLRGKRLAAKGLLVTFSTPESIGKMRKANNTDD-QATPVGGFIRFEFFEDGWDE
ArUGT5 1 MATEA-PTHVLLVSFPAQGNINPLLRGKRLAAKGLLVTFSTPESIGKMRKANNTDD-QATPVGDGFIIRFEFFEDGWDE

UGT84A13 79 DEPRRODLQOYLFQLELIGKDIIPKMIKRNAAEMGRPVSCLINNFFIPWVSDVAESLGLPSAMLWVQSCAFCAYYHYHVG
UGT84A22 79 NEPRRODLQOYLFQLELVGKDLLEPKMTQKHADQDRPVSCLINNFFIPWVSDVAETLGLPSAMLWVQSCAFCFSAYYHYHVG
UGT84A23 80 DEPRRODLQOYLFQLEKVGKVLIPQMIQKNAEQGRPVSCLINNFFIPWVSDVAETLGLPSAMLWVQSCAFLAYYHYHVG
UGT84A24 79 DEPRRODLQOYLFQLEKVGKEVTPRMIKKNNEQNRPVSCLINNFFIPWVSDVAESLGLPSAMLWVQSCAFAAYYHYHVG
ArUGT1 79 NETRRODLQOYLFQLELVGKEVTPMIRRNAAEQGRPVSCLINNFFIPWVSDVAESLGLPSAMLWVQSCAFCSSYYHYHVG
ArUGT2 79 NETRRODLQOYLFQLELVGKEVTPMIRRNAAEQGRPVSCLINNFFIPWVSDVAESLGLPSAMLWVQSCAFCSSYYHYHVG
ArUGT3 80 DDPRRKDLDECYLRFQLELAGKDALHOIKKHAKHDHPVSCLINNFAFISWVQDVAELGLIPNAVLWVQSCAFCFSAYYHYHVN
ArUGT4 77 DDPRRKDLDEFLRHLPEAGKDALHOIKKHAKHDHPVSCLINNFFIPWVQDVAELGLIPNAVLWVQSCAFCFSAYYHYHVN
ArUGT5 79 DDPRRKDLDEYLFQLELAGKDALHOIKKHAKHEDRPVSCLINNFFIPWVQDVAELGLIPNAVLWVQSCAFCFSAYYHYHVN

UGT84A13 159 LVFPFSEAEFFIDVQLPOMPLLYKDEVPFLYPTTPYFLRRAILGQYRNLDPFCILMDTFQELEHEVIEYMSKICPIK
UGT84A22 159 IVFPFSEDNMEIDVQLPOMPLLYKDEVPFLYPTTPYFLRRAILGQYRNLDPFCILMDTFQELEHEVIEYMSKISPIK
UGT84A23 160 LVFPFSENAMEIDVQLPOMPLLYKDEVPFLYPTTPYFLRRAILGQYRNLDPFCILMDTFQELEHEIIEYTSKICPIK
UGT84A24 159 LVFPFSEAMEIDVQLPOMPLLYKDEVPFLYPTTPYFLRRAILGQYRNLDPFCILMDTFQELEHEIIEYMSKICPIK
ArUGT1 159 LVFPFSEAEFFIDVQLPOMPLLYKDEVPFLYPTTPYFLRRAILGQYRNLDPFCILMDTFQELEHVDVIEYASKLCPIK
ArUGT2 159 LVFPFSEAEFFIDVQLPOMPLLYKDEVPFLYPTTPYFLRRAILGQYRNLDPFCILMDTFQELEHVDVIEYASKLCPIK
ArUGT3 160 LVFPFSEQPMIDVQLPOMPLLYKDEVPFLHPSPPFVFLGRVILGQYRNLDPFCILMDTFQELEHEIIDYMAKLCVIK
ArUGT4 157 LVFPFSETEPKLDVQLPOMPLLYKDEVPFLHPSPPFVFLGRVILGQYRNLDPFCILMDTFQELEHEIIDYMAKLCMIK
ArUGT5 159 LASFPSETEPKLDVQLPOMPLLYKDEVPFLHPSPPFVFLGRVILGQYRNLDPFCILMDTFQELEPEIINYSKLCVMIK

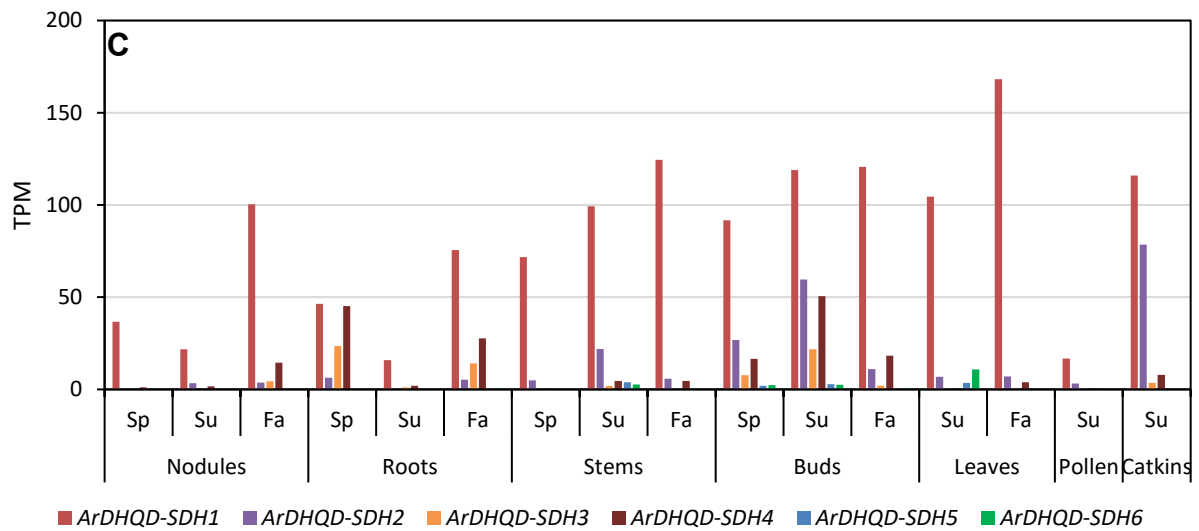
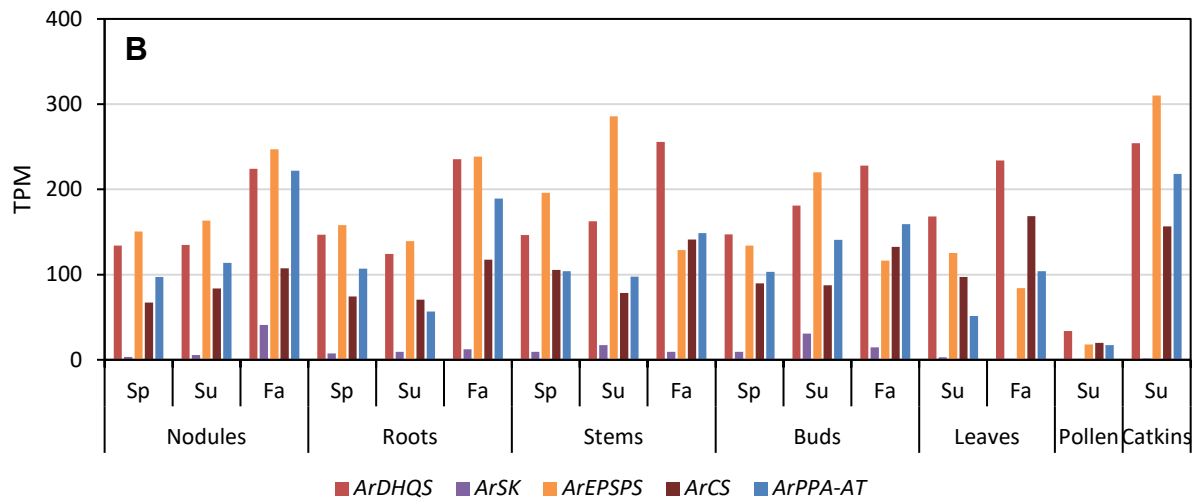
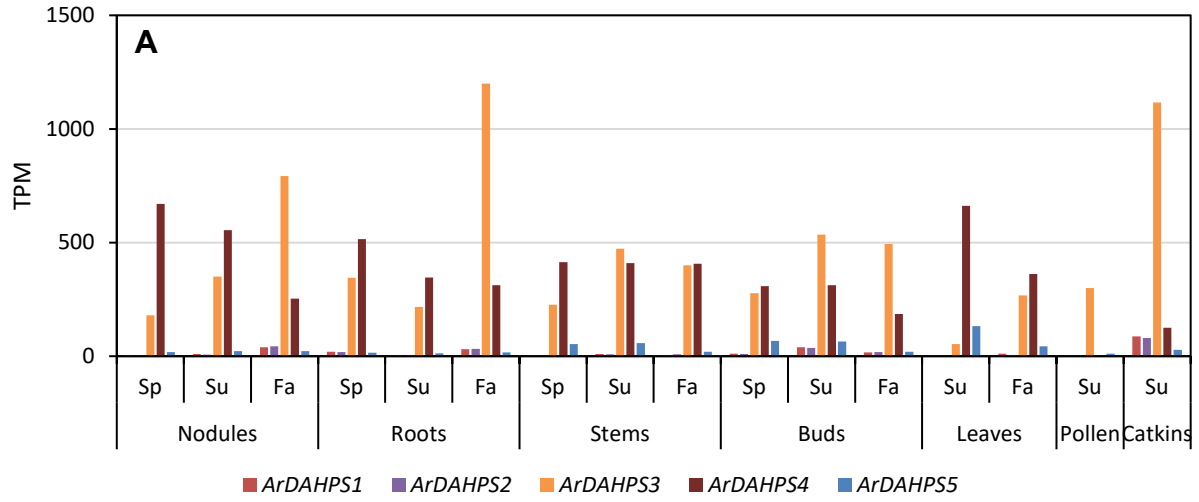
UGT84A13 239 TVGPLFKNPKAPNSVVRGDFMKADDCLEWLDKSKPPQSVVYISFGSVVYLTQKQVDEIAGLLSGVSVFLVWMKPPHKDAG
UGT84A22 239 PVGPLFKNPKAPNSVVRGDFMKADDCMEWLDKSKPPGSSVYISFGSVVYLTQEHVDEIAGILSGSVSVFLVWMKPPHKDAG
UGT84A23 240 TVGPLFKNPKAPNNTVKGDFMKADDCIGWLDKSKPPASSVYVYISFGSVVYLTQDQVDEIAGLLSGVSVFLVWMKPPHKDSG
UGT84A24 239 TVGPLFKNPKAPNANVRGDFMKADDCISWLDKSKPPASVYVYISFGSVVYLTQDQVDEIAGLLSGLNFLVWMKPPHKDSG
ArUGT1 239 TVGPLFKNPKAQNNTVVRGDFMKADDCMEWLDKSKPPQSVVYISFGSVVYLTQPVDEIAGLLSGVSVFLVWMKPPHKDSG
ArUGT2 239 TVGPLFKNPHRTQPCAATS-----
ArUGT3 240 PVGPLFINPKAPNNTSVRGDILKADDCIEWLDKSKPPASVYISFGSVVYLTQEQVDEIAGLLSGVSVFLVWIKPPGKDTD
ArUGT4 237 PVGPLFINPKAPNNTSIRGDVLKADDCIEWLDKSKPPASVYISFGTVAHVQEQVDEIAGLLSGVSVFLVWMKPPHKDID
ArUGT5 239 PVGPLFINPKAPTTSIRGDFLKADDCIEWLDKSKPPASVYISFGSSANIGQAQADELAGLFGVSVFLVWMKPPQKDAF

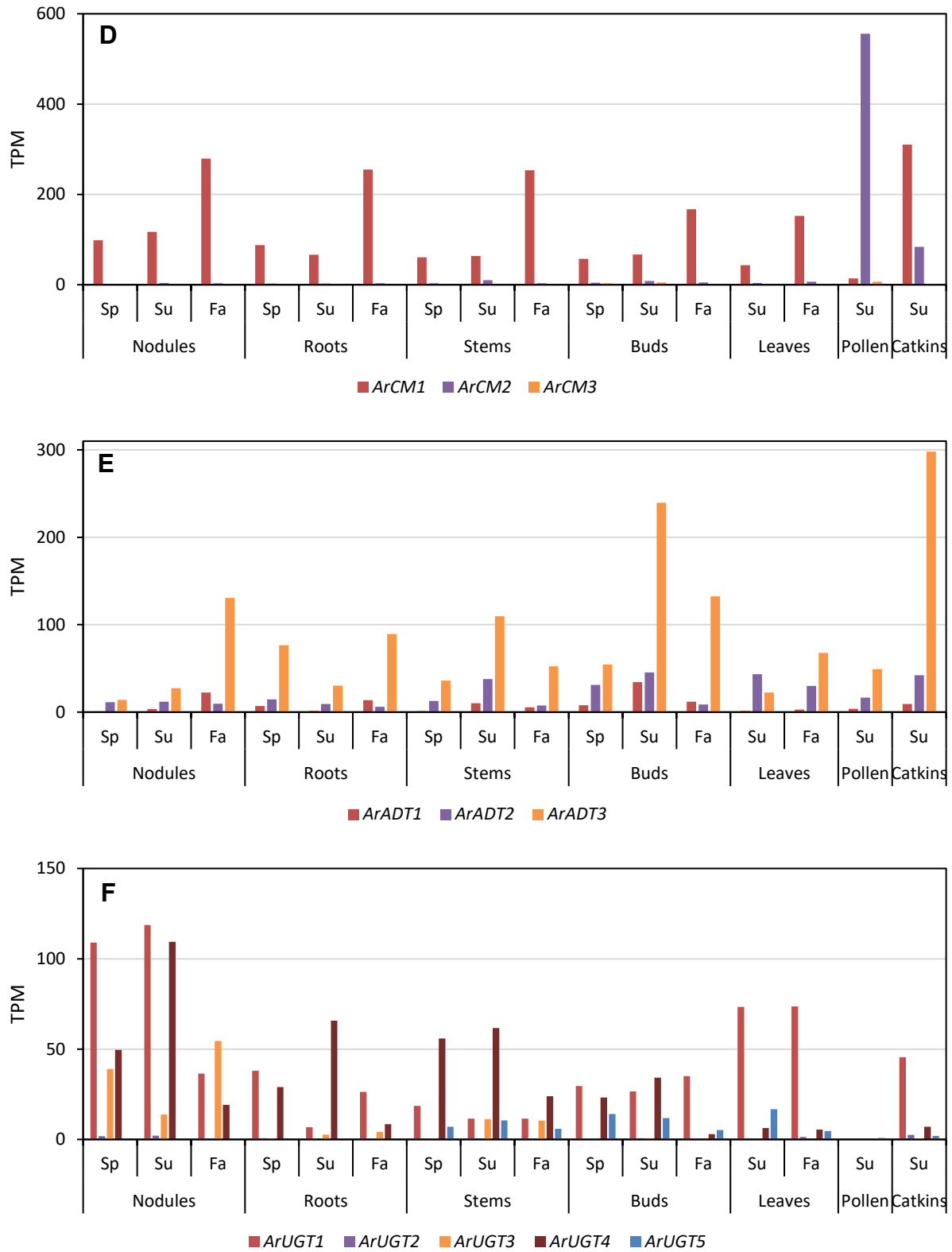
UGT84A13 319 LELLVLPDGFLEKAGDNGRVVQWSPQEOVLAHPSVACFVTHCGWNSMTEALTSGLMPVVAFFPOWGDQVTDVAVYLVDFVFKTG
UGT84A22 319 LELLVLPDGFLEKAGDKGRVVQWSPQEOVLVHVAACFVTHCGWNSMTEALSLASGLMPVVAFFPOWGDQVTDVAVYLVDFVFKTG
UGT84A23 320 YTVLTLPEGFLEKAGDRGKVVQWSPQEOVLAHAPATACFVTHCGWNSMTEALTSGLMPVVAFFPOWGDQVTDVAVYLVDFVFKTG
UGT84A24 319 YQLTLTLPEGFLEKAGDKGRVVQWSPQEOVLAHPSVACFVTHCGWNSMTEALSSGLMPVVAFFPOWGDQVTDVAVYLVDFVFKTG
ArUGT1 319 YELLVLPDGFLEKAGDKGRVVHWSQPEOVLAHPSVACFVTHCGWNSMTEALTSGLMPVVAFFPOWGDQVTDVAVYLVDFVFKTG
ArUGT2 -----
ArUGT3 320 FKLHVLPDGFLEKVGNRGKLIQWSPQEOVLAHPSVACFVTHCGWNSMTEALTLGVPVVAFFPOWGDQVTDVAVYLVDFVFKTG
ArUGT4 317 FKLHVLPDGFLEKVGDRGKVIQWSPQEOVLAHPSVACFVTHCGWNSMTEALTLGVPVVAFFPOWGDQVTDVAVYLVDFVFKTG
ArUGT5 319 FKLHVLPDGFMEKVGNGKVIQWSPQEOVLAHPSVACFVTHCGWNSMTEALTLGVPVVAFFPHWGDQVTDVAVYLVDFVFKTG

UGT84A13 399 VRMORGEAENRVIIRDEVEKCLLEATVGPKAVEKQNAKWKAAAEAAAFSEGGSSDRNIQAFVDEVVRRRSVAITGKSTAN
UGT84A22 399 IRMORGEAEDRVIIRDEVEKCLREATVGPKAAEMKANAKWKETAEEAAVAEGGSSDRNMQAFVDEVVRRRSIGITSKSTTP
UGT84A23 400 VRMORGEAEDKLIIRDVVEKCLREATVGPKAAEMKANAKWKAAAEAAAFSEGGSSDRNIQAFVDEVVRRRSIETITASKPAV
UGT84A24 399 VRMORGEAENKLIIRDVVEKCLLEATVGPKAAEMKANAQWKAAAEAAVAEGGSSDRNIQAFVDEVVRRRSIAIQSNKSEF
ArUGT1 399 LRMORGEAENRVIIRDETEKCLVEATVGPKAVEKQNAKWKAAAEAAVAEGGSSDRNIQAFVDEVVRRRSIALTSKSAAS
ArUGT2 -----
ArUGT3 400 VRLSRGLAENRVIIRDEVEKCLLEATVGPKAVEKQHAQWKWTAAGEAAVAEGGSSDONIQSFVDEIRKRCVAATSA----
ArUGT4 397 VRLGRGOAENRVIIRDEVEKCLLEATVGPKAVEKQHAQWKWTAAGEAAVAEGGSSDONIQSFVDEIRKRCVAVTSTSTITQ
ArUGT5 399 LRLSRGLAENRVIIRDEVEKCLLDAMVGPKAVEKHNAAKWKRAAKEVVAEGGSSDONVKAIVDEIRKRCVAATSKSTTP

```

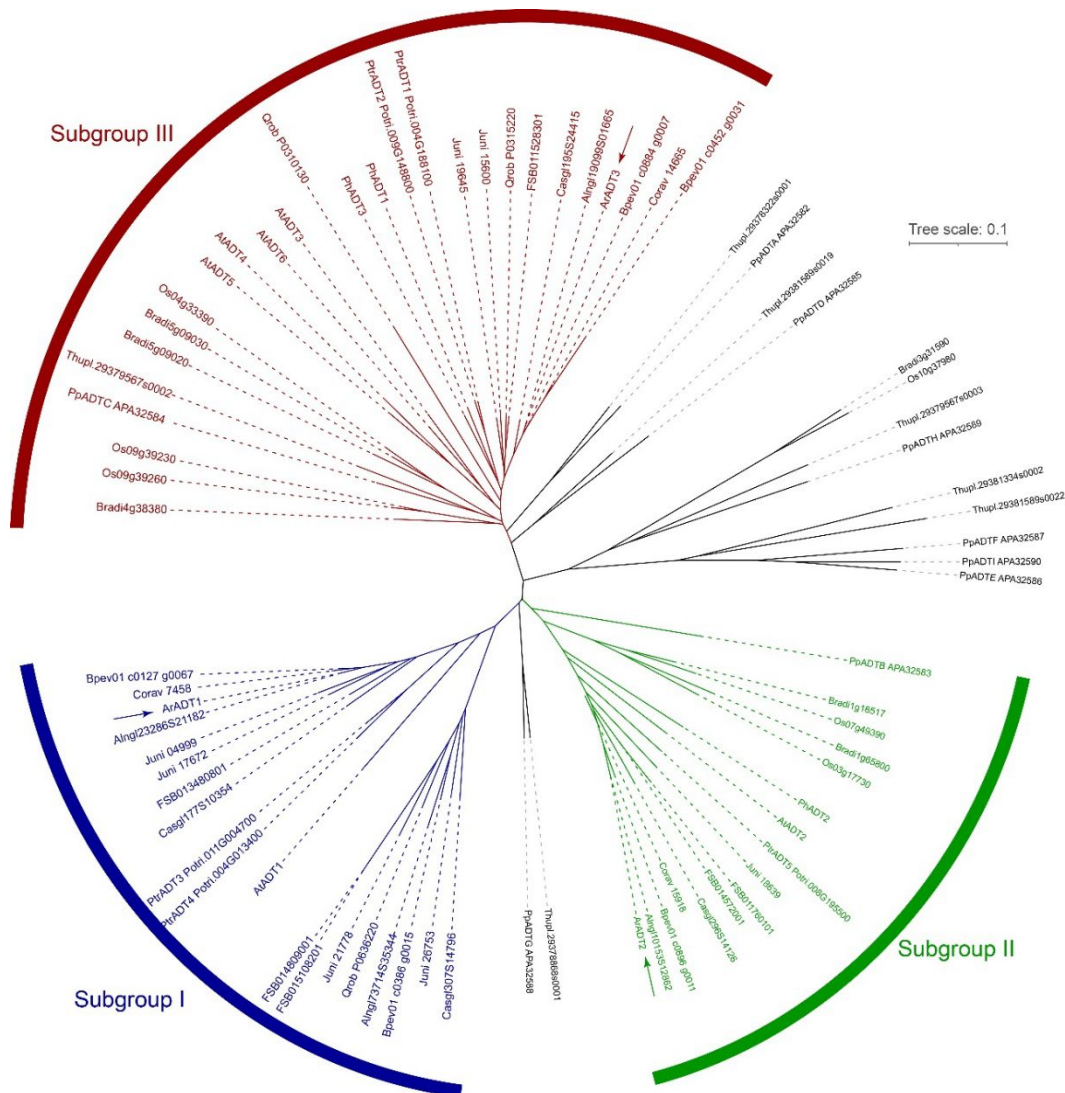
Supplementary Figure S11. Protein sequence alignment of *bona fide* β -glucogallin forming UDP glucosyltransferases from oak [*Quercus robur*, UGT84A13 (13)], tea [*Camellia sinensis*, UGT84A22 (55)], pomegranate [*Punica granatum*, UGT84A23 and UGT84A24 (56)] and that of putative homologs (ArUGT1 to ArUGT5) in red alder.





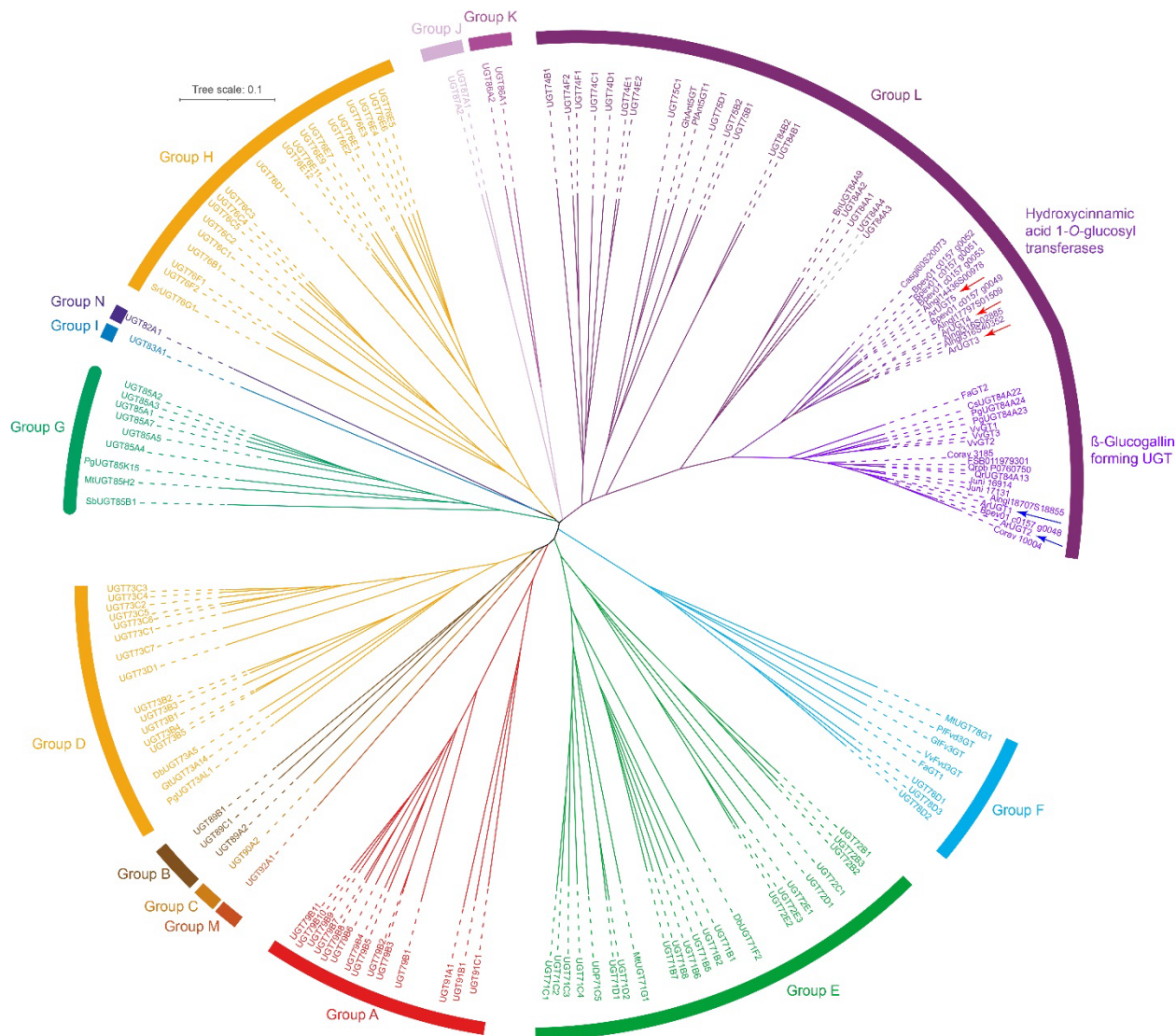
Supplementary Figure S12. Expression levels of red alder homologs of *bona fide* shikimate, chorismate and phenylalanine pathway genes, as well as β -glucogallin forming genes. (A) 3-

Deoxy-D-arabinoheptulosonate 7-phosphate synthase (*ArDAHPS*). **(B)** Dehydroquinate synthase (*ArDHQS*), shikimate kinase (*ArSK*), 5-enolpyruvylshikimate-3-phosphate synthase (*ArEPSPS*), chorismate synthase (*ArCS*), and prephenate aminotransferase (*ArPPA-AT*). **(C)** Bifunctional 3-dehydroquinate dehydratase/shikimate dehydrogenase (*ArDHQD-SDH*). **(D)** Chorismate mutase (*ArCM*). **(E)** Arogenate dehydratase (*ArADT*). **(F)** β -glucogallin forming UDP glucosyltransferase (*ArUGT*). Abbreviations: TPM, transcripts per million; Sp, spring; Su, summer; Fa, fall.



Supplementary Figure S13. Unrooted phylogenetic tree of arogenate dehydratases (ADTs) from *Alnus glutinosa* (Alngl), *Alnus rubra* (Ar), *Arabidopsis thaliana* (At), *Betula pendula* (Bpev), *Brachypodium distachyon* (Bradi), *Casuarina glauca* (Casgl), *Corylus avellana* (Corav), *Fagus sylvatica* (FSB), *Juglans nigra* (Juni), *Oryza sativa* (Os), *Petunia hybrida* (Ph), *Pinus pinaster* (Pp), *Populus trichocarpa* (Ptr), *Quercus robur* (Qrob), and *Thuja plicata* (Thupl). A multiple-sequence alignment was built using Clustal Omega (57), and the unrooted phylogenetic tree was rendered with iTOL (58).

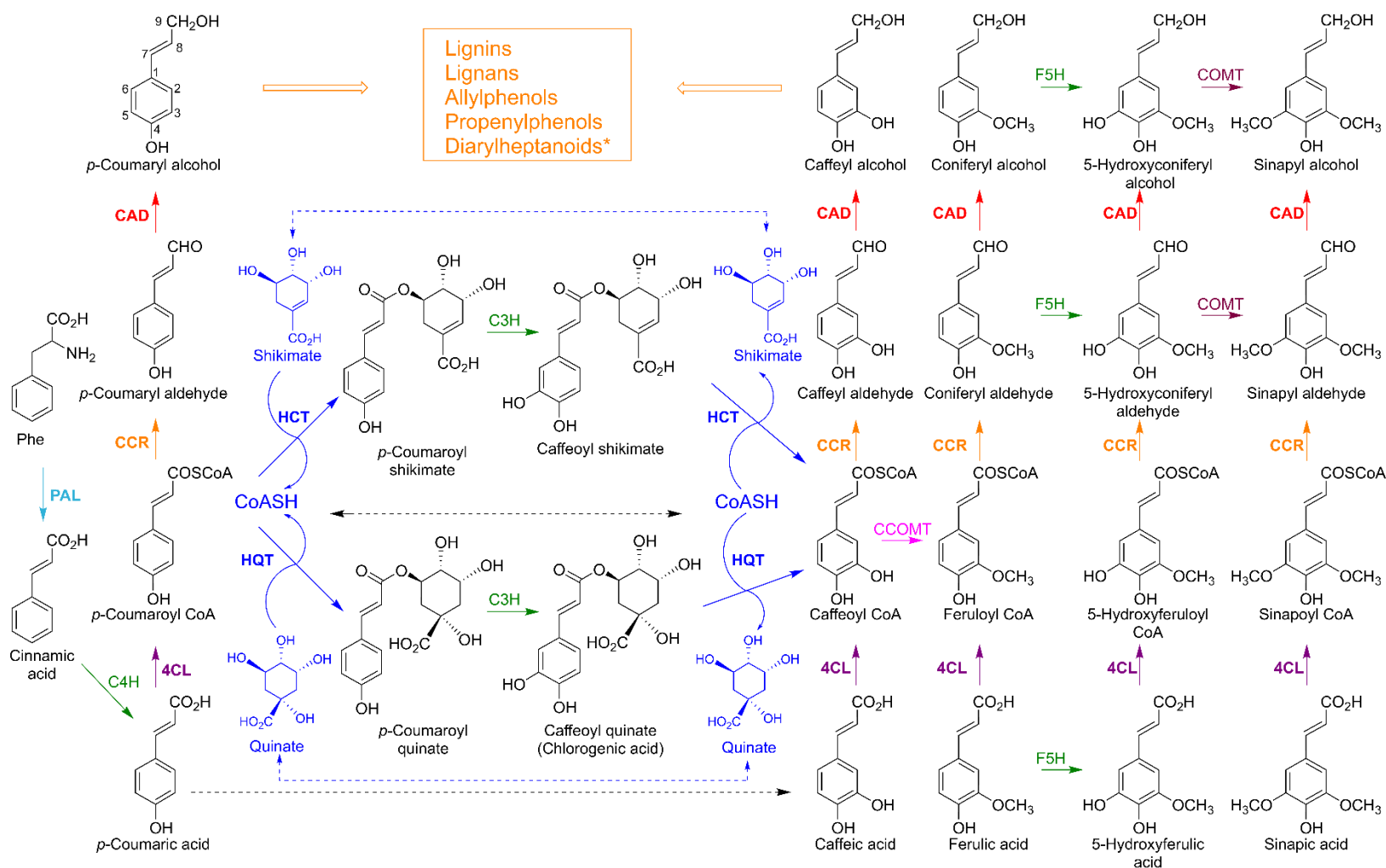
AtADT1 (At1g11790), AtADT2 (At3g07630), AtADT3 (At2g27820), AtADT4 (At3g44720), AtADT5 (At5g22630), AtADT6 (At1g08250), PhADT1 (ACY79502), PhADT2 (ACY79503), PhADT3 (ACY79504), PpADTA (APA32582), PpADTB (APA32583), PpADTC (APA32584), PpADTD (APA32585), PpADTE (APA32586), PpADTF (APA32587), PpADTG (APA32588), PpADTH (APA32589), PpADTI (APA32590), PtrADT1 (Potri.004G188100), PtrADT2 (Potri.009G148800), PtrADT3 (Potri.011G004700), PtrADT4 (Potri.004G013400), PtrADT5 (Potri.008G195500). Sequences were obtained from NCBI and from the Phytozome (v13) databases, whereas sequences from other Fagales species were retrieved (using BLASTP) from the corresponding databases as indicated in Supplementary Table S1.



Supplementary Figure S14. Unrooted phylogenetic tree of UDP-glucosyltransferases (UGTs). Analysis used the *Arabidopsis* superfamily UGTs (59) and other UGTs of established functions. A multiple-sequence alignment was built using Clustal Omega (57), and the unrooted phylogenetic tree was rendered with iTOL (58). Red alder UGTs, ArUGT1 and 2, cluster with β -glucogallin forming UGTs in subgroup L and are indicated with blue arrows. ArUGT3–ArUGT5 are shown with red arrows.

Arabidopsis thaliana: UGT71B1 (AT3G23330), UGT71B2 (AT3G21760), UGT71B5 (AT4G15280), UGT71B6 (AT3G21780), UGT71B7 (AT3G21790), UGT71B8 (AT3G23390), UGT71C1 (AT2G29750), UGT71C2 (AT2G25780), UGT71C3 (AT1G07260), UGT71C4 (AT1G06880), UGT71C5 (AT1G07240), UGT71D1 (AT2G25790), UGT71D2 (AT2G29710), UGT72B1 (AT4G01070), UGT72B2 (AT1G01390), UGT72B3 (AT1G01420), UGT72C1 (AT4G36770), UGT72D1 (AT2G18570), UGT72E1 (AT3G50740), UGT72E2 (AT5G66690), UGT72E3 (AT5G26310), UGT73B1 (AT4G39040), UGT73B2 (AT4G39030), UGT73B3 (AT4G34131), UGT73B4 (AT2G15490), UGT73B5 (AT2G10670), UGT73C1 (AT2G33480), UGT73C2 (AT2G36760), UGT73C3 (AT2G36780), UGT73C4 (AT2G36770), UGT73C5 (AT2G36800), UGT73C6 (AT2G36790), UGT73C7 (AT3G53160), UGT73D1 (AT3G53150),

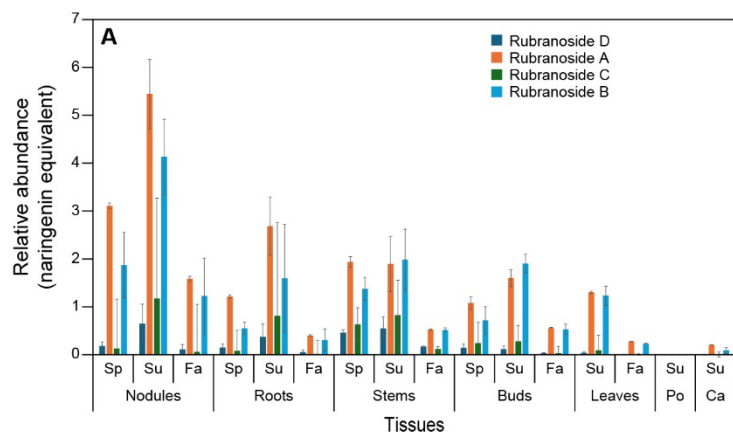
UGT74B1 (AT1G25270), UGT74C1 (AT2G31790), UGT74D1 (AT2G31750), UGT74E1 (AT1G05675), UGT74E2 (AT1G05160), UGT74F1 (AT2G43840), UGT74F2 (AT2G43820), UGT75B1 (AT1G05560), UGT75B2 (AT1G05530), UGT75C1 (AT4G14090), UGT75D1 (AT4G15550), UGT76B1 (AT3G11340), UGT76C1 (AT5G05870), UGT76C2 (AT5G05860), UGT76C3 (AT5G05900), UGT76C4 (AT5G05880), UGT76C5 (AT5G05890), UGT76D1 (AT2G26480), UGT76E1 (AT5G59580), UGT76E2 (AT5G59590), UGT76E3 (AT3G46700), UGT76E4 (AT3G46690), UGT76E5 (AT3G46720), UGT76E6 (AT3G46680), UGT76E7 (AT5G38040), UGT76E9 (AT5G38010), UGT76E11 (AT3G46670), UGT76E12 (AT3G46660), UGT76F1 (AT3G55700), UGT76F2 (AT3G55710), UGT78D1 (AT1G30530), UGT78D2 (AT5G16690), UGT78D3 (AT5G17030), UGT79B1 (AT5G53060), UGT79B2 (AT4G27560), UGT79B3 (AT4G27570), UGT79B4 (AT3G29630), UGT79B5 (AT1G50580), UGT79B6 (AT5G54010), UGT79B7 (AT4G09500), UGT79B8 (AT2G22930), UGT79B9 (AT5G53990), UGT79B10 (AT1G64910), UGT79B11 (AT1G64920), UGT82A1 (AT3G22250), UGT83A1 (AT3G02100), UGT84A1 (AT4G15480), UGT84A2 (AT3G21560), UGT84A3 (AT4G15490), UGT84A4 (AT4G15500), UGT84B1 (AT2G23260), UGT84B2 (AT2G18900), UGT85A1 (AT1G22400), UGT85A2 (AT1G22360), UGT85A3 (AT1G22380), UGT85A4 (AT1G78270), UGT85A5 (AT1G22370), UGT85A7 (AT1G22340), UGT86A1 (AT2G36970), UGT86A2 (AT2G28080), UGT87A1 (AT2G30150), UGT87A2 (AT2G26230), UGT89A2 (AT5G03490), UGT89B1 (AT1G73880), UGT89C1 (AT1G06000), UGT90A2 (AT1G10400), UGT91A1 (AT2G22590), UGT91B1 (AT5G65550), UGT91C1 (AT5G49690) and UGT92A1 (AT5G12890). *Brassica napus*: BnUGT84A9 (AAF98390). *Camellia sinensis*: CsUGT84A22 (ALO19890). *Cleretum bellidiforme*: DbUGT71F2 (AAL57240) and DbUGT73A5 (CAB56231). *Fragaria × ananassa*: FaGT1 (AAU09442) and FaGT2 (AAU09443). *Gentiana triflora*: GtUGT73A14 (BAC54092) and GtFv3GT (Q96493). *Glandularia hybrida*: GhAnt5GT (BAA36423). *Medicago truncatula*: MtUGT71G1 (AAW56092), MtUGT78G1 (ABI94025) and MtUGT85H2 (ABI94024). *Perilla frutescens*: PfAnt5GT1 (BAA36421) and PfFvd3GT (BAA19659). *Punica granatum*: PgUGT84A24 (ANN02877), PgUGT85K15 (ANN02874), PgUGT73AL1 (ANN02876) and PgUGT84A23 (ANN02875). *Quercus robur*: QrUGT84A13 (XP_050262807). *Sorghum bicolor*: SbUGT85B1 (AAF17077). *Stevia rebaudiana*, SrUGT76G1 (AAR06912). *Vitis vinifera*, VvGT1 (AEW31187), VvGT2 (AEW31188), VvGT3 (AEW31189) and VvFvd3GT (BAB41026). Accession numbers correspond to the NCBI database. Finally, sequences from other Fagales species were retrieved (using BLASTP) from the corresponding databases as indicated in Supplementary Table S1.



Supplementary Figure S15. Simplified angiosperm phenylpropanoid pathway to diarylheptanoids, lignins, lignans, *etc.*

4CL: 4-Coumarate CoA ligase; C4H: Cinnamate 4-hydroxylase; CAD: Cinnamyl alcohol dehydrogenase; CCOMT: Caffeoyl CoA *O*-methyltransferase; CCR: Cinnamoyl CoA reductase; COMT: Caffeic acid *O*-methyltransferase; F5H: Ferulate 5-hydroxylase; HCT: Hydroxycinnamoyl CoA:shikimate hydroxycinnamoyl transferase; HQT: Hydroxycinnamoyl CoA: quinate hydroxycinnamoyl transferase; PAL: Phenylalanine ammonia lyase; C3H: *p*-Coumarate 3-hydroxylase. NOTE: PAL, C4H, and 4CL enzymes are also common to the first three biochemical steps leading to the flavonoid metabolic class.

*Diarylheptanoids are thought to be derived from CoA esters (e.g. *p*-coumaroyl or caffeoyl CoAs).



Supplementary Figure S16. Relative abundance of diarylheptanoids in naringenin equivalents.

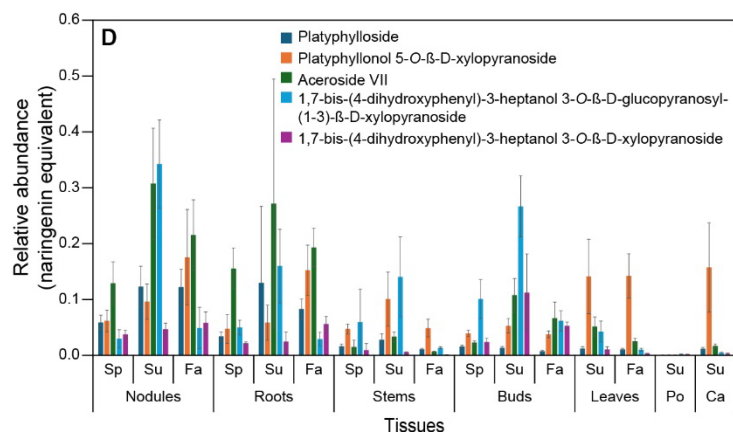
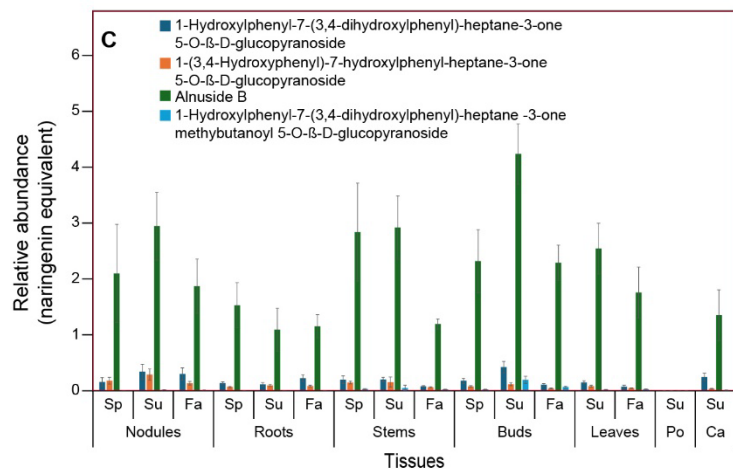
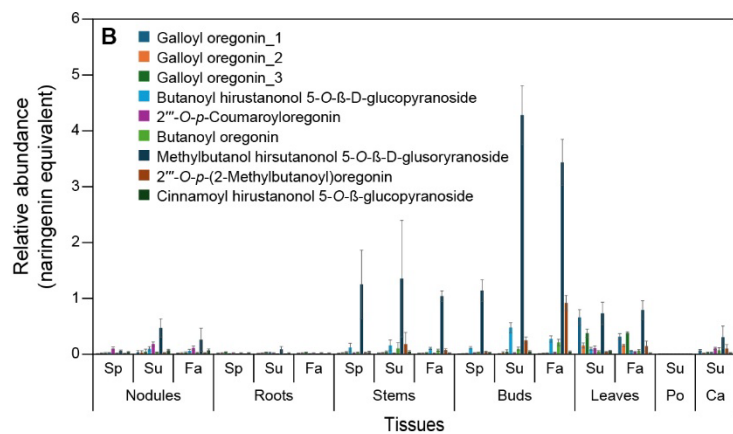
(A) Rubranosides A to D.

(B) Hirsutanonol derived metabolites.

(C) Alnuside B and related diarylheptanoids.

(D) Platyphyloside and related diarylheptanoids.

Abbreviations: Ca, catkins; Fa, fall; Po, pollen; Sp, spring; Su, summer.



```

AtPAL1 1 MEINGAHKSNNGGVDAMLCGGDIKTKNMVINAEDPLNWGAAAEQKGSHLDEVKRMVAEFRKPVVNLGGETLTIISQVAAI
ArPAL1 1 -----METTPKNGETLQVSHDP--LNWGVAAESLKGSHLDEVKRMVAEVRKPVVRLGGETLTIISQVAAI
ArPAL3 1 -----METTPNNGETLQVSHDPSLNWGVAAESLKGSHLDEVKRMVAEVRKPVVRLGGETLTIISQVAAI
ArPAL2 1 -----MAACNNGTA-TNFCGTGDPNLWGAAAEALNGSHLDEVKRMVAEFRKPVVRLGGETLTIISQVAAI
ArPAL4 1 -----MKICNNV-----SESGDPLNWGAAAEALNGSHLDEVKRMVAEFRKPVVRLGGETLTIISQVAAI

AtPAL1 81 STIGNSVKVELSEARAGVNASSDWVMSMNKGTDSYGVTGFGATSHRRTKQGAALQKELIRFLNAGIFGSTRKETSHTL
ArPAL1 63 ATCDSSGVKVELSEARAGVKASSDWVMSMASGTDSYGVTGFGATSHRRTKQGAALQKELIRFLNAGIFGNGTESCHTL
ArPAL3 65 ATCDSSGVKVELSEARAGVKASSDWVMSVSGTDIYGTGFGAASHRRTKQGAALQKELIRFLNAGIFGNGTESCHTL
ArPAL2 64 ATHDAGVKVELSEARAPGVKASSDWVMSMNKGTDSYGVTGFGATSHRRTKQGAALQKELIRFLNSGIFGNGTESCHTL
ArPAL4 59 ATHDAGVKVELSEARAPGVKASSDWVMSMNKGTDSYGVTGFGANSHRRTKQGAALQKELIRFLNSGIFGNGTESCHTL

AtPAL1 161 PHSATRAAMLVRINTLLQGSIRFEIMEAITSFLNHNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNKATGPNGEA
ArPAL1 143 PHSATRAAMLVRINTLLQGSIRFEIMEAITKFLNHNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNKATGPNGES
ArPAL3 145 PYSATRAAMLVRINTLLQGSIRFEIMEAITKFLNHNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNKATGPNGES
ArPAL2 144 PQATRAAMLVRINTLLQGSIRFEIMEAITKFLNHNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNKATGPNGES
ArPAL4 139 PQATRAAMLVRINTLLQGSIRFEIMEAITKFLNHNITPCLPLRGTITASGDLVPLSYIAGLLTGRPNKATGPNGES

AtPAL1 241 LTADEAFRLAGIDSSGFFELQPKEGLALVNGTAVGSGMASMVLFEINLSVLAEILSAFAEVMQKPEFTDHLTHLKLHH
ArPAL1 223 LDAVEAFRLAGIDSSGFFELQPKEGLALVNGTAVGSGMASMVLFEANILSVLSEILSAFAEVMQKPEFTDHLTHLKLHH
ArPAL3 225 LDAVEAFRLAGIDSSGFFELQPKEGLALVNGTAVGSGMASMVLFEANILSVLSEILSAFAEVMQKPEFTDHLTHLKLHH
ArPAL2 224 LDAVKAFFELAGINGGFFELQPKEGLALVNGTAVGSGMASMVLFEANIQAVALSEITSAFAEVMQKPEFTDHLTHLKLHH
ArPAL4 219 LDAVKAFFELAGINGGFFELQPKEGLALVNGTAVGSGMASMVLFEANIQAVALSEITSAFAEVMQKPEFTDHLTHLKLHH

AtPAL1 321 PGQIEAAAIMEHILDGSSYKLAQKLHEMDPLQPKQDRYALRTSPQWLGQIEVIRYATKSIEREINSVNDNPLIDVSR
ArPAL1 303 PGQIEAAAIMEHILDGSAVYKAAKLAQKLHEMDPLQPKQDRYALRTSPQWLGQIEVIRYSTKSIEREINSVNDNPLIDVSR
ArPAL3 305 PGQIEAAAIMEHILDGSAVYKAAKLAQKLHEMDPLQPKQDRYALRTSPQWLGQIEVIRYSTKSIEREINSVNDNPLIDVSR
ArPAL2 304 PGQIEAAAIMEHILDGSAVYKAAQKLHEMDPLQPKQDRYALRTSPQWLGQIEVIRYSTKMIEREINSVNDNPLIDVSR
ArPAL4 299 PGQIEAAAIMEHILDGSAVYKAAQKLHEMDPLQPKQDRYALRTSPQWLGQIEVIRYSTKMIEREINSVNGNPLIDVSR

AtPAL1 401 NKAHGGNFQGTPIGVSMNDNRLAIAAIGKLMFAQFSELVNDFYNNGLPSNLASRNPSLDYGFGKAEIAMASYCSELQF
ArPAL1 383 NKALHGGNFQGTPIGVSMNDNRLAIAASIGKLMFAQFSELVNDFYNNGLPSNLSASRNPSLDYGFGKAEIAMASYCSELQF
ArPAL3 385 NKALHGGNFQGTPIGVSMNDNRLAIAASIGKLMFAQFSELVNDFYNNGLPSNLSASRNPSLDYGFGKAEIAMASYCSELQF
ArPAL2 384 NKALHGGNFQGTPIGVSMNDNRLAIAASIGKLMFAQFSELVNDFYNNGLPSNLSGRNPSLDYGFGKAEIAMASYCSELQF
ArPAL4 379 NKALVGGNFQGTPIGVSMNDNRLAIAASIGKLMFAQFSELVNDFYNNGLPSNLSASRNPSLDYGFGKAEIAMASYCSELQF

AtPAL1 481 LANPVTSHVQSAAEQHNQDVNSLGLISSRKTAEAVDILKLMSTFLVALCOAIDLRHLEENLKSTVKNTVSOVAKRVLTTG
ArPAL1 463 MANPVTSHVQSAAEQHNQDVNSLGLISSRKTAEAVDILKLMSTFLVALCOAIDLRHLEENLKSTVKNTVSOVAKRVLTTG
ArPAL3 465 MANPVTSHVQSAAEQHNQDVNSLGLISSRKTAEAVDILKLMSTFLVALCOAIDLRHLEENLKSTVKNTVSOVAKRVLTTG
ArPAL2 464 LANPVTSHVQSAAEQHNQDVNSLGLISSRKTAEAVDILKLMSTFLVALCOAIDLRHLEENLKSTVKNTVSOVAKRVLTTG
ArPAL4 459 LANPVTSHVQSAAEQHNQDVNSLGLISSRKTAEAVDILKLMSTFLVALCOAIDLRHLEENLKSTVKNTVSOVAKRVLTTG

AtPAL1 561 VNGELHPSRFCEKDLLKVVVDREHVFAYIDDPSCSATYPLMQKLRQVLVDHALNNGEKEKERNLSTSIFOKIGFEELKVLPL
ArPAL1 543 VNGELHPSRFCEKDLLKVVVDREHVFAYIDDPSCSATYPLMQKLRQVLVDHALNNGEKEKERNLSTSIFOKIGFEELKVLPL
ArPAL3 545 VNGELHPSRFCEKDLLKVVVDREHVFAYIDDPSCSATYPLMQKLRQVLVDHALNNGEKEKERNLSTSIFOKIGFEELKVLPL
ArPAL2 544 SNGELHPSRFCEKDLLKVVVDREHVFAYIDDPSCSATYPLMQKLRQVLVDHALNNGEKEKERNLSTSIFOKIGFEELKVLPL
ArPAL4 539 SNGEHPHPSRFCEKDLLKVVVDREHVFAYIDDPSCSATYPLMQKLRQVLVDHALNNGEKEKERNLSTSIFOKIGFEELKVLPL

AtPAL1 641 KEVESARVAHESGNPAVILNRIKECRSYPLYRFVREELGTGLLTGKVTSPGEEFDKVFAMCAGKIIDPMLDCLRDWNGA
ArPAL1 623 KEVESARVAHESGNPAVILNRIKECRSYPLYRFVREELGTGLLTGKDRSPGEEFDKVFAMCAGKIIDPMLDCLRDWNGA
ArPAL3 625 KEVESARVAHESGNPAVILNRIKECRSYPLYRFVREELGTGLLTGKDRSPGEEFDKVFAMCAGKIIDPMLDCLRDWNGA
ArPAL2 624 KEVESARTELENEKAAIPNLIKNCRSYPLYRFVREELGTGLLTGKARSPEGEEFDKVFAMCAGKIIDPMLDCLRDWNGA
ArPAL4 619 KEVESARTELENGKAAIPNLIKNCRSYPLYRFVREELGTGLLTGKDRSPDEEFDKVFAMCAGKIIDPMLDCLRDWNGA

AtPAL1 721 PLPIC
ArPAL1 703 PLPIC
ArPAL3 705 PLPIC
ArPAL2 704 PLPIN
ArPAL4 699 PLPIN

```

Supplementary Figure S17. Protein sequence alignment of *bona fide* phenylalanine ammonia lyase from *Arabidopsis thaliana* (AtPAL1 (17)) and that of putative homologs (ArPAL1 – ArPAL4) in red alder.

```

HtC4H      1  -----MDL L L I E K T L V A L F A A I I G A I L I -----SKLRGKFKFLPPGPIPVPIFGNWLQVGD DLNH
ArC4H1     1  -----MDL L L L E K T L L A L F L A V I V A I T I -----SKLRGKRFKFLPPGPIPVPIFGNWLQVGD DLNH
ArC4H2     1  MAHLVATPIFYTLITVFLISATNFFFSFTLLS AALPLIPLIAYYFGS TSGRRGSGALPPGPLSFPIFGNWLQVGN DLNH
ArC4H3     1  MAHLVATPIFYTLITVFLISATNFFFSFTLLS AALPLIPLIAYYFGS TSGRRGSGALPPGPLSEPIFGNWLQVGN DLNH

HtC4H      56  RNLFDLAKRFGEIILLR MGQRNLV VSSPELAKV LHTQGV EFGSRTRNVVFDIFTGKGQDMVFTVYGEHWRKMRRIMTV
ArC4H1     56  RNLFDLAKRFGDI FLLR MGQRNLV VSSPELAKV LHTQGV EFGSRTRNVVFDIFTGKGQDMVFTVYGEHWRKMRRIMTV
ArC4H2     81  GVLATMAQKYGPVLLKLGSKNLA VSSPELASQV LHTQGV EFGSRPRNVVFDIFTGNGQDMVFTVYGEHWRKMRRIMTL
ArC4H3     81  GVLATMAQKYGPVLLKLGSKNLA VSSPELASQV LHTQGV EFGSRPRNVVFDIFTGNGQDMVFTVYGEHWRKMRRIMTL

HtC4H      136  PFFTnkVVQCYRHW EAEAAAVVDVKKNEEAAATGTVIRRR LQIMMYNMMFRIMFDRRFSEDDPLFLK LKAINGERSR
ArC4H1     136  PFFTnkVVQCYRHW EAEAAASVVDVKKNEEAAATGTVIRRR LQIMMYNMMYRIMFDRRFSEDDPLFMKLKAINGERSR
ArC4H2     161  PFFTnkVVHSYSGM WEEEMDLVVHDLKRDERVVRTKGLVIRKRLQIMLYNIMYRIMFDAKFESQEDPLFIEATKFN AERSR
ArC4H3     161  PFFTnkVVHSYSGM WEEEMDLVVHDLKRDERVVRTKGLVIRKRLQIMLYNIMYRIMFDAKFESQEDPLFIEATKFN AERSR

HtC4H      216  LAQSF EYNYGDFIPLLRPFLR NYLKLCKEVKDR LQLFKDYFVDERKKIGSTKMDNNQLKCAIDHILBAKEKGEINEON
ArC4H1     216  LAQSF DYNYGDFIPLLRPFLRS YLKI CKEVKERLQLFKDYFVDERKKLASTKPTDNEGLKCAIDHILDAQKGEINEON
ArC4H2     241  LAQSF EYNYGDFIPLLRPFLR GYLKCRDLQSRRLAFNNFYVEKRRKIMAANG-EKHKINCAIDHII DAEMKGEISEAN
ArC4H3     241  LAQSF EYNYGDFIPLLRPFLR GYLKCRDLQSRRLAFNNFYVEKRRKIMAANG-EKHEINCAIDHII DAEMKGEISEAN

HtC4H      296  VLYIVENINVA A IETTLWSI EWGTAELVNHPETQAKLRHELDTRKLGPGVQITEPDVQNL PYLQAVKETLRLRMAIPLLV
ArC4H1     296  VLYIVENINVA A IETTLWSI EWGLAELVNHPKIQKKVRDEIDTVLGLGVQVTEPDIQKLPYLQAVKETLRLRMAIPLLV
ArC4H2     320  VLYIVENINVA A IETTLWSI EWATAELVNHP TVQKIRDEISAVL-KGQPVTESNLHEL PYLQAVKETLRLRHSIPLLV
ArC4H3     320  VLYIVENINVA A IETTLWSI EWATAELVNHP TVQKIRDEISAVL-KGQPVTESNLHEL PYLQAVKETLRLRHSIPLLV

HtC4H      376  PHMNLHDAKLG GFDIPAESKILVN AWLANNPDQWKKPEEFRPERFLEEEAKVEANGN---DFRYLPFGWGRRSCP GIIIL
ArC4H1     376  PHMNLHDAKLG GFDIPAESKILVN AWLANNPANWKNPEEFRPERFLEEE SKVEANGN---DFRYLPFGWGRRSCP GIIIL
ArC4H2     399  PHMNL EEAKLGGYTI PKESKVVVNA WLSNNPEEAWKNPEEFRPERFLEEE GSTDAVAGKVD FRYLPFGWGRRSCP GIIIL
ArC4H3     399  PHMNL EEAKLGGYTI PKESKVVVNA WLSNNPEEAWKNPEEFRPERFLEEE GSTDAVAGKVD FRYLPFGWGRRSCP GIIIL

HtC4H      453  ALPILGITIGRLV ONFELL EPPGQSKIDTAEKGGQFSLHILKHS TIVAKPRSF
ArC4H1     453  ALPILGITIGRLV ONFELL EPPGQSKIDTAEKGGQFSLHILKHS NIVAKPRSF
ArC4H2     479  ALPILGLVIAKLV ONFEMNAPVGM EKIDVSGKGGQFSLHIASHSTVAFDPITA
ArC4H3     479  ALPILGLVIAKLV ONFEMNAPVGM EKIDVSGKGGQFSLHIASHSTVAFDPITA

```

Supplementary Figure S18. Protein sequence alignment of *bona fide* cinnamate 4-hydroxylase from *Helianthus tuberosus* (HtC4H (18)) and that of putative homologs (ArC4H1 – ArC4H3) in red alder.

```

NtHCT      1  -----MKTEVKESTMVPAAETPQORLWNSNLDLVVPRFHFTPSVYFYRPTGSPNFFDCKVLKEALSALVVPFY
ArHCT1    1  -----MIINVKESTMVREPGETPQORALWNSNLDLVVPRFHFTPSVYFYRPTGAPNFFDANVLKEALSALVHFY
ArHCT2    1  -----MIINVKESTMVREPGETPKRGLWNSNLDLVVTRFHFTPSVYFYRPTGAPNFFDASVLKEALSALVVPFY
ArHCT3    1  MQVGKSEKRRAKMIISVKESTMVRPGETPKHRLWNSNLDLVVTRFHFTPSVYFYRPTGAPNFFDASVLKEALSALVVPFY

NtHCT     69  PMAGRLCRDEEGRIEIDCKQGVLFVEAETGVDVDDFGDFAPTTELRLQIIPAVDYSQGIQSYALVLIQITLHFKCGGVSIG
ArHCT1    69  PMAARLKRDEEGRIEIDCNAEGVLFVEAETTSVDDFGDFAPTTELRLQIIPAVDYSKGIQSYPLVLIQVTFYFKCGGVSIG
ArHCT2    69  PMAGRLKLDDEEGRIEIDCNAEGVLFVEAETGSLVDDFGDFAPTTELRLQIIPAVDYSKGIQSYPLVLIQVTFYFKCGGVSIG
ArHCT3    81  PMAGRLKLDDEEGRIEIDCNAEGVLFVEAETGSLVDDFGDFAPTTELRLQIIPAVDYSKGIQSYPLVLIQVTFYFKCGGVSIG

NtHCT    149  VGMQHHAADGASGLHFINTWSDMARGLDLTIPPFID-RTLRLRARDPPQPAFQHVVEYQPPETLKVTPENTFISEVPETSV
ArHCT1    149  VGMQHHAADGASGLHFVNTWSDMARGLDLTIPPFIDRTLRLRARDPPQPAFQHVVEYQPPETLKVTPENTFISEVPETSV
ArHCT2    149  IGTQHLLADGASGLHFVNTWSDMARGLDLTIPPFIN-RTLRLRARDPPQPAFQHVVEYQPPETLKVTPENTFISEVPETSV
ArHCT3    161  IGTQHLLADGASGLHFVNTWSDMARGLDLTIPPFIN-RTLRLRARDPPQPAFQHVVEYQPPETLKVTPENTFISEVPETSV

NtHCT    228  SMFKITRDQNLTLKAKSKDGNITVNYTSYEMLSGHVWRSTCMARGLAHDQETKLYTATDGRSRLAPPLPPGYFGNVIFIT
ArHCT1    227  SMFKITRDQNLTLKAKSKDGNITVNYTSYEMLSGHVWRSTCMARGLAHDQETKLYTATDGRSRLAPPLPPGYFGNVIFIT
ArHCT2    222  SMFKITRDQNLTLKAKSKDGNITVNYTSYEMLS-----ETKLYAVDGRSRLAPPLPPGYFGNVIFIT
ArHCT3    234  SMFKITRDQNLTLKAKSKDGNITVNYTSYEMLSGHLWKCVSARALPDDQETKLYAVDGRSRLAPPLPPGYFGNVIFIT

NtHCT    308  TPIALAGDLQSKPTWYAASKLHDAALARMNDYLRSAIDYLELQPDLSALVIRGAHTYRCPNLGITSWIRLPIHDADFGWGR
ArHCT1    307  TPIALAGDLQSKPTWYAASRIHNAALARMNDYLRSAIDYLELQPDLSALVIRGAHTYRCPNLGITSWIRLPIHDADFGWGR
ArHCT2    285  TPIALAGDLQSKPTSYAASMIHNAALARMNDYLRSAIDYLELQPDLSALVIRGAHTYRCPNLGITSWIRLPIHDADFGWGR
ArHCT3    314  TPIALAGDLQSKPTSYAASMIHNAALARMNDYLRSAIDYLELQPDLSALVIRGAHTYRCPNLGITSWIRLPIHDADFGWGR

NtHCT    388  PIFMGGGIPYEGLSIFILPSPTNDGSLVAIALQSEHMKVFEKFLYDF
ArHCT1    387  PIFMGGGIPYEGLSIFILPSPTNDGSLVAIALQSEHMKVFEKFLYDF
ArHCT2    365  PIFMGGGIPYEGLSIFILPSPTNDGSLVAIALQSEHMKVFEKFLYDF
ArHCT3    394  PIFMGGGIPYEGLSIFILPSPTNDGSLVAIALQSEHMKVFEKFLYDF

```

Supplementary Figure S19. Protein sequence alignment of *bona fide* hydroxycinnamoyl CoA:shikimate hydroxycinnamoyl transferase from *Nicotiana tabacum* (NtHCT (19)) and that of putative homologs (ArHCT1 – ArHCT3) in red alder.

```

AtC3H      1  MSWRLTAVATIAAVVSYKLIQRLRQKFPFPPGSPKPIVGNLYDIPVRFRCFAEWAQSYGPTISVWFSGSLNVIVVSSAELA
ArC3H1    1  MSLPLIPISLIVLLLAYKLYQWLRFLKPPGPRPWPIVGNLYDIKPVRFRCFADWAQAYGPTISVWFGSLNVIVVSNSELA
ArC3H2    1  MALPLIPVSLIILLLLAYKLYQRLRFLKPPGPRPWPIVGNLYDVKPVRFRCAEWAQAYGPTISVWFGSLNVIVVSNSELA
ArC3H3    1  MALPLIPVSLIILLLLAYKLYQRLRFLKPPGPRPWPIVGNLYVNVKPVRFRCFAEWAQAYGPTISVWFGSLNVIVVSNSELA
ArC3H4    1  MALPLIPVSLIILLLLAYKLYQRLRFLKPPGPRPWPIVGNLYVNVKPVRFRCFAEWAQAYGPTISVWFGSLNVIVVSNSELA

AtC3H     81  KEVLKEHDQMLADRFRNRSTEFASRNQGLDIWADYGPHYVKVRKVCTLELFTPKRLESIRPIREDEVITAMVESVFRDCNL
ArC3H1    81  KEVLKEHDQMLADRFRSRSAAKFSRDGKDLIADYGPHYVKVRKVCTLELFSPKRLEALRPIREDEVITAMVESIFKDCIN
ArC3H2    81  KEVLKEHDQMLADRFRNRSSARFSRDGKDLIADYGPHYVKVRKVCTLELFSPKRLEALRPIREDEVITAMVESIFNDCAH
ArC3H3    81  KEVLKEHDQMLADRFRNRSSATMFSRDGKDLIADYGPHYVKVRKVCTLELFSAKRLEALRPIREDEVITAMVKSIFNDCIH
ArC3H4    81  KEVLKEHDQMLADRFRNRSSATMFSRDGKDLIADYGPHYVKVRKVCTLELFSAKRLEALRPIREDEVITAMVKSIFNDCIH

AtC3H    161  PENRKGKGLQLRKYLGAVAFNNITRLAFGKRFMNAEGVYDEQGLEFKAIVSNGLKGLASLSIAADHIPWRWMPFADEKAF
ArC3H1    161  PENNGKSLLVKKYLGAVAFNNITRLAFGKRFVNSEGVMDQGLEFKAIVANGKLGASLAMAADHIPWRWMPFDEEEAF
ArC3H2    161  PYNKGKRLLVKKYLGAVAFNNITRLTFGKRFMNSEGVMDQGLEFKAIVANGKGGSLPIADHIPWRWMPFRHEEAF
ArC3H3    161  PDNKGKRLLVKKYLGAVAFNNITRLTFGKRFMNSEGVMDQGLEFKAIVANGKLGGLPIADHIPWRWMPFQDEEAF
ArC3H4    161  PDNKGKRLLVKKYLGAVAFNNITRLTFGKRFMNSEGVMDQGLEFKAIVANGKLGGLPIADHIPWRWMPFQDEEAF

AtC3H    240  AEHGARRDRLTRIMEEHTTARQKSSGAKQHFVDALLTLKDYDLS EDTIIGLLWDMITAGMDTTAITAEWAMAEIKNP
ArC3H1    240  AKHGARRDRLTRIMEEHTQAREKSGGAKQHFVDALLTLKDEYDLS EDTIIGLLWDMITAGMDTTAISVEWAMAEIKNP
ArC3H2    240  AKHGERRDRLTKTMEEHTQARNKSGSAGKQHFVDALFTLQDKYDLSDDTVIGLLWDMITAGTDTTAISVEWAMAEIKNP
ArC3H3    241  EKHWEIRDRLTKTMEEHTQARNKSGEAKQHFADALFTLQEKYDLCDDTVIGLLWDMMGAGTDTTAISVEWAMAEIKNP
ArC3H4    241  EKHWEIRDRLTKTMEEHTQARNKSGEAKQHFADALFTLQEKYDLCDDTVIGLLWDMMGAGTDTTAISVEWAMAEIKNP

AtC3H    320  RVQHKVQEEEDRVVGLDRLITEADFSRLPYLQCVVKESERLHPPTPLMLPHRSNADVKIGGYDIPKGSNVHVNVWAVARD
ArC3H1    320  RVQHKVQEEEDRVVGLERVI TEADFTLPLPYLQCVVKESERLHPPTPLMLPHRANANVKIGGYDIPKGSNVHVNVWAVARD
ArC3H2    320  RVQHKVQEEEDSVIGFKRFMTEADFSRLPYLQCVVKESERLHPPTPLMLPHRANANVKIGGYDIPKGSNVHVNVWAVARD
ArC3H3    321  RVQHKVQEEEDRVVIGFERFMTEADFSRLPYLQCVVKESERLHPPTPLMLPHRANANVKIGGYDIPKGSNVHVNVWAVARD
ArC3H4    321  RVQHKVQEEEDRVVIGFERFMTEADFSRLPYLQCVVKESERLHPPTPLMLPHRANANVKIGGYDIPKGSNVHVNVWAVARD

AtC3H    400  PAVWKNPLEFRPERFLEEDVDMKGFDFRLLPFGAGRRMCPGAQLGINLVTSMIGHLLHHFVWVWPPPEGVKLEEEIDMSENPG
ArC3H1    400  PAVWKNPMQFRPERFLEEDVDMKGFDFRLLPFGAGRRMCPGAQLGINLVTSMIGHLLHHFVWVWPPPEGVKLEEEIDMSENPG
ArC3H2    400  PAVWKNPLEFRPERFLEEDVDMKGFDFRLLPFGAGRRMCPGAQLA INLVTSMIGHLLHHFVWVWPPPEGVKLEEEIDMSENPG
ArC3H3    401  PAVWKNPLEFRPERFLEEDVDMKGFDFRLLPFGAGRRMCPGVQLA INLVTSMIGHLLHHFVWVWPPPEGVKLEEEIDMSENPG
ArC3H4    401  PAVWKNPLEFRPERFLEEDVDMKGFDFRLLPFGAGRRMCPGVQLA INLVTSMIGHLLHHFVWVWPPPEGVKLEEEIDMSENPG

AtC3H    480  LVTYMRTPVQAVATPRLPSILYKRVVYDMD
ArC3H1    480  LVTYMRTPVQAVANPRLPSHLYKRVVAVDMD
ArC3H2    480  LVTYMRTPVQAVANPRLPSHLYKRVVAVDMD
ArC3H3    481  LVTYMRTPVQAVANPRLPSHLYKRVVAVDMD
ArC3H4    481  LVTYMRTPVQAVANPRLPSHLYKRVVAVDMD

```

Supplementary Figure S20. Protein sequence alignment of *bona fide* *p*-coumarate 3-hydroxylase from *Arabidopsis thaliana* (AtC3H (20)) and that of putative homologs (ArC3H1 – ArC3H4) in red alder.


```

At4CL1 1 -----MAPQEQAVSQVMEKQSNNNSVIFRSKLPDIYIPNHLSLHDIYIFQNISEFATKPCLINGPTGHVITYSDVH
Ar4CL1 1 -----MATCAEQQEFIFRSKLPDIYIPKHLPHSYCFENISKFGSRPCLINGSTGEVHTYYDVE
Ar4CL2 1 -----MATCAEQQEFIFRSKLPDIYIPKHLPHSYCFENISKFGSRPCLINGSTGEVHTYYDVE
Ar4CL3 1 -----MEPCKDLQEFIFRSKLPDIYIPNHLPLHYCFENISQFKDRPCLINGANGVTITYSEVE
Ar4CL4 1 MISVASNSTEPQNFSPKIPP-TPPSQACVFSKLPFIPISNHLPLHYCFEHLAEFDRPCLIVGSTGKTFSAETH

At4CL1 73 VISRQLAANFHKLGVNQNDVVMLLLPNCPEFVLSFLAASRGGATAAANPFFTAEIAKQAKASNTKLIITEARVVDKLR
Ar4CL1 60 LTARKVASGLSKLGIKGDVIMLLLPNSPEFAFVFLGASYLGAMTTAANPFFTAEVSKQKASNAKLVTQSQYVDKVK
Ar4CL2 60 LTARKVASGLSKLGIKGDVIMLLLPNSPEFAFVFLGASYLGAMTTAANPFFTAEVSKQKASNAKLVTQSQYVDKVK
Ar4CL3 60 LTARKVASGLDKVGIKQGVIMLLLPNCPEFAFVFLGASYLGAVSTTANPFFTAEVAKQAKASNAKLITQALYVDKVK
Ar4CL4 80 LVSRKLAAGLSNLGIKKGDVVMVLLQNCAEFIFSLGASMLGAMATTANPFFTAEVFKQLNSRAKLIITQSQNVDKLR

At4CL1 153 PLQNDGVE-----VIICIDNESVPIPGCLRFELTQSTTEASEVDSVEISPDVVVALPYSSGTTGLPKGVMLTHKG
Ar4CL1 140 DYAENGVE-----KIICIDDSP----PKDLHFSELTOADE---NDIPEVDISPDVVVALPYSSGTTGLPKGVMLTHKG
Ar4CL2 140 DYAENGVE-----KIICIDDSP----PKDLHFSELTOADE---NDIPEVDISPDVVVALPYSSGTTGLPKGVMLTHKG
Ar4CL3 140 DFAENGVE-----KVMCIDS-Q----LDLFLHFSQLTOADE---NDIPAVKINPDDVVVALPYSSGTTGLPKGVMLTHCG
Ar4CL4 160 EPEGENFPKLGEDFTVITVDD-----PPKCLHFSEVISEANE---DEFITVSLPEPDDVVALPSSGTTGLPKGVMLTHKS

At4CL1 227 LVTSAQQVDGENPNLYFHSQDVLICVLPPLFHIYALNSVFLCGLRVGAAILIMPKFEINLLELILQRCVIVAVNVPPIV
Ar4CL1 207 LVTSAQQVDGENPNLYFHSQDVLICVLPPLFHIYSLNSVFLCGLRAGAAILIVQKFEISLLELILQYKVSVMPIVVPPIV
Ar4CL2 207 LVTSAQQVDGENPNLYFHSQDVLICVLPPLFHIYSLNSVFLCGLRAGAAILIVQKFEISLLELILQYKVSVMPIVVPPIV
Ar4CL3 206 LVTSAQQVDGENPNLYFHSQDVLICVLPPLFHIYSLNSVFLCGLRVGAAILIMPKFEIVKLELQCYKVIAPFVPPIV
Ar4CL4 232 LTTSAQQVDGENPNLYLNEEDVILCVLPPLFHIYSLNSVFLCGLRAGAAILIMQKFEIALLELIRHIVSVAIVVPPIV

At4CL1 307 LAIAKSSTEKYDLSSIRIVKSGAAPLKGLEDAVNAKFFNAKLGQGYGMTEAGPVLAMSLCFAKEPFEVKSGACGTVVR
Ar4CL1 287 LSIKRFPHKYDLSSIKMLKSGAPLKGLEDIVKAKFFKALFGQGYGMTEAGPVLAMCLAFAKEPMEVKSGACGTVVR
Ar4CL2 287 LSIKRFPHKYDLSSIKMLKSGAPLKGLEDIVKAKFFKALFGQGYGMTEAGPVLAMCLAFAKEPMEVKSGACGTVVR
Ar4CL3 286 LAIAKSPDAQWYDSSIRVMVMSGAAPLKGLEDVRAKLFNAKLGQGYGMTEAGSVLTMCLAFAKEPFEVKSGACGTVVR
Ar4CL4 312 LAIAKNEVIANFDLSSIRIVLSGAAPLKGLEDAVRNRVPEQAILGQGYGMTEAGPVLMSCLCFAKQPFPTKSGSCGTVVR

At4CL1 387 NAEMKIVDPDTGDSISRNQFGEICIRGQIMKGYLNDAEATATIDKIGWLHTGDIGLIDDDDELFIVDRLKELIKYKGF
Ar4CL1 367 NAEMKIVDPDTGASLPRNQSGEICIRGQIMKGYLNDPEATARTIDKEGWLHTGDIGLIDDDDELFIVDRLKELIKYKGF
Ar4CL2 367 NAEMKIVDPDTGASLPRNQSGEICIRGQIMKGYLNDPEATARTIDKEGWLHTGDIGLIDDDDELFIVDRLKELIKYKGF
Ar4CL3 366 NAEMKIVDPDTGASLPRNQGETIYIRGSQIMKGYLNDPEATERTIDKIGWLQTDIGLIDGDELFIVDRLKELIKYKGF
Ar4CL4 392 NAEIKVIDPDTGCSLGRNQFGEICIRGSQIMKGYLNDAEATATIDVEGWLHTGDIGYVDDDELFIVDRLV-----

At4CL1 467 QVAPAELEALLIGHDIIDVAVVAMKEFAAGEVPVAFVVVSKDSELSDDVVKQFVSKQVSKCVLQENQOQSVLH-----
Ar4CL1 447 QVAPAELEALLLTHPNISDAAVVPMKDDIAGEVPVAFVVRNSGSOITEDLTKQFVSKQVVFYKRL---SRVFFIDVVPKS
Ar4CL2 447 QVAPAELEALLLTHPNISDAAVVPMKDDIAGEVPVAFVVRNSGSOITEDLTKQFVSKQVVFYKRL---SRVFFIDVVPKS
Ar4CL3 446 QVAPAELEALLLTHPNISDAAVVPMKDEAAGEVPVAFVVVNSGSKISEDDLIKQYTSQVVFYKRL---KRVFFIDVVPKA
Ar4CL4 463 --PPAELEALLVSHSIAADAAVVPMKDDIAGEVPVAFVVRSDFELTEAVKFEIAKQVVFYKRL---YKVVFFIDVVPKS

At4CL1 -----
Ar4CL1 524 PSGKILRKLRAKLAAGFSN-
Ar4CL2 524 PSGKILRKLRAKLAAGFSN-
Ar4CL3 523 PSGKILRKLRAKLAAGGLPD
Ar4CL4 538 PSGKILRKLRAKLAASPVP

```

Supplementary Figure S21. Protein sequence alignment of *bona fide* 4-coumarate CoA ligase from *Arabidopsis thaliana* (At4CL1 (21)) and that of putative homologs (Ar4CL1 – Ar4CL4) in red alder.

```

AtF5H      1  MESSISQILSKLSDPTTSLVIIVS LFFIFISFITRRRRE PYPGPRGPIIIGNMIMMDQLTHRGLANLAKKYGGLOHLRMG
ArF5H      1  -MSSLILQALOPLH--MAAFFIIELELFLFLGLIAFRIRQRRE PYPGPRGPIIIGNMIMMDQLTHRGLAKLAKOYGGIHLRIG

AtF5H      81  FLHMYAVSSPEVARQVLOVQDSVFSNRPATIAISYLTYDRADMAFAHYGPFWRQMRKLCVMKVFSSRKAESWASVRDEVD
ArF5H      78  YLHMVAVSSPDVARQVLOVQDNI FSNRPATIAISYLTYDRADMAFAHYGPFWRQMRKLCVMKVFSSRKAESWASVRDEVD

AtF5H     161  KMVRSVSCNVGKPINVGEQIFALTRNIYRAAFGSACEKGDQDEFIRILQEFSKLFGAFNVADFIPIYFCWIDPQGINRRLV
ArF5H     158  LTVREVAANIIGKPVNI GELVFSLTKNIIYRAAFGTSSHDGQDEFI GILQEFSKLFGAFNIADFIPIWLSWVDPOGLYIRLP

AtF5H     241  KARNDLDGFIDDIIDEHMKKKENQNAVDDGDVVDTDMVDDLAFYSEEAKLVSETADLQNSIKLTRDNIAKIIIMDVMFGG
ArF5H     238  R-----LASHWTG

AtF5H     321  TETVASAIEWALTELLRS PEDLKRVOQELAEVVGGLDRRVEESDIEKLYLTKCTLKETLRMHPIPLLHETAEDTSDIGF
ArF5H     246  SSTRSSITITWRRGRKIMAP-----

AtF5H     401  FIPKKS RVMINAFAIGRDPTSWTDPDTPFRPSRFLEPGVPDFKGSNFEFIPFGSGRRSCPGMQLGLYALDLAVAHILHCFE
ArF5H     265  -----LLKVRLIWMMTC--

AtF5H     481  WKLPDGMKPSSELDMNDVFGLTAPKATRLFAVPTTRLICAL
ArF5H     -----

```

Supplementary Figure S22. Protein sequence alignment of *bona fide* ferulate 5-hydroxylase from *Arabidopsis thaliana* (AtF5H (22)) and that of a putative homolog (ArF5H) in red alder.

```

AtCOMT1    1  MGSIAETQITP-----VOVIDEALFAMQLASASVLPMLKSALELDLLEIMAKNG--SPMSPIEIASRLPTKNPDA
ArCOMT1    1  MSSASETQITP-----TQVSDDEEANLFAMQLASASVLPMLKSALELDLLEIMATAGPGA LSPSETASQLPTTNPDA
ArCOMT2    1  MSSIEQTIPNPIITQQQELDQGEIEEVKLA VRLANGV VLPMLKSALELNLEIISDAGTGA LSPSETAGRLPTKNPDA

AtCOMT1    72  PVMLDRILRLLTYSVLTCSNRKLSGDCVERLYGLGPVCKY LTKNEDGVSIAALCLMNQDKVLMESWYHLKDAI LGGIP
ArCOMT1    74  PVMLDRILRLLASYSVLTYSRLTIPDGRVERLYGLGPVCKFLTKNEDGVSIAALNLMNQDKVLMESWYHLKDAI LEGGIP
ArCOMT2    81  PVMLDRILRLLASYSILRCSLVIREDEGVERLYGVGPKCKFLVKNPHGGSVAPLFLHHDKVLMESWYHLNDVILEGGIP

AtCOMT1    152  FNKAYGMFAFEYHGTDPRFNKVFNNCGMSNHSTITMKKILETYKGFEGLTSLVDVGGGIGATLRMIVSKYPNLRGINFDLP
ArCOMT1    154  FNKAHGMTSFEYHGKDLRFNKVFNKGMSDHSITITMKKILETYKGFEGLTSLVDVGGGTGAVLSMIVSKYPSIRGINFDLP
ArCOMT2    161  FNKAYGMTAFEYEGTDQRFNRFVNKAMSNHSTITMKKILLVYKGFEGIKVLVDVGGGIGVTLNITTSKYPQIRGINFDLP

AtCOMT1    232  HVIEDAPSHPGIEHVGGDMFVSVPKGDAIFMKWICHDWSDEHCVKFLKNCYESLPEDGKVI LAECILETPDSSLSTKQV
ArCOMT1    234  HVIEDAPSYPGVIEHVGGDMFVSVPKGDAIFMKWICHDWSDEHCLKFLKNCYALPNNKVI LAECILPVAPDLSLTKQV
ArCOMT2    241  HVIEDAPSYPGVEHIEGGDMFDSVEEGDAIFMKWILLHDWSDEHCLKLLKNCNKA LKALPNSGKVI LAESIRPVAPSSV SANIV

AtCOMT1    312  VHVDCIMLAHNPGGKERTKEFEALAKLSGFKGKIKVVCDAFGVNLIELLKKL
ArCOMT1    314  IHLVDVIMLAHNPGGKERTKEFEALAKLAGFQGFQVVCDAFNNTYIMEFLKKL
ArCOMT2    321  FEQDLIMLAQNPGGKERTKEFEALALQSGFSCEVICCAVNSWVMEFHK-

```

Supplementary Figure S23. Protein sequence alignment of *bona fide* caffeic acid *O*-methyltransferase from *Arabidopsis thaliana* (AtCOMT1 (23, 24)) and that of putative homologs (ArCOMT1 and ArCOMT2) in red alder.

```

AtCCOMT1 1 MATTTTEATKTSSTNGEDQKQSNLHQEVGHKSLQSDDLQYQIILETSVYPREPESMKELREVTAKHPWNIMTTSADEG
ArCCOMT1 1 -----MAFAYELADLEQAGSHQGFHGKSLQSDALYQYIILTSVYPREHEAMKELREMTAKHPWNIMTTSADEG
ArCCMOT2 1 -----MAFAYGLADLEQAGSHQGFHGKSLQSDALYQYIILTSVYPREHEAMKELREMTAKHPWNIMTTSADEG
ArCCOMT3 1 -----NQNLHFIPSRSSKSLQSDDLQYQIILETSVYPREPELTKELRHITAGHPRRAMATAFDAA

AtCCOMT1 81 QFLNMLEKLVNAKNTMEIGVYTGYSLLATALALPDDGKILAMDVNREHYELGLPIIKKAGVAHKIDFREGPALPVLDEIV
ArCCOMT1 70 QFLSMLLKLINAKNTMEIGVYTGYSLLATALALPDDGKILAMDVNREHYELGLPIIKKAGVAHKIDFREGPALPVLDKLV
ArCCMOT2 70 QFLSMLLKLINAKNTMEIGVYTGYSLLATALALPDDGKILAMDVNREHYELGLPIIKKAGVAHKIDFREGPALPVLDKLV
ArCCOMT3 61 QLLAMLLKLVNPKKTEVGVITGYSLLLATALALPDDGKILAMDVNREHYELGLPIIKKAGVADKINEIESEALPVLDDQLL

AtCCOMT1 161 ADEKNHGTDFDFVVDADKDNYNHYHRLIDLVKIGVIGYDNTLWNGSVVAPPDAPMRKYVRYRDFVLELNKALAADPR
ArCCOMT1 150 EDEKNHGTDFDFVVDADKDNYNHYHERLIKLVKVGGLIGYDNTLWNGSVAAPADAPPEYLYYRQFVVELNNALAVDPK
ArCCMOT2 150 EDEKNHGTDFDFVVDADKDNYNHYHERLIKLVKVGGLIGYDNTLWNGSVAAPADAPPEYLYYRQFVVELNNALAVDPK
ArCCOMT3 141 QTHENECSEFFAFDADKVNYSYHERLVKLVKVGGVVYDNTLWNGSVAAPADAPPEYLYYRQFVVELNNALAVDPK

AtCCOMT1 241 IEICMLFVGDGITICRRIS
ArCCOMT1 230 IEICMLFVGDGITICRRIK
ArCCMOT2 -----
ArCCOMT3 220 VQISHASHGDGITICRRLY

```

Supplementary Figure S24. Protein sequence alignment of *bona fide* caffeoyl CoA *O*-methyltransferase from *Arabidopsis thaliana* (AtCCOMT1 (25)) and that of putative homologs (ArCCOMT1 – ArCCOMT3) in red alder.

```

AtCCR1 1 MPVDVASPAGKTVCVTGAGGYIASWIVKLLERGYTVKGTVRNPDDPKNTHLRELEGKERRITLCKADLDYEAALKAA
ArCCR1 1 MPVDCCSASGLSTVVCVTGAGGFIASWIVKLLLEKGYTVKGTVRNPDDPKNAHLRELEGKERLITLCKADLDYEAALKVA
ArCCR2 1 -----MSGAEKVVVCVTGASGYIASWIVKLLLRQGYTVKASVRDENDPKMEHLLVFDGAKERLQLCKADLLEGAFFDSV

AtCCR1 79 IDGCDGVFHTASPVT----DDPEQMVPAVNGAKFVINAAA-EAKVKRVVITSSIGAVYM-DPNRDFEAVVDESCWSDLD
ArCCR1 80 IDGCDGVFHTASPVT----DDPELMVPAVNGTKNVIIAAA-EAKLFRVVITSSIGAVYM-DPNRDFPDVVDESCWSDLE
ArCCR2 75 VHGCEGVFHTASPCYHNVTDPQAEFLIPALRGTLNVVRSACKVPTVKRVVITSSVQAVAFNGKSLAPDVIIDETWSDPA

AtCCR1 153 FCKNTKNWYCYGKMAEQAWEVTAKEKGVDLVVNPVVLVGLPELQPTINASVYHVLKYLTSAGTYANIQAAYVDVRDVA
ArCCR1 154 FCKNTKNWYCYGKMAEQAWEVTAKEKGVDLVVNPVVLVGLPQLQPTNINASVYHVLKYLTSAGTYANSVQAAYVHVRDVA
ArCCR2 155 VGEKSKSWYRLSKTIAEEAAMKFAKENGIDVVINPGLVIGPLLQPTINSSAETVLKLVNCA-EREFNVTYRLVDVRDVA

AtCCR1 233 LAHVLVLEAPSASGRYLLAESARHRGEVVELLAKLFPEYPLPTKCSDEKNPRAKPYKFNQKTKLGLLEFTSTKQSLYDT
ArCCR1 234 LAHILVLETPSASGRYLCAESVLRGVVVOILLAKLFPEYPLPTKCSDEKNPRAKPYKFSNQKTKLGLLEFTPVKQCLYET
ArCCR2 234 NAHILALEKPSASGRYCLVGRVLRHCEVVKLRLRELFPLNLLEKCAADKPEFLSTYKVSKEVAESLGLNFTPVEVSLKDT

AtCCR1 313 VKSLQEKGHIAFPFPPPSASQESVENGKIKGS
ArCCR1 314 VKSLQEKGVLPVPTQQQE----SI-----RIGS
ArCCR2 313 VESLKEKNFLCA-----

```

Supplementary Figure S25. Protein sequence alignment of *bona fide* cinnamoyl CoA reductase from *Arabidopsis thaliana* (AtCCR1 (26)) and that of putative homologs (ArCCR1 and ArCCR2) in red alder.

C47 H69

```

AtCAD5 1 ---MGIM--AERKTTGWAARDPSGLSPMITYLRETEGPEDEVNIRIICCGICHIDLHQTKNDLG-MSNYPMVPGHEVVGVEV
AtCAD4 1 ---MGSV--AGEKALGWAARDPSGVLSPMITYLIRSTGADVYIKVICCGICHIDLHQTKNDLG-MSNYPMVPGHEVVGVEV
ArCAD1 1 ---MGSLE--TERKTTGWAARDPSGVLSPMITYLIRDITGPEDEVYIKVICCGVCHIDLHQTKNDLG-MSNYPMVPGHEVVGVEV
ArCAD2 1 -----MEGT--VI--GWAARDSSGHLSPMISFLLREAGEPEDEVYFKVLYCGMDHIDLHQMNEIHS-STNYPLVPVPGHEVVGVEV
ArCAD3 1 -MAKSP--EHPFKAFGWAARDSSGVLSPKFSRRATGDEEDVTF--VLYCGICHSDLHMIKNEWG-TSFYPLIPGHEFVGVV
ArCAD4 1 -MAKLP--EHPFKAFGWAARDSSGVLSPKFSRRATGDEEDVTF--VLYCGICHSDLHMIKNEWG-MSTYPLVPVPGHEFVGVV
ArCAD5 1 MARKSP--EHPFKAFGWAARDSSGVLSPKFSRRATGDEEDVTF--VLYCGICHSDLHMIKNEWG-YTSYPLVPVPGHEFVGVV
ArCAD6 1 -MAKSP--EHPFKAFGWAARDSSGHLSPKFSRRATGDEEDVTF--VLYCGICHSDLHMIKNEWG-GTSTYPLVPVPGHEFVGVV
ArCAD7 1 -MDKAP--EHPFKAFGWAARDSSGHLSPFNFSRRATKDDVRFKVLVYCGICHSDLHMIKNEWG-FSTYPLVPVPGHEFVGVV
ArCAD8 1 MVTKAP--EHPFKAFGWAARDSSGHLSPFNFSRRATKDDVRFKVLVYCGICHSDLHMIKNEWG-ISTYPLVPVPGHEFVGVV
ArCAD9 1 -MAKSP--EHPFKAFGWAARDSSGHLSPKFSRRATGDEEDVTF--VLYCGICHSDLHMIKNEWG-GTSTYPLVPVPGHEFVGVV
ArCAD10 1 -MAKSP--EHPFKAFGWAARDSSGHLSPKFSRRATGDEEDVTF--VLYCGICHSDLHMIKNEWG-GTSTYPLVPVPGHEFVGVV
    
```

Zn binding signature: **GHEXXGXX**

C103

XXXGXV C100 C106 C114

```

AtCAD5 76 VEVGSDVSKFTVGDIVVGVGVGCGGGSPCEERDLEIYCPKKEIWSYNDVYINQOPTQGGFAKATVWHQKXVVIPEGVAV
AtCAD4 77 LEVGSDVSKFTVGDIVVGVGVVGGCGGSKPCSSLEIYCNKRKWSYNDVYIDCKPTQGGFADTMVNOXKXVVIPEGVAV
ArCAD1 76 VEVGSDVSKFRVGTIVVGVGATVGCORNORPCNADTEIYCNKRKWSYNDVYIDCKPTQGGFAGAMVWHQKXVVIPEGVAA
ArCAD2 72 VEIIGSDVKKFRVKTIVVGVGIVGSCGECFSCKSNEEYCNHRILLTYNGTYKDKRPTQGGHSSAMVWHQKXVVIPEGVAP
ArCAD3 79 TEVGSKVQRKFKVGDVGVGCMVGAACHSCNCSNNLEIYCPKSLTYGAKYDGTITTYGGYSDTMVADEHFEVVIPEVNIPL
ArCAD4 79 TEVGSKVKKVKGDKVGVGCMVGAACHSCENCNDLEIYCPKEMILTYSSVYSDGTITTYGGYSDTMVANERYIIHFPEVNIPL
ArCAD5 80 TKIKGTVKRFKVGQVGVGVVGVSCKNCESCNDLEIYCPKMLTYNSTCHDGTIKTYGGYSDTMVVEERYVLRFPENIPL
ArCAD6 79 TEVGSKVKKVKGDKVGVGCMVGAACHSCNMDLEIYCPKMLTYNSFYSDGTITTYGGYSDTMVANERYIIHFPEVNIPL
ArCAD7 79 TEVGSKVKKVKGDKVGVGCMVGAACHSCENCNDLEIYCSQWILTYGFVYHDGTITTYGGYSDTMVANERYIIHFPEVNIPL
ArCAD8 80 TEVGSKVKKVKGDKVGVGCMVGAACHSCENCNDLEIYCSQWILTYGFVYHDGTITTYGGYSDTMVANERYIIHFPEVNIPL
ArCAD9 79 TEVGSKVKKVKGDKVGVGCMVGAACHSCENCNDLEIYCPKMLTYNSFYSDGTITTYGGYSDTMVANERYIIHFPEVNIPL
ArCAD10 80 TEVGSKVKKVKGDKVGVGCMVGAACHSCENCNDLEIYCPKMLTYNSFYSDGTITTYGGYSDTMVANERYIIHFPEVNIPL
    
```

L

C163 GLGGV

```

AtCAD5 156 EQAPLLCAGVTVYSPLSHFGLKQPLRGGIILGGVGHMVKITAKAMGHHVTVISSDKKREALEHLGADDFVYVSSDQ
AtCAD4 157 EQAPLLCAGVTVYSPLSHFGLMASGLRGGIILGGVGHMVKITAKAMGHHVTVISSDKKREALEHLGADDFVYVSSDF
ArCAD1 156 EQAPLLCAGVTVYSPLSHFGLKQSLRGGIILGGVGHMVKITAKAMGHHVTVISSDKKREALEHLGADDFVYVSSDS
ArCAD2 152 AQVAPLLCAGVTVYSPLKQKGSNKALKAGIILGGVGHMVKITAKAMGHHVTVISSDKKREALEHLGADDFVYVSSNT
ArCAD3 159 DAAPLLCAGITVYSPLKYYGLDKPGMIGVVLGGLGHLAVKFAKAMCVKVTVISTSEPKKEALEHLGADDFVYVSSIDQ
ArCAD4 159 DAAPLLCAGITVYSPLKYYGLAEPGRHIGVVLGGLGHLAVKFAKAFKAKVTVISTSEPKKEALEHLGADDFVYVSSRDQ
ArCAD5 160 DAAPLLCAGITVYSPLKYYGLAEPGRHIGVVLGGLGHLAVKFAKAFKAKVTVISTSEPKKEALEHLGADDFVYVSSDA
ArCAD6 159 DAAPLLCAGITVYSPLKYYGLAEPGRHIGVVLGGLGHLAVKFAKAFKAKVTVISTSEPKKEALEHLGADDFVYVSSRDQ
ArCAD7 159 DAAPLLCAGITVYSPLKYYGLAEPGRHIGVVLGGLGHLAVKFAKAFKAKVTVISTSEPKKEALEHLGADDFVYVSSHDQ
ArCAD8 160 DAAPLLCAGITVYSPLKYYGLAEPGRHIGVVLGGLGHLAVKFAKAFKAKVTVISTSEPKKEALEHLGADDFVYVSSRDQ
ArCAD9 159 DAAPLLCAGITVYSPLKYYGLAEPGRHIGVVLGGLGHLAVKFAKAFKAKVTVISTSEPKKEALEHLGADDFVYVSSHDQ
ArCAD10 160 DAAPLLCAGITVYSPLKYYGLAEPGRHIGVVLGGLGHLAVKFAKAFKAKVTVISTSEPKKEALEHLGADDFVYVSSRDQ
    
```

```

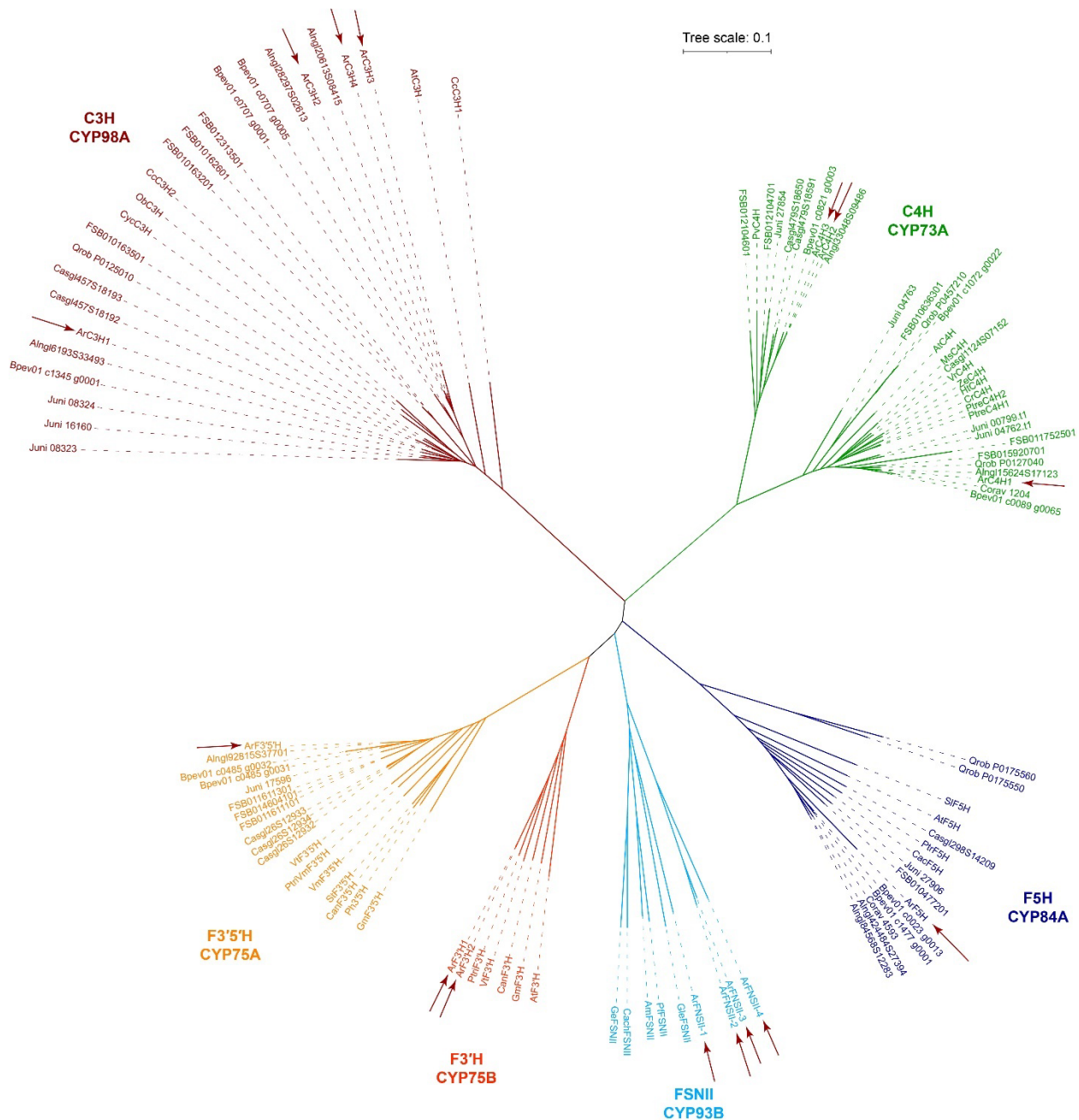
AtCAD5 236 AKMSELADSIDYIDTVPVHHALEPYLSLLKLDGKLLMVGVINNPLQFLTPLMLG-----EKVITGSEFTGSMKE
AtCAD4 237 AEMQRLADSIDYIDTVPVHHALEPYLAQLKLDGKLLMVGVINNPLQFVTPVLILG-----RKMVTSSTFTGSMKE
ArCAD1 236 TRMQEVADSIDYIDTVPVHHALEPYLSLLRDLGKLLMVGVINNPLQKFRQRKKNNGYLISECVWERDNHRELGRSMKE
ArCAD2 232 AEMERATNMDYIDTVPVAVHPQSYLPLKVDGKLLMVGIAKPLQFDSVDMILG-----RKTITGSEFTGSMKE
ArCAD3 239 EQMQAAGTIDYIDTVSAAKHPLEPLDLLKSHGKLVLMGAPEKPLELPVFPLLMG-----RKMVGGSGTGGMKE
ArCAD4 239 EQQAAKNTMDGIDTVSAAHSLPLDLLKSHGKLVLMGAPEKPLELPVFPLLMG-----RKTIVAGSATGGMKE
ArCAD5 240 AKMKAAMTMDYIDTVSAAVHAPLEPLDLLKLNGLKLVMLGPEKPLELPVFPLLMG-----RKMVGGSDVGGMKE
ArCAD6 239 EQQAAKNTMDGIDTVSAAHSLPLDLLKSHGKLVLMGAPEKPLELPVFPLLMG-----RKTIVAGSCTGGMKE
ArCAD7 239 EQMQAAGTIDYIDTVSAAHHPLEPLDLLKSHGKLVMLGPEKPLELPVFPLLMG-----RKMVTVGSGTGGMKE
ArCAD8 240 EQMQAAMTMDGIDTVSAAVHAPLEPLDLLKSHGKLVMLGAPKPLELPVFPLLMG-----RKMVVGSAAGGMKE
ArCAD9 239 EQQAAKNTMDGIDTVSAAHSLPLDLLKSHGKLVLMGAPEKPLELPVFPLLMG-----RKTIVAGSCTGGMKE
ArCAD10 240 EQQAAKNTMDGIDTVSAAHSLPLDLLKSHGKLVLMGAPEKPLELPVFPLLMG-----RKTIVAGSCTGGMKE
    
```

```

AtCAD5 306 TEEMEFCKEKGKLSITIEVVKMYYVNTAFERLEKNDVRYRFVVDVAGSNLDA-----
AtCAD4 307 TEEMEFCKEKGKISTIEVTKIDEINIAERLRKNDVRYRFVVDVAGSNLVEEAATTTN
ArCAD1 316 TEEMEFCKEKGKCSMIEVVKMYYVNTAKAERLEKNDVRYRFVVDVAGSNLDQ-----
ArCAD2 302 TQEMLEFALKGKISMEVVKMYYVNTAFERLEKNDVRYRFVVDVAGSNLE-----
ArCAD3 309 TQEMVDFAAKHNIADIEVVISMEYVNTAMERLSKADVRYRFVIDIGNILKSSSF-----
ArCAD4 309 TQEMVDFAAKHNIADTEIIPMEYVNTAMERLAKGDVRYRFVIDIGNILAAATKSSS--
ArCAD5 310 TQEMVDFSAKHGICADIEIIRMEDINSAMARLAKSDVRYRFVIDVANSLSQ-----
ArCAD6 309 TQEMVDFAAKHNIADTEIIPMEYVNTAMERLAKGDVRYRFVIDIGNILAAATK----
ArCAD7 309 TQEMVDFAAKHNIADTEIIPMEYVNTAMERLAKGDVRYRFVIDIGNILDATK----
ArCAD8 310 TQEMVDFAAKHNIADIEVVISMEYVNTAMERLAKGDVRYRFVIDIGNILAAATK----
ArCAD9 309 TQEMVDFAAKHNIADTEIIPMEYVNTAMERLAKGDVRYRFVIDIGNILAAATK----
ArCAD10 310 TQEMVDFAAKHNIADTEIIPMEYVNTAMERLAKGDVRYRFVIDIGNILAAATK----
    
```

Supplementary Figure S26. Protein sequence alignment of *bona fide* cinnamyl alcohol

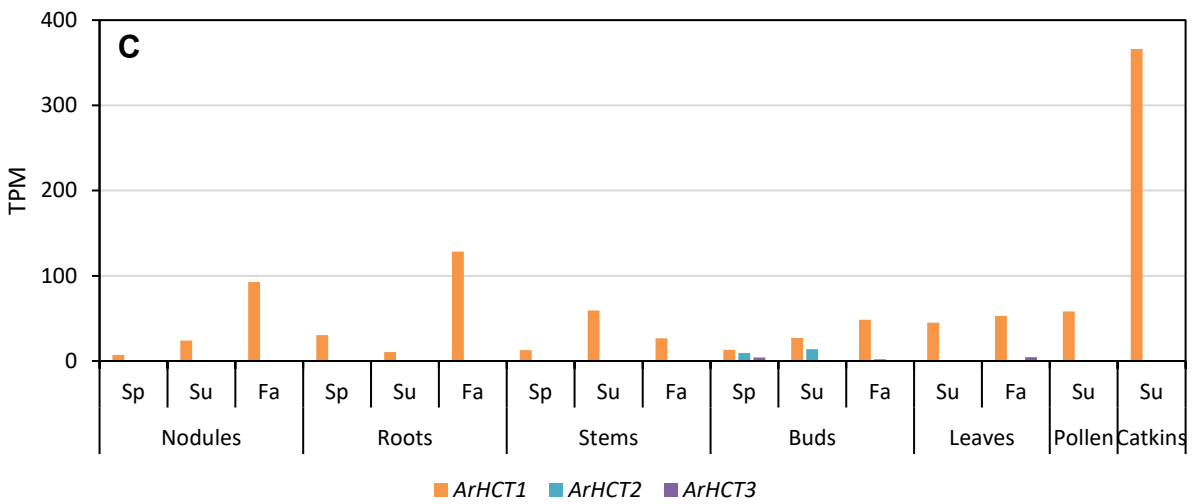
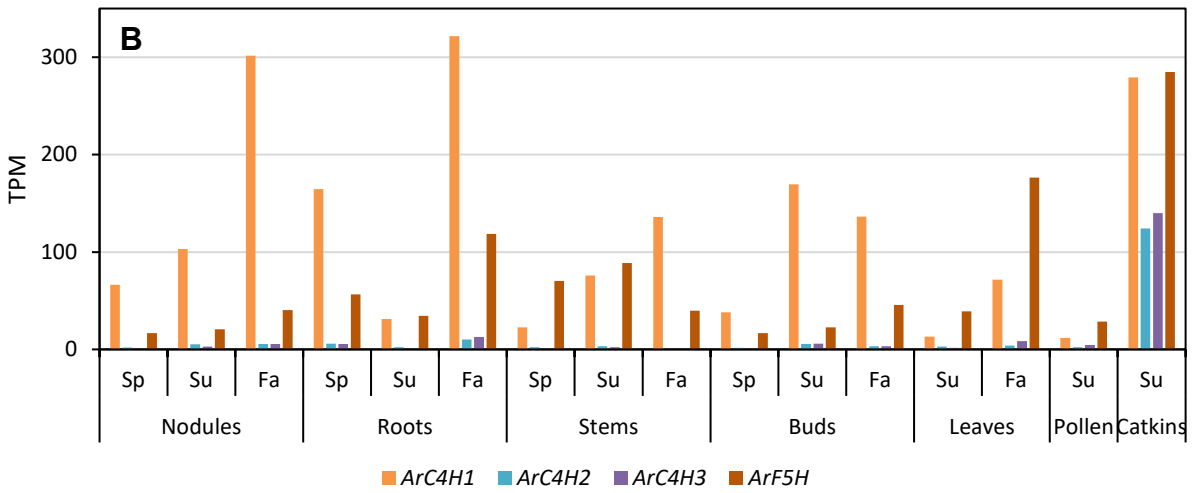
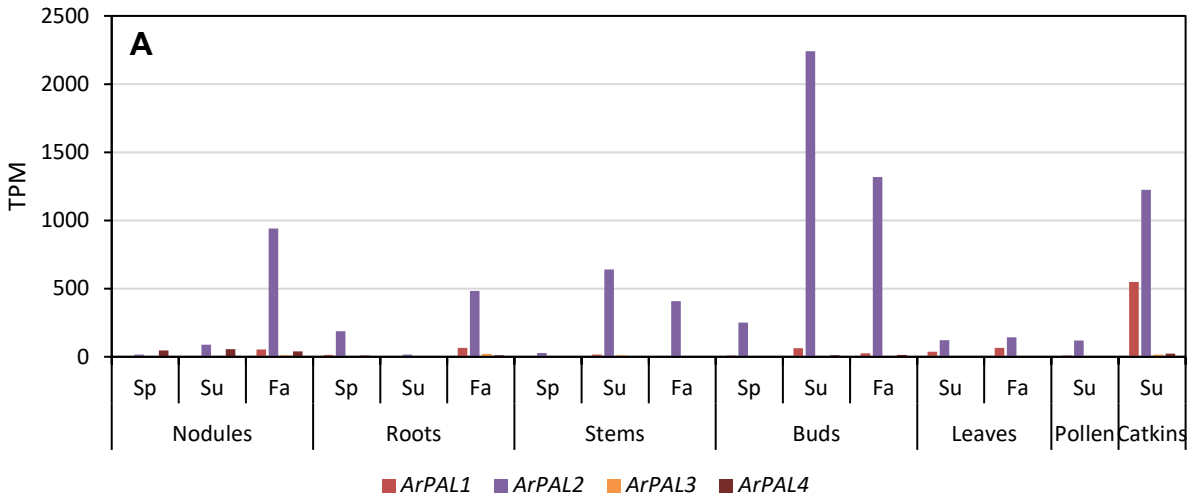
dehydrogenases from *Arabidopsis thaliana* (AtCAD4 and AtCAD5 (27)) and that of putative homologues (ArCAD1 – ArCAD10) in red alder. Highly conserved Zn1 catalytic center (C47, H69, and C163), the Zn-binding signature GHEXXGXXXXXGXXV, the Zn2 structural motif (C100, C103, C106, and C114), and the NADPH-binding domain [GLGGV(L)G] motif are highlighted.

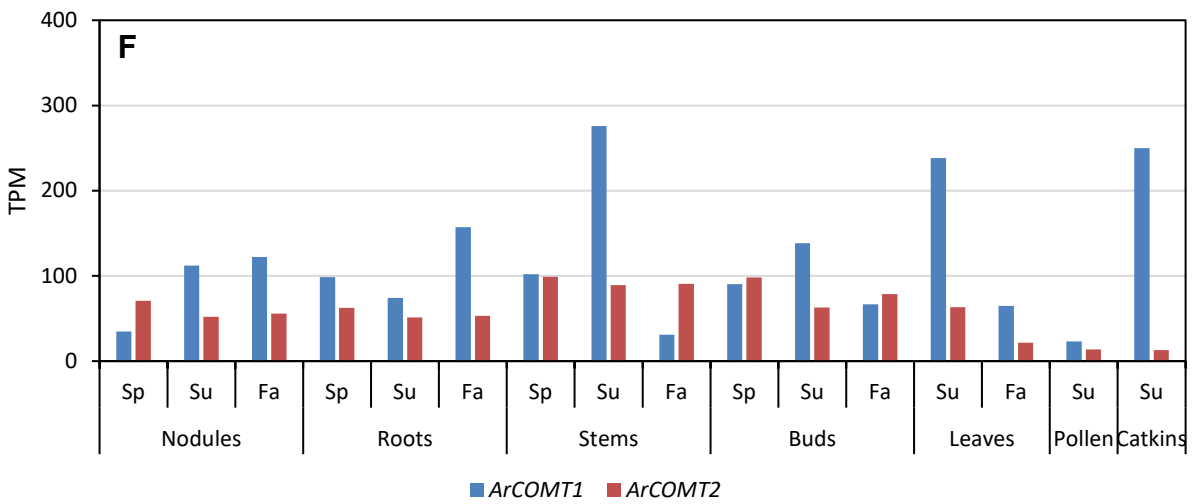
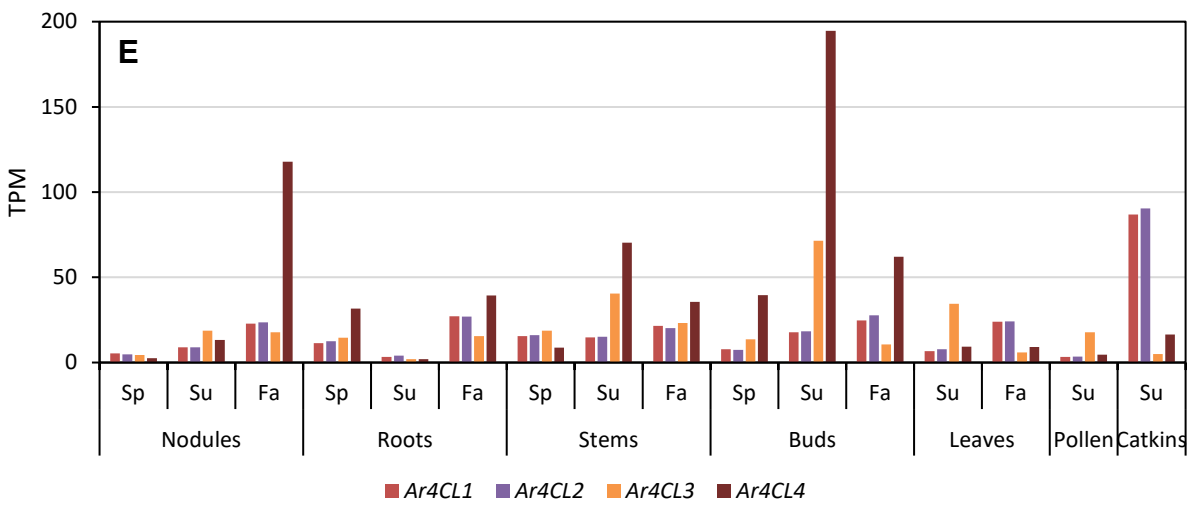
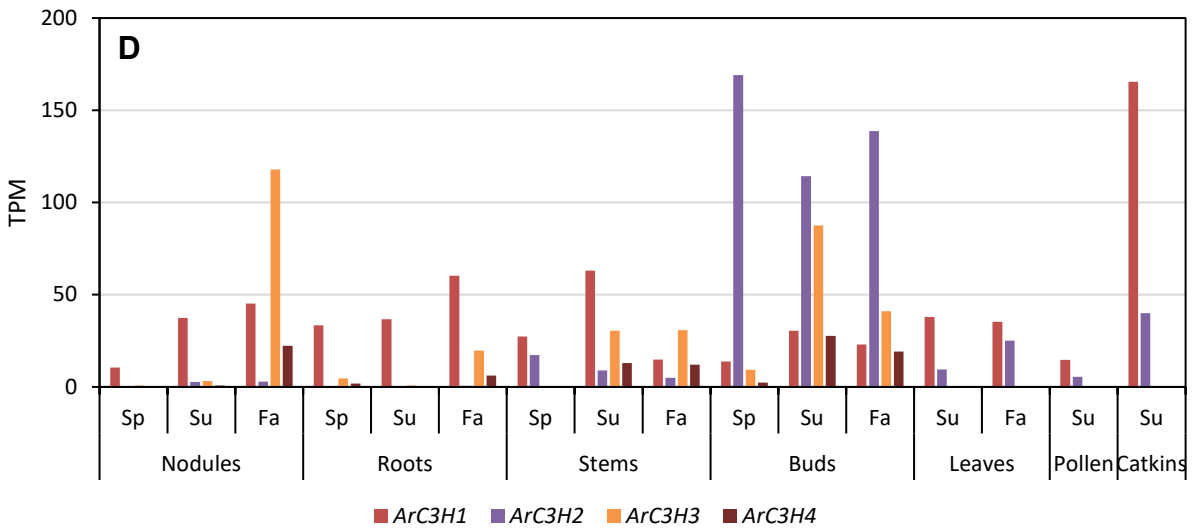


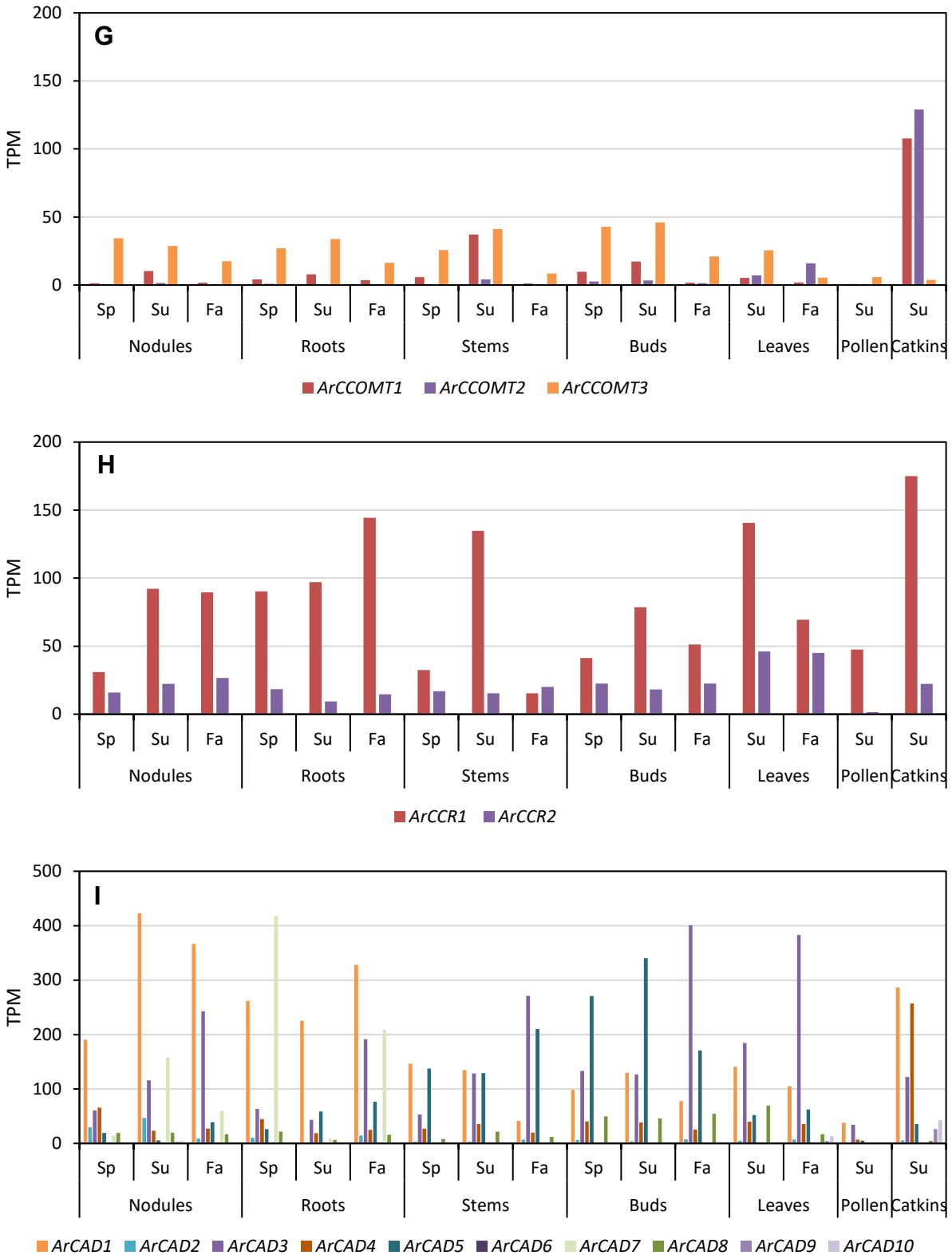
Supplementary Figure S27. Unrooted phylogenetic tree of hydroxylases involved in phenylpropanoid (C4H, C3H and F5H) and phenylpropanoid-acetate (F3'H, F3'5'H and FSNII) pathways from *Alnus glutinosa* (Alngl), *Alnus rubra* (Ar), *Antirrhinum majus* (Am), *Arabidopsis thaliana* (At), *Betula pendula* (Bpev), *Callistephus chinensis* (Cach), *Camptotheca acuminata* (Cac), *Capsicum annuum* (Can), *Casuarina glauca* (Casgl), *Catharanthus roseus* (Cr), *Coffea canephora* (Cc), *Corylus avellana* (Corav), *Cynara cardunculus* (Cyc), *Fagus sylvatica* (FSB), *Gerbera* hybrid cultivar (Ge), *Glycine max* (Gm), *Glycyrrhiza echinata* (Gle), *Helianthus tuberosus* (Ht), *Juglans nigra* (Juni), *Medicago sativa* (Ms), *Ocimum basilicum* (Ob), *Perilla frutescens* (Pf), *Petunia × hybrida* (Ph), *Phaseolus vulgaris* (Pv), *Populus tremuloides* (Ptre), *Populus trichocarpa* (Ptri), *Quercus robur* (Qrob), *Solanum lycopersicum* × *S. peruvianum* (Sl), *Solanum tuberosum* (St), *Vigna radiata* (Vr), *Vinca major* (Vm), *Vitis vinifera* (Vt) and *Zinnia elegans* (Ze). Sequences were obtained from NCBI and a multiple-

sequence alignment was built using Clustal Omega (57), with the unrooted phylogenetic tree rendered with iTOL (58). *Dark red* arrows show red alder hydroxylases.

AmFSNII (CYP93B3; BAA84071), AtC3H (CYP98A3; At2g40890), AtC4H (CYP73A5, At2g30490), AtF3'H (CYP75B1, At5g07990), AtF5H (CYP84A1), CacF5H (AAT39511), CachFSNII (CYP93B5; AAF04115), CanF3',5'H (QGZ19237), CanF3'H (XP_016563358), CcC3H1 (CYP98A35; ABB83676), CcC3H2 (CYP98A36; ABB83677), CrC4H (CAA83552), CycC3H1 (CYP98A49; ACO25188), GeFSNII (CYP93B2; AAD39549), GleFSNII (CYP93B1; P93149), GmF3',5'H (AAM51564), GmF3'H (BAB83261), HtC4H (CAA78982), MsC4H (L11046), ObC3H (CYP98A13; AAL99200), PfFSNII (CYP93B6; BAB59004), PhF3',5'H (CAA80266), PtreC4H1 (ABF69099), PtreC4H2 (ABF69101), PtrF5H (CAB65335), PtriF3'H (XP_002319761), PtriVmF3',5'H (XP_002314004), PvC4H (CAA70595), SIF5H (AAD37433), StF3',5'H (NP_001274807), VmF3',5'H (BAC97831), VrC4H (P37115), VtF3',5'H (BAE47007), VtF3'H (BAE47004), ZeC4H (Q43240). Sequences from other Fagales species were retrieved (using BLASTP) from the corresponding databases as indicated in Supplementary Table S1.

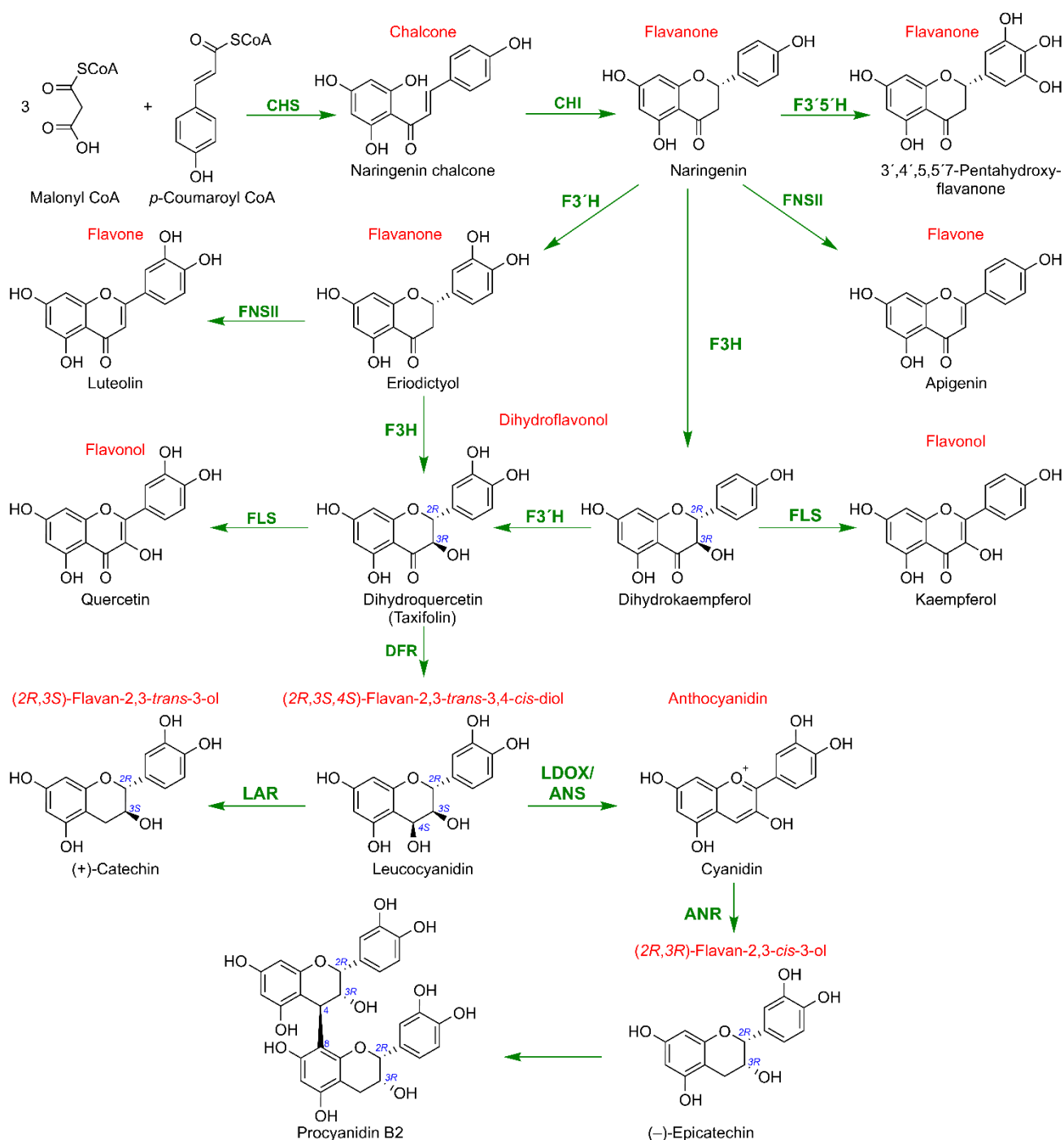






Supplementary Figure S28. Expression levels of red alder homologs of *bona fide* phenylpropanoid pathway genes. **(A)** Phenylalanine ammonia lyase (*ArPAL*). **(B)** Cinnamate 4-hydroxylase (*ArC4H*) and ferulate 5-hydroxylase (*ArF5H*). **(C)** Hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl

transferase (*ArHCT*). **(D)** *p*-Coumarate 3-hydroxylase (*ArC3H*). **(E)** 4-Coumarate CoA ligase (*Ar4CL*). **(F)** Caffeic acid *O*-methyltransferase (*ArCOMT*). **(G)** Caffeoyl CoA *O*-methyltransferase (*ArCCOMT*). **(H)** Cinnamoyl CoA reductase (*ArCCR*). **(I)** Cinnamyl alcohol dehydrogenase (*ArCAD*). Abbreviations: TPM, transcripts per million; Sp, spring; Su, summer; Fa, fall.



Supplementary Figure S29. Simplified flavonoid biochemical pathway. ANR: Anthocyanidin reductase; CHI: Chalcone isomerase; CHS: Chalcone synthase; DFR: Dihydroflavonol 4-reductase; F3',5'H: Flavonoid 3',5'-hydroxylase; F3'H: Flavonoid 3'-hydroxylase; F3H: Flavonoid 3-hydroxylase; FLS: Flavonol synthase; FNSII: Flavone synthase II; LAR: Leucoanthocyanidin reductase; LDOX/ANS: Leucoanthocyanidin dioxygenase/anthocyanidin synthase.

AmCHS 1 MVTVEVRKAQRAQGPATVLAIGTATPANCVDQSTYPDYFRITNSEHMTELKEKFKRMCQKSNIKRRYMHLTEEILKEN
ArCHS1 1 MVTVEVRKAQRAQGPATVLAIGTATPPNCVEQSTYPDYFRITKSEHKTTELKEKFKRMCQKSMIKKRYMYLTEEILKEH
ArCHS2 1 MASVDEIFKAQRAQGPATVLAIGTANPSHCYQDDFPDYFRITKSEHMTDLKRFKRCQKSMIKKRYTHLTEILKEN
ArCHS3 1 MASVDEIFKAQRAQGPATVLAIGTATPNSCVSQADYDPDYFRITKSHMTELKEKFKRMCQKSMIKKRYMYLNEEILKEN
ArCHS4 1 -----

AmCHS 81 PAMCEYMAPSLDARQDIVVVEVPLKGEAAQKAIKEWQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLRPSVKRFMM
ArCHS1 81 PNIQAYMAPSLDARQDIVVVEIPKLGKEAATKAIKEWQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLRPSVKRIMM
ArCHS2 81 PSMCTFMEPSLDARQDIVVVEIPKLAREAAKAIKEWQPKSKITHLFESTTLCTSDMPGADYQLTKLLGLKRSVKRIMM
ArCHS3 81 PNMCAVMAPSLDARQDIVVVEVPLKLGKEAATKAIKEWQPKSKITHLVFCTTSGVDMPGADYQLTKLLGLRPSVKRIMM
ArCHS4 1 -----MPGADYQLTKLLGLKRSVKRIMM

AmCHS 160 YQQGCFAGGTVLRVAKDLAENNAAGARVLVVCSEITAVTFRGPDTHLDSL VGQALFGDGAAAIVGSDVVGVERPLFQI
ArCHS1 160 YQQGCFAGGTVLRVAKDLAENNKARVLVVCSEITAVTFRGPDSDHDSL VGQALFGDGAAAIVGADPVEVEPLFEI
ArCHS2 161 YQHGCFCGGTVLRVAKDFAENHKGARVLAVCAEITASNFRGPDSSHFDALVGOQLFGDGAAAIVGADPDTSVRPLFQI
ArCHS3 160 YQQGCFAGGTVLRVAKDLAENNKARVLVVCSEITAVTFRGPDTHLDSL VGQALFGDGAAAIVGADPDTVERPLFEI
ArCHS4 24 YQHGCFCGGTVLRVAKDFAENHKGARVLAVCAEITASNFRGPDYASHLFDALVGOQLFGDGAAAIVGADPDTGVERPLFQI

AmCHS 240 VTAQAQTLLPDSHGAI DGHLEVGLTFHLLKDVPGIISKNIEKSLKEAFDPLGISDWNSEFWIAHPGGPAILDQVEKLG
ArCHS1 240 VSAQAQTLPDSDGAI DGHLEVGLTFHLLKDVPGIISKNIEKSLVEAFDPLGIDWNSEFWIAHPGGPAILDQVESKLG
ArCHS2 241 VSAQAQTLLPDSGAI DGHLEVGLTFHLLKDVPGIISKNIESSLVESEFETGISDWNSEFWIHPGGPAILDQVEARLGL
ArCHS3 240 VTAQAQTLPDSDGAI DGHLEVGLTFHLLKDVPGIISKNIEKSLAEAFDPLGISDWNSEFWIAHPGGPAILDQVESKLG
ArCHS4 104 VSAQAQTLLPDSGAI DGHLEVGLTFHLLKDVPGIISKNIESSLVESEFETGISDWNSEFWIHPGGPAIL--VEARLGL

AmCHS 320 KPEKLRSTRQVLSEYGNMSSACVLFILDEMRKSSAKEGMS TTGEGLDWGVLFQFGPGLTVETVVLHSPIN---
ArCHS1 320 KAEKLAATRHLSEYGNMSSACVLFILDEMRKSSAENGLKTTGEGLEWGVLFQFGPGLTVETVVLHSPST---
ArCHS2 321 EEEKLRARHVLSEYGNMSSACVLFILDETRKRSVEEGKATTCGLDWGVLFQFGPGLTVETVVLHSPLETLH
ArCHS3 320 KEEKLRATRHLSEYGNMSSACVLFILDEMRKSSAEKQKVTGEGLEWGVLFQFGPGLTVETVVLHSPVPV--
ArCHS4 182 EEEKLRARHVLSEYGNMSSACVLFIFG-----

Supplementary Figure S30. Protein sequence alignment of *bona fide* chalcone synthase from *Antirrhinum majus* (AmCHS (28)) and that of putative homologs (ArCHS1 – ArCHS4) in red alder.

PhCHI 1 MSPSISVTKVQVENYVFPPTVKPPASTKTLFLLGGAGHRLDVEGKFKVFTVIGVYLEESAVQFLAPKWKGKSAEELIHSV
ArCHI 1 MVSPISLSGVQVENVEFPPTVKPPASTKTLFLLGGAGERLEIQGKFKVFTVIGVYLEDTAVPSLAVKWKGKSAEELTESV

PhCHI 81 DFFRDIVTGPFKEFTRVRFILPLTGQKQSEKVAENCVAHWKATCTYS DAGSRAIEKFLNIVKSEITFLPGASILFTQSPNG
ArCHI 81 EFFRDIVTGPFKEFTRVTMILPLTGQQQSEKVSENCVAHWKSVGVYIDAEKKAIEKFLVFKDKKFPKSSSILFTQSPNG

PhCHI 161 SLITTSFTKDDSISEAGNAVIENKQFSEAVLEIITGEGHGVSPAARKSIAARMSEIFKNSLF-----
ArCHI 161 SLAIGFSGKQSIPEIIGNAVIENKLLSGAVLESIIGKHGVSPAARKSLAARLSEILAEGKGPETEKV

Supplementary Figure S31. Protein sequence alignment of *bona fide* chalcone isomerase from *Petunia × hybrida* (PhCHI (29)) and that of a putative homolog (ArCHI) in red alder.

AtF3H 1 MAPCFTLTELAGEKSLNSKFVRDEDERPKVAYNVFSDEIPVISLAGIDVDGKRREICRQIVEACENWGFQVVDHGVDIN
ArF3H 1 MAPFTLTELAGEKTLQTSFVRDEDERPKVAYNVFSSEIPVISLAGIDVDGRRREICRQIVDACEWGFQVVDHGVDAN

AtF3H 81 LVADMTRLARDFALPPEIKLRFDMSSGGKGGFIVSSHLQGEAVQDWREIVTYFSYPTIRNRDYSRWPDKPEAGWVAVTEBY
ArF3H 81 LISDMTRLARDFALPPEIKLRFDMSSGGKGGFIVSSHLQGEAVQDWREIVTYFSYPTIRSRDYSRWPDKPEAWRAVTEQY

AtF3H 161 SERLMSLACKLLEVLSEAMGLEKESLTNACVDMQKIVVNYYPKCPQPDLLTGLKRHTDPGTITLLLQDQVGLQATRDN
ArF3H 161 SDRLEMLACKLLEVLSEAMGLEKESLTKACVDMQKIVVNYYPKCPQPDLLTGLKRHTDPGTITLLLQDQVGLQATRDN

AtF3H 241 GKTWITVQPEVGAFFVNLGDHGHFLSNGRFKNADHQAVVNSNSRSLSIATFQNPAPDATVYPLKREGEKAVLEAPITFA
ArF3H 241 GKTWITVQPEVGAFFVNLGDHGHFLSNGRFKNADHQAVVNSNYSRSLSIATFQNPAPDAKVYPLKREGEKAVLEAPITFA

AtF3H 321 EMYKRKMGSDTELARLKKLAKEERDHL--REVAKPVQDIFA
ArF3H 321 EMYKRKMSKSDTELARLKKLAKEQDLEKDRMDAKPEDIIFA

Supplementary Figure S32. Protein sequence alignment of *bona fide* flavanone 3-hydroxylase from *Arabidopsis thaliana* (AtF3H (30)) and that of a putative homolog (ArF3H) in red alder.

AtF3'H MATLFLITLLATVLFLLRIFSHRRNRSHNNRLLPPGPNPWPIIGNLPHMGTVPHRTLSAMVITYGPIIHLRLGFVDVVVA
 ArF3'H1 -MSPFSLILGTVALAIFLYYLLNLRTHRHRLFLPPGPKPWPIIGNLPHMGTVPHHSLAAMARTYGPLMHLRLGFVDVVVA
 ArF3'H2 -MSPFSLILGTVALAIFLYYLLNLRTHRHRLFLPPGPKPWPIIGNLPHMGTVPHHSLAAMARTYGPLMHLRLGFVDVVVA

AtF3'H ASASVAHQFLKHDSNFSRPPNSGAKHMAYNYQDLVFAPYGERWRRLRKISSVHLFSGKALDEFRHRVQEEVAVLNRAL
 ArF3'H1 ASASVAHQFLKHDSNFSRPPNSGAKYIAYNYQDLVFAPYGERWRRLRKISSVHLFSGKALDEFRHRVQEEVAVLNRAL
 ArF3'H2 ASASVAHQFLKHDSNFSRPPNSGAKYIAYNYQDLVFAPYGERWRRLRKISSVHLFSGKALDEFRHRVQEEVAVLNRAL

AtF3'H VRVGTKPVNLGQLLVNMCVFNALGRVMEGRRLFCADA---DHKADEFKSMVTEMMLAGVFNIGDFIPSLDWLDLQGVAGK
 ArF3'H1 ANAGQKPVNLGQLLVNMCVFNALGRVMEGRRVFGDGSGGGDAKADEFKSMVTEMMLAGVFNIGDFIPALEWLDLQGVAAK
 ArF3'H2 ANAGQKPVNLGQLLVNMCVFNALGRVMEGRRVFGDGSGGGDAKADEFKSMVTEMMLAGVFNIGDFIPALEWLDLQGVAAK

AtF3'H MKRLHHRFDFAFLSSILKEHEMNGQDKHTDMLSTLISLKGTDIDGGGSLTDTEIKALLLNMFAGTDTTSASTVDWAIAE
 ArF3'H1 MKKLHNRFDAFLTGIVEEHKKNNGG-KHMDLSTLISLK-ENADGEGKLTDEIKALLLNMFAGTDTTSSTVEWAIAE
 ArF3'H2 MKKLHNRFDAFLTGIVEEHKKNNGG-KHTDMLSTLISLK-ENADGEGKLTDEIKALLLNMFAGTDTTSSTVEWAIAE

AtF3'H LIRHPDIMVKAQEELDTVVGDRREVNESDIAQLPYLQAVIKENFRLHPPSTPLSLPHIASESCEINGYHIPKGSTLLINIV
 ArF3'H1 LIRHPKILAQVRQELDSVVGQERLVSELDVAQLTYLQAVVKEIFRLHPPSTPLSLPRMADDDCEIAGYHIPKGSTLLINIV
 ArF3'H2 LIRHPKILAQVRQELDSVVGQERLVSELDVAQLTYLQAVVKEIFRLHPPSTPLSLPRMADDDCEIDGYHIPKGSTLLINIV

AtF3'H AIARDPDQWSNPLAFKPERFLPGGEKSGVDVKGSDFELIPFGAGRRICAGMSLGLRITQFLTATLVQCFDWFVLAGGVTP
 ArF3'H1 AIARDPDQWSNPLFRPERFLPGSEKAGVDVKGNDDEFIPFGAGRRICAGMSLGLRMVHLLTASLVHAFDWLWLANGLMPE
 ArF3'H2 AIARDPDQWSNPLFRPERFLPGSEKASVDVKGNDDEFIPFGAGRRICAGMSLGLRMVHLLTASLVHAFDWLWLANGLMPE

AtF3'H KLNMEESYGLTLQRAEPLMVHPRPRLAPNVYGLGSG
 ArF3'H1 KLNMD EAYGLTLQRAEPLMVHPRPRLSPHVYRASS-
 ArF3'H2 KLNMD EAYGLTLQRAEPLMVHPRPRLSPHVYRASS-

Supplementary Figure S33. Protein sequence alignment of *bona fide* flavonoid 3'-hydroxylase from *Arabidopsis thaliana* (AtF3'H (31)) and that of putative homologs (ArF3'H1 and ArF3'H2) in red alder.

PhF3'5'H 1 ----MMILTELCAATSFLIAHIIISLISKTTGRLLPPGPRGWPVIGALPLLGMPHVSLAKMAKKYGAMYLKIGTC
 ArF3'5'H 1 MAALDIFLFRRLATAITLFFITLFSVRSLL-GKPSRLPPGPRGWPVIGALPLLGMPHVSLANMAKKYGPMYLKIGTC

PhF3'5'H 76 GMMAVASTPDAAF AFLKTLDFNSNRPPNAGATHLAYNAQDMVFAHYGPRWKLRLKLSNLHMLGGKALDHWANVRANELGH
 ArF3'5'H 80 NMVVASTPDAAF AFLKTLDFNSNRPPNAGATHLAYNAQDMVFADYGRWKLRLKLSNLHMLGGKALDWAHVRGANELGH

PhF3'5'H 156 MLKSMSSMSREQRVVVEMLTAMANMIGQVLSRVRVFDKGEVNEFKDMVVELMTIAGFNIGDFIPCLAWMDLQGI
 ArF3'5'H 160 MLRAMNLSKQCEAVVVEMLTAMANMIGQVLSRVRVFTKGEESNEFKDMVVELMTSAGFNIGDFIPCLAWMDLQGI

PhF3'5'H 236 EKRMKRLHKKFDALLTKMFEHKAATYERKGGKPDFLDVVMENGDNSEGERLSTTNIKALLLNLFAGTDTSSSAIEWALA
 ArF3'5'H 240 EGMKRLHKKFDVLLTKMFEHTASAHQRKGGKPDFLDVVMANRENSAGEKLLSTNIKALLLNLFAGTDTSSSAIEWALA

PhF3'5'H 316 EMNKNPATLKKQAQEMDQVIGRNRRLLESDIPLNLPYLAICKETFRKHPSTPLNLPRLSNPEPCIVDGYIIPKNTRLSVNI
 ArF3'5'H 320 EMNKNPATLKKRAHEEMDKVIGRNRRLLESDIQKLPYLAICKETMRKHPSTPLNLPRLSTVAGEVDGYIIPKNTRLSVNI

PhF3'5'H 396 WAIGRDPQVWENPLDFNPERFLSGRNSKIDPRGNDFELIPFGAGRRICAGTRMGIVVEVIFILGTLVHVSFDWKLPESEVIEL
 ArF3'5'H 400 WAIGRDPQVWENPLDFNPERFLSGRNSKIDPRGNDFELIPFGAGRRICAGTRMGIVVEVIFILGTLVHVSFDWKLPESEVIEL

PhF3'5'H 476 NMEEAFLGLAQKAVPLAAMVTPRLQLDVIYVP
 ArF3'5'H 479 NMEESFGLAQKAVPLAAMVSPRLSLISAYAS

Supplementary Figure S34. Protein sequence alignment of *bona fide* flavonoid 3',5'-hydroxylase from *Petunia × hybrida* (PhF3',5'H (32)) and that of a putative homolog (ArF3',5'H) in red alder.

```

PhFLS      1  --MKTAAQGV-----SATLTMEVAVRQAIASLSKCMDTIPSEVIRSENEQPAATLHGVLQVPEVIDLRDPDENKMKVRIAD
ArFLS1    1  MGIERVQGIASLDGHKKLMGVERVQGIASLS--KDEI PAEFIRSVNEQPGITTVHGEVLEVPIDLDPDEEKVIRATAD
ArFLS2    1  -----MAVESVQAIASLSLTNTNIPAEFIRPEKEQPAITTFHGPAPEIPTIDLDPDQENLVRSTAD

PhFLS     75  ASKEWGI FQLINHGIPDEAIADLQKVGKEFFEHVPOEKEELIAKTPGSDNIEGYGTSLOKELEGKKVWDHDLFHKIWPPS
ArFLS1    79  ASINWGI FQIVNHGIPVEAITKLAQSKKEFFE-LPOEKEEVYAKPPGPHSMGYGTLQKELEGKKVWDHDLFHKIWPPS
ArFLS2    63  ASKEWGI FQVVNHGIPTEVVSNLQAVGKEFFQ-LSPFEEKEAYAKPPGACSMEGYGNKLOKDLNKKKSWVDHDLFHKIWPPS

PhFLS    155  AVNYRWF PKNPPSYREANE EYLRKREVVDRIFKLSLSLGLGLEGHEMI EAAGDEIVYLLKINYYPPCPRPDALGVVAH
ArFLS1   158  AVNYRWF PKNPPSYREANE EYAKQLHGCVVFKLFSLSLGLGLEGHELKEAASGDIMVYLLKINYYPPCPRPDALGVVAH
ArFLS2   142  LKNYHFW PKNPPSYREANE EYVARYVREVADKLLTRLSLGLGLQGDALKEAGGHEMVLKINYYPPCPRPDALGVVAH

PhFLS    235  TDMSYITILVPNEVQGLQVFKDGHWYDVKYIPNALIVHIGDQVEILSNGKYKSVYHRTTVNKLKTRMSWPVFLEPPSEHE
ArFLS1   238  TDMSYITILVPNEVQGLQASFDGHWYDVKYIPNALIVHIGDQVEILSNGKYKAVLHRTTVNKEQTRMSWPVFLEPPSEFE
ArFLS2   222  TDSISITILVPNEVPGLOVLRDWRVIAKYIPNALIVHIGDQVEILSNGKYKSVLHRSVVDKEKARMSWPVFLEPPSEFV

PhFLS    315  VGPTRKLLSEANPPKFKTKKYKDYVYCKLNKLPQ
ArFLS1   318  VGRHPKLV DHENPPKYTKKYSDYVYCKLNKLPQ
ArFLS2   302  VGRHPQLVNEENPKKYARKKDYAYCKLNKLPQ

```

Supplementary Figure S35. Protein sequence alignment of *bona fide* flavonol synthase from *Petunia × hybrida* (PhFLS (33)) and that of putative homologs (ArFLS1 and ArFLS2) in red alder.

```

AmDFR     1  MSPTSLN TSETAPPSSTFVCVGTGAAGFIGSWLMRLLERGYTVRATVRDPGNMKKVKHLIELPKADTLLTLWKADMVTE
ArDFR1    1  -----MGA ESETVCVTGAAGFIGSWLMRLLERGYTVRATVRDPEENMKKVKHLELPKAKTLLTLWKADLADE
ArDFR2    1  -----MGESESVCVTGA SFIGSWLMRLLERNYSVRATVRDPGKYKVEGQDPNIGKKRQ-----

AmDFR    81  GSFDEAIQ GCEGVFHLATSMEFDSVDPNEVIKPTIDGVLNIRKSCVQAKTVKFTFTISCGTVNVEEHQKPVYDEIDS
ArDFR1    69  GSFDEAIK GCGSGVFHVATPMDFFESODPNEVIKPTVNGVLSIMKACVEAKTVRLVFTSSAGTIDVSGHQKPVNDESCWS
ArDFR2    58  GNR-----REDKI IQLKLFLSLCGTAKMNEVIKPTINGVILGIMNACVKAKTVRKLVFTSSAGSVNVEEHQKPVYDESCWS

AmDFR   161  DDFIN SKKMTGWMYFVSKTLAEQA AAKENND FISIIPPLVVGPFIMPFP PPSLITALSPTITGNEAHYSIIKQCGQ
ArDFR1   149  DVDFCR I KKM TGWMYFVSKTLAEQA AWFAKENND FISIIPPLVVGPFIMPSMPPSLITALSPTITGNEAHYSIIKQGF
ArDFR2   133  DVDFCRA KKM TGWMYFVSKTLAEQA AWFAKENND FDIIPPLVVGPFIMPSPMPPSLITALSPTITGNEAHYSIIKQGF

AmDFR   241  VHLDDLCE GHIFLF EYPKAEGRYICSSH DATIYDI AKLITENPEYHIFDEFEGIDKDFVVSFSSKKVIGMGFIFKYIL
ArDFR1   229  VHLDDLCA HIFLF EYHPKAEGRYICSA D ATILDAKLLREKYPEYNYPTKFKVDENLFVVSFVSKKVKDLGFOFKYSL
ArDFR2   213  VHLDDLCA HIFLF EYHPKAEGRYICSA D ATIHDI AKLLREKYPEYNIPTKLGIEKLDIVSFSSKKVKELGFOFKYSL

AmDFR   321  EDMVRGAI DTCREKGLPEYSTKNNK GDEKEPI LNSLENNYNIQDKELFPISEEKHINGQENALLSNTQDKELLPTSEEKR
ArDFR1   309  EDMFV EAVETCRAGLIP LA AENS VSGTS-----
ArDFR2   293  EDMFV GAVETCRAGLLP PAAGN-----

AmDFR   401  VNGLESALLSKIQDK EVLPTSGVKHAKGQENALLPDIANDHTDGRI
ArDFR1   -----
ArDFR2   -----

```

Supplementary Figure S36. Protein sequence alignment of *bona fide* dihydroflavonol 4-reductase from *Antirrhinum majus* (AmDFR (34)) and that of putative homologs (ArDFR1 and ArDFR2) in red alder.

```

DuLAR      1  M TVSGAIPSMTRKNTLVGGTGFIGQF TKASLGF GYPT LLVRPGPVS --PSKAVLIK FQDKGAK IYGVINDKEOME
ArLAR1    1  M TVIT-SVSTAKGRVLIAGATGFIGQFVAEASLDAGRPTY LLVRPGVGGGPKASTI AFQDKGATL HGLINDKEFME
ArLAR2    1  M CGSGPNVSEAGGRAMIIGSTGFIGQFVAEASLDSGRPTYLLVQSSQAS --PSKASTIKSLQDKGAILIYGSIEDRDLLE

DuLAR      79  KILKEYEIDVVISLVGGARLDQLTLEAIKSVKTIKRFLPSEFGHDVDRTPDVEPGLTMYKEKRLVRRAVEEYGIPTN
ArLAR1    80  KILKEHEIDLVISAVGGNLLDQLTLEAIKSVGTIKRFLPSEFGHEVDRADPVEPGLDMYQAKRKVRRLLIEYGVPHTY
ArLAR2    79  KILKEHKLIVVISAIGDRISDQLTLVDAIKAVGTIKRFLPSEFGHDDRADPVEPALTFYNSKRLVRRVTEAAGIANTY

DuLAR     159  ICCNSIASWPYYDNCCHPSQVPPDQFQIYGDGNTKAYFDGNDIGKFTMKTIDDIRTLNKNVHFRPSSNCYSINELASL
ArLAR1    160  ICCNSIASWPYYDNTHPSVVPPLDQFQIYGDGNVKAYFVAGADIGKLTIKTIDDIRANKCVHFRPAGNLYFINELASL
ArLAR2    159  ICCNSIAAWPYHDNTHPAVLPPLDRFYIYGDGTVKAYFVAGADIGKFTMKSIDDIRTLNKCWHFRPSSNLLNINELASL

DuLAR     239  WEKKIGRTLPRFTVADKLLAAHAENIIPESIVS SFTHDFINGCQVNFSDI EHS DVE DILY PDEKFRLLDDCEDFP
ArLAR1    240  WEKKIGRTLPRVTVSENDLLAAEAENRIPQ SIVASETHDFIKGCQVNFIDGPD DVEVGALYPDETFRLLDFCNDFLV
ArLAR2    239  WEKKIKRELPRVSI SEDDLLAAAKMIRIPESIVAAL THDFIKGCQINAMDKN DVEVCSLYPDPFRFTIDDCFNDLTK

DuLAR     319  MVHDKIHAGKSGE----KIKDKKPE--LVQGTSTIEEINKDIKTLVETQPNEEIKKDMKALVEAVPISAMG
ArLAR1    320  NIDSKSKSTFEIT---TASEVVEP---MATAAYA-----
ArLAR2    319  NSVGAPKVVVEEPISSNDALIAFNSKEALVITA-----

```

Supplementary Figure S37. Protein sequence alignment of *bona fide* leucoanthocyanidin reductase from *Desmodium uncinatum* (DuLAR (35)) and that of putative homologs (ArLAR1 and ArLAR2) in red alder.

```

AtLDOX    1  -----MVAVERVESLAKSGTISIPKEYIRKKEELENSINDV FLEE KEDGPOVPTIDLKNIESDDEKIRENCIEELKKA
ArLDOX    1  MVTNSSSMAERVESLSSSGIQAIKEYVRKQEELNSIGNV FEEEKKEGPOVPTIDLKNIESDDEAVREKCREELKRAAV

AtLDOX    76  DWGVMHLINHGIPADL MERVKKAEEFFSLSVVEEKEYANDQALGKIQGYGSKLANNASGQLEWEDYFFHLAYPEKRD
ArLDOX    81  EWGVMHIVNHGIPDELL ERVKEACKVFFDLPVVEEKEYANDQALGKIQGYGSKLANNASGQLEWEDYFFHLAYPEKRD

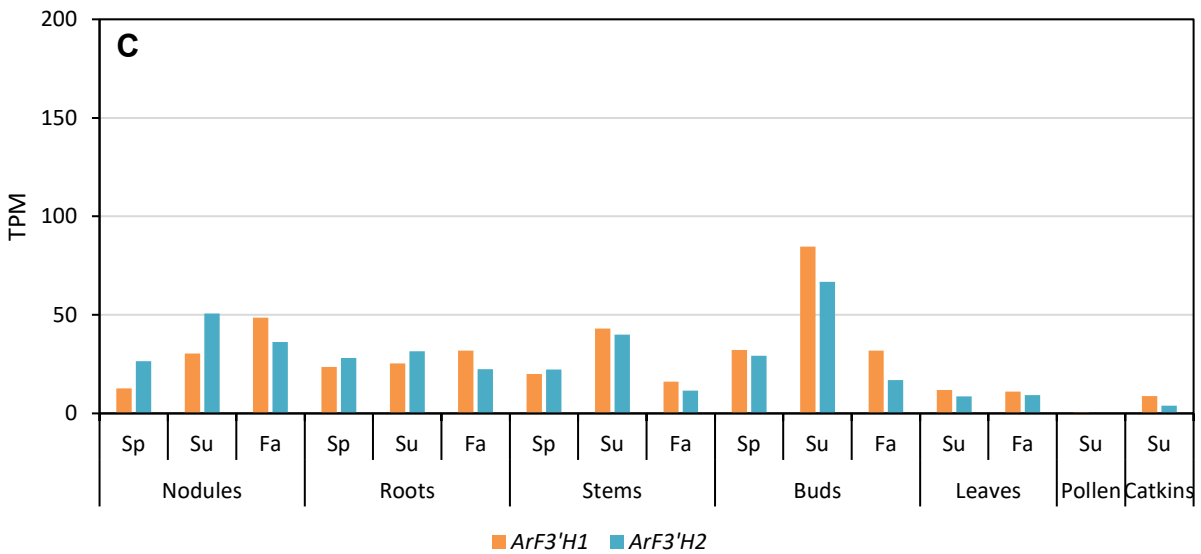
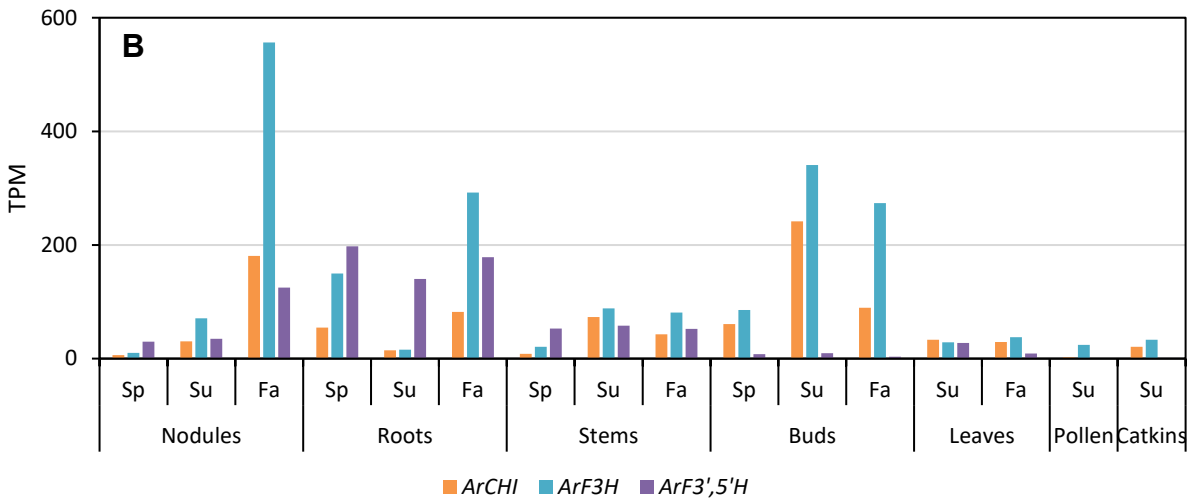
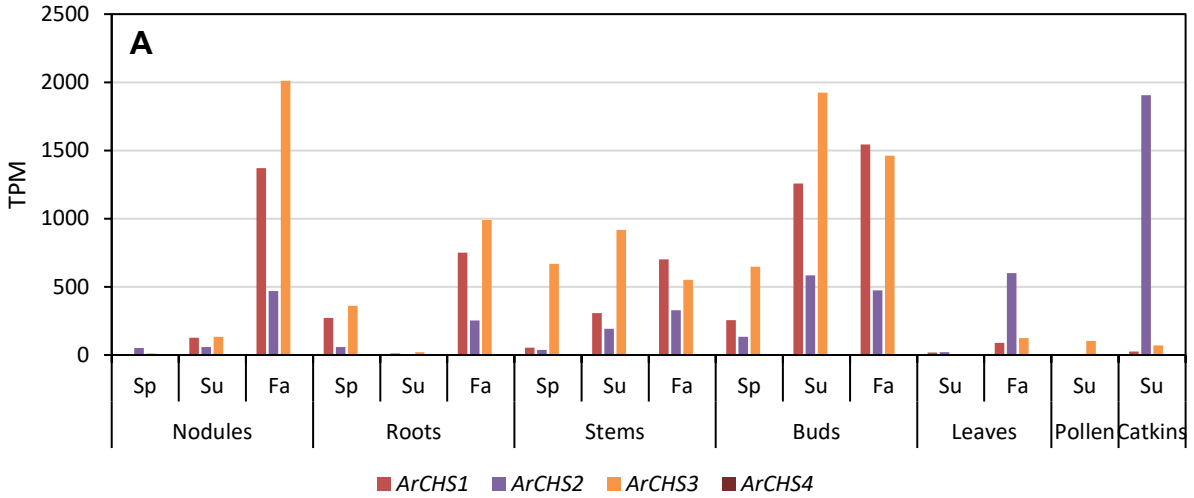
AtLDOX    156  SIWPKTPSDYIEATSEYACLRLLATKVFKALS GLGLEPDRLEKEVGGLEELLQMKINYYPCPQPELALGVEAHTDV
ArLDOX    161  SIWPKTPSDYIAATSEYACLRKLASKVLSVLS GLGLEEGRLEKEVGGLEELLQMKINYYPCPQPELALGVEAHTDI

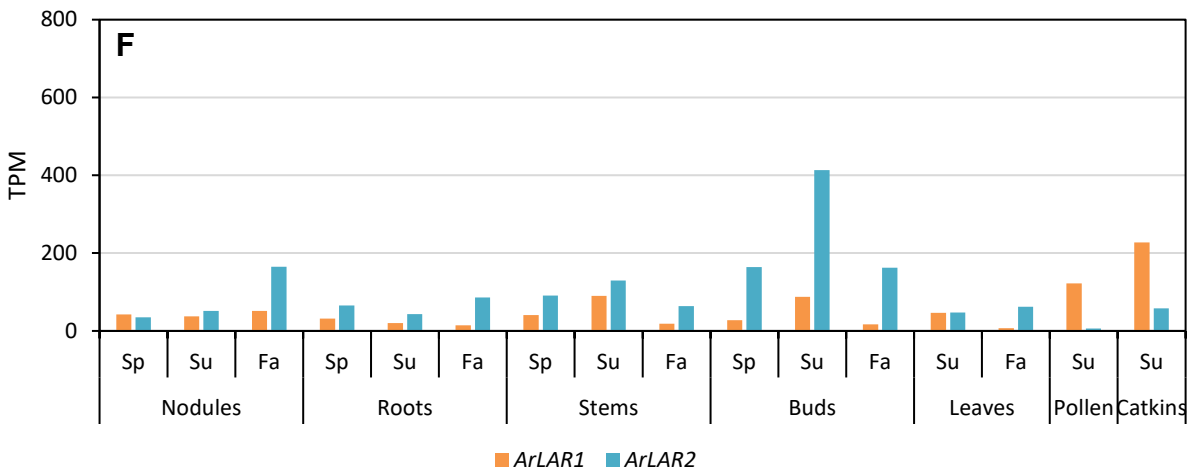
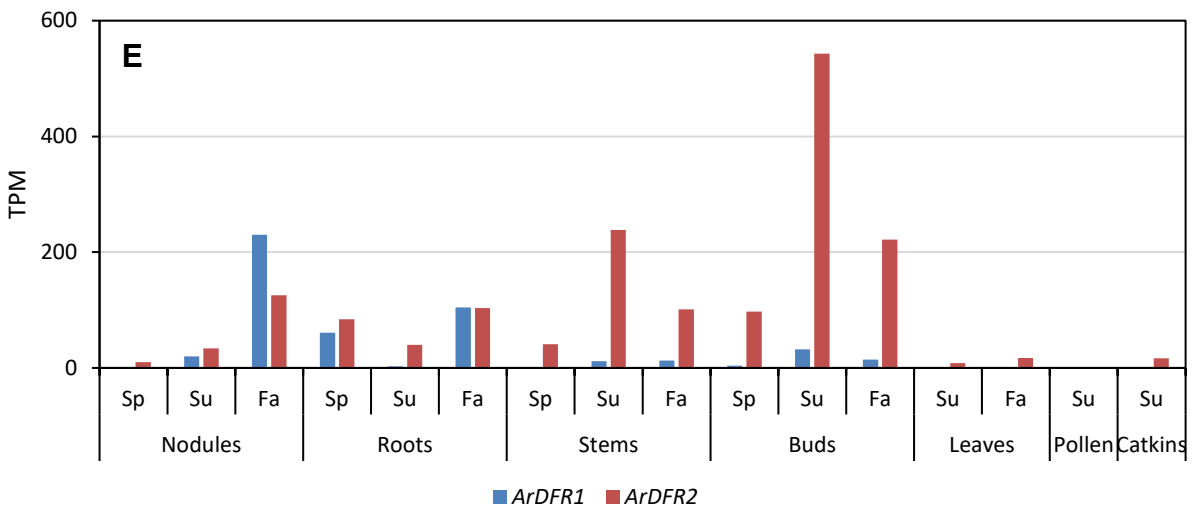
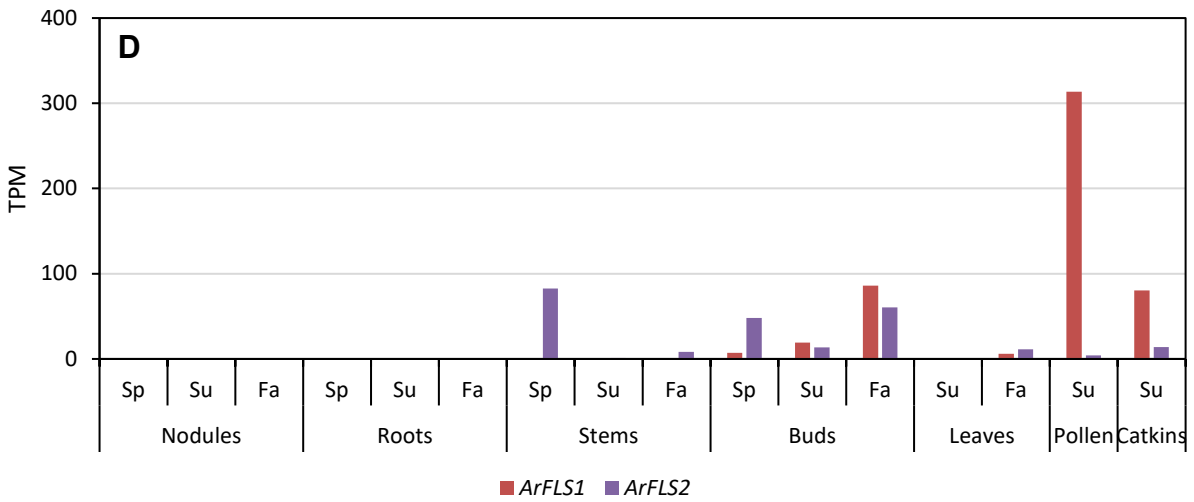
AtLDOX    236  SALTFILHNMPGLQLFYEGKQVTAACVPSIIMHIGDTLEILSNCKYKYSILHRGLVNKEKVRISWAVFCEPPKIKIILK
ArLDOX    241  SALTFILHNMPGLQLFYQKQVTAACVPSIIMHIGDTLEILSNRKYKYSILHRGLVNKEKVRISWAVFCEPPKIKIILK

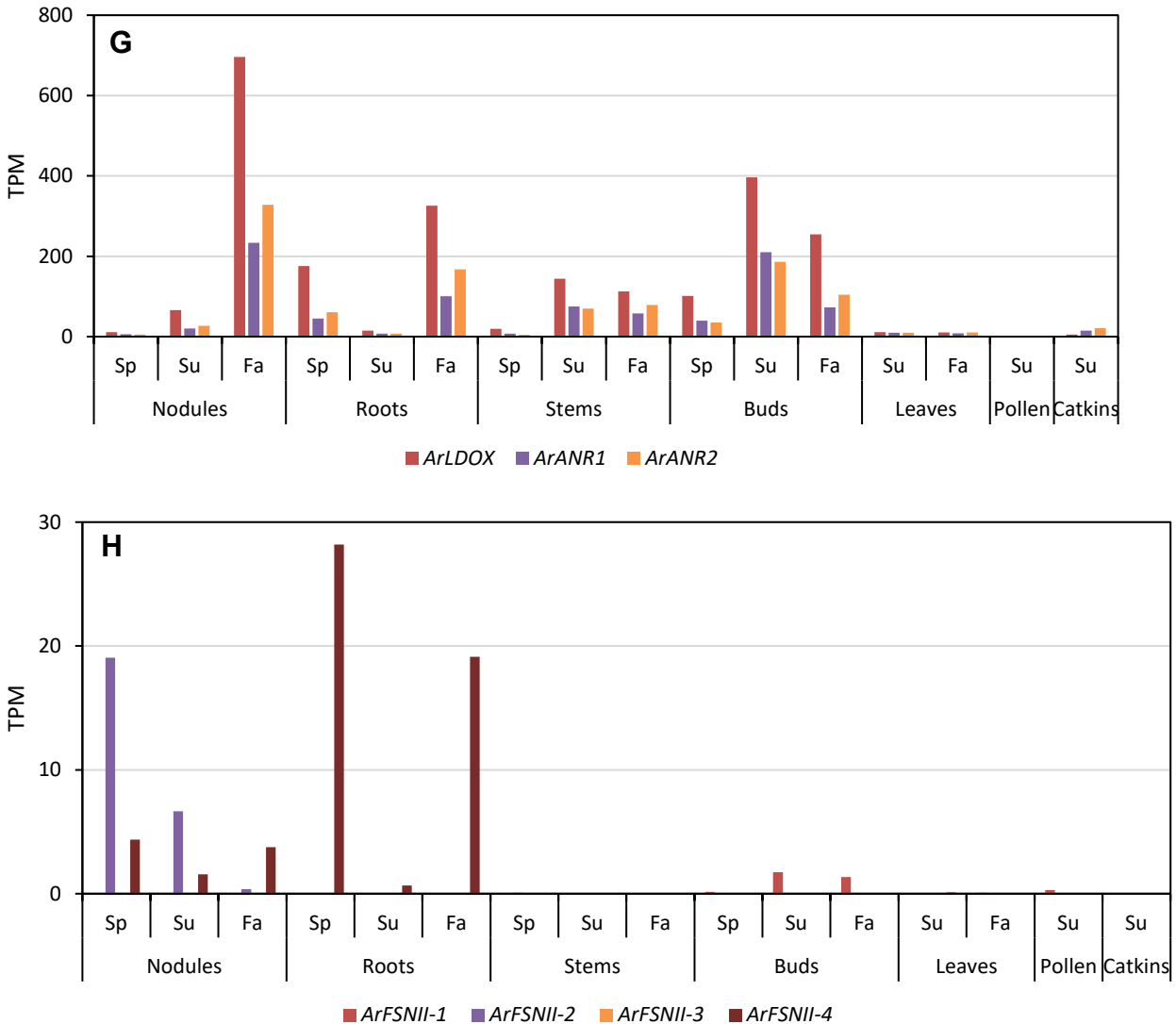
AtLDOX    316  PLPEVSVSESPAKFPPRTFAQHIQHKLFGKEQDEIVSEKND
ArLDOX    321  PLPEVSEEEPAIFPPRTFAQHIQHKLFRKIQDALDAK---

```

Supplementary Figure S38. Protein sequence alignment of *bona fide* leucoanthocyanidin dioxygenase [also called anthocyanidin synthase (ANS)] from *Arabidopsis thaliana* (AtLDOX (36)) and that of a putative homolog (ArLDOX) in red alder.







Supplementary Figure S41. Expression levels of red alder homologs of *bona fide* flavonoid pathway genes. **(A)** Chalcone synthase (*ArCHS*). **(B)** Chalcone isomerase (*ArCHI*), flavanone 3-hydroxylase (*ArF3H*), flavonoid 3',5'-hydroxylase (*ArF3',5'H*). **(C)** Flavonoid 3'-hydroxylase (*ArF3'H*). **(D)** Flavonol synthase (*ArFLS*). **(E)** Dihydroflavonol 4-reductase (*ArDFR*). **(F)** Leucoanthocyanidin reductase (*ArLAR*). **(G)** Leucoanthocyanidin dioxygenase (*ArLDOX*) and anthocyanidin reductase (ANR). **(H)** Flavone synthase II (*ArFNSII*). Abbreviations: TPM, transcripts per million; Sp, spring; Su, summer; Fa, fall.

```

DRR206      1 --MG---SKLL--VLFVFMVLFAL----SSAIPNKRKPKYKPCKNLWYFHDILYNGKNAANATSAIVAAPPCVSLTKL
ArDIR1     1 --MG---GRYYYVVFQRLTL-LAV-----SSAF-GGKKQYKPCKRSLYFHDITYNGENAANATSAIVAAPPGANLTIL
ArDIR2     1 --ME---SNIQILAFFLFFLL-LGSMAAPAQRKDKAKDRPCCRRLWYFHDIVYNGKNLRNATSAIVGAPAWANTITM
AtDIR6     1 --MAFLVEKQLFKALFSFLLVLFSDTVLSF--RKTIDQKPKCKHFSFYFHDILYDGDNVAANATSAIVSPFG-----
ArDIR3     1 MAMGELVQ---KSCFLLFLMLIC-QSVLASQKYYSLRHRKPKCRLELYYHDILENGTDMANATSAKATNQTAE-----

DRR206     68 APQSHFGNLTIVFDDPITLSHSLSSKQVGRAQGFYIYDKNTYTSWLSFTFVLSNSTHQQGTITFAGADPIVAKTRDISVTG
ArDIR1     69 AGQFHFGNLTAVFDDPITLDNNLFSKPVGRAQGLYIYDKNTFTAWLGFTFVLSNSTHQQGTINENGADPIIAKTRDISVVG
ArDIR2     75 AGQNHFGDLVVFDDPITLDNNLSTPVGRAQGFYVYDKKDIETAWLGFSTFVLSNSTQHKGSINEAGADPLMNAKTRDISVVG
AtDIR6     71 LGNFKFKGFVFDGPITLDKNYLSPKPVRAQGFYFYDKMDENSWFSITLVFNSTEHKGTINIMGADLMMEPTRDISVVG
ArDIR3     70 LGNFFHFGMLVVFDDPITLDSHLSPPLEARAQGFYFYNNKDDENAWFSITLVFNSSQHKGTINIMGADLMMEPTRDISVVG

DRR206     148 GTGDFFMHARGIATITDAFEGEAYFRLLGVYIKFEECW
ArDIR1     149 GTGDFFMHARGIATIMTDAFEGEVYFRLLVVDIKFYECS
ArDIR2     155 GTGDFFMHARGIATIMTDAFEGEVYFRLLVVDIQLYECS
AtDIR6     151 GTGDFFMHARGIATFVTDLFDQAKYFRVMDIKLYECS
ArDIR3     150 GTGDFFMHARGIATIRTDLFDQGSYFRLLVMDIKLYECS

```

Supplementary Figure S42. Protein sequence alignment of *bona fide* dirigent proteins from the Dir-a sub-family, DRR206 from *Pisum sativum* (39) and AtDIR6 from *A. thaliana* (60) and that of putative homologs (ArDIR1 – ArDIR3) in red alder.

```

PsPTS1      1  -----MAK-SSS-----FFITFF-----ISLISFATATKIQSLSPTMLGF--QDEKFLHLHFYFHD
ArDIR4      1  -----MAK-SISKITPILFEL--HSTYVT----AETHNFSRNLSPSSLGL--KKEKLSHLHFYFHD
ArDIR5      1  -----MAK-TLPKVTPTVLELFTLFSFSAF---TtagsFSRALSPAKLGL--KKEKLSHLHFYFHD
ArDIR6      1  -----MAK-TLPKVTPTFLELFTLISFATF---ATAGGFSRTLSPAKLGL--KKEKLSHLHFYFHD
ArDIR7      1  -----MAK-TLPKVIPTFLELFSLFSFSAF---ATAGGFSRTLSPVKFGL--KKEKLSHLHFYFHD
ArDIR8      1  -----MAK-ILPIAFQLLEFFLSSFSSTVSVAGDQDEGFSRSLDRKLLGL--KKEKLSHFRFYIHD
ArDIR9      1  -----MQALYSIIVP---MGR-ILPIASQLLEFFLSSISTVSVTGE-DEGFSRSLDRKLLGL--KKEKLSHFRFYIHD
ArDIR10     1  -----MAR-ILPIASQLLEFFLSSFSSTVSVAGE-DGGFSRSLDRKLLGL--KKEKLSHFRFYIHD
ArDIR11     1  -----MAR-ILPIASQLLEFFLSSFSSTVSVAGE-DGGFSRSLDRKLLGL--KKEKLSHFRFYIHD
ArDIR12     1  -----MAR-ILPA---FIVFFLHYFSATLV-YGEDKGFGTPLDPKLLHKV--KKEKLSHFNFFIHD
ArDIR13     1  -----MAR-ILPIFPSKFIIFSLISFLAIIIVTG-ENHGFVRNLDLRRLLGQIKKKEKLSHIFHYIHD
ArDIR15     1  -----MAR-ILPA---FIVFFLIGSFAATSMASAREEQGFCRPLDRKKHKL--KKEKLSHFNFFIHD
ArDIR21     1  MYSNPMQALYSIIVP---MAK-ILPIAFQLLEFFLSSFSSTVSVAGDQDEGFSRSLDRKLLGL--KKEKLSHFRFYIHD

PsPTS1      50  IIVIGPKPSMVFVAEPNGKVENALEFGTIVAMDPLTAGPERDSKLVGIAQGLYITSIISOEMGLMVMVMTAFSGEENGST
ArDIR4      54  IVSGENPFAVRVAEAATNTSATGFGAVVMDDPLTLEPSSSKLVGRAQGYALASQHEAGLLMVNFAFEGKYSGST
ArDIR5      56  IVSGRNPFAVRVAEAAMNMSLTGFGAVFMDDPLTTRPELSSKLVGRAQGYASASQTEVGLMVMNFAFEGKYNGSS
ArDIR6      56  IVSGENPFAVRVAEAPMNTSLTGFGAVVMDDPLTTRPELISKHVGRAQGYASASQTEVGLMVMNFAFEGKYNGSS
ArDIR7      56  IVSGRNPFAVRVAEAPMNTSLTGFGAVVMDDPLTTRPELISKHVGRAQGYASASQTELGLLMVMNFAFEGKYNGSS
ArDIR8      60  IVSGRNPASAVLVVP-PA-NT-STGFGLVSMIDPLTIGPKLSSKLVGRAQGFYASASQELGLLMVMNFAFEGKYNGST
ArDIR9      69  IVSGRNPASAVMVP-PA-NT-STGFGLVSMIDPLTIGPKLSSKLVGRAQGFYASASQELGLLMVMNFAFEGKYNGST
ArDIR10     59  IVSGRNPASAVMVP-PV-NT-STGFGLVSMIDPLTIGENLSSKLVGRAQGFYASASQELGLLMVMNFAFEGKYNGST
ArDIR11     59  IVSGRNPASAVMVP-PV-NT-STGFGLVSMIDPLTIGENLSSKLVGRAQGFYASASQELGLLMVMNFAFEGKYNGST
ArDIR12     55  IVSGRNPFAVQVVP-PP-NTSKTGFGLAMFDPLTIGPELSSKTIQIAQGLYISASQEEFLLMVMNFAFEGKYNGST
ArDIR13     61  IVSGRNPFAVQVVP-PV-NTSASATSFGANLIDPLTIGPKLSSKLVGRAQGFYALTSQEEVGLLMVMNFAFEGKYNGST
ArDIR15     56  IVSGRNPFAVQVVP-PP-NTSKTGFGLVSMIDPLTIGPELSSKTIQIAQGLYISASQEEFLLMVMNFAFEGKYNGST
ArDIR21     75  IVSGRNPASAVLVVP-PA-NT-STGFGLVSMIDPLTIGPKLSSKLVGRAQGFYASASQELGLLMVMNFAFEGKYNGST

PsPTS1      130  LSLGRNMMSETIREMAIVGGIGAFREVRGYAIAKFFVDFITGDAIVEYDLEVFHY
ArDIR4      134  LSVLGRNSV-FSAVREMPVIGGSGLFRFARGYAARHTHFDIKTGDATIVEYNVYAFHY
ArDIR5      136  LSVLGRNVI-LSAVRELPVIGGSGLFRFARGYAARHTHFDIKTGDATIVEYNVYAFHY
ArDIR6      136  LSVLGRNVI-LSAVRELPVIGGSGLFRFARGYAARHTHFDIGSGDAIVEYNVYVFHY
ArDIR7      136  LSVLGRNVI-LSAVRELPVIGGSGLFRFARGYAARHTHLDIETGDATIVEYNVYVFHY
ArDIR8      138  ISVLGRNAV-FSAVREMPVIGGSGLFRFARGYVCASTHFDIKTGDATIVEYNVYVLIHY
ArDIR9      147  ISVLGRNAV-FSAVREMPVIGGSGLFRFARGYVCASTHLDIKTGDATIVEYNVYVLIHY
ArDIR10     137  ISVLGRNAV-FSAVREMPVIGGSGLFRFARGYVCASTHFDIKTGDATIVEYNVYVLIHY
ArDIR11     137  ISVLGRNAV-FSAVREMPVIGGSGLFRFARGYVCASTHFDIKTGDATIVEYNVYVLIHY
ArDIR12     134  LSLGRNPGATSKVREFPVIGGSGLFRFARGYALATHELNMTGDATIVEYNYVLIHY
ArDIR13     140  ISVLGRNVI-FSKVREMPVIGGSGLFRFARGYVLSLTKFNSTGDATIVEYNVYVLIHY
ArDIR15     135  LSLGRNPGATSRVREFPVIGGSGLFRFARGYALATHELNMTGDATIVEYNVYVFHY
ArDIR21     153  ISVLGRNAV-FSAVREMPVIGGSGLFRFARGYVCASTHLDIKTGDATIVEYNVYVLIHY

```

Supplementary Figure S43. Protein sequence alignment of *bona fide* dirigent proteins from the Dir-b sub-family, PsPTS1 from *Pisum sativum* (40) and that of putative homologs (ArDIR4 – 13, 15 and 21) in red alder.

```

GhDIR4      1  -----MRGTSVLSW-----IILLCVCQA-----AVRSQYSDTLPYQPRPV
ArDIR14    1  ---MHSEERRAR-----EFSKELSIMKLMRVNFVS----FCALLLCMVSIALL-----RQYCTMPLLPYHMLE
ArDIR16    1  -----M--HMAKVAP-----VLLLFSIAMAIQLEKQSTVAEPNEEVDLW--Q-NLQTSKQ
ArDIR17    1  -----MACFVTFITLYSFLIFS-AS-----LNTISG--AFVVEEFS-EA-MAIKREE
ArDIR20    1  -----M--QMAKLIQV----VLMFLFSLVAAIPWAQS-----SEESWATSQ-EFQKKET
ArDIR22    1  ---MHQPGYSTNRAESTLLETIPAE--LMDSFVTF----LFFSASL-----INTIDG--AFIEKFS-EP-RAVKREQ
ArDIR25    1  -----MASFLSLFY-FCLILF-STG-----STTVFG--VFSEESL-IT-MYAKRME
ArDIR27    1  MFLMHSEERRAR-----EFSKELSIMKLMRVNFVS----FCALLLCMVTIALL-----RQYCTMPLLPYHMLE

GhDIR4      37  LVTNLHFYMHFTGTT--AVVLTQAN-ITSNNSSVPFATLVAVNDPLRTGPEEDSELTGNVQGISLLAGSNA-SSTQYI
ArDIR14    56  KETVLHFFLHDTPGSSPSAVRVVQANGTVHNTSLTPFGTVVWVDDPLTEGPEDEKSTILGNSRCVYVSAGQDELMLVAYL
ArDIR16    47  TISRLNFYFHDITLGGKQPSAVQVAQASNT--DKSPTLFGAVYIFDDPLTEGPEETSRLVGRAQGLYGSACQQLSLLVAM
ArDIR17    42  KTHLHFFHFDITLGGKQPSAVRIAGPT---ESTFVNFNTMMIDDLTEGPEETSRLVGRAQGLYGSACQQLSLLVAM
ArDIR20    43  VITNLNFYFHDIVCGKNPFAVRVAQAANT--NKSPTSFCVVMADDPLTKTPDFNSLLVGRAQGLYGSACQQLSLLVAM
ArDIR22    60  --SHLHFFHFDILGCEHPTAVFIAGPP---KSTILDGNTMMIDDLTEGPEETSRLVGRAQGLYSRADQNDVALLMVV
ArDIR25    41  KTRRLHFFHFDVLDGNTPTAIFIVTPQ---NLSVGGFCATYIMDDLTEGLEPESKPVGRAQGLYSRADQNDVALLMVI
ArDIR27    62  KETVLHFFLHDTPGSSPSAVRVVQANGTAHNTSLTPFGTVVWVDDPLTEGPEDEKSTILGNSRCVYVSAGQDELMLVAYL

GhDIR4      113  EFCENTGKFNNGSSLSVFSRCD-----PGLAVVGGRGQFAMATGHALFNPILINA--TNVIMEFNFTVIHY-----
ArDIR14    137  DFCFTSGEFNGSSLSVLSRNTVTVQIEREFVAVGGRGKFRMARGFANLKCIFFNESNGDSIVEYTMVVRHY-----
ArDIR16    126  TVVFTAGFNASSLTILGRNAALQEVREMPITIGGSGVFRLLARGEAIKTHSLNFTSGDAIVEYNVAVHY-----
ArDIR17    119  NFAFVEDKYNGSSISMLGRNPVLDVREMPIVGGSGLFRNRCGMALAHNTNMDPNTGDAIVEYDVYVSNFQVDAG-LATA
ArDIR20    122  NFGFVDGIYNGSSISLLGKNPASHPVREMAIVGGTGIFFRLARGYAIATHWFNATSGDAIVRYNVTVVH-----
ArDIR22    135  NFAFVEDKYKNGSSISMLGRNPVANDVREMPIVGGSGLFRNRCGMALAHNTNSMD-NAGDAIVEYNVYVSNVEVDASLATQ
ArDIR25    118  NFAFLEGIYNGSSLSILGRNPVFHAVREMPIVGGSGVFRLLARGYALAKTVRFNKKSGDAIVEYNVSVLHF-----
ArDIR27    143  DFCFTSGEFNGSSLSVLSRNTVTVQIEREFVAVGGRGKFRMARGFANLKCIFFNESNGDSIVEYTMVVRHY-----

```

Supplementary Figure S44. Protein sequence alignment of *bona fide* dirigent proteins from the Dir-d sub-family, GhDIR4 from *Gossypium hirsutum* (41) and that of putative homologs (ArDIR14, 16, 17, 20, 22, 25 and 27) in red alder.

```

AtDIR10 1 -----MAGQKILSLLVIALVVTFAAAARLLDEENAFSATTTTLGSGSGSTGIGFGAGTGSSGSGSTGFVGFAGSGSS
ArDIR18 1 -----
ArDIR24 1 -----
ArDIR19 1 -----
ArDIR23 1 MAKLAITANTLKGFCLVLLAITIGCANSARILDEVDPQPPILPYPPQSTNPTATPV--
ArDIR26 1 -----

AtDIR10 73 GSGSTGSGLGAGTGSIPSSGSGPGLPTASSVPGSLAGGSGSLPTTGSATGAGAGTGSAL-----GGGPGAGS
ArDIR18 1 -----MHRHRT-----FPISSMAA-----SEPSNIF
ArDIR24 1 -----
ArDIR19 1 -----
ArDIR23 58 -PDVTPQVVGPPTA-----ITP-----SQQIPASTVADDVDVDVDVAPVASPDVEPDVDSFQPEPEVE
ArDIR26 1 -----

AtDIR10 142 AIGGAGAGPALGGAGAGPALGGAGAGSALGGGGAACAPALGGG--AGAGFALGGVA---GSGSALGGGASAGPDN
ArDIR18 22 PIK-----NIL-----RLIL-----LALTTICTASARILDEETTPTTTEPVSS-----VFPVVTAAGR
ArDIR24 1 -----VSKGL-----SSASFFLTI-----IAVVHVGMIMAAVDVAATGKEP
ArDIR19 3 THK-----ATA-----NLIF-----LTLTILSATASARILDGQPAVAVVPEEATNGV-----DTPVTNADPSH
ArDIR23 114 PVDSPQPEPEEP-----D-----TVPTFVTTVTPVVGPF--TIP---APLGFVAPVTTAATVAKPGGPQAS
ArDIR26 1 -----MA-----KLPSISMLTITACALLFMFIHQ-----ASSRTIIN-PTPTPRKHYH

AtDIR10 217 TLVFFMHDLGG-SNPTARAVTGVVNPALSGQLPFAKPNGANLPSVNGVPSN---NNNNGVNNNNPFLVGLGGTIA-
ArDIR18 71 PMSFFMHDLGG-SNPSANAVSGVSNPAMNGQVPPFAKPNGAVLPVNMNGVPSN-NNNNNGVNNNNIPFLTGLSGNIP-
ArDIR24 39 ILELYMHDLGG-NSPTAREPITGLLGN-LYNGQVPAFAPIGCFVPPQG-GVAIPNANGAIPTVNGVTGIFLGTGLGAGN-
ArDIR19 55 TLVFFLHDLGG-SNPTAREAVTGVVNNPALNGQLPFAKPNGAVLPVSNNGNGLT-DN-----NNNNNIPFLTGLSGATA-
ArDIR23 173 GLVFFMHDLGG-IHPSARVVTGLIANTQING-LPFTTPNNDLFPVSGVPISN--ANLNGLNNNNPFLAGLNGVGFQ
ArDIR26 45 KITFLMRDVLINVTQPSSEIPATEK---VTSTQLPFPKELGLFPNG-CHPVL--GSNPTVPG-----TAAGLSTQIL-

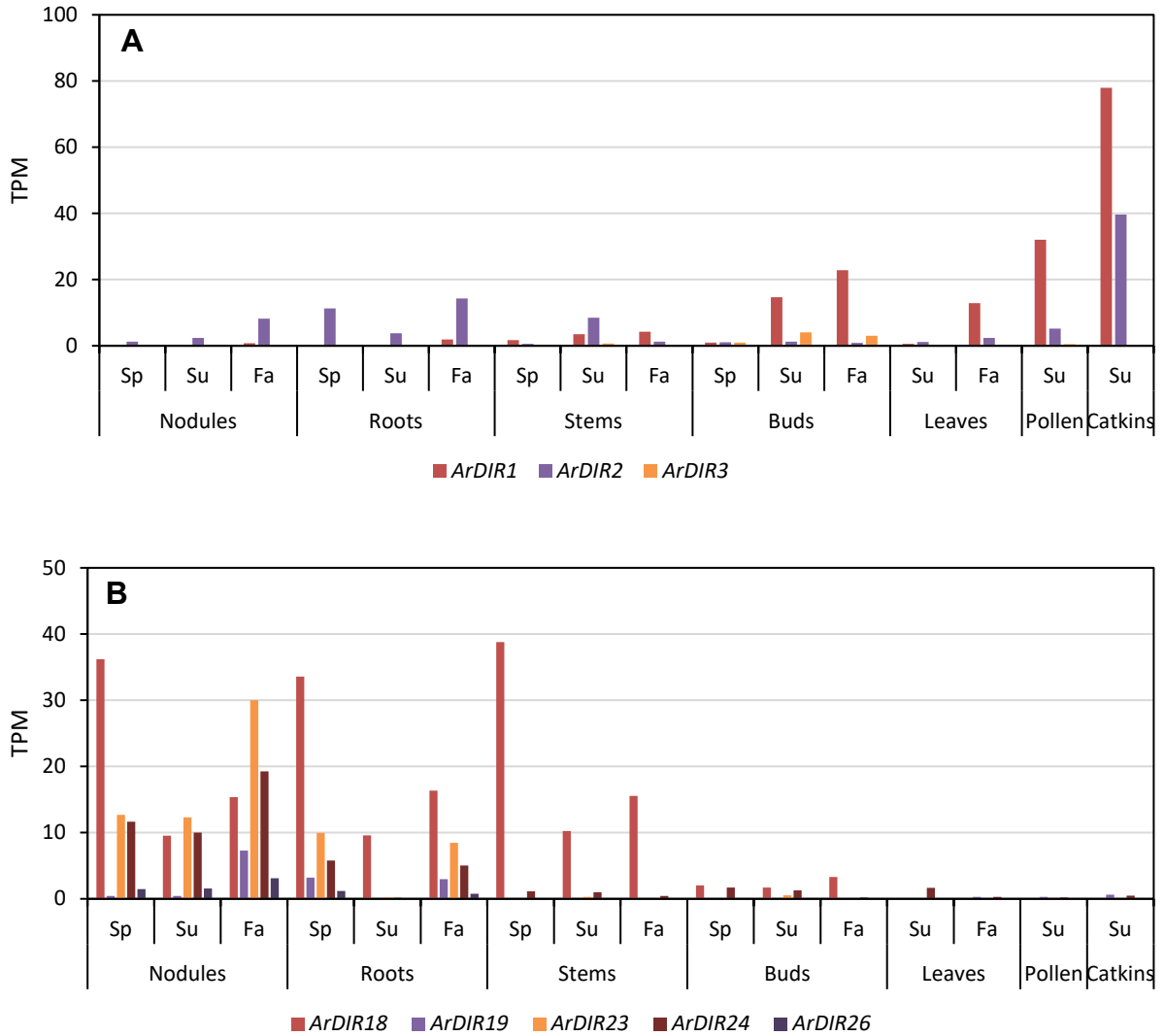
AtDIR10 292 --NLLQNNNGNNI--LNGFPVASGGQLPSGSALQMLMFGTMTVIDDELTEGHEL--GSGLVGKAQGYVASAIDGTSQT
ArDIR18 148 --NVLQNNNGNC-----NN-GENGAQLBAGMTLQKLMFGTLSVFDNDELTEGHEL--GSGLVGKAQGFYVASSDGTSTQ
ArDIR24 115 -----A-----NNQNPQQLGPDGLGLCFGTITVIDDLTSGDEL--GSGLVGKAQGFYVASSADGITQ
ArDIR19 127 --AVMNNNGGFC-----NVFVNGGQLPSGNTLQKLMFGTMTVIDDELTEGHEL--RSGLVGKAQGFYVSSDGTGQT
ArDIR23 249 ASTVQLQNSGNGIVNNGNLPFVTAAGQLPSGATLQKLMFGSTVIDDELTEGHVL--GSAVVGKAQGFYVASSDGTSTQ
ArDIR26 112 -----DLSATGLSFEASATLQELEFGTMTVIDDELTEGHELFKSSKNMSPMLGKAQGFYVASSDGTSSHM

AtDIR10 366 MAFTAMFESGG---YEDSLSFFGVLRRTAVSESHIGVMGGTGKYNARCFALIKFTFGSGTQQNPHQFTDGLTETVVE
ArDIR18 217 MAFTTMEFQSGG---YADSLSFFGVHRTAASESOLAVMGGTGKYVDKGYATVKTFPAIN---QQ--NNGVETLLQ
ArDIR24 174 MAFTAVIEGGE---YGDSLNFCTYKICSTLSYLSVIGGTGKKNASCFAEVRSLIP-----PGQHVLDCAETLLRI
ArDIR19 197 MAFTVMFESGR---YADSLSFFGTHRTAVSESOIAMGGTGKYVNAKGFATVKTYAEIN---NQQHETDGLTLLQ
ArDIR23 327 MAVTVLHGCQHEVVEDLISFFGVHRTASLESQAVVGGTGKYENAKGFATVETLHS-----EDQHTDGDVDTLQ
ArDIR26 173 MAMTAYFAGSD---GFKDGLRFFGVHQTDPVPSHIAVGGTGKYQGANCYAKVKAVDHVGSHSD---EEGKTNKLLLF

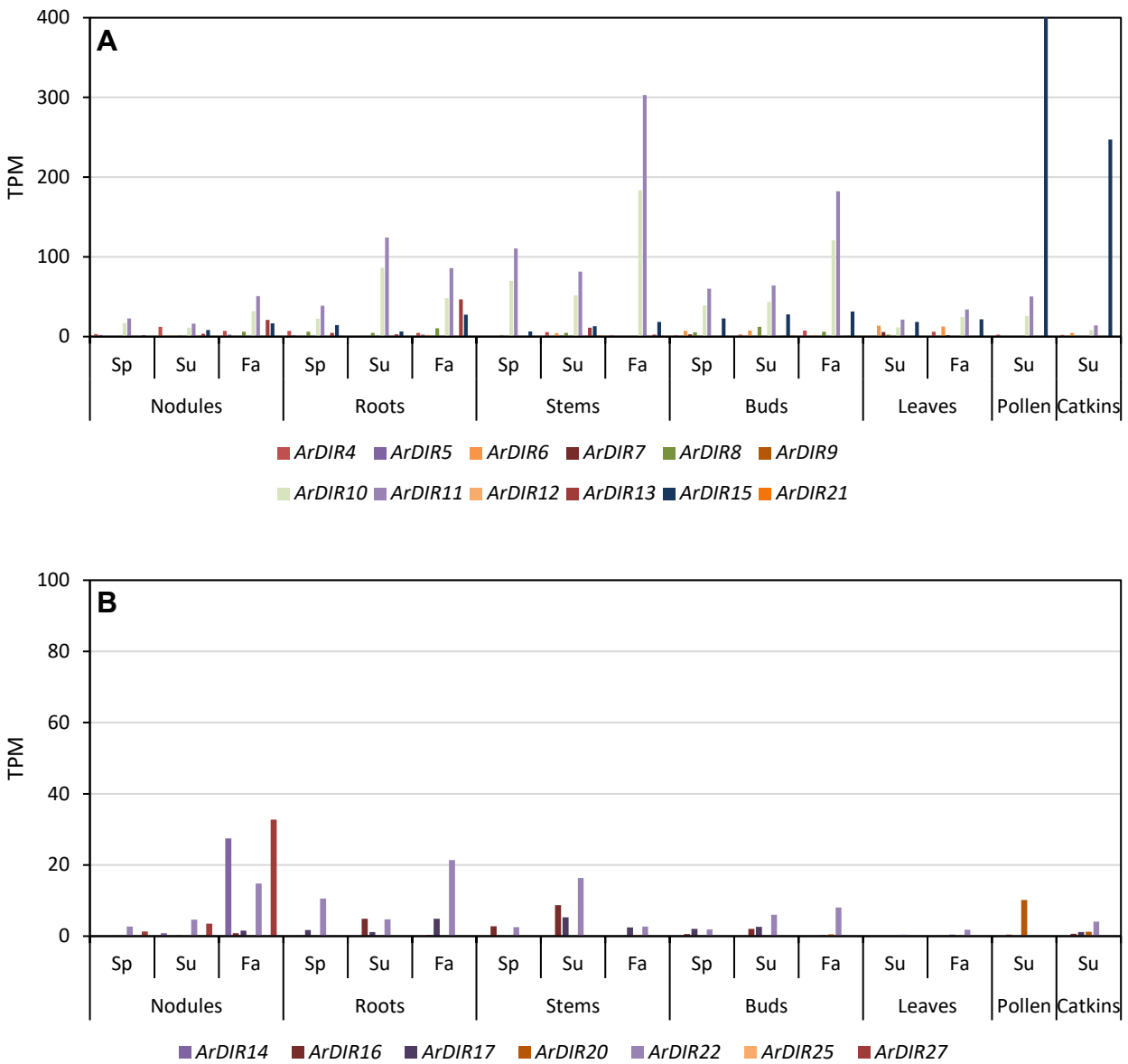
AtDIR10 442 TVYLSY
ArDIR18 286 TVYLLY
ArDIR24 243 TVHLLY
ArDIR19 268 TVYVLY
ArDIR23 400 SVYLSE
ArDIR26 247 HVYLS-

```

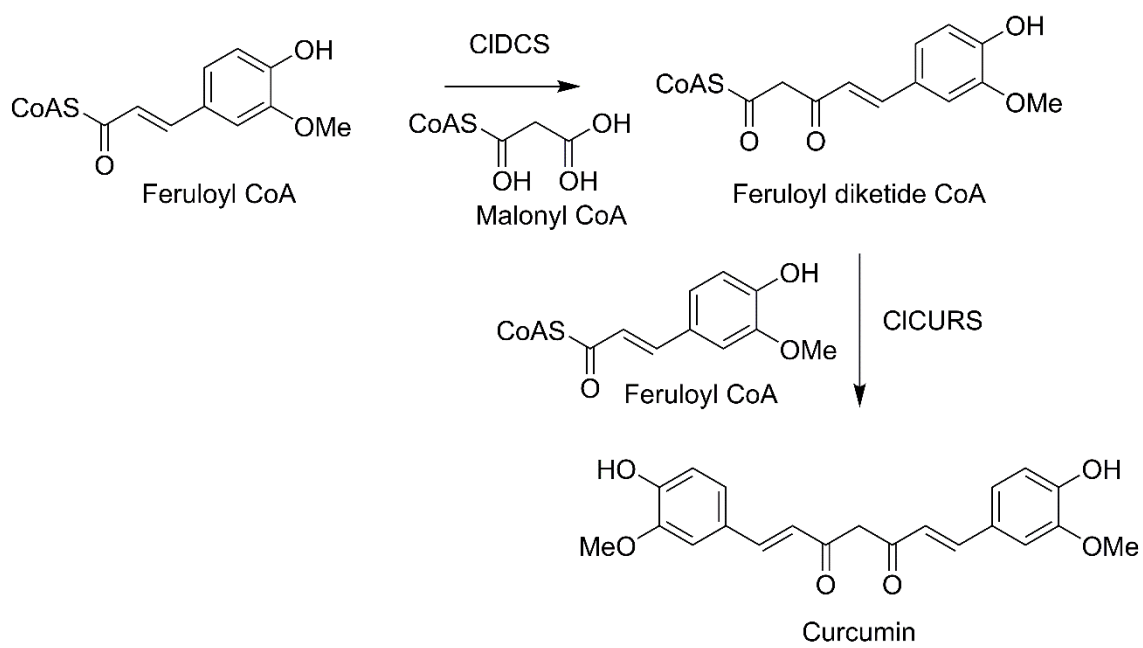
Supplementary Figure S45. Protein sequence alignment of *bona fide* dirigent proteins from Dir-e sub-family, AtDIR10 from *A. thaliana* (42) and that of putative homologs (ArDIR18, 19, 23, 24 and 26) in red alder.



Supplementary Figure S46. Expression levels of red alder homologs of *bona fide* dirigent protein genes in the DIR-a (**A**) and DIR-e (**B**) sub-families. Abbreviations: TPM, transcripts per million; Sp, spring; Su, summer; Fa, fall.



Supplementary Figure S47. Expression levels of red alder homologs of *bona fide* dirigent protein genes in the DIR-b (**A**) and Dir-d (**B**) sub-families. Abbreviations: TPM, transcripts per million; Sp, spring; Su, summer; Fa, fall.



Supplementary Figure S48. Pathway to curcumin in *Curcuma longa* (Cl). CIDCS, diketide-CoA synthase; CICURS, curcumin synthase (61).

```

FiPLR1 1 --MCKSKVLLIIGGTCYLGKRLVKASLAQGHEITYILHRPEI-GVDIDKVEMLISFKMQGAHLVSGSFKDENSILVDAVKLVD
ArPLR1 1 --MEKSRVLLVGGTCYLGKRLVKASLDQGHETYVLQRPEI-GVDIEKVQMLLSFKKQGAHLVSGSFTDHDQISIVNAVKLVD
ArAPS1 1 -MAHKSKILFIGGTCYIGKFIVEASAKSGHPTFALVREST-VSDPVKGGKLVHFKSLGVTLHGDLYDHASLVKAIKHVD
ArAPS2 1 -MAHKSKILFIGGTCYIGKFIVEASAKSGHPTFALVREST-VSDPVKGGKLVHFKSLGVTLHGDLYDHASLVKAIKHVD
ArAPS3 1 -MAOKSRILLIIGGTCYIGKFIVASAKAGHPTFALIREST-RSDPVKGGKLVENFNKLVGVTLHGDLDHDSGLVKAIKQVD
ArAPS4 1 -MAOKSKILLIIGSTCFYIGKFIVEASAKSGHPTFALVREST-VSDPIKGGKLVQHFKSLGVTLHGDLYDHASLVKAIKQVD
ArAPS5 1 -MAOKSKILLIIGSTCFYIGKFIVEASAKSGHPTFALVREST-VSDPIKGGKLVQHFKSLGVTLHGDLYDHASLVKAIKQVD
ArPPS1 1 MANEKSILIFGGTCYIGNYLARASCLLGHETYVYARPLTPQTTPSKLELHKQLHSMGATFVLGELEEEHKIVSAVREVD
ArPPS2 1 MENEKSILIFGGTCYIGNYLAKASCLLGHETYVYARPLTPQTTPSKLELHKELHSMGATFVLGELEEEHKIVSALREVD

FiPLR1 78 VVISAIISGVHIRSHQILLQIKLVEAIKEAGNVKRFPSDFGMDPAKFMOTAMEPGKVTLDKEMVVRKATIEKAGIPFTYVS
ArPLR1 78 VVIQAIISGVHIRSHQILLQIKLVDAIKEAGNVKRFPSDFGTDPAERMPNAILPGRVTFDDKMWVRKATQDAKIPFTYVS
ArAPS1 79 VVISIVGHLQIAD----QVKIIAAIKEAGNVKRFPSDFGNDVDR--VHAVEPAKTAFAFATKVEIRRKIEAEAGIPYTVVS
ArAPS2 79 VVISIVGHLQIAD----QVKIIAAIKEAGNVKRFPSDFGNDVDR--VHAVEPAKTAFAFATKVEIRRKIEAEAGIPYTVVS
ArAPS3 79 VVISIVGRLQIFGD----QVKIIAAIKEAGNVKRFPSDFGNDVDR--VHAIGPAKTAFAFATKAKIRRTIEAEAGIPYTVVT
ArAPS4 79 VVISAVGHPOVED----QDKIIAAIKEAGNVKRFPSDFGNDVDR--VHAVGPAQAAFDRVETIRRKIETAGIPYTVVA
ArAPS5 79 VVISAVGHPOVED----QDKIIAAIKEAGNVKRFPSDFGNDVDR--VHAVGPAKAAFDRVETIRRKIETAGIPYTVVA
ArPPS1 81 IVISALAYPOVLD----QVKIIDAIKVGAGNIKRFPSDFGVEEDR--VTPLPPEAFLDKRRKIRRAAEAGIPYTFVS
ArPPS2 81 IVISALAYPOVLE----QVKIIDAIKVGAGNIKRFPSDFGVEEDR--VTPLPPEAFLDKRRKIRRVVEAGIPYTFVS

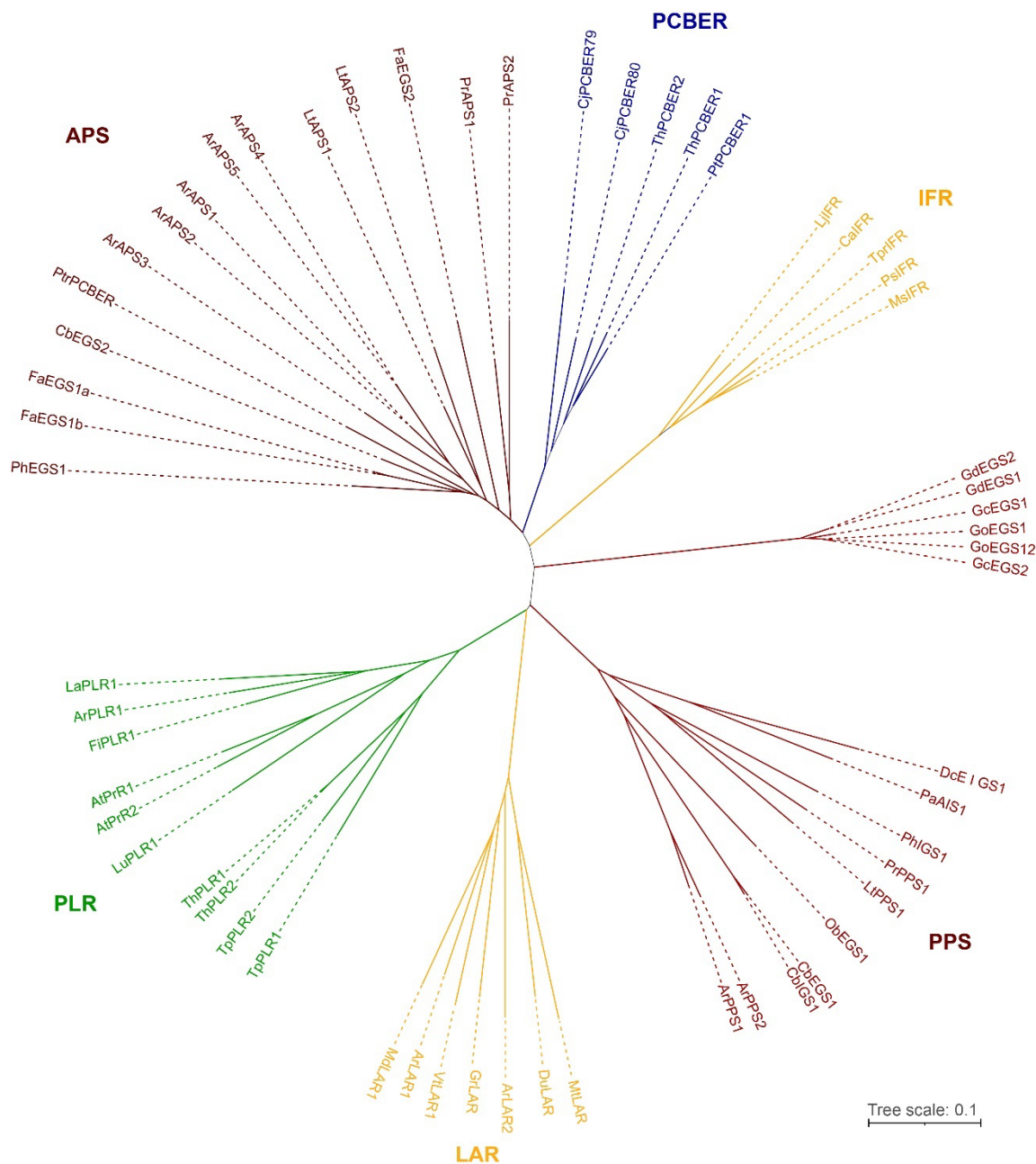
FiPLR1 158 ANCFAGYFLGGLCOQFCKI-----LPSRDFVIHGDGNKKAALYNNEDDIATYAIKKTINDPRTLNKTIYISPPKNILSQ
ArPLR1 157 ANCFAGYFLGGLCOPGYI-----LPSRDSVVLFGDGNTKALYVDEDDIAMYTIKTIIDDPRTLNKTVYIRPPKNILSQ
ArAPS1 152 SNYFAGYFLPTLAQPGILT-----SPPRDKITIFGDGHVKALFNKEDDIGTYTIRTVDDPRTLNKIVYLRPPKNIYSF
ArAPS2 152 SNYFAGYFLPTLAQPGILT-----SPPRDKITIFGDGHVKALFNKEDDIGTYTIRTVDDPRTLNKIVYLRPPKNIYSF
ArAPS3 152 SNYFAGYFLPTLAQPGAR-----APSRDKVIFGDGNAKALFNKEDDIGSYTIRAVDDPRTLNKILYLRIPKNIYSL
ArAPS4 152 SNYFAGYFLPTLAQPGFTAPPIEKVTPPREKVTIFGDGNAKALFNKEDDIGTYTIRAVDDPRTLNKIVYLRPPKNIYSF
ArAPS5 152 SNYFAGYFLPTLAQPGFTAPPIEKVTPPREKVTIFGDGNAKALFNKEDDIGTYTIRAVDDPRTLNKIVYLRPPKNIYSF
ArPPS1 154 ANCFASYFLSHLLHPHEK-----GDDIAVYGSGEAKAVLNYEEDIATYTIKVANDPRTGNRIVYIRPPKNILSQ
ArPPS2 154 ANCFASYFLNYLLHPHEK-----GDDIAVYGSGEAKAVLNYEEDIATYTIKVANDPRTGNRVVIYIRPPKNILSQ

FiPLR1 230 REVVQIWEKLIQKELQKITSKEDFLASVKELEYAQVGLSHMHDVNYQCCLTSFEIG--DEEEASKLYPEVKYTSVEEY
ArPLR1 229 REVVEIWEKLIQELHKKSSISKEDFLASIKGLEAYEQVGLIHYHVCYEGCLTNFEIGK-DAGEASELYPEIDYTSVHDY
ArAPS1 224 NELVALWEKKIGKTLKTYVSEEKILKDIQEAIPINVLAINHSVFKGDHTNFEIEASFGVEASELYPDVRYTTVEEY
ArAPS2 224 NELVALWEKKIGKTLKTYVSEEKILKDIQEAIPINVLAINHSVFKGDHTNFEIEASFGVEASELYPDVRYTTVEEY
ArAPS4 224 NELVALWEKKIGKTLKTIHVSEENILKDKISESEFPVNAILSIHHSVFKGDHETSFEIEASFGVEASELYPDFNYTTVDEY
ArAPS5 232 NELVALWEKKIGKTLKTIYVPEEKLLKDIQESPIPTNLILAIYHSVLRKGDHTNFEIEASFGVEASELYPDVRYTTVEEY
ArPLR6 232 NELVALWEKKIGKTLKTIYVPEEKLLKDIQESPIPTNLILAIYHSVLRKGDHTNFEIEASFGVEASELYPDVRYTTVEEY
ArPPS1 224 LELISLWEKKTGRSFKVYVSEEEIVKLSSETLPTAEALPISILHSIFIKG-QMVFELRE-DDLEASRLYPELEYTTVDKL
ArPPS2 224 LELISLWEKKTGRSFRNIHVPEEEIVKLSSETLPPKEDIPMSILHSIFIKG-QTVFELRE-DDMEASRLYPDFKYTTIDQL

FiPLR1 308 LKRVV-----
ArPLR1 308 MKRYL-----
ArAPS1 304 LQQFV-----
ArAPS2 304 LQQFV-----
ArAPS3 304 LDPFVGDVMEKGLSVPALFGYCNEP
ArAPS4 312 LQQFV-----
ArAPS5 312 LQQFV-----
ArPPS1 302 LDIFLVNPCKPTMGSE-----
ArPPS2 302 LDIFLVNPCKFAAAAFE-----

```

Supplementary Figure S49. Protein sequence alignment of *bona fide* pinoresinol-lariciresinol reductase from *Forsythia intermedia* (FiPLR1 (43)) and that a putative homolog (ArPLR1), as well as that of allylphenol and propenylphenol synthases (ArAPS and ArPPS) in red alder.



Supplementary Figure S50. Unrooted phylogenetic tree of known pinoresinol-lariciresinol reductases (PLRs), phenylcoumaran benzylic ether reductases (PCBERs), allylphenol and propenylphenol synthases (APS, PPS), and isoflavone reductases (IFRs), as well as leucoanthocyanidin reductases (LARs) from *Alnus rubra* (Ar), *Arabidopsis thaliana* (At), *Cicer arietinum* (Ca), *Clarkia breweri* (Cb), *Cryptomeria japonica* (Cj), *Daucus carota* (Dc), *Desmodium uncinatum* (Du), *Forsythia intermedia* (Fi), *Fragaria × ananassa* (Fa), *Gossypium raimondii* (Gr), *Gymnadenia conopsea* (Gc), *G. densiflora* (Gd), *G. odoratissima* (Go), *Larrea tridentata* (Lt), *Linum album* (La), *Linum corymbulosum* (Lc), *Linum usitatissimum* (Lu), *Lotus japonicus* (Lj), *Malus domestica* (Md), *Medicago sativa* (Ms), *Medicago truncatula* (Mt), *Ocimum basilicum* (Ob), *Petunia hybrida* (Ph), *Pimpinella anisum* (Pa), *Pinus taeda* (Pt), *Piper regnellii* (Pr), *Pisum sativum* (Ps), *Populus trichocarpa* (Ptri), *Thuja plicata* (Tp), *Trifolium pratense* (Tpr), *Tsuga heterophylla* (Th), and *Vitis vinifera* (Vt). Sequences were

obtained from NCBI and a multiple sequence alignment was built with Clustal Omega (57). The tree was rendered with iTOL (58).

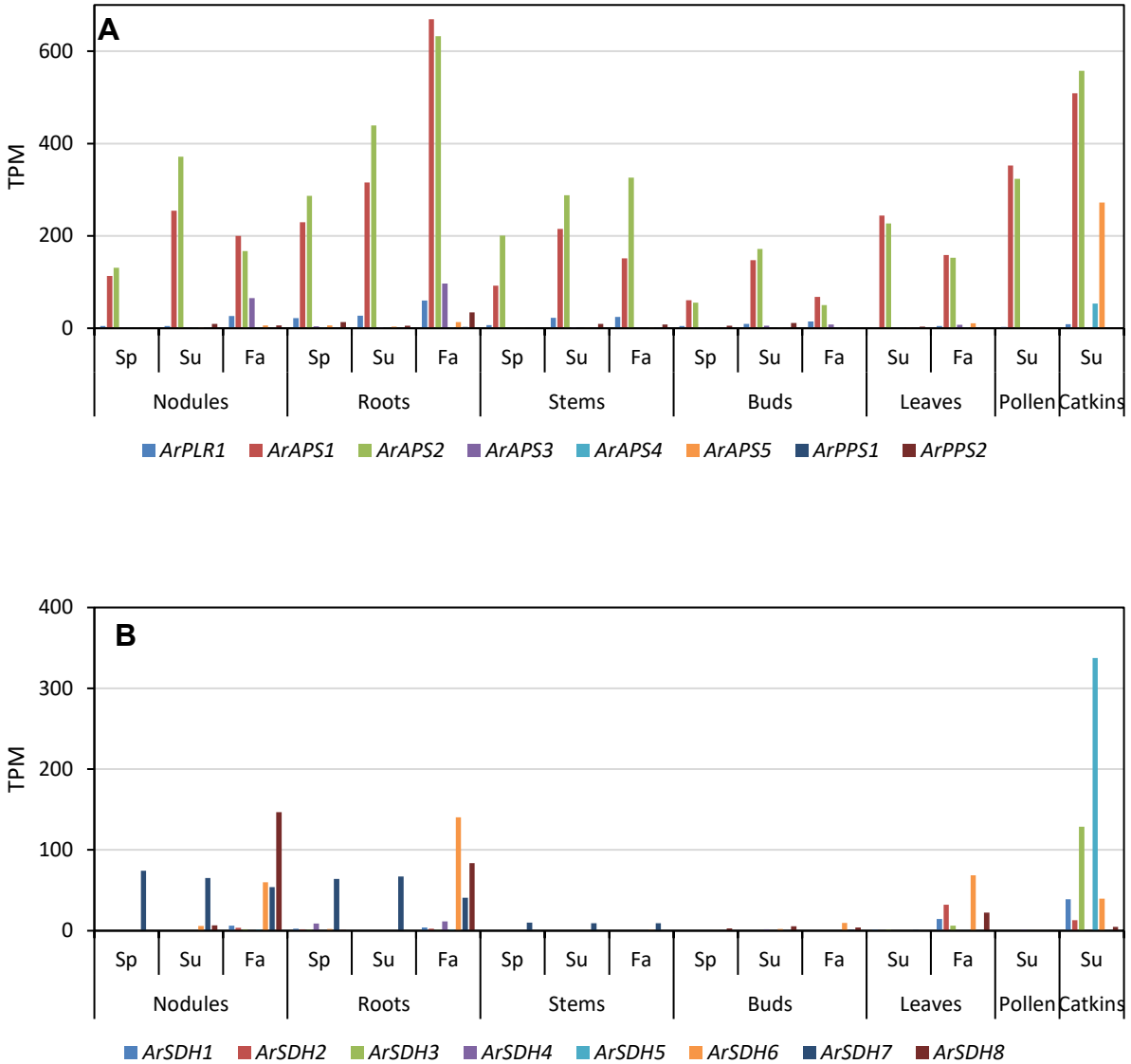
PLRs: AtPrR1 (Q9FVQ6), AtPrR2 (Q9SVP6), FiPLR (AAC49608), LaPLR1 (Q4R0I0), LuPLR1 (CAH60858), TpPLR2 (Q9LD13), TpPLR1 (Q9LD14), ThPL1R (AAF64184), and ThPLR2 (AAF64185).

PCBERs: CjPCBER79, CjPCBER80, LcPCBER1 (ACA60729), LcPCBER2 (ACA60730), PtPCBER (AAC32591), PtrPCBER (CAA06706), ThPCBER1 (AAF64176), and ThPCBER2 (AAF64177).

APs and PPs: CbIGS1 (ABR24112), CbEGS1 (ABR24113), CbEGS2 (ABR24114), *DcE(I)GS1* (XP_017241251), FaEGS1a (AGV02006), FaEGS1b (AGV02007), FaEGS2 (AGV02008), GcEGS1 (AKB11747), GcEGS2 (AKB11748); GdEGS1 (AKB11749), GdEGS2 (AKB11750), GoEGS1 (AKB11751), GoEGS2 (AKB11752), LtAPS1 (AHA90804), LtAPS2 (AHA90805), LtPPS1 (AHA90806), ObEGS1 (ABD17321), PhIGS1 (ABD17322), PhEGS1 (ABR24115), PaAIS1 (ACL13526), PrAPS1 (AHA90807), PrAPS2 (AHA90808), and PrPPS1 (AHA90809).

IFRs: CaIFR (Q00016), LjIFR (QBF58800), MsIFR (AAC48976), PsIFR (P52576), and TprIFR (QBF58799).

LARs: DuLAR (Q84V83), GrLAR (CAI56324), MdLAR1 (AAZ79364), MtLAR (CAI56327), and VtLAR1 (CAI26309).



Supplementary Figure S51. Expression levels of red alder homologs of (A) *bona fide* pinoresinol lariciresinol reductase (*ArPLR*), allylphenol synthase (*ArAPS*) and propenylphenol synthase (*ArPPS*) genes, and (B) secoisolariciresinol dehydrogenase (*ArSDH*) genes. Abbreviations: TPM, transcripts per million; Sp, spring; Su, summer; Fa, fall.

```

PpSDH      1  -----MGSSTPDSSTNRLODKVAITGGAGGIGETTA L FVRYGAKV IADIA
ArSDH1     1  HSLFYKRPSPADLHIEKHSIILTNNKLMEGASLPPAPIFKRLEGKVALITGGASGIGETTARLFVQHGAKEIADVQ
ArSDH2     1  -----MEGALPPAPIFKRLEGKVALITGGASGIGETTARLFVQHGAKEIADVQ
ArSDH3     1  HSLFYKRPSPADLHIEKHSIILTNNKLMEGASLPPAPIFKRLEGKVALITGGASGIGETTARLFVQHGAKEIADVQ
ArSDH4     1  -----MEGASLPPAPIFKRLEGKVALITGGASGIGETTARLFVQHGAKEIADVQ
ArSDH5     1  -----MKGASLPPAPIFKRLEGKVALITGGASGIGETTARLFVQHGAKEIADVQ
ArSDH6     1  -----MGSFSLLSAAARLEGKVALITGGASGIGESTVRLFSKHGAKVIADIQ
ArSDH7     1  -----MATSLSLAVTRRLEGKVALITGGASGIGECTAKVFAFHGAKVIADIQ
ArSDH8     1  -----MGSVSLVSAARLEGKVALITGGASGIGESTARLFVSKHGAKVIADIQ

PpSDH      50  DDHCGQKVCNNIGSPDVISVHCDVTKEDVRRNLVDTAKHGKLDIMFCNVGLSTTPYSILEAGNEDEKRVMDINVLG
ArSDH1     81  DDIGHSVCDIGTADDISYVHCDVTNESDVQNAVKTAVSKHGKLDILFSNAGITNADDGILEEFDSEKNRNVFNVLG
ArSDH2     51  DDIGHSVCDIGTADDISYVHCDVTNESDVQNAVKTAVSKHGKLDILFSNAGITNADDGILEEFDSEKNRNVFNVLG
ArSDH3     81  DDIGHSVCDIGTADDISYVHCDVTNESDVQNAVKTAVSKHGKLDILFSNAGITNADDGILEEFDREKNRNVFNVLG
ArSDH4     50  DELGHSVCDIGTNTDDISYVHCDVTNESDVQNAVETAVSKHGKLDILFSNAGITNADNSILDFLEKNRNVFNVLG
ArSDH5     50  DDLGHSICSDISTGDDMISYVHCDVTNESDVQNAVNIYVSKHAKLDILFSNAGITNADGILEEFDREKSRNVFNVLG
ArSDH6     50  DDLGHSVCDLNSQ--STSVHCDVTKEDVNAVNLA VSKFGKLDIMFNAGIAGVAKTDIREITKABFEQVIGINLVG
ArSDH7     50  DELGHAVNEALGPS--NSLYVHCDVTDEAHKNAVEKAVDTYKGLDIMFNAGIADDNKARITDNEKSDFERVLSVNVG
ArSDH8     50  DDLGHSVCDLNSQ--STSVHCDVTKEDVNAVNLA VSKFGKLDIMFNAGIAGVAKTNILEITKABFEQVIGINLVG

PpSDH      129  AFLVAKHAAVMIPAKKGSI VFTASISSFTAGEGVSHVYTAIKHAVI GLTSLCTELGEYGIRVNCVSPYIVA PLLITDV
ArSDH1     159  GFLAAKHAAKVMIPAKKGSI LFTSSTASVIHGE TPHIYAASKHALVGLTKNLCVELGEYGIRVNCVSPSGVATPLLRNV
ArSDH2     129  GFLAAKHAAKVMIPAKKGSI LFTSSTASVIHGE TPHIYAASKHALVGLTKNLCVELGEYGIRVNCVSPSGVATPLLRNV
ArSDH3     159  GFLAAKHAAKVMIPAKKGSI LFTSSTASVIHGE TPHIYAASKHALVGLTKNLCVELGEYGIRVNCVSPSGVATPLLRNV
ArSDH4     129  GFLAAKHAAKVMIPAKKGSI LFTSSTASVIHGE TPHIYASKHAVVGLTKNLCVELGEYGIRVNCISPSGVATPLLRNV
ArSDH5     129  GFLAAKYAKVMIPAKKGSI LFTSSIASVIHGT TPHIYASKHALVGLSKNLCVELGEYGIRVNCISPSGVATPLLRNV
ArSDH6     128  VFLGTHKHAAVMIPAHHGSITTTASVCSITVGGTASHGYTSSKHVVGLMNTAVELGQEGIRVNCVSPYLVLTPLAKDF
ArSDH7     128  VFLGTHKHAAVMIPAHHGSITTTASVCSITVGGTASHGYTSSKHVVGLMNTAVELGQEGIRVNCISPYLATPLAKDF
ArSDH8     128  VFLGTHKHAAVMIPACHGSIITTASAASTVGGTASHGYTSSKHVVGLMNTAVELGQEGIRVNCVSPYVVPPLAKDF

PpSDH      209  FGV DSSRVEELAHQAANLKGTLRAEDVAEAVL LAGDESKYVSGLNVLVDGGYTRTNPAFPTLKHGLA-
ArSDH1     238  SGIQDKKTLEDMICSAANLKGVVLETDVVAEAAFLVSDSKYVSGLNVLVDGGYSTTNKAIKDKKEKHS-
ArSDH2     208  SGIQDKKTLEDMICSAANLKGVVLETDVVAEAAFLVSDSKYVSGLNVLVDGGYSTTNKAIKDKKEKHS-
ArSDH3     238  SGIQDKKTLEDMICSAANLKGVVLETDVVAEAAFLVSDSKYVSGLNVLVDGGYSTTNKAIKDKKEKHS-
ArSDH4     208  CGIQDKKTLEDMICSAANLKGVVLETDVVAEAAFLVSDSKYVSGLNVLVDGGYSTTNKAIKDKKEKHS-
ArSDH5     208  SGIQDKKTLEDMICSAANLKGVVLETDVVAEAAFLVSDSKYVSGLNVLVDGGYSTTNKAIKDKKEKHS-
ArSDH6     207  LKIDDDGWHRA---YSNLGKGLAEADVAVL LAGDESKYVSGHNLVDGGFTIENSQFSMYA-----
ArSDH7     207  VGLDDEGLNINSLANLKGVTLEADVANAALMLASNESRYVSGHNLVDGGFTIENSQFQIFQYPPDS--
ArSDH8     207  YKIDNDGWHRA---YSNLGKGLAEADVAVL LAGDESKYVSGHNLVDGGFTIENSQFSMYA-----

```

Supplementary Figure S52. Protein sequence alignment of *bona fide* secoisolariciresinol dehydrogenase from *Podophyllum peltatum* (PpSDH1 (45)) and that of putative homologs (ArSDH1 – ArSDH8) in red alder.

Supplementary Material References

1. Sati, S. C., Sati, N., and Sati, O. P. (2011) Bioactive constituents and medicinal importance of genus *Alnus*. *Pharmacogn. Rev.* **5**, 174–183
2. Jackrel, S. L. (2016) Drivers and consequences of intraspecific trait variation: Ecology of a forest-stream community. *Ph.D. Thesis, The University of Chicago*
3. Jackrel, S. L., Morton, T. C., and Wootton, J. T. (2016) Intraspecific leaf chemistry drives locally accelerated ecosystem function in aquatic and terrestrial communities. *Ecology* **97**, 2125–2135
4. Keith, B., Dong, X., Ausubel, F. M., and Fink, G. R. (1991) Differential induction of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase genes in *Arabidopsis thaliana* by wounding and pathogenic attack. *Proc. Natl. Acad. Sci. USA* **88**, 8821–8825
5. Bischoff, M., Roesler, J., Raesecke, H.-R., Goerlach, J., Amrhein, N., and Schmid, J. (1996) Cloning of a cDNA encoding a 3-dehydroquinate synthase from a higher plant, and analysis of the organ-specific and elicitor-induced expression of the corresponding gene. *Plant Mol. Biol.* **31**, 69–76
6. Singh, S. A., and Christendat, D. (2006) Structure of *Arabidopsis* dehydroquinate dehydratase-shikimate dehydrogenase and implications for metabolic channeling in the shikimate pathway. *Biochemistry* **45**, 7787–7796
7. Fucile, G., Garcia, C., Carlsson, J., Sunnerhagen, M., and Christendat, D. (2011) Structural and biochemical investigation of two *Arabidopsis* shikimate kinases: The heat-inducible isoform is thermostable. *Protein Sci.* **20**, 1125–1136
8. Klee, H. J., Muskopf, Y. M., and Gasser, C. S. (1987) Cloning of an *Arabidopsis thaliana* gene encoding 5-enolpyruvylshikimate-3-phosphate synthase: Sequence analysis and manipulation to obtain glyphosate-tolerant plants. *Mol. Gen. Genet.* **210**, 437–442
9. Schaller, A., Schmid, J., Leibinger, U., and Amrhein, N. (1991) Molecular cloning and analysis of a cDNA coding for chorismate synthase from the higher plant *Corydalis sempervirens* Pers. *J. Biol. Chem.* **266**, 21434–21438
10. Eberhard, J., Ehrler, T. T., Epple, P., Felix, G., Raesecke, H.-R., Amrhein, N., and Schmid, J. (1996) Cytosolic and plastidic chorismate mutase isoenzymes from *Arabidopsis thaliana*: Molecular characterization and enzymic properties. *Plant J.* **10**, 815–821
11. Maeda, H., Yoo, H., and Dudareva, N. (2011) Prephenate aminotransferase directs plant phenylalanine biosynthesis via aroenate. *Nat. Chem. Biol.* **7**, 19–21
12. Cho, M.-H., Corea, O. R. A., Yang, H., Bedgar, D. L., Laskar, D. D., Anterola, A. M., Moog-Anterola, F. A., Hood, R. L., Kohalmi, S. E., Bernards, M. A., Kang, C., Davin, L. B., and Lewis, N. G. (2007) Phenylalanine biosynthesis in *Arabidopsis thaliana*: Identification and characterization of aroenate dehydratases. *J. Biol. Chem.* **282**, 30827–30835
13. Mittasch, J., Böttcher, C., Frolova, N., Bönn, M., and Milkowski, C. (2014) Identification of UGT84A13 as a candidate enzyme for the first committed step of gallotannin biosynthesis in pedunculate oak (*Quercus robur*). *Phytochemistry* **99**, 44–51
14. Madeira, F., Park, Y. m., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D., and Lopez, R. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641
15. Sperschneider, J., Catanzariti, A. M., DeBoer, K., Petre, B., Gardiner, D. M., Singh, K. B., Dodds, P. N., and Taylor, J. M. (2017) LOCALIZER: Subcellular localization prediction of both plant and effector proteins in the plant cell. *Sci Rep* **7**, 44598

16. Thumuluri, V., Almagro Armenteros, J. J., Johansen, Alexander R., Nielsen, H., and Winther, O. (2022) DeepLoc 2.0: Multi-label subcellular localization prediction using protein language models. *Nucleic Acids Res.* **50**, W228–W234
17. Cochrane, F. C., Davin, L. B., and Lewis, N. G. (2004) The *Arabidopsis* phenylalanine ammonia-lyase gene family. Kinetic characterization of the four PAL isoforms. *Phytochemistry* **65**, 1557–1564
18. Teutsch, H. G., Hasenfratz, M. P., Lesot, A., Stoltz, C., Garnier, J. M., Jeltsch, J. M., Durst, F., and Werck-Reichhart, D. (1993) Isolation and sequence of a cDNA encoding the Jerusalem artichoke cinnamate 4-hydroxylase, a major plant cytochrome P450 involved in the general phenylpropanoid pathway. *Proc. Natl. Acad. Sci. USA* **90**, 4102–4106
19. Hoffmann, L., Maury, S., Martz, F., Geoffroy, P., and Legrand, M. (2003) Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *J. Biol. Chem.* **278**, 95–103
20. Schoch, G., Goepfert, S., Morant, M., Hehn, A., Meyer, D., Ullmann, P., and Werck-Reichhart, D. (2001) CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *J. Biol. Chem.* **276**, 36566–36574
21. Costa, M. A., Bedgar, D. L., Moinuddin, S. G. A., Kim, K.-W., Cardenas, C. L., Cochrane, F. C., Shockey, J. M., Helms, G. L., Amakura, Y., Takahashi, H., Milhollan, J. K., Davin, L. B., Browse, J., and Lewis, N. G. (2005) Characterization *in vitro* and *in vivo* of the putative multigene 4-coumarate CoA ligase network in *Arabidopsis*: Syringyl lignin and sinapate/sinapyl alcohol derivative formation. *Phytochemistry* **66**, 2072–2091
22. Meyer, K., Cusumano, J. C., Somerville, C., and Chapple, C. C. S. (1996) Ferulate-5-hydroxylase from *Arabidopsis thaliana* defines a new family of cytochrome P450-dependent monooxygenases. *Proc. Natl. Acad. Sci. USA* **93**, 6869–6874
23. Goujon, T., Sibout, R., Pollet, B., Maba, B., Nussaume, L., Bechtold, N., Lu, F., Ralph, J., Mila, I., Barriere, Y., Lapierre, C., and Jouanin, L. (2003) A new *Arabidopsis thaliana* mutant deficient in the expression of *O*-methyltransferase impacts lignins and sinapoyl esters. *Plant Mol. Biol.* **51**, 973–989
24. Moinuddin, S. G. A., Jourdes, M., Laskar, D. D., Ki, C., Cardenas, C. L., Zhang, D., Davin, L. B., and Lewis, N. G. (2010) Insights into lignin primary structure and deconstruction from *Arabidopsis thaliana* COMT (caffeic acid *O*-methyl transferase) mutant *Atomt1*. *Org. Biomol. Chem.* **8**, 3928–3946
25. Schmitt, D., Pakusch, A.-E., and Matern, U. (1991) Molecular cloning, induction, and taxonomic distribution of caffeoyl-CoA 3-*O*-methyltransferase, an enzyme involved in disease resistance. *J. Biol. Chem.* **266**, 17416–17423
26. Patten, A. M., Cardenas, C. L., Cochrane, F. C., Laskar, D. D., Bedgar, D. L., Davin, L. B., and Lewis, N. G. (2005) Reassessment of effects on lignification and vascular development in the *irx4* *Arabidopsis* mutant. *Phytochemistry* **66**, 2092–2107
27. Kim, S.-J., Kim, M.-R., Bedgar, D. L., Moinuddin, S. G. A., Cardenas, C. L., Davin, L. B., Kang, C.-H., and Lewis, N. G. (2004) Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **101**, 1455–1460
28. Sommer, H., and Saedler, H. (1986) Structure of the chalcone synthase gene of *Antirrhinum majus*. *Mol. Gen. Genet.* **202**, 429–434
29. van Tunen, A. J., Hartman, S. A., Mur, L. A., and Mol, J. N. (1989) Regulation of chalcone flavanone isomerase (CHI) gene expression in *Petunia hybrida*: The use of alternative promoters in corolla, anthers and pollen. *Plant Mol. Biol.* **12**, 539–551

30. Pelletier, M. K., and Shirley, B. W. (1995) A genomic clone encoding flavanone 3-hydroxylase (Accession No. U33932) from *Arabidopsis thaliana*. *Plant Physiol.* **109**, 1125
31. Schoenbohm, C., Martens, S., Eder, C., Forkmann, G., and Weisshaar, B. (2000) Identification of the *Arabidopsis thaliana* flavonoid 3'-hydroxylase gene and functional expression of the encoded P450 enzyme. *Biol. Chem.* **381**, 749–753
32. Holton, T. A., Brugliera, F., Lester, D. R., Tanaka, Y., Hyland, C. D., Menting, J. G. T., Lu, C.-Y., Farcy, E., Stevenson, T. W., and Cornish, E. C. (1993) Cloning and expression of cytochrome P450 genes controlling flower colour. *Nature* **366**, 276–279
33. Holton, T. A., Brugliera, F., and Tanaka, Y. (1993) Cloning and expression of flavonol synthase from *Petunia hybrida*. *Plant J.* **4**, 1003–1010
34. Beld, M., Martin, C., Huits, H., Stuitje, A. R., and Gerats, A. G. (1989) Flavonoid synthesis in *Petunia hybrida*: Partial characterization of dihydroflavonol-4-reductase genes. *Plant Mol. Biol.* **13**, 491–502
35. Tanner, G. J., Francki, K. T., Abrahams, S., Watson, J. M., Larkin, P. J., and Ashton, A. R. (2003) Proanthocyanidin biosynthesis in plants: Purification of legume leucoanthocyanidin reductase and molecular cloning of its cDNA. *J. Biol. Chem.* **278**, 31647–31656
36. Pelletier, M. K., Burbulis, I. E., and Winkel-Shirley, B. (1999) Disruption of specific flavonoid genes enhances the accumulation of flavonoid enzymes and end-products in *Arabidopsis* seedlings. *Plant Mol. Biol.* **40**, 45–54
37. Xie, D. Y., Sharma, S. B., Paiva, N. L., Ferreira, D., and Dixon, R. A. (2003) Role of anthocyanidin reductase, encoded by *BANYULS* in plant flavonoid biosynthesis. *Science* **299**, 396–399
38. Martens, S., and Forkmann, G. (1999) Cloning and expression of flavone synthase II from *Gerbera* hybrids. *Plant J.* **20**, 611–618
39. Seneviratne, H. K., Dalisay, D. S., Kim, K. W., Moinuddin, S. G. A., Yang, H., Hartshorn, C. M., Davin, L. B., and Lewis, N. G. (2015) Non-host disease resistance response in pea (*Pisum sativum*) pods: Biochemical function of DRR206 and phytoalexin pathway localization. *Phytochemistry* **113**, 140–148
40. Meng, Q., Moinuddin, S. G. A., Kim, S.-J., Bedgar, D. L., Costa, M. A., Thomas, D. G., Young, R. P., Smith, C. A., Cort, J. R., Davin, L. B., and Lewis, N. G. (2020) Pterocarpan synthase (PTS) structures suggest a common quinone methide-stabilizing function in dirigent proteins and proteins with dirigent-like domains. *J. Biol. Chem.* **295**, 11584–11601
41. Effenberger, I., Zhang, B., Li, L., Wang, Q., Liu, Y., Klaiber, I., Pfannstiel, J., Wang, Q., and Schaller, A. (2015) Dirigent proteins from cotton (*Gossypium* sp.) for the atropselective synthesis of gossypol. *Angew. Chem. Int. Ed. Engl.* **54**, 14660–14663
42. Hosmani, P. S., Kamiya, T., Danku, J., Naseer, S., Geldner, N., Guerinot, M. L., and Salt, D. E. (2013) Dirigent domain-containing protein is part of the machinery required for formation of the lignin-based Casparian strip in the root. *Proc. Natl. Acad. Sci. USA* **110**, 14498–14503
43. Dinkova-Kostova, A. T., Gang, D. R., Davin, L. B., Bedgar, D. L., Chu, A., and Lewis, N. G. (1996) (+)-Pinoresinol/(+)-lariciresinol reductase from *Forsythia intermedia*: Protein purification, cDNA cloning, heterologous expression and comparison to isoflavone reductase. *J. Biol. Chem.* **271**, 29473–29482
44. Kim, S.-J., Vassão, D. G., Moinuddin, S. G. A., Bedgar, D. L., Davin, L. B., and Lewis, N. G. (2014) Allyl/propenyl phenol synthases from the creosote bush and engineering production of specialty/commodity chemicals, eugenol/isoegenol, in *Escherichia coli*. *Arch. Biochem. Biophys.* **541**, 37–46

45. Xia, Z.-Q., Costa, M. A., Péliissier, H. C., Davin, L. B., and Lewis, N. G. (2001) Secoisolariciresinol dehydrogenase purification, cloning and functional expression: Implications for human health protection. *J. Biol. Chem.* **276**, 12614–12623
46. Huang, K., Li, M., Liu, Y., Zhu, M., Zhao, G., Zhou, Y., Zhang, L., Wu, Y., Dai, X., Xia, T., and Gao, L. (2019) Functional analysis of 3-dehydroquinate dehydratase/shikimate dehydrogenases involved in shikimate pathway in *Camellia sinensis*. *Front. Plant Sci.* **10**, 1268
47. Akagi, T., Ikegami, A., Suzuki, Y., Yoshida, J., Yamada, M., Sato, A., and Yonemori, K. (2009) Expression balances of structural genes in shikimate and flavonoid biosynthesis cause a difference in proanthocyanidin accumulation in persimmon (*Diospyros kaki* Thunb.) fruit. *Planta* **230**, 899–915
48. Tahara, K., Nishiguchi, M., Funke, E., Miyazawa, S. I., Miyama, T., and Milkowski, C. (2021) ehydroquinate dehydratase/shikimate dehydrogenases involved in gallate biosynthesis of the aluminum-tolerant tree species *Eucalyptus camaldulensis*. *Planta* **253**, 3
49. Muir, R. M., Ibanez, A. M., Uratsu, S. L., Ingham, E. S., Leslie, C. A., McGranahan, G. H., Batra, N., Goyal, S., Joseph, J., Jemmis, E. D., and Dandekar, A. M. (2011) Mechanism of gallic acid biosynthesis in bacteria (*Escherichia coli*) and walnut (*Juglans regia*). *Plant Mol. Biol.* **75**, 555–565
50. Ding, L., Hofius, D., Hajirezaei, M. R., Fernie, A. R., Börnke, F., and Sonnewald, U. (2007) Functional analysis of the essential bifunctional tobacco enzyme 3-dehydroquinate dehydratase/shikimate dehydrogenase in transgenic tobacco plants. *J. Exp. Bot.* **58**, 2053–2067
51. Guo, J., Carrington, Y., Alber, A., and Ehling, J. (2014) Molecular characterization of quinate and shikimate metabolism in *Populus trichocarpa*. *J. Biol. Chem.* **289**, 23846–23858
52. Bischoff, M., Schaller, A., Bieri, F., Kessler, F., Amrhein, N., and Schmid, J. (2001) Molecular characterization of tomato 3-dehydroquinate dehydratase-shikimate:NADP oxidoreductase. *Plant Physiol.* **125**, 1891–1900
53. Bontpart, T., Marlin, T., Vialet, S., Guiraud, J.-L., Pinasseau, L., Meudec, E., Sommerer, N., Cheynier, V., and Terrier, N. (2016) Two shikimate dehydrogenases, *VvSDH3* and *VvSDH4*, are involved in gallic acid biosynthesis in grapevine. *J. Exp. Bot.* **67**, 3537–3550
54. Gritsunov, A., Peek, J., Diaz Caballero, J., Guttman, D., and Christendat, D. (2018) Structural and biochemical approaches uncover multiple evolutionary trajectories of plant quinate dehydrogenases. *Plant J.* **95**, 812–822
55. Cui, L., Yao, S., Dai, X., Yin, Q., Liu, Y., Jiang, X., Wu, Y., Qian, Y., Pang, Y., Gao, L., and Xia, T. (2016) Identification of UDP-glycosyltransferases involved in the biosynthesis of astringent taste compounds in tea (*Camellia sinensis*). *J. Exp. Bot.* **67**, 2285–2297
56. Ono, N. N., Qin, X., Wilson, A. E., Li, G., and Tian, L. (2016) Two UGT84 family glycosyltransferases catalyze a critical reaction of hydrolyzable tannin biosynthesis in pomegranate (*Punica granatum*). *PLOS ONE* **11**, e0156319
57. Madeira, F., Madhusoodanan, N., Lee, J., Eusebi, A., Niewielska, A., Tivey, A. R. N., Lopez, R., and Butcher, S. (2024) The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res.* **52**, W521–W525
58. Letunic, I., and Bork, P. (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296
59. Ross, J., Li, Y., Lim, E., and Bowles, D. J. (2001) Higher plant glycosyltransferases. *Genome Biol* **2**, REVIEWS3004
60. Kim, K.-W., Moinuddin, S. G. A., Atwell, K. M., Costa, M. A., Davin, L. B., and Lewis, N. G. (2012) Opposite stereoselectivities of dirigent proteins in *Arabidopsis* and *Schizandra* species. *J. Biol. Chem.* **287**, 33957–33972

61. Katsuyama, Y., Kita, T., Funa, N., and Horinouchi, S. (2009) Curcuminoid biosynthesis by two type III polyketide synthases in the herb *Curcuma longa*. *J. Biol. Chem.* **284**, 11160–11170