#### **Supplementary Figures**



**Supplementary Figure 1: Multiple amino acid sequence alignment of the FS and HFSs.** Protein sequences of CYP76AH1 (*S. miltiorrhiza*), CYP76AH24 (SfFS) (*S. fruticosa*), CYP76AH4 and CYP76AH22-23 (RoFS1 and RoFS2) (*R. officinalis*) are given. Conserved CYP domains (green) and amino acids which were mutagenized (purple) are highlighted.



Supplementary Figure 2: Multiple amino acid sequence alignment of the  $C_{20}Oxs$ .

CYP76AK6 (*S. fruticosa*) and CYP76AK7-8 (*R. officinalis*) are shown with highlighted

conserved CYP domains (green).



**Supplementary Figure 3: GC-MS analysis of yeast expressing FS and HFS.** (**a**) *R.* 

*officinalis* and *S. fruticosa* leaf surface extracts, and extracts of yeast strains coexpressing GGPPS, CPS, MiS, ATR1 and indicated CYPs (selected *m/z* signals: 270, 272, 286, 300, 302). Miltiradiene (**1**), abietatriene (**2**), ferruginol (**4**), 11 hydroxyferruginol (**5**) and hydroxyferruginol quinone (**6**). (**b**) EI mass spectra of abietatriene (**1**) and miltiradiene (**2**).





**Supplementary Figure 4: Total ion chromatograms of engineered yeast expressing the indicated enzymes and rosemary leaf surface extracts.** (**a**) GC-MS analysis. Miltiradiene (**1**), abietatriene (**2**), ferruginol (**4**), 11-hydroxyferruginol (**5**) and pisiferal (**8**). (**b**) LC-MS analysis. CA (**9**), CO (**10**) and PA (**11**). \*Originates from yeast.



**Supplementary Figure 5: LC-MS analysis of yeast expressing FS and HFS.** (**a**) Total ion chromatogram of leaf surface extracts from rosemary and sage, and of extracts from yeast co-expressing GGPPS, CPS, MiS, ATR1 and indicated CYPs. Ferruginol (**4**) and 11-hydroxyferruginol (**5**). (selected *m/z* signals: 285.221, 301.217) (**b**) ESI mass spectra of ferruginol (**4**) 11-hydroxyferruginol (**5**) from yeast and rosemary.



**Supplementary Figure 6: GC-MS analysis (selected** *m/z* **signals: 270, 272, 286, 302) of** *in vitro* **enzyme assays from microsomal preparations.** Indicated enzymes were incubated with the given substrates. Miltiradiene (**1**), abietatriene (**2**), ferruginol (**4**) and 11-hydroxyferruginol (**5**).



**Supplementary Figure 7: LC-MS analysis (selected** *m/z* **signals in the negative mode: 285.221, 301.217, 331.191, 329.175) of** *in vitro* **enzyme assays from microsomal preparations of indicated enzymes.** The assays with the given enzymes were incubated with different substrates. Ferruginol (**4**), 11-hydroxyferruginol (**5**), CA (**9**) and CO (**10**). Note: Ferruginol is poorly detected by LC-MS, while 11-hydroxyferruginol is poorly detected by GC-MS.





**Supplementary Figure 8: NMR analysis of 11-hydroxyferruginol.** (**a**) Structures of 11-hydroxyferruginol and hydroxyferruginol quinone with labeled carbon atoms.  $(b)$  <sup>1</sup>H NMR spectrum of 11-hydroxyferruginol immediately after dissolving in CDCl<sub>3</sub>. (c)  ${}^{1}H$ NMR spectrum of 11-hydroxyferruginol 24 h after dissolving in CDCl<sub>3</sub> showing signals of 11-hydroxyferruginol and hydroxyferruginol quinone. (**d**) Labeled: HMB correlations via  ${}^{2}J_{CH}$  and  ${}^{3}J_{CH}$  of methyl group proton signals of 11-hydroxyferruginol. Vertical trace: <sup>1</sup>H 1D NMR spectrum. (e) Low-field part of the HMBC spectrum of 11hydroxyferruginol. Horizontal trace: <sup>1</sup>H 1D NMR spectrum. (**f**) Methyl group region of the HSQC spectrum of 11-hydroxyferruginol (5) after partial conversion to hydroxyferruginol quinone (6). Vertical trace: <sup>1</sup>H 1D NMR spectrum.



**Supplementary Figure 9: 3D-model of CYP76AH1 as representative of all other FS and HFSs.** The heme (space fill representation with orange colored iron ion) and the substrate (magenta carbon atoms) are located in the center.



**Supplementary Figure 10: 3D-model of the active sites of CYP76AH1 and CYP76AH22 with bound substrate (green carbon atoms) and heme (orange carbon atoms).** Red labelled amino acid residues influence the product specificity which was proven by site directed mutagenesis. In the active center of the enzymes the substrate is mainly recognized by hydrophobic amino acid residues (V297, L367, V210 and V479 or F478, respectively). Active sites of (**a**) CYP76AH1 and (**b**) of CYP76AH22 with bound

abietatriene. The hydrogen atom of the aromatic ring system of the substrate is abstracted by the reactive oxygen atom bound to the iron ion. They are in closest proximity (3.6 Å in CYP76AH1 and 3.4 Å in CYP76AH22, red dashed line) which supports the oxidation reaction. (**c**) Active site of CYP76AH1 with bound ferruginol. The substrate cannot be efficiently oxidized because its distance to the reactive oxygen atom is 4.6 Å and therefore too far to support oxidation at C12 (red dashed line). In contrast, the substitution of V479 in CYP76AH1 by the more bulky F478 in CYP76AH22 leads to a slightly different docking arrangement of ferruginol (**d**). Therefore, the hydrogen atom at C12 can be abstracted by the reactive oxygen atom which is in a distance of 3.1 Å.



**Supplementary Figure 11: 3D-model of CYP76AH1 with bound heme in the center.** Labelled amino acids A117 and S118 are distinct from those in HFS. They may influence the secondary structure in the colored region and restrict access to the active site cleft.



**Supplementary Figure 12: 3D-model of CYP76AH22 with bound heme in the center.** Labelled amino acids G117 and G118 disturb the α-helix present in CYP76AH1 and influence the secondary structure in the highlighted region.



**Supplementary Figure 13: Chromatographic analysis of** e**xtracts from yeast strains expressing wild type and mutagenized CYP76AH1.** They were co-expressed in yeast with GGPPS, CPS, MiS and ATR1. Part of the chromatograms showing elution of ferruginol (**4**) and 11-hydroxyferruginol (**5**) from (**a**) GC-MS (selected *m/z* signals: 286, 300, 302) and (**b**) LC-MS analysis (selected *m/z* signals in the negative mode: 285.221, 301.217) are given.



**Supplementary Figure 14: Chromatographic analysis of** e**xtracts from yeast strains expressing wild type and mutagenized CYP76AH22.** They were co-expressed in yeast with GGPPS, CPS, MiS and ATR1. Part of the chromatograms showing elution of ferruginol (**4**) and 11-hydroxyferruginol (**5**) from (**a**) GC-MS (selected *m/z* signals: 286, 300, 302) and (**b**) LC-MS analysis (selected *m/z* signals in the negative mode: 285.221, 301.217) are given.



Supplementary Figure 15: GC-MS analysis of C<sub>20</sub>Ox expression in yeast and N. *benthamiana* **producing miltiradiene/abietatriene.** Part of the GC-MS chromatogram containing miltiradien-20-al (3) is shown (selected *m/z* signal: 257). (**a**) Transient coexpression of CPS, MiS and indicated enzymes in *N. benthamiana*. (**b**) Co-expression of GGPPS, CPS, MiS, ATR1 and indicated CYPs in yeast.



**Supplementary Figure 16: GC-MS analysis of yeast strains co-expressing GGPPS,** 

**CPS, MiS, ATR1 and indicated CYPs.** (Selected *m/z* signals: 270,272, 286, 300, 302,

316). Miltiradiene (**1**), abietatriene (**2**), ferruginol (**4**), 11-hydroxyferruginol (**5**),

hydroxyferruginol quinone (**6**) and carnosaldehyde (**7**).



**Supplementary Figure 17: LC-MS analysis of extracts obtained from rosemary and sage leaves, authentic standards and yeast strains expressing GGPPS, CPS, MiS, ATR1 and indicated CYPs.** (Selected *m/z* signals in the negative mode: 315.196, 331.191, 329.175). Carnosaldehyde (7), CA (9) and CO (10).





**rosemary and sage as well as of yeast strains co-expressing GGPPS, CPS, MiS,** 

**ATR1 and indicated CYPs**. (Selected *m/z* signals: 270, 272, 286, 300). Miltiradiene (**1**),

abietatriene (**2**), ferruginol (**4**) and pisiferal (**8**).



**Supplementary Figure 19: LC-MS total ion chromatogram of extracts from rosemary, sage and from yeast strains co-expressing GGPPS, CPS, MiS, ATR1 and indicated CYPs.** CA (9) and PA (11).



**Supplementary Figure 20: Chromatographic analysis of engineered yeast strains which were used for absolute NMR quantification.** (**a**) GC-MS analysis of indicated strains (selected *m/z* signals: 270, 272, 286, 300, 302, 316). (**b**) LC-MS analysis of yeast strains expressing the indicated enzymes (selected *m/z* signals in the negative mode: 285.221, 301.217, 315.196, 331.191, 329.175). Miltiradiene (**1**), abietatriene (**2**), ferruginol (**4**), 11-hydroxyferruginol (**5**), pisiferal (**8**), CA (**9**), CO (**10**) and PA (**11**).









**Supplementary Figure 21: <sup>1</sup> H NMR spectra of engineered yeast for quantification.** The strains expressed the enzymes GGPPS, CPS, MiS (**a**) together with ATR1 and CYP76AH1 (**b**), CYP76AH22 (**c**), CYP76AH1D301EN303SV479F (**d**),

CYP76AH22E301DS303NF478V (**e**), CYP76AH22 and CYP76AK8 (**f**), or CYP76AH1 and CYP76AK8 (**g**). Product specific signals, which were used for quantification, are highlighted. The corresponding integrals including the peak areas are given. As internal standard hexamethyldisiloxane (HMDS) was used for quantification.



30

### **Supplementary Figure 22: Phylogenetic analysis of CYPs involved in the**

**biosynthesis of CA and PA.** The maximum likelihood tree was generated using MEGA (version 6) from protein sequence alignment and illustrates the relationship of CYP enzymes of the PD biosynthesis to other proteins from the CYP76 and CYP71 clan. Bootstrap values (1000 repeats) are indicated for each branch of the tree.

## **Supplementary Tables**

# **Supplementary Table 1: Primers used for cloning, introduction of single point mutations and quantitative real-time PCR.**



















**Supplementary Table 2: Sequences of synthetic galactose inducible promoters and** 

### **assignment to their assembled genes and terminators.**



![](_page_40_Picture_99.jpeg)

![](_page_41_Picture_74.jpeg)

![](_page_42_Picture_50.jpeg)

![](_page_43_Picture_120.jpeg)

# **Supplementary Table 3: Constructs and corresponding yeast strains.**

# *GGPPS:CPS:MiS:ATR1:CYP76AH4:CYP76AK8* 3,4,7

## *GGPPS:CPS:MiS:ATR1:CYP76AH1:CYP76AK8* 3-5

CYP76AH22, CYP76AH23 and CYP76AH24 were formerly known as RoFS1, RoFS2 and SfFS, respectively.

## **Supplementary Table 4: NMR data of 11-hydroxyferruginol and hydroxyferruginol**

**quinone.** (<sup>13</sup>C: δ [ppm], <sup>1</sup>H: δ [ppm] m (J [Hz]) (CDCl<sub>3</sub>) obtained from a Agilent VNMRS 600, CDCl<sub>3</sub>, <sup>1</sup>H @ 599.829 MHz (<sup>1</sup>H; <sup>1</sup>H, <sup>13</sup>C HSQC; <sup>1</sup>H, <sup>13</sup>C HMBC), reference: <sup>1</sup>H: tetramethylsilane int. = 0 ppm; concentration: ca. 0.2 mmol/L. Carbon positions refer to Fig. S5A.

![](_page_44_Picture_212.jpeg)

![](_page_45_Picture_212.jpeg)

 $\frac{1}{a}$  chemical shifts of HSQC or HMBC correlation peaks;  $\frac{1}{b}$  may be interchanged; n.d. not detected; s singlet; d doublet; t triplet; sp septet; ddd doublet of doublet of doublet.

**Supplementary Table 5: NMR data of miltiradien-20-al**. (<sup>13</sup>C: δ [ppm], <sup>1</sup>H: δ [ppm] m (J [Hz]) (CDCl<sub>3</sub>) obtained from a Agilent VNMRS 600,  $C_6D_{12}$ , <sup>1</sup>H @ 599.829 MHz (<sup>1</sup>H; <sup>1</sup>H,<sup>13</sup>C HSQC; <sup>1</sup>H,<sup>13</sup>C HMBC), reference: <sup>1</sup>H: tetramethylsilane int. = 0 ppm. Carbon positions refer to Fig. S5A.

![](_page_46_Picture_237.jpeg)

![](_page_47_Picture_216.jpeg)

 $\frac{1}{a}$  chemical shifts of HSQC or HMBC correlation peaks;  $\frac{1}{b}$  may be interchanged;

s singlet; d doublet; t triplet; { } additional information.

**Supplementary Table 6: NMR data of carnosaldehyde.** ( 13C: δ [ppm]¸ <sup>1</sup> H: δ [ppm] m (J [Hz]) (CDCl<sub>3</sub>) obtained from a Agilent VNMRS 600, CDCl<sub>3</sub>, <sup>1</sup>H @ 599.829 MHz (<sup>1</sup>H; <sup>1</sup>H,<sup>13</sup>C HSQC; <sup>1</sup>H,<sup>13</sup>C HMBC), reference: <sup>1</sup>H: tetramethylsilane int. = 0 ppm. Carbon positions refer to Fig. S5A.

![](_page_47_Picture_217.jpeg)

![](_page_48_Picture_173.jpeg)

 $\frac{1}{a}$  chemical shifts of HSQC or HMBC correlation peaks; n.d. not detected; s singlet,; d

doublet; t triplet; sp septet; { } additional information.

**Supplementary Table 7: Protein sequences used for phylogenetic analysis with their corresponding accession numbers.**

![](_page_49_Picture_160.jpeg)

![](_page_50_Picture_157.jpeg)

![](_page_51_Picture_180.jpeg)

<sup>a</sup> not characterized yet

### **Supplementary References**

- 1 Yamanishi, M. *et al.* A Genome-Wide Activity Assessment of Terminator Regions in *Saccharomyces cerevisiae* Provides a ″Terminatome″ Toolbox. *ACS Synthetic Biology* **2**, 337-347, (2013).
- 2 Zaret, K. S. & Sherman, F. DNA sequence required for efficient transcription termination in yeast. *Cell* **28**, 563-573 (1982).
- 3 Brückner, K. *et al.* Characterization of two genes for the biosynthesis of abietanetype diterpenes in rosemary (*Rosmarinus officinalis*) glandular trichomes. *Phytochemistry* **101**, 52-64 (2014).
- 4 Mizutani, M. & Ohta, D. Two Isoforms of NADPH:Cytochrome P450 Reductase in *Arabidopsis thaliana*: Gene Structure, Heterologous Expression in Insect Cells, and Differential Regulation. *Plant Physiology* **116**, 357-367, (1998).
- 5 Guo, J. *et al.* CYP76AH1 catalyzes turnover of miltiradiene in tanshinones biosynthesis and enables heterologous production of ferruginol in yeasts. *Proc Natl Acad Sci U S A* **110**, 12108-12113, (2013).
- 6 Božić, D. *et al.* Towards Elucidating Carnosic Acid Biosynthesis in Lamiaceae: Functional Characterization of the Three First Steps of the Pathway in *Salvia fruticosa* and *Rosmarinus officinalis*. *PLoS One* **10**, e0124106, (2015).
- 7 Zi, J. & Peters, R. J. Characterization of CYP76AH4 clarifies phenolic diterpenoid biosynthesis in the Lamiaceae. *Org Biomol Chem* **11**, 7650-7652 (2013).