Supplementary information for:

Thermostable fatty acid hydroxylases from ancestral reconstruction of cytochrome P450 family 4 enzymes

Kurt L. Harris[1], Yichi Zhang[1], Jade Yang[2], Maxwell B. Zeigler[2], Raine E.S. Thomson [1], Saskya E. Carrera-Pacheco [3], Drake Russell [2], Shoko Okada[4], Silja J. Strohmaier[1], Yosephine Gumulya[1], Colin Scott[4], Rheem A. Totah[2], and Elizabeth M.J. Gillam[1]*

[1] School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, Brisbane, 4072 Australia

[2] Department of Medicinal Chemistry, School of Pharmacy, University of Washington, Seattle, WA, 98195-7631, USA

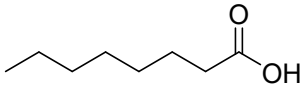
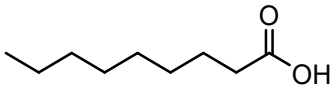
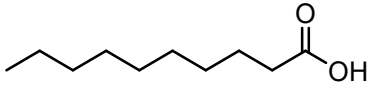
[3] Centro de Investigación Biomédica (CENBIO), Facultad de Ciencias de la Salud Eugenio Espejo, Universidad UTE, Quito 170527, Ecuador

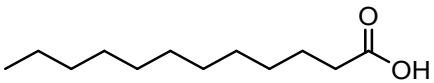
[4] CSIRO Synthetic Biology Future Science Platform, Black Mountain Science and Innovation Park, Clunies Ross St, Canberra, 2601, Australia

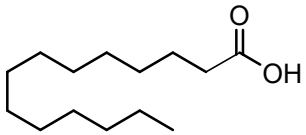
*Corresponding author: e.gillam@uq.edu.au

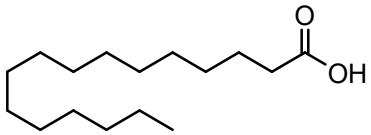
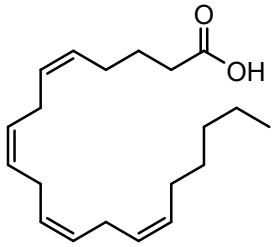
Supplementary Table 1. Substrate and hydroxylation regioselectivities of mammalian CYP4 enzymes towards FAs

Common FA substrates of the CYP4s are listed, along with the forms that metabolise them and regioselectivities if known (in the form of terminal to subterminal i.e., ω:ω-1 ratio, unless otherwise stated).

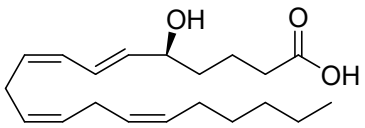
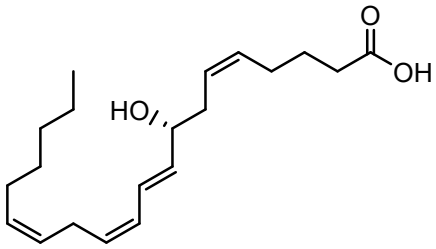
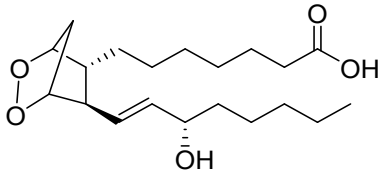
FA Substrate	Form	Regioselectivity	Source
Caprylic acid (C8) 	Human CYP4B1 (S427P) Rabbit CYP4B1	>1.5:1 7.5:1	[1]
Nonanoic acid (C9) 	Human CYP4B1 (S427P) Rabbit CYP4B1	1.6:1 1.4:1	[1]
Capric acid (C10) 	Human CYP4B1 (S427P) Rabbit CYP4B1	3.1:1 1.1:1	[1]
Lauric acid (C12)	Rat CYP4A1	40:1	[2]

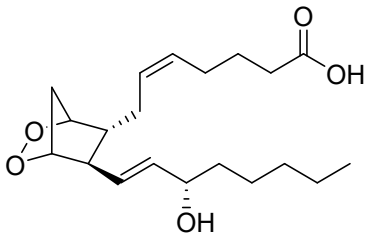
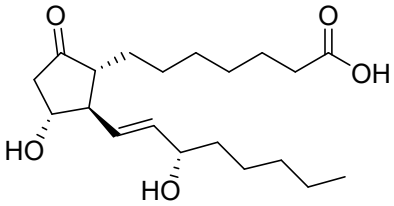
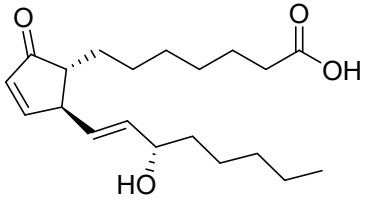
	Rat CYP4A8	25:1	
	Rat CYP4A2 Rat CYP4A3	>6:1 >3:1	[2,3]
	Rabbit CYP4A5 Rabbit CYP4A6 Rabbit CYP4A7	3:1 12:1 18:1	[4]
	Human CYP4A11	>15:1	[2,5]
	Mouse CYP4A10 Mouse CYP4A12a Mouse CYP4A12b Mouse CYP4A14	16:1 5:1 1:1.1 1.6:1	[6]
	Human CYP4B1 (S427P)	1.2:1	[1]

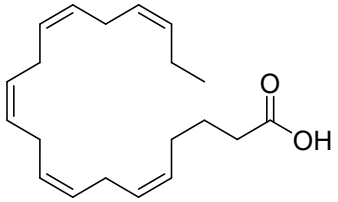
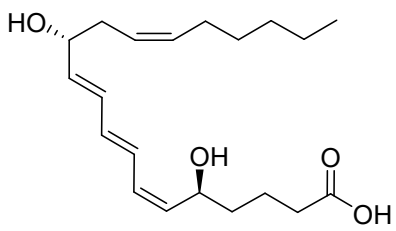
	Rabbit CYP4B1	3.5:1	
	Rat CYP4F4	ω	[7]
	Human CYP4Z1	ω-4>(2,3,5)	[8]
Myristic acid (C14)  	Rat CYP4A1 Rat CYP4A2 Rat CYP4A3 Rat CYP4A8 Human CYP4A11	3:1 3:1 1.2:1 1.6:1 4:1	[2]
	Human CYP4V2	ω-2>(3,4,5)	[9]
	Human CYP4Z1	ω-2	[8]
Palmitic acid (C16)	Rat CYP4A1 Rat CYP4A8	1:1 1.6:1	[2]

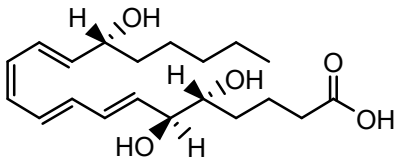
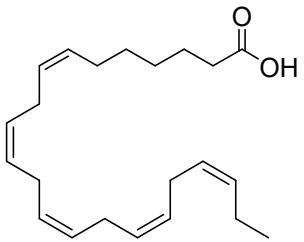
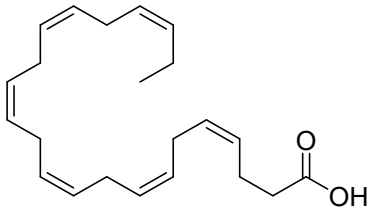
	Rabbit CYP4A4 Rabbit CYP4A5 Rabbit CYP4A6 Rabbit CYP4A7	4:1 1.4:1 2.7:1 2.5:1	[4]
	Human CYP4A11	>2.2:1	[2,5]
Arachidonic acid (C20:4) 	Rat CYP4A1 Rat CYP4A2 Rat CYP4A3	6:1 2.4:1 2:1	[2]
	Rabbit CYP4A4 Rabbit CYP4A6 Rabbit CYP4A7	ω ω ω	[4]
	Human CYP4A11	3.7:1	[5]

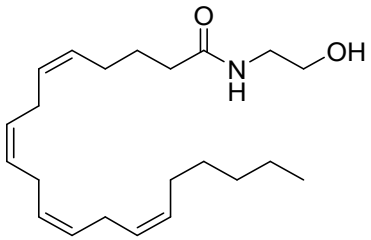
	xMouse CYP4A10 Mouse CYP4A12a Mouse CYP4A12b	4:1 7:1 8:1	[6]
	Rabbit CYP4B1	ω-8	[10-12]
	Rat CYP4F1 Rat CYP4F4	ω ω	[7]
	Human CYP4F2	26:1	[13,14]
	Human CYP4F3a Human CYP4F3b	25:1 20:1	[14]
	Human CYP4F8	ω-2	[15,16]
	Human CYP4F12	ω-2	[16,17]
	Human CYP4X1	internal epoxidation	[18]

5-HETE 	Mouse CYP4A14	ω	[19]
8-HETE 	Human CYP4F2	ω	[20]
	Mouse CYP4A14	ω	[19]
Prostaglandin H_1 	Human CYP4F8	ω-1>ω-2	[15]
Prostaglandin H_2	Human CYP4F8	ω-1>ω-2	[15]

			
Prostaglandin E_1 	Rabbit CYP4A4	ω	[4]
Prostaglandin A_1 	Rabbit CYP4A4 Rabbit CYP4A7	ω	[4]
Eicosapentaenoic acid (C20:5)	Mouse CYP4A12a	(ω-2, ω-3)-epoxide	[6]

	Mouse CYP4A12b		
	Human CYP4F3a Human CYP4F3b	2:1 1.7:1	[14]
	Human CYP4V2	ω>ω-1	[9]
Leukotriene B_4 (C20) 	Rat CYP4F1 Rat CYP4F4	ω ω	[7]
	Rat CYP4F5	ω-2	[21]
	Rat CYP4F6	ω-1, ω-2	[21]
	Human CYP4F2	ω	[20]
	Human CYP4F3a Human CYP4F3b	ω ω	[22]
	Mouse CYP4F14	ω	[19]

	Mouse CYP4F18	ω-1>ω-2	[23]
Lipoxin A_4 	Rat CYP4F1	ω	[24]
	Human CYP4F2	ω	[20]
	Mouse CYP4A14	ω	[19]
Docosapentaenoic acid (C22:5) 	Human CYP4F8 Human CYP4F12	(ω-2, ω-3)-epoxide	[16]
Docosahexaenoic acid (C22:6) 	Human CYP4F3a Human CYP4F3b	1.1:1 2.1:1	[14]
	Human CYP4F12	(ω-2, ω-3)-epoxide	[16]
	Human CYP4V2	ω-1>ω-2	[9]

Anandamide (C22) 	Human CYP4X1	internal epoxidation	[18]

23

24

Supplementary Table 2. N-terminal modifications used for CYP4 expression

The known N-terminal modifications applied to extant members of the CYP4ABTXZ clade (hCYP4A11[25], rCYP4B1[26], hCYP4X1[18]) to facilitate recombinant expression in E. coli are compared with the corresponding native form of each sequence as well as the unmodified "native" sequences for each node of interest. The existing N-terminal modifications in the literature were used to help identify possible modifications to apply to each ancestor. Ultimately, the hCYP4A11-like modification was applied to CYP4ABTXZ, as the hCYP4A11 sequence showed the greatest sequence similarity (77%) to the ancestor. The rCYP4B1-like modification was applied to all intermediate nodes, as this sequence retained the greatest length of native sequence, allowing for subsequent shortening to apply alternative N-terminal modifications if required for expression.

Form	N-terminal sequence	Reference
hCYP4A11_Native	LLGDVSGILQAASLLILLLLLIKAVQLYLHRQWLLKALQQFPCPPSHWLFGH	[27]
hCYP4A11_Modified	--------MALLLAVFLLLLLIKAVQLYLHRQWLLKALQQFPCPPSHWLFGH	[25]
rCYP4B1_Native	MLGFLSRLGLWASGLILILGFLKLLRLLLRRQRLARAMDSFPGPPTHWLFGH	[28]
rCYP4B1_Modified	MALLLAVFGLWASGLILILGFLKLLRLLLRRQRLARAMDSFPGPPTHWLFGH	[26]
hCYP4X1_Native	LETRWARPFYLAFVFCLALGLLQAIKLYLRRQRLLRDLRPFPAPPTHWFLGH	[29]
hCYP4X1_Modified	----------------------------MAKKTSSKGKL-PFPAPPTHWFLGH	[18]
CYP4ABTXZ_Native	----------------LALLLLKAIQLYLRRQRLLRALQLFPGPPTHWLYGH	

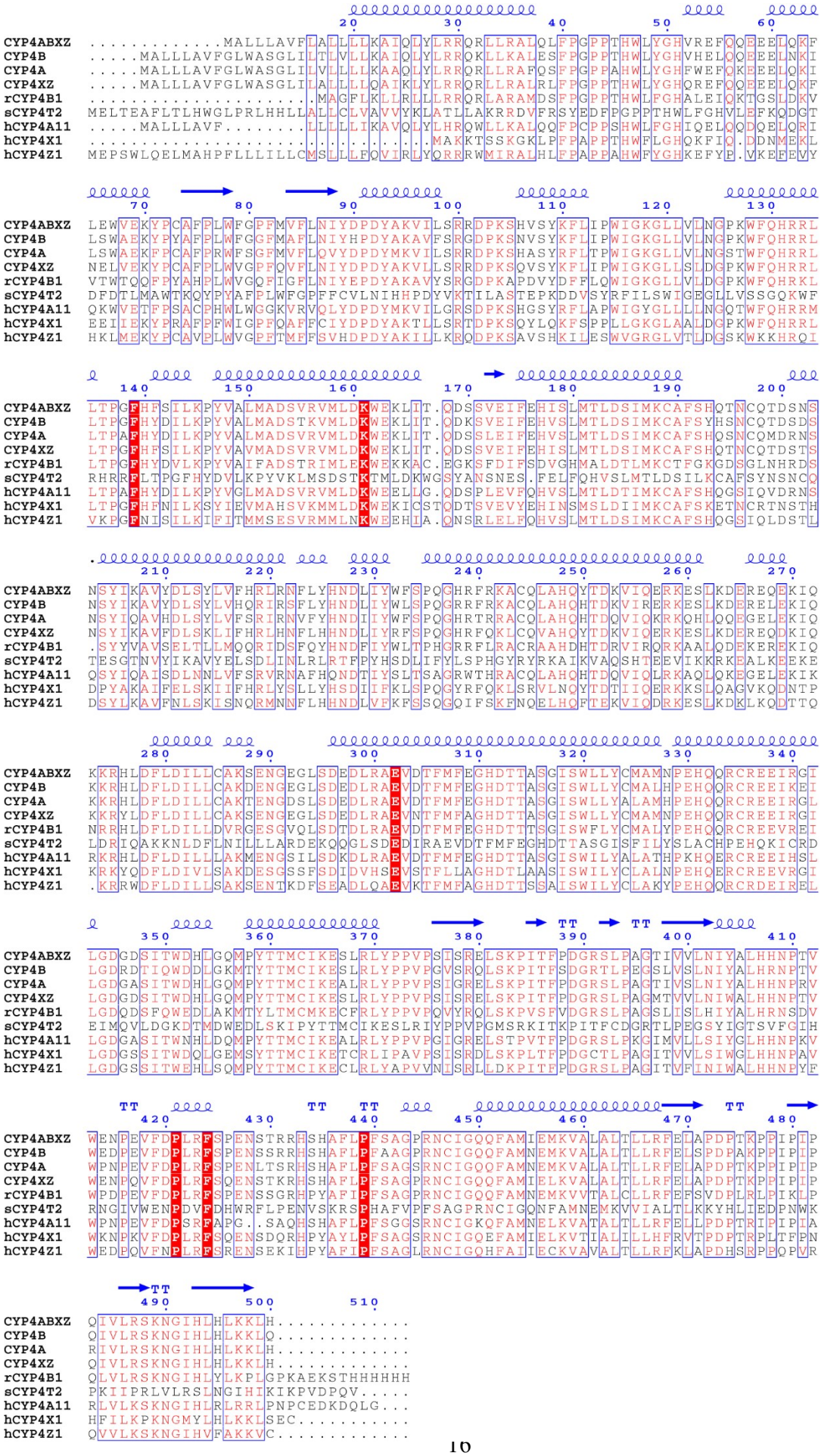
```
CYP4ABTXZ_Modified (4A11-like)  --------MALLLAVFLALLLLKAIQLYLRRQRLLRALQLFPGPPTHWLYGH
CYP4ABTXZ_Modified (4B1-like)   MALLLAVFGLWASGLILALLLLKAIQLYLRRQRLLRALQLFPGPPTHWLYGH
CYP4ABTXZ_Modified (4X1-like)   ---------------------------MAKKTSSKGKLQLFPGPPTHWLYGH
CYP4B_Native                    ----------------LTLVLLKAIQLYLRRQKLLKALESFPGPPTHWLYGH
CYP4B_Modified (4B1-like)       MALLLAVFGLWASGLILTLVLLKAIQLYLRRQKLLKALESFPGPPTHWLYGH
CYP4A_Native                    ----------------LVLLLLKAAQLYLRRQRLLRAFQSFPGPPAHWLYGH
CYP4A_Modified (4A11-like)      --------MALLLAVFLVLLLLKAAQLYLRRQRLLRAFQSFPGPPAHWLYGH
CYP4A_Modified (4B1-like)       MALLLAVFGLWASGLILVLLLLKAAQLYLRRQRLLRAFQSFPGPPAHWLYGH
CYP4XZ_Native                   ----------------LALLLLQAIKLYLRRQRLLRALRLFPGPPTHWLYGH
CYP4XZ_Modified (4X1-like)      ---------------------------MAKKTSSKGKLRLFPGPPTHWLYGH
CYP4XZ_Modified (4B1-like)      MALLLAVFGLWASGLILALLLLQAIKLYLRRQRLLRALRLFPGPPTHWLYGH
```

Supplementary Table 3. Mass transitions used to detect the AA metabolites and the internal standards.

Fatty Acid	**Mass Transition**
20- HETE	487>183.1
19- HETE	487>183
15- HETE	487>183
12-HETE	487>347
11- HETE	487>335
9- HETE	487>307
5-HETE	487>283
14,15- EET	487>333
11,12- EET	487>333
8,9- EET	487>307
5,6- EET	487>283
12- HETE d8	495>183.1
14,15- EET d11	498>183.1
8,9- EET d11	498>183.1
5,6- EET d11	498>183.1

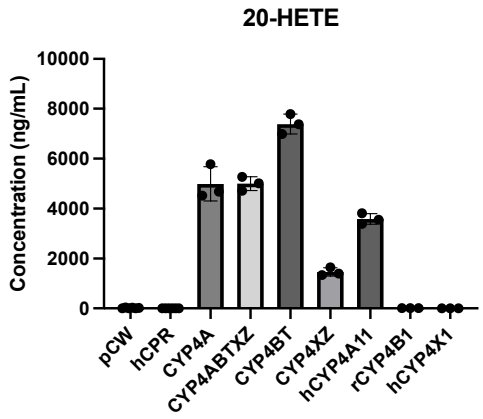
Supplementary Table 4. Estimated limit of detection (LOD) and limit of quantitation (LOQ) for the measured EETs and HETEs

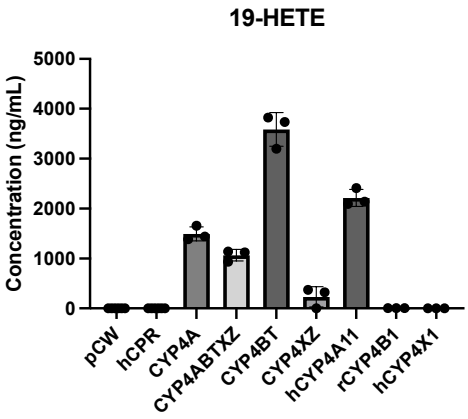
	5,6-EET	8,9-EET	11,12-EET	14,15-EET	5-HETE	9-HETE	8-HETE	12-HETE	11-HETE	15-HETE	19-HETE	20-HETE
LOD (fmol/mL; S/N = 3)	1597	349	127	57	343	179	61	167	121	45	89	43
LOQ (fmol/mL; S/N = 10)	5324	1164	423	190	545	598	202	558	405	150	297	144

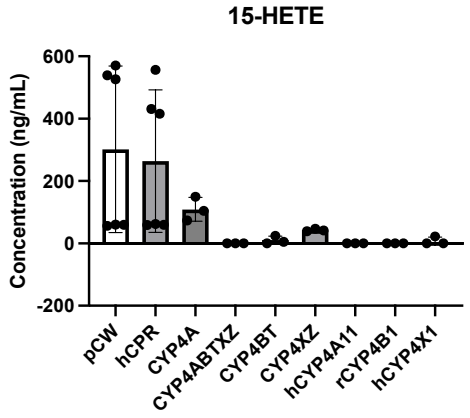

CYP4ABXZ
CYP4B
CYP4A
CYP4XZ
rCYP4B1
sCYP4T2
hCYP4A11
hCYP4X1
hCYP4Z1

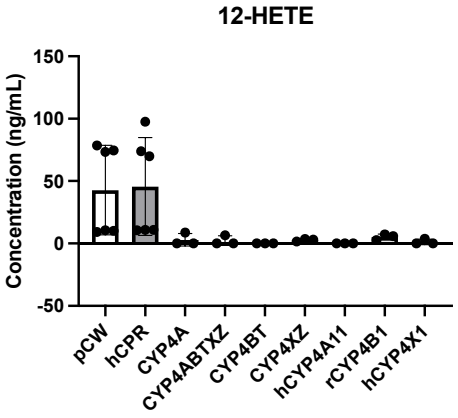
Supplementary Figure 1. Multiple sequence alignment of the CYP4 ancestors and corresponding extant forms.

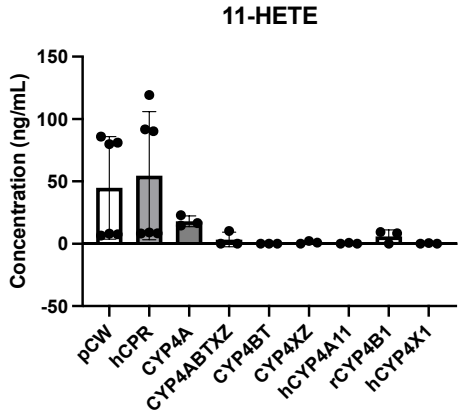
The figure was created using ENDscript 3.0 [30]. Red boxes indicate amino acid identity, red characters indicate similarity, and blue frames surround regions of homology. The secondary structure of rCYP4B1 is represented above the alignment in blue, with α-helices depicted as coils and β-sheets as arrows. Residue numbering is with respect to the numbering used for the rCYP4B1 crystal structure (PDB: 5T6Q).

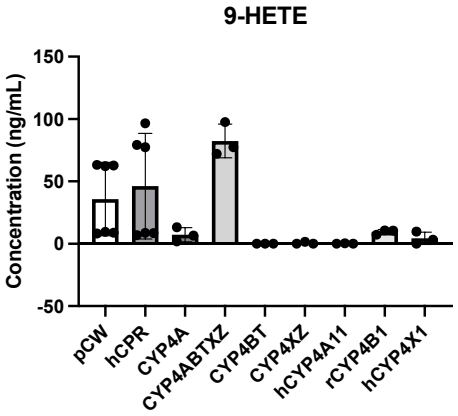

20-HETE
Concentration (ng/mL)
10000
8000
6000
4000
2000
0
pCW
hCPR
CYP4A
CYP4ABTXZ
CYP4BT
CYP4XZ
hCYP4A11
rCYP4B1
hCYP4X1

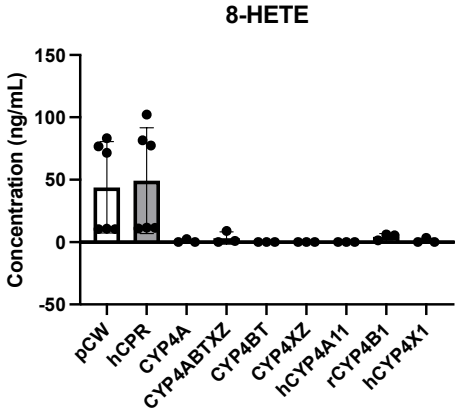

19-HETE
Concentration (ng/mL)
5000
4000
3000
2000
1000
0
pCW
hCPR
CYP4A
CYP4ABTXZ
CYP4BT
CYP4XZ
hCYP4A11
rCYP4B1
hCYP4X1

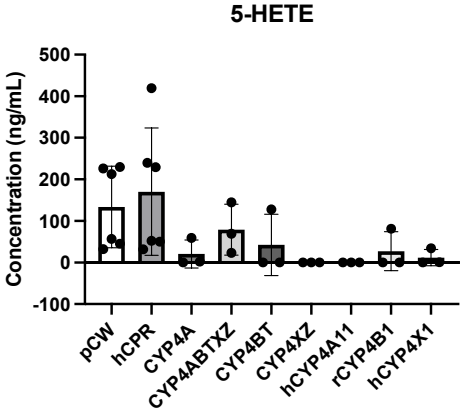

15-HETE
Concentration (ng/mL)
600
400
200
0
-200
pCW
hCPR
CYP4A
CYP4ABTXZ
CYP4BT
CYP4XZ
hCYP4A11
rCYP4B1
hCYP4X1


12-HETE
Concentration (ng/mL)
150
100
50
0
-50
pCW
hCPR
CYP4A
CYP4ABTXZ
CYP4BT
CYP4XZ
hCYP4A11
rCYP4B1
hCYP4X1


11-HETE
Concentration (ng/mL)
150
100
50
0
-50
pCW
hCPR
CYP4A
CYP4ABTXZ
CYP4BT
CYP4XZ
hCYP4A11
rCYP4B1
hCYP4X1


9-HETE
Concentration (ng/mL)
150
100
50
0
-50
pCW
hCPR
CYP4A
CYP4ABTXZ
CYP4BT
CYP4XZ
hCYP4A11
rCYP4B1
hCYP4X1

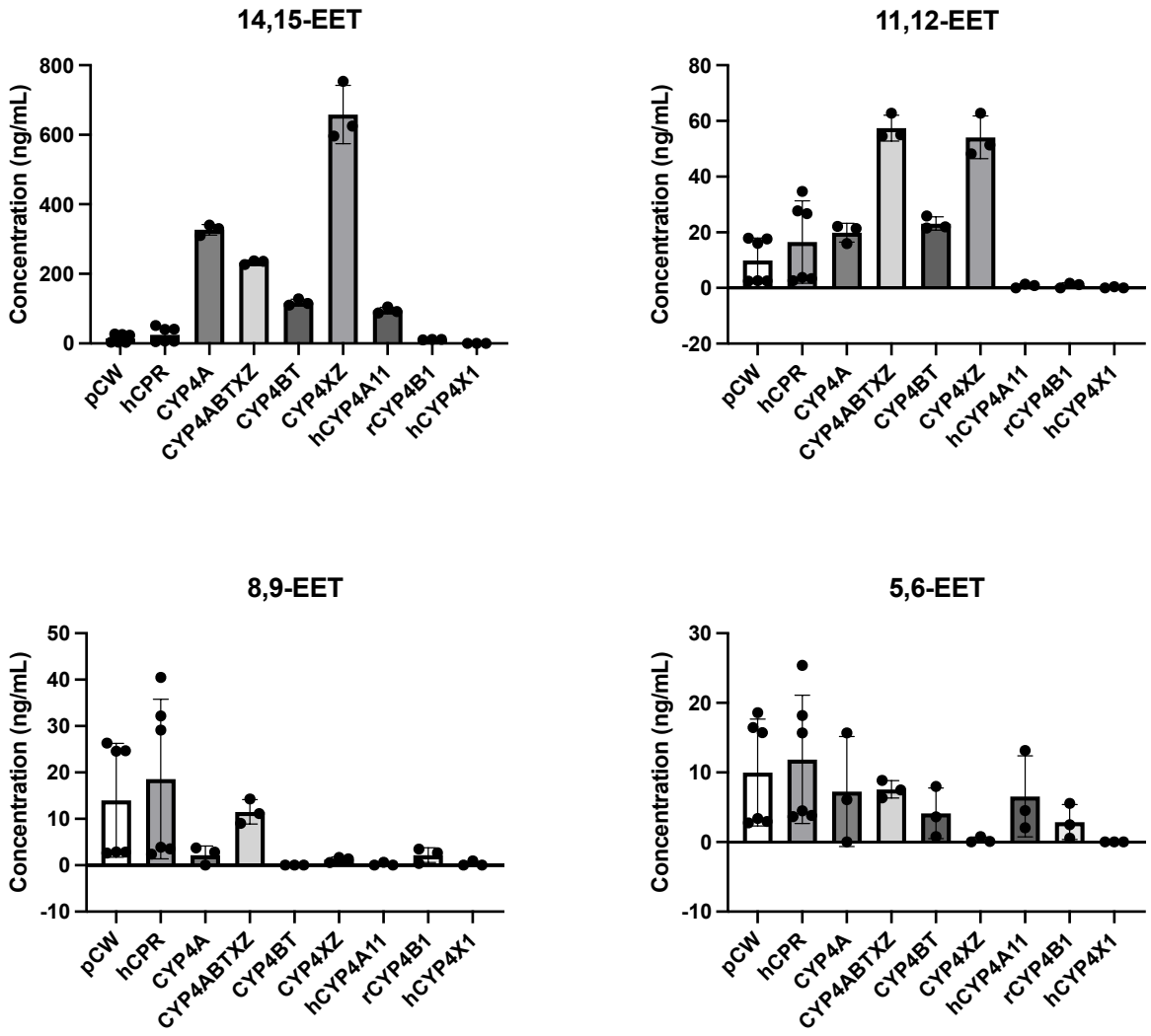

8-HETE
Concentration (ng/mL)
150
100
50
0
-50
pCW
hCPR
CYP4A
CYP4ABTXZ
CYP4BT
CYP4XZ
hCYP4A11
rCYP4B1
hCYP4X1


5-HETE
Concentration (ng/mL)
500
400
300
200
100
0
-100
pCW
hCPR
CYP4A
CYP4ABTXZ
CYP4BT
CYP4XZ
hCYP4A11
rCYP4B1
hCYP4X1

57

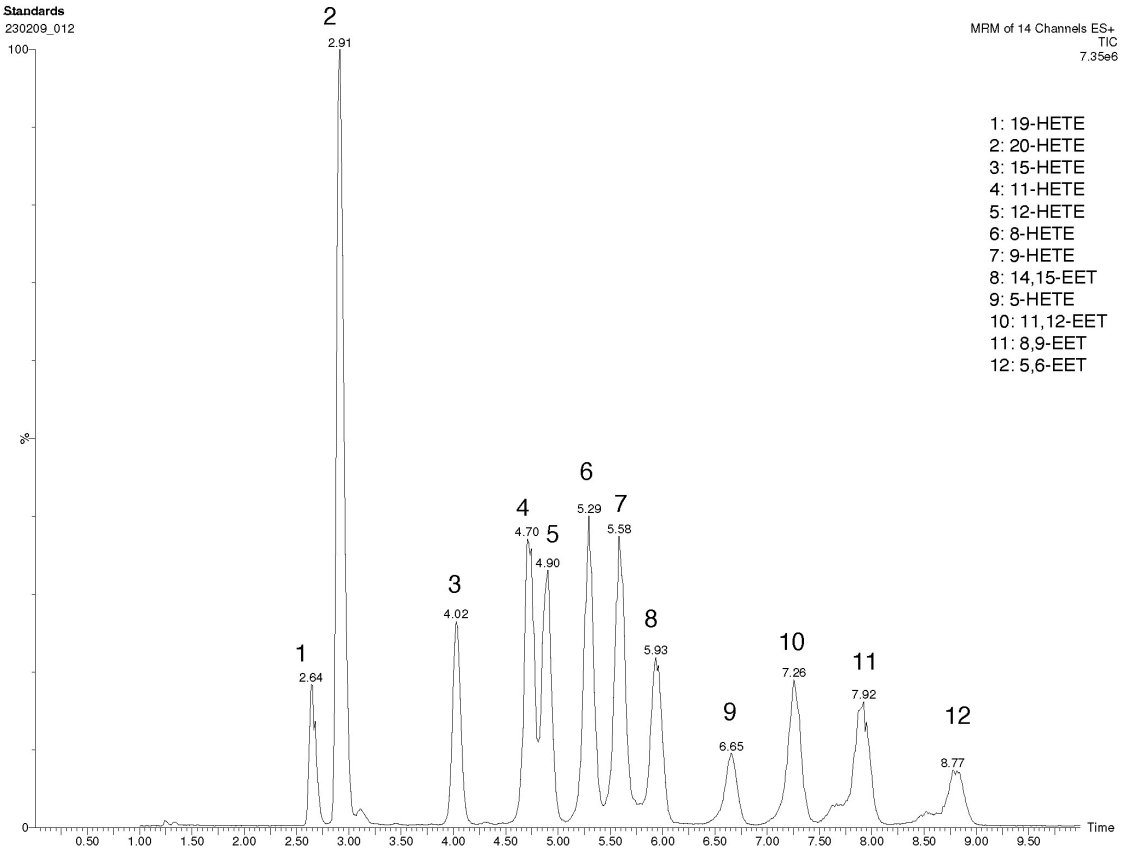
Supplementary Figure 2. HETE formation by ancestral and extant CYP4s

Incubations of the ancestral and extant CYP4s with AA were carried out and metabolites were analyzed with LC-MS. The data presented are the means +/- SD of three technical replicates for all CYPs and six technical replicates for CYP-free controls (data points collected on two separate days) .

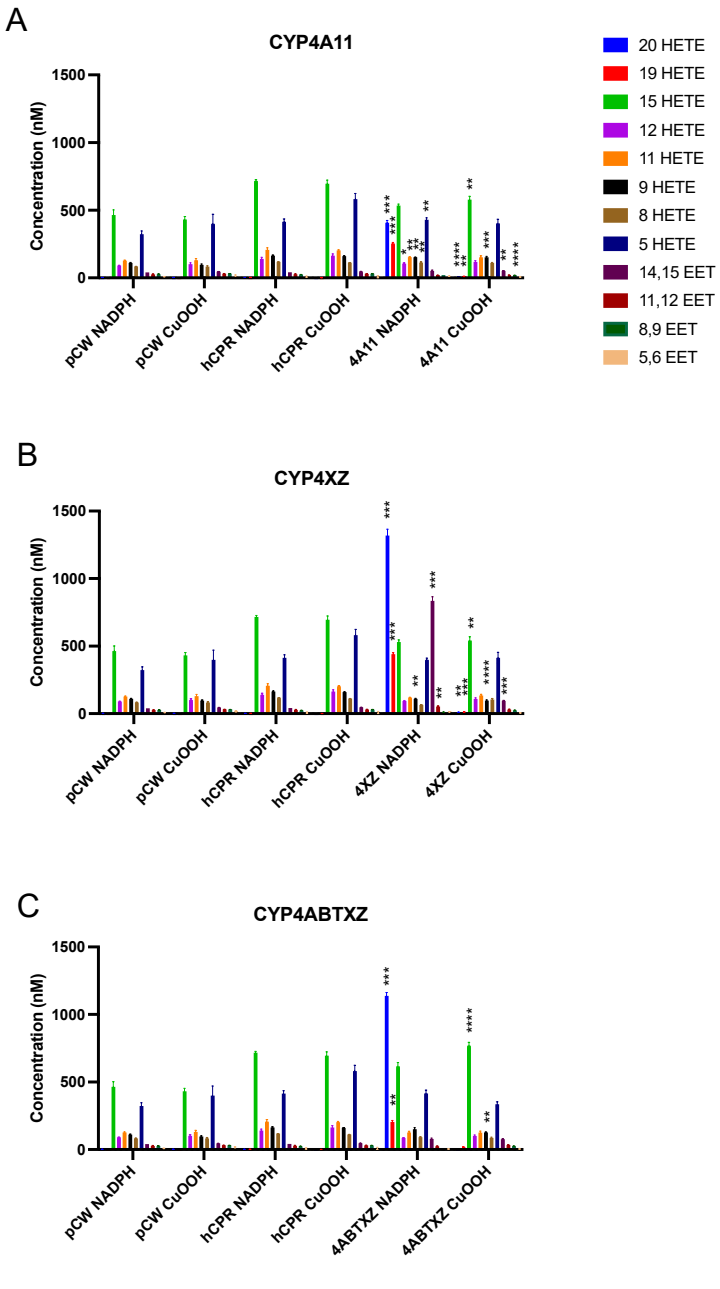


Supplementary Figure 3. EET formation by ancestral and extant CYP4s.

Incubations of the ancestral and extant CYP4s with AA were carried out and metabolites were analyzed with LC-MS. The data presented are the means +/- SD of three technical replicates for all CYPs and six technical replicates for CYP-free controls (data points collected on two separate days).

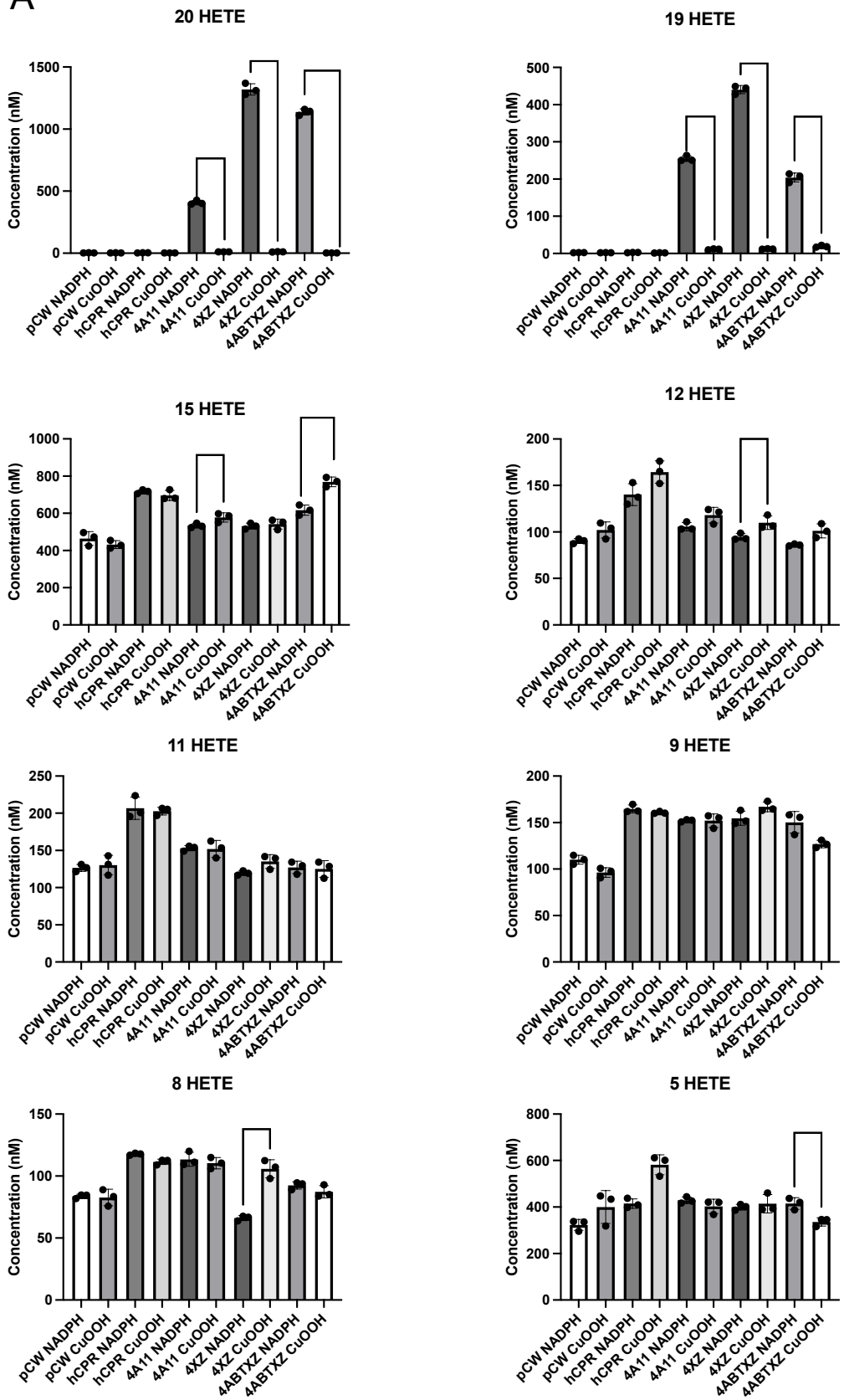


Supplementary Figure 4. Representative LC-MS chromatogram of AA metabolites in a standard showing chromatographic separation.



Supplementary Figure 5. AA oxidation by CYP4A11 (A), CYP4XZ (B),and CYP4ABTXZ (C). The data presented are the concentrations of 5-, 8-, 9-, 11-, 12-, 15-, 19-, and 20- HETE and 5,6-; 8,9-; 11,12-; and 14,15- EET from AA incubations in the presence of NADPH or CuOOH. Asterisks indicate statistical significance compared to respective NADPH or CuOOH pCW control for each metabolite using an unpaired t-test (*$p < 0.05$; ** $p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$).

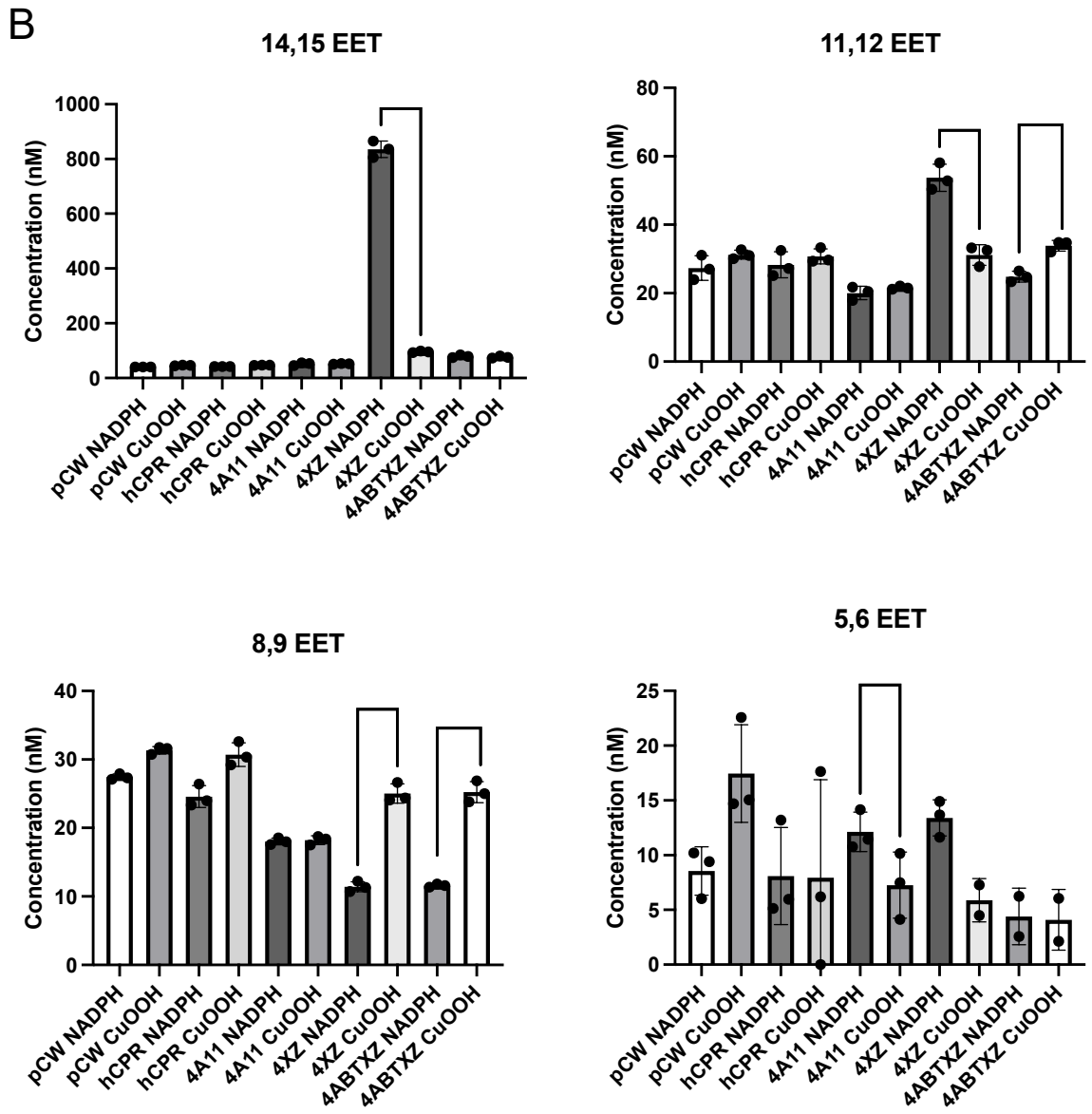
A


20 HETE
19 HETE
15 HETE
12 HETE
11 HETE
9 HETE
8 HETE
5 HETE
Concentration (nM)
pCW NADPH
pCW CuOOH
hCPR NADPH
hCPR CuOOH
4A11 NADPH
4A11 CuOOH
4XZ NADPH
4XZ CuOOH
4ABTXZ NADPH
4ABTXZ CuOOH

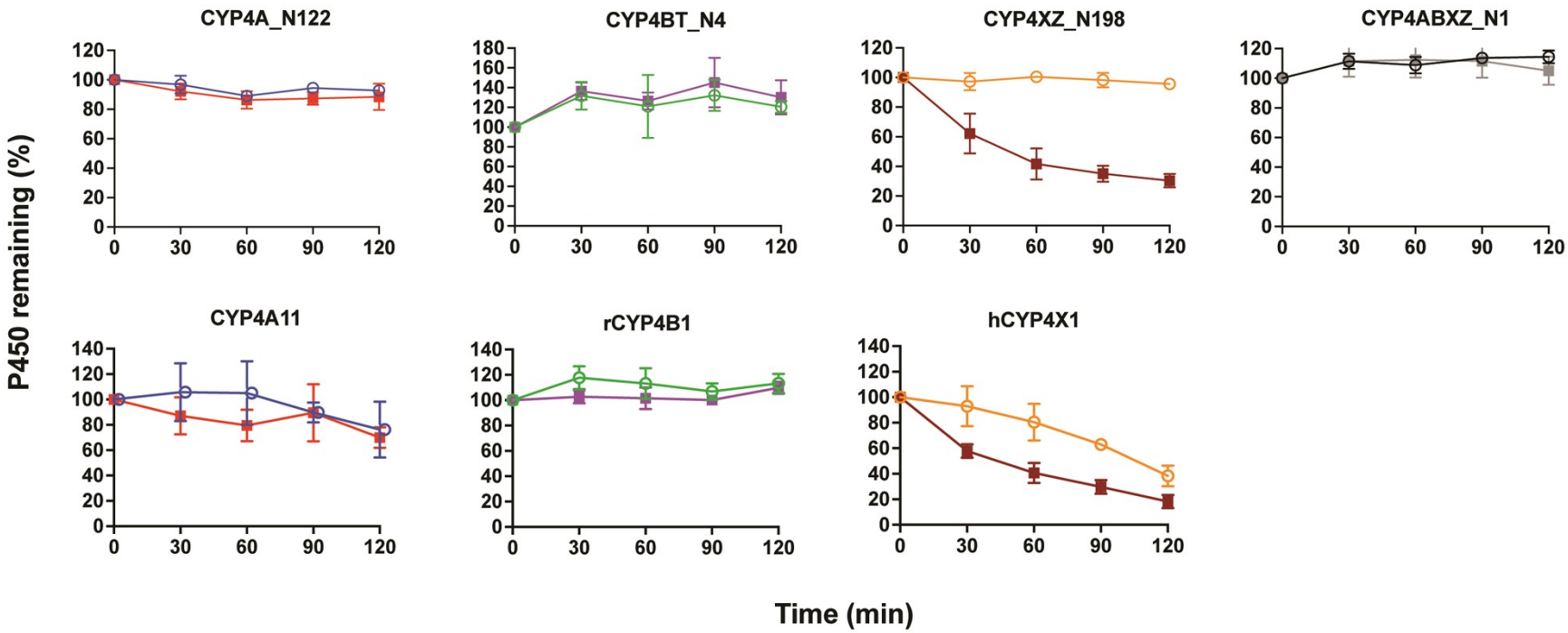
87

88

89



Supplementary Figure 6. Differences in AA metabolite formation between a standard NGS and O2 surrogate (CuOOH). HETE (A) and EET (B) AA oxidation products were quantified via LC-MS/MS. Asterisks indicate statistically significant differences in the amount of metabolite formed between NADPH and CUOOH incubation conditions using a one-tailed, t-test (*$p < 0.05$; ** $p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$).



Supplementary Figure 7. Stability of extant and ancestral P450s to incubation with CuOOH. Bacterial membrane preparations containing 0.6-1.0 µM of the P450s indicated were incubated with 500 µM CuOOH in 100 mM potassium phosphate buffer, pH 7.4. At the times indicated, aliquots of incubations were removed and the residual P450 hemoprotein concentration was quantified by Fe(II).CO *vs.* Fe(II) difference spectroscopy. Data represent the mean +/- SD of n=3 replicates.

References

1 Zheng, Y.-M., Fisher, M. B., Yokotani, N., Fujii-Kuriyama, Y. & Rettie, A. E. Identification of a meander region proline residue critical for heme binding to cytochrome P450: Implications for the catalytic function of human CYP4B1. *Biochemistry* **37**, 12847-12851, doi:10.1021/bi981280m (1998).

2 Hoch, U., Zhang, Z., Kroetz, D. L. & Ortiz de Montellano, P. R. Structural determination of the substrate specificities and regioselectivities of the rat and human fatty acid ω-hydroxylases. *Arch. Biochem. Biophys.* **373**, 63-71, doi:http://dx.doi.org/10.1006/abbi.1999.1504 (2000).

3 Helvig, C., Dishman, E. & Capdevila, J. H. Molecular, enzymatic, and regulatory characterization of rat kidney cytochromes P450 4A2 and 4A3. *Biochemistry* **37**, 12546-12558, doi:10.1021/bi981048g (1998).

4 Roman, L. J. *et al.* Expression of rabbit cytochromes P4504A which catalyze the ω-hydroxylation of arachidonic acid, fatty acids, and prostaglandins. *Arch. Biochem. Biophys.* **307**, 57-65, doi:https://doi.org/10.1006/abbi.1993.1560 (1993).

5 Palmer, C. N. *et al.* Characterization of a cDNA encoding a human kidney, cytochrome P-450 4A fatty acid omega-hydroxylase and the cognate enzyme expressed in Escherichia coli. *Biochim. Biophys. Acta* **1172**, 161-166 (1993).

6 Muller, Dominik N. *et al.* Mouse CYP4A isoforms: enzymatic properties, gender- and strain-specific expression, and role in renal 20-hydroxyeicosatetraenoic acid formation. *Biochem. J.* **403**, 109-118, doi:10.1042/bj20061328 (2007).

7 Xu, F., Falck, J. R., Ortiz de Montellano, P. R. & Kroetz, D. L. Catalytic Activity and Isoform-Specific Inhibition of Rat Cytochrome P450 4F Enzymes. *J. Pharmacol. Exp. Ther.* **308**, 887, doi:10.1124/jpet.103.059626 (2004).

8 Zöllner, A. *et al.* Human CYP4Z1 catalyzes the in-chain hydroxylation of lauric acid and myristic acid. *Biol. Chem.* **390**, 313, doi:10.1515/BC.2009.030 (2009).

9 Nakano, M., Kelly, E. J., Wiek, C., Hanenberg, H. & Rettie, A. E. CYP4V2 in Bietti's crystalline dystrophy: ocular localization, metabolism of ω-3-polyunsaturated fatty acids, and functional deficit of the p.H331P variant. *Mol. Pharmacol.* **82**, 679-686, doi:10.1124/mol.112.080085 (2012).

10 Ashkar, S. *et al.* Retinoic acid induces corneal epithelial CYP4B1 gene expression and stimulates the synthesis of inflammatory 12-hydroxyeicosanoids. *J. Ocul. Pharmacol. Ther.* **20**, 65-74, doi:10.1089/108076804772745473 (2004).

11 Vafeas, C. *et al.* Hypoxia stimulates the synthesis of cytochrome P450-derived inflammatory eicosanoids in rabbit corneal epithelium. *J. Pharmacol. Exp. Ther.* **287**, 903-910 (1998).

12 Mastyugin, V. *et al.* Hypoxia-induced production of 12-hydroxyeicosanoids in the corneal epithelium: involvement of a cytochrome P-450 4B1 isoform. *J. Pharmacol. Exp. Ther.* **289**, 1611-1619 (1999).

13 Powell, P. K., Wolf, I., Jin, R. & Lasker, J. M. Metabolism of arachidonic acid to 20-hydroxy-5,8,11,14-eicosatetraenoic acid by P450 Enzymes in human liver: Involvement of CYP4F2 and CYP4A11. *J. Pharmacol. Exp. Ther.* **285**, 1327-1336 (1998).

14 Fer, M. *et al.* Cytochromes P450 from family 4 are the main omega hydroxylating enzymes in humans: CYP4F3B is the prominent player in PUFA metabolism. *J. Lipid Res.* **49**, 2379-2389, doi:10.1194/jlr.M800199-JLR200 (2008).

15 Bylund, J., Hidestrand, M., Ingelman-Sundberg, M. & Oliw, E. H. Identification of CYP4F8 in human seminal vesicles as a prominent 19-hydroxylase of prostaglandin

endoperoxides. *J. Biol. Chem.* **275**, 21844-21849, doi:10.1074/jbc.M001712200 (2000).

16 Stark, K., Wongsud, B., Burman, R. & Oliw, E. H. Oxygenation of polyunsaturated long chain fatty acids by recombinant CYP4F8 and CYP4F12 and catalytic importance of Tyr-125 and Gly-328 of CYP4F8. *Arch Biochem Biophys* **441**, 174-181, doi:10.1016/j.abb.2005.07.003 (2005).

17 Bylund, J., Bylund, M. & Oliw, E. H. cDNA cloning and expression of CYP4F12, a novel human cytochrome P450. *Biochem. Biophys. Res. Comm.* **280**, 892-897, doi:https://doi.org/10.1006/bbrc.2000.4191 (2001).

18 Stark, K., Dostalek, M. & Guengerich, F. P. Expression and purification of orphan cytochrome P450 4X1 and oxidation of anandamide. *FEBS J.* **275**, 3706-3717, doi:10.1111/j.1742-4658.2008.06518.x (2008).

19 Kikuta, Y., Kasyu, H., Kusunose, E. & Kusunose, M. Expression and catalytic activity of mouse leukotriene B4 ω-hydroxylase, CYP4F14. *Arch. Biochem. Biophys.* **383**, 225-232, doi:https://doi.org/10.1006/abbi.2000.2078 (2000).

20 Kikuta, Y., Kusunose, E. & Kusunose, M. Characterization of human liver leukotriene B4 ω-hydroxylase P450 (CYP4F2). *J. Biochem.* **127**, 1047-1052, doi:10.1093/oxfordjournals.jbchem.a022696 (2000).

21 Bylund, J., Harder, A. G., Maier, K. G., Roman, R. J. & Harder, D. R. Leukotriene B4 ω-side chain hydroxylation by CYP4F5 and CYP4F6. *Arch. Biochem. Biophys.* **412**, 34-41, doi:https://doi.org/10.1016/S0003-9861(03)00030-4 (2003).

22 Christmas, P. *et al.* Alternative splicing determines the function of CYP4F3 by switching substrate specificity. *J. Biol. Chem.* **276**, 38166-38172, doi:10.1074/jbc.M104818200 (2001).

23 Christmas, P. *et al.* CYP4F18 is the LTB4 omega-1/omega-2 hydroxylase in mouse polymorphonuclear leukocytes: Identification as the functional orthologue of human PMN CYP4F3A in the down-regulation of responses to LTB4. *J. Biol. Chem.* (2005).

24 Kikuta, Y., Kusunose, E., Ito, M. & Kusunose, M. Purification and characterization of recombinant rat hepatic CYP4F1. *Arch Biochem Biophys* **369**, 193-196, doi:10.1006/abbi.1999.1271 (1999).

25 Songhee, H. A. N. *et al.* Self-sufficient catalytic system of human cytochrome P450 4A11 and NADPH-P450 reductase. *Biomol. Ther.* **17**, 156-161 (2009).

26 Cheesman, M. J., Baer, B. R., Zheng, Y.-M., Gillam, E. M. J. & Rettie, A. E. Rabbit CYP4B1 engineered for high-level expression in Escherichia coli: ligand stabilization and processing of the N-terminus and heme prosthetic group. *Arch. Biochem. Biophys.* **416**, 17-24, doi:http://dx.doi.org/10.1016/S0003-9861(03)00278-9 (2003).

27 Bellamine, A. *et al.* Characterization of the CYP4A11 gene, a second CYP4A gene in humans. *Arch. Biochem. Biophys.* **409**, 221-227, doi:https://doi.org/10.1016/S0003-9861(02)00545-3 (2003).

28 Gasser, R. & Philpot, R. M. Primary structures of cytochrome P-450 isozyme 5 from rabbit and rat and regulation of species-dependent expression and induction in lung and liver: Identification of cytochrome P-450 gene subfamily IVB. *Mol. Pharmacol.* **35**, 617-625 (1989).

29 Strausberg, R. L. *et al.* Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 16899-16903, doi:10.1073/pnas.242603899 (2002).

30 Gouet, P. & Robert, X. Deciphering key features in protein structures with the new ENDscript server. *Nucl. Acids Res.* **42**, W320-W324, doi:10.1093/nar/gku316 (2014).