

Supplementary Material

Alternative oxidase gene family in *Hypericum perforatum* L.: characterization and expression at the post-germinative phase

Isabel Velada^{1†}, Hélia G. Cardoso^{1†*}, Carla Ragonezi¹, Amaia Nogales², Alexandre Ferreira¹, Vera Valadas¹
and Birgit Arnholdt-Schmitt^{1*}

¹EU Marie Curie Chair, ICAAM - Instituto de Ciências Agrárias e Ambientais Mediterrânicas, Universidade de Évora, Pólo da Mitra, Ap. 94, 7002-554 Évora, Portugal, ²Linking Landscape, Environment, Agriculture and Food, Instituto Superior de Agronomia-Universidade de Lisboa, Lisboa, Portugal

[†]These authors have contributed equally to this work.

* **Correspondence:** Corresponding Authors: eu_chair@uevora.pt and hcardoso@uevora.pt

1 Supplementary Figures and Tables

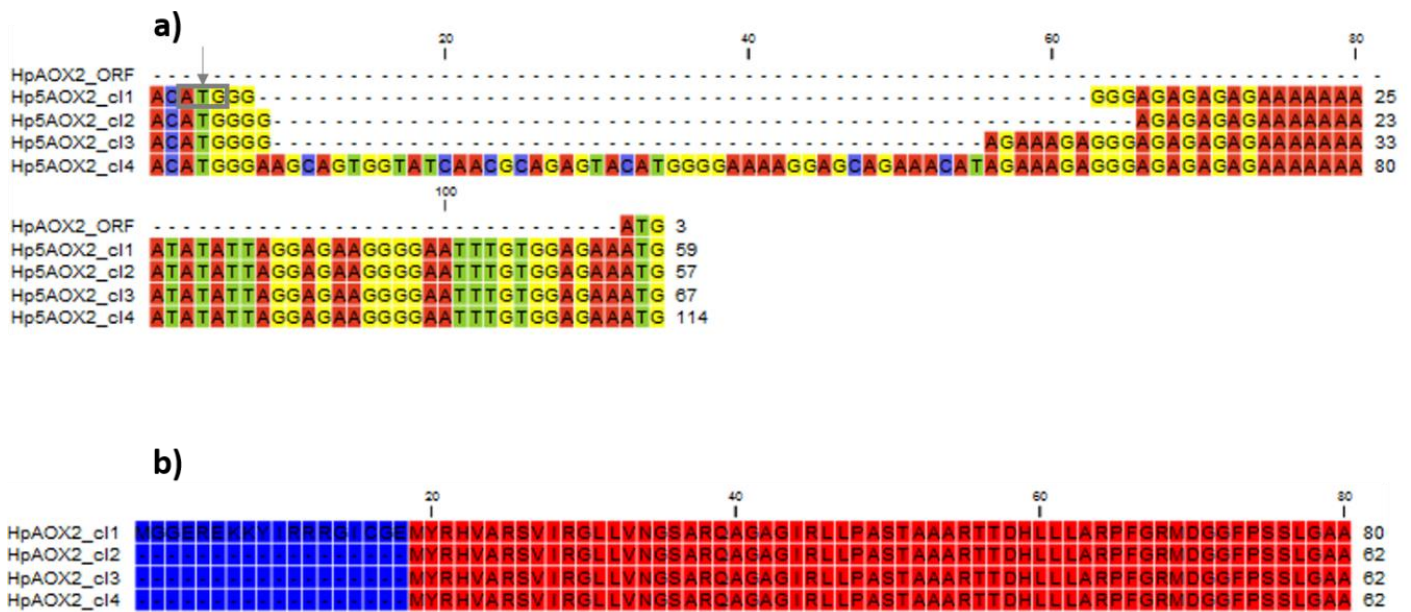


Figure S2. Alignment of the four isolated sequences corresponding to 5' end of *HpAOX2* gene. The sequences correspond to the region from the primer adapter until ATG codon (included in the alignment). **a)** Alignment of the nucleotide sequences; **b)** Alignment of the corresponding putative translated peptide where an alternative ATG codon was identified in the sequence HpAOX2_c11 (region in blue).

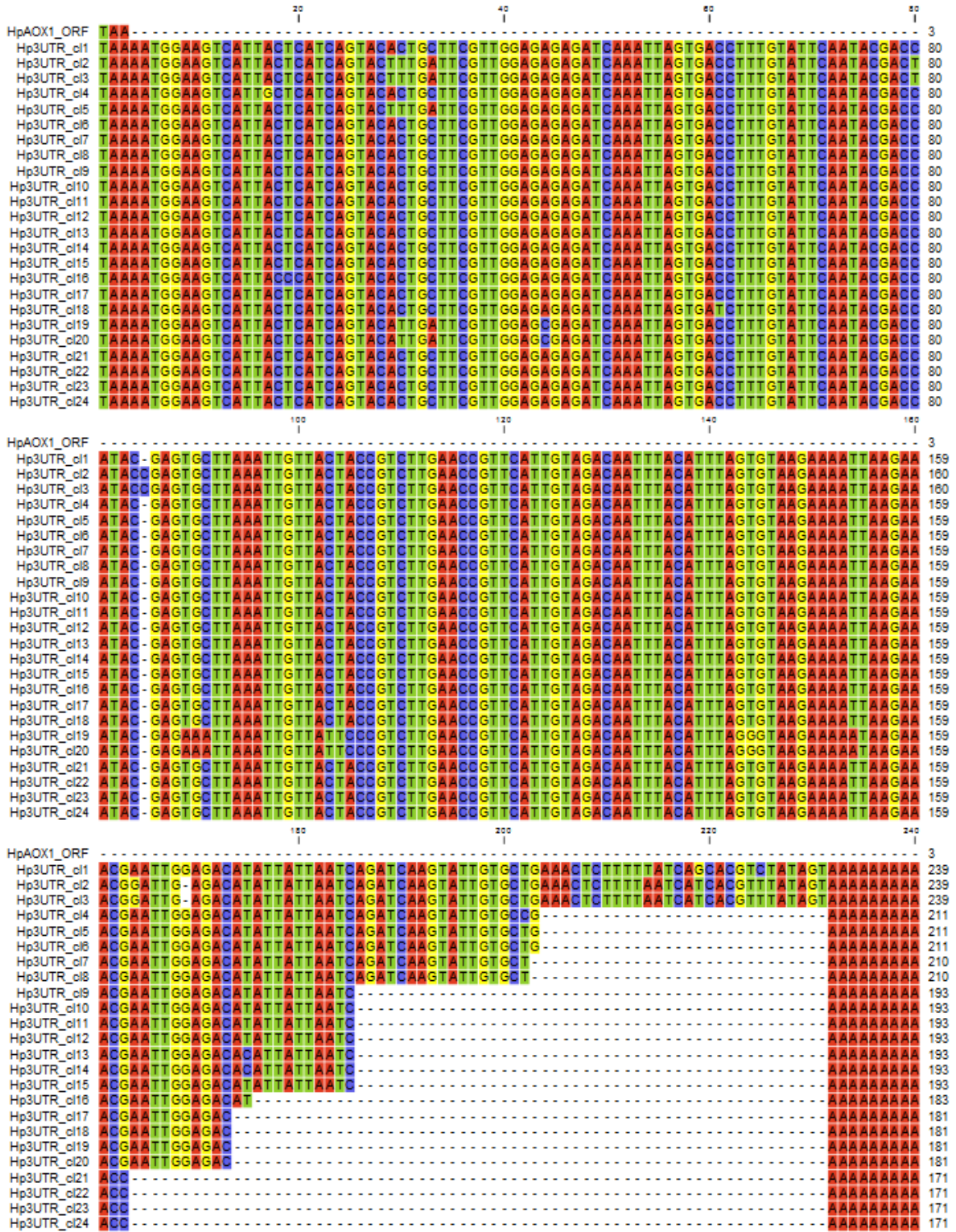


Figure S3. Alignment of the twenty four isolated sequences corresponding to 3'end of *HpAOX1* gene. The sequences correspond to the region from the stop codon (TAA, included in the alignment) to the poly(A)-tail (only nine common adenines are represented).

ACATGGGGATCCCCAAGAAACACGAAGCAAAAAAAAAACCCTAGAAGACAAAACCTAACTGCTCTCCTAAAAAG
AAATTAAG 80
TGAAAACAAAGAAATTGTAGAGATTATTGAATCGAGATGATAAGCCGCGGGCTACTAGGCTGGCGAGTTCG
GCGGTTGC 160

M I S R G G T R L A S S A V A

GGCGGCCGCTGCCAATGCTCGGGCCTTCTCGACTGCCGGAGCAACGGAGAAGTCTGTTCTCAGATACTATGC
TACTCCGG 240
A A A A N A R R F S T A G A T E K S V L R Y Y A T P

CAAAGGGCGGAAACAGTGCGGTGTCCCGTGGGCGAGGGGCCGATCTTGACGAGGAGTGCAACGACGATG
AGCTTGGGA 320
A K G G N S A V S P W A R G P I L T R S A T T M S L G

GAGAAGCAAGTGAAGGAGGAGGCTGAAGAGAATAAGTTCTTCGGAGCGTCCACCGGCGCCGGAGGGAA
CAAGGATGG 400
E K Q V K E E A E E N K V P S E R S T G A G G N K D G

CGGCGATGAGAAGGCCGTTGTTAGCTACTGGGGTATTAACCCCGGGAAGATCACTAAGGAAGACGGCACCGT
TTGGACCT 480
G D E K A V V S Y W G I N P G K I T K E D G T V W T

GGACTTGCTTTAGGCCATGGGAAGCATACCAACCAAACATGTTCGATTGATCTGAAGAAGCATCACGCTCCGA
CTACTTTC 560
W T C F R P W E A Y Q P N M S I D L K K H H A P T T F

TTGGACAAACTGGCTTTTCGGACCGTTAAGTCTCTCCGATGGCCACCGACATATTCTTTCAGAGGAGGTATG
GTTGTTCG 640
L D K L A F R T V K S L R W P T D I F F Q R R Y G C R

GGCAATGATGCTCGAAACAGTGGCCGCGGTGCCAGGCATGGTAGGAGGAATGCTCCTCCACTGCAAGTCACT
GAGGCGCT 720
A M M L E T V A A V P G M V G G M L L H C K S L R R

TCGAGCACAGCGGAGGATGGATCAAAACCCTACTGGACGAGGCAGAGAACGAGCGGATGCACTTGATGACA
TTCATGGAG 800
F E H S G G W I K T L L D E A E N E R M H L M T F M E

GTTGCCAAGCCCAAATGGTACGAGCGTGCTCTGGTCTTACAGTGCAGGGCGCCTTCTTCAACATCTACTTTC
TGGGGTA 880
V A K P K W Y E R A L V F T V Q G A F F N I Y F L G Y

TCTCATTTCGCCAAGTTCGCGCATAGGGTGGTTGGATACCTCGAGGAAGAAGCAATCCACTCCTACACCGA
GTTTCTCA 960
L I S P K F A H R V V G Y L E E E A I H S Y T E F L

AGGAGTTGGACAATGGCAACATCGAGAATGTGCCTGCTCCTGCAATTGCTGTCTGACTACTGGCGCCTACCTC
CCAACCTCC 1040
K E L D N G N I E N V P A P A I A V D Y W R L P P N S

ACTCTAAAGGATGTGGTTCAGGTTGTCAGGGCCGACGAGGCTCATCACCGCGACGTCAATCATTTCATCG
GACATACA 1120
T L K D V V Q V V R A D E A H H R D V N H F A S D I H

CTACCAAGGACGTGAGCTGAGGGAATGCCCTGCTCCTGCTGGGTTATCACTAAAATGGAAGTCATTGCTCATC
AGTACACT 1200

Y Q G R E L R E C P A P L G Y H *

GCTTCGTTGGAGAGAGATCAAATTAGTGACCTTTGTATTCAATACGACCATACGAGTGCTTAAATTGTTACTA
CCGTCTT 1280

GAACCGTTCATTGTAGACAATTTACATTTAGTGTAAGAAAATTAAGAAACGAATTGGAGACATATTATTAATC
AGATCAA 1360

GTATTGTGCTGAAACTCTTTTAATCATCACGTTTATAGTAAAAAAAAAAAAAAAAAAAAA 1415

Figure S4. Nucleotide and deduced amino acid sequences of cDNA encoding *Hypericum perforatum* L. alternative oxidase 1 (*HpAOX1*, acc. no. EU330415.1). ▼ indicates the position of the three introns, * indicates stop codon.

ACATGGGAAGCAGTGGTATCAACGCAGAGTACATGGGGAAAAGGAGCAGAAACATAGAAAAGAGGGAGAGAG
AGAAAAAAA 80

ATATATTAGGAGAAGGGGAATTTGTGGAGAAATGTACCGCCACGTGGCAAGGTCAGTGATCAGGGGACTCCT
GGTCAACG 160

M Y R H V A R S V I R G L L V N

GGAGTGCCCGCCAGGCTGGCGCCGGGATCCGGCTGCTTCCAGCTTCCACAGCCGCCGCGCGGACGACCGAC
CACCTCCTC 240

G S A R Q A G A G I R L L P A S T A A A R T T D H L L

CTCGCGAGGCCTTTCGGCCGGATGGACGGTGGATTCCCGTCTCGCTCGGCGCGGGCAGTTACTTGAGGATG
TTGAGCAG 320

L A R P F G R M D G G F P S S L G A A S Y L R M L S S

CTCATCGTCGTCGGAGGGGAAGCCGGTGGCGGAGGAGGAGAAGAAGGCGGTGGCGTCCAGTTACTGGGGAA
TCTCGAGGC 400

S S S S E G K P V A E E E K K A V A S S Y W G I S R

CGAAGGTCACTAGGGAGGACGGCACCGACTGGCCTTGGAACTGCTTCATGCCTTGGGAGTCTTATAGGACAG
ATACGTCA 480

P K V T R E D G T D W P W N C F M P W E S Y R T D T S

ATAGATCTGACGAAGCACCACGTTCCAAAGACATTTCTTGACAAGTTTGCTTACTGGACGGTCAAGATCCTTC
GAATTCC 560

I D L T K H H V P K T F L D K F A Y W T V K I L R I P

TACAGATATATTCTTTCAGAGACGGTACGGCTGCCGCGCAATGATGTTGGAAACAGTTGCAGCTGTACCGGG
TATGGTTG 640

T D I F F Q R R Y G C R A M M L E T V A A V P G M V

GTGGGATGCTTCTGCACTTGAGGTCGCTCCGCAAGTTCGAGCAAACCTGGGGGTTGGGTCAAGGCCTTGCTCG
AAGAAGCA 720

G G M L L H L R S L R K F E Q T G G W V K A L L E E A

GAGAACGAAAGAATGCACTTGATGACAATGGTAGAGCTCGTTCAGCCAAAATGGTACGAGAGGATGCTGGTT
CTTATCGT 800

E N E R M H L M T M V E L V Q P K W Y E R M L V L I V

GCAGGGGCTGTTCTTCAACGCCTACTTTGTGATCTATCTCGTGTCCCGAAGCTCGCTCACAGGGTAACAGG
CTATTTGG 880

Q G L F F N A Y F V I Y L V S P K L A H R V T G Y L

AAGAAGAGGCCATTTCATATACCGAGTACCTCAAAGACATAAGGAGTGGGAAAATCGAGAATGTCAAGG
CTCCCGCT 960

E E E A I H S Y T E Y L K D I R S G K I E N V K A P A

ATTGCGATCGACTACTGGAGGTTGCCAAGGAAGCTACACTCGAGGATGTCATAACTGTGATCCGAGCCGAT
GAAGCACA 1040

I A I D Y W R L P K E A T L E D V I T V I R A D E A H

TCACCGTGATGTCAACCACTTTGCTTCTGATGTTCACTATCAGGGCAAGAACTGAAAGAGGCACCTGCTCCT
ATTGGCT 1120

H R D V N H F A S D V H Y Q G K K L K E A P A P I G

ATCATTGAGGGGTTTTTTTTTTTTTTAACTTATGTTTTTCAGAGTCAAGTAATATCGAAATTTATTCGAGACTT
ATTTAG 1200

Y H *

TTCTTACTTCAGCTTGTGTTGTATATGCTATGTATTGTAAACGTATATTATTGTAAAATATAAGGCGCATATAA
TATACA 1280
TTGTTTTGTTGACATAAAAAAAAAAAAAAAAAAAAA 1311

Figure S5. Nucleotide and deduced amino acid sequences of cDNA encoding *Hypericum perforatum* L. alternative oxidase 2 (*HpAOX2*, acc. no. EU330413.1). ▼ indicates the position of the three introns, * indicates stop codon.

		20		40		60								
EU330415.1	MIS	--	RG	TRL	ASS	AVAAAAAN	ARRF	STAGATEKSVLRY	YATPAKGGNSAVSPWARGP	IL	--	TR	SATTM	65
AT3G22370	MMIT	--	RG	GAKAAKSL	--	LVAAGPRLF	STVTRVSSHEAL	SASH	LKPGVTS	SAWI	WTRAPT	IGGMR	EASTI	66
AT3G22360	MMMS	--	RR	--	YG	AKLM	--	ETAVTHS	-----	HLLN	PRVPLVTEN	IRVPAM	GVVRF	SKM
AT3G27620	MIT	TL	LRSL	LD	ASKQ	--	ATSING	ILF	-----	HQLAP	----	AKY	FRVPAV	GGLRD
AT1G32350	MSYR	SI	YR	TL	RPVLS	SSVQS	-----	-----	-----	-----	-----	-----	-----	34
TC03G031300	MMS	--	RG	TRL	ASSM	--	LAFV	GP	PRY	F	STA	-----	SARTLS	NEGM
EU330413.1	MY	--	RH	VARS	IR	GLLVNG	-----	SAR	QA	GAG	IRLL	PASTAAART	TDHLL	LARPF
AT5G64210	M	--	SQL	IT	KAAAL	RVLV	CGR	--	GNC	NMF	VSSV	SST	VMKSP	YETAP
CP00042G00490	M	--	NR	FVAR	S	IRAL	MVN	-----	SR	SCGS	NRML	YAATTA	TVP	PWE
ME10292G00060	MMN	NL	V	IR	SV	RGLLV	NGSR	GGRY	I	S	TATA	ATTAG	V	LK
RC30063G00030	M	--	NR	VV	IR	SV	LR	GLQV	NGG	--	GGRC	I	STA	AS
TC02G011670	M	--	NR	FL	V	R	S	V	M	QGL	--	IN	GR	-----
				80		100		120		140				
EU330415.1	S	L	G	E	K	Q	V	K	E	E	-----	A	E	N
AT3G22370	T	L	G	E	K	T	P	M	K	E	-----	E	D	A
AT3G22360	T	F	E	K	K	-----	-----	-----	-----	-----	-----	K	T	T
AT3G27620	T	F	E	K	K	-----	-----	-----	-----	-----	-----	K	T	T
AT1G32350	H	L	P	N	V	R	L	L	S	-----	S	D	T	S
TC03G031300	A	L	G	D	K	-----	-----	-----	-----	-----	-----	E	Q	E
EU330413.1	S	-----	Y	L	R	M	L	S	-----	S	S	S	S	E
AT5G64210	N	L	R	W	M	G	M	S	A	S	A	M	-----	E
CP00042G00490	N	-----	W	R	R	M	M	S	-----	S	Q	E	S	S
ME10292G00060	Y	-----	W	R	R	M	M	S	-----	S	T	S	A	E
RC30063G00030	C	-----	W	R	R	M	M	S	-----	S	T	S	Q	L
TC02G011670	E	-----	W	R	R	M	M	N	-----	S	A	P	-----	-----
				160		180		200						
EU330415.1	W	T	W	T	C	F	R	P	W	E	A	Y	Q	P
AT3G22370	W	K	W	C	F	R	P	W	E	T	Y	K	A	D
AT3G22360	W	K	W	C	F	R	P	W	E	T	Y	K	S	D
AT3G27620	W	K	W	C	F	R	P	W	E	T	Y	K	A	D
AT1G32350	W	K	W	C	F	R	P	W	E	S	Y	K	A	D
TC03G031300	W	K	W	T	C	F	R	P	W	E	A	Y	K	A
EU330413.1	W	P	W	N	C	F	R	P	W	E	S	Y	R	T
AT5G64210	W	P	W	N	C	F	R	P	W	E	T	Y	Q	A
CP00042G00490	W	P	W	N	C	F	R	P	W	E	A	Y	G	D
ME10292G00060	W	P	W	N	C	F	R	P	W	E	T	Y	R	A
RC30063G00030	W	P	W	N	C	F	R	P	W	E	T	Y	Q	S
TC02G011670	W	P	W	N	C	F	R	P	W	E	T	Y	K	A
				220		240		260		280				
EU330415.1	G	G	M	L	L	H	C	K	S	L	R	R	F	E
AT3G22370	G	G	M	L	L	H	C	K	S	L	R	R	F	E
AT3G22360	G	G	M	L	V	H	C	K	S	L	R	R	F	E
AT3G27620	G	G	M	L	M	H	F	K	S	L	R	R	F	E
AT1G32350	G	G	M	L	L	H	L	K	S	L	R	R	F	E
TC03G031300	G	G	M	L	L	H	C	K	S	L	R	R	F	E
EU330413.1	G	G	M	L	L	H	L	R	S	L	R	K	F	E
AT5G64210	G	G	M	L	L	H	L	K	S	L	R	K	F	E
CP00042G00490	G	G	M	L	L	H	L	K	S	L	R	K	F	E
ME10292G00060	A	G	M	L	L	H	L	R	S	L	R	K	F	E
RC30063G00030	G	G	M	L	L	H	L	R	S	L	R	K	F	E
TC02G011670	G	G	M	L	L	H	L	R	S	L	R	K	F	E
				300		320		340						
EU330415.1	K	F	A	H	R	V	V	G	L	E	E	E	A	I
AT3G22370	K	F	A	H	R	M	V	G	L	E	E	E	A	I
AT3G22360	K	F	A	H	R	M	V	G	L	E	E	E	A	I
AT3G27620	K	F	A	H	R	M	V	G	L	E	E	E	A	I
AT1G32350	K	L	A	H	R	I	T	G	Y	L	E	E	E	A
TC03G031300	K	F	A	H	R	M	V	G	L	E	E	E	A	I
EU330413.1	K	L	A	H	R	V	T	G	Y	L	E	E	E	A
AT5G64210	R	L	A	H	R	V	V	G	L	E	E	E	A	I
CP00042G00490	K	L	A	H	R	I	V	G	Y	L	E	E	E	A
ME10292G00060	K	L	A	H	R	I	T	G	Y	L	E	E	E	A
RC30063G00030	K	L	A	H	R	I	T	G	Y	L	E	E	E	A
TC02G011670	K	L	A	H	R	I	V	G	Y	L	E	E	E	A
				360										
EU330415.1	F	A	S	D	I	H	Y	Q	G	R	E	L	R	E
AT3G22370	F	A	S	D	I	H	Y	Q	G	R	E	L	R	E
AT3G22360	Y	A	S	D	I	H	Y	Q	G	R	E	L	R	E
AT3G27620	Y	A	S	D	I	H	Y	Q	G	H	E	L	R	E
AT1G32350	Y	A	S	D	I	Q	F	K	G	H	E	L	R	E
TC03G031300	F	A	S	D	I	H	Y	Q	G	R	Q	L	R	E
EU330413.1	F	A	S	D	V	H	Y	Q	G	K	K	L	R	E
AT5G64210	F	A	S	D	I	R	N	Q	G	K	E	L	R	E
CP00042G00490	F	A	S	D	I	H	Y	Q	G	K	E	L	R	E
ME10292G00060	F	A	S	D	I	H	Y	Q	G	K	E	L	R	E
RC30063G00030	F	A	S	D	I	H	Y	Q	G	K	E	L	R	E
TC02G011670	F	A	S	D	I	H	F	Q	G	K	E	L	R	E

Figure S6. Multiple alignment of translated amino acid sequences of previously reported AOX proteins from, *Arabidopsis thaliana* (AtAOX1a_AT3G22370, AtAOX1b_AT3G22360, AtAOX1c_AT3G27620, AtAOX1d_AT1G32350 and AtAOX2_AT5G64210), *Theobroma cacao* (TcAOX1_TC03G031300 and TcAOX2_TC02G011670) *Carica papaya* (CpAOX2_CP00042G00490), *Manihot esculenta* (MeAOX2_ME10292G00060), *Ricinus communis* (RcAOX2_RC30063G00030) and AOX from *Hypericum perforatum* (HpAOX1_EU330415.2 and HpAOX2_EU330413.2). The alignment was performed using CLC Main Workbench 6.7.1 software. The data were retrieved from public web-based database Plaza v2.5, freely available at http://bioinformatics.psb.ugent.be/plaza/versions/plaza_v2_5/. Amino acids residues differing are shown in red, deletions are shown by minus signs. The putative mitochondrial transit peptides (mTP) are shown in blue boxes. The sites of two conserved cysteins (CysI and CysII) involved in dimerization of the AOX protein by S–S bond formation (Umbach and Siedow, 1993) are indicated in dark grey boxes. Helices $\alpha 1$ and $\alpha 4$, which form the hydrophobic region on the AOX molecular surface and are involved in membrane binding, are shown in red; helices $\alpha 2$, $\alpha 3$, $\alpha 5$ and $\alpha 6$, which form the four-helix bundle accommodating the diiron center, are shown in green (Moore et al. 2013). Amino acids residues that coordinate the diiron center (E, glutamate and H, histidine) and those that interact with the inhibitor are in yellow and light pink boxes, respectively.

ATGATAAGC CGCGGCGGTACTAGGCTGGCGAGTTCGGCGGTTGCGGCGGCCGCTGCCAATGCTCGGTGCT
TCTCGACGGCCGGAGCCACGGAGAGGTCTGTTCTCAGATACCATGCTACTCCGGCGAAGGGCGGAAACAG
TGCGGTGTCGCCGTGGGCGAGGGGCCGATCTTGACGAGGAGTGTAACGACGATGAGCTTGGGAGAGAAG
CAAGTGAAGGAGGAGGCTGAAGAGAATAAGGTTCTTCGGAGCGTTCACCGGCGCCGGAGGGAACAAGG
ATGGCGGCGATGAGAAGGCCGTTGTTAGCTACTGGGGTATTAACCCGGGAAGATCACTAAGGAAGACGG
CACCGTTTGGACGTGGATCTGTTTTAGGTAAGCTTGATTTTATTTTATTTTTCTTATTTTTTTTTGGT
AGTAATATTTTATTTTTTCTTATTTCTTGACCTTCGTCAATAGAGATTTTGACTTCCTTTTGTACGC
GGCGTTTGAATCTTTGACTTTGTTATTTAGGGATTTTTTTTGTGTGGATTAATTTATATTTAATTGATAT
GATTTAGATGTTCTAATATATTAGGGAATTTTTGTTTTGTGATGTATATTTAATTAATAATGATTGGTA
AATGTGAAGCCATGGGAAGCATACCAACCAACATGTCGATTGATCTGAAGAAGCATCACGCTCCGACTA
CTTTCTTGGACAAACTGGCTTTTTGGACCGTTAAGTCTCTCCGATGGCCACCGACATATCTTTTCAGGT
AATATTTAATTTCTTCAATATTATTTTTTCGCATATCTTCTTAGGTTGAGATTACTACTGGGAGACTT
TTTTACTGAGAAAAATCTAAATAACATTGTTATAAAATTTTGGAGTGAACATGAAAATTTGAGTTTGGTA
TATCTTGGACATCGGTAGTATGTTACATTGTGACTTATTCAATGTTACAATTACCATCAAAAAAAAAA
TTAGGGTAAATTTCACTTGGTCATTGACCAGATTGTATAAAGAGACTCTAACACTACAATATGTATGTAT
ATATTGAATAATATATATAATTTCAACCCAACAAAAATATATTGAACCCTTCTATAATTTGTTGTGTTGC
ATCTGACGCCATTTTCTGTTCTGTTTTTTTTGGTGTAATCCTGGTTCTGTGTTGTTTACTAACTT
TCCCCAAAAAATTGTAAATGTGAATTAATGAGAAAAATTCTCGTTAAGTCATCGAAAGAAGGATTTTTTT
TAAATCTTCTTCAAGAGGAGATCTATATTATTAGAAAGAGGAGGACAATGGGTGGCAGCTTAATGGT
TCCTGGCGCCACCTTTTCTTGCCCTCTTTTACGTGAATCACTTGTCCATCAGTTGAGAACAGAAGGCTTT
GCTGTACCATGTAAGTCTCTAGTTAATGCAAAAAATTCGAGAACTTTGGGTTGAATATGAAACCAAGT
TATACTTAGATGAGAATCTGGTTTATTGTTTGTGTTTTATCACAACTCTATGGACTTAAAAGACTTCCA
TTTAGAGTATGATTGTAGATAATTCCTTCTGATTAATCTCTGGCATTATGTCATGCAGAGGAGGTATGG
TTGTGCGGCAATGATGCTCGAAACAGTGGCCGCGGTGCCAGGCATGGTAGGAGGAATGCTCCTCCACTGC
AAGTCACTGAGGCGCTTCGAGCACAGCGGAGGATGGATCAAACCCCTACTGGACGAGGCAGAGAACGAGC
GGATGCACTTGATGACATTCATGGAGGTTGCCAAGCCCAAATGGTACGAGCGTGCTCTGGTCTTTCACAGT
GCAGGGCGCCTTCTTCAACATCTACTTTCTGGGGTATCTCATTTCCCCAAGTTTCGCGCATAGGGTGGT
GGATACCTCGAGGAAGAAGCAATCCACTCCTACACCGAGTTTCTCAAGGAGTTGGACAATGGCAACATCG
AGAATGTGCTGCTCCTGCAATTGCTGTGCTGACTACTGGCGCTACCTCCCAACTCCACTCTAAAGGATGT
GGTTCAGGTTGTGAGGGCCGACGAGGCTCATCACCGCGACGTCAATCATTTTGCATCGGTAATCAAGTT
CCCTTGTTTTGTTTTCCCTAACCAACCTGTGGTTTCTGCAGATTCTCGATGTGATACAAACGGTTGATAC
TATCGTCTTAAATGCAGGACATACACTACCAAGGACGTGAGCTGAGGGAATGCCCTGCTCCACTGGGTTA
TCACTAA

Figure S7. Genomic sequence of *Hypericum perforatum* L. alternative oxidase 1 (HpAOX1, acc. no. KU674355). Exons are in yellow, stop codon in pink, repetitive sequences in red, pre-miRNA sequences in grey. In blue there is the second predicted pre-miRNA sequence which is two base pairs shorter and was predicted as in reverse sequence.

ATGTACCGCCACGTGGCAAGGTCAGTGATCAGGGGACTCCTGGTCAACGGGAGTGCCCGCCAGGCTGGCC
CCGGGATCCGGCTGCTTCCAGCTTCCACAGCCGCCGCGCGGACGACCGACCACCTCCTCCTCGCGAGGCC
TTTCGGCCGGATGGACGGTGGATTCCCGTCTCGCTCGGCGCGGCGAGTACTTGAGGATGTTGAGCAGC
TCATCGTCGTCCGAGGGGAAGCCGGTGGCGGAGGAGGAGAAGAAGGCCGGTGGCGTCCAGTACTGGGGAA
TCTCGAGGCCGAAGGTCAC TAGGGAGGACGGCACCGACTGGCCTTGGAAC TGCTTCATGGTATTAACATA
TCTTAGAAAAGCTTGATACGATGTGATTAATGATT TGATTGTTATTAATCGATAATTAACAAGTTTAGCGT
GGTTAGCCTCGGATAATTTAGAAATGGATAATCTGAGAAGTTGTTAGGGTGTGTGAACGAGGACAAAACA
CGCTGGAAAAAAGTTATCTAGATTGTGTAGACACTCGGT CAGATCGCGGGATGTTACAGTATTAATAGT
AGAAATTTGAATTACATGCTACAAGCTTTACATGTTGTTTCGTTACGCAGTAGATCAGGTT CATAGCAGA
ATGCAGAAATCTTATTAGTTCAAT TCCGTGCTTGTA TGAATAACGCATGCGATCAGC CGGACAATTATAT
TTCGAGTTCAACAAGTATGGGAATGATGAAGTCTGTACTAGT TCAAAACTGCTTTATCAGATGTTAAACA
ACTTTTTTAGTTACGTGGAAATTTCTTATTAGAAA TGAACGAATTAAGTACATGGATGTGGATTTCTTT
CATTGTCTGAGACAGTTGGCCTACGATCTTGACGAGTCTTAGAGTATCGGGTTTGTCTAAATCTTTGT
GGCTACTTTGCCTGCTAAACCAATATGTAGCTTATATGAATTGCTCGTTAACTAAGTGTATGTGTTTTG
TGGGGAACTTTGATTAAGAAGTGGAGTTGATTGAAGTTATGTTTATGGGTGTTGGTAGCCTTGGGAGTCT
TATAGGACAGATACGTCAATAGATCTGACGAAGCACCACGTTCCAAAGACATTTCTTGACAAGTTTGCTT
ACTGGACGGTCAAGATCCTTCGAATTCCTACAGATA TATTCTTTCAAGT CAGGTTGTTTAAATCTTCATG
GGTTGT TCAAAAATTTGGCATGAGTATTGGTGGTCTGGATAATCAATACAGAGCAAAGACTGTTTTCGATG
ATTGTCGTTAACTTGTGCCGCTGAATTAGT TTTGATCATCCACTGCAAAGTTTACTACACTTTGTTCCCTC
TTGTCAAAGTTTTCACAAATCTTGTGAATCGTGTCTCAGAGACGGTACGGCTGCCGCGCAATGATGTTGG
AAACAGTTGCAGCTGTACCGGGTATGGTTGGTGGGATGCTTCTGCACCTGAGGTCACTCCGCAAGTTCGA
GCAAACCGGGGGTTGGGTCAAGGCCTTGCTCGAAGAAGC CGAGAACGAAAGAATGCACTTGATGACAATG
GTAGAGCTCGTT CAGCCAAAGTGGTACGAGAGGATGCTGGTTCTTATCGTGCAGGGGCTGTTCTTCAACG
CCTACTTTGTGATCTATCTCGTGTCCC GAAGCTCGCTCACAGGGTAACAGGCTATTTGGAAGAAGAGGC
CATTCAATCATATACCGAGTACCTCAAAGACATAAGGAGTGGGAAAATCGAGAAATGTCAAGGCTCCCGCT
ATTGCGATCGACTACTGGAGGTTGCC TAAGGAAGCTACACTCGAGGATGTCATAACTGTGATCCGAGCCG
ATGAAGCACATCACCGTGATGTCAACCACTTTGCTTCTGTAAAGTATTAGCTCTCAATCTGAATTTATATT
TGGCATGCCAGTTACGAGATCGATAAGATTGATATACCAAAC TTTCTTCGTGAACCAGTATTTGTTGTAA
AAAGCGATTTCTTTGATTAACAGAATTCATTT CAGATGGAATTCGATTTACTCAATCTACTCGGTCTT
TGGGTGCAGGATGTTCACTATCAGGGCAAGAAACTGAAAGAGGCACCTGCTCCTATTGGCTATCATTTGA

Figure S8. Genomic sequence of HpAOX2 (acc. no. KU674356). Exons are in yellow, stop codon in pink, repetitive sequences in red, predicted LTR region homologous of a *Cassandra* MT-LTR retrotransposon sequence in green.

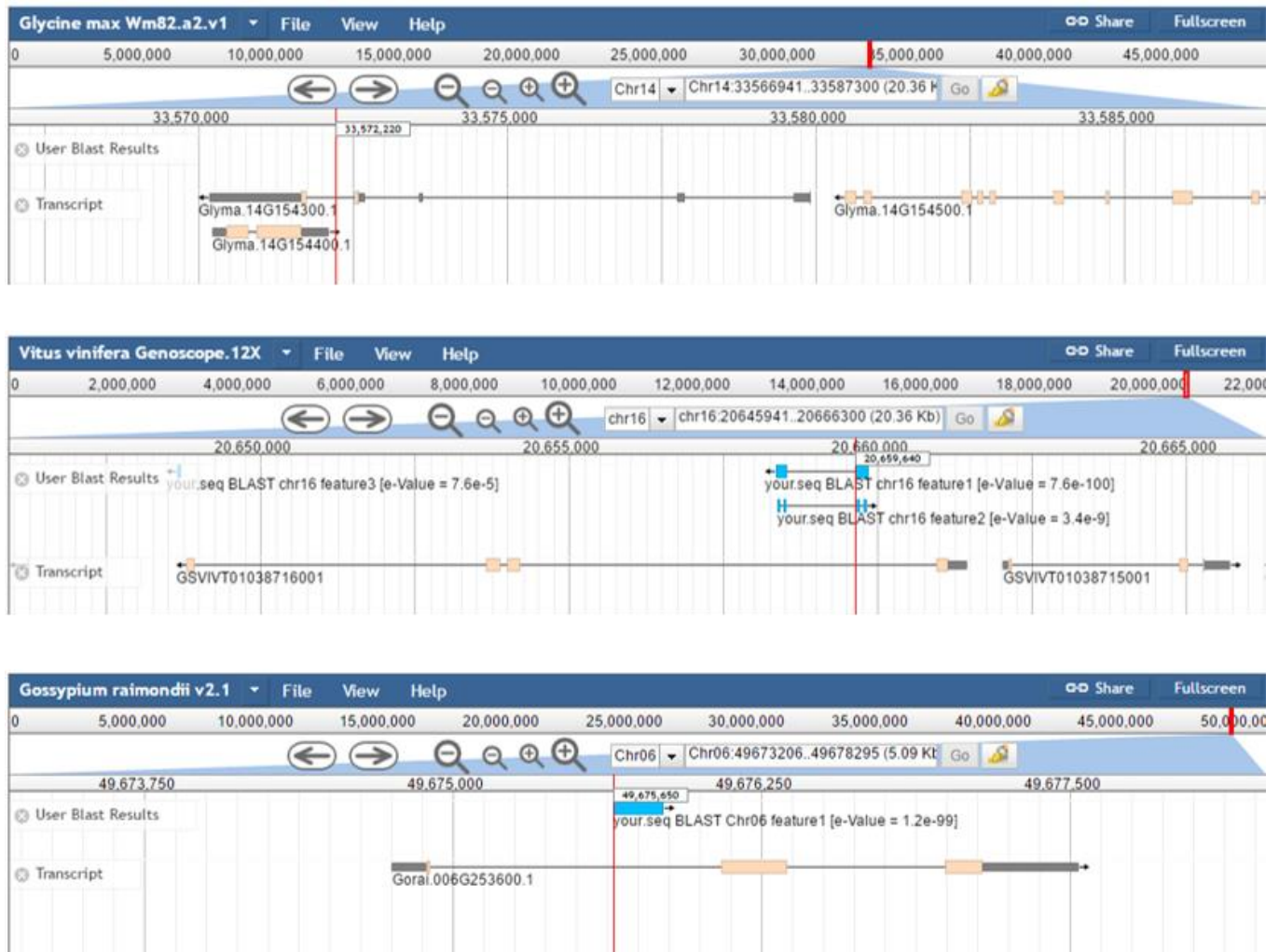


Figure S9. Genome mapping showing the location of the pre-miRNA sequences that encode the homologous miRNAs.

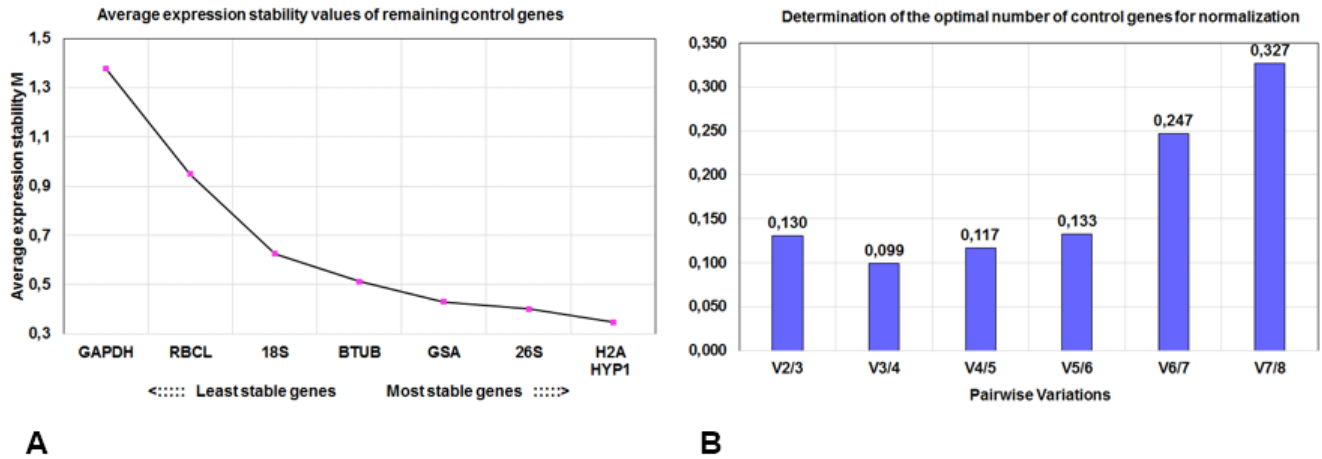


Figure S10. Expression stability analysis of candidate reference genes by GeNorm algorithm. **(A)** Average expression stability M. **(B)** Determination of the optimal number of reference genes for normalization by pairwise variation (V).

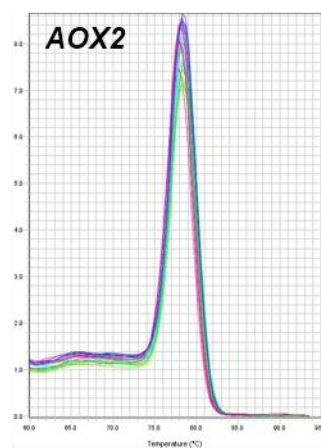
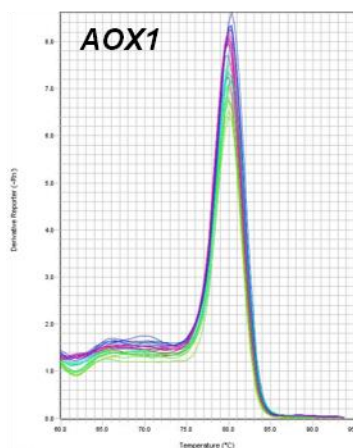
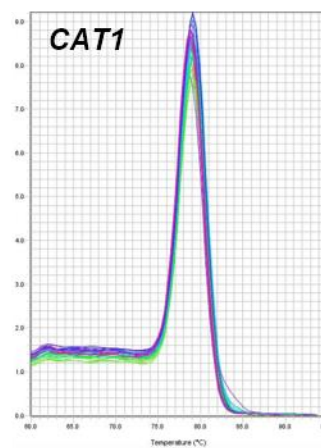
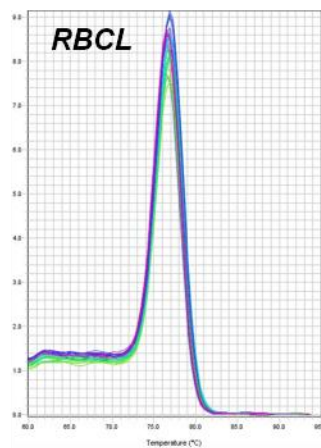
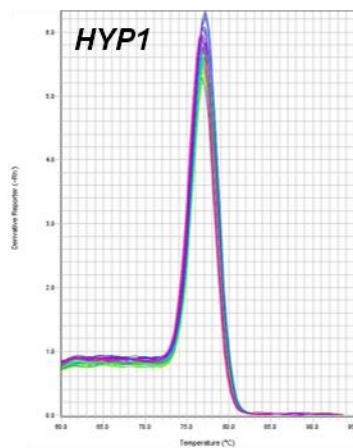
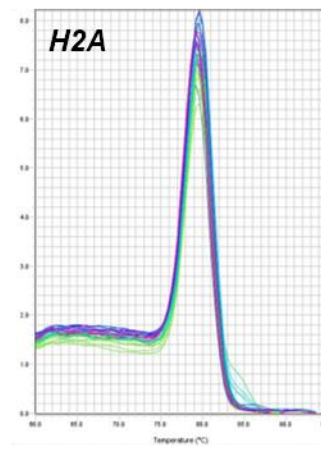
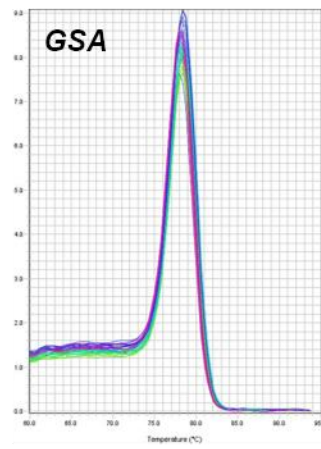
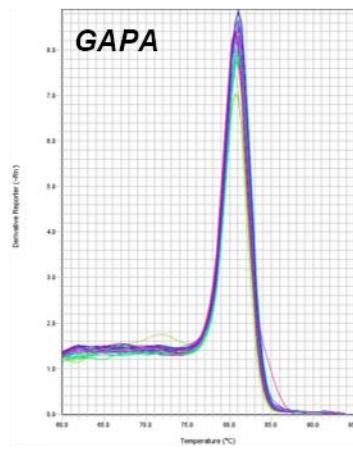
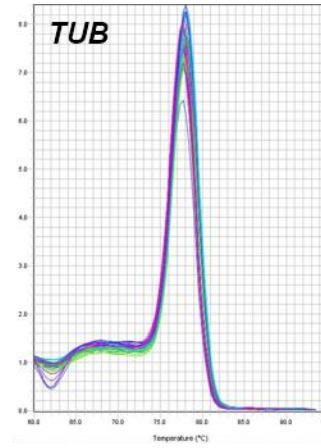
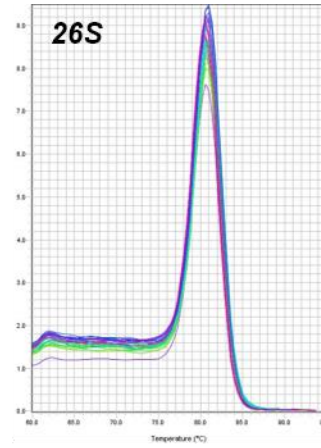
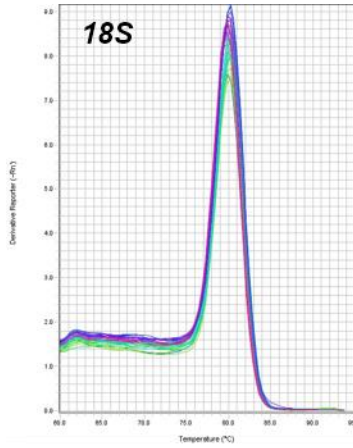


Figure S11. Melting curves of the genes analyzed showing specific peaks.

1.2 Supplementary Tables

Table S1. Primer sequences used in *HpAOX* genes isolation.

Gene	Aim	Primer sequence (5'-3')	Ta (°C)
<i>HpAOX1</i>	3'RACE-PCR	Fw: GCAGGGCGCCTTCTTCAACATCTAC	58
	5'RACE-PCR	Rev: CATCAAGTGCATCCGCTCGTTCTCT	68
	Whole gene	Fw: CCTAACTGCTCTCCTAAAAAGA Rev: CTCTCTCCAACGAAGCAATG	55
<i>HpAOX2</i>	3'RACE-PCR	Fw: AAGCTCGCTCACAGGGTAACAGG	55
	5'RACE-PCR	Rev: GAACAGCCCCTGCACGATAAGAACC	68
	Whole gene	Fw: AGAAGGGGAATTTGTGGAG Rev: CAACACAAGCTGAAGTAAGAAC	60

Ta: annealing temperature, Fw: forward, Rev: reverse.

Table S2. Eudicot plant species used in the NJ analysis.

O	F	Species	Accession number
Brassicales	Brassicaceae	<i>Arabidopsis lyrata</i>	AL1G33660 AL3G24680 AL3G24690 AL5G06730 AL0G08000 AL8G30860
		<i>Arabidopsis thaliana</i>	AT1G32350 AT3G22360 AT3G22370 AT3G27620 AT5G64210
		<i>Boechera stricta</i>	Bostr.19424s0421.1 Bostr.19424s0422.1 Bostr.3359s0124.1 Bostr.0568s0383.1
		<i>Brassica rapa</i>	Bra010153 Bra001865 Bra031351 Bra023835 Bra037768
		<i>Capsella grandiflora</i>	Cagra.1189s0011.1 Cagra.1189s0012.1 Cagra.5575s0007.1 Cagra.3957s0019.1 Cagra.0248s0102.1
		<i>Capsella rubella</i>	Carubv10014065m Carubv10019557m Carubv10009739m Carubv10026681m
		<i>Eutrema salsugineum</i>	Thhalv10021013m Thhalv10009342m Thhalv10005697m
		<i>Thellungiella parvula</i>	TP3G20130 TP3G20140 TP2G27270
		C	<i>Carica papaya</i>
Cucurbiti	Cucur	<i>Citrullus lanatus</i>	CL10G07420
		<i>Cucumis melo</i>	CM00136G00010
		<i>Cucumis sativus</i>	Cucsa.398150.1
Fabales	Fabaceae	<i>Glycine max</i>	GM04G14800 GM08G07690 GM08G07700 GM05G24455
		<i>Lotus japonicus</i>	LJ2G020780 LJ4G005280 LJ4G005290
		<i>Medicago truncatula</i>	MT5G026620 MT5G070680 MT5G070870 MT5G070880
		<i>Phaseolus vulgaris</i>	Phvul.002G127100.1 Phvul.002G209100.1 Phvul.002G209200.1

O	F	Species	Accession number
Malvales	Malvaceae	<i>Gossypium raimondii</i>	Gorai.008G296600.1 Gorai.012G142200.1 Gorai.005G220400.1 Gorai.005G220500.1
		<i>Theobroma cacao</i>	TC03G031300 TC02G011670
Malvales	Lina	<i>Linum usitatissimum</i>	Lus10035670 Lus10005372 Lus10020523
		<i>Populus trichocarpa</i>	PT03G09340 PT12G01430 PT12G01440 PT15G01960
	Salicaceae	<i>Salix purpurea</i>	SapurV1A.1470s0080.1 SapurV1A.0346s0170.1 SapurV1A.0377s0140.1 SapurV1A.0377s0150.1 SapurV1A.3352s0030.1 SapurV1A.0894s0160.1
		<i>Manihot esculenta</i>	ME10292G00060
	E	<i>Ricinus communis</i>	RC30063G00030
Ranunc	R	<i>Aquilegia coerulea</i> Goldsmith	Aquca_105_00003.1 Aquca_043_00024.1 Aquca_008_00134.1 Aquca_033_00110.1
Rosales	Rosaceae	<i>Fragaria vesca</i>	FV5G29310 FV5G21950
		<i>Malus domestica</i>	MD00G028680 MD00G081720 MD13G026910 MD16G016620
		<i>Prunus persica</i>	Prupe.5G018700.1 Prupe.1G061800.1 Prupe.1G061900.1 Prupe.1G061400.1
Sapindales	Rutaceae	<i>Citrus clementina</i>	Ciclev10001766m Ciclev10003687m Ciclev10028835m
		<i>Citrus sinensis</i>	orange1.1g037339m orange1.1g019765m orange1.1g020532m
Solanales	Solanaceae	<i>Solanum lycopersicum</i>	Solyc08g005550 Solyc08g075540 Solyc08g075550 Solyc01g105220
		<i>Solanum tuberosum</i>	PGSC0003DMT400019708 PGSC0003DMT400019707 PGSC0003DMT400047562 PGSC0003DMG400012558
A	a	<i>Amborella trichopoda</i>	ATR_00038G01180 ATR_00048G01570
B	b	<i>Beta vulgaris</i>	BV5G19180 BV9G03180
C	c	<i>Kalanchoe laxiflora</i>	Kalax.0453s0008.1 Kalax.1476s0004.1 Kalax.0496s0017.1 Kalax.0907s0011.1 Kalax.0414s0014.1

<i>D</i>	<i>d</i>	<i>Mimulus guttatus</i>	Migut.J01127.1 Migut.E01358.1 Migut.N01067.1
<i>E</i>	<i>e</i>	<i>Eucalyptus grandis</i>	Eucgr.E01214.1 Eucgr.E01213.1 Eucgr.I02663.1
<i>F</i>	<i>f</i>	<i>Vitis vinifera</i>	VV02G09030 VV02G09050 VV00G00110

O: order, F: family; C: *Caricaceae*, Curcubi: *Curcubiales*, Curc: *Curcubitaceae*, Lina: *Linaceae*, E: *Euphorbiaceae*, Ranun: *Ranunculales*, R: *Ranunculaceae*, A: *Amborellales*, B: *Caryophyllales*, C: *Saxifragales*, D: *Lamiales*, E: *Myrtales*, F: *Vitales*, a: *Amborellaceae*, b: *Amarantaceae*, c: *Crassulaceae*, d: *Phrymaceae*, e: *Myrtaceae*, f: *Vitaceae*.

In grey the AOX2 sequences.

Table S3. Monocot plant species used in the NJ analysis.

Species	Accession number
<i>Brachypodium distachyon</i>	BD3G52505 BD5G20540 BD5G20547 BD5G20557
<i>Hordeum vulgare</i>	CAJW010038523 CAJW011587016 CAJW010099492
<i>Musa acuminata</i>	GSMUA_Achr5G03810_001 GSMUA_Achr6G01170_001 GSMUA_Achr6G01300_001 GSMUA_Achr1G27800_001
<i>Oryza brachyantha</i>	OB02G22630 OB02G36280 OB04G30980 OB04G30990
<i>Oryza glaberrima</i>	ORGLA02G0249500 ORGLA04G0206000 ORGLA04G0206100
<i>Oryza sativa</i>	BGOSGA008063 BGOSGA005788 BGOSGA014421 BGOSGA014422
<i>Sorghum bicolor</i>	SB04G030820 SB06G027410 SB06G027420 SB06G027430
<i>Zea mays</i>	ZM02G05480 ZM02G05490 ZM02G05500 ZM05G37570