

Appendix A Protein sequences used in conserved motif analysis

>ApisOr23

MNLNDEQNYIVNLKLMKITGFYHLISPRAPKYFGFNVYKVTA AIEVMTGIFSIIMLFLSSYY
YLDNTNELMSHFMLVVAIFFSTLKIFWVSRNSETIWNNMDMTCINFLLYTGHKKEILKKA
RAKSISTTILFVILWSSVTVAWSISPFVVDVYLNKFKDETRRFYNSLNYVYPISEEFYNE
HFLYFYVVEMLSVVFWGHGTVA YDTFVISICITIAFQLKTI AVSYISLNDKKGDIKNLKDND
LEAMFNKLLIQDQQNMFKKIKEIYKIFEPVTFVQLAAQSMLIILQAYMIFINHYNGFSLLS
VPIIKLIVTVAPNIIHLFITCYLYTNINHQQDSMN FALYSSDWTAMSINYKKMLLFTMRMND
AEKLLKISLRKIVNLEMFASVMHLTYSIISVLAKSYGNTNTK

>AglyOr14

MNPNDENYIINLKLMKITGFYQLINPHTSKYLGFN VYKVGAAFEVMFGIISMLLLCLSSYY
YLDNTNELMSHFMLVVAIFFSIFKISWASKKSEMIWNNLDMTSINFLSYTGHKQEILQTARA
KSISTTIIFVILWSSVTVAWSISPFKIDVYLN VKFNDEIRRFYNSLNYVYPISEESYNENFL
YFYVVEMLQVIFWGHGTVA YDTFVISICISIAFQLKTI AVSYTSLNDIKGDIKNLKHNDLEAI
LNLKLV IQDQQKMFKKIKEIYKIFQPVT FVQLVAQSMLIILQAYMIFINYYNGFSLLSVPIIKL
IVTVAPNIIHLFITCYLYSNINDQKDSMN FALYSGDWTAMSIKYKNMLLFAMRMNDAEKL
KLKISLRKIVNLEMFANVMHLTYRIISVLSKSYGNTKSK

>AgosOr23

MNPNDENYIINLKLMKITGFYQLINPHTSKYLGFN VYKVGAGLEV MFGIISMLLLFLSSYY
YLDNTNELMSHFMLVVAIFFSTFKISWVSKKSEMIWNNLDMTSINFLSYTGHKQEILQTAR
AKSISTTIIFVILWSSVTVAWSISPFKIDVYLN VKFNDEIRRFYNSLNYVYPISEESYNENF
LYFYVVEMLQVIFWGHGTVA YDTFVISICISIAFQLKTI AVSYTSLNDIKGDIKNLKHNDLE
AILNLKLV IQDQQKMFKKIKEIYKIFQPVT FVQLAAQSMLIILQAYMIFINHYNGFSLLSVPII
KLIVTVAPNIIHLFITCYLYSNINDQKDSMN FALYSCDWTAMSIKYKNMLLFAMRMNDAE
KLKISLRKIVNLEMFASVMHLTYSIISVLSKSYGNTKSK

>DnoxOr23

MILNDEQNYLINLKLMKITGFYQLIHPR TTKYFGSNAYNAVA AIEVMAGVFSISLLFLSSYY
YLDNTNELMNHFMLVVAIFFSTFKIFWVSKNSKRIWNNMDITSINFLTYTGHKREILHN GR
AKSIFTTILFVILWSSVTVAWSISPFKIDVYLN VKFKDGIRRLRYNSLNYVYPISEEFYNGH
FLYFYVIELLQVILWGHGTVA YDTFVISICISIAFQLKTI AVSYISLNDRKCDEKNFKDDLE
AMFNKLLIQDQQNMFKKIKKIYQIFQPVTYVQLAAQSMLIIFQAYMIFINHYNGFSLISVP
ILKLVVTVAPNIMHLFTTCYLYSNINYQKDSMN FALYSSDWTAMSINYKKMLLFAMRMND
AEKLLKQKISLRKIVNLEMFASVMHLTYSIISV LAKSYGKTNTK

>McerOr23

MDLNDEQNYIINLKLMKITGFYQLINPCTPKHFGFN VFKVAAAIEVMTGVISVSLFLSSSY
YLNNTNELMSHFMLVVAIFFSTFKIFWVSRNSETIWNNMDMTCINFLSYTGHKKEILQNAR
AKSISTTILFIILWSSVTVAWCISPFVVDVYLN VKFKDDEIRQFRYNSLNYVYPISEGEFYNE
HFLYFYVVEMLQVFWGHGTVA YDTFVISICISIAFQLKTI AVSYISLNDRKC DIKH FEDND
LEAMFNKLLIQDQQNIKIKEIYKIFQPVTYVQLAAQSMLIILQAYMIFINHYNGFSLLSVPII
KLVVTVAPNIIHLFITCYLYSDINYQKDSMN FALYSSDWTAMSISYKKMLLFTMRMNDAE
KLKISLRKIVNLEMFASVMHLTYSIISV LAKSYGNTNTK

>MperOr23

MNFNDEKNYIFNLRLMKITGFYQLIYPSAPKCFGFNAYKVAAAIEVMTGVLSV SFLFSSSY
YYLDNTNELMSHFMLVVAIFFSTFKIFWVSRNSKTIWNNLDMTSINFLSYTGHKKEILQNA

RAKSISTTILFVILWSSVTVAWCISPFVVDVYLVNFKDKDDEIRRFYNSLNYVYPISGESYN
 EHFLYFYVVEMLQVVFVWGHGTVAYDTFVISICISIAFQLKTIASYSISLNDRKGDVKNLKD
 NDLEAIFNLKLLIQDQQNMFKKIKEIYKIFQPVTYVQLAAQSMLIILQAYMIFINHYNGFSL
 LSVPIIKLVVTVAPNIIHLFITCYLYSDINYQKDSMNFALYSSDWTAMSISYKMLLFTMRM
 NDAEKLKLRKIVNLEMFASVMHLTYSIISVLAKSYGNTNTK

>RmaiOr23

MTLDDEQNYIVNLKLMKITGFYQLINPGTPKYFGFNVYKIGAAIEVMMSGIISILLCSSSY
 YLDNTNELMSHFMLVVAIFFSTFKISWVSRNSETIWNNDMTSVDFLSYTGHKKETLQIAR
 AKSISTTILFVILWSSVTVAWSISPFVVDVYLVNFKFNDEIRRFYNSLNYVYPINEESYNKH
 FLYFYVVEMLQVVFVWGHGTVAYDTFVISICISIAFQLKTIASYSISLNDIKRDVKNLTHNDL
 EAIFNLKLLIQDQQNMFKKIKETYKIFQPVTYVQLAAQSMLIILQAYMIFINHFNGFSLLSVP
 IIKLVVTVAPNIIHLFITCYLYHNINDQKESMNFALYSSDWTAMSIKYKNMLLFTMRMNDAE
 KLKLRKIVNLEMFASVMHLTYSIISVLAKSYGNIKSK

>SmisOr23

MNLNDEQNYIVNLKLMKITGFYHLINPRAPKYFGFNVYKVTAALEVMTGLFSITMLFLSS
 YYYLDNTNELTSHFMLVVAIFFSTLKFVWVSRNSETIWNNDMTMTCINFLSYTGHKKEILKK
 ARAKSMSTTILFVILWSSVTVAWCISPFVVDVYLDVFKDKDETRRFYNSLNYVYPISEEFY
 NQHFLYFYVIEMLSVVFVWGHGTVAYDTFVISICITIAFQLKTIASYSISLNGRKGDIKNFKDN
 DLEAMFNLKLLIQDQQNMFKKIKEIYKIFEPVTYVQLAAQSMLIILQAYMIFINHYNGYSLL
 SVPIIKLVVTVAPNIIHLFITCYLYTNINHQQDSMNFALYSGDWTAMNINYYKMLLFTMRM
 NDAEKLKLRKIVNLEMFASVMHLTYSIISVLAKSYGNTNTK

Appendix B Primers used in this study

Primer name	Primer sequence (5'-3')
Gene cloning	
ApisOr23-F	ATGAATCTCAATGATGAGCAAAAC
ApisOr23-R	TTATTTAGTATTAGTATTTCCATAACTTTTTG
Expression vector construction	
ApisOr23-F	TCAACTAGT <u>GCCACCAT</u> GAAATCTCAATGATGAGCAA (Spe I)
ApisOr23-R	TCACTCGAGTGAACCTATCAGGTGATAGAAACC (Xho I)
RT-PCR	
ApisOr23-F	AAATTACTGGTTTCTATCACCTGATAAGTCC
ApisOr23-R	GGCGTGTTTCATCTTTGAACTTAAT
ApisSDHB-F	CAGAAACTCCCGAAGTGAAGC
ApisSDHB-R	TAATCCAACGATACGCCTGC

Note: The restriction enzyme sites are marked with bold fronts; Kozak sequences are marked with underline.

Appendix C All 57 chemicals used for functional characterization of ApisOr23

Number	chemical name	CAS
1	2-Phenylethanol	60-12-8
2	cis-3-Hexen-1-ol	928-96-1
3	β -Citronellol	106-22-9
4	Geraniol	106-24-1
5	(1S)-(-)-Verbenone	1196-01-6
6	1-Hexanol	111-27-3
7	(S)-cis-Verbenol	18881-04-4
8	trans-3-Hexen-1-ol	928-97-2
9	3,7-Dimethyl-3-octanol	78-69-3
10	(-)-Borneol	464-45-9
11	(+)-Borneol	464-43-7
12	(1R)-(-)-Myrtenol	19894-97-4
13	(-)-trans-Pinocarveol	547-61-5
14	(-)-Linalool	126-91-0
15	Linalool	78-70-6
16	Methyl benzoate	93-58-3
17	Myrcene	123-35-3
18	(R)-(+)-Limonene	5989-27-5
19	α -Pinene	80-56-8
20	(-)- β -Pinene	18172-67-3
21	Camphene	79-92-5
22	α -Humulene	6753-98-6
23	(S)-(-)-Limonene	5989-54-8
24	α -Terpinene	99-86-5
25	(-)-trans-Caryophyllene	87-44-5
26	Farnesene, mixture of isomers	
27	cis-2-Hexen-1-ol	928-94-9
28	1-Heptanol	111-70-6
29	4'-Ethylacetophenone	937-30-4
30	Hexyl acetate	142-92-7
31	trans-2-Hexen-1-al	6728-26-3
32	4-Ethylbenzaldehyde	4748-78-1
33	3-Vinylbenzaldehyde	19955-99-8
34	(1R)-(-)-Myrtenal	18486-69-6
35	Benzaldehyde	100-52-7
36	Heptanal	111-71-7

37	cis-3-Hexenyl acetate	3681-71-8
38	trans-2-Hexenyl acetate	2497-18-9
39	cis-3-Hexenyl acetate	3681-71-8
40	1,4-Diethylbenzene	105-05-5
41	Tetradecane	629-59-4
42	Nonyl acetate	143-13-5
43	Ocimene	13877-91-3
44	Tridecane	629-50-5
45	Methyl salicylate	119-36-8
46	(±)-Camphor	76-22-2
47	Nerolidol	7212-44-4
48	2,6-Di-tert-butylphenol	128-39-2
49	2-Pentadecanone	2345-28-0
50	Acetophenone	98-86-2
51	(+)-Cedrol	77-53-2
52	Ethyl butyrate	105-54-4
53	Ethyl hexanoate	123-66-0
54	Salicylaldehyde	90-02-8
55	Methyl 2-methoxybenzoate	606-45-1
56	Eugenol	97-53-0
57	Methyl phenylacetate	101-41-7



Appendix D Ten motifs predicted from Or23 clade in aphids