

Supplementary Material

Promoter deletion analysis reveals root-specific expression of the alkenal reductase gene (*OS-AER1*) in *Oryza sativa*

Aniversari Apriana^{A,B}, *Atmitri Sisharmini*^B, *Hajrial Aswidinnoor*^A, *Kurniawan R. Trijatmiko*^{B,C} and *Sudarsono Sudarsono*^{A,C}

^APMB Lab, Department of Agronomy and Horticulture, Faculty of Agriculture, Bogor Agricultural University, Jalan Raya Ciampea, Bogor, Indonesia.

^BIndonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development, Jalan Tentara Pelajar 3A, Bogor, Indonesia.

^CCorresponding authors. Emails: s_sudarsono@ymail.com; krtrijatmiko@gmail.com

attggtt	aat	at	ggattaat	tactacaact	gttttactgc	tttcaactac	tgcttttaaa	-3008
tgctagcttt	atgcaaatga	acctctagcc	tcctttgggt	ataccctgca	tcataccttc			-2948
ttttccggta	tgacttgctt	agtacagtgg	gtagtactca	gtcttgctct	cttttcccc			-2888
acaccagagt	tgaagtcctt	ctcagttgaa	gatgtctcga	agaggttggt	ttcgtcgcgg			-2828
ccgtcaagga	tgccgtgga	atagagtcgc	ccgctgcagg	agtcaagtct	tcaggcttag			-2768
ctcgttttta	tcttttccgc	tgatattgta	atctttt	ata	tt	tttttgtaa	gacgtggata	-2708
tgtatgtcaa	caattatcgt	ttgtgtatcc	tggtagttc	tgacagagg	tttaatac	at		-2648
att	cagctta	gaaattctag	ttcgtgaatt	tttggtcgtg	acaaccacct	agctcaccac		-2588
ctctctcctg	gtgccctga	ctcctccgcc	ccaacctcca	tactaccgg	cctcccgtg			-2528
ttgccgatgt	gctacgcaa	gccgtgtgcc	ctactcatcg	gccgggttac	cggtccctt			-2468
gcttccgcca	gcctcaaac	tcccccccc	ccctcctct	ctctgtttag	tgttcagaca			-2408
tggagttagt	agcgggagga	taaggctgat	gtgggctcg	actgtgatt	ctgacattgg	gttttgacag		-2348
gtcgtggact	tggtctatgg	accattgagc	actgtggatt	actgggctgg	ttactgtgct			-2288
gctgctagct	tgctgtacct	tttgtttttt	ctctcgttca	tagagtgggt	ttgcctctct			-2228
cagcgcctca	tcgatatctg	cagctctccg	acatctacat	agactttgct	tcgacttcat			-2168
ccttctgact	cgcatatctt	cgacttgatc	acccattcaa	gtgcttggct	acaatgactt			-2108
ccgcaaatga	tcatcatcaa	gtgtctacgc	catgaggctc	cgggttgccg	ctgctcacc			-2048
ctctcgggct	tcagt	gccac	ccacctctat	tagttgttgc	agtgtcgaga			-1988
ggtgccctcc	tcagcatgac	ttcatcaccg	atggacttct	cgccgctgca	ttgacatgga			-1928
gctacaactg	tgatcatcct	ttggaccaca	gcgtcaccgc	tctcgtcgtc	tccatcgccg			-1868
ccgcacatc	gacatcctcg	tcgtctgcac	tggtgatctt	cagctgctag	gttcttcttt			-1808
gtctacttctg	agcatcatcg	ccgcgtctcc	acgctgccat	catcactgct	ctagggcggt			-1748
ggtccagcta	cctctacatg	gctaccgaca	tcgccgtcca	agccgttgggt	cccactactt			-1688
catcttcatc	tcggtccaga	atgcctcatt	ggcagcgtcg	tcgtatcttc	ctcgactaca			-1628
atltgctctt	ctccgacaag	tgcggtgttc	tcaatcaatt	ctctctcttc	actgctctcg			-1568
gccac	gatc	atctcggagg	ccttctctgc	tagtctctc	tg	atatt	gcc	-1508
ggtggtctcc	tacctcgccc	gtgcgggtatt	gacaacatca	ttgcgcgtgt	ttgtcctccc			-1448
gggtctggca	aaccggcgg	aacgtctcgt	cctccaccgt	tcggcaacat	cgactctccg			-1388
gtgtcatctt	ccttaacgac	caccgcattc	tcgtcaccgt	catcatcttt	agttcgtctt			-1328
cgctcacgac	tactctatgg	tgccccatgc	acgcccgtcg	gctcggctac	ctcgacatca			-1268
gctttccgat	ttcggctaca	tcgatcatgg	ctactctacg	cacggcttca	tcgaccacgg			-1208
ctccctcaca	accttcgttt	tggtacgctc	gactatggca	caaagggtca	tgatatgtat			-1148
tgagtactct	cgtcggtttt	ttctctagtc	taagcgttca	cactgctcac	gctcggaccg			-1088
tggggag	ata	tt	agagtata	tatggtagtt	ctagaataa	taggatataa	ttgatttgat	-1028
tagagtcttt	ccttatctct	agtacctttc	taatctttac	ccgtgtaatc	tctatatacc			-968
agcctacgag	gctcaataca	aatat	tcgat	ttacgcagta	ttcctctctt	tctatactat		-908
tcaacagcat	gatccatctc	tttcatcctg	tacatcagct	tcactagatt	aattggcaca			-848
tcactctgact	tcttgaattg	attgcgcaat	tgagttatgt	aaccaaagtg	tatactttga			-788
aaatlttgac	acggttaagg	tacgtagccg	gtgacgttag	glttgataga	tttttctaga			-728
gctacctctg	accggtctct	gcagatgtag	ggcccacaca	caaataaata	aactttgaat			-668
tgaagaaaat	gtgtcaataa	cggacaaaac	ttctacgaga	taaataggta	aaggacgatg			-608
aatat	atat	gtaaaaacat	atcgtaagta	tgaggctttt	tctactatgg	cccttttttt		-548
ttgtgtgtgc	ggtgctttat	atacagttca	aattgttgaa	gcaacattta	ctgtcatgga			-488
ttatgtcaac	agagatgtca	gtacat	aata	taagctttta	tgaaatactc	tctccgggtca		-428
caaataataa	gtgtcacttt	agaatcagtg	tagtcaacaa	gagtttaact	tgactatctt			-368
tgltttgaa	acatcacttg	tgatttgga	tagccaaact	tttaaaaaca	ataactttca			-308
aatataata	ttagtttaca	cacaattcgt	tttgatttta	tacaaatgaa	ttaagtgatc			-248
aaattaagtt	atctcctgat	gactttgcaa	agtttaaaac	aacacttatt	tatgattgat			-188
tggagggatc	aattaattac	tcattctcca	ttcttaccga	tcgagcgtca	tgagtgcag			-128
agglttggtc	attcaaactg	taaacctcat	ctatlttaaac	ccaatcgcga	gtgctaagct			-68
tgcaactccc	cagcttgaac	ttagaggtgg	glttgcttagc	ctggttttac	aagaatctaa			-8
accagca	*							-1

Fig. S1. Sequence of the 3082 bp putative promoter isolated from *Oryza sativa AERI* (*OsAERI*) gene. The positions of selected cis-acting elements within the putative promoter were highlighted in different color. The AATAT or ATATT was consensus sequences of the ROOTMOTIFTAPOX-1 cis-acting element, whereas the GCCAC was the SORLI1AT cis-acting element. * start codon position.

Table S1. List of motif name, consensus motif, and nucleotide positions of cis-acting elements existed in the 3082 bp *Oryza sativa alkenal reductase 1 (OsAER1)* promoter region and their potential functions (NEW PLACE: <https://sogo.dna.affrc.go.jp/>)

No.	Motif name	Consensus motif	Nucleotide position	Potential function
1.	TATA-BOX	TATTTAAA	-89	Rice TATA binding protein interacts functionally with transcription factor IIB and the RF2a bZIP transcriptional activator in an enhanced transcription system
2.	ROOTMOTIFTAPOX1	ATATT AATAT	-296, -1076, -1521, -2645,-2726 -302, -457, -602, -943, -3056	Motif found in promoters of <i>rolD</i> , involved in root specific expression
3.	SORLIP1AT	GCCAC	-1562,-2028	Sequences over-represented in light-induced promoters in <i>Arabidopsis</i> , root specific gene
4.	ASF1MOTIFCAMV	TGACG CGTCA	-752 -138,-1346, -1352, -1892, -2050, -2822	TGACG motifs are found in many promoters and are involved in transcriptional activation of several genes by auxin and/or salicylic acid. Abiotic and biotic stress differentially stimulate AS-1 element activity. It is also involved in root specific expression
5.	RAV1	CAACA	-204, -389, -477, -502,-902, -1399, -1471, -2696	Binding consensus sequence of <i>Arabidopsis (At)</i> transcription factor, RAV1. RAV protein involved in some biological processes
6.	SURECOREATSULTR11	TGTTG GTCTC	-510, -1599, -1999, -2525 -709,-1420, -1499,-1531,-1779,-1876, -2200,-2850	Core of sulfur-responsive element (SURE) found in the promoter of SULTR1;1 high-affinity sulfate transporter gene in <i>Arabidopsis</i> ; SURE contains auxin response factor (ARF) binding sequence (GAGACA) and involved in abiotic stress responses
7.	RHERPATEXPA7	TCACGA GCACGT TCGTGA	-1319 -2203 -2607,-2621	Right part of RHEs (Root Hair-specific cis-Elements) conserved among the <i>Arabidopsis thaliana</i> A7 (AtEXPA7) orthologous (and paralogous) genes from diverse angiosperm species with different hair distribution patterns
8.	OSE2ROOTNODULE	CTCTT	-886,-918,-1390,-1579,-1618,-2895	One of the consensus sequence motifs of organ-specific elements (OSE) characteristic of the promoters activated in infected cells of root nodules. It's involved in controlling of root specific expression
9.	OSE1ROOTNODULE	AAGAG AAAGAT ATCTTT	-385,-2844 -1512 -367,-990,-1338,-2732,-2753	One of the consensus sequence motifs of organ-specific elements (OSE) characteristic of the promoters activated in infected cells of root nodules. It's involved in controlling of root specific expression
10.	DOFCOREZM	AAAG	-216,-615,-662,-800,-1163,-1514	Core site required for binding of Dof proteins in maize. It's involved in some biological processes and responds to biotic and abiotic stresses

		CTTT	-222,-310,-326,-367,- 377,-408,-450,-530, -552,-569,-672,-790,- 885,-990,-917,-999, -1018,-1263,-1338,- 1808,-1906,-2181,-2266, -2732,-2753,-2760,- 2894,-2945,-2972, -2998,-3012,-3025	
11.	WRKY71OS	TGAC	-130,-225,-374,-716,- 753,-778,-1475,-1933, -1968,-2110,-2159,- 2606,-2567,-2934	A core of TGAC-containing W-box of the gibberellin signaling pathway transcriptional repressor
		GTCA	-117,-138,-392,-412,- 428,-468,-492,-652, -1346,-1352,-1382	
12.	GTGANTG10	GTGA	-131,-250,-754,-1832,- 2607,-2621	GTGA motif found in the promoter of the tobacco (<i>Nt</i>) late pollen gene g10 which shows homology to pectate lyase and is the putative homologue of the tomato gene lat56
		TCAC	-351,-411,-427,-864,- 1098,-1107,-1199, -1321,-1351',-1576,- 1762,-1891,-1960, -2049,-2136,-2544,-2591	
13.	ABRELATERD1	ACGTG	-2712	It's involved in response to ABA signaling
		CACGT	-2203	
14.	ARFAT	TGTCTC	-2850	ARF (auxin response factor) binding site found in the promoters of primary/early auxin response genes of <i>Arabidopsis thaliana</i> (A.t.); AuxRE. It's involved in response to auxin.
15.	DREDR1ATRD29AB	TACCGACAT	-1717	Binding site of <i>DREB1</i> and <i>DREB2</i> - binding site of <i>Arabidopsis CBF1</i> . It is related to responsiveness to drought, low-temperature or high-salt stress
	2SSEEDPROTBANAPA	GTGTTTG	-1455	Conserved in many storage-protein gene promoters; May be important for high activity of the <i>napA</i> promoter in seed.
16.	GC-BOX	GCCGCC	-1866	Core of GCC-box found in many pathogen-responsive genes such as <i>PDF1.2</i> , <i>Thi2.1</i> , and <i>PR4</i> ; Has been shown to function as ethylene-responsive element.
17.	W-BOX	TGAC	-130,-225,-374,-716,- 753,-778,-839,-1475,- 1933,-1968,-2110,-2159,- 2350,-2363,-2567,-2606,- 2934,-2934	W box found in the promoter region of a transcriptional repressor ERF3 gene in tobacco; It may be involved in activation of <i>ERF3</i> gene by wounding

		GTCA	-117,-138,-392,-412,- 428,-468,-480,-492, -652,-1346,-1352,-1382,- 1892,-1913,-2050, -2699,-2783,-2822	
18.	AREs (ANAERO3CONSENSUS)	GCTGCT TCATCAC	-1820,-2285 -1762,-1960	Motifs found in silico in promoters of 13 anaerobic genes involved in the fermentative pathway (anaerobic set 1) (Mohanty et al., 2005)
19.	TATCCAMAYMOTIF OSRAMY3D	GTGGATA	-2708	
20.	SEF4MOTIFGM7S	ATTTTTG TAAAAAC	-2613 -319,-591	TATCCAY motif found in rice (<i>Os</i>) <i>RAmy3D</i> alpha-amylase gene promoter; Y=T/C; a GATA motif as its antisense sequence; TATCCAY motif and G motif are responsible for sugar repression (Toyofuku et al. 1998) SEF4 binding site; "Binding with SEF4 (soybean embryo factor 4)"; R=A/G;
21.	MYCCONSENSUSAT	CAAATG CAGGTC CACCTC	-260,-2099,-2989 -2345 -2011,-2282,-2285	
22.	MYB	CAACTG TAACGG CCGTTG CAGTTG	-1918,-3037 -645 -1700 -2860	MYC recognition site found in the promoters of the dehydration-responsive gene rd22 and many other genes in Arabidopsis. This sequence is also known as RRE (R response element) (Hartmann et al., 2005) MYB recognition site found in the promoters of the dehydration-responsive gene rd22 and many other genes in Arabidopsis.
23.	ANAERO1CONSENSUS	TTTGTTT	-365,-2261	Element involved in the fermentative pathway (anaerobic set 1) (Mohanty et al., 2005)
24.	ABRERATCAL	GCGCGTG	-1459	"ABRE-related sequence" or "Repeated sequence motifs" identified in the upstream regions of 162 Ca(2+)-responsive upregulated genes
25.	TGACGTVMAMY	TGACGT	-751	TGACGT motif" found in the <i>Vigna mungo</i> (V.m.) alpha-Amylase (Amy) gene promoter; Required for high level expression of alpha-Amylase in the cotyledons of the germinated seeds
26.	LTR	CCGAC GTCGG	-1611,-1719,-2196 -1132,-2368	LTRE-1" (low-temperature-responsive element) in barley (H.v.) blt4.9 gene promoter. Promoter elements in a low-temperature-responsive gene
27.	GT1CONSENSUS	TTTTTC	-566,-732,-1125,-2257	Consensus GT-1 binding site in many light-regulated genes
28.	EBOXBNNAPA	CAAATG CAATTG CATCTG CAGCTG	-260,-2099,-2989 -816 -841 -1822	E-box of napA storage-protein gene of <i>Brassica napus</i> (B.n.)
29.	CACTFTPPCA1	CCAT ATGG	-157,-891,-1300,-1768,- 1873,-2075,-2132, -2323,-2547 -489,-558,-1063,-1169,- 1170,-1238,	Tetranucleotide (CACT) is a key component of Mem1 (mesophyll expression module 1)

TCAT	-1308,-1506,-1727,- 1929,-1954,-2328, -2405,-3004 -98,-116,-137,-164,-307,- 491,-844,-882, -939,-1217,-1240,-1345,- 1381,-1467,-1508, -1556,-1659,-1679,- 1685,-1765,-1790, -1912,-1963,-2094,- 2168,-2217,-2247,-2490, -2954
ATGA	-135,-193,-226,-259,- 445,-575,-896,-1155, -1969,-2073,-2098,- 2111,-2935,-2988
