

Supplemental Tables and Figures to

Genome Structures and Halophyte-Specific Gene Expression of the Extremophile *Thellungiella parvula* in Comparison with *Thellungiella salsuginea* (*Thellungiella halophila*) and *Arabidopsis*

Dong-Ha Oh^{1*}, Maheshi Dassanayake¹, Jeffrey S. Haas, Anna Kropornika, Chris Wright, Matilde Paino d'Urzo, Hyewon Hong, Shahjahan Ali, Alvaro Hernandez, Georgina M. Lambert, Gunsu Inan, David W. Galbraith, Ray A. Bressan, Dae-Jin Yun, Jian-Kang Zhu, John M. Cheeseman, and Hans J. Bohnert

¹ These authors contributed equally to the article.

* Corresponding author; e-mail ohdongha@life.illinois.edu.

Supplemental Figure S1. Comparison of relative DNA contents by flow cytometry.

Supplemental Figure S2. Analysis of scaffold 00254.

Supplemental Figure S3. Comparison between *T. parvula* genomic sequences TP-1 and TP-2 with three orthologous genomic regions in *A. thaliana* and *T. salsuginea*.

Supplemental Figure S4. Syntenic relationships and homologies of sequences represented by percent identity plots (pip).

Supplemental Figure S5. Hairpin structures in the 5'UTR of the halophytic *Thellungiella* species.

Supplemental Figure S6. Comparison of CT tracts in the 5'UTRs of various genes between the three species.

Supplemental Table S1. ORFs identified in TP-1, AT-1 and TS-1.

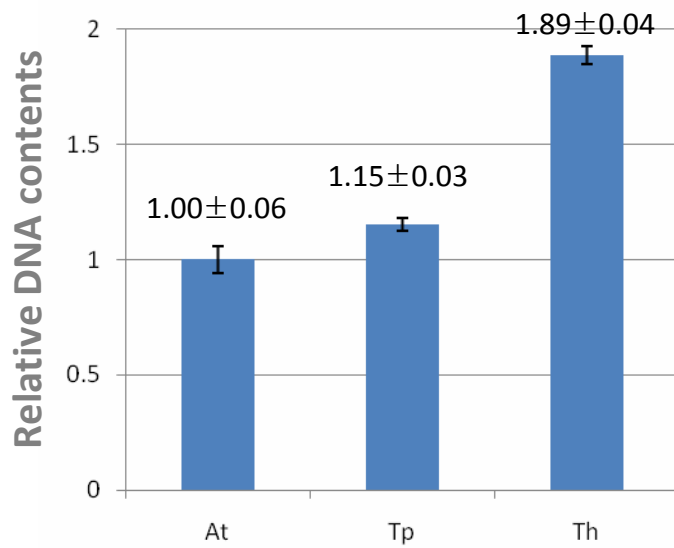
Supplemental Table S2. ORFs identified in TP-2, AT-2 and TS-2.

Supplemental Table S3. Comparison of ratio of non-synonymous (**Ka**) vs. synonymous (**Ks**) substitution rates in orthologous genes and the estimated divergence time of three species.

Supplemental Table S4. Comparison of relative gene expression levels by qPCR.

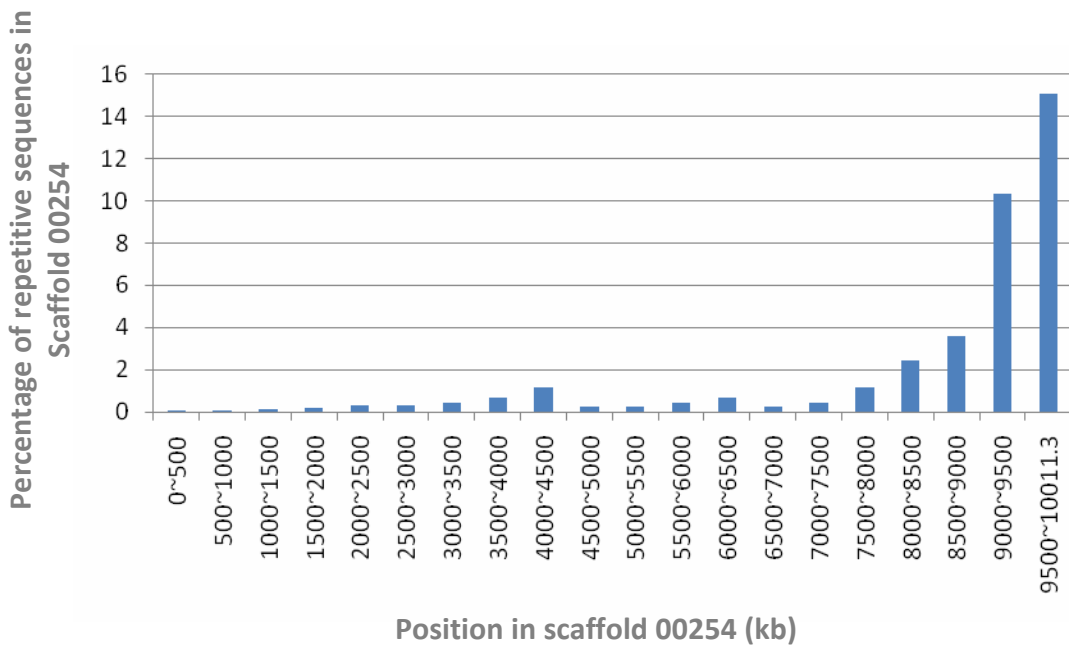
Supplemental Table S5. Known *cis*-elements in SOS1 promoter regions up to 2000bp upstream of the translation start site.

Supplemental Table S6. List of PCR primers.

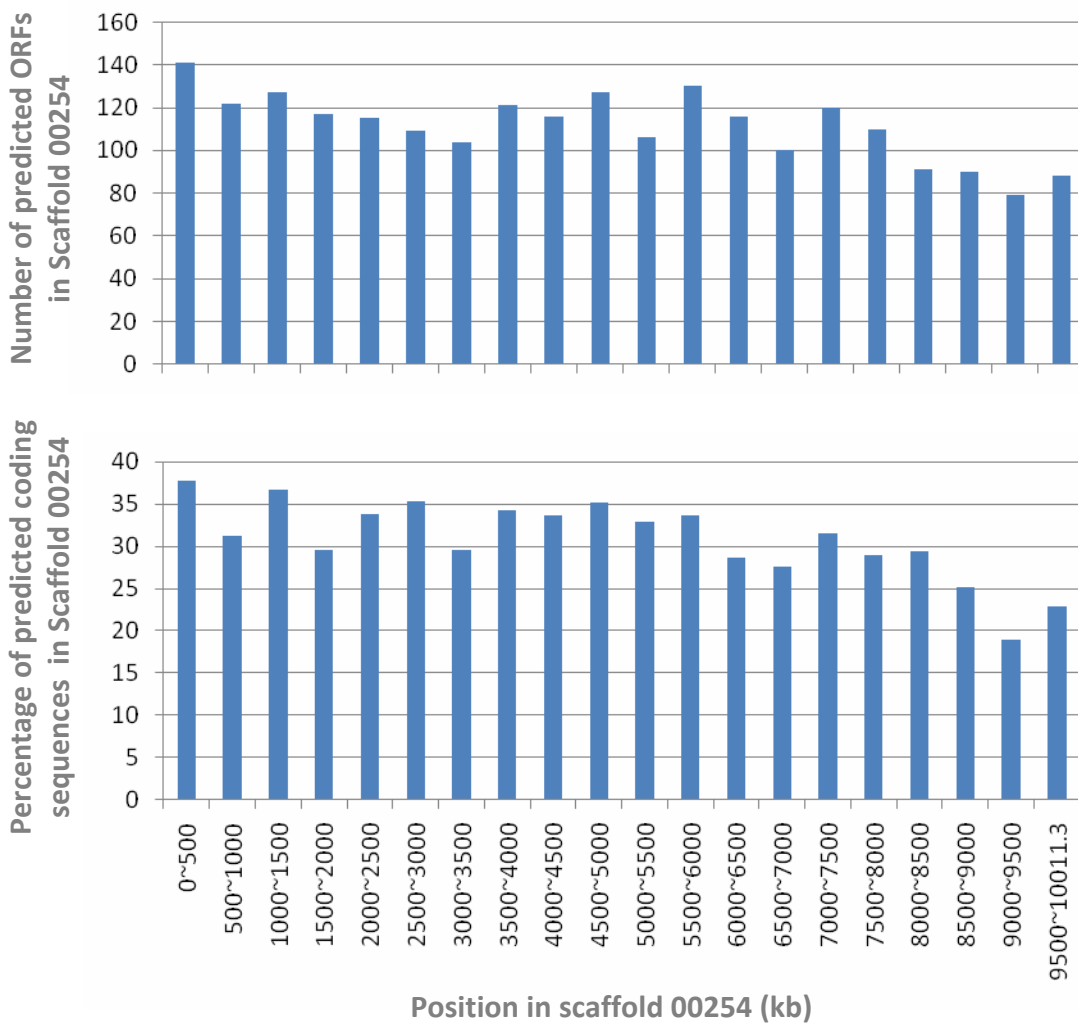


Supplemental Figure S1. Comparison of relative DNA contents by flow cytometry. Flow cytometry was performed as described in the method. Relative values to Arabidopsis are represented. Error bars indicate standard deviation from three repeats.

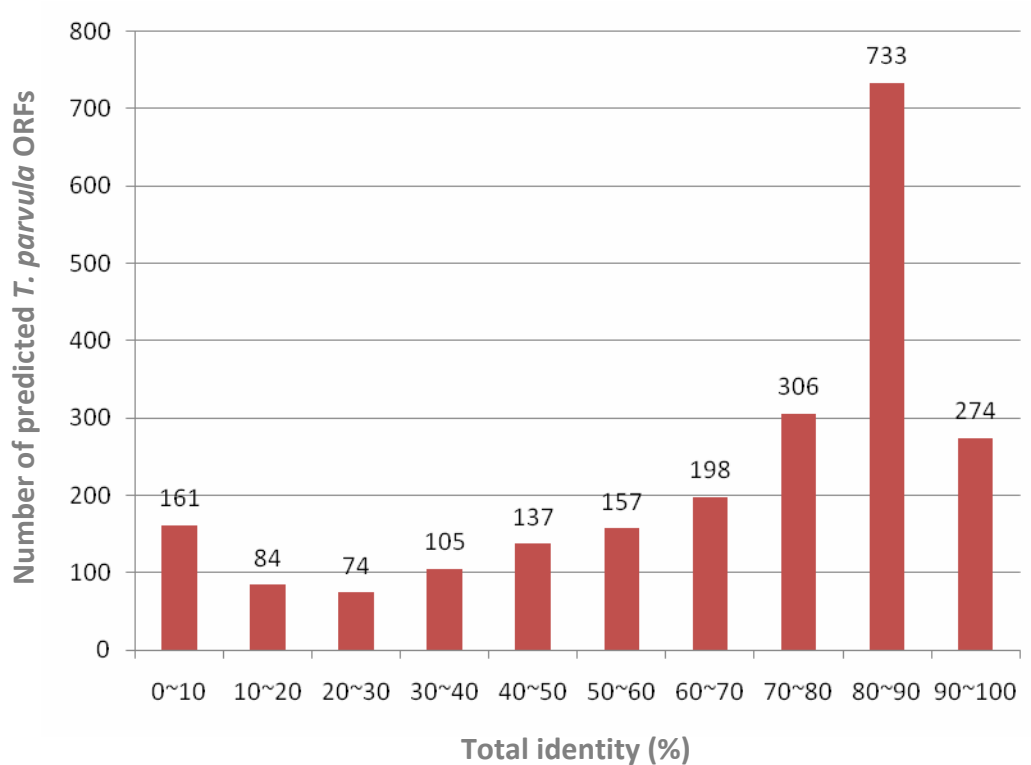
(A)



(B)



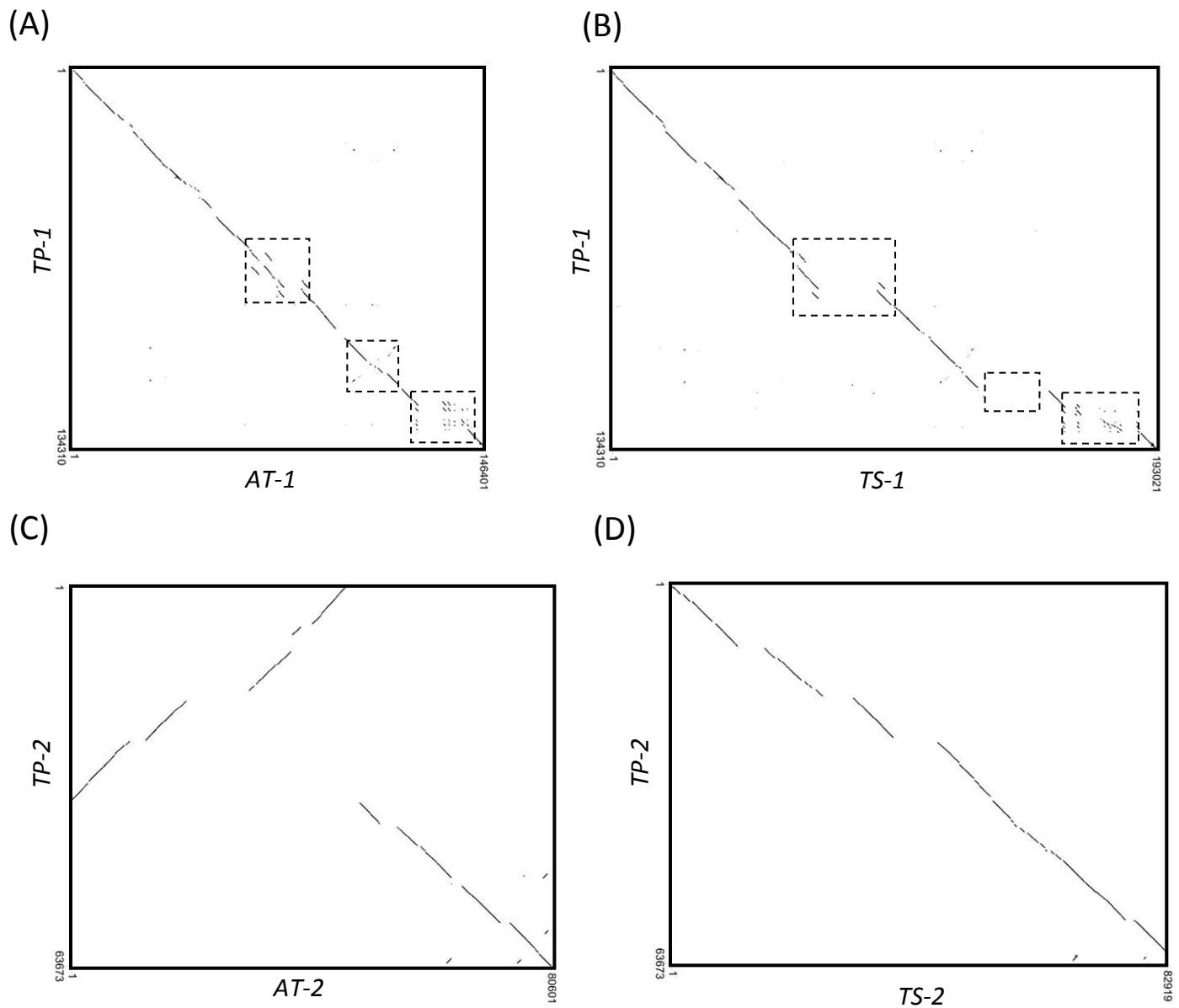
(C)



Supplemental Figure S2. Analysis of scaffold 00254.

Distribution of repetitive sequences (A) and ORFs/coding sequences (B) predicted by Repeatmasker and FGENESH, respectively, in scaffold 00254.

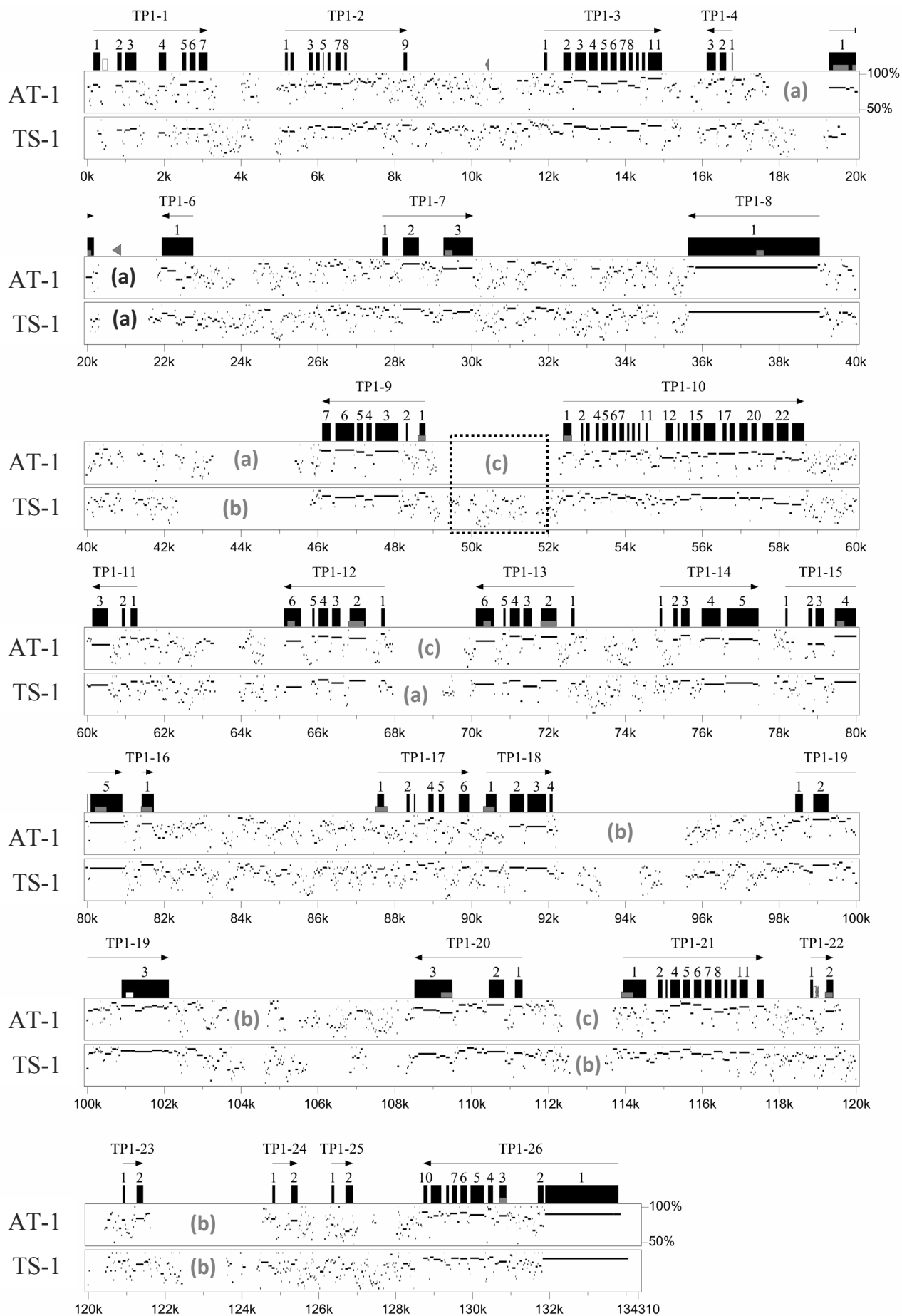
(C) Homology of predicted ORFs in scaffold 00254 compared to *A. thaliana* reference mRNAs. Total identity was calculated by multiplying percentage of the *T. parvula* predicted ORF sequence covered by the blastn alignment with percent nucleotide identity.



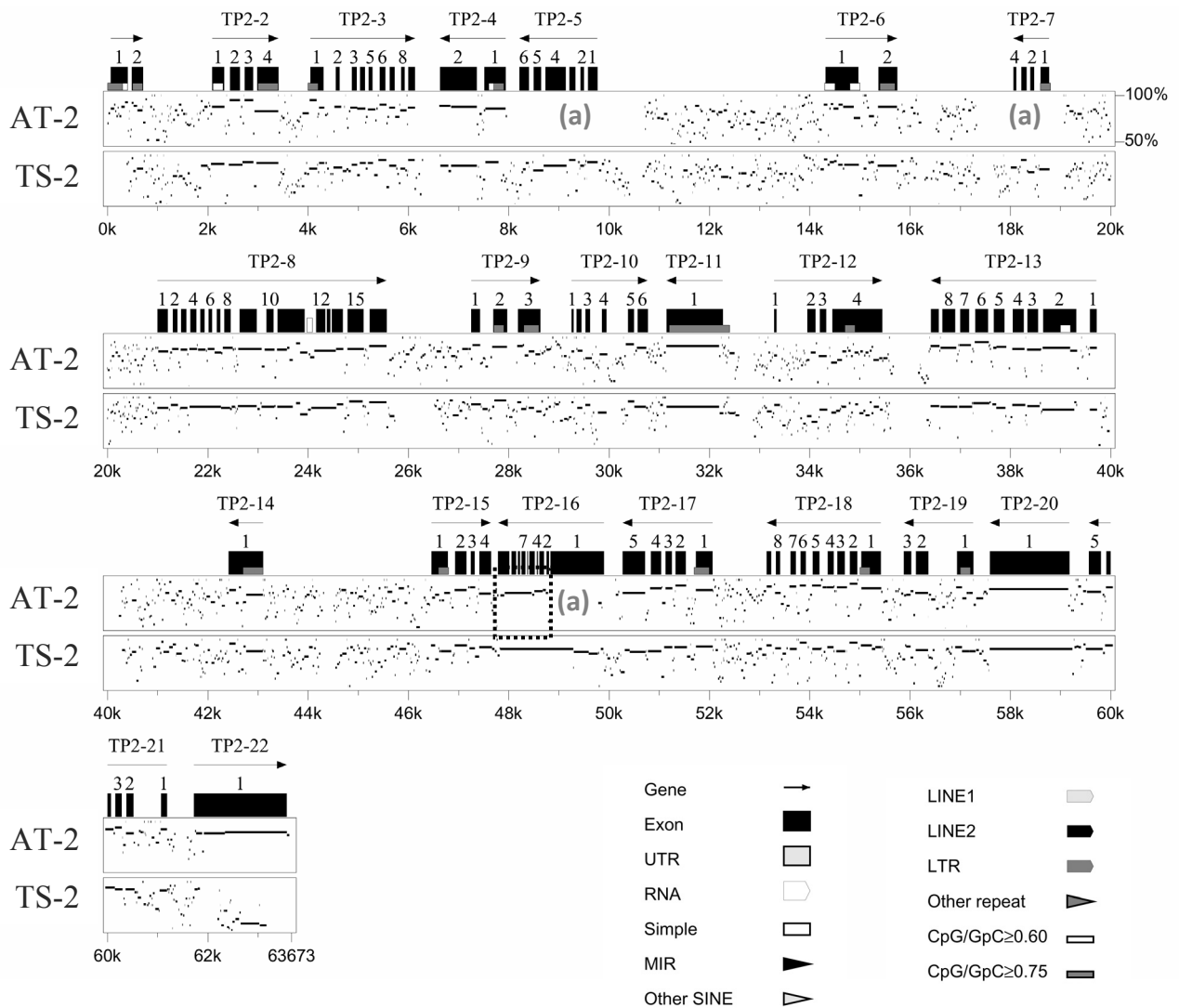
Supplemental Figure S3. Comparison between *T. parvula* genomic sequences TP-1 and TP-2 with three orthologous genomic regions in *A. thaliana* and *T. salsuginea*.

AT-1 (A) and AT-2 (B) are from *A. thaliana* and TS-1 (C), TS-2 (D) from *T. salsuginea*. Comparisons were made using dot-plot analyses with word size 7. In (A) and (B), boxes indicate regions where linearity is broken. In (C), a and b denote an inversion and the extent of the collinear region, respectively.

(A)



(B)



Supplemental Figure S4. Syntenic relationships and homologies of sequences represented by percent identity plots (pip).

TP-1 (A) and TP-2 (B) were compared to their orthologous regions in *A. thaliana* and *T. salsuginea* using PiPmaker plot with a word size of 7 (Schwartz et al., 2000). Predicted exons for the ORFs in TP sequences and the direction of transcription are indicated. Dashed box in (A) denote the intergenic space between SOS1 and its 5' neighboring gene. For the explanation of symbols (a-c) see the text.

Supplemental Table S1. ORFs identified in TP-1, AT-1 and TS-1.

TP-1 ID ^a	Evalue to AT ^b	Evalue to TS ^b	AT-1 ID ^c	TS-1 ID ^c	Homologous AT locus	AT annotation
Tp01(+)	5E-130	3E-125	At01 (+)	Th01 (+)	AT2G01900	Putative phosphatase family protein
Tp02(+)	6E-60	1E-64	At02 (+)	Th02 (+)	AT2G01905	CYC18; cyclin
Tp03(+)	0	0	At03 (+)	Th03 (+)	AT2G01910	ATMAP65-6; microtubule binding
Tp04(-)	8E-35*	2E-38*	At04 (-)	Th04 (-)	AT2G01913	unknown protein
-	-	-	At_a (-)	-	AT2G01918	calcium ion binding
Tp05(+)	4E-90	-	At_b (+)	-	AT2G01920	clathrin assembly protein-related
Tp06(-)	8E-80	1E-34*	At05 (-)	Th05 (-)	AT2G01930	BPC1; transcription factor
Tp07(+)	7E-151	2E-149	At06 (+)	Th06 (+)	AT2G01940	transcription factor/ zinc ion binding
Tp08(-)	0	0	At07 (-)	Th07 (-)	AT2G01950	BRL2; transmembrane receptor kinase
-	-	-	At_c (+)	-	AT2G01960	TET14; function unknown
Tp09(-)	0	0	At08 (-)	Th08 (-)	AT2G01970	endomembrane protein 70, putative
Tp10(+)	0	0	At09 (+)	Th09 (+)	AT2G01980	SOS1; sodium:hydrogen antiporter
Tp11(-)	4E-83	5E-97	At10 (-)	Th10 (-)	AT2G01990	unknown protein
-	-	-	-	Th_A (+)	(AT1G14610)	TWN2; valine-tRNA ligase
Tp12(-)	0	0	At11a (-)	Th11 (-)	AT2G02000	GAD3; glutamate decarboxylase
Tp13(-)	0	0	At11b (-)	-	AT2G02010	GAD4; glutamate decarboxylase
Tp14(+)	0	0	At12 (+)	Th12 (+)	AT2G02020	oligopeptide transport-related
-	-	-	At13 (+)	Th13 (-)	AT2G02030	F-box family protein
Tp15(+)	0	0	At14 (+)	Th14 (+)	AT2G02040	PTR2; peptide transporter
Tp16(+)	7E-59*	8E-19	At15 (+)	Th15 (+)	AT2G02050	NADH-ubiquinone oxidoreductase
Tp17(+)	8E-57	3E-72	At16 (+)	Th16-1 (+)	AT2G02060	transcription factor
-	-	-	-	Th16-2 (+)	AT2G02060	transcription factor
Tp18(+)	1e-149*	0	At_d ^d	-	AT2G02061	unknown protein
Tp19(+)	0	0	At17a (+)	Th17a (+)	AT2G02070	AtIDD5; transcription factor
Tp20(-)	0	0	At17b (-)	Th17b (-)	AT2G02080	AtIDD4; transcription factor
-	-	-	-	Th_B (+)	(AT2G02030)	F-box family protein
Tp21(+)	0	0	At18 (+)	Th18 (+)	AT2G02090	ETL1; helicase
Tp22(-)	2E-20*	1E-24*	AT19a (+)	Th19a (+)	AT2G02100	LCR69; peptidase inhibitor
Tp23(+)	2E-20*	9E-25*	-	Th19b (+)	AT2G02100	LCR69; peptidase inhibitor
-	-	-	At19b (+)	-	AT2G02120	LCR70; peptidase inhibitor
-	-	-	At19c (+)	-	AT2G02130	LCR68; peptidase inhibitor
-	-	-	At19d (+)	-	AT2G02135	LCR71; peptidase inhibitor
Tp24(+)	1E-18	1E-23	At19e (+)	Th19c (+)	AT2G02140	LCR72; peptidase inhibitor
Tp25(+)	5E-10	2E-22	At19f (+)	Th19d (+)	AT2G02147	LCR73; peptidase inhibitor
Tp26(-)	0	0	At20 (-)	Th20 (-)	AT2G02150	pentatricopeptide repeat-containing

a , c : signs in parentheses denote orientation of DNA strands.

b: results from blastx to the most highly related *A. thaliana* or *T. salsuginea* protein.

c : named as in Nah et al., 2009.

d: not present in Nah et al. 2009; present in TAIR9.

*: homology found only in part of an ORF.

Supplemental Table S2. ORFs identified in TP-2, AT-2 and TS-2.

TP-2 ID ^a	Evalue to AT ^b	Evalue to TS ^b	AT-2 ID ^c	TS-2 ID ^d	Homologous AT locus	AT annotation
Tp01(+)	4E-6*	N/A	At14(-)	Ts01(+)	AT5G66400	RAB18; responsive to ABA
Tp02(+)	9E-72	2E-73	At13(-)	Ts02(+)	AT5G66390	PER72; peroxidase
Tp03(+)	3E-150	4E-150	At12(-)	Ts03(+)	AT5G66380	ATFOLT1; folic acid transporter
-	-	-	At11(-)	-	AT5G66370	unknown protein
Tp04(-)	2E-164	3E-125	At10(-)	Ts04(-)	AT5G66360	rRNA adenine dimethylase family
Tp05(-)	3E-152	2E-147	-	Ts05(-)	(AT3G53210)	nodulin MtN21 family
Tp06(+)	5E-122	2E-123	At09(-)	Ts06(+)	AT5G66350	SHI; transcription factor
-	-	-	At08(-)	-	AT5G66340	unknown protein
Tp07(-)	4E-43*	2E-43*	-	Ts07(-)	(AT1G77370)	glutaredoxin, putative
-	-	-	At07(-)	-	AT5G66330	leucine-rich repeat family protein
-	-	-	At06(-)	Ts08(+)	AT5G66320	zinc finger (GATA type) family protein
Tp08(+)	0	0	At05(-)	Ts09(+)	AT5G66310	kinesin motor family protein
Tp09(+)	3E-65	4E-59	At04(-)	Ts10(+)	AT5G66300	NAC105; transcription factor
Tp10(+)	6E-58	8E-60	At03(-)	Ts11(+)	AT5G66290	unknown protein
Tp11(-)	9E-45*	2E-44*	At02(-)	Ts12(-)	AT5G66280	GMD1; GDP-mannose 4,6-dehydratase
Tp12(+)	6E-110	4E-125	At01(-)	Ts13(+)	AT5G66270	zinc finger (CCCH-type) family protein
-	-	-	At15(-)	-	AT5G66410	PLP3b; beta-tubulin binding
Tp13(-)	0	0	At16(-)	Ts14(-)	AT5G66420	unknown protein
-	-	-	At17(-)	-	AT5G66430	SAM methyltransferase family
Tp14(-)	2E-71	1E-72	At18(-)	Ts15(-)	AT5G66440	unknown protein
Tp15(+)	2E-76	3E-86	At19(-)	Ts16(+)	AT5G66450	phosphatidic acid phosphatase-related
Tp16(-)	8E-158	4E-159	-	Ts17(-)	(AT3G47840)	pentatricopeptide repeat-containing
Tp17(-)	0	0	At20(-)	Ts18(-)	AT5G66460	(1-4)-beta-mannan endohydrolase
Tp18(-)	0	0	At21(-)	Ts19(-)	AT5G66470	GTP binding / RNA binding
Tp19(-)	1E-43*	8E-25*	At22(-)	Ts20(-)	AT5G66480	unknown protein
-	-	-	At23(-)	Ts21(-)	AT5G66490	unknown protein
Tp20(-)	0	0	At24(-)	Ts22(-)	AT5G66500	pentatricopeptide repeat-containing
Tp21(-)	3E-84	7E-87	At25(-)	Ts23(-)	AT5G66510	GAMMA CA3; carbonate dehydratase
Tp22(+)	0	N/A	At26(-)	Ts24(+)	AT5G66520	pentatricopeptide repeat-containing

a, c, d : signs in parentheses denote orientation of DNA strands.

b : results from blastx to the most highly related *A. thaliana* or *T. salsuginea* protein.

c: named as in Deng et al., 2009.

*: homology found only in part of an ORF.

Supplemental Table S3. Comparison of ratio of non-synonymous (**Ka**) vs. synonymous (**Ks**) substitution rates in orthologous genes and the estimated divergence time of three species.

	TP ID	AT ID	TS ID	Ka/Ks			Divergence time		
				TP vs AT	TP vs TS	AT vs TS	TP vs AT	TP vs TS	AT vs TS
TP-1	Tp01	At02	Th02	0.1840	0.2547	0.2007	14,010,200	7,015,467	12,621,400
	Tp03	At03	Th03	0.1055	0.1124	0.1345	12,500,933	9,075,533	10,680,833
	Tp07	At06	Th06	0.1164	0.2177	0.1536	12,928,533	6,630,133	12,806,067
	Tp08	At07	Th07	0.0719	0.0678	0.0668	17,092,333	12,536,000	16,977,800
	Tp09	At08	Th08	0.0248	0.0325	0.0278	15,025,200	11,721,333	14,432,167
	Tp10	At09	Th09	0.3300	0.2666	0.2732	10,193,500	7,853,767	10,862,967
	Tp11	At10	Th10	0.3228	0.3657	0.3418	11,602,333	6,453,433	10,397,533
	Tp14	At12	Th12	0.1282	0.1529	0.0878	14,891,200	9,491,033	17,632,533
	Tp15	At14	Th14	0.0532	0.0539	0.0337	11,589,433	11,366,833	15,499,500
	Tp21	At18	Th18	0.1343	0.1145	0.0965	10,093,333	7,624,300	11,521,067
	Tp26	At20	Th20	0.1623	0.1749	0.1592	13,631,033	10,078,567	14,603,867
TP-2	Tp02	At13	Th02	0.0856	0.1082	0.0852	13,796,200	8,436,267	13,231,800
	Tp03	At12	Th03	0.0762	0.1226	0.1109	10,599,900	7,403,033	9,129,333
	Tp04	At10	Th04	0.2565	0.5008	0.2903	11,557,933	9,478,600	16,217,800
	Tp06	At09	Th06	0.2304	0.2786	0.2483	13,112,333	8,876,467	10,081,433
	Tp08	At05	Th09	0.2535	0.2438	0.3386	10,012,700	6,591,800	9,393,300
	Tp09	At04	Th10	0.1375	0.2381	0.2022	17,666,700	10,089,567	16,346,767
	Tp10	At03	Th11	0.1291	0.1799	0.1028	11,821,433	6,101,967	10,465,267
	Tp15	At19	Th16	0.2879	0.4046	0.2865	12,458,433	6,426,433	13,667,167
	Tp17	At20	Th18	0.1734	0.1138	0.1572	11,090,600	12,083,967	11,764,800
	Tp18	At21	Th19	0.0866	0.1075	0.0872	9,021,300	7,735,767	10,273,633
	Tp20	At24	Th22	0.2350	0.2818	0.2094	9,366,867	8,366,067	10,009,867
Tp21	At25	Th23	0.2372	0.2178	0.2499	8,014,600	4,788,100	7,309,500	
Average						12,264,219	8,531,497	12,431,583	
Sd						2,452,416	2,083,382	2,821,010	

Supplemental Table S4. Comparison of relative gene expression levels by qPCR.

Tissue	Species	NaCl (mM)	Relative expression levels							
			AT2G01950	AT2G01970	AT2G01980	AT2G01990	AT3G53210	AT1G77370	AT3G47840	AT5G66460
Shoot	At	0	1.00±0.07	1.00±0.20	1.00±0.09	1.00±0.08	1.00±0.23	1.00±0.14	1.00±0.12	1.00±0.10
		200	0.15±0.03	1.87±0.14	1.59±0.16	0.39±0.09	0.50±0.09	1.58±0.10	1.06±0.14	0.97±0.09
	Tp	0	0.60±0.07	0.45±0.04	2.63±0.17	1.10±0.20	2.47±0.17	2.49±0.48	1.09±0.26	0.43±0.03
		200	0.77±0.17	0.83±0.05	8.62±0.88	1.53±0.19	5.71±1.10	3.41±0.31	1.37±0.30	1.37±0.16
		350	0.37±0.12	1.18±0.11	25.10±3.06	0.76±0.06	11.22±0.64	2.44±0.11	1.23±0.25	0.89±0.24
Root	At	0	1.00±0.13	1.00±0.03	1.00±0.25	1.00±0.06	1.00±0.17	1.00±0.03	1.00±0.08	1.00±0.09
		200	0.19±0.07	1.00±0.14	2.30±0.33	0.47±0.19	1.49±0.52	1.20±0.08	1.70±0.23	0.35±0.05
	Tp	0	0.47±0.04	0.56±0.03	1.48±0.90	0.90±0.04	0.84±0.08	3.24±0.80	1.30±0.19	0.11±0.01
		200	0.69±0.09	0.78±0.06	6.97±2.59	3.50±0.91	0.53±0.07	3.96±0.86	1.79±0.30	0.42±0.04
		350	0.12±0.02	0.62±0.02	5.22±1.40	2.64±0.23	0.21±0.19	1.80±0.31	1.47±0.16	0.65±0.10

Relative mRNA levels compared to those of an EF1 α gene (orthologs of AT5G04810)

Values represented were fold differences normalized with the untreated At samples.

Shaded are relative mRNA levels of SOS1 (orthologs of AT2G01980).

Supplemental Table S5. Known *cis*-elements in SOS1 promoter regions up to 2000bp upstream of the translation start site.

Frequencies in each species followed by positions are given in brackets (for frequencies of four or fewer only); negative numbers show sites on the opposite strand. Positions are counted from 5' end of the sequence and position 2000 refers to the base pair immediately upstream to the start codon. Promoter sequences were searched against PlantCare database (Rombauts et al., 1999), PlantProm database (Shahmuradov et al., 2003), PlantPAN (Chang et al., 2008), and PLACE database (Higo et al., 1999).

Regulatory element	AT	TS	TP	Sequence	Function
ABRE RATCAL	1 (-902)	1 (583)	0	MACGYGB	"ABRE-related sequence" identified in Ca(2+)-responsive genes
L1BOXATPDF1	1 (374)	0	0	TAAATGYA	"L1 box" in AT PROTODERMAL FACTOR1 (PDF1)
TBOXATGAPB	1 (226)	0	0	ACTTTG	"Tbox" AtGAPB
DPBFCORECDC3	3 (95,902, - 905)	3 (-224, - 426, 828)	0	ACACNNG	ABA induced; embryo specific; binds DPBF-1 and 2
ABRE	1 (1883)	2 (1562, - 1617)	0	CACGTG	abscisic acid responsiveness
E5-core	0	1 (-772)	0	ACATCTG	abscisic acid responsiveness; bDPBF-1; DPBF-2
ANAERO1CONSENSUS	2 (680, 852)	0	5	AAACAAA	anaerobic genes involved in the fermentative pathway
ARE	3 (-528, - 1526, - 1490)	1 (- 1942)	0	TGGTTT	anaerobic induction
ASF1 motif CAMV	0	2 (-243, - 975)	1 (822)	TGACG	auxin and/or salicylic acid responsiveness
TGA-box	0	0	1 (1791)	TGACGTAA	auxin-responsiveness
ARF	0	0	1 (754)	TGTCTC	auxin-responsiveness; binds AT1G59750 (ARF1)
AGL3	1 (614)	1 (576)	1 (665)	NNTNCCAWWWATAGNWN	binds Agamous AT2G03710; F19B11.16.
AG	8	5	2 (191, - 372)	NTTWCCWAAWNNNGNAAN	binds Agamous AT4G18960.
C8GCARGAT	0	1 (147)	1 (538)	CWNNNNNNW	binds AGL15
PY-Box	0	0	1676	CCTTTTC	binds aleurone layer nuclear proteins
AP1	1 (382)	3 (326, - 333, 509)	0	TTTTTRG	binds AP1 At1g69120
ARR1AT	8	6	9	NGATT	binds ARR1

PIF3	2 (-203, 880)	1 (-558)	1 (-159)	GKRGGMCACGTGRMSWCK	binds AT1G09530 /BHLH8/ PAP3 in phytochrome signalling pathways
CDC5	4 (164, 245, -782, 922)	0	1 (916)	NGCTCAGCGCN	binds AT1G09770
RAV1	5	3 (402, 445, -682)	2 (-716, -719)	NNGCAACAKAWN	binds AT1G13260
Athb-1	8	13	12	NYNCAATTATTGSA	binds AT3G01470/ HAT5/ HD-ZIP-1
ANT	0	1 (529)	1 (89)	CACANWTCCRAKG	binds AT4G37750
Bellringer	5	0	1 (502)	AAATTARW	binds At5g02030/replumless/pennywis
Athb-5	7	8	9	CAATTATTG	binds AT5G65310/ATHB5/ MNA5.4
AtMYB77	1 (1024)	0	0	ATCGACGGTTGG	binds AtMYB77; induced by light, cold, and wounding
box b	1 (410)	0	0	AGCAAAGCAA	binds DOF1
Athb-9	5	15	5	NNNNGTAATGATTRCNYS	binds homeobox HD-ZIPIII proteins
MYB4	2 (-49, 681)	0	1 (214)	AMCWAMC	binds myb At4g38620
Myba	1 (1029)	0	0	CGGTTG	binds Myb proteins
PtMYB4	1 (-1095)	0	0	ACCTACA	binds PtMYB4
p33TCP20 BS1	0	0	1 (-1254)	GCCCCG	binds TCP proteins involved in growth and development
EVENINGAT	0	1 (-840)	1 (-891)	AAAATATCT	circadian control; "EE (evening element) motif"
CCA1 motif (2)	0	0	2 (889-, -928)	AAAAATCTA	circadian regulation; binds CCA1
LTR	1	1 (1535)	0	CCGAAA	cold induction
LTRECOREATCOR15	1 (-798)	0	2 (116, 183)	CCGAC	cold responsiveness
GA-3	0	0	1 (1202)	AGAAAGGAA	control ovule identity; binds BPC1
ARR10	0	2 (353, 840)	3 (546, 763, 891)	AGATHHKNN	cytokinin induction; binds At4g31920
TC-rich repeats	0	2 (560, 890)	1 (1057)	GTTTTCTTAC	defense and stress responsiveness
AtMYC2	0	3 (-426, 788, 829)	1 (871)	CACATG	dehydration and abscisic acid responsiveness; binds AT1G32640
DRE 2	1 (-1784)	0	0	GCCGAC	dehydration responsiveness
MYCATERD1	0	6	1 (-871)	CATGTG	dehydration responsiveness; Myc recognition
MYB1AT	2 (510, 546)	1 (462)	0	WAACCA	dehydration-responsive; binds myb proteins
MYCCONSENSUSAT	4 (96, 424, 430, 903)	4 (426, 788, 829,	2 (248, 871)	CANNITG	dehydration-responsiveness; Myc recognition

	893)				
WA	1 (336)	1 (27)	1 (435)	GGTCAAA	disease resistance; binds WRKY1
PB1	0	1 (-1429)	0	TGCAAAAG	endosperm specific induction; binds BPBF
TA-rich region	0	1 (1460)	0	TATATATATATATATATATATA	enhancer
Motif VI	0	0	1 (-1670)	TGGTCC	epicotyl-specific nuclear factor
ERE	0	2 (-935, -1152)	1 (-818)	ATTTCAAA	ethylene responsiveness
ACGTAT ERD1	2 (227, 904)	2 (415, 583)	4 (71, 99, 186, 824)	ACGT	etiolation-induced expression of erd1 (early responsive to dehydration)
CARG AT CONSENSUS	1 (534)	1 (506)	0	CCWWWWWGG	found in AtSOC1 MADS-box flowering-time gene
GAREAT	0	1 (-846)	1 (-897)	TAACAAR	GA-responsiveness
GARE-motif	0	2 (614, 1507)	0	AAACAGA	gibberellin-responsive element
HSE	0	0	2 (-1601, -1732)	AAAAAATTTTC	heat stress responsiveness
HEXAMERATH4	1 (-45)	0	0	CCGTCG	hexamer motif of At histone H4
circadian	1 (1451)	1 (-1679)	2 (1210, -1584)	CAANNNNATC	involved in circadian control
GT2	0	0	1 (-784)	GGTAATTA	light down regulation and dark induced; binds DF1
ACE	0	0	2 (598, 1075)	AAAACGTTTA	light responsiveness
GA-motif	0	1 (652)	1 (-854)	AAAGATGA	light responsiveness
GT1 motif	1343	0	0	AATCCACA	light responsiveness
ATCT-motif	0	0	-551, -876, -915	AATCTAATCT	light responsiveness
GAG motif	1 (-875)	0	0	AGAGATG	light responsiveness
box 4	4 (1038, 1712, 1231, 1839)	2 (845, 1762)	1 (938)	ATTAAT	light responsiveness
G-box	0	1 (-1562)	1 (-1162)	CACGTA	light responsiveness
GATA-motif	0	1 (-727)	0	GATAGGA	light responsiveness
CATT motif	1 (1949)	0	0	GCATTC	light responsiveness
GT1CONSENSUS	14	11	10	GRWAAW	light responsiveness
TCT-motif	1 (-1437)	1 (894)	0	TCTTAC	light responsiveness
I-box	1 (1817)	1 (1678)	-787	TGATAATGT	light responsiveness
box 1	0	3 (-935, -1152, -1144)	5	TTTCAAA	light responsiveness
AT1-motif	0	1 (845)	0	ATTAATTTTACA	light responsiveness and shoot specific expression

as-2-box	1 (-1697)	8	0	GATAATGATG	light responsiveness and shoot specific expression
AE box 2	0	0	1 (-1985)	AGAAACAA	light responsiveness; binds AEF
AE box 1	1 (84)	0	0	AGAAACTT	light responsiveness; binds AEF
GATABOX	12	9	14	GATA	light responsiveness; binds ASF-2; in LHClI type I Cab genes
GBF1 BS3	1883	0	0	ACACGTGT	light responsiveness; GBF1
GBF1 BS8	0	1 (1562)	0	TTACGTGT	light responsiveness; GBF1
CGTCA-motif	0	3 (-657, 1955, 1223)	2 (-520, -1791)	CGTCA	MeJA-responsiveness
MBS	0	4 (641, 1038, -1058, -1088)	0	TAACTG	MYB binding site involved in drought-inducibility
MYBCORE	2 (48, -452)	3 (-58, 78, 108)	0	CNGTTR	myb recognition
S1	0	1 (435)	0	ATGGTATT	negative regulation, binds S1F
SORLIP3AT	0	2 (492, -495)	1 (289)	TGTATATAT	one of "Sequences Over-Represented in Light-Induced Promoters (SORLIPs)
CCA1	0	0	1 (763)	AAMAATCT	phytochrome regulation of an AtLhcb; binds myb AT2G46830
RHERPATEXPA7	1 (-227)	0	0	KCACGW	Root Hair-specificity
WBOXATNPR1	1 (-504)	0	2 (-129, 821)	TTGAC	salicylic acid (SA)-induction by WRKY
TCA-element	0	1 (-1909)	0	GAGAAGAATA	salicylic acid responsiveness
SEF3 BS 1	0	0	0 (-1387)	CACCCA	seed specific enhancer; binds SEF3
SURECOREATSULTR11	0	1 (-417)	1 (-755)	GAGAC	sulfur-responsive element (SURE) in SULTR1
SV40COREENHAN	0	1 (388)	0	GTGGWWHG	SV40 core enhancer

Chang, W., Lee, T., Huang, H., Huang, H., & Pan, R. (2008). PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. *BMC Genomics*, 9, 561.

doi:10.1186/1471-2164-9-561

Higo, K., Ugawa, Y., Iwamoto, M., & Korenaga, T. (1999). Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research*, 27(1), 297-300.

Rombauts, S., Déhais, P., Van Montagu, M., & Rouzé, P. (1999). PlantCARE, a plant cis-acting regulatory element

database. *Nucleic Acids Research*, 27(1), 295-296.

Shahmuradov, I. A., Gammerman, A. J., Hancock, J. M., Bramley, P. M., & Solovyev, V. V. (2003). PlantProm: a database of plant promoter sequences. *Nucleic Acids Research*, 31(1), 114-117.

Supplemental Table S6 . List of PCR primers.

Primer	Sequence	Used for
SOS1(1052F)	GTGTTGTCATTGCTGAAGGCATT	Fig 5 (A)
SOS1(1399R)	CAAATTGGGTAGTGGATCCATTAA	Fig 5 (A)
1G77370(181F)	CGTTCAGAACGCCATCTTGCCAA	Fig 6
1G77370(314R)	CCTCTCTCTGATCAAGCTCCACAACAA	Fig 6
2G01950(2057F)	GGTCAAATCCCTGAATCTTTCTCGAA	Fig 5 (B)
2G01950(2217R)	TCCGGTAAAGGAACACCGCAAA	Fig 5 (B)
2G01970(1612F)	GAATGGTGGTGGAGATCATTCCCTATG	Fig 5 (B)
2G01970(1763R)	AAACCGTAACAAATGCAAGCCATGTA	Fig 5 (B)
SOS1(925F)	ATGACTTTGGGCATGTTTTATGCTGCA	Fig 5 (B)
SOS1(1075R)	CCTTCAGCAATGACAACACCACTGAGGA	Fig 5 (B)
2G01990(274F)	AGTCAGTTTGTCACTCCACATGTTA	Fig 5 (B)
2G01990(425R)	GATGGTTCTGTTGGAGATTTTTCTCG	Fig 5 (B)
3G47840(816F)	GAGGGATGTGGTTTCGTGGACAAG	Fig 6
3G47840(994R)	GCTGCTCGCCCCACACAAGTCT	Fig 6
3G53210(166F)	TCTCTGTTCTTGCTCCCTCTGCTT	Fig 6
3G53210(313R)	GAGGTATTGTCCAATCCAAAGATGTA	Fig 6
5G66460(786F)	TGGCTTGGGAGCTCATGAACGA	Fig 6
5G66460(920R)	CTTCGAGGCCAGCTTCAAGCA	Fig 6
5G60390(953F) ^a	CAATGTCAAGAATGTTGCTGTCAAGGAT	Fig 5 (B) and 6
5G60390(1090R) ^a	TCTGACCAGGGTGGTTCATGATGAT	Fig 5 (B) and 6

^a used for the reference gene in all qPCR.

Primers for the reference genes shown in Figure 5 (A) were the same as Oh et al. 2009.

Oh DH, Leidi E, Zhang Q, Hwang SM, Li Y, Quintero FJ, Jiang X, D'Urzo MP, Lee SY, Zhao Y, Bahk JD, Bressan RA, Yun DJ, Pardo JM, Bohnert HJ (2009) Loss of Halophytism by Interference with *SOS1* Expression. *Plant Physiol* 151: 210-222