A computational analysis of Salt Overly Sensitive 1 homologs in halophytes and glycophytes

Cherin E. Kim¹ and Ray A. Bressan²

¹ West Lafayette Jr./Sr. High School, West Lafayette, IN, USA

² Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, IN, USA

Corresponding Author: Ray A. Bressan Department of Horticulture and Landscape Architecture, Purdue University, 625 Agriculture Mall, West Lafayette, IN, 47907, USA bressan@purdue.edu

Abstract

Soil salinity is one of the most serious impediments to global agricultural productivity. Although most terrestrial plants are glycophytes which cannot tolerate high salt concentrations, a small fraction of species are halophytes. Exactly what allows these extremophile plants to survive in saline conditions is not yet well understood. Several studies have established the Salt Overly Sensitive (SOS) pathway as the canonical model for the mechanism responsible for salt tolerance. The SOS pathway involves interplay among Na⁺-H⁺ antiporters for transporting sodium, and the activation of the kinase that phosphorylates the transporter. Among them, SOS1, a plasma membrane Na^+ -H⁺ antiporter, has been shown to be a critical component for maintaining salt homeostasis by pumping sodium out of cells upon activation. Therefore, it is of great interest to evaluate any differences of SOS1 in halophytes as compared to glycophytes. Here we report a computational analysis of the primary and secondary structures of eight halophytes and seven glycophytes. ClustalW alignment of the protein sequences as a whole reveals no regions conserved specifically in only halophytes or in only glycophytes. In addition, the key regulatory residues at the C-terminus of SOS1, S1136 and S1138, which were shown to be the phosphorylation sites by the kinase SOS2, were completely conserved in all 15 halophytes and glycophytes. The four amino acids, G136, R365, G777, and G784, in which alterations affect the function of SOS1, are mostly conserved in the 15 species. The 14-3-3 binding site in the C-terminus which is important in the phosphorylation step of SOS1 in the SOS signal transduction cascade is also well conserved. Furthermore, the number of transmembrane helices for each species is between 9 and 12 and there is no significant difference between halophytes and glycophytes. If halophytes present any special feature of SOS1, it likely involves the presence (halophytes) or absence (glycophytes) of a SOS1-interacting component.

Introduction

Salt stress is a major abiotic stress factor that significantly impacts agricultural output. Approximately 20% of the world's cultivated land and nearly half of all irrigated lands are affected by salinity (Zhu, 2001). High salinity can adversely affect plants by inducing physiological responses at the cellular level such as ion imbalance and hyperosmotic stress as well as by altering outward physical characteristics through wilting, drying, and death (Zhu, 2001). However, plants differ in their ability to tolerate high salt concentrations. Halophytes, which make up around 1% of the world's flora, are an extremophile group of plants that are able to survive in highly saline soils (Flowers and Colmer, 2008; Flowers and Colmer, 2015). In contrast, the majority of plants, including most commercial crops, are glycophytes, which are not necessarily salt-intolerant but cannot survive in environments where the salt concentration is too high (Garg et al., 2013; Ji et al., 2013). Even so, the line separating halophytes from glycophytes is not discrete. There is actually a continuum of salt tolerance degrees with very salt-sensitive glycophytes at the low end and highly salt-tolerant halophytes at the high end (Rozema and Schat, 2012). Furthermore, within halophytes, many dicotyledonous halophytes grow optimally in 50-250 mM NaCl while monocotyledonous halophytes prefer less than 50 mM NaCl (Flowers and Colmer, 2008). We limit our species to dicotyledonous halophytes only in order to minimize any conflicting effects due to phylogenetics on salt tolerance ability.

The factors that make halophytes different from glycophytes are not yet well understood. However, the mechanism(s) by which plants respond to increased salt concentrations has been extensively studied. The protein kinase cascade involving Salt Overly Sensitive (SOS) 1,2, and 3 has emerged as a predominant model for salt stress responses, as shown in Fig. 1. The SOS signal transduction pathway is activated when SOS3, a calcium-binding protein, binds with Ca²⁺, which is well-known as a "warning" signal for Na⁺ toxicity (Ji et al., 2013; Zhu, 2001). SOS3 then interacts with and activates SOS2, a serine/threonine kinase, and SOS2 phosphorylates SOS1, a plasma membrane Na⁺-H⁺ antiporter, resulting in Na⁺ being transported out from the cell.

SOS1 in particular has been identified as playing a key role in the reactions of plants to salinity stress. Mutants of SOS1 display stunted growth and significant leaf discoloration in environments with high Na⁺, whereas transgenic *sos1-1* mutants containing the wild-type *SOS1* gene controlled by the cauliflower mosaic virus 35s promoter show no or little difference in

growth in salt compared with the wild type, demonstrating that the *SOS1* locus is necessary, but not necessarily sufficient, for salt tolerance (Shi et al., 2000). In addition, *SOS1* expression is upregulated in plants in response to NaCl stress (Shi et al., 2000). The C-terminal region, where phosphorylation sites and 14-3-3- binding sites are located, is also proposed to be important for SOS1 function in plant salt tolerance (Shi et al., 2000).

In this study we present the comparison of SOS1 protein sequences of eight halophytes and seven glycophytes. We looked for distinctly conserved regions as a whole, at key mutation positions, and near the C-terminus, and we determined the number of transmembrane helices in each species to analyze secondary structure.

Materials & Methods

SOS1 and SOS1 homologs were identified using the BLASTp search of the NCBI protein database, using previously identified SOS1 from Arabidopsis from The Arabidopsis Information Resource (TAIR) gene database (arabidopsis.org). The top BLASTp hit was used as the homolog for each species. Protein sequences were aligned using ClustalW alignment in the sequence alignment editor BioEdit. The number of transmembrane helices were determined using the TMHMM Server version 2.0 (http://www.cbs.dtu.dk/services/TMHMM/), which predicts transmembrane helices in proteins.

Results & Discussion

To identify any regions conserved uniquely in either halophytes or glycophytes in SOS1, we first collected the SOS1 protein sequences of eight halophytes and seven glycophytes, using NCBI Blast with *Arabidopsis thaliana* SOS1 as a query. The halophyte species analyzed were *Eutrema parvulum, Eutrema halophila, Salicornia dolichostachya, Aeluropus littoralis, Suaeda salsa, Salicornia brachiata, Mesembryanthemum crystallinum,* and *Distichlis spicata* (Table 1). The glycophyte species analyzed were *Arabidopsis thaliana, Solanum lycopersicum, Oryza sativa* Indica group, *Sorghum bicolor, Gossypium hirsutum, Selaginella moellendorffii*, and *Physcomitrella patens* (Table 1). ClustalW alignment was used to compare the sequences at the primary level. Although there is variation among species, no differentiating regions between halophyte and glycophyte protein sequences were found at the primary level. That is, there were

no regions where all of the halophyte sequences had the same (or similar) amino acid and all of the glycophyte sequences did not have that amino acid, or vice versa (Fig. 2).

Several *sos1* mutant alleles in *Arabidopsis thaliana* were initially isolated from a mutant screen looking for mutants having altered salt response. When treated with 100mM NaCl, *sos1* mutants showed sensitive responses such as arrested growth, chlorosis in older leaves, and dark color in young leaves., The genetic mapping study revealed that four of these alleles (*sos1-3*, *sos1-8*, *sos1-9*, and *sos1-12*) had missense mutations in SOS1 which result in amino acid substitutions such as Gly 136 to Glu, Arg 365 to Cys, Gly 777 to Glu, and Gly 784 to Asp in *sos1-3*, *sos1-8*, *sos1-9*, and *sos1-12*, respectively. This indicates that the amino acids at these particular positions are crucial (Shi et al., 2000). Sequence alignment showed that amino acids are mostly conserved at these crucial missense mutation areas (Fig. 3A). At region A, only *S*. *brachiata* showed a deviation from the amino acids as its sequence had arginine instead of glycine at this position. Mutation regions B and D were completely conserved in all species and at region C, there was another substitution of arginine for glycine, this time with *S. lycopersicum*.

It has been proposed that there are key regulatory residues at the C-terminus of SOS1. Previously, 1136 Serine and 1138 Serine in the C-terminus of Arabidopsis SOS1 were shown to be the phosphorylation sites by SOS2 (Quintero et al., 2010). These two amino acid positions were completely conserved in all 15 halophytes and glycophytes (Fig. 3B). Arabidopsis SOS2 phosphorylates these serine residues and facilitates binding by a 14-3-3 protein. An alignment of the SOS1 protein sequences of *A. thaliana* and *P. patens* shows strong conservation, suggesting a common mechanism of sodium tolerance shared by plants (Kleist et al., 2014). The sequences were also strongly conserved at this 14-3-3 binding site in the C-terminus (Fig. 3B). These data show that since both the sequences as a whole, key mutation areas, and important binding sites are all mostly totally conserved, halophyte-glycophyte differences are most likely not evident at the primary sequence level.

The number of transmembrane helices was determined for each species. All species had between 9 and 12 transmembrane helices but no distinct pattern in either glycophytes or halophytes was found.

In conclusion, there appears to be no major differences in SOS1 between halophytes and glycophytes at the primary and secondary levels. Considering that SOS1 is required but not sufficient for improved salt tolerance in other species as well as SOS1's pivotal role in the

signaling cascade, the regulation interplay of *SOS1* and related genes, rather than their sequences or structures, should be further investigated to differentiate halophytes and glycophytes from a molecular perspective. Alternatively, a comparison of SOS1-interacting proteins obtained from halophytic and glycophytic species may be a more promising strategy to find factors that affect the differential role of SOS1 in halophytes and glycophytes.

With the explosion of the size of the human population in recent years, it has become increasingly relevant to explore alternatives to traditional food sources. Considering that a large percentage of land cannot be used to cultivate agricultural crops due to high soil salinity, halophytes may be used as models to transgenetically improve conventional crops or be domesticated for mass consumption in their own right (Flowers and Colmer, 2015). Therefore, further research involving halophytes is necessary in order to understand the underlying mechanisms of plant salt tolerance.

References

- Flower TJ, Colmer TD. 2015. Plant salt tolerance: adaptations in halophytes. *Annals of Botany* 115:327-331. DOI: 10.1093/aob/mcu267.
- Flowers TJ, Colmer TD. 2008. Salinity tolerance in halophytes. *New Phytologist* 179:945-963. DOI: 10.1111/j.1469-8137.2008.02531.x.
- Garg R, M Verma, S Agrawal, R Shankar, M Majee, M Jain. 2013, Deep Transcriptome Sequencing of Wild Halophyte Rice, *Porteresia coarctata*, Provides Novel Insights into the Salinity and Submergence Tolerance Factors. *DNA Research*. DOI: 10.1093/dnares/dst042.
- Ji H, JM Pardo, G Batelli, MJ Van Oosten, RA Bressan, X Li. 2013. The Salt Overly Sensitive (SOS) Pathway: Established and Emerging Roles. *Molecular Plant* 6:275-286. DOI: 1 0.1093/mp/sst017.
- Kleist TJ, AL Spencley, S Luan. 2014. Comparative phylogenomics of the CBL-CIPK calciumdecoding network in the moss *Physcomitrella*, Arabidopsis, and other green lineages. *Frontiers in Plant Science* 5:187. DOI: 10.3389/fpls.2014.00187.
- Quintero FJ, J Martinez-Atienza, I Villalta, X Jiang, WY Kim, Z Ali, H Fujii, I Mendoza, DJ Yun, JK Zhu, JM Pardo. 2010. Activation of the plasma membrane Na/H antiporter Salt-

Overly-Sensitive 1 (SOS1) by phosphorylation of an auto-inhibitory C-terminal domain. *PNAS* 108:2611-2616. DOI: 10.1073/pnas.1018921108.

- Rozema J, H Schat. 2013. Salt tolerance of halophytes, research questions reviewed in the perspective of saline agriculture. *Environmental and Experimental Botany* 92:83-95. DOI: 10.1016/j.envexpbot.2012.08.004.
- Shi H, M Ishitani, C Kim, JK Zhu. 2000. The Arabidopsis thaliana salt tolerance gene SOS1 encodes a putative Na⁺/H⁺ antiporter. PNAS 97:6896-6901. DOI: 10.1073ypnas.120170197.

Zhu JK. 2001. Plant salt tolerance. Trends in Plant Science 6:66-71.

Table 1. SOS1 homologs of 9 halophytic species and 8 glycophytic species.

Species Name	Reference	Number of transmembrane helices	Gene Name	Resources
Halophytes				
Eutrema parvulum	Oh et al. (2010)	12	>gi 312190386 gb ADQ43186.1 salt overly sensitive 1 [Schrenkiella parvula]	http://thellungiella.org, NCBI
Eutrema halophila	Wang et al. (2006)	12	>gi 229914867 gb ACQ90592.1 salt overly sensitive 1-like protein [Eutrema halophilum]	http://thellungiella.org, NCBI
Salicornia dolichostachya	Rozmea and Schat (2013)	11	>gi 781786798 emb CDL70804 .1 plasma membrane Na+/H+ antiporter [Salicornia dolichostachya]	NCBI
Aeluropus littoralis	Rozmea and Schat (2013)	13	>gi 359843922 gb AEV89922.1 plasma membrane Na+/H+ antiporter [Aeluropus littoralis]	NCBI
Suaeda salsa	Rozmea and Schat (2013)	12	>gi 584609172 gb AHJ14584.1 Na+/H+ antiporter [Suaeda salsa]	NCBI
Salicornia brachiata	Rozmea and Schat (2013)	10	>gi 214028396 gb ACJ63441.1 salt overly sensitive 1 [Salicornia brachiata]	NCBI
Mesembryanthemum crystallinum		10	>gi 124301347 gb ABN04858.1 salt-overly-sensitive 1 [Mesembryanthemum crystallinum]	NCBI
Distichlis spicata		12	>gi 301087449 gb ADK60916.1 plasma membrane Na+/H+ transporter [Distichlis spicata]	NCBI
Glycophytes				
Arabidopsis thaliana	Essah et al. (2003)	9	>gi 22325422 ref NP_178307.2 sodium/hydrogen exchanger 7 [Arabidopsis thaliana]	https://www.arabidopsis.org
Solanum lycopersicum		12	>gi 350535282 ref NP_0012346 98.1 plasmalemma Na+/H+ antiporter [Solanum lycopersicum]	NCBI
Oryza sativa Indica group	Rozmea and Schat (2013)	12	>gi 218187326 gb EEC69753.1 hypothetical protein OsI_39290 [Oryza sativa Indica Group]	NCBI
Sorghum bicolor	Wang et al. (2014)	12	>gi 242086498 ref XP_0024436 74.1 hypothetical protein SORBIDRAFT_08g023290 [Sorghum bicolor]	NCBI
Gossypium hirsutum	Rozmea and Schat (2013)	12	>gi 846328855 gb AKN19929.1 salt overly sensitive 1 [Gossypium hirsutum]	NCBI
Selaginella moellendorffii		11	>gi 302753756 ref XP_0029603 02.1 hypothetical protein SELMODRAFT_75049, partial [Selaginella moellendorffii]	NCBI
Physcomitrella patens		12	>gi 33352144 emb CAD91921. 1 putative Na/H antiporter [Physcomitrella patens]	NCBI

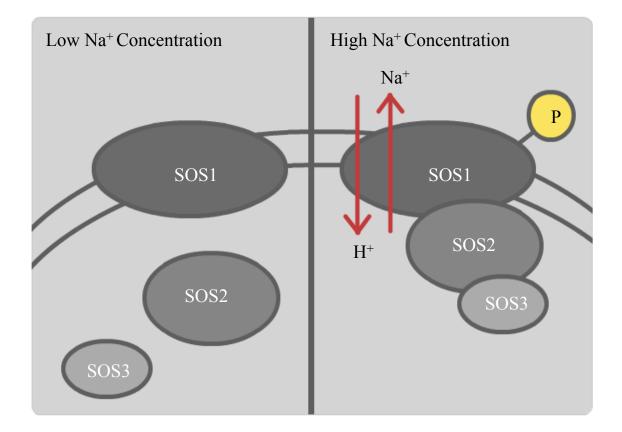


Figure 1. Model for salt response in plants.

Under highly saline conditions, the Salt Overly Sensitive (SOS) pathway is triggered when SOS3, a calcium-binding protein, activates SOS2, a serine/ threonine kinase, which in turn phosphorylates SOS1, a Na⁺/H⁺ antiporter.

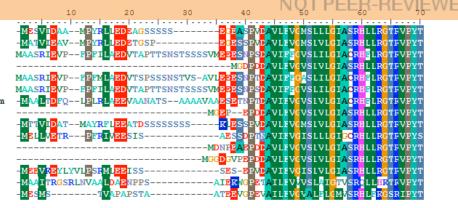
Figure 2 Peer Preprints

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella_patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella patens



80	90	100	110	120	130	140
				ALLEBSAESMEV		
VALIVIGIALGS-11GT						
VALLITGIGLGSLEVGT				ALLEESSESMET		
VALLVVGVALGSLEYGT						
VALLVIGIGLGSLRYGT			_			
VALLLIGIGLGSLEYGT	KHG <mark>A</mark> GREGI	GIRIWENID	PELLLAVELE	ALLFESSFSMEI	HQIRR<mark>CIA</mark>Ç	NVLL
N VALLII <mark>GIGLG</mark> SLEYGT	KHGLG <mark>R</mark> G <mark>D</mark>	GIRLWEHID	P <mark>E</mark> LLL <mark>A</mark> VELE	ALLEESSESMEV	HQIKR <mark>C</mark> I <mark>I</mark> Ç	MLIL
V <mark>A</mark> LLVL <mark>GIALGSLEYG</mark> T						
VALLVIGIALGSLEYGA						
VALIVI <mark>GIGICALEY</mark> GT						
VALIVLEV ^A LGSLEFGT VALIVLEV ^A LGSLEVGT				ALLFESSFSMEI ALLFESSFSMEV		
VALLIIGLGLGSLEYGT						
VALLVIGMGLGALEYGT						
VILLLE GIGIGGLEYGV	~			ALLFESSFAMD		

150	160	170	180	190	200	210
<u> </u>	<u> </u>				.	·
AGPCVLISTFCLAT	LIKLTFPYCWD	WRTSLLEGGI	LSAT <mark>DEVA</mark> VV.	ALLRELCASR	RISTVI <mark>EGE</mark> SI	MND-
AGPEVLISTFCLAS	IVRLTEPYDWD	WRTSLLLGGL	LSATDEVAVV.	ALLRELCASE	RLSTVI <mark>EGE</mark> SI	MNDG
AGPOVLISTFCIGN	ALKLSEPYDWN	WITSLLLGGI	LSATDEVAVV.	ALLERICASE	RLSTIIEGESI	MNDG
AGPGVIMSTFLLG						
AGPGVLL STCILG ⁴						
AGPEVLISTECIGE						
AGPGVLISTECLGS						
AGPCVVISTFLLGT	LIRVTFPYNWS	WRISLLLGGI	LSAT <mark>DEVA</mark> VV.	ALLRELCASR	RLSTII <mark>EGE</mark> SI	MNDG
AVPCVLISTACLGS	IVEVTEPYEWD	WRTSLLLGGL	LSATDEVAVV.	ALLRELCASE	RLSTII <mark>EGE</mark> SI	MNDG
AGPCVLISTFELG4	ALKIAFPYNWS	WSTSLLLGGL	LSATDEVAVV	ALLE	RLGTIIEGESI	MNDG
AGPGVLISTFELGS						
AGPEVVVSTVLLG						
AGPGVIISTFCLGE						
AGPGVVISTFCLGL						
A PCVLISTFLLGV	<mark>A IVKV F PY</mark> NWN	WSTGLLLGGI	LSATDEVAVV.	ALLRELCASE	RLSTLI <mark>EGE</mark> SI	MNDG

220	230	240	250	260	270	280
TATVVEOLEIRNVMGES	STWASTTARI					 V = A
TAIVVFOLFLERVMGHS						
IAIVVYOLFYCMALGRS						
TAIVVYQLFYFMVLG	FN <mark>AG</mark> ÇVIRFL	TEVSLCAVEL	GLAFGIA SVL	WL <mark>GFIFND</mark> TI	IEIALTVAVS	AI
T <mark>A</mark> IVVYQLFYÇMVL <mark>GR</mark> S						
IAIVVY <mark>ILFY</mark> CMVLGRS						
TAIVVYQLFFCMVLGRS						
TAIVVYQLFLFNVLGR TAIVVFOLFLFMMMGON						
TAIVVYOLLIFNVIGY						
TAIVVYOLFYFRVLGR						
TAIVVYQLFYFMVLGR1	F <mark>D</mark> AGSIIRFL	SEVSLCAVAL	GLAFGINSIL	WL <mark>GFIFND</mark> TI	IEIALTLAVS	AIY
T <mark>A</mark> IVVYQLFFRMVN <mark>GE</mark> S						
TANVVFELFECNVLGR	IF SAGD I CRF L					
TAIVVERLE<mark>L</mark>ÇMVL<mark>G</mark>KI	FSAGDVVSFL	SFVVF <mark>CA</mark> VAL	GL <mark>LFG</mark> IVSVI	NV <mark>G</mark> VENDIV	IEITLTL <mark>TA</mark> S	INA

Figure 2. Amino acid sequence alignment of 8 halophytic and 7 glycophytic species

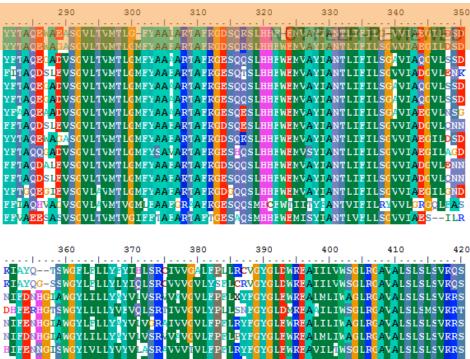
Figure 2 (continued)

Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella_patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum SelaginelPeen@Pentesdortfst//tioi.org Physcomitrella_patens



NIFENHGIAWGYLELLYYVVIVER IVVCVLFPELRYFGYGLEWREALILIWAGLRGAVALSLSLSVRRS NIFDNHGIAWGYLILLYYVVIVER IVVCVLFPELRYFGYGLEWREALMLIWAGLRGAVALSLSLSVRRS HIFENNGISWGYLVLLYVVIASR VVVCVLYPLLRFGYGLEWREAVILIWSGLRGAVALSLSLSVRRS VHFERHGTSWGFLLLLYVYTGLSR VVVCVLYPLLRFGYGLEWRESIIIVWSGLRGAVALSLSLSVRRS NIFKTYDNSWGYLILLYVIGLSR VVVCVLYPLLCRFGYGLEWRESIIIVWSGLRGAVALSLSVRRS NFFRHGASWGFLLLLYVYTGLSR VVVCVLYPLLRFGYGLEWRESIIIVWSGLRGAVALSLSVRRS NFFRHGASWGFLLLLYVYTGLSR VVVCVLYPLLRFGYGLEWRESIIIVWSGLRGAVALSLSVRRS NFFRHGASWGFLLLLYVYTGLSR VVVCVLYPLLRFGYGLEWRESIIIVWSGLRGAVALSLSVRRS NFFRHGASWGFLLLLYVYTGLSR VVVCVLYPLLRFGYGLEWREATIIVWSGLRGAVALSLSVRRS SSORSTGRUWFYLLLYFFLGLSR IVVCVLYPLLRFGYGLEWREATIIVWSGLRGAVALSLSVRRS SSORSTGRUWFYLLGUSR IVVCVLYPFLRYFGYGLEWREATIIVWSGLRGAVALSLSVRRS SSORSTGRUWFYLLGUSR IVVCVLYPFLRYFGYGLEWREATIIVWSGLRGAVALSLSVRRS SSORSTGRUWFYLLGUSR IVVCVLYPFLRYFGYGLEWREATIIVWSGLRGAVALSLSVRVF SSORSTGRUWFYLLGUSR IVVCULYPFLRYFGYGLEWREATIIVWSGLRGAVALSLSVRVF

430	440	450	460	470	480	490
· · · · · · · · · · · · · · ·						
SGNSFLSTETGTI	L <mark>CIFFT</mark> GGIVI	LTLIINGST	TQFALRLL <mark>C</mark> M	DGLE <mark>ATKI</mark> RI	I L <mark>DYTRFE</mark> MLN	<mark>F</mark> A L
SGNSFLSTETGT	AFIFFTGGIVI	LTLIVN <mark>G</mark> ST	TQFALRLLEM	D <mark>gleaskir</mark> i	IL <mark>dytrye</mark> mln	<mark>F</mark> A L
SGDPT-LLSIQTGTI						
SDAVHPYLKPEDGT1						
SGDPA-LLSTQTGTI						
SGDPT-LLSAQTGTI						
SGDAS-LLSAQTGTI						
SDAVQSYLKPEVGT1						
SGNSHISRETGTI						
SDGSQ-YISSDTGTI						
SDAVQTHLKPVDGI						
SDAVQP <mark>YLKEE</mark> NGTI						
S <mark>G</mark> GSF-NLSS <mark>E</mark> TGS						
YLSQEEYKQRL-DI						
GQHDPTQNTYLSKLTEAI	REVEET <mark>cc</mark> vv	LTLIINCST	TQELLLELKM	DTITETKSEN	ILEFARYEMMS	F AL

	500	510	520	530	540	550	560
		ADWRTVENYI	SSLKDSEC	ECVH-PHSGS	KTGNLD	HTSLEDTRIR	
	AFEDLGDDEELG-1						
I	AFGDLGDDEELG-1	E <mark>AD</mark> WPTVKRYI	KSLNTLDO	ERIH-PHEND	GPSETDG <mark>YL</mark> D	EMNLK <mark>CMFVR</mark> I	LN <mark>C</mark> V
1	E <mark>AFGELR</mark> DDEEL <mark>G</mark> -1	ADWVTVRRYI	RCLDEDI	EFVH-PHQVS	GT <mark>DDYM</mark> H	ITRNIK <mark>DIFVR</mark> I	LN <mark>C</mark> V
I	AFGDLGEDEELG-1	ADWATVRRYI	KSLNTLDE	ERIH-PHEAS	G-TEN <mark>DGYL</mark> D	EMNLK <mark>E</mark> MFVRI	. LN <mark>C</mark> V
I	D <mark>AFGDLGDDEELG</mark> -1	E <mark>ADWPTVRRY</mark> I	KSLNTLDC	ERIH-PHEND	GPSETDG <mark>YL</mark> D	EMNLK <mark>EMF</mark> VRI	.LN <mark>C</mark> V
a I	E <mark>AFGDLGEDEELG</mark> -1	E <mark>ADWPTVRRY</mark> I	KSLNSG-C	EQIH-PHDNS	TS <mark>GGDLD</mark>	EMSLT <mark>DIFVR</mark> I	LNCV.
_	E <mark>AFGELR</mark> DDEEL <mark>G</mark> -1						
1	R <mark>AFQ</mark> DLGDDEELG-I	E <mark>ADWPIVE</mark> SYI	SS <mark>L</mark> KGS <mark>E</mark> C	E <mark>lvhh</mark> phngs	KI <mark>G</mark> SLD	PRSLEDIFME	LNCV
	E <mark>AFGDL</mark> GDDEELG-1	E <mark>ADWPTVKRY</mark> I	TSLN <mark>DVE</mark> C	EFVH-PHTSS	GN <mark>DDNV</mark> D	EM <mark>HLEDIRIR</mark> I	LNCV.
	E <mark>AFGDLR</mark> DDEEL <mark>G</mark> FI	ADWVTVRRYI	TCLNDLDI	EE <mark>VH-PHAV</mark> S	DRNDRM	(TMNL <mark>RDIF</mark> VRI	LNCV.
	ESFGEL <mark>R</mark> DDEELG-1	ADWITVRRYI	TCLNDLDN	EPEH-PHDVS	G <mark>R</mark> DDE <mark>M</mark> E	IMNLT <mark>DIFVR</mark> I	LNCV.
	E <mark>AFE</mark> DLMDDEELG-1	E <mark>AD</mark> WPTVKRYI	TSLNDLEC	DFVH-PHTES	EADNDLD	PSNLRDIRIRI	LNCV
rg/1	0A728FLodEDIBELofini	AD668VVRCC-E	3YC4.0 Open Ac	icess rec: 21 ja	an-2016 <mark>2 publ</mark>	21 Jah 2016 LR	LNCV
	E <mark>AFGELGDDEEL</mark> G-1	ADWNTIRRYV	SCLSYEEN	KPT <mark>H-PH</mark> GPP	SFALP <mark>DG</mark> EYA	KQQMIDIRIR	LNGL

Figure 2 (continued)

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella_patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella_patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella patens

S	570	580 NTANILMRSV	590 DEALDEVSTE	600 S- <mark>LCDWRG</mark> LK	610 ·····		
	CAAYWEMLDEGRISE CAAYWYMLDEGRITC						IP <mark>G</mark> RLV PKRLV
	CAAYWCMLEEGRINO CAAYWYMEDEGRITO	TTANILMRSV	DE <mark>AI</mark> DIVSTI	P-LCEWNGLR	SSVRFPSYYR	YLQMSRI	P <mark>RR</mark> LI PRRLV
	CAAYWMLDEGRITO CAAYWMLDEGRISO	NTANALMQSV	DE <mark>ALDEV</mark> DHE	P-L <mark>CDWR</mark> GLR	NSVQFP <mark>TYY</mark> R	FLQSSI	PRRLV
	Ç <mark>aaywe</mark> ml <mark>eegr</mark> itq	AT <mark>A</mark> NILM <mark>R</mark> SV	DE <mark>AMD</mark> LIS <mark>G</mark> Q	P-LCDWRGLQ	SNVQFPSYYR	FLQMSRI	PRRLI
	Ç <mark>AI</mark> YWEMLDEGRISE Ç <mark>AAYWE</mark> MLNEGRIEQ	TIANLLMQSV	EEANDVVSHE	P-LCCWRGLR	SMVNI PNYYR	FLQTSE	PRRLV P <mark>RRLV</mark>
	Ç <mark>AAYWE</mark> ML <mark>EEGR</mark> ITQ Ç <mark>AAYWE</mark> ML <mark>EEGR</mark> ITQ	AT <mark>A</mark> NILM <mark>R</mark> SV	DE <mark>AMD</mark> IVSEQ	K-LCDWRGLR	SNVQFPNYYF	FLQMSRI	
	Ç <mark>AAYW</mark> GMLDEGRI <mark>4</mark> Q Ç <mark>AAYWY</mark> ML <mark>EEGR</mark> ITQ						
	Ç <mark>AAY</mark> W <mark>MLDEGR</mark> ITQ	TA <mark>AL</mark> ILMQSI	DE <mark>ALD</mark> EV KHH	TALO <mark>DWE</mark> GLR	EHI FPGY R	YSWRSTSVI	.PRRAS

640	650	660	670	680	690	700
TYF ⁴ VERLESACYI						
IYF ⁴ VDRLESACYI						
TFFTVERLESACSI						
TNETVERLESCYI						
TFFTVERLESACST						
TFFTVERLESACSI	AAFLE <mark>AH</mark> RI	VRGQLODEVGDS	EVS <mark>FA</mark> IINES	E <mark>A</mark> EGEE <mark>A</mark> RK	ELEDGRITEP(VLFV
TFFTVERLESCYI	AAFLF <mark>AHRI</mark>	ARRQLYDFIGES	EI <mark>A</mark> TAVINES	ETEGEE <mark>AR</mark> R	ELECVRITEPE	VLFV
TYFTVERLESCY1						
TYF ² VERLESACYI						
TYFTVERLESACY1						
TYFIVERLESCYI						
TYFIVERLESCCY1						
TFFIVERLENCCI TY VERLEISCYI						
TYPUTOOLELACYT	AAFLEAHEM		EVERAVIORS	FACACEARU	ELEGIRLIES	VIRA

710	720	730	740	750	760	770
VRTRÇVTYAVLNHLLI						
VRTRÇVTYSVLSHLL						
VRTRQATYAVLOHLI						
LRTRÇVTYSVLTHLS						
VETRQATEAVLOHLI						
VKTRQATYVLQRLI						
VKTRÇVTHSVLKHLI	YIESLEK (SLL <mark>EEKE</mark> IHH	L <mark>HDA</mark> VQT <mark>DLR</mark> I	VLRNPPLVRI	.PKARDLVSTHI	LLCA
LKTRÇVTYSVLTHLS	YIQNLQK I	SLL <mark>EERE</mark> MIHI	L <mark>DDA</mark> LQT <mark>DLR</mark> I	KL <mark>KR</mark> NPEMVRM	IPFVCDLLNTHI	PLV <mark>G</mark> V
VRTRÇVTYSVLNHLL	YIENLER/	GLL <mark>EER</mark> EIAHI	L <mark>HDA</mark> VQT <mark>GLR</mark> I	RLLRNPPIVRI	.PRL <mark>SDMI</mark> IS <mark>H</mark> I	LSVA
VRTRÇVTYAVLNHLI	YVENLEK (SIL <mark>EERE</mark> MAHI	L <mark>HDA</mark> VQT <mark>DLR</mark> I	RLVRNPPLVK	PRIRDLISVNI	PLL <mark>CA</mark>
LRTRCVTYSVLTHLSE	YIQNLQK I	GLL <mark>EERE</mark> MAHI	L <mark>DDA</mark> LQT <mark>DLR</mark> I	K <mark>ekr</mark> npplvri	IPFVS <mark>DLLNTH</mark> I	PLV <mark>CA</mark>
LRTRCVTYSVLTHLSE	YIQNLQK (LLEERENVOI	L <mark>DDA</mark> LQT <mark>DLR</mark> I	RL <mark>OR</mark> NPPIVRM	IPFVS <mark>DLLNTH</mark> I	PLV <mark>CA</mark>
VRTRCVTYSVLNHLI	YLENLEKA	LLEEREMLHI	LHDAVQTDLR	RLLRNPPLVRI	PRINDIISAHI	PELCA
VRTRÇVVEAILLRISI	YIQSLEK ^S	GLLED <mark>SE</mark> IMHI	LHD <mark>A</mark> VQ <mark>ADLRI</mark>	RLLRNPPSVSM	IP <mark>KGADLL</mark> RTOI	PEIUS
VKRCVTHAILLHLT	YVESLENA	LFESRE4VHI	LHDAVQIDLRI	RLLRDPPIVEN	IPSAEQT LONGI	PEICA

780	790	800	810	820	830	840
I EAAMOREL KESKE		VRECSKATCUW	TODETVRW		HPTFSHCST	
LPSAICEPLKE SKRE						
LPEALRNULVESTRE	EVEVR <mark>G</mark> MSL	YREGERPIGIWI	LISNCVVRW	A-SKSSENERS	LHP <mark>AF</mark> THGTT	LGLYE
LP <mark>ETVRD</mark> PLLSN <mark>I</mark> RE	TVR <mark>AH</mark> GTIL	YREGSRPTGIWI	LVS <mark>LGIVRW</mark> I	F-SQ <mark>RLSSR</mark> HS	L <mark>DPILSHG</mark> ST	LGLYE
LP <mark>EGLRN</mark> RLV <mark>G</mark> STRE						
LP <mark>EALRN</mark> LLV <mark>G</mark> STRE						
LEATARDULVNSTRE						
LE <mark>AAT</mark> RDPLLSNTKE LPEAFCEPLKESKKE						
LPRIVRENLICSTRE						
LEAAMRDELLISTRE						
LE <mark>AA</mark> VRDPLLSNTRE	IVREQ <mark>G</mark> ITL	YREGSRPTGIWI	LVSI <mark>CVVRW</mark> I	r-sõrls <mark>rrhc</mark> i	L <mark>DPILSHG</mark> ST	LGLYE
LPSSVRESLECSTRE	RMK I RGMT L	Y <mark>REGSR</mark> PN <mark>GIW</mark> I	LISNCVVKSC	GRVRAYETSTQ	CIQLLLMGVR	WACM K
LPSOIQALIENSSRE					LHPTFS <mark>HG</mark> ST	LGLYE
MPISVQKPLLDVSKE	MRLQGSVI	YREDMK5DG1W	LITNCVVRW	SKKAVGG- <mark>Q</mark> RLI	RQPTESEGST	LGLYE

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.1668v1 | CC-BY 4.0 Open Access | rec: 21 Jan 2016, publ: 21 Jan 2016

NOT PEER-REVIEWED

Figure 2 (continued) Peer Preprints

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella_patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella_patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella_patens

850	860	870	880	890	900 910
VLT <mark>GRPYMCDMITD</mark> SV-					EEEVAMHELRALA
VLT <mark>GRPYMCE</mark> VITDSVV					
VLI <mark>GRPY CDNVTDSV</mark> A					
VLV <mark>GRPYICE</mark> MVTDSVV					
VLI <mark>GRPY CDMITDSVA</mark>					
VLIGRPY CDNVTDSVA					
VLIGRPYLCDMITDSVV VLIGRPYICDMITDSVV					
VLIGRPYLCDLITDSNV					
VLVGRPYICDIITDSVA					
VLIGRPYICDMITDSVV					
VLIGRPYICCM TDSV					
YWLENHTSATWSQIPWF	SVFFIESDRI	SMLRSDEAL	EDFLWQESAIV	VL <mark>ARLLE</mark> PQI	FERIALHDLRALV
VIM <mark>GKSFLCDLIAD</mark> SVV					
VLT <mark>GRPYLCD</mark> VV <mark>AD</mark> SVV	ECFFVE <mark>RS</mark> RI	.FI <mark>Q</mark> K <mark>BRAD</mark> L	EDFLWQESAL	V <mark>A</mark> RIVL <mark>G</mark> PQ	EDMATLODIRALV

920	930	940	950	960	970	980
SAESSELATYVSCESI	ETAYNSTCL		TOBEL-AASE	AVI I P-SNRN	OSERNSSE	S-TM
SAESSKQTTYVSGESI						
AERST-LNTYLRGETI	EVPPHSIGE	LL <mark>EGFIRSH</mark> S	SLVEEL-ITSE	APLNP-AQGN	ASFLNQNG	GY-K
AERST-INVYIRGEDI						
AERST-LNTYLRGETI						
AERST-LNTYLRGETI						
IERST-MNTYLSGENV					~	
AERST-VNMYIRGEDI SHESSKLHTYVHGESI						
AERST-MSIYIRGES						
TERST-MNIYIRGE 21						
AERST-MNIYIRGED					CADLGLE	
AERSS-METYIIGESI	EVCHQSVGE1	ML <mark>EGFIR</mark> PSE	IAEGEL-IKSE	AVLLP-SQGN	QSFEHADR	S <mark>GS-T</mark>
QN-SS-MISFIRGEII	EVPVGE1 <mark>G</mark> LI	LI <mark>EGF</mark> LK <mark>O</mark> D-	GENEIICAE	AGIVCPS	HETGSF	
M <mark>E</mark> GSS-MRI <mark>FLRGE</mark> VE	E L <mark>RHKE</mark> MG LI	LL <mark>EGEVR</mark> OE-	-NSTELITAE	AGLTARILLP	SPTGSWS	

990	1000	1010	1020	1030	1040	1050
RMSESROAIR	YIVETE	ABUTTENTGAL	I-GVORTLORK	PSALSSPRGE	PTSDHKLPRSS	SRE
RVSFSRQAAR		AFVIIFN <mark>HGA</mark> I				
SASES <mark>HQ</mark> GES		AFVLL <mark>ID</mark> MSQS				
HVDYCYAAPV		ARILIFEI VR'I				
STSESHQGAS		AFVLLI <mark>D</mark> MSQS				
SASESHQGES		AFVLL <mark>ID</mark> MSQS AFVIVEDINSE				
HVDYCYNAPG		ARIIEFEIGRU			~	
RVSFS00ATQ		ARAIIFNIGAE				
AASESHQPS		AFVIME <mark>DIAG</mark> E				
RIDYCYTAPS		ARILEVEI GRI				
DYCHTAPR		ARIIFLE				
TA <mark>SESHQ</mark> RSGYLLETRO	SVITOVETR	AFVIIPDISH	<mark>BGNRVLR</mark> NN	S <mark>S</mark>	SENLSBRSD	- RE
<mark>NISLQQPT</mark>	V <mark>FHAE</mark> AR	SFVLLF <mark>E</mark> SSK	7 <mark>GE</mark> IL <mark>SQG</mark> SSQ	LL <mark>A</mark> VPRKTSF	RSMLRLSS <mark>R</mark> TS	STNT

Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella_patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum lycopersicum

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella patens

1270 | | | . . . Eutrema parvulum Eutrema halophila _____ Salicornia_dolichostachya RXSXSGDGLLLCILKEKI Aeluropus littoralis _____ Suaeda salsa _____ Salicornia_brachiata Mesembryanthemum crystallinum _____ Distichlis spicata _____ Arabidopsis thaliana _____ Solanum lycopersicum Oryza sativa Indica _____ Sorghum bicolor _____ Gossypium hirsutum _____ Selagi Physco

>					U.I.PEH	K-PIEVIEVVEL	J
	HSGLM, WPENIIYQA	EEQ	3D I SKKALN	SEFAIQL	SIFGSNVD	LY <mark>RR</mark> SESFG	
	HRGLM, WPENIIYKA	EQE	EE MNGKT LN I	LS <mark>EFA</mark> MQL	SIFGSTEN	LY <mark>KR</mark> SVSF <mark>G</mark>	
	HSGLLSWPDTQYKSH-	QHLPVG	EE I EDDEN-I	LS <mark>ARA</mark> MQL	SIYGSNV	KDAPL <mark>RGQSF</mark> K	
	HSNLLSWPQSFRR RDS	IPRSPRRPNASL	SEIRNHPSSI	SS <mark>ka</mark> lql	SMYGGMIDDTI	TGQRRQ-RSR	
	HSGLLSWPETQYRSH-	QQLPTG	Q <mark>EI</mark> EESEN-I	LS <mark>ARA</mark> MQL	SIYGSMVPSR	JESFKEAPM<mark>R</mark>GQSFK	
	HSGLLSWPDTQYKSH-	QHLPVG	EEIEDDEN-I	LS <mark>AR</mark> AMQL	SIYGSMV	KDAPI <mark>RGQSF</mark> K	
	HCGLVSWPENARPER-	EGEDEE	EIDEHEQN-N	IS <mark>AFAR</mark> QL	SIYGSMLKE	KDPLLKTVSFH	
	HSGLLSWPESFRK	-PRGPHNVSL	AEIRSQPGSI	LSI <mark>FA</mark> LQL	SMYGSMMDDM	IPGQGQ <mark>RR</mark> Q-RHR	
	HRGLMSWPENIYAK	QQQ	E-INKTTLSI	LS <mark>EFA</mark> MQL	SIFGSNVN	VY <mark>RR</mark> SVSF <mark>G</mark>	
	LGGLMSWPENTYKAM-	QHRQDV	ERTGQQETN	IST <mark>FA</mark> MQL	NIFGSMIS	DTRS <mark>R</mark> SRSFP	
	HSGLLSWPESFRK	-SRGAQNG-ASL	TEIRDHPASI	SAFALQL	SMYGSMINDM	KSGQGQGQ <mark>RR</mark> Q-RHR	
	HSGLLSWPESFR	-SRGNLG	LAAEMLPGG	LSS <mark>FA</mark> LQL	SMYGSNV-ILS	SSGQGHSH <mark>RR</mark> QGRHR	
	HCGLMSWPEHFFRGR-	QHTQNB	DATDQQVNR	LS <mark>AFA</mark> MQL	SIFGSEVN	LPQ- <mark>R</mark> SWSLS	
	RTEVORHSSHNEIRIA	SDHSIPIAQAWG	SGSSCVPDK	T <mark>ARA</mark> L <mark>E</mark> I	SVFGSEVD	V <mark>R</mark> DKRRAD	

1130	1140	1150	1160	1170	1180	1190
	-					
GINNNKAQGN-SSY						
GLNNNKAQDN-LSY						
GDNLGNPSRSY						
LQATSQT <mark>H</mark> SS <mark>SY</mark>						
GDNLGNPSHV-RSY						
GDNLGNPSRSY						
ESVLNKPSS <mark>H</mark> SLSY						
RIQVT-NPR <mark>H</mark> SSSY						
GIYNNKLQDN-LLY						
GISAAKTSHS-QSY	PEVRSDGV	QILVSVRS	GSTTLRKK	VQGENKDN	ISIQLPSA	PIEES

1200		1220			
	DECIIVRIDS				
TMAEE-SSDDE	- <mark>DE</mark> SIIVRIDS-	PSTI <mark>VFR</mark> NDV		 	
	G <mark>DDDVIVRID</mark> S-				
	AR <mark>DE</mark> 4VVIVDS-				
	GDECVIVRIDS-				
	GDDDVIVRIDS-				
	G <mark>EEDVIVRID</mark> S-				
	AGEEVIVEVDS-				
	-DE <mark>GIVVRID</mark> S-				
	GEDERLIRID SC				
	AGEEVIVFVDS-				
	AGEEVIVEVDS-				
HVNDDSSDESG	G <mark>EDEILVRID</mark> S-	-PSWLSFHQTS		 	
		DC		 	
EFREDDSSDDS	DR <mark>EE</mark> IVRIDS-	PSELFHNSRR	<u>E</u> L	 	

inella_moellendorf:	ii	
omitrella patensis	ttps://doi.org/10.7287/peeri.preprints.1668v1 CC-BY 4.0 Open Access rec: 21 Jan 2016, publ: 21 Jan 2016	5

Figure 3A Peer Preprints

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella_patens

Eutrema_parvulum Eutrema_halophila Salicornia_dolichostachya Aeluropus_littoralis Suaeda_salsa Salicornia_brachiata Mesembryanthemum_crystallinum Distichlis_spicata Arabidopsis_thaliana Solanum_lycopersicum Oryza_sativa_Indica Sorghum_bicolor Gossypium_hirsutum Selaginella_moellendorffii Physcomitrella_patens

150 160 Α . . . <u>.</u> | | | | | . . . AGE<mark>GVLISTFCLATI</mark>IKLTFPYCWDWK¹ AGEGVLISTFCLASIVKLTFPYDWDWK AGE GVLISTECIGNALKLSEPYDWNWM AGEGVIMSTELLGJAVKLEFPYKWNWK AGEGVLLSTCILGEALKLTEPYEWNWK AGEEVLISTECIGVALKLSEPYDWNWM AGECVLISTECLGSALKLSEPYDWNWK AGE GVVISTE<mark>H</mark>LG<mark>TH</mark>IKVTEPYNWSWK AVEG VLISTOCLGSIVKVTFPYSWDWK AGEG VLISTFELGPALKIGFPYNWSWS AGECVLISTEBLGSALKLTEPYNWNWK AGEGUVVSTVLLGEAVKLTEPYNWSWK AGEGVIISTECLGEALKITEPYEWNWK AGE GVVISTECLGLESHFILPYCWSWSV A BEGVLISTFILGVATVKVFPYNWNWS

360	370	B	380	390
		· <u>·</u> · · · ·		.
KIAYQTSWGFLELLYE	YIHL	SREIVV	GALFPILR	CVGYC
KIAYQG-SSWGYLFLLYI				
NIFDNHGIAWGYLILLYG	YV V	SRUVEV	GVLFPEL <mark>X</mark>	FGYC
DHFERHGASWGYLLLLYV				
NIFENHGIAWGYLELLYS	1 1000		STATUTE AND ADDRESS OF	CONTRACTOR DESIGNATION
NIFDNHGIAWGYLILLY				
HIFENNGISWGYLVLLYV		and the second se		Contraction of the local division of the loc
VEFERHGISWGFLLLLYV	EVLI	SRIVVV	SVLYPILR	FGYG
KIAYQG-NSWRELELLYV				Deline Coll Problem 101
NIFKIYDNSWGYLILLYV				
VHFERHGASWGFLLLLYV	FVQI	SRILVV	VILYPILR	IFGYG
AHFERHGSSWGFLLLLYV	FVQI	SRIIVV	GVLYPILR	FGYC
KIFQNNGNSWGYLILLYI			and the second se	Contraction of the local distance of the loc
SSQRSTGRDWAYLLCLYE	FLQL	SRILVV	IALYPGLS	YGYG
SQDKIQGRDWGYLLLLYI	FVQI	SRMVV	GLLYP <mark>C</mark> LKY	FGYG

780	790	800 C	D
	.		· · · <u>· ·</u> · · · · · · ·
LFAAMCEPLKHSKKE	MKLRGVTLYF	KEGSK <mark>a</mark> lgvwl	ICIGIVKWQ-S
LPSAICEPLKHSKKE	MKLRGVTLYF	EGSKPIGVWL	IBIGIVKWK-S
LPDALFNLIVCSTKE	VKVRGMSLYF	EGCKPIGIWL	ISNGVVKWA-S
LP DI VRDPLLSNIKE	VRAHGATLYF	EGSRP.GIWL	VSIGIVKWI-S
LP <mark>BC</mark> LRNRLVCSTKE	VKVRGMTLYF	KEG <mark>GKPIGIWL</mark>	ISNGVVKWA-S
LP ^D ALFNLIVCSTKE	VKVRGMSLYF	KEG <mark>C</mark> KPIGIWL	ISNGVVKWA-S
LFATARDVLVNSTKE	VKVRGLTLYF	EGSRP GIWL	ISNGVVKW <mark>g-</mark> s
LFAA _B RDPLLSNTKE	URCEGTVLYF	KEGSRP <mark>1</mark> 'G <mark>VWL</mark>	VSIGVVKWI-S
LPEAFCEPLKESKKE	MKLRGVTLYF	KEGSKP <mark>I</mark> GVWL	IBIGIVKWK-S
LPPNVREALICSTKE	MKLRGATLYE	EGSKATEVWL	ISNGVVKWS-S
LFAAMRDPLLNSTKE	VKCEGTILYF	REGSRPIGIWL	VSIGVVKWI-S
LFAAVRDPLLSNTKE	VREQG ITLYF	REGSRPIGIWL	VSIGVVKWE-S
LPSSVRESLEC STKE	MKERGMTLYF	EGSRPIGIWL	ISIGVVKSGRV
LPSOIQALIFNSSKE	ELKLRG <mark>SV</mark> LYF	EGSR <mark>M</mark> GLWL	IANGIVKWN-T
MPHSVQRPLLDVSKE	MKLQG <mark>SV</mark> IYF	REDMKSIGIWL	I TI IG VVKWSKK

NOT PEER-REVIEWED

Liguro 2D				
Figure 3B				
Peer Preprints		ľ	NOT PEER-F	REVIEWED
	1200	1210 * *	1220	1230
		· · · · · · ·		.
Eutrema_parvulum	RIVESSDDE-DECII	NRIDSPS	IVFRNDP	
Eutrema_halophila	TMAEE-SSDDE-DEGII	NRIDS PS	IVFRNDV	
Salicornia_dolichostachya	-AVDDSSSESGCDDDVI	NRIDS PS	LTFRQAPXWDF	LPTPIVLLY
Aeluropus_littoralis	-HGIPCSEDEAARDEAV	<mark>∖</mark> IVDSPS	LSVPHHSSAMF	HSPPQDK
Suaeda_salsa	-ARDDSSSESGGDEDVI	NRIDSPS	LSFRQAP	
Salicornia_brachiata	-AVDDSSSESGCDDDVI	NRIDS PS	LTFRQAP	
Mesembryanthemum_crystallinum	-LPE-SSDESGCEEDVI	NRIDS PS	LSFRHAP	
Distichlis_spicata	EDDNSSEE SACEEVI	VRVDS PS	LSFRQSSGAAV	RSPPPREN-
Arabidopsis_thaliana	TMVESSDBEDB-DECIV	NRIDSPS	IVFRNDL	
Solanum_lycopersicum	DTREYSSDDSGGEDEHL			
Oryza_sativa_Indica	BBDEDNSSDESAGEEVI	NRVDSPS	LTFROPSSAAD	R
Sorghum_bicolor	HLQEDNSSDDSACEEVI	NRVDSPS	LSFRQSAAAPP	PQDQ
Gossypium_hirsutum	HVNDDSSDESGGEDEII	NRIDSPS	LSFHQTS	
Selaginella_moellendorffii				
Physcomitrella_patens	EPKPDDSSDDSDKEEHI	NRIDSPS	LFHNSRREL	

Figure 3. Crucial amino acid positions.

3A. A-D. *sos1* missense mutation positions.3B. SOS2 phosphorylation sites.