Title

Arabidopsis thaliana SPF1 and SPF2 are nuclear-located ULP2-like SUMO proteases that act downstream of SIZ1 in plant development

Authors

Pedro Humberto Castro Miguel Ângelo Santos Sara Freitas Pepe Cana-Quijada Tiago Lourenço Mafalda A. A. Rodrigues Fátima Fonseca Javier Ruiz-Albert Jorge E. Azevedo Rui Manuel Tavares Araceli G. Castillo Eduardo R. Bejarano Herlânder Azevedo

Supplementary data

Fig. S1. Protein sequence alignment of the catalytic domain of SPF1/2 subgroup members.

Fig. S2. Purification elution of recombinant proteins with the SPF2 and SPP1 catalytic domains with an N-terminus GST-tag.

Fig. S3. Schematic representation of Arabidopsis T-DNA insertion mutants for *SPF2* and *SPF1* and semi-quantitative RT-PCR.

Fig. S4. Plant and silique size of the wild-type Col-0 and *spf1/2* mutant.

Fig. S5. Morphology of 1-month-old plants of the SPF2 and SPF1 second allele T-DNA mutant.

Fig. S6. In silico analysis of SPF2 and SPF1 expression patterns.

Table S1. List of primers used for genotyping Arabidopsis T-DNA insertion lines.

Table S2. List of primers used in semi-quantitative and quantitative RT-PCR.

Table S3. List of primers used for plasmid constructs.

Table S4. *Cis*-elements over-represented in the promoter region of differentially expressed genes in *spf1/2*.

	410			. 430		450
AT1G09730, SPF1/ASP1/ULP2b	VWNTDLEDDV	EVSGYNLNQQ	KRYFPSFD	DEP	FEDVVYPKGD	PDAVSICKRD
ATR 00033G02480	CNSEKINDDD	SMGTOAEMSN	NPYLTRCN	JE D	FEDLVYPKGD	PDAVSTCKRD
0905011770	The Party of the P	NUNTERSTER	DEC VERDE			TRANSTORES
0305011770	FASE	NWNIEPSLFF	PRQYFSDI	LEE	FEDVIYPKGD	HDAVSISKRD
AT4G33620, SPF2/ULP2a	FDTITESEEI	AFSGHDLGTS	LTNLA	ADS	FEDLVYPQGE	PDAVVVRKQD
PP00024G00240	ERLTRYP	LEELSINANL	ASWLSPEI	LRT	FETLVYPQGD	PDAVTITSKD
SM00000G08840	AIEELKAE	KORVKEOPRI	LLOEALSM		LHSLVYPODD	PDPVTITSND
Consistency	0012223433	3244323233	23342213	253	97699**779	7 . 7 . 6 . 3 7 6 .
consistency	0012223433	5244525255	23342223	, <mark>, ,</mark> ,	0/00940//0	/ / / 0 0 0 / 0
		(12154)				1.202
	460	470		. 480	490	500
AT1G09730, SPF1/ASP1/ULP2b	VELLOPETFV	NDTIIDFYIN	YLKNOIOT	r e e	KHRFHFFNSF	FFRKLADLDK
ATR 00033G02480	VELLOPEVEV	NDTIDEXIK	TOCKLER	C H D	KNREHEENSE	FFRELADLDK
0905011770		NEWTTERNUT	THE OWNER THE			
	VELLIPETTV	NUTITURIVK	HISTRIEF	PAE	KHRIHFFNSF	FFREEADLOK
AT4G33620, SPF2/ULP2a	IELLKPRRFI	NDTIIDFYIK	YLKNRISE	PKE	RGRFHFFNCF	FFRELANLDE
PP00024G00240	V D V L R P MG F L	NDTIIDFYIK	YLQNQLSI	DKD	K G R F Y F F N S F	FFRKLADSSI
SM00000G08840	IDLLRPSAFL	NDTIIDFYIK	HLOMTMSE	DDE	KAKTYFFNSF	FFPKLVDLSA
Consistency	988 + 5 + 43 + 7	*******	7 6 4 5 7 6 4	4 4 8	83867 * * * 7 *	7 88775
consistency						
					+	
	510	520		. 530	540	550
AT1G09730, SPF1/ASP1/ULP2b	D P S S I A D G K A	AFLRVRKWTR	KVDMFGKD	IYC	FVPVNYNLHW	SLIVICHP GE
ATR 00033G02480	NPGCAFESKA	AFORVERWTE	KVNMFK K	IYO	FIPVNFNYHW	SLMVICHPGE
OS05G11770	DOGRAPECRA	AFTPUPKWTP	KINTETKE	PPT.	FTPUNENT	STATCAPCE
	DUGKAFLOKA	AFLAVKKAIK	KINI FI KI			SLIVICIFGE
AT4G33620, SPF2/ULP2a	GTPSTCGGRE	AYQRVQKWTK	NVDLFEKI	IYC	FIPINCSFHW	SLVIICHPGE
PP00024G00240	KYTGP <mark>NRGKI</mark>	VYD <mark>RVR</mark> KWTR	KVSLFEKI	VYC	FIPVNQSLHW	SLIIVCHLGN
SM00000G08840	LPADEARA	AFARLEKWTK	KEDIFOKD	IYC	FIPVSRSLHW	SLLVICYL SD
Consistency	3322415676	883 * 87 * * * 8	8767 4 4 8	888	* 9 * 9 8 3 7 6 * *	* * 7 9 9 * 7 5 8 7
conclosed						
	= (0		*	500	500	(00
		5/0		. 580		600
AT1G09730, SPF1/ASP1/ULP2b	VANRTDLD	<mark>L</mark> D D S K K	VPCILHME	DSI	KGS-HAGLKN	LVQTYLCE
ATR 00033G02480	VVKSED	DNDHK	VPCVLHL	DSI	KGS-HKDLKD	LVQGYLWE
0805G11770	VETEKDGD		IPCILHME	DSL	KGS-HSGLKD	IIOSYLWE
AT4C22620 SPE2/ULP20	TUDCUUNDUC	EDDEVENDOD	WD C T T UT		V C C U V C C I IN	TE DOVIDE
A14G33020, SFF2/0LF2a	LVPSHVNFHS	FUDEVENPOR	VPCILHLL	191	KGSHKGGLIN	IFPSILKE
PP00024G00240	LGSSPDE	SRG	TPCILHL	DSM	EGN-HEGIEE	HIRNYILQ
SM00000G08840	MLPTDSDL	HTV	STRILHF	DSL	TGF-HSDIEP	FVRKLEWNRR
Consistency	7344351100	0000021434	5769**6*	• • 7	6 6 0 7 4 6 8 5 4	4700448627
· · · · · · · · · · · · · · · · · · ·						
	(10	(20		(20)	C10	(50
				. 650		
AT1G09730, SPF1/ASP1/ULP2b	EWKERHKET -	SDDISSRFM-	NLRFVS	SLE	LPQQENSFDC	GLFLLHYLEL
ATR 00033G02480	EWKERHAEE -	K <mark>E</mark> S C E N A L V <mark>N</mark>	FSNMRFI	PLE	LPQQENSFDC	GLFLLHYVEL
OS05G11770	EWKERHPES-	ASDCSDKFL-	NLRFIS	LE	LPQODNSFDC	GLFLLHYVEL
AT4G33620 SPE2/ULP2a	EWKARHENT	NDSSRAPNM-	0 S I 5	LE	LPOORNSEDC	GLELLHYLDL
DD00024C00240	ANTERNORS	A D P A T C U M			UD O O T NUC D C	CTVI I BYURI
FF00024G00240	ANTEENQ35-	ADEALSHM		AI	VEQUINACUC	GIILLHIVEL
SM00000G08840	KTSEKEDRK-	YHFDQIKFV-		RVE	VPQQDNLHDC	GLFLLHYVEL
Consistency	6757773550	3 <mark>5 4 3 4 4 3 4 4</mark> 0	00423214	477	8 * * * <mark>5 * 5 5</mark> * *	98 88
	660	670		680	690	700
	PI NENDINES	DEVIVNACNE	TYTNMEDI			TERTIENDO
AT 1609730, SPF 1/ASP 1/ULP20	TLALAPLNES	PERIINASNE	LILINWEFE	PAL	N 2 T V WILLIGK	LIFELLENKS
ATR_00033G02480	FLRDVPTNFS	PFKITKYSNF	LTKDWFPF	PVE	AFSKRFHLQK	LIYDLHDDHI
OS05G11770	FLMDTPRSFN	PLKIDSFANY	LSDDWFPF	PAE	A S L K R S L I R K	LIHKLLKEPS
AT4G33620, SPF2/ULP2a	FVAQAPAKFN	PSLISRSANF	LTRNWFP	AKE	ASLKERNILE	LLYNLHKGHD
PP00024G00240	FLKGAPOHES	LASFEGEPYF	LTRTWFKS	S S E	ASAKESATEN	LLIRLYODYH
SM00000C09940	PIPPOPTOF	CT CT TTWNHT	DP		A CANEROLLE	THOMITARI
51000000000040	LACTISKS	S L S L I V N W E			SAKATOLLK	HOMETAKL
Consistency	844564577	5457344548	64236643	35	/ 5 - 3 3 8 4 6	843944433
				Consi	stency legend:	
				Uncon	served 0123456	Conserved

Fig. S1. Protein sequence alignment of the catalytic domain of SPF1/2 subgroup members. Protein sequences were retrieved from *Arabidopsis thaliana*, *Physcomitrella patens*, *Selaginella moellendorffii*, *Amborella trichopoda*, and *Oryza sativa* ssp. Japonica genomes, presenting the three conserved catalytic residues (arrows). Consistency between sequences (colour gradient) indicates the level of conservation of each residue. Protein sequence alignment was performed using PRALINE (http://www.ibi.vu.nl/programs/pralinewww/).



Fig. S2. Purification elution of recombinant proteins with the SPF2 and SPF1 catalytic domains with an N-terminus GST-tag. A protein elution aliquot ran in 10% SDS-PAGE and was stained with *Xpert Safe Protein Stain* (Grisp). Arrow indicates approximately the expected molecular weight for GST-SPFs recombinant proteins.



Fig. S3. Schematic representation of Arabidopsis T-DNA insertion mutants for *SPF2* and *SPF1* and semiquantitative RT-PCR. (**A**) Schematic representation of *SPF2* and *SPF1* displaying exons (grey boxes), introns (thin lines), and UTRs (black boxes). The site and orientation of T-DNA insertions are represented by triangles with the respective SALK line code. Scale bar indicates 1 Kbp. (**B**) Semi-quantitative RT-PCR for wild-type (Col), *spf2-2 (spf2)*, *spf1-1 (spf1)* and *spf2-2 spf1-1 (spf1/2)*. *ACT2* was used as a loading control and the total extracted RNA, used as template for reverse transcription, served as a quality control.



Fig. S4. Plant and silique size of the wild-type Col-0 and *spf1/2* mutant. Plant size at the end of life cycle (**A**), silique size (**B**) and number of seeds normalized by silique size (**C**). Error bars represent standard error of the means (SEM), n > 10 (A), and n = 9 (B, C). Asterisks indicate statistically significant differences with respect to the wild-type (unpaired t test; ***, P < 0.001; ****, P < 0.0001).



Fig. S5. Morphology of 1-month-old plants of the SPF2 and SPF1 second allele T-DNA mutants.



Fig. S6. *In silico* analysis of *SPF2* and *SPF1* expression patterns. Expression profile of Arabidopsis *SPF2* (red) and *SPF1* (blue) during development (**A**) and in different plant organs (**B**). Analysis was conducted in Genevestigator (genevestigator.com).

Primer name	Primer sequence (5' to 3')	Description
SPF2-2 LP SPF2-2 RP	ACCCACAAAGGGTTCCCTGCAT TCTCTTGCTGCGGGGAGCTGA	Genotyping of <i>spf2-2</i> (SALK_090744)
SPF2-3 LP SPF2-3 RP	CCAACTTTACAAGCTGCTTCG TGCTGACAGGAATTGATCTCC	Genotyping of <i>spf2-3</i> (SALK_140824)
SPF2-1 LP SPF2-1 RP	AATCAGTTTTGGTCGGTAGGC AAAGCATGCATCCCTCTTTTC	Genotyping of <i>spf2-1</i> (SALK_023493)
SPF1-1 RP SPF1-1 LP	TCCGCCTAGCTGGCAAGAGA CCGGTTCTGCAACGCCAACG	Genotyping of <i>spf1-1</i> (SALK_040576)
SPF1-2 RP SPF1-2 LP	GTCAACGCCAGCTAAACTCAC CCAAAATCGAATGAAAGCATG	Genotyping of <i>spf1-2</i> (SALK_022079)
SIZ1-2 RP SIZ1-2 LP	CACGACAGATGAAGCATTGTG GAGCTGAAGCATCTGGTTTTG	Genotyping of <i>siz1-2</i> (SALK_065397)
LBb1.3	ATTTTGCCGATTTCGGAAC	Left border primer for genotyping SALK T-DNA insertion lines

 Table S1. List of primers used for genotyping Arabidopsis T-DNA insertion lines.

Gene (AGI code)	Primer name	Primer sequence (5' to 3')	Tm	GC (%)	Product size (bp)
<i>ACT2</i> (At3g18780)	ActinF ActinR	CTAAGCTCTCAAGATCAAAGGCTTA ACTAAAACGCAAAACGAAAGCGGTT	52.7 57.2	40.0 40.0	211
<i>SPF2</i> (At4g33620)	SPF2 RT F SPF2 RT R	CTTTTGACTGTGGCCTCTT CTTTGTGGAGGTTGTAAAGC	49.6 49.1	47.4 45.0	183
SPF1	SPF1 RT F	GGAAGAAGAAATGGAAGGTC	47.8	45.0	152
(At1g09730)	SPF1 RT R	CTAAATGGTCAGTGGTTTCC	48.2	45.0	
<i>CAD7</i>	CAD7 RT F1	TTTCCTCTCATCTTTGGGCG	57.9	50.0	144
(At4g37980)	CAD7 RT R1	GGCGGTGTTGACATAATCCG	59.3	55.0	
<i>KNAT1</i>	KNAT1 RT F1	AGTGGCCATATCCTTCTGAGTC	59.3	50.0	181
(At4g08150)	KNAT1 RT R1	TCCATGTACAGAGCTGCGTG	60.1	55.0	
<i>NIA1</i>	NIA1 RT F1	CCACCAGGAGAAACCGAACA	59.9	55.0	167
(At1g77760)	NIA1 RT R1	TCATCCCCATGAGGTTCCAG	58.9	55.0	
<i>PER1</i> (At1g48130)	PER1 RT F1 PER1 RT R1	TTCGCCAATTCTTGGACCGT CCTTGCTTCCGTGATTAAAGGC	60.2 60.2	50.0 50.0	199
<i>SOC1</i> (At2g45660)	SOC1 RT F1 SOC1 RT R1	GAGCAGCTCAAGCAAAAGGAGA GGGCTACTCTCTTCATCACCTC	61.1 59.6	50.0 54.6	137
<i>XTH6</i>	XTH6 RT Fw1	TGATCAGAGCACTGGATGTGG	59.8	52.4	153
(At5g65730)	XTH6 RT Rv1	TCTAGCTCGTCTCTCACCGT	59.8	55.0	
<i>XTH31</i>	XTH31 RT F1	TCCACTGGGAGTGGGTTCA	60.1	57.9	191
(At3g44990)	XTH31 RT R1	GAATAAGGCTTCCCTGGCGT	60.1	55.0	

 Table S2. List of primers used in semi-quantitative and quantitative RT-PCR.

Tm - Melting temperature; *bp* - Base pair

Primer name	Primer sequence (5' to 3')	Description
SPF2 pENTR Fw1 SPF2 pENTR Rv1	TTGCGGCCGCCATGACTCTCCGGTCAGTTC AATCC TTGGCGCGCCCAGTTTTTGGCTTGGC	Cloning into <i>pENTR</i> ; RS: <i>Not</i> I Cloning into <i>pENTR</i> ; RS: <i>Asc</i> I
SPF1 pENTR Fw1 SPF1 pENTR Rv1	AAGCGGCCGCCATGAAGAAAAACTTTGAA GTATTCG AAGGCGCGCCCTTTCTCCATCTCCTCAGCTT CGCC	Cloning into <i>pENTR</i> ; RS: <i>Not</i> I Cloning into <i>pENTR</i> ; RS: <i>Asc</i> I
SPF2 pCM190 Fw1 SPF2 pCM190 Rv1	AGCTTTgtttaaacATGACTCTCCGGTCAGTTC ATAGTTTAGCGGCCGCTCAAGTTTTTGGCTT GGCC	Cloning into <i>pCM190</i> ; RS: <i>Pme</i> I Cloning into <i>pCM190</i> ; RS; <i>Not</i> I
cSPF2 pGEX5 Fw1 cSPF2 pGEX5 Rv1	TTGGATCCAAGATCTAGTTTACCCTCAAGG AG AATAGCGGCCGCTTAACCTTTGTGGAGGTT GTAA	Cloning into <i>pGEX-5X-1</i> ; RS: <i>Bam</i> HI Cloning into <i>pGEX-5X-1</i> ; RS: <i>Not</i> I
cSPF1 pGEX5 Fw1 cSPF1 pGEX5 Rv1	AAGAATTCGAGGATGTTGTCTATCCAAAGG GTG AACTGCGGCCGCTTAGTTTTCAAGGAGTTC AAAT	Cloning into <i>pGEX-5X-1</i> ; RS: <i>Eco</i> RI Cloning into <i>pGEX-5X-1</i> ; RS: <i>Not</i> I

 Table S3.
 List of primers used for plasmid constructs.

RS – Restriction site

Table S4. *Cis*-elements over-represented in the promoter region of differentially expressed genes in *spf1/2*. The DEGs were submitted to Athena analysis (O'Connor *et al.*, 2005; DOI: 10.1093/bioinformatics/bti714), to scan for binding site enrichment.

<i>Cis</i> -element name	<i>Cis</i> -element sequence*	Nr. of genes	Predicted in the genome	Found in the genes	<i>p</i> -value	Corresponding TFs
Down-regulated						
AtMYC2 BS in	CACATG	61	35%	53%	< 10e-6	MYC2
RD22 MYCATERD1	CATGTG	61	35%	53%	< 10e-6	MYC2
Up-regulated						
AtMYC2 BS in RD22	CACATG	47	35%	47%	< 10e-3	MYC2
MYCATERDI CARGCW8GAT TATA-box Motif	CATGTG CWWWWWWWWG TATAAA	47 70 91	35% 59% 91%	47% 70% 82%	< 10e-3 < 10e-3 < 10e-4	MYC2 AGL15

* R (A/G), M (A/C), W (A/T), K (G/T), B (C/G/T), N (A/C/G/T)