

Structural analysis and evolution of specificity of the SUMO UFD E1-E2 interactions

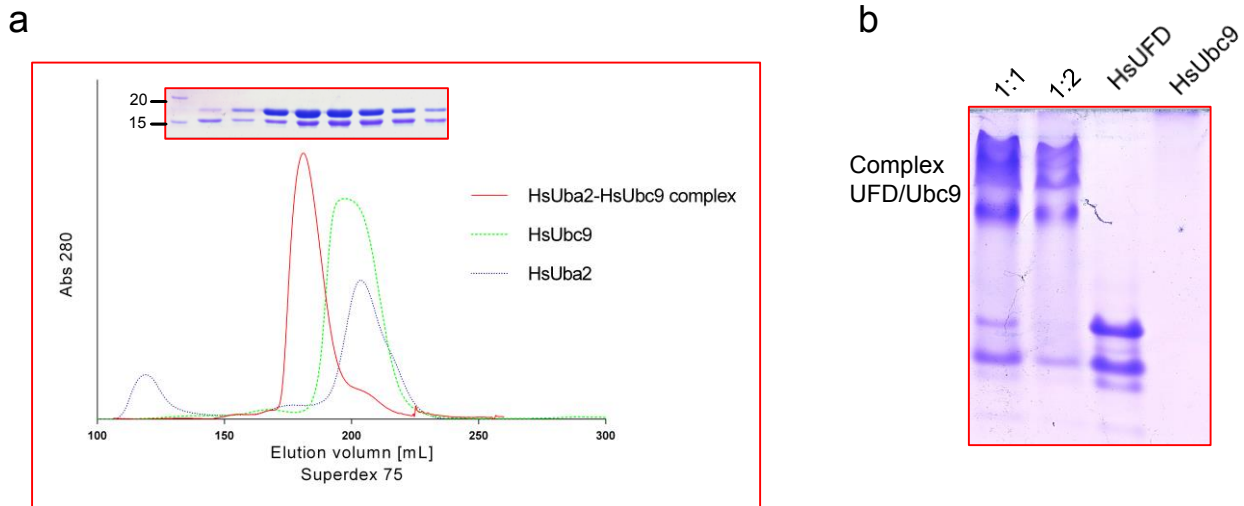
Bing Liu¹, L. Maria Lois^{2*} and David Reverter^{1*}

¹Institut de Biotecnologia i de Biomedicina and Departamento de Bioquímica i Biologia Molecular, Universitat Autònoma de Barcelona, 08193 Barcelona, Bellaterra, Spain.

²Center for Research in Agricultural Genomics-CRAG. Edifici CRAG-Campus UAB, Bellaterra 08193 Barcelona, Spain

*To whom the correspondence should be addressed: David Reverter, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain, Ph: +34 93 5868955, FAX: (93)5812011, E-mail: david.reverter@uab.cat and L. Maria Lois, Center for Research in Agricultural Genomics-CRAG, 08193 Bellaterra, Barcelona, Spain, Ph: +34 93 5636600 ext.3215, Fax. +34 93 5636601, email: maria.lois@cragenomica.es.

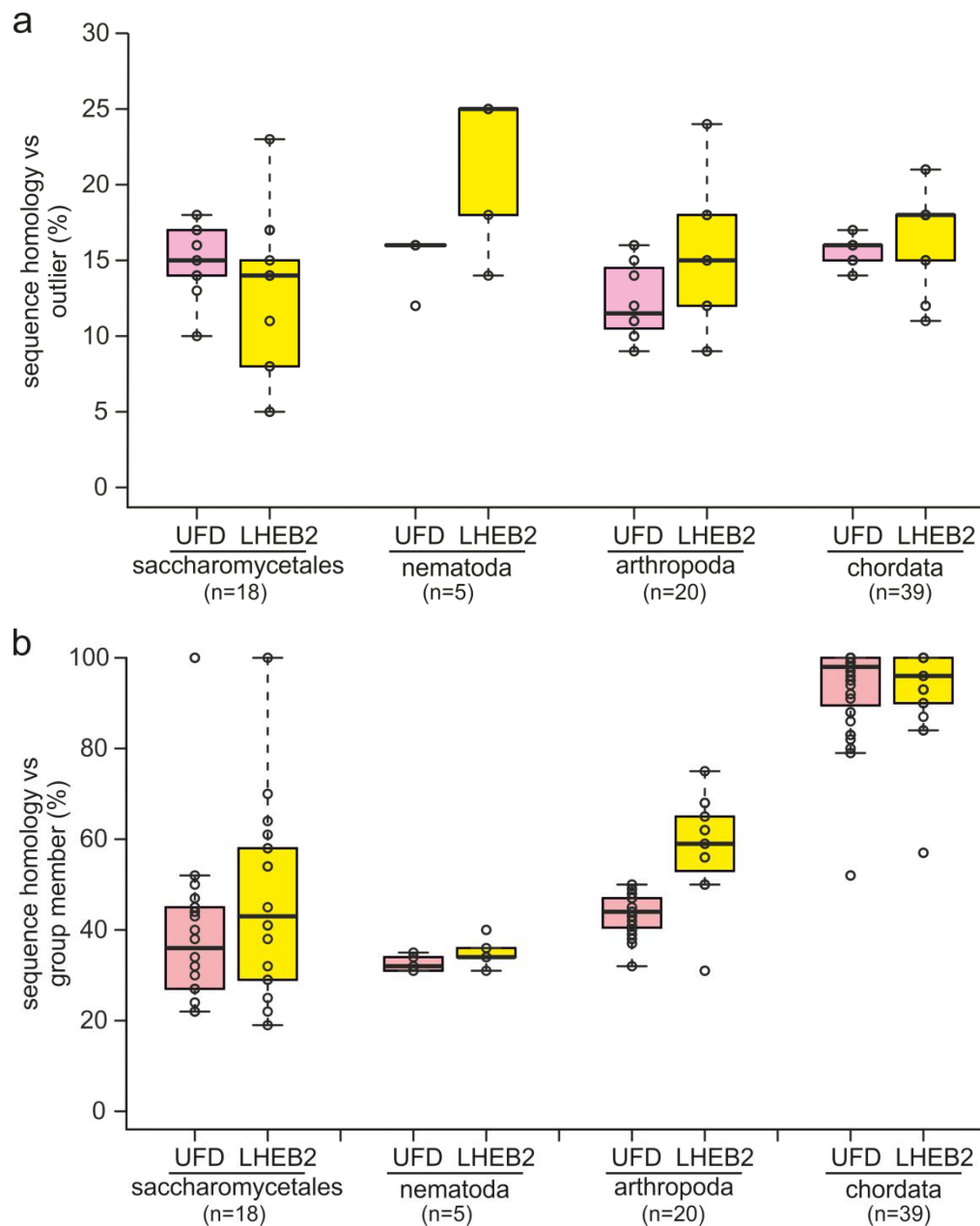
Supplementary Information



Supplementary Figure S1. Purification of the complex between human E1 UFD domain and Ubc9.

(a) Size exclusion chromatography profiles of human Ubc9, human E1 UFD domain and the complex with a buffer containing 100mM NaCl, 20mM Tris 8, 1mM BME. Inset, SDS-PAGE of the peak fractions.

(b) Native gel electrophoresis of the complex formation between human UFD domain and human Ubc.



Supplementary Figure S2. Analysis of pair homology distribution between UFD or LHEB2 sequences within each phylogenetic group analyzed. (a) Distribution of identity percentage of UFD (pink boxes) or LHEB2 (yellow boxes) domains between each sequence in the analyzed group and an outlier sequence (human, in the case of saccharomycetales, and yeast, in the case of nematoda, arthropoda and chordata). (b) Distribution identity percentage of UFD (pink boxes) or LHEB2 (yellow boxes) domains between each sequence in the analyzed group and a sequence from the same group (yeast, in the case of saccharomycetales, and human, in the case of nematoda, arthropoda and chordata). Center lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots; data points are plotted as open circles.

* 20 * 40 * 60 * 80 *
Api : CNVCS--KGE-IVFVDMHNFVROEELV--LKKKLNVAPDVTMLD---RVLTISSEDETGDVD-MYDMTISEA-GMTNCGSFDADD : 79
Cbri : CYVCS--TKE-VFLYVNLAEMLSGLCEKV--LKQELNLAADPAI--ESNTNSIVLSDGEMDP-LMSKLAEL-SMDGAILACDDF : 80
Cbr : CFCAD---VRS-VFLYVNPDEMVGGLCEKV--LKQELNLAADPAI--GGTFNMIISDPEDKMDEMLT-KISEV-SDNCAILACDDY : 82
Cre : CYVCS--TRE-VFIYVNPDEMVGALRDKV--LMQELNLAADPAI--TNTANVVLSDGEMDP-LMSKLAEL-SMDGAILACDDY : 80
Cel : CFVCS--KRE-VFIYVNPDEMVGGLCEKV--LKQELNLAADPAI--SATSRIIVSDGEMDP-LMSKLAEL-SMDGAILACDDY : 80
Cjap : CYVCS--TRE-VFIYVNPDEMVGGLCEKV--LKQELNLAADPAI--NASSRIIVSDGEMDP-LMSKLAEL-SMDGAILACDDY : 80
Sma : CLVCS--LHLLCAPELLLRILRDRILIRHLGLAPDVEV--PDRGIIISSEDEDEETLNLTADFLKLTHTCLQCDDF : 85
Cin : CYVCS--KPE-ITLKLNTKTLCEQFRDKI--LKSHLGLAPDVEIL--GRGTLISSEDEEENQSLGQTLDSFNTHCSRRAADF : 83
Dpu : CYVCS--QPE-VTVLLNVSQMVQSFEKV--LRGALHISPDVMD--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Hma : CFCAK---KPE-VTIFLNLQTVVKQLEDKI--LKDKLCLVAPDVEID--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Aqu : CYVCS--CPE-ASVKNLNTNSTIATLRDKI--IIGHFGIAPDVEID--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Nve : CYVCS--KPE-VTVVFNVTETMIOALEEKV--LKERFGLVAPDVEID--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 83
Lcha : CYVCS--KPE-VTVKLVNHSKVLQTLQDKI--LKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Aca : CYVCS--KPE-VTVKLVNHSKVLQTLQDKI--LKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Tga a : CYVCS--KPE-VTVKLVNHSKVLQTLQDKI--LKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Mlu b : CFCAS---KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Oan : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Pca : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Mmus : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Rno : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Itr : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Oga : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Laf a : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Hsa : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Mlu a : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Ocu : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Bta : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Amel : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Mpu : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Eca : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Fca : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Cpo : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Nle : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Cjac : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Ptr : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Mmul : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Clu : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Gga : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Mga : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Psi : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Mdo : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Sha : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Dre : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Gmo : CYVCS--RPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Oni : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Tni : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Tru : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Gac : CYVCS--QPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Ola : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Xma : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Hme : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Bmo : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Dpl : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Tca : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 82
Phu : CFCSE---KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Dps : CYVCS--DPA-IITLRIDTKRVHKEIRDDV--LIKTILNLPDVTV--VGSNLSISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Dwi : CYVCS--DPA-IITLRIDTKRVHKEIRDDV--LIKTILNLPDVTV--VGSNLSISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Dgr : CYVCS--DPA-IITLRIDTKRVHKEIRDDV--LIKTILNLPDVTV--VGSNLSISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Dvi : CYVCS--DPA-IITLRIDTKRVHKEIRDDV--LIKTILNLPDVTV--VGSNLSISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Nan : CYVCS--DPA-IITLRIDTKRVHKEIRDDV--LIKTILNLPDVTV--VGSNLSISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Dme a : CHVCS--DPA-IITLRIDTKRVHKEIRDDV--LIKTILNLPDVTV--VGSNLSISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Dya a : CYVCS--DPA-IITLRIDTKRVHKEIRDDV--LIKTILNLPDVTV--VGSNLSISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Aga : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Aae : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Cqu : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Nvi : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Ace : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Ame : CYVCS--KPE-VTVVRLNVHKKVVLTLQDKI--VKEKFAVAPDVOIE--GKGTILISSEGETE-NMPLTISDF-NVSDCSRRCEDF : 81
Sst : CPSSSL---LSR-GIMNLTNQELQENTLRWLVDLVEKKGYYEDGLSI-IVGKSRVYDVFDD---NIDSSISELGGFDLVLQDDE : 82
Lel : CPTEAG---TAR-GVFKVRDSDMGNLTLNWLLEQLKVSYGSRSISI-QAGKMLIYDVFDD---YVEAKIKVDPGFQNALILQDDE : 82
Mgu b : CPAESL---SAR-GILHVSQKDAP--TLGQLVEIKKKYGY--EIAL-ILGKAKLIYDVFDD---NDRSVLEVAGFVSGELLQVQDD : 78
Cdu : CASDSL---TAR-GVVTISIDDLVLTLDGFIKLDKADSYGIKLSISI-QIGKSLIYDVFDD---YLOTKIKVDPGFQNALILQDDE : 82
Ctr : CASDSL---TSR-AVLRISNDLRLITLGLVDFEYKGYNKEISISI-QLGKSLIYDVFDD---HLDPLNKVDPGFQNALILQDDE : 82
Kla : CSVSSN---IQR-AVLKIPKAIVEHMSLYDLNQIKISKYEFTGEYSVL--TNTNSLYDVFDD---LECRKISRNFNRTGTFLLYKDE : 81
Cgla : CAVCSH---VIR-GVFKVRDSDMGNLTLNWLLEQLKVSYGSRSISI-QAGKMLIYDVFDD---YVEAKIKVDPGFQNALILQDDE : 81
Ecy : CPVCSH---YQR-VTITLNVQESWQSMKGLDILIKIRDRYGFSEMSVM--MTSNR-LADDFD---LYSQTLOGLNINTITLAVSDS : 81
Ego : CSVCAH---YQR-CALTVGAVRCKTLTIGELIEAVQEKYGFSEVSVV--TSTNR-LADDFD---LNGNTIEKLELRTGTFLLYKDE : 81
Lth : CPVCSH---IAR-GVLTLSQNDLDMRSELVGAIVQEKYGFSEVSVV--KSTQR-LADDFD---LLNLTAEV-KVARTITLISDE : 80
Tph : CSVCSM---VVR-GILKVSTEKLTKLIDLITLAKTYGTLISVIV--TKDQR-LADDFD---LYE-STHDL-QSDSITLISDE : 81
Vpo : CAVCSK---VCR-GVLEVSTVKLQSTIGELIELIRTTFKYPNDVSVL--VSRQR-LADDFD---LLD-TIEDL-KLGEQDVLFTDD : 81
Nca : CAVCSK---VTR-GVLSINDMNIIGELIEGK--LREKYGFSEDISLL--SSEQR-LADDFD---LLDRSTKVD-NLRNCSVILLTDE : 78
Nda : CAVCSK---VTR-GVIVISDPTILKEFIAE--LRSKYGFSEDISLL--LSEQR-LADDFD---MVDRTSDV-NLRNCSVILLTDE : 78
Sce : CPVCSK---VCR-GVILKSSDCLNKMKLSDFVVLIREKYSYPODISLL--ASNQR-LADDFD---LNDRTLSEI-NLGNCSVILLTDE : 81
Tde : CPVCSK---VCR-GVILKTEKGLNEINISKLIDFIREKYSYPODISLL--TSNQR-LADDFD---LECTIADC-KLHAGSVVLFSD : 81
Zro : CAVCSK---VAR-GVVKITEKGLKLTLSQIDLLKKNYNSPSTSLI--TSDQR-LADDFD---LLDRTLSQA-KLHAGSVVLFSD : 81

	100	*	120	*	140	*	
Api	: H NFR	-----VKIIA	-----YDKE	-----KSGEEDLD	-----EILNN	-----	: 107
Cbri	: M DFG	-----VKVVF	-----CKNS	-----NFRGDQ	-----EIVRE	-----	: 106
Chr	: M ELE	-----KLFY	-----RSAD	-----HLKGDQ	-----EVARE	-----	: 108
Cre	: M DMV	-----KLFY	-----RRAN	-----HLRGDN	-----EIARS	-----	: 106
Cel	: Q EME	-----IKLFI	-----KKGD	-----RLAGDD	-----EVARS	-----	: 106
Cjap	: Q DMV	-----KLFY	-----RRQN	-----DLKGE	-----KILRE	-----	: 106
Sma	: R DFT	-----IRLILSCISVELANKLP	-----TSDLSTLNLNLSSRLEE	-----QLTSEELQSET	-----TWIRVGD	-----	: 142
Cin	: L N D	-----IVVNI	-----LHSE	-----N LPDDMI	-----QVSET	-----	: 110
Dpu	: H E D	-----NITI	-----AHRD	-----Q LEDNRE	-----SIAD	-----	: 109
Hma	: L NFE	-----IAITL	-----RDKK	-----D INVQD	-----FLEGE	-----	: 109
Aqu	: L N Q	-----ILNL	-----FHSD	-----EALSDEKE	-----EVIDS	-----	: 110
Nve	: L N E	-----VINI	-----KHRT	-----D LETDQ	-----EVEGD	-----	: 110
Lcha	: L D T	-----LINV	-----LNSE	-----D LEKDVE	-----EUVGD	-----	: 109
Aca	: L D T	-----LINV	-----LHCE	-----E LAKDVE	-----EUVGD	-----	: 109
Tgu a	: L D T	-----LINV	-----LHSE	-----D LEKDVE	-----EUVGD	-----	: 109
Mlu b	: L D T	-----LINI	-----LHSE	-----D LGKDIE	-----EUVGD	-----	: 109
Oan	: L D T	-----LINL	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Pca	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Mmus	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Rno	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Itr	: L D T	-----LINI	-----LHSE	-----D LEKDVE	-----EUVGD	-----	: 109
Oga	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Laf a	: L D T	-----LINI	-----LHSE	-----D LAKDVE	-----EUVGD	-----	: 109
Hsa	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Mlu a	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Ocu	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Bta	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Amel	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Mpu	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Eca	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Fca	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Cpo	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Nle	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Cjac	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Ptr	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Mmul	: L D T	-----LINI	-----LHSE	-----D LGKDVE	-----EUVGD	-----	: 109
Clu	: L D T	-----LINI	-----LHSE	-----D LGKDIE	-----EUVGD	-----	: 109
Gga	: L D T	-----LINV	-----LHSE	-----D LEKDVE	-----EUVGD	-----	: 109
Mga	: L D T	-----LINV	-----LHSE	-----D LEKDVE	-----EUVGD	-----	: 109
Psi	: L D T	-----LINV	-----LHSE	-----D LEKDVE	-----EUVGD	-----	: 109
Mdo	: L D T	-----LINV	-----LHSD	-----D LEKDVE	-----EUVGD	-----	: 109
Sha	: L D T	-----LINV	-----LHSD	-----D LEKDVE	-----EUVGD	-----	: 109
Dre	: L D T	-----LVNV	-----IHSE	-----E LEKDVE	-----EUVGD	-----	: 109
Gmo	: L D T	-----LINV	-----LHAE	-----D LEKDVE	-----EUVGD	-----	: 109
Oni	: L D T	-----LINV	-----LHVE	-----D LERDVE	-----EUVGE	-----	: 109
Tni	: L D T	-----LVNV	-----LHTE	-----E LERDVE	-----EUVGE	-----	: 109
Tru	: L D T	-----LVNV	-----LHTE	-----E LERDVE	-----EUVGD	-----	: 109
Gac	: L D A	-----LVNV	-----LHTE	-----E LERDVE	-----EUVGE	-----	: 109
Ola	: L D T	-----LINV	-----LHTE	-----E LERDVE	-----EUVGE	-----	: 109
Xma	: L D T	-----LINV	-----LHTE	-----E LERDVE	-----EUVGE	-----	: 109
Hme	: L N E	-----VRVRL	-----QQDD	-----DEKSWKLITD	-----	-----	: 104
Bmo	: L N E	-----VRVRL	-----QQED	-----EENSRLVTD	-----	-----	: 104
Dpl	: L N E	-----VRVRL	-----QQED	-----EKTWRLVTD	-----	-----	: 104
Tca	: L N E	-----IINV	-----NQYE	-----AKEKDPP	-----KFIAN	-----	: 109
Phu	: L N E	-----TVYV	-----NHYE	-----APKDEPD	-----KIAD	-----	: 108
Dps	: H N S	-----SVII	-----SHFD	-----A ERDSDL	-----EVSAD	-----	: 108
Dwi	: H N K	-----RIII	-----NHFD	-----A DREDNL	-----EUVAD	-----	: 108
Dgr	: F N E	-----SIII	-----AHFD	-----A EREDVL	-----EVIAD	-----	: 108
Dvi	: F S E	-----SIII	-----AHFD	-----A DREDVL	-----EVIAD	-----	: 108
Nan	: F N E	-----SIII	-----SHFD	-----S ERDEAV	-----EUVAD	-----	: 108
Dme a	: F N E	-----SIII	-----SHFD	-----A ERDENL	-----EUVAD	-----	: 108
Dya a	: F N E	-----SIII	-----SHFD	-----A ERDENL	-----EUVAD	-----	: 108
Aga	: L N E	-----SITV	-----LHKE	-----A GREEAP	-----EIVAD	-----	: 108
Aae	: V N E	-----TFTV	-----IHKD	-----P GRDESS	-----DIVAD	-----	: 108
Cqu	: V N E	-----TFTI	-----VHKD	-----P ARDEPS	-----DIVAD	-----	: 108
Nvi	: L N S	-----KVTV	-----VHRE	-----KPLPNSDEPE	-----VITAD	-----	: 111
Ace	: Q N S	-----TVYI	-----VYREKDPK	-----DQDSPO	-----LILAD	-----	: 112
Ame	: Q N S	-----TITI	-----IYRE	-----RPSLKGDSPD	-----LILAD	-----	: 111
Sst	: NDELE	-----N ELYI	-----	-----TVVN	-----EPTTEK	-----	: 103
Lel	: DDLE	-----K ELLI	-----	-----NVVDPGPEGDYNGIVLTKLPSV	-----	-----	: 115
Mgu b	: NDELE	-----N EFYI	-----	-----SLGD	-----KNE	-----	: 96
Mgu	: NDELE	-----N EFYI	-----	-----SLGD	-----KNE	-----	: 96
Cdu	: ADELE	-----N ELYL	-----	-----NIKD	-----DTVE	-----R	: 103
Ctr	: DDQLE	-----N QLYL	-----	-----SIVD	-----EQCDLELPLKLR	-----	: 110
Kla	: EMEDG	-----AQLRQSIEFYI	-----	-----EFVP	-----ENSGEDLQLPNLPD	-----	: 116
Cgla	: EEEAG	-----LTKQIMEYYI	-----	-----EISD	-----ESEDKIEQLSLPQI	-----	: 114
Ecy	: ISKED	-----D TVRKPIEFYL	-----	-----ELDN	-----LVNEID	-----	: 107
Ego	: VPSEEDN	-----TVRKPIEFYL	-----	-----ELSD	-----SATG	-----ELIE	: 111
Lth	: QDSEE	-----SIRAPVELYI	-----	-----EQGE	-----PEGINLPDIE	-----	: 109
Tph	: EDEND	-----MVRKSMELYI	-----	-----DVVDK	-----STATNLIELPKID	-----	: 114
Vpo	: EGDSDSNQFRKS	-----ELYI	-----	-----AVVS	-----DKSDDYIKLPSIE	-----	: 115
Nca	: EGNEETGMVRKP	-----ELYI	-----	-----NVTD	-----DNSIKLTLPPID	-----	: 111
Nda	: EENDA	-----GMCRRK	-----ELYL	-----ELRN	-----EISGQNI	-----PKIILPELD	: 114
Sce	: EGDTM	-----IRKAIELFL	-----	-----DVDD	-----ELPCNTCSLPDVE	-----	: 112
Tde	: EGVGD	-----GSYRKAIELYI	-----	-----ETVK	-----GPLENEINLPDVV	-----	: 114
Zro	: EGDES	-----EMVRKP	-----EYI	-----DVVE	-----ESAINSEIELPALE	-----	: 115

Supplementary Figure S3. Protein sequence alignment of the Sae2 UFD domain from metazoa and Saccharomycetales. Human Sae2 homolog sequences were retrieved from EggNOG database. After removal of incomplete sequences or sequences exceeding in length, 86 sequences were retained. Selected sequences were aligned using MUSCLE at EMBL-EBI. The sequence alignment fragment corresponding to the LHEB2 domain is enclosed in a red rectangle. Residue shading correspond to 90% (white letter and dark background), 70% (white letter and gray background), and 50% (black letter and light gray background) of sequence identity.

Phylum	Symbol	Sequence code	Species
Arthropoda	Aae	>7159.AAEL010641-PA	<i>Aedes aegypti</i> (yellow fever mosquito), species, mosquitos
Arthropoda	Ace	>12957.ACEP_00011683-PA	<i>Atta cephalotes</i> (Leafcutter ant)
Arthropoda	Ag	>7165.AGAP008637-PA	<i>Anopheles gambiae</i> (African malaria mosquito), species, mosquitos
Arthropoda	Ame	>7460.GB12173-PA	<i>Apis mellifera</i> (honey bee), species, bees
Arthropoda	Api	>7029.ACYP1006138-PA	<i>Acyrtosiphon pisum</i> (pea aphid), species, aphids
Arthropoda	Cqu	>7176.CPJJ007700-PA	<i>Culex quinquefasciatus</i> (southern house mosquito), species, mosquitos
Arthropoda	Dgr	>7222.FBpp0149193	<i>Drosophila grimshawi</i> (species, flies)
Arthropoda	Dme_a	>7227.FBpp0076457	<i>Drosophila melanogaster</i> (fruit fly), species, flies
Arthropoda	Dpl	>13037.EHJ64540	<i>Danaus plexippus</i> (Monarch butterfly)
Arthropoda	Dps	>7237.FBpp0276753	<i>Drosophila pseudoobscura</i> (species, flies)
Arthropoda	Dpu	>6669.DappuP300104	<i>Daphnia pulex</i> (common water flea), species, crustaceans
Arthropoda	Dvi	>7244.FBpp0226639	<i>Drosophila virilis</i> (species, flies)
Arthropoda	Dwi	>7260.FBpp0249132	<i>Drosophila willistoni</i> (species, flies)
Arthropoda	Dya_a	>7245.FBpp0265382	<i>Drosophila yakuba</i> (species, flies)
Arthropoda	Dan	>7217.FBpp0113673	<i>Drosophila ananassae</i> (species, flies)
Arthropoda	Tca	>7070.TCO02294-PA	<i>Tribolium castaneum</i> (red flour beetle), species, beetles
Arthropoda	Bmo	>7091.BGIBMGA012116-TA	<i>Bombyx mori</i> (domestic silkworm), species, moths
Arthropoda	Hme	>34740.HMEL017289-PA	<i>Heliconius melpomene</i>
Arthropoda	Nvi	>7425.NV10104-PA	<i>Nasonia vitripennis</i> (jewel wasp), species, wasps &c
Arthropoda	Phu	>121225.PHUM280660-PA	<i>Pediculus humanus</i>
Chordata	Aca	>28377.ENSACAP00000001485	<i>Anolis carolinensis</i> (Green anole) (American chameleon)
Chordata	Ame1	>9646.ENSAMEP000000017826	<i>Ailuropoda melanoleuca</i> (giant panda), species, carnivores
Chordata	Bta	>9913.ENSBTAP000000003333	<i>Bos taurus</i> (cattle), species, even-toed ungulates
Chordata	Cjac	>9483.ENSJCJAP000000020075	<i>Callithrix jacchus</i> (white-tufted-ear marmoset), species, primates
Chordata	Clu	>9615.ENSJCAF000000010792	<i>Canis lupus familiaris</i> (dog), subspecies, carnivores
Chordata	Cpo	>10141.ENSJCP000000009465	<i>Cavia porcellus</i> (Guinea pig)
Chordata	Dre	>7955.ENSJARP0000000102182	<i>Danio rerio</i> (zebrafish), species, bony fishes
Chordata	Eca	>9796.ENSECAP000000000109	<i>Equus caballus</i> (horse), species, odd-toed ungulates
Chordata	Fca	>9685.ENSJFCA0000000020786	<i>Felis catus</i> (domestic cat), species, carnivores
Chordata	Gac	>69293.ENSJGAC000000004237	<i>Gasterosteus aculeatus</i> (three-spined stickleback), species, bony fishes
Chordata	Gga	>9031.ENSJGALP0000000038520	<i>Gallus gallus</i> (chicken), species, birds
Chordata	Gmo	>8049.ENSJMOP000000004426	<i>Gadus morhua</i> (Atlantic cod), species, bony fishes
Chordata	Hsa	>9606.ENSJHSP0000000246548	<i>Homo sapiens</i>
Chordata	Itr	>43179.ENSJSTP0000000002332	<i>Ictidomys tridecemlineatus</i> (thirteen-lined ground squirrel), species, rode
Chordata	Laf_a	>9785.ENSJLAFP0000000011542	<i>Loxodonta africana</i> (African savanna elephant), species, placentals
Chordata	Mga	>9103.ENSJMGAP000000006132	<i>Meleagris gallopavo</i> (turkey), species, birds
Chordata	Mlu_a	>59463.ENSJMLUP000000002773	<i>Myotis lucifugus</i> (little brown bat), species, bats
Chordata	Mlu_b	>59463.ENSJMLUP0000000018571	<i>Myotis lucifugus</i> (little brown bat), species, bats
Chordata	Mmul	>9544.ENSJMUP000000007742	<i>Macaca mulatta</i> (Rhesus monkey), species, primates
Chordata	Mmus	>10090.ENSJMUSP0000000099807	<i>Mus musculus</i>
Chordata	Mpu	>9669.ENSJMUP0000000007408	<i>Mustela putorius furo</i> (domestic ferret), subspecies, carnivores
Chordata	Nle	>61853.ENSJNLEP0000000013124	<i>Nomascus leucogenys</i> (northern white-cheeked gibbon), species, primates
Chordata	Ocu	>9986.ENSJOCUP0000000001260	<i>Oryctolagus cuniculus</i> (rabbit), species, rabbits & hares
Chordata	Ola	>8090.ENSJORLP0000000015752	<i>Oryzias latipes</i> (Japanese medaka), species, bony fishes
Chordata	Oni	>8128.ENSJONIP0000000013939	<i>Oreochromis niloticus</i> (Nile tilapia), species, bony fishes
Chordata	Pca	>9813.ENSJPCAP0000000009534	<i>Procavia capensis</i> (Cape rock hyrax), species, placentals
Chordata	Ptr	>9598.ENSJPTRP0000000018512	<i>Pan troglodytes</i> (chimpanzee), species, primates
Chordata	Rno	>10116.ENSJRNOP0000000028672	<i>Rattus norvegicus</i>
Chordata	Sha	>9305.ENSJSHAP0000000002779	<i>Sarcophilus harrisi</i> (Tasmanian devil), species, marsupials
Chordata	Tgu_a	>59729.ENSJTGUP000000009909	<i>Taeniopygia guttata</i> (zebra finch), species, birds
Chordata	Tni	>99883.ENSJTNP0000000005165	<i>Tetraodon nigroviridis</i> (spotted green pufferfish), species, bony fishes
Chordata	Xma	>8083.ENSJXMAP0000000016139	<i>Xiphophorus maculatus</i> (southern platyfish), species, bony fishes
Chordata	Cin	>7719.ENSJCINP0000000030774	<i>Ciona intestinalis</i> (vase tunicate), species, tunicates
Chordata	Lcha	>7897.ENSJLACP000000007856	<i>Latimeria chalumnae</i> (coelacanth), species, coelacanths
Chordata	Mdo	>13616.ENSJMODP0000000014955	<i>Monodelphis domestica</i>
Chordata	Oan	>9258.ENSJOANP0000000015209	<i>Ornithorhynchus anatinus</i> (platypus), species, monotremes
Chordata	Oga	>30611.ENSJOGAP000000007933	<i>Otolemur garnettii</i> (Small-eared galago) (Garnett's greater bushbaby)
Chordata	Psi	>13735.ENSJPSIP0000000010107	<i>Pelodiscus sinensis</i>
Chordata	Tru	>31033.ENSJTRUP0000000016721	<i>Takifugu rubripes</i> (Japanese pufferfish) (Fugu rubripes)
Cnidaria	Nve	>45351.NEMVEDRAFT_vlg161397-PA	<i>Nematostella vectensis</i> (starlet sea anemone), species, sea anemones
Cnidaria	Hma	>6085.XP_002165708	<i>Hydra magnipapillata</i>
Nematoda	Cbr	>135651.CBN10217	<i>Caenorhabditis brenneri</i> (Nematode worm)
Nematoda	Cbri	>6238.CBG13604	<i>Caenorhabditis briggsae</i> (nematodes)
Nematoda	Cel	>6239.W02A11.4	<i>Caenorhabditis elegans</i> (roundworm), species, nematodes
Nematoda	Cjap	>281687.CJA09929	<i>Caenorhabditis japonica</i>
Nematoda	Cre	>31234.CRE28085	<i>Caenorhabditis remanei</i> (<i>Caenorhabditis vulgaris</i>)
Platyhelminthes	Sma	>6183.Smp_166220_mRNA	<i>Schistosoma mansoni</i> (flatworms)
Porifera	Aqu	>400682.PAC_15728046	<i>Amphimedon queenslandica</i> (species, sponges)

Supplementary Table S1. Phylum, symbol, sequence code and species used in the alignment of the Sae2 UFD domain from metazoa.