Article Addendum

A large family of genes that share homology with CLE domain in Arabidopsis and rice

Shinichiro Sawa,* Atsuko Kinoshita, Shigeyuki Betsuyaku and Hiroo Fukuda

University of Tokyo; Tokyo, Japan

Key words: CLE, CLAVATA, meristem, SAM, RAM, peptide

Thirty-one <u>CL</u>AVATA3/<u>ENDOSPERM</u> SURROUNDING REGION (ESR)-related (CLE) proteins are encoded in the Arabidopsis genome, and they are supposed to function as dodecapeptides with two hydroxyproline residues. Twenty-six synthetic CLE peptides, corresponding to the predicted products of the 31 *CLE* genes, were examined in Arabidopsis and rice. Nineteen CLE peptides induced root meristem consumption, resulting in the short root phenotype in Arabidopsis and rice, whereas no CLE peptides affected the shoot apical meristem in rice. Database searches revealed 47 putative *CLE* genes in the rice genome. Three of the rice *CLE* genes, *OsCLE502*, *OsCLE504* and *OsCLE506*, encode CLE proteins with multiple CLE domains, which are not found in the Arabidopsis genome, and polyproline region was found between these CLE domains. These results indicate conserved and/or diverse CLE functions in each plant species.

Intercellular communication is a fundamental mechanism for coordinating the development of complex bodies of multicellular organisms such as plants and animals. Peptide signaling in plants has been largely overlooked for many years, despite the importance of peptide signaling in animals, yeast and other organisms. The recent identification of several peptide hormones indicated the importance of cell-cell communication^{1,2} in defense responses,³ cell proliferation,⁴ cell differentiation,⁵ shoot apical meristem (SAM) size regulation,⁶ self-incompatibility in crucifer species,⁷ and stomatal patterning.⁸

CLAVATA3 (CLV3), and tracheary element differentiation inhibitory factor (TDIF) were shown to be involved as <u>CL</u>AVATA3/ <u>ENDOSPERM SURROUNDING REGION (ESR)-related (CLE)</u> members, and they function as dodecapeptides.^{5,6} Chemically biosynthesized peptides could be a powerful tool for examining the

*Correspondence to: Shinichiro Sawa; University of Tokyo; Graduate School of Science; Hongo 7-3-1; Bunkyo-ku, Tokyo 113-0033 Japan; Tel.: +81.3.5841.4464; Fax: +81.3.5841.4464; Email: sawa@biol.s.u-tokyo.ac.jp

Submitted: 11/23/07; Accepted: 11/28/07

Previously published online as a *Plant Signaling & Behavior* E-publication: http://www.landesbioscience.com/journals/psb/article/5344

Addendum to: Kinoshita A, Nakamura Y, Sasaki E, Kyozuka J, Fukuda H, Sawa S. Gain-of-function phenotypes of chemically synthetic CLAVATA3/ESR-related (CLE) peptides in *Arabidopsis thaliana* and *Oryza sativa*. Plant Cell Physiol 2007; 48:1821–5; PMID: 17991631; DOI: 10.1093/pcp/pcm154. CLE peptide functions. Twenty-six putative CLE peptides encoded in the Arabidopsis genome were investigated. Nineteen CLE peptides functioned not only in Arabidopsis but also in rice to reduce root apical meristem (RAM) size, whereas no Arabidopsis peptides affected rice SAM size in our assay system. However, 10 CLE peptides exhibited a strong effect on the Arabidopsis SAM.⁹ This may indicate that the CLE peptides function less redundantly in the SAM than in the RAM, and that some Arabidopsis CLE peptides can bind less effectively to rice receptors due to the sequence differences between Arabidopsis and rice.

Sequencing of the rice genome has finished, and rice genes encoding putative CLE domains were database searched from RAP-DB (http://rapdb.lab.nig.ac.jp) using the Arabidopsis CLE sequences of 12 amino acid residues as queries. The resulting rice sequences were used as queries to repeat the step in an iterative manner. The search was terminated when no novel sequences were retrieved. A total of 47 putative *CLE* genes were found in the database search (Fig. 1).

The coding regions of *CLV3* and *CLE40* genes are interrupted by two introns in Arabidopsis,¹⁰ and the application of these synthetic peptides to wild type plants induced SAM consumption. In rice, nine *CLE* genes, *OsCLE201*, *OsCLE305*, *OsCLE402*, *OsCLE502*, *OsCLE506*, *OsCLE507*, *OsCLE47*, *OsCLE603* and *FON2/FON4*, have multiple exons. *FON2/FON4* has been reported to regulate floral meristem size in a similar manner to the *CLV3* in Arabidopsis,^{11,12} and another eight *CLE* genes might be involved in meristem size regulation.

Three rice CLE genes, OsCLE502, OsCLE504 and OsCLE506, encode CLE proteins carrying multiple CLE domains (Fig. 1), although no Arabidopsis gene encodes such a CLE protein. In addition to the results described in Kinoshita et al.,⁹ four more putative CLE domains, OsCLE502C (REVPSGPDPITS), OsCLE502D (REVPSGPDPITS), OsCLE502E (RKVHHKALGIAS) and OsCLE506F (RLTPIGPDPIHN), were found in these rice CLEs. The two sequence-related rice CLE genes, CLE504 and CLE505, are located at interval of no more than about 2 kb on chromosome 5, suggesting that they have arisen by local gene duplication events. However, OsCLE504 has two putative CLE domains, whereas OsCLE505 has one apparent CLE domain and another CLE-like motif composed of an unusual 13 amino acid residues, HDVPSSGPSPVHN, which should correspond to OsCLE504A (Fig. 1). The CLE-like peptide might not function any longer as a CLE peptide due to an additional amino acid. It is possible that

		20 40 60 80 100
		+++++
	FON2/FON4	MQSQRGEEIFDLVFDLV
	OSCLE501	MA VERNVTAAANASLVIL-
	OSCLE502	MARAVEKNIN VAAAAAAALAALLAALLYMIMILSSI TAITAAKYVEGOPPOELEPIIITIPPIAISGESDISVILKKVPIGPDPITSDPPPPPPSTPTQFSVLKKVPIGPDPITSDPPPPP
GCC2001 NMGAAMERCILLANAMEL	OSCLE504	MATTHAORRECLEMALUF
	OsCLE505	MATGAAANHGRCSLLLFAAIAMVLLVFVATTTAAAARDVRR
	OsCLE206	MRRFSKQHLVPFILLLLLVMSHLPISSLGSRRAFREE-
	OsCLE508	MGGLAPSAPWRWSWSVARAVFLASLLVLASAQQQPRPPRAPEMSAVDVDAILARVCGGGSSRQAAPVPPLPLCHEL
	OSCLE104	M
	OsCLE101	MV VIDITYTERENCONSDED
ed:ED:5 H188 188 188 188 ed:ED:5 H188 188 188 188 ed:ED:5 Head Constructure (Constructure (Construe (Construe (Constructure (Construe (Constructure (Construe (Const	OsCLE301	MSSWSWSVRLMVLTLTLTLGAELAGIRHGRRIIPSLEG-DSPSPSPSASKGCGLLTEEMAAPPGHYYS
CHC2001 HEBRICAL WILLIAM/TRADUCTORIZA ADJUBRICAL MULTIPACIAN DE LIAMANTA DE LA MARTICALORES (CTTURNE) CHC2001 HEBRICAL WILLIAMANTA DE LIAMANTA DE LI	OsCLE15	MLRSRKSRVMVMLVTAALLLTDMAGVSYGRRLIPDLDAMAVVGGSPPAKG-GYMSR-LQVPPSDS
International Construction Linkawaru	OsCLE202	MSRRLGAAAAVLLLWLAVLTFALHGYYGGRLGSARRNILLQHPALALHLPTRKML-AVASFDDASSPSSLTTTDRHHH-
International (International International Internatione Internatinte International International International Internat	OSCLE403 OgCLE204	MRUTGAA
GCL2001 PML28EQ(FAAAAM2LC2/PML2/WAAA THEFUG PYLAVEAALTSAVVELUMALEAAVEAU GCL2001 HERRENTALACIEV/PML2AAAA CONSTRUCTORY AND	OSCLE303	MVMLLRRGLISAAFLLTLVLLFLLQLGIISTTVGCCSCCSFCGDEQQKQQQQHGVGDLRPGRRLLIGHH0HQVVLAKGSMEL
declassi H - BESINTREGNALCIANYOLCAMA AMELIGUIDALEDIALITIAN DELCO. declassi PPPE - ATTENTIANA DELANACIANA DELCO. AMELIGUIDALEDIALITIANA DELCO. declassi PPPE - ATTENTIANA DELANACIANA DELCO. AMELIGUIDALEDIALITIANA DELCO. declassi PPPE - ATTENTIANA DELANACIANA DELCO. CELLEC. declassi PRESENTATION TENTIANA DELANA DELANACIANA DELCO. CELLEC. declassi PRESENTATION TENTIANA DELANA DELANACIANA DELCO. CELLEC. declassi PRESENTATION TENTIANA DELANA DELANACIANA DELCO. CELLEC. CELLEC. declassi PRESENTATION TENTIANA DELANA DELANACIANA DELANA DELANACIANA DELANA DELANACIANA DELANA DELANACIANA DELANA DELANA DELANACIANA DELANA DEL	OSCLE302	MVLRRQRPAAAAAGGLLGPWLL-VAVLV
	OsCLE603	MRRRRWPARCAAALCIAVVVLQLAAAA
CHEESE H	OsCLE103	MRPARRGWGVGALARAFLVILLLLIAAAATTTT
International Control Internatinternatinterenational Control International Con	OSCLE404	MFFFFAIIFDERKKUAALIEUHAALBENG
eccl.bs/s MMANUSCREPAA	OsCLE201	${\tt MSTGRVRHKCTTQPMIFATSLRVKLVGATPLPARLSPPAPPASVVRACACLPRLGIGASAMWRRGRAMRWRRRWKYSLPAACLALVLLAAAAARGVPIGGGGMAHTGGGRVVACACLPRLGIGASAMWRRGRAMRWRRRWKYSLPAACLALVLLAAAAARGVPIGGGGMAHTGGGRVVACACLPRLGIGASAMWRRGRAMRWRRRWKYSLPAACLALVLLAAAAARGVPIGGGGMAHTGGGRVVACACLPRLGIGASAMWRRGRAMRWRRRWKYSLPAACLALVLLAAAAARGVPIGGGGMAHTGGGRVVACACLPRLGIGASAMWRRGRAMRWRRRWKYSLPAACLALVLLAAAAARGVPIGGGGMAHTGGGRVVACACLPRLGIGASAMWRRGRAMRWRRRWKYSLPAACLALVLLAAAAARGVPIGGGGMAHTGGGRVVACACLPRLGIGASAMWRRGRAMRWRRRWKYSLPAACLALVLLAAAAARGVPIGGGGMAHTGGGRVVACACLPRLGIGASAMWRRGRAMRWRRRWKYSLPAACLALVLLAAAAARGVPIGGGGMAHTGGGRVVACACLPRLGIGASAMWRRGRAMRWRRRWKYSLPAACLALVLLAAAAARGVPIGGGGMAHTGGGRVVACACLPRLGIGASAMWRRGRAMRWRRRWKYSLPAACLALVLLAAAAARGVPIGGGGMAHTGGGRVVACACLPRLGIGASAMWRRGRAMWRRRWKYSLPAACLALVLLAAAAARGVPIGGGMAHTGGGRVACACLPRLGIGASAWWRRGRAMWRRRWKYSLPAACLALVLLAAAAARGVPIGGGMAHTGGGRWWRGRAMWRRGRAMWRRRWKYGGGWAATGAAGAAGAARGVPIGGGMAHTGGGRWWRGRAMWRRGRAMWRRGRAMWRGRAMWRGRAMWRGRAMWRGRAMWRGRAWWKYGGGGMAHTGGGGWAATG$
doct_biol MCC_ADMUNICATION CONTRACT	OSCLE205	MARARWIGDGRRPAAALPLLGLCAFLLOVSLAPPPGEEEEEAKVRSSSCAVMLVVSLAPPPGEEEEEAKVRSSSCAVMLVVSLAPPPGEEEEEAKVRSSS
000120910 N=045 1000110 1000110 1000110 00012071 N=055 985 1000110 1000110 00012071 N=055 985 1000110 1000110 00012071 N=055 985 1000110 1000110 1000110 00012071 N=0501120110 1000110 1000110 1000110 1000110 000120810 N=0501120110 1000110 1000110 1000110 1000110 1000110 000120810 N=0501120110 1000	OsCLE102	MDTARPVHPLRVHGESIRGLLLLLLFVVQCSLLSCCLAHAAAAADAVDRDDPVVTATAG
GGELED1 M BEYARBARKUNT, TLALLED - GLIL VI. LLMTENALPRE- GGELED15 M BEYARBARKUNT, TLALLED - GULIL VI. LLMTENALPRE- GGELED15 M BEYARBARKUNT, TLALLED - GULIL VI. VI. GGELED15 M BEYARBARKUNT, VI.LAS - TEVIA. HIM ANNALATA GGELED1 M BEYARBARKUNT, VILLAS - TEVIA. HIM ANNALATA GGELED1 M BARLANDUNT, VILLAS - TEVIA. HIM ANNALATA GGELED1 M ANNALATUNT, VILLAS - TEVIA. HIM ANNALATUNT, VILLAS - TEVIA. GGELED1 M ANNALATUNT, VILLAS - TEVIA. HIM ANNALATUNT, VILLAS - TEVIA. GGELED1 M ANNALATUNT, VILLAS - TEVIA. HIM ANNALATUNT, VILLAS - TEVIA. GGELED1 M ANNALATUNT, VILLAS - TEVIA. HIM ANNALATUNT, VILLAS - TEVIA. GGELED1 M ANNALATUNT, VILLAS - TEVIA. HIM ANNALATUNT, VILLAS - TEVIA. GGELED1 M ANNALATUNT, VILLAS - TEVIA. HIM ANNALATUNT, VILLAS - TEVIA. GGELED1 M ANNALATUNT, VILLAS - TEVIA. HIM ANNALATUNT, VILLAS - TEVIA. GGELED1 M ANNALATUNT, VILLAS - TEVIA. HIM ANNALATUNT, VILLAS - TEVIA. GGELED1 M ANNALATUNT, VILLAS - TEVIA.	OsCLE507	MACC - RIVLVFFF - FFFF - FFFF - FCVCV - LVATTTTMDAV
OPECED35 M CODDG COUNT - CTVL. - VEXAMORATION/LASSIS. OPECED36 M VEXAMORATION/LASSIS. - VEXAMORATION/LASSIS. OPECED36 M ALIABULLT/LLILBIA COUNT - VEXAMORATION/LASSIS. OPECED361 M ALIABULT/LLILBIA COUNT - VEXAMORATION/LASSIS. OPECED362 M ALIABULT/LLILBIA COUNTALIA - VEXAMORATION/LASSIS. OPECED363 M ALIABULT/LLILBIA COUNT	OSCLE71	M SRPARSARVSAT TLALIFG-GLILVS
Oct.Eb91 MYUITSILLLIVITUWIYUWIYUUU	OsCLE305	MCGGGGGGQMVVVVLLVCVVLVSSAAMAGGARTGPVHLAGGTASSSSA
doct.2801 M ALABSLICTULLUSI- QOVIDI. C213-ALABSLICTULLUSI- QOVIDI. doct.2801 M ALABSLICTULLUSI- QOVIDI. C213-ALABSLICTULUSI- QOVIDI. doct.2801 M ALABSLICTULUSI- QOVIDI. FERRICA- ALABSLICTULUSI- QOVIDI. doct.2801 M ALABSLICTULUSI- QOVIDI. FERRICA- SAN- doct.2801 M ALABSLICTULUSI- QUVIDIAL FERRICA- SAN- doct.2801 M ALABSLICTULUSI- QUVIDIAL FERRICA- SAN- doct.2801 M ALABSLICTULUSI- QUVIDIAL FERRICA- SAN- doct.2803 M ALBSLICTULUSI- QUVIDIAL FERRICA- SAN- doct.2804 M ALBSLICTULUSI- ANA FERRICA- SAN- doct.2805	OsCLE304	MVQHTSRLDLFLVCTVVVVVVALHHATAVHGLTRAELVLAPAPAMAPAPA
Non-19951 PE-2004.04.07.07.07.07.04.06.07.00.07.07.00.00.07.07.00.00.07.00.00.	OsCLE801	MALSANRLLTVLLCLLLSH-QCKVYGLKGIS-LAFGGEEDEVPEKKPRVLAQSN
0-012001 I	OSCLE401 OSCLE203	m
OPEEB02 HRKLAM-CPT2UVFLUVVTPL	OsCLE601	MAYSSSTLRSPLVRAAVVLMLLLVVMSAAVSRGEPDHDHVQLQLAV
doct.bs1 HAKLAI-CPUVUVUVUVUASSAPERIODERAAGI-LGREGLQODAI	OsCLE602	MAKLAM-CFFICAVFLLVVTTPGLPRLAGSVPLGRRWLQDSAV
00121886 H<	OsCLE51	MAKLAL-CFCVVLVLVLVLASSPAPLSDDRRAAGL-LGRRGLQQDAI
Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion Conclusion	OsCLE46	MKLITLSCLCLCLLLUV-TOSSSPVSVSVSGDRCPVLHHRRLHDMVAAA-
OctEB39 H	OSCLE402 OSCLE44	M ABPHCAVKPCCCPCCCLLVAAVLAVA
OBJELESS MARXANUMALFLACTUD	OsCLE49	M
Oct:1837 MDOQUERINGLICERPQRUMA-HTC	OsCLE50	MAAKAAVVMALFLACTLDRAWTPTGAARTMTMMMVQRGEVSAVVA
Decl:B8509 MNKAREKKULCVVVV.: J.L.LLAAAU	OsCLE47	MDGQDRSNGKLCRRPQRQRMRTQ
0010183 N	OsCLE509	MNKAGKKVL/CCVVVV-LL/CAAAVQASAARPLLRETAVVGGDBGSQVGGG-
0:0:0:2333 H	OSCLE45 OSCLE48	M AA-KIILL AVFAF ALIMISAVAK
0erLB42 N=AR-R-BASI TURAVIA-CVLIVOTTSS	OsCLE43	MAR-RVGVVVA-VIAA-CVLLVVGMTAPPVDAAAAAARRL-GNGRDA-AVTDPA
120 140 140 140 200 DOUL/FEMA TELECONSTRUCTORENTSENSENSENSENSENSENSENSENSENSENSENSENSENS	OsCLE42	MAR-RASIIVAAVIAA-CVLLVCMTTSSVVDAAAAAPARRLLGSGRDDDAVAAPV
1.10 1.10 1.00 1.00 2.00 0.012/F001		
PONL/PON4 TEXPTCOPEON/COTERPOSESTARS		
GetL8501	FON2/FON4	$ {\tt TEFKTKQPRGVKGTRRPSWSSWSSTASRSSPPPgRGAPSAAAAAELRSVPAGPDPMHHGSPRRPEHARSTGRP$
GelL8502 LSEFYUREUYSGIDDDILSE -DEPPEPEPER LSEFYUREUYSGIDDILSE	OsCLE501	PPGBLPPIRTPPPSVISGPRVSPVLRSVPTGPNPITSDPPPPPNHERFQADEDKGVLFPSKPRGHVPPSGPSKPPPSYHLS
OBLEB306 PRINCIPAL OF LINE AND	OSCLE502	LSEFFVIREVPSGPDPITS-DPPPPPPPLSEFFVIREVPSGPDFITSDPPPPPPPITEFFILREVHHKALGIASKOGGLRPAKP
OutL2506 PAAVEKKMANNE PAAVEKKMANNE OutL2506	OSCLE504	PROGRATOGRATIFICEPUPITREFQUPPPEPERERENDAATOGRATIFICEDEPITREFERERENDAATOGORGATIFICEDEPAREREQUPPEPERERATA
0eCL2206	OsCLE505	PAAVEKKMAAVMR <u>HDVPSSGPSPVHN</u> GAPTPPAAADEPTVAVTERLVPTGSNPLHNMPSPLQGRTPTKKTVN
GelCLES06	OsCLE206	AVSGFRSHELAPTMAPSQEKEAGVVAGAGSICGQKYAVSRRMVPQGPNPLHN
0stL1200	OsCLE508	WRHRGGVRHHRRPAPPGR-DEEVDLRYGVARLVPTGPNPLHN
OctEB 01 SASSYRINGERMEQUIAANAVAFHARPRPHOGOG FEAD BULLAPTOSE VYPEGENPLAR OctEB 01 SERMQ FEED MUYDASE VYPEGENPLAR OctEB 01 SERMQ FEED MUYDASE VYPEGENPLAR OctEB 02	OSCLEI04 OSCLE503	KRKINGGGGGGVGVGVGVGVGVFFFSSRAGELIDAKIUVSSRVVFSGVFLINI SSSAHGVHVDOSBHAAFTREFGTGTGGDDCFFSDVKFSFTGSNPLINI
OetCLE301	OsCLE101	SASSYRNNRRMEGGLAAAAVAFHARRFRPHGGGGFEADKRLAPTGSNPLHNLR
OeCLE25	OsCLE301	SRRMQEEEEDHGVSASKRVVPEGPNPLHN
OBCLE202 HHRHHHHHHHHHHHRUDDWNRKQYPTAAGFGEEUDPR-FQVQKUVPQGNPLINH OBCLE203	OsCLE15	GHHVGDEYRSMHAVSKRLVPQGPNPLHN
OccLE201 Linuxiani L	OSCLE202	HHKHHGHHHKGHDKWNKGUPPTAAGGEEVUPR-FGVQKLUPTGPNPLHH SBAATTPTTPTTPTUAGGGEVUPR-FGVQKLUPTGPNPLHH
OscLE303	OsCLE204	TPKHGTYLMRAAAPGGESRAAM-YGESKLVFYGGNPLHN
OcCLB302 PTARELVG	OsCLE303	
OdcLE603 PVHRAVAKAGGRSTAFDAGGVPCKEKSGGHGGAPSCSDDDUKKVVPTGVPNLIHRKNINCPHQLPLHGVFVSIPQVEDSAVHHV OdcLE201 APGSARSSKGHGRRSHINARAVVDAAMPVGSTPVFALSPAADESSKILTSCDPLHRS OdcLE201 VASSTAFDAAAAASCKKQGSSSGAGTOPAAAGRGKWSPPAADABKVVPTGVPNLIHR OdcLE201 VASSTAFDAAAAASCKKQGSSSGAGTOPAAAGRGKWSPPAADABKVVPTGVNLIHR OdcLE205	OsCLE302	PTAAELVGDSKRKVPTGANPLHNR
Artsmitsingsmit	OSCLE603	PVHRAVAKAKGGGRSTAFDAGGGVPCKEKSGGHGGAPSPCSDDDDKKVVPTGPNDLHKNKKCPHQLPLHGVFVSIPQVEDSAVHHV a dosage schurauberuna ambuygter duydal (saka findervede in beggende i under schure)
OscLE404 APDVVVAGLUGIGLGHRMRPGGGDLVDDEXRVPTGPNPLINR OscLE201 VASSTAAFDAAAAASGCKKQRSSRGGTGPAAAGRGWSPPAAAARDSKVVPTGSNPLINR OscLE205	OscLE24	AE USANGUNANDIARAD DIWAWAY VAANAY VAANAY VAANAY VAANAY AA A
OsCLE201 VASSTAAFDAAAAASRCKKQRSSRGGATGPAAAGRGWSPPAAAADERVVPTGSNPLHRK OsCLE205 LPAAATSVPAGGRRLLLPAARTRHFRFRWRWAADIDSKVEVPSGPNPDSN OsCLE302 RGRFLFPSPALQLHSVQVNVAAHPWRKERRSRARTRRRARATLMAVSKGVVPTGANPDSN OsCLE306 PAIAAGRRHLVGAADAQWRTLEDFKADDFFRGSKRRIPKGPDPINNRYCKACFILSLILLICVSNPCSKY OsCLE306 PAIAAGRRHLVGAADAQWRTLEDFKADDFFGDSKRRVPNOPDPINNRYCKACFILSLILLICVSNPCSKY OsCLE305 PGPAVATPRODAAGATTMTATTTTTTAAATTATFAADPYKDSKKKVPNOPDPINNRYCKGCCSLKKTRFGVEWKVSCQVDDDDDDDHIMDDGE OsCLE304 PPANNVGVDAAKERPAATTAAQT	OsCLE404	A FDVVVAGLVG I GLGHRWR PGGGDLVDEKKRVPTGPNPLHNR
OscLE205 LPAARTSVPAGGRELLLPAARTREFERERMAATLDAVSKEQVPTGANPDSN OscLE507 RGREFLPSPALQLHSVQVNVAAHPWSKERRSRRRRRAATLMAVSKEQVPTGANPDSN OscLE507 KGTSAAAAGGGGGGGRWTELTAGSPARYSAGADEFRGSKRRTPKOPDPTINNRRAGKTTVAPRRD OscLE306 PAIAAGRRHLVGAADAGQMETLEDPKADDPFSGSKRRVPNOPDPTINNRYCKACFILSLLILICVSNPCSKY OscLE305 PGPAVATPRGDAAGATMTATTTTMTAAATTAFFAADPYKDSKRVPNOPDPTINNR OscLE304 PPANNVGVDAAKEFFAATTAAAQTSKMRRVPNOPDPTINNR OscLE304 PANNVGVDAAKEFFAATTAAAQTSKMRRVPROSDPTINNRS OscLE401 AAFDASVKAAAATATATGSSPSKVFDPDRMSKRRVPRGSDPTINNC OscLE203 AAVNVTAAAAAAARPTSTEKNTAATGSSPSVVEDPDRMSKRRVPRGSDPTINNC OscLE203 AAVNVTAAAAAAARPTSTEKNTAATGSSPSVVEDPDRMSKRRVPRGSDPTINNC OscLE203 AAVNVTAAAAAAARPTSTEKNTAATGSSPSVVEDPDRMSKRRVPRGSDPTINNC OscLE401 NAVAATATFDASVKAAAATATTGSSPSVVEDPDRMSKRRVPRGSDPTINNC OscLE51	OsCLE201	vasstaafdaaaaaasrckkorssrggatgpaaagrgwsppaaaaadervvptgsnpl h r
USLDIUZ RGRKPLPSFAULILS VUW VARHPTSEKKRSKRKRKRKATA JWSKQVPTGANFDSN OSCLES10 RGRKPLPSFAULILS VUW VARHPTSEKKRSKRKRKRKATA JWSKQVPTGANFDSN OSCLE310 PALAAGRRHLVGADAQQMRTLEDFKADDFFQGSKRRVPNGPDPINKRACKTTVAPRRD OSCLE310 PGPAVATFELTAGSPARYSAGF-SSKRKRVPNGPDPINKRYCKACFILSLILILOVSNFCSKY OSCLE301 PGPAVATPRGDAAGATTMTATTTTTAAATTAFAA-DFYGDSKKRVPNGPDPINKRYCKACGSLKKTRFGVEWKVSCQVDDDDDDDDHIMDDGE OSCLE301 AA	OsCLE205	LPAAATSVPAGGRELLEPARTTREFPERWNSAGIDDSKHEVPSGPNDSSNR
OscLEB10 Definition Definition Definition OscLEP1 PGSSVLGVGGRRMLVQAADAQQMRTLEDFKAD-DDFODSKRRVPNOPDPINKWYCKACFILSLLILCSUNFCSKY OscLEB1 PGSSVLGVGGRRMLVQAADAQQMRTLEDFKAD-DDFSSMKRVPNOPDPINKWYCKACGRSLKKTRFGVEWKVSCQVDDDDDDDHIMDDGE OscLEB10 PGDNAUTPREDAAGATMTATTTTTAAATTAFAA-DDPYKDSKKRVPNOPDPINKRYCKGCRSLKKTRFGVEWKVSCQVDDDDDDDHIMDDGE OscLEB10 AANURMKGYSASPSSADPNMKSERVPROSDPINKYC OscLEB11 AA	OSCLEI02 OSCLE507	RUKRE LESSFLUGLES VUNVAAHTMSKEKKRSKKRKKRKRAFILMAVSKRUVFIGANDUSSN Kotskalagogggaptielingsprysa
OSCLE71 PGSSVLGVGGRRMMINGGLVHSHRSLEDFNAGDAFSSMKRVPNGPDPIHMR OsCLE305 PGPANATPRGDAAGATTMTATTTTTTAAATTAFFAA-DPVSDSKKVPNGPDPIHMR OsCLE304 PPANNVUGUDAAKEEFAATTAAAQ7SKVKRVERGSDPIHMR OsCLE801 AANLNNKGGYSASPSSADPNRMSERVRGSDPIHMRY OsCLE401 NAVAATTATPDASVKAAAATATATGSSPSKVPDPDRMSKRVRRGSDPIHMRY OsCLE401 NAVAATAAAAARPTSTSTEKTNATGSSPSKVPDPDRMSKRVRRGSDPIHMRY OsCLE601 ITGRRMLVVAGSNTATTISSOTAVAAAMPYSESKRSSPGGPDPOHH OsCLE602 VVSGGRLTPAITAA	OsCLE306	PAIAAGRMLVQAADAGQMRTLEDFKADDPFQDSKRVPNGFDFIHWRVCKACFILSLLILICVSNFCSKY
OscLE305 PGPAVATPREDAAGATTMTATTTTTTAAATTATFAADPYKDSKKKVPNOPDPINNRCKGCRSLKKTRFGVEWKVSCQVDDDDDDDDHIMDDGE OscLE304 PPANNVVGVDAAKERPAATTAAQTSKWRVRGSDPINNRCS OscLE801 AA	OsCLE71	PGSSVLGVGGRRMMINGGLVHSHRSLEDFNAGDAFSSMKRVPNGPDPIHR
OSCLESU4 PPANNVGVDAAKERFAATTAAQTSKRVRVERGSDPINRS OSCLESU3 AA	OsCLE305	PGRAVATPRGDAGGATTWITATTTTWITAAATTATFAADPYKDSKKKVPNGPDPIHNKPCKRGCRSLKKTRFGVEWKVSCQVDDDDDDDDHIMDDGE
OscLE401 NAVATATFDASVKAAAATATATGSSPSKVPDPDRNSKRVPRGSDPINKC OscLE401 ANAVTAAAAAAAARFDTSTEKNTAATGSSPSKVPDPDRNSKRVPRGSDPINKC OscLE601 ITGRRMLVVAGSNTATTISSOTAVAAA-MPYSESKSSPGGPDPOHH OscLE602 VVSGGRLTPAITAA OscLE403 VVDGSPTAAATATTTTSNPRPYSESKSSPGGPDPOHH OscLE402 VVSGGRLTPAITAA OscLE403 VVDGSPTAAATATTTTTAMPRPYSESKSSPGGPDPOHH OscLE404 VVSGPPRPPPPPPAPAARAARTSGTAVETULPRQENDCBEIDETVYEOSKLESPGGPNPOHH OscLE402 RAETTATMYTAKELREKQDVTKGAEEDVTTTTTTGFGAESEREVPTGPDPIHHKGRGPRQSP OscLE44 GGGGGGEGGENNOGAQPEQRKEPVCMAAFTSLXVAPPPPSVKGDKEVPGGPDPOHH OscLE44 GGGGGGGGGGGEGNNOGAQPEQRKEPVCMAAFTSLXVAPPPPPSVKGDKEVPGGPDPOHH OscLE40 VNGGGGGGGGEGRNOGAQPEQRKEPVCMAAFTSLXVAPPPPPSVKGDKEVPGGPDPOHH OscLE47 VAGN-GGGGAQPRFKWWTRREGGKKEVPGGPDPOHH OscLE45 OVNGGKAQREFK	OSCLE304 OSCLE801	PYANNY VYUJAAKEKPAATTAAAQT
OBCLE203 AAVAVTAAAAAAARPDTS - TEKNTAATGSSSPSTVPDP DRMSKRVRRGSDPINKC OBCLE601	OSCLE401	NAVAATATFDASYKAAATATATATGSPSKVFD2DRNSKRVFRKSD21HMKC
OsCLE601 ITGRRMLVVAGSNTATTISSOTAVAAAMPYSESKASSPGGPDPOHH OsCLE602 VVSGGRLTPAITAA	OsCLE203	AAVAVTAAAAAAARFDTS TEKNTAATGSSS PSTVFDP DRMSKRVRRGSDP I HNKC
OBCLEB02 VVSGGRLTPAITAA	OsCLE601	ITGRRMLVVAGSNTATTISSQTAVAAMPYSESKRSSPGGPDPOHH
OscLB41 VVSQPPRPPPPAPAAAAARTSGTAVETVLPRQKDDOERIDETVYEOSKLSPGOPRVQHH OscLE44 VVSQPPRPPPAPAAPAARTSGTAVETVLPRQKDDOERIDETVYEOSKLSPGOPRVQHH OscLE44 GGGGFGIRGRRRPGR	OSCLE602	VVSGGRLTPAITAA
OsCLE402 RASTIATMYTAKELREKQDVTKGAREDVTTTTTTGFGA DSSLEVPGOPDPIHHGRGPRQSP OsCLE44 GGGGGIGGRGRNQGAQPEQRKEPVGMANFTRSLAVAPPDPSVHGDREVPGOPDPOHH OsCLE50 VNGGGGGGGGGGGGGRGNNQGAQPEQRKEPVGMANFTRSLAVAPPDPSVHGDREVPSGPDPIHHGASFSSASP OsCLE50 VNGGGGGGGGGRGNNQGAQPEQRKEPVGMANFTRSLAVAPPDPSVHGDREVPSGPDPIHHGASFSSASP OsCLE50 VNGG-GGGGAQPRKWITRREIGGDKXTVPGGPDQHH OsCLE53 OVAG	OSCLES1 OSCLE46	VVSOPPPRPPPAPAPAARTSGTAVETVLPRORDGREIDETVYGGSKELSPGGPNDOHH
OsCLE44 GGGGFGIRGRRRPGR	OsCLE402	RAETTATMYTAKELREKQDVTKGAEEDVTTTTTTTGFGAESEREVPTGPDPIH#GRGPRRQSP
OscLE49 AWNEVSYTAEIVGGGGKWEVPGGPDPCHH# OscLE50 VNGGGGGGEGRONNGGAQPEQRKEFVGMAAFTRSLAVAPPPPSVHGDEVPSGPDPCHH# OscLE47 VAGN-GGGGAQPERFKWITRREIGGDKTVPGGPPPQHH OscLE509 VVAGGNQSPA	OsCLE44	GGGGFGIRGRRRPGRWNVR-SLQGGKREVPGGPDPQHHY
OSCLESU VNGGGGGGGERERKErVGMAAFIKSLAVAPPPPSHGDREVPSGPDPIHHGASPSSASP OSCLESU VAGN-GGGGADEREKWITREEIGGDKTVPGGPPDCHH OSCLESU VVAGGNQSPA	OsCLE49	ANNRVSVTAELVGG
OscLE\$109 VVAGONSPA	OSCLE50	VNGGGGGGGEKKGNNGGAOPEORKEVGMAAFTRSLAVAPPPPSVHGDREVPSGPDTHHEGASPSSASP VAGN_GGGGADAPEK
Oscle45 AGVAGMMKVPAAAAGEAAGRRSGGGAVVVVGGAAAHESKELSPGOPDPQHH Oscle48 VAARGRFRKIMREETTLDDGGAAIGEKRSPGGPDPQHH Oscle43 LEAMMPAQTTVAPVVADGGDVDV-SGSKELSPGGPDPQHH Oscle42 VNVAAAAEPIMQQPAQMVAPVVADGDDGGVVPAGSKELSPGGPDPQHH	OSCLE509	VVAGGNOSPA
OSCLE48 VAARGRFRKIMREETTLDD-GGAAIGEKKRSPGGPDQHH OSCLE43 LEAMMPAQTTVAPVVADGGDVDV-SGSKELSPGGPDPQHH OSCLE42 VNVAAAAEPIMQQPAQMVAPVVADGDDGGVVPAGSKELSPGGPDPQHH	OsCLE45	AGVAGMMKVPAAAAGEAAGRRS3GGAVVVVGGAAAHESKKLSPGGPDPQHI
OSCLE43 LEAMMPAQTTVAPVVADGGDVDV-SGSKELSPGGPDPQHH OSCLE42 VNVAAAEPIMQQPAQMVAPVVADGDDGGVVPAGSKELSPGGPDPQHH	OsCLE48	VAARGRFRKIMREETTLDDGGAAIGESKRRSPGGPDPQHH
OPCTD45 AMAMATRIMÁŠKAŘMAKAAMDADD2AAAKW2SKATBAGADNÁHH	OsCLE43	LEAMMPAQTTVAPVVADGGDVDVSGSKKLSFGGPDPOHH
	0800042	TATABASHI ANG

Figure 1. For figure legend, see page 339.

Figure 1. Alignment of the deduced polypeptides of the rice *CLE* gene family. The conserved dodecapeptide CLE region is boxed, and a CLE-related 13 amino acid sequence in OsCLE505 is underlined. Proline rich region is shown in gray. N-terminal 31 amino acid residues of CLE506, MSSISYFLVAMLLCN GFGFIVSAQVVGGGSS, are not shown because of limited space.

these sister *CLE* genes might acquire different functions during their evolutionary steps. *OsCLE506* encodes six CLE domains and it might function in rice-specific physiological events, but not in morphological events, because the transgenic plants harboring the RNAi construct of the OsCLE506 gene did not show any obvious morphological abnormalities (Sawa et al., unpublished results).

Most of the conserved CLE domains were located at or near the C-terminal end of the sequences we identified, but every CLE domain located at the middle region of the OsCLE502, OsCLE504 and OsCLE506 proteins was followed by an acidic amino acid residue and a characteristic polyproline region (Fig. 1). The polyproline sequence often adopts a rigid, rod-like secondary structure called as a polyproline type II helix, which is capable of serving as a proteinprotein interaction domain or organizing unfolded polypeptides.^{13,14} In the case of the yeast peptide pheromone, alpha factor, four copies of 13 amino acid peptides are produced from one precursor.¹⁵ The KEX2-encoded endoprotease cleaves the precursor after pairs of basic residues such as lysine and arginine and these basic residues are chopped out by a carboxypeptidase, KEX1.13 Thus, CLE proteins with multiple CLE domains might be precursors and require a similar maturation steps. In this context, the polyproline region might serve as or provide a recognition site for the first endoprotease. Many CLE precursors have lysine and/or arginine residues just after the CLE domain, and this also supports the idea that carboxypeptidases are responsible for the C-terminal maturation step.

CLE proteins are currently one of the best described families of small polypeptides in plants; however, precise molecular details, such as CLE peptide maturation, movement, reception and signaling in a target cell, remain to be solved. Further genetic and biochemical analyses of the CLE family would give insights to help unveil not only the molecular mechanisms, but also the diversity and evolution of intercellular communication in plants.

References

- 1. Sawa S, Kinoshita A, Nakanomyo I, Fukuda H. CLV3/ESR-related (CLE) Peptides as Intercellular Signaling Molecules in Plant. Chem Rec 2006; 6:303-10.
- Fukuda H, Hirakawa Y and Sawa S. Peptide signaling in vascular development. Curr Opin Plant Biol 2007; 10:477-82.
- Pearce G, Moura DS, Stratmann J, Ryan CA. Production of multiple plant hormones from a single polyprotein precursor. Nature 2001; 411:817-20.
- Matsubayashi Y, Sakagami Y. Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of *Asparagus officinalis* L. Proc Natl Acad Sci USA 1996; 93:7623-7.
- Ito Y, Nakanomyo I, Motose H, Iwamoto K, Sawa S, Dohmae N, Fukuda H. Dodeca-CLE peptides as suppressors of plant stem cell. Science 2006; 313:842-5.
- Kondo T, Sawa S, Kinoshita A, Mizuno S, Kakimoto T, Fukuda H, Sakagami Y. A plant peptide encoded by *CLV3* identified by in situ MALDI TOF-MS analysis. Science 2006; 313:845-8.
- Schopfer CR, Nasrallah ME, Nasrallah JB. The male determinant of self-incompatibility in Brassica. Science 1999; 286:1697-700.
- Hara K, Kajita R, Torii KU, Bergmann DC, Kakimoto T. The secretory peptide gene *EPF1* enforces the stomatal one-cell-spacing rule. Genes Dev 2007; 21:1720-5.
- Kinoshita A, Nakamura Y, Sasaki E, Kyozuka J, Fukuda H and Sawa S. Gain-of-function phenotypes of chemically synthetic CLAVATA3/ESR-related (CLE) peptides in *Arabidopsis thaliana* and *Oryza sativa*. Plant Cell Physiol 2007; 48:1821-5.
- Hobe M, Muller R, Grunewald M, Brand U, Simon R. Loss of CLE40, a protein functionally equivalent to the stem cell restricting signal CLV3, enhances root waving in Arabidopsis. Dev Genes Evol 2003; 213:371-81.
- Suzaki T, Toriba T, Fujimoto M, Tsutsumi N, Kitano H, Hirano H.Y. Conservation and diversification of meristem maintenance mechanism in *Oryza sativa*: Function of the *FLORAL ORGAN NUMBER2* gene. Plant Cell Physiol 2006; 47: 1591-602.
- Chu H, Qian,Q, Liang W, Yin C, Tan H, Yao X, Yuan Z, Yang J, Huang H, Luo D, Ma H, Zhang D. The *floral organ number4* gene encoding a putative ortholog of Arabidopsis CLAVATA3 regulates apical meristem size in rice. Plant Physiol 2006; 142:1039-52.
- Dmochowska A, Dignard D, Henning D, Thomas D. Y, Bussey H. Yeast KEX1 gene encodes a putative protease with a carboxypeptidase B-like function involved in killer toxin and alpha factor precursor processing. Cell 1987; 50:573-84.
- Rath A, Davidson AR, Deber CM. The structure of "unstructured" regions in peptides and proteins: role of the polyproline II helix in protein folding and recognition. Biopolymers 2005; 80:179-85.
- Kay BK, Williamson MP, Sudol M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. FASEB J 2000; 14:231-41.