Additional files 3 for Joo et al.

Article title: Common ancestry of heterodimerizing TALE homeobox transcription factors across Metazoa and Archaeplastida.

Authors: Sunjoo Joo, Ming Hsiu Wang, Gary Lui, Jenny Lee, Andrew Barnas, Eunsoo Kim, Sebastian Sudek, Alexandra Z. Worden, and Jae-Hyeok Lee.

The following Supporting Information is available for this article:

Fig. S1. Alignment of TALE homeodomain sequences of the Archae-algal collection.

Fig. S2. Homology domain alignment of KNOX sequences.

Fig. S3. GLX class is defined by PBL-Chloro domain.

Fig. S4. Extensive homology of Picsa_04995 and Klefl_00021_0250 to Chocr_41034 indicates their classification as PBX-Red.

Fig. S5. Intron-retention pattern suggests parallel evolution of KNOX and non-KNOX group classes from common duplicated TALE ancestors.

Fig. S6. Alignment of Class-B TALE proteins in volvocales.

Fig. S7. Alignment of Mam-A TALE proteins in mamiellophyceae.

Fig. S8. Alignment of Mam-B TALE proteins in mamiellophyceae.

Fig. S9. Full-length proteins are necessary for mamiellophyceae non-KNOX TALE proteins to form heterodimers.

Fig. S10. TALE interaction network defined by this study using yeast-two-hybrid assays.

Fig. S11. Identification of MEIS homologs in choanoflagellates.

Fig. S12. Identification of PBX homologs in choanoflagellates.

Fig. S13. Alignment of the MEIS homologs in Amorphea.

Fig. S14. Alignment of the PBX homologs in Amorphea.

Fig. S15. ELK-domain alignment.

Fig. S1. Alignment of TALE homeodomain sequences of the Archaeplastida-algal collection. The 106 sequences were made into an alignment after excluding 20 near identical sequences to reduce redundancy. Animal/amoeba/haptophyte outgroup sequences are included as they share homology with Archaeplastida TALEs outside the homeodomain. The three bars above the sequence numbers show predicted alpha helices. Discarded insertions are noted in red arrowheads.

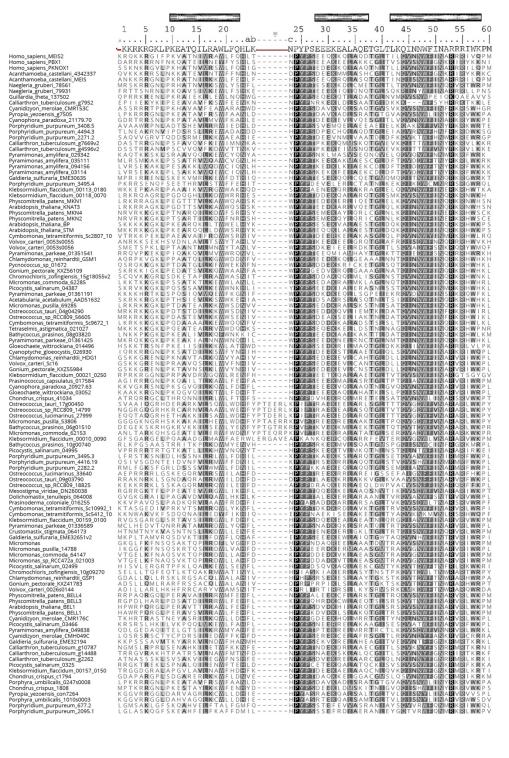


Fig. S2. Homology domain alignment of KNOX sequences. MEIS class outgroup sequences are included at the bottom. Class label is on the left. KN-A, KN-B, ELK, HOMEOBOX, KN-C1, and KN-C2 domains are labeled on the top. Class groups are labeled by colored bars on the left next to the gene names. Yellow, light green, and green shades in sequences show more than 60%, 80%, or 100% similarity in each column. Gaps between KN-A and KN-B and between KN-B and ELK have been eliminated.

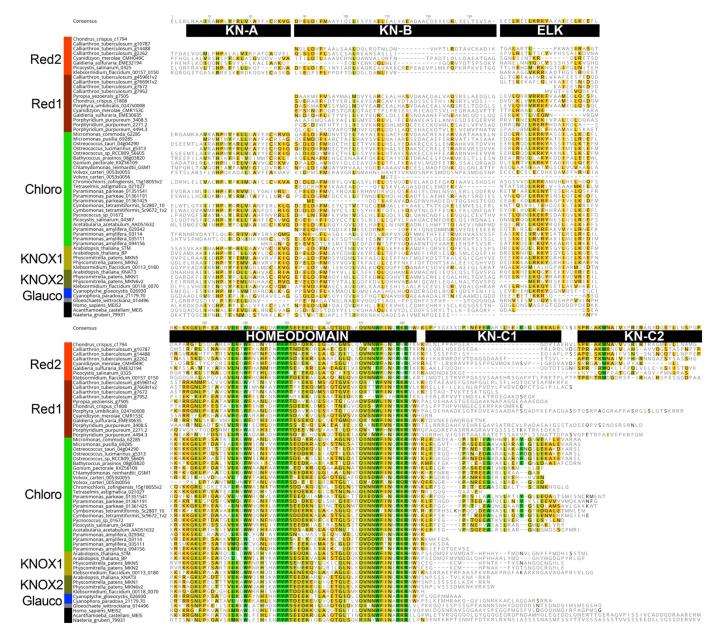


Fig. S3. GLX class is defined by PBL-Chloro domain. (A-C) Three alignments are adjusted with inserting gaps for direct comparison among different PBL-Chloro domains. (A) GLX-Chloro class members. (B) GLX-Basal class members. (C) Three Viridiplantae sequences with strong MEINOX homology domain. PBC-homology is shared among the Chlorophyta non-KNOX sequences.

(a)

(u)																				
Picocystis_salinarum_02499 Chlamydomonas_reinhardtii_GSP1 Gonium_pectorale_KXZ41783 Volvox_carteri_0026s0144 Chromochloris_zofingiensis_15g18055v2 Klebsormidium_flaccidum_00010_0090	K K K A G S P P V L R K	GAAAA	A A G G H A G T V				A G D T D G T	AQPP LAAA LPLP FSSP	P P P L P A P V P P P P	L P R T / L V W T P S P S E	EMDR	QGLCF RAALS RAAQS MAKY1	GAAG							
(b)																				
Crustomastix_stigmata_064173 Dolichomastix_tenuilepis_064008 Cymbomonas_tetramitiformis_sc10992_1 Pyramimonas_amylifera_049838 Pyramimonas_parkeae_01336589	MPS LAV	GGVN(TGAP(PGFAI	GDL GA DGS			- L R R	К L E S S G G T K	RVER QYYP FRSP	TRNA PPPN GKAT	ALCAA MA-VR	E E A E L E D V O R A P	G - R G E	E G E Q N P S O							
(c)																				
Picocystis_salinarum_02499 Prasinococcus_capsulatus_017584 Klebsormidium_flaccidum_00159_0100	VIK	Q S DM E	EPLKA	AIQA	NASL	V E	EVKH	LMSI	YKKE	LEKV	DRTC	Q G L C F Y E Q C F V P Q R E	RLLD	ESNRI	MDA	GADVEN RTDVEL	1			
Guillardia_theta_137502	VKK	T S E S I	e s s l 7	FLQL	NTP	/-MQ	QIKL	I E N Q	FNSE	KCKV	DQVC	EDVCO	2 R M							
							MEI	NOX	homol	logy		Cys-ric	h							
Picocystis_salinarum_02499 Chlamydomonas_reinhardtii_GSP1 Gonium_pectorale_KX241783 Volvox_carteri_002650144 Chromochloris_zofingiensis_15g18055v2 Klebsormidium_flaccidum_00010_0090	AFM RFM HFM EYM	AACDE VQSDF EKSDE RRODO			AMN HLG AAC ALSI	SDH SEA GKA	DAEA QAEA RIEE DLHA	- T - A Q - A S - L A	G S A G			AAP - L SA - - L SG - - OL T -	PQ 	E H A A F P G A A E A F S D L K K	SGV ANV PFTI AAVI	RQMLLD /SLVLR /NLVLR QMLYD QAVLD (QSQLD	L-K L-R VAR	– S S V – S S L – S S L – L R F	SHYI SAYI MLFL ELYK	KHLN RHLD EYLE KLLG
Crustomastix_stigmata_064173 Dolichomastix_tenuilepis_064008 Cymbomonas_tetramitiformis_sc10992_1 Pyramimonas_amylifera_049838 Pyramimonas_parkeae_01336589	E E E E Anl Evyf	EEAEE KQYEA ROLKE		QMP VTL	P P R ARQE		P P P K RMAT H A F	PVAR LAST MTLE	P - Q P Q	SNNP	R R S D K N	RLI KLFN-		E H L A Y T N G H A G S	AAAI PGQI VRVI	DSLV EAAFE TELIV DQMVS SSLTE	D - T T - V M - D	- GHF - - RRY - - EKL -	SLYN NAMV ASHI	GDDE NTLR Kyfo
Picocystis_salinarum_02499 Prasinococcus_capsulatus_017584 Klebsormidium_flaccidum_00159_0100	TAVE	E R A	- I A L A	ANS	TTQ	ENT	RANT	-HRQ	FLER	VGQI	RVQT	KVADP	AAAT	TGQR	TSNS	RQMLLD SQILPG MEYTHE	S - N	ITSY	KAYD	TIVA
Guillardia_theta_137502							_													
	1								PBC	C-B ho	moloa	V								

Fig. S4. Extensive homology of Picsa_04995 and Klefl_00021_0250 to Chocr_41034 indicates their classification as PBX-Red. MEINOX-homology and PBL-Red domains are indicated by red bars below the alignment.

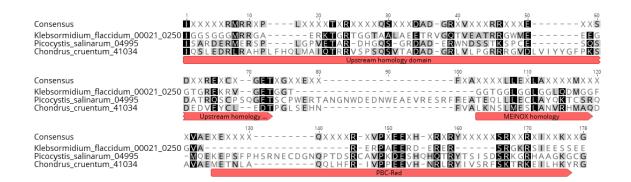
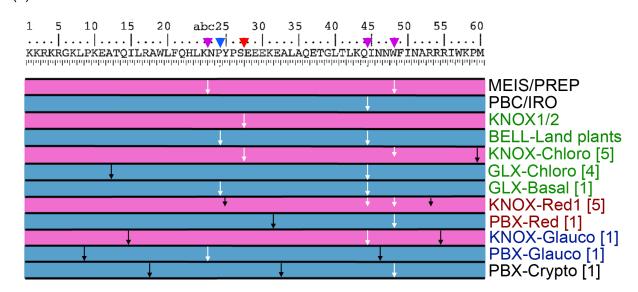


Fig. S5. Intron-retention pattern suggests parallel evolution of KNOX and non-KNOX group classes from common duplicated TALE ancestors. (A) Intron locations collected from 12 TALE classes are shown with arrows. Half arrows indicate cases where not all the class members share the position. White arrows indicate shared positions in at least two different classes, and black arrows indicate class-specific positions. The numbers above the consensus sequence show 60 amino acid positions; the three-amino-acid extension is denoted as 'abc.' Row color depicts two alternative domain configurations: purple for MEIS/KNOX types, and navy for PBX/GLX types. Class names are colored according to their phylogenetic groups: green for Viridiplantae, red for Rhodophyta, blue for Glaucophyta, and black for outgroups. The numbers following the class names show how many genes provided the intron information. Of the shared positions, purple triangles on the top mark those shared between MEIS/KNOX and PBX/GLX classes, blue triangles mark those shared between GLX and BELL classes, and red triangles mark those shared between KNOX classes. A notable exception is the KNOX-Red1 class, for which three Rhodophyta clades show different intron locations (44/45, 48[2/3] or 53[2/3]), indicating that the 44/45 intron position can indeed be displaced to 48[2/3] or elsewhere, albeit infrequently.-The unique 46/47 intron in the PBX-Glauco (Cyapa 20927) would presumably have resulted from a similar displacement in intron position. (B) Distribution of conserved introns among the TALE homeobox classes. Identified TALE classes are mapped on the Arachaeplastida phylogeny. The 44/45 intron is marked by blue outline and the 48[2/3] intron is marked by red outline. Underlines of the class names indicate the presence of a PBC-homology domain. The Archaeplastida phylogeny is modified from figure 1 of Jackson et al. (2015). (a)



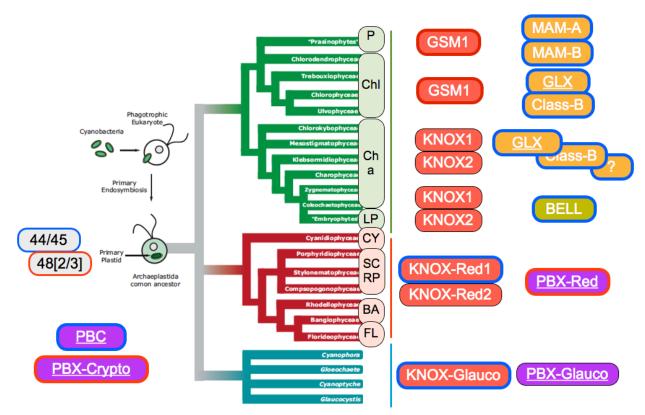


Fig. S6. Alignment of Class-B TALE proteins in volvocales. Short motifs are conserved among all members in this class over the entire length of the sequence.

Consensus	l Maxaeyshpxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
ldentity ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	МЕ А Е У SHP А А А ААААААААААААААААААААААААААААА
Consensus	ϔ sapaalnlp Χχχχχχχχχχχχχρτςροφοφήgepgs μχορραθχχχρηχχχρηχχχρηλαααταggvgwaxsχορχ
ldentity ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	VSAPAALNLPGPGHRFGGSAAWPTSPQQQHGFPGSHNODDPARSOMPPSDMOHPTAWAPYAAAATAGGVGWAGSAOOP
Consensus	PİXXXGHSPHPYÄAWHAHAHAHQHHPXHQRXLHİXXXXAPXXXXİRXXXÖLQPSİQPYGTGAVAÄTDMRGVATAİĞXXXXPX
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	PASAGHSPHPYAAWHAAHAHAHANGHMHHHPPOORHHSEHOOLGPOOROPODENGTGAVAATOMRGVATSGWGVP DH
Consensus	X BÍX X XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	HEIN HOOPPOLIPAAACSEPSSTIKKKE ENISSAATILITERELITALAAGPDGRAGGAEAAAPPGARQPPCSAGPPDGSTGVGAGG AAGAGGGGGATGSPGSSGGGGTKKKE ENISSAATILITERELITALAAGPDGRAGGAEAAAPPGARQPPCSAGPDGSTGVGAGG AROQROQQQCH
Consensus	GVGÝGMGGGGGGGGGGSAVQRRYGPAAAVAGRPANÁXX MENDERGHX XX MENDERGX MIXNÁX MITAAAAAC É A E AAARAQT NAMAXNAX
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	CVGYGMGGGGGAGRSAVQRYGPAAAVAGPAMGA GPU PPL PGGPQDV ,ARTITAAAACEA E AAARAQTXAAWAP
Consensus	AXXX Ó G XAXXXXX DÉXX G XXAAXAÁXXXXXXXXA A A A A A A A A A A A
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	A SA A GG A MA A A T Y T P G V V A MA V AWD A T A G U G SA A A A P A A A A P D S H U S S T R A A A T D O A V G A GG S V D G S G P R A T G G G G A S G M A D F A S G C T T A A A A A C Y T M V A A A A A A E G A M O A T D G E R S O U S S T G L M U II G G S S G G G A T A Q A A G O P O Q P A Q P A Q Q C Q O Q O Q O Q O Q O Q O Q O U Q O U H O H M U II G G S S G G G G A T A Q A A G A G X X X A A A G X K X G X K G G G X X X X X X X X Y E R G S A D G G E S W S V T P L P T S L P P L L K P O D V G S G X X X X X G G G X X
Identity	
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	A A A VAGA KAGGESGGESGESGESGESGESGESGESGESGESGESGESGE
Identity	
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	TTUTTSMAAVAAAAAAA AAAAAAAAAAAAAAAAAAAAATGTWGAAAWABOQHAAATBSDBSTSHSMARINGHGRHSAGGGAGGGGGRRRGSTTG AAMAAAAAAAAAAAAAAAAA
Identity ChrHD1_exonstr	GGGAAGGGGEWGRGRGTGSTAATSMOLLHLGEPIN AGDGGGGGA AGAT A AMEH SRGSAEG GGGG G GG GG GG G G
GonpeHDG1_exonstr VcaHDG1_ExonStr	- GAGHAGYGCH HAT JINCH LET CHI THAT JINCH LET CHI THAT GAGA AN AN AN AN AN AN AN AN AN AN AN AN AN
Consensus	WEXXXXXX ÜPINXXGXXXXXXIII QODIN PLTPPPAAAASPATLAWAAAAHAALAASAADXX ÜXXEXHXRQXXXXXXQQXLIQ
ldentity ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	АААНЖИР О СИРВИХ 50 БАСЕР Г. 1930 ГОЛИР Г. ТР. РРАААА 5 РАТ LAWAAAAHAA LAA 5 ААО РИРИРИ В И РИРИО ОО ОО ОО ОО О АРСИРУ 5 АРЕУРУК СЕСАА ЖАЕКИЭ. GEOREEON LSENH G - O TOS COURD - II
Identity ChrHD1_exonstr GonpeHDG1_exonstr	
Identity ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	АААНЖИР О СИРВИХ 50 БАСЕР Г. 1930 ГОЛИР Г. ТР. РРАААА 5 РАТ LAWAAAAHAA LAA 5 ААО РИРИРИ В И РИРИО ОО ОО ОО ОО О АРСИРУ 5 АРЕУРУК СЕСАА ЖАЕКИЭ. GEOREEON LSENH G - O TOS COURD - II
Identity ChrHO1_exonstr GonpeHCG1_exonstr VexHOG1_Exonstr Consensus Identity ChrHO1_exonstr GonpeHCG1_exonstr VexHOG1_Exonstr Consensus Identity Consensus	A A HWP OQ PRNS DGA GNPL P 100 CM P L TP PPA AAAS PAT L AWAAAAHAA L AA SAAD P PP PP PP PP HH HOLP EU V P OQ VEG APG PV SA PP V PM GGAAA AEMA E 100
Identity ChriefD 2849047 GraphelfGCI 280047 VeaHDGI 280057 Consensus Identity Consensus Control 280057 Consensus Identity Lidentity	A A HWP O O BON S DGA GR PL 2 103 O B PL T P P P A A A S P A T LAWAA A A HAA LAA S A A D P P PL P B H H OU P U O V O O V LO APC P V S A P P Y P M GGA A A A B M D T T T P P P A A A S P A T LAWAA A A HAA LAA S A A D P P P P B P H H OU P U O V O O V LO GE O REION L S P Y P M GGA A A B M D T T T P P P A A A S P A T LAWAA A A HAA LAA S A A D P P P P B P P P B S Y A P B S T A P P B S Y A P B GE O REION L S P Y M G GA A A B M D T T T P P P A A A S P A T LAWAA A A HAA LAA S A A D P P P B P P B P P B S Y A P B S T A P P B S Y A P B GE O REION L S P Y M G GA A A B M D T T P P P A A A S P A T LAWAA A HAA LAA S A A D P P P B S Y A P S T P P B S Y A P B P X B GU L S P V G GA A B A B M D T S P Y T A A P P G X G R C FISDUIS X M X X X X X S A A VA P P P P A L S P Y M A P À QQ X X X Q X X X X Q X X X X Q P H H GU L Q P QQ S W F I S P Y T A A P P G L S G R C FISDUIS X M X X X X X S A A VA P P P P A L S P Y M A P À QQ X X X Q X X X X Q X X X X Q P H M GU L Q P QQ S W F I S P Y T A A P P G L S G R C FISDUIS X M M X Y X X X X X A A VA P P P P A L S P Y M A P À QQ S Y H Q H Q P Q L Q P L P L Q T G M GG U
Identity ChriefD assnarr VaHDG1_Exonstr VaHDG1_Exonstr Consensus Identity ChriefD assnarr Consensus Identity ChriefD assnarr Consensus Identity ChriefD assnarr Consensus Identity ChriefD assnarr	A A ANW P O O PONS D CA GRO L DIGO CALP L TP P P AAAA S P AT L AWAAAAHAA L AA SAAD P PP P P P P P P PH HOME P MOV P O OO V LOG APC DY S S PP Y P M O G CA AR A S D O DIGO DIGO DIGO DIGO DIGO DIGO DIGO
Identity ChrHD1_exonstr GorperDC1_exonstr VeaHDG1_bonstr Consensus Identity VeaHDG1_bonstr VeaHDG1_bonstr Consensus Identity ChrHD1_sconstr Consensus Identity ChrHD1_sconstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_bonstr Consensus Identity	AA AHWP OO PENS DOAGEN LETUS ON PLTP PPAAAAS PATLAWAAAAHAALAASAAD PPP PP PPP PPHHHOUP WOV OO QOULA BARAANAAAHAALAASAAD PPP PPP PPP PPHHHOUP WOV OO QOULA BARAANAAAHAALAASAAD PPP PPP PPPP PPHHHOUP WOV OO QOULA BARAANAAHAALAASAAD PPP PPPPPPPPPPPPPPPPPPPPPPPPPPPPP
Identity Chrief) zwardtr GrappeHGL zwardtr VeaHDGL Exonstr VeaHDGL Exonstr Consensus Identity Chrief) zwardtr GonpeHGL zwardtr GongetHGL zwardtr Consensus Identity Chrief) zwardtr GongetHGL Zeonstr VeaHGL Exonstr GongetHGL Zeonstr VeaHGL Exonstr VeaHGL Zeonstr VeaHGL Zeonstr	AA AHWP OO PRINS DGAGNPL P 100 CMP L TP PPAAAAS PAT LAWAAAAHAA LAA SAAD P PP P P P PH HHOUP BU V P OO VEG APC PV SAPP V PM GGAAAAAEMI GBOR FE ML S PAPI S STAP P DTS S VA PS GBOR FE ML S PN HA O TO SOCIES - BT
Identity Chriefo J. sexnatr VarHDG1_ExonStr VarHDG1_ExonStr Consensus Identity Chrief0_soxnatr GenpeHDG1_exonstr VarHDG1_ExonStr VarHDG1_ExonStr VarHDG1_ExonStr VarHDG1_exonstr VarHDG1_exonstr VarHDG1_exonStr VarHDG1_exonStr VarHDG1_exonStr VarHDG1_exonStr VarHDG1_exonStr VarHDG1_exonStr VarHDG1_exonStr VarHDG1_exonStr VarHDG1_exonStr VarHDG1_exonStr VarHDG1_exonStr VarHDG1_exonStr VarHDG1_exonStr	AA AHWP OO BINS DO GAOR PLETING ON PLETP PEAAAAS PATLAWAAAAHAA LAASAAD PP PP PP PP PH HOUP EUG VP OO VEG ABC OF VS APB YPM GO GAAMAEMS IN THE SAR VS TAP PINS VAAP GEOREIONLS RUNNEQ OT OS CONSTITUTION OF THE SAR VS APP PROFESSION OF THE SAR VS APP OF THE SAR VS APP OF THE SAR VS APP OF THE SAR VS APP PLAN BOUL SPOND OF THE SAR VS APP OF THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND THE SAR VS AND THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AND VS AR VS AS THE SAR VS AS THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR VS AS THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR VS AS THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR VS AND THE SAR
Identity ChritOL secontr VaHDG1_Exontr VaHDG1_Exontr Consensus Identity ChritOL secontr GonpeHDG1_exontr VaHDG1_Exontr Consensus Identity ChritOL secontr GonpeHDG1_exontr VaHDG1_Exontr Consensus Identity ChritOL secontr VaHDG1_Exontr VaHDG1_Exontr VaHDG1_Exontr VaHDG1_Exontr VaHDG1_Exontr VaHDG1_Exontr VaHDG1_Exontr VaHDG1_Exontr VaHDG1_Exontr VaHDG1_Exontr VaHDG1_Exontr VaHDG1_Exontr	AA AHWP OO PENS DIG AGEN LIGTORE LIGTORE LIGTORE AND A A A SPATLAWAAAAHAA LAA SAAD PEN PENENHHOUP BUOY DOOVER ABCAP VS OG PARA A BEM OMET HAD PENENHHOUP BUOY DOOVER GROREONLISEN VS AA ABM
Ginefity Ginefity Ginefity Ginefity Ginefity Ginefity Gonemus Gonemus Identity Gonemus Identity Consensus	AA A HWP OO PENS OG A GRP L ENG OM PLETP PPAAAAS PATLAWAAAAHAA LAASAAD PP PP PP PP PPH HHOUP EU O YO OG VEG BEOR EGMLS RUN O O TOS OO DE BETTO ON WEST AND STATE AWAAAAHAA LAASAAD PP PP PP PP PPH HHOUP EU O YO OG VEG BEOR EGMLS RUN O O TOS OO DE BETTO ON B
Identity Chrido Lasonatr VarHOG Lasonatr VarHOG Lasonatr VarHOG Lasonatr (ArHOL Lasonatr GengeHOG Lasonatr VarHOG Lasonatr GengeHOG Lasonatr VarHOG Lasonatr GengeHOG Lasonatr VarHOG Lasonatr Consensus Identity Chrido Lasonatr GengeHOG Lasonatr VarHOG Lasonatr	AA A HWP OO PENS DO GAGEP LETING CERP LETP PPAAAAS PATLAWAAAAHAA LAASAAD PPEP PEPEH HHOUP EU O YO OO OU CER BY DE PPEY PYEGGAAAAEMSE GEORE CHLSEN HY DA OT DIS OCTUBE EX JUGGUX QXX QX WF I SP VTAAP PG X GEREESDUS XWAX XXX X SAAVAPPP PALS PY MAP AQQ XX XQ X QX XX XQ XXX XQ PHEBGUL Q PQQ SWF I SP VTAAP PG LS GREESDUS XWAX XXX X SAAVAPPP PALS PY MAP AQQ XX XQ X QX XX XQ QX XX XQ PHEBGUL Q PQQ SWF I SP VTAAP PG LS GREESDUS XWAX XX XX SAAVAPPP PALS PY MAP AQQ XX QX QX XX QQ QQ QQ QQ QQ QQ QQ QQ Q
Identity Christol, seconstr VeaHDG1_Exonstr VeaHDG1_Exonstr VeaHDG1_Exonstr VeaHDG1_Exonstr Consensus Identity Christol, seconstr Consensus Identity Christol, seconstr Consensus Identity Christol, seconstr Consensus Identity Christol, seconstr VeaHDG1_Exonstr VeaHDG1_Exonstr VeaHDG1_Exonstr VeaHDG1_Exonstr VeaHDG1_Exonstr VeaHDG1_Exonstr Consensus Identity Christol, seconstr VeaHDG1_Exonstr Consensus Identity Christol, seconstr Consensus Identity Christol, seconstr VeaHDG1_Exonstr Consensus Identity Christol, seconstr Consensus Identity Christol, seconstr Consensus Identity Christol, seconstr Consensus Identity Christol, seconstr Consensus	AA AHWP OO PENS DO AGEN PLETER DE LETERE PERSAAAS PAT LAWAAAAHAA LAASAAD PEN PENENHHOUP BUOY DO OO VEG BERGEORIE SALESSA PETSEN YAASA PETSESSA PETS
Identity Chriefo J. asonatr VeaHDG1_ExonStr VeaHDG1_ExonStr VeaHDG1_ExonStr Consensus Identity Chriefo J. asonstr GonpeHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr Consensus Identity Chriefo J. asonstr GonpeHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr Consensus Identity Chriefo J. asonstr GonpeHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr VeaHDG1_exonstr	AA A HWP OO PENS DO GAGEP LETING CERP LETP PPAAAAS PATLAWAAAAHAA LAASAAD PPEP PEPEH HHOUP EU O YO OO OU CER BY DE PPEY PYEGGAAAAEMSE GEORE CHLSEN HY DA OT DIS OCTUBE EX JUGGUX QXX QX WF I SP VTAAP PG X GEREESDUS XWAX XXX X SAAVAPPP PALS PY MAP AQQ XX XQ X QX XX XQ XXX XQ PHEBGUL Q PQQ SWF I SP VTAAP PG LS GREESDUS XWAX XXX X SAAVAPPP PALS PY MAP AQQ XX XQ X QX XX XQ QX XX XQ PHEBGUL Q PQQ SWF I SP VTAAP PG LS GREESDUS XWAX XX XX SAAVAPPP PALS PY MAP AQQ XX QX QX XX QQ QQ QQ QQ QQ QQ QQ QQ Q
Identity ChritO1, savnatr GongeHGG1, savnatr VaHDG1, ExonStr VaHDG1, ExonStr GongeHGG1, savnatr VaHDG1, ExonStr Consensus Identity ChritO1, savnatr GongeHGG1, savnatr VaHDG1, ExonStr VaHDG1, ExonStr Consensus Identity ChritO1, savnatr GongeHGG1, savnatr VaHDG1, ExonStr Consensus Identity ChritO1, savnatr GongeHGG1, savnatr VaHDG1, ExonStr Consensus Identity ChritO1, savnatr GongeHGG1, savnatr Consensus Identity ChritO1, savnatr GongeHGG1, savnatr Consensus Identity ChritO1, savnatr GongeHGG1, savnatr Consensus Identity ChritO1, savnatr GongeHGG1, savnatr VaHDG1, ExonStr Consensus Identity ChritO1, savnatr GongeHGG1, savnatr VaHDG1, ExonStr VaHDG1, ExonStr VaHDG1, ExonStr	A A A HWY CO OPEN S DG AG RELE PLOYED PLATE PLATE PLATE PLANA AS PAT LAWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Consensus	Láno X SX X UX X X UX X GX GX X GX T P G E A Q GQ P G Q Q X X A A A X A X X X X X X X X X I UX X P L UX P R X X SPOLIFIX X WAX A X X FA A X X A X X X X X X X X X X
ldentity ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	GISS SERRQUR OG I GGGGGDMAAAVAHVVIR RUSGIDHTPLIPVESPDIFS SWAAAEAFAAACAAR PSP EULIRUAG GGGT DGRRTPGEA QGQDPGQQAVAAAAAA VQEVAXIBUSELUPRUD SPDIFANWAVA DEFAAQAASN SSON
Consensus	Ligio Ligio
ldentity ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	SGSGGGAAGAGADSAGGTALAAGATAGLPFGAIGSTAVUHSSTAAAATAAAUGAHTGR-VEVAAVGRASVDARGNSLG NTASDGGGGSGSG-GGCGATSGGAGAAAEWIRTPQLDSUVADWVRGTPPPPSULGSD0FVEVAAVGRQPAP- 1.590 1.000 1.00 1.00 1.00 1.00 1.00 1.00 1
Consensus	S L F R S S G G G Ś G S N G G S G G G G G G G G G G G G G G
ldentity ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	SLFRSSGGGSGSGGGGGGGGGGGGGGGGGGGGGGGGGGGG
Consensus	XPXXXXXXIXXXXGXGXXXXXXXXXXXXXXXXXXXGXXGXXG
Identity ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	P BLQUQADITAGTAGGSGGSGAWEPHGISPATAGLDAAGHRGGVAAPAYAAQQTAWAAAPAYPEBADTDPLVAAATATA SRQDWHPGTUPKRAGAGPASVUELAAAGAIGDRNAGPGSSGAUTVATADAGARPWANGLSSLBTSAAAAQYATEATAAT uga uja uja uja uja uja uja uja uja uja uj
Consensus	XŘGGHGAMAXXXMEXEMPRXXXMAXXXMAXXXMAXXXXATELXALLXALLXALLXXXXDRXMAXXXXLEELXXITXXRXMXXXXXXXX
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	A AAAP PD OLP PWL SATD GATA SAAGYD SATUAAUTAU FEHSMD PAAAATA LEEUSS LVP SAVAAAAAA VAA GRGGGAAVTALU LUP PAGGAAAAAAAYAAYAAS SALATUE AU OTUM TNID PGAAASALUEEU QYITAO PAAAYGGGGGGGGGG
Identity	
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	A & SAVVPAAAAGPGPAAGQPETLAAPMPPAWQAPPAAMIGOHUQORARSOS GAGGISTIZGHUHHOHHYOHHIGAPPPIPHAPPAAPP GSGGSDELWIRNPEDAAADWEANVAAAGAAAARDZAESDVGSTMERQOTPIPOALPPP
Consensus	
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	IHVGTATHPHP2LPHHQFHHQHHHQHGHHDSAHDRPLPPPBQQQTLQPPQQQHQQQHPPPPPBQQQQCQPPHPHPAYHPH
Identity	
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	ВНФИ/ ЮНО 1984 ЖААААЕ GO P Y HL P V P ЖРА G ШЕ Q P H T H H H Y P Q P H P Q H G Q B Q H OH OH Q H Q H O H OH OH OH H H OQ V P WAA A S Q P R DE H
Identity	
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	GAABUP PAHHISIPHEDOOD212121121121200000000000000000000000
Identity	
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	DDGGTGGGGGB OYGYQHHQHHQHPXTASATDAAAALWAXYAAWATAAAAAXDGGAGGGAGAGGGGGATYSSAAEAQAKAAA AAAAAAAAAQQWAEGLSDSSPLHRASSWGSMVVQVPQSSXAAAAAGDGRQRHTAYLAAPRADAPRADAQ QTRTYAGGTQMA QQELGGIGGGMHQLE-PXAAWAVAERQPSGAQQAVTLBFSLDAADAMA 2,200 2
Identity	
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	Υ ΕΨΟ Q Υ Q L Q LEG YH Q Q Q Q Y E Q Y Q Q Q Q Q Q Y E L Y Q Q Q Q Q Q Q Y D X Y G Y MA A T XA A T EH P P R E G D HQ P JP P D -
Identity ChrHD1_exonstr GonpeHDG1_exonstr	GVERVISSA
VcaHDG1_ExonStr Consensus	Land State
ldentity ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	I AGNL SGGGATAGAAGP GGGR P GAWNWI SLOGRVAAGALAYGI GAGGG SISINWI SLOGRVAAGALAYGI GGGGG SONWI TLOGRVAAGALAYGI GGGGGGSONWI TLOGRVAAGALAYGI GGGGGGSONWI TU GGRVAAGALAYGI GGGGGGGSONWI TU GGRVAAGALAYGI SVAAP SEMINGRGL (I KSSS
Consensus	24%0 24%0 24%0 24%0 25%0 25%0 25%0 25%0 25%0 25%0 25%0 25
Identity ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	สุดสุสุสุด สุดสุสุสุด สุดสุสุสุด สุดสุสุลสุด สุดสุสุลสุด สุดสุสุดสุ
Consensus	
Identity ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	NAVTALIS (WYVAHLIVH PYPSIEDEKEVIC GAHTGLDIL CUNNWEINAR VEIWKET FIOVENSI- NH PRMAA EAGARGDGELL TAVNISTRAWYVAHLIVH PYPSIEDEKEVICAN SIGTDIL CUNNWEINAR VEIWKET FIOVENSI- NH PRMAA EAGARGDGELL AAVOSTRAWYVAHLIVH PYPSIEDEKEVICAN BURGLOUINNWEINAR VEIWKET FIDIS VENSI- NH PRMAA EAGARGDITELL 2000 2000 2000 2000 2000 2000 2000 20
Identity	
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	TIKM CAARNS PTAOLALVAADASAN ON NE CAAQWASE DHAADACAADA GGA KRGSGGTGAKR SKASGA LAGAG PAGGELGS GKAMAARNS PTAOLALVAADATATTCICHE CAARAAAE ELEMMAAAKATTAGGG- AKAMAARNS PTAOLALVAADATSCHE CHE CAARAAABKASIEVUE SARRAAAGKSGK GSISSRAYSMP WRGRRRR GRGRE ELGI 2000 2,200 2,
Identity	
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr Consensus	AA EA ANA BA GE CARR'S BRACEGOEGEGES RS RG TIGGGLGE QD QY QQ QY QQ HQ HQ QQ HQ HQ QQ QG AD F GAA E F LGD GGAALA E YMA - GG SK KE CGG SS SR RR GEGGGG WR GR RR RR PW RF GED V V GR GGLLSP C 2/0 2/0 2/0 2/0 2/0 2/0 2/0 2/0
Identity	
ChrHD1_exonstr GonpeHDG1_exonstr VcaHDG1_ExonStr	ΤΥ GHLAGD FAAAMAS FAAAGAG VMD PHAAAAAAA VAAAAGAA GA GAGGA GAGGAGGA YGDA G TDD P II G Y X - GGA GAC GGCGGG GS GBS V SME G X

Fig. S7. Alignment of Mam-A TALE proteins in mamiellophyceae. Short motifs (Box1-4) are conserved among all members in this class over the entire length of the sequence. Red reverse triangle at 548-549 shows the truncation position of Micco_Mam-A-tr used in Yeast-two-hybrid analysis.

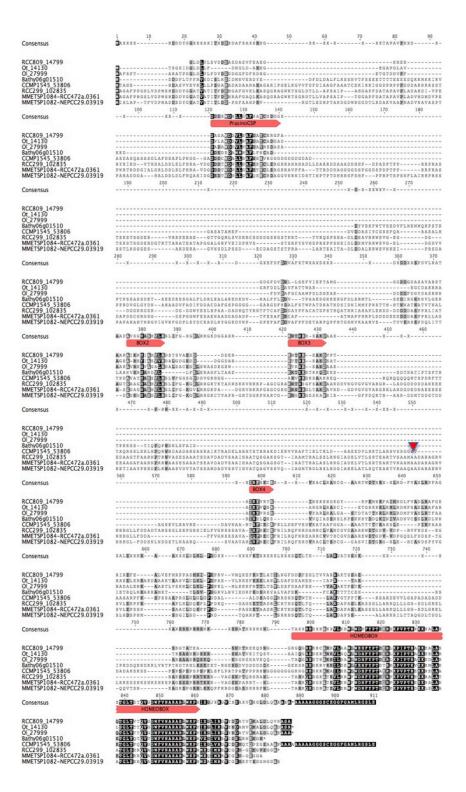
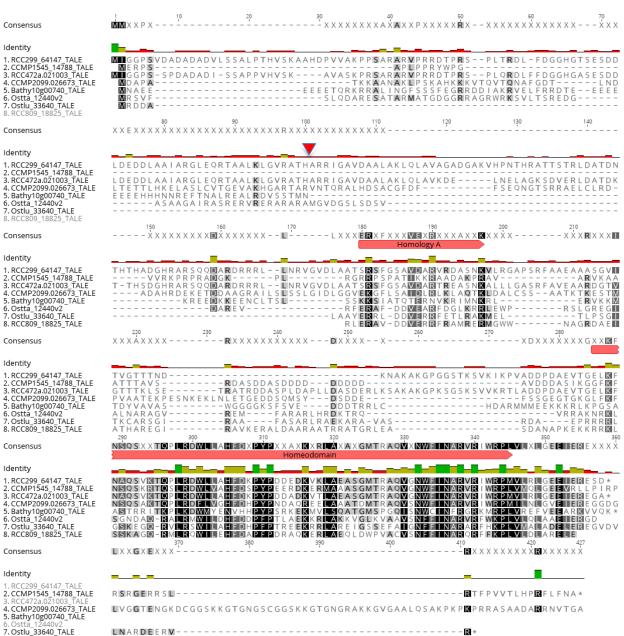


Fig. S8. Alignment of Mam-B TALE proteins in mamiellophyceae. A conserved motif is found between 180-197 amino acids in the alignment. Red reverse triangle at 100-101 shows the truncation position of Ostta Mam-B-tr used in Yeast-two-hybrid analysis. Homology is restricted to a single homology-A domain ouside the homeodomain.



Consensus

Consensus Identity

Identity 1. RCC299_64147_TALE 2. CCMP1545_14788_TALE 3. RCC472a.021003_TALE 4. CCMP2099.026673_TALE 5. Bathy10g00740_TALE 6. Ostta_12440v2 7. Ostlu_33640_TALE 8. RCC800_18825_TALE

Consensus

Identity

1. RCC299_64147_TALE 2. CCMP1545_14788_TALE 3. RCC472a.021003_TALE 4. CCMP2099.026673_TALE 5. Bathy10g00740_TALE 6. Ostta_12440V2 7. Osttu_33640_TALE 8. RCC809_18825_TALE

Consensus

Identity

1. RCC299_64147_TALE 2. CCMP1545_14788_TALE 3. RCC472a.021003_TALE 4. CCMP2099.026673_TALE 4. CCMP2099.026073_17 5. Bathy10g00740_TALE 6. Ostta_12440v2 7. Ostlu_33640_TALE 8. RCC809_18825_TALE

Consensus

Identity

1. RCC299 64147 TALE
2. CCMP1545_14788_TALE
3. RCC472a.021003_TALE
 4. CCMP2099.026673_TALE
Bathy10g00740_TALE
6. Ostta_12440v2
7. Ostlu_33640_TALE
8 RCC800 18825 TALE

Consensus

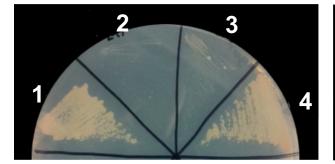
Identity

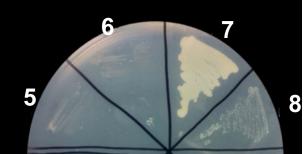
2. CCMP1545_14788_TALE
3. RCC472a.021003_TALE
 4. CCMP2099.026673_TALE
5. Bathy10g00740_TALE

7	Octlui	33640	TALL	-

8. RCC809_18825_TALE

Fig. S9. Full-length proteins are necessary for mamiellophyceae non-KNOX TALE proteins to form heterodimers. Left and Right: Yeast-two-hybrid assays on Ade-/His-/Leu-/Trp- medium. The construct information for the prey conjugated with the GAL4 DNA-binding domain and for the bait conjugated with the GAL4 transcriptional activation domain is given in the table below.

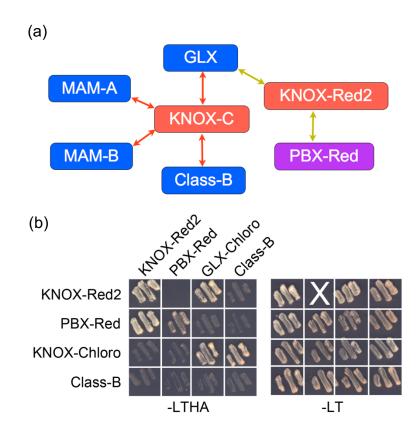




Sector	DBD	AD	Sector	DBD	AD
1	Ostta_KNOX	Ostta_Mam-A	5	Micco_Mam-B	T-Ag
2	Ostta_KNOX	Ostta_Mam-B-tr	6	Micco_Mam-A-tr	Micco_KNOX
3	Ostta_Mam-A	Ostta_Mam-B-tr	7	Micco_Mam-B	Micco_KNOX
4	Ostta_Mam-A	Ostta_KNOX	8	Micco_Mam-B	Micco_Mam-A-tr

Fig. S10. TALE interaction network defined by this study using yeast-two-hybrid assays.

(A) Summary diagram for the TALE interaction network. (B) Yeast-two-hybrid assays for the cross-species interaction of TALE proteins. Only one of the possible reciprocal combinations of GAL4 domain conjugations is provided for simplicity. Large X indicates no yeast in the sector. - LTHA: Leu-/Trp-/His-/Ade- medium; -LT: Leu-/Trp- medium.



Consensus	
ldentity Monbr_27982_MEIS EGD73338_SalpMEIS HsMEIS2	MISDLLPNUAPDLATFLNPNLEVALDSIKDDPTMNWPPATGPASVRTSMAHERYYSFNTRUDWPVGLNGOPATUCANISLISTP MI
Consensus Identity	
Monbr_27982_MEIS EGD73338_SalpMEIS HsMEIS2	DFLAUNCSTMSEGPAMPLPPKS-QADPTGSSLTPADLPEALLTDALLSMNPTIDRDSGNSSGSGSGSGSGSGSAPDP NTSTNFFDTTFDAPFTSEANWHLKQ-QGUQTMRPFTLDNDGDGFENACAPQFTPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
Consensus Identity	240 240 240 240 240 240 240 240 240 240 240
Monbr_27982_MEIS EGD73338_SalpMEIS HsMEIS2	DUQADUN PNESPDLK PASISTNDSFSAPSSSASH NASSUSAADU SAT DEAULAAAQAQADTNLLSSEEUQTAAQHAD RUQSIMDASERPKNIIRTASDGTDRPQQAAGMCGTRVTPKTHEMCVAIQDETRTA-TATLPIDDTHPHNN PUFPLUALVEKCELATCTDREPGVAGGDVCSSDSFNEDUAVFAKQV-RAEKPUFSSNPELD MEIS
Consensus Identity	240 240 240 240 240 240 240 240 240 240
Monbr_27982_MEIS EGD73338_SalpMEIS HsMEIS2	GFAFELEDFVVSLNDWADTMMNATKAETSPAPA-TTSPVTSAAITPTNTSTLTPALAISGQAQPPAAAPAVT SATSTILTGPTAPTSC-AVCNRGGAVGTTEVACSCGPSARTMHSGARRNSVSDUKD
Consensus Identity	300 300 400 EXDYSKGXXXXXXXRSGXGXXXXXXXXXXXXXXXXXXXXXXXXX
Monbr_27982_MEIS EGD73338_SalpMEIS HsMEIS2	GSSKSGHELLSGSSTNUAOHNPSSWRDHOBAUSTHSAGTPGSVARSISUN-QSLPPSEPGWHRAADLQSSD EKDYSKGEKDGSKDGRGGGGRAUAUKKIKTGHGNSGRRTDGIOTAGHSAAMTRQQQQQQQQQNNKKKNRKKK ERDGSSKSDHELLSGSSTNUAOHNPSSWRDHODAUSTHSAGTPG
Consensus Identity	410 410 400 400 400 400 400 400 400 400
Monbr_27982_MEIS EGD73338_SalpMEIS HsMEIS2	SCLISDSDTE GAPGAGSAHAARSPORKRPAHSTHTIGTDIST KRRSQKRSKE ADNATTAAATRATQGRDTVATPPVNGDCDTTSGGDTAGATEQGHVARDGGTKLERQQRIVRPAADVVARA PSSGGHASQSGDNSSEQGDGLDNSVASPGTGDDDD
Consensus Identity	
Monbr_27982_MEIS EGD73338_SalpMEIS HsMEIS2	RHCTKREASSSIDTIKEWIEAHTDRPYPTDQDKTEIMQQTGLDLMQINNWEINARRELLVKWN
Consensus Identity	sio XKITSXISXXIXXVMY2XXXXXXXIQRXXXIQRXXIQRXXIGAXXXXXGMNMGMDXXXXXIX
Monbr_27982_MEIS EGD73338_SalpMEIS HsMEIS2	NKITISTI BIRPSSPTAFCILLE LIDFARVRURLHNAVITFCDQLSRCF* VKITISTI HNR GDVM YPATAGATMRRORCT THHWRGAPVL

Fig. S11. Identification of MEIS homologs in choanoflagellates.

Consensus Identity	¹ X T A T N S D S G G D V A Q V L L Å F K K K Q H D C E ² T G A G S P S S P ^M I E V X X X X X M ¹ S E Q X X G X X M ² X X X X X X X X X X X X X X X X X X X
Monbr_4837_PBX EGD73507_SalpiPBX HsPBX1	MATATNSDSGGDDVAQVLLAFKKKQHDCETRGAGSPSSPMIEVDDGKISMHISEQDSGSEVASGSSSSSSSSSSSSSSS MIS
Consensus Identity	100 110 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Monbr_4837_PBX EGD73507_SalpiPBX HsPBX1	SVTSAERLDADCPSIHGSRUHRRRRPHGNGKHKANSMSGVNNDPTSUNAAVTNDVAGK QPPPPPMEDDAQQPPOKMAKVTPOQETISTHTEDPLVRLISUPVRUAPS EGEGGRKQDIGDILQQIMTITDOSLDEAQARKEIAUNCHRMKDALFNVLCEIKEKTVISIRGAQEEEPTDPQLMR PBCA
Consensus Identity	$\overset{140}{\longrightarrow} \times \times \times \overset{170}{\boxtimes} \times \times \times \times \overset{170}{\boxtimes} \times \times \times \times \overset{170}{\boxtimes} \times \times \times \times \overset{170}{\boxtimes} \times \times \times \times \overset{170}{\times} \times \times \times \times \times \times \times \times \times \times \times \times \times \times \times \times \times \times$
Monbr_4837_PBX EGD73507_SalpiPBX HsPBX1	IGK SGTLNPGASLKQIRSSDNQGLQPTAPPQKAEAIUAQNDSDDTPDRSDASSNSGNSGDASSSTNS DSDSTUPTSASGPTPPHSQPQQAPPHPSSQQHPQQQQQQHNNQHPPQLPHQHHPQPTPPQPUQQQQQQQQ UDNLUAEGVAGPEKGGGSAAAAAAAASGGGGGSDNSVEHSDYRAKLSQIRQTYHTBLEKYEQA PBCA
Consensus Identity	2200 2200 2200 2200 200 200 200 200 200
Monbr_4837_PBX EGD73507_SalpiPBX HsPBX1	SKP STEDVASRVSAAGLSRAPASASDRTEKAEL-GQULDKDDDHASMDTAOTQSTRRSNTGGR QQPQQLSAAQUQAHQPHHPDTSPKHVPPOMHA-QAMAISTDGHGASDRRK
Consensus Identity	310 XIXXIIXXIIXXIIXEWEEXXXXIIIIXXXXXXXXXXXX
Monbr_4837_PBX EGD73507_SalpiPBX HsPBX1	NNMPHEVTISRUK EWFFALTISH PYPSEQKKRELASOCOLTLOQINNWEINARRRUNNPPKRS
Consensus Identity	²⁹⁰ XIX X X L AQVX X X X X X X X X X X X X X X X X X X
Monbr_4837_PBX EGD73507_SalpiPBX HsPBX1	- NG TA LIR QVTRINAAS R VE TSE SSE SSG PRAKHVKRGE EGLQELDSSE ASE GNDHSEKP
Consensus Identity	S S A QWP S G Q P V S V V G V L P G X X X S S A G D X S X X X X X X X X X X X X X X X X X
Monbr_4837_PBX EGD73507_SalpiPBX HsPBX1	– SSSE ESSMIND YAR QMR LUEQEV – VALRS UN QQQEQIIIR QQEIUSHLQQQEQER SSAQWPSGQPVSVVGVLPGTSDASSGDASILAATAPATATIVQSQUPUDARLVAVHDDDAATRQQQOHEQ-QQHEQEQ - SAQWPSGQPVSVVGVLPGTSDASSGDLFMSIVQSUNGDS-YQGAQVGANVQSQVDTURH
Consensus Identity	540 XXRXXXXXXXZZXXQZXXQQQXXXXXXJXXPXXXXXXXXXXXX
Monbr_4837_PBX EGD73507_SalpiPBX HsPBX1	LMRERASMMMHQEMLQQQNRYMGSIPKPNVPSP-GMPG-GPMVAVIVPNGVGSARLRPGQASYLGGPTLPPVSMGA RQRSQQDVEEQQQQQQQQQQQQQPAVVSVQDDAQSRLGQQLHQHPLLDATRLPTSSSHA VSVQDAQSRLGQQLAASQMYSPQGISANGGWQDAUTPSSVTSPTEGPGS
Consensus Identity	X X X X X X Z Z X X X X X X X X X X X X
Monbr_4837_PBX EGD73507_SalpiPBX HsPBX1	DPTGLKMIGSMINN MISYW PIMGMQSWPLVPQAPQGGKIFGPWGSSESHSGSDGGRSUVQAGMHMIRSG YTVAPSUUHQHQQLPTIAVPIQQQQQQQQQQQQQQQQQQQQQQHQQQHHQQQQHVWVLASGHGGGGGGVVWSSAPEHLVPV
Consensus Identity	
Monbr_4837_PBX EGD73507_SalpiPBX HsPBX1	SESDARARTSMSPDSGTN* SASFVDAPSSSVTTPLTPPSRATNINMHPQQQ VHSDTSN

Fig. S12. Identification of PBX homologs in choanoflagellates.

Fig. S13. Alignment of the MEIS homologs in Amorphea.

1. Capowc_EFW44798_TALE 2. Thetra_KNC56447_TALE 3. Monbr_27982_MEIS 4. Salros_EGD73338_MEIS	1 DTESDGTPSKRARL NGSAGGPGNEQAP SSGSGGLSSARSD DNDGDGFEENACAF	20 DNGWSÅTSAG QQVP DLQADLNPNFSPD QFTP	40 NNNNNSSNNNN LSPTHTSSPTSA LKPAISITNPSFSAPSS ISSSSSSSSSS	50 NNNN NN PLPPPGM SSNAPLSALNMVSGEA SASH HNASSLSAAD SSSS SSASTINRHRU 120	70 SAYPPPPHAMQ AS SA TDE QS IMD
1. Capowc_EFW44798_TALE 2. Thetra_KNC56447_TALE 3. Monbr_27982_MEIS 4. Salros_EGD73338_MEIS	A P P Q P M Q M Ă Y A S N A L L A A A Q A Q A D T N L A S F R P K N I R T A S P	Y S G H P Y Y M P S P Y P I L S S E E L Q T A A Q H A I G T D R P Q Q A A G M C G	D Q S M P Y N Y S Y G M P V S Q 	P V P P V H H T H A E Y K P T H A 	SANSVA AAATA TTSETKVTTTA TTSPVTSAAIT TTPPNNSAIS
3. Monbr_27982_MEIS 4. Salros EGD73338 MEIS	- P T N T S T L T P A L A S T S L R G P T A P T S C A N	5 G Q A Q P P A A A P A ` / C N R G G A V G T T E V A (V TT A G P S R R S S H C S C G P S A R T M H S G A R R I	190 ER VI I EK P A S S E EH M GDM A P VN ES Y S V A G S S D I K D E K D - Y S K G E 250 220 220 220 220 220 220 220 220 220	K D G S K D G R G G G
1. Capowc_EFW44798_TALE 2. Thetra_KNC56447_TALE 3. Monbr_27982_MEIS 4. Salros_EGD73338_MEIS	NVQHPPPQSGQQQ DRARAAKRSKLSS GKQMPLPKAAAPP GRALALKKHKTGH(2 P SQ A LHQNG A AVTRQT TTRRLTQPGSGVN NSGRRTD	- SEAPVTPNTLPHAST - AGVDSGSDSDYV E LTGPSAAPSISLN-QS - GEQTAGHSAAMTRQQ - 320	270 F L D S P R G A G K S P R E D G P G K K Q D G G N G K R K R S S A S L P P S E P G W H R A A D L Q S S Q Q Q Q Q Q N N K K K N R K K K 330	NPDPSNG TKRGSNG DSCLSDSDTEG KKRRSQKRSKE
2. Thetra_KNC56447_TALE 3. Monbr_27982_MEIS 4. Salros EGD73338 MEIS	APGAGSAHAARSPO ADNATTAAATRATO	R K R P A H S T H T T G T I G R D T V A T P P V N G D	G G A A A K S R	330 K D G Q H T V L T R G R ŠK Q R P R WG G NR A S S R R P K SR H C T H V A R D G G T K L E R Q R T N 400	R L F P P Q A K A V F K F A S S S I D T L K R F P R R S T K V L
1. Capowc_EFW44798_TALE 2. Thetra_KNC56447_TALE 3. Monbr_27982_MEIS 4. Salros_EGD73338_MEIS	K EWL F AHT DR P Y P I	DQDKTELMQQTGL EEEKQELRAASGL	DLMQINNWFINARRL	400 L Q K P D P L L VNR N F K R C R R R Y S T L V K V N	- K P S R N K T S T S
1. Capowc_EFW44798_TALE 2. Thetra_KNC56447_TALE 3. Monbr_27982_MEIS 4. Salros_EGD73338_MEIS	SORRTRSPAEPATA RPSSPTAFCIE	I.			

Fig. S14. Alignment of the PBX homologs in Amorphea.

	1 10	20 30	40 50	60
1. Capowc_EFW47516_TALE	T P T N R V G F A S H D V T A A Q L	DVC AQERAACAASETPS	SLAEIRGLFAFLDMQLLK	VARRL
2. Spharc_KNC86410_TALE		H – – Q 🛛 L E L L A D G P E R N A R [
Salros_EGD73507_PBX	P Q P L Q Q Q Q P Q Q Q Q Q P Q Q P Q			
4. Monbr_4837_PBX		SAAG LSR AP ASA S D RTEK -		
5. HsPBX1	HTELEKYEQACNEFTTHV	// – – N 🗖 L 🖪 E Q 🕏 R T – R P I S P -	K E I E RMV S	 R K – F
	70 8		100 110	120
	SSLEYQLKRKVYTKVIFLK			
Spharc_KNC86410_TALE	DSIDEALQKHC			
Salros_EGD73507_PBX	MAESTD			
4. Monbr_4837_PBX	ASMDTAQTC	QSTRRSNTGG RNNM PHE V T	SRLKEWFFAHTSHPYPS	
5. HsPBX1	SSIQMQLKQSTCEAVMILF	RSRFLDARRKRRNFNKQAT	EILNEYFYSHLSNPYPS	EEAKEEL
	130 140	150 152		
	ADQTNLTMRQISTWFANKE			
2. Spharc_KNC86410_TALE	AEECQLEIQQIRNWFANTE			
3. Salros_EGD73507_PBX	ATKENLSMTQVNNWF I NAF			
4. Monbr_4837_PBX	ASQCDLTLQQINNWFINA			
5. HsPBX1	AKKCGITVSQVSNWFGNKF	K I K Y K K N		

Fig. S15. ELK-domain alignment.

Consensus Identity	1 10 20 25 E X X J R X X L X R K Y X X X J X X L X X - E F X
1. CCMP720_035111_KNOX-R1_site_1 2. Ostlu_5313_KNOX_site_1 3. Ostta_04g04290_KNOX_site_1 4. RCC809_56605_KNOX_site_1 5. AT4G08150_BP_site_1 6. CCMP720_094156_KNOX-R1_site_1 7. CCMP720_03114_KNOX-R1_site_1 8. AT1G62360_STM_site_1 9. Picsa_04387_KNOX_site_1 10. RCC299_62285_KNOX_site_1 11. Cyagl_026930_KNOX_site_1 12. CCMP726_01361425_KNOX_site_1 13. CCMP726_01361425_KNOX_site_1 14. CCMP1545_69285_KNOX_site_1 15. CCMP726_013611425_KNOX_site_1 15. CCMP726_013611425_KNOX_site_1 16. Bathy08g03820_KNOX_site_1 17. Cyapa_20927.63_PBC 18. Cyapa_20927.64_PBC 19. Glowi_03052_PBC	E E S K E E E RMRYANTIME E E - DIM E E R E A LKRKYASSIMTIKS - E FM E E R R Q D K K Y ASSIMTIKT - E FM E E R R Q D K K Y ASSITMIKT - E FM D R E K N H L K Y S G Y S S K Q - E L S E A T K D E C K Q Y AN T R E R K - E I L D R E K G Q L R K Y S G Y G S K Q - E FM D NM R Q Q L K K Y S G Y G S K Q - E F D NM R Q Q L K K Y S G Y G S K Q - E F D NM R Q Q L K K Y S G Y G S K Q - E F D NM R Q Q L K K Y S G Y G S K Q - E F D NM R Q Q L K K Y S G Y G S K Q - E F D D R K S K K K Y A S S L A K D - E F E D D R K S K K K Y A S S L A K D - E F E D D R K S K K K Y A S S L A K D - E F E D D R K S K K K Y A S S L A K D - E F E Q T I R S H K R F A N T K V V Y K N - E F K E D D R K S K K K Y A T S S S R D - E F L L V N Q Q F K R K Y S N S Q T L K Q - D F L D E A F R N Q L A K Y K D D P A L E E - E W L F K T V R A E L E R K Y L E G K V Q K P K K V F R A V R L N R R K Y L E G G L F N R S G K A