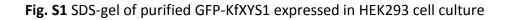
New Phytologist Supporting Information

Article title: Identification of an algal xylan synthase KfXYS1 proves that there is functional orthology between algal and plant cell wall biosynthesis

Authors: Jacob Krüger Jensen, Marta Busse-Wicher, Christian Peter Poulsen, Jonatan Ulrik Fangel, Peter James Smith, Jeong-Yeh Yang, Maria-Jesus Peña, Malene Hessellund Dinesen, Helle Juel Martens, Michael Melkonian, Gane Ka-Shu Wong, Kelley W. Moremen, Curtis Gene Wilkerson, Henrik Vibe Scheller, Paul Dupree, Peter Ulvskov, Breeanna Rae Urbanowicz and Jesper Harholt

Article acceptance date: 14 January 2018

The following Supporting Information is available for this article:



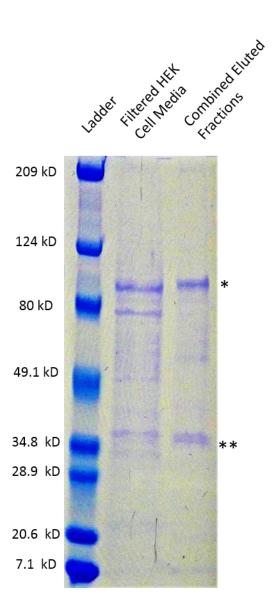


Fig. S1 Purification of recombinant GFP-KfXYS1. Lane 1, Broad Range Prestained SDS-PAGE Standard (Bio-Rad); Lane 2, HEK293 cell culture medium; Lane 3, pooled eluted protein. The gel was stained with coomassie brilliant blue R-250 (Bio-Rad). *KfXYS1(MW 85.46kD) and **His-Avi-GFP (MW 33.2kD).

Fig. S2 Phylogenetic tree of PARVUS orthologs.

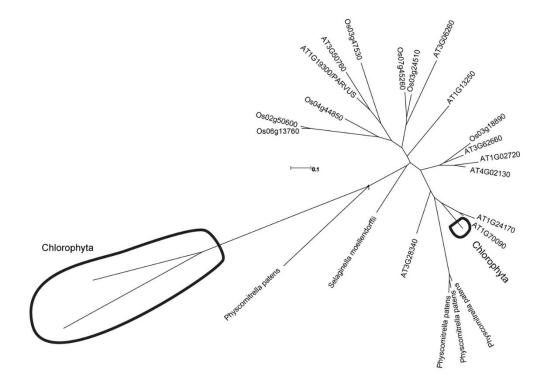


Fig. S2. Phylogenetic tree of PARVUS orthologs. Sequences included from chlorophytes, *Physcomitrella patens, Selaginella mollendorffii,* rice and Arabidopsis. The two encirclings indicate sequences from chlorophytes (black). Branches with no name correspond to chlorophyte sequences from 1kp dataset and are probably not full length. All branches with names are sequences from complete genome projects. Branch support was evaluated with Approximate Likelihood-Ratio Test and is given as a value between 0 and 1. Supporting Information Notes S1 holds tree files where all sequences used are labeled.

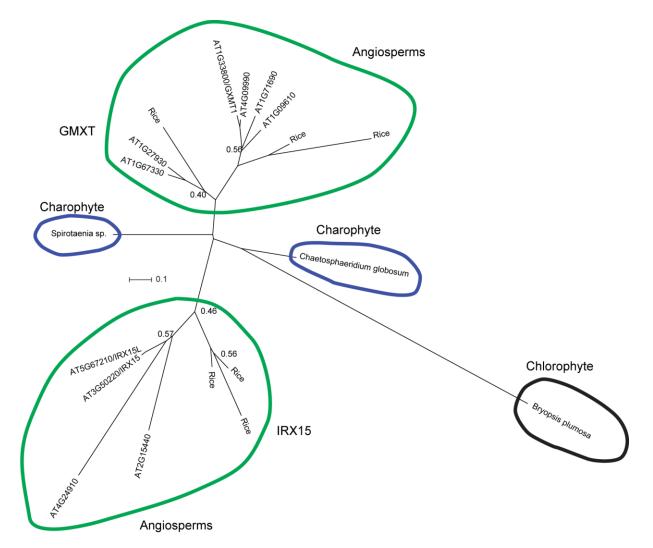


Fig. S3 Phylogenetic tree of IRX15/GXMT1 orthologs.

Fig. S3. Phylogenetic tree of IRX15/GXMT1 orthologs. Sequences included from chlorophyte and charophyte green algae, rice and Arabidopsis. Encircling indicate sequences from angiosperms (green), chlorophytes (black) and charophytes (dark blue). For angiosperms, the IRX15 and GXMT1 orthologs group in two separate groups, respectively. Chlorophyte and charophyte branches correspond to sequences from 1kp dataset and are probably not full length. Branch support was evaluated with Approximate Likelihood-Ratio Test and is given as a value between 0 and 1. Supporting Information Notes S1 holds tree files where all sequences used are labeled.

Fig. S4 Phylogenetic tree of IRX9 and IRX14 orthologs.

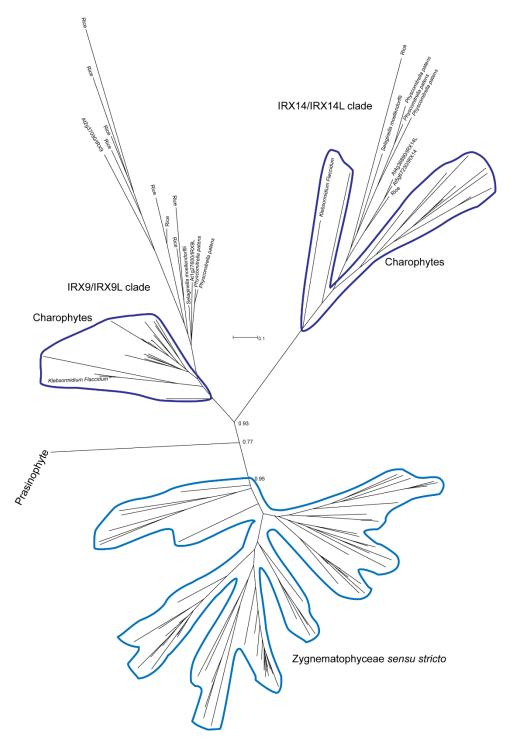


Fig. S4. Phylogenetic tree of IRX9 and IRX14 orthologs. Sequences included from Ostreococcus lucimarinus (prasinophyte), charophytes, Klebsormidium flaccidum, Physcomitrella patens,

Selaginella moellendorffii, rice and Arabidopsis. Encircling indicate charophytes (dark blue) and Zygnematophyceae sensu stricto (light blue). Branches with no name correspond to charophyte sequences from 1kp dataset and are probably not full length. All branches with names are sequences from complete genome projects. Branch support was evaluated with Approximate Likelihood-Ratio Test and is given as a value between 0 and 1. Supporting Information Notes S1 holds tree files where all sequences used are labeled. Fig. S5 Phylogenetic tree of IRX8 orthologs.

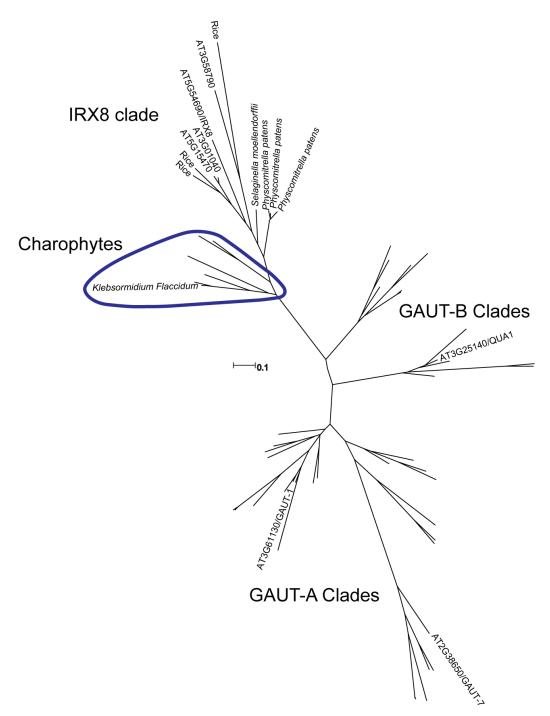


Fig. S5. Phylogenetic tree of IRX8 orthologs. Sequences included from charophytes, *Klebsormidium flaccidum, Physcomitrella patens, Selaginella moellendorffii*, rice and Arabidopsis. Encircling indicates sequences from a group of several species of charophytes (dark blue). Branches in encircling with no name correspond to charophyte sequences from 1kp dataset and are probably not full length. Branches in GAUT-A and GAUT-B clades are full length sequences from *Physcomitrella patens, Selaginella moellendorffii*, rice and Arabidopsis. Branch support was evaluated with Approximate Likelihood-Ratio Test and is given as a value between 0 and 1. Supporting Information Notes S1 holds tree files where all sequences used are labeled.

Fig. S6 Phylogenetic tree of GUX1-5 orthologs.

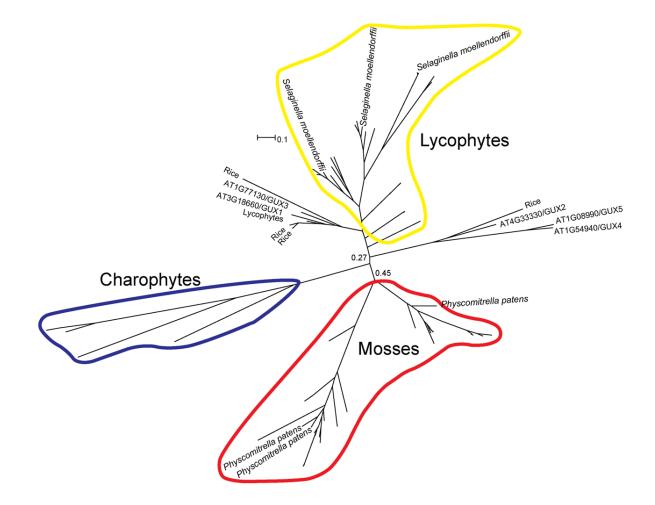


Fig. S6. Phylogenetic tree of GUX1-5 orthologs. Sequences included from charophytes, *Klebsormidium flaccidum, Physcomitrella patens, Selaginella moellendorffii,* rice and Arabidopsis. Encircling indicate sequences from Charophytes (dark blue), mosses (red) and lycophytes (yellow). Branches with no name correspond to charophyte, moss and lycophyte sequences from 1kp dataset and are probably not full length. All branches with names are sequences from complete genome projects. Branch support was evaluated with Approximate Likelihood-Ratio Test and is given as a value between 0 and 1. Supporting Information Notes S1 holds tree files where all sequences used are labeled.

Fig. S7 Phylogenetic tree of ESK1 orthologs.

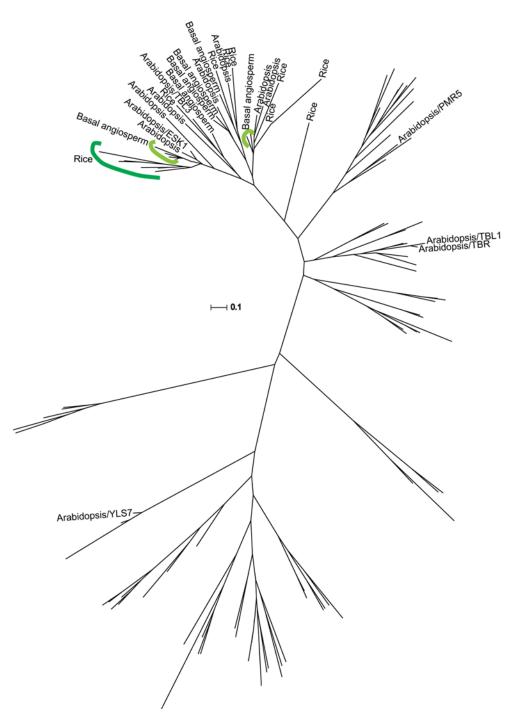


Fig. S7. Phylogenetic tree of ESK1 orthologs. Sequences included from basal angiosperms, rice, Arabidopsis and from the 1kp dataset. Encircling indicates sequences from rice (dark green) and basal angiosperms (light green). Branches in the TBL3 clade named basal angiosperm correspond to basal angiosperm sequences from 1kp dataset and are probably not full length. Branches with no name are either rice or Arabidopsis full length sequences. Branch support was evaluated with Approximate Likelihood-Ratio Test and is given as a value between 0 and 1. Supporting Information Notes S1 holds tree files where all sequences used are labeled. Fig. S8 Phylogenetic tree of GT61 orthologs.

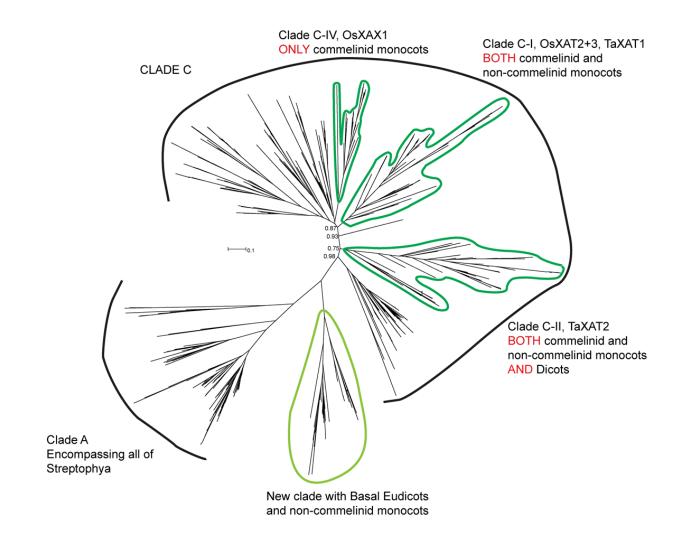


Fig. S8. Phylogenetic tree of GT61 orthologs. Sequences included from *Physcomitrella patens*, *Selaginella moellendorffii*, rice, Arabidopsis and angiosperms from the 1kp dataset. Clade B, encompassing AtXYLT, which is involved in N-glycan biosynthesis, is not included in the tree. Three subclades in the Clade C are individually encircled (green) and each hold at least one member of a proven xylan arabinosyltransferases (XA) from either rice or wheat. As indicated, these three clades contain commelinid monocots only (Clade C-IV), non-commelinid and commelinid monocots (Clade C-I) and non-commelinid and commelinid monocots as well as dicots (Clade C-II). A new clade with basal eudicots and non-commelinid monocots were also identified (light green encircling). Branch support was evaluated with Approximate Likelihood-Ratio Test and is given as a value between 0 and 1. Supporting Information Notes S1 holds tree files where all sequences used are labeled. Fig. S9 Phylogenetic tree of OsAT10 orthologs.

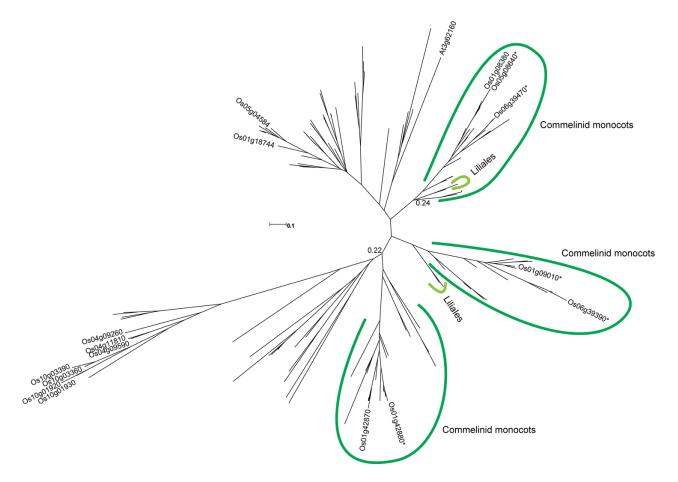


Fig. S9. Phylogenetic tree of OsAT10 orthologs. Sequences included from rice, Arabidopsis and monocots from the 1kp dataset. Rice OsAT10 is a putative cell wall related acyltransferase (BAHD) involved in ferulic or coumaric acid substitution of xylan. Sequences marked with asterisk are putative xylan ferulic or coumaric acid transferases. Encircling indicates sequences from a group of several species of commelinid monocots (dark green) and of lilioid monocots (light green). Branches with no name correspond to monocot sequences from 1kp dataset and are probably not full length. All branches with names are sequences from complete genome projects. Branch support was evaluated with Approximate Likelihood-Ratio Test and is given as a value between 0 and 1. Supporting Information Notes S1 holds tree files where all sequences

used are labeled.

Fig. S10 Enzymatic specificity of KfXYS1.

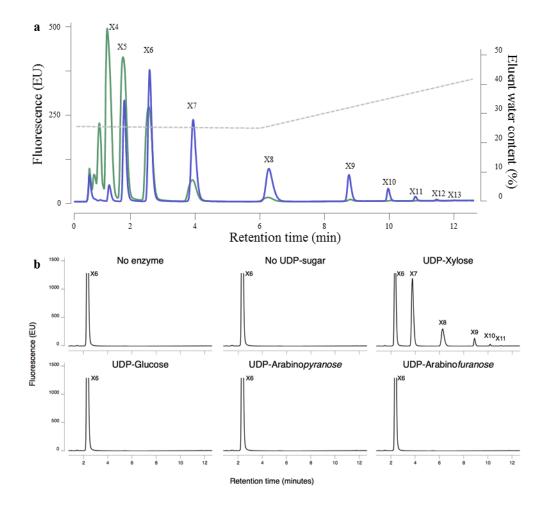


Fig. S10 Enzymatic specificity of KfXYS1. (a) KfXYS1 enzymatic product analysis by use of 1,4-βxylanase. KfXYS1 protein was heterologously expressed in Saccharomyces cerevisiae. Reaction mixtures with UDP-xylose, fluorescently labeled xylotetraose (X4) and KfXYS1 preparation before (blue) and after xylanase treatment (green) were analyzed by HPLC. The shift in distribution towards X4-X6 shows that the xylanase is degrading areas of the oligos formed by KfXYS1 (from X4 and UDP-xylose) and so proves that the linkages formed are 1,4-β-linkages. Gray dashed line, eluent water content (right y-axis). X4-X13, progressively longer labeled xylan oligosaccharides. (b) Donor substrate specificity of KfXYS1 protein heterologously expressed in *S. cerevisiae*. Reactions were performed with mixtures of the indicated UDP-sugars in combination with labeled xylohexaose (X6) and recombinant KfIRX10 protein preparation. Reactions were analyzed by HPLC.

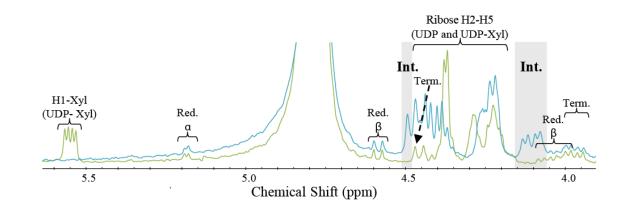


Fig. S11 Characterization of reaction mixtures with or without KfXYS1 by 1D 1H NMR.

Fig. S11. Characterization of reaction mixtures with or without KfXYS1 by 1D 1H NMR. The control spectrum of reaction mixture containing xylobiose and UDP-xylose only (green) contained signals of terminal (Term.) and reducing (Red.) xylose. KfXYS1 reaction spectrum (blue) after 16 h incubation containing diagnostic signals of internal β -1,4-xylose linkages (Int.; grey shade) that are absent in the control. Signals corresponding to H1 of the xylose of UDPxylose in the control spectrum, are absent in the spectrum of the KfXYS1 reaction and only signals of UDP are present.

IDFunctionAGIFamilySurfamilyNameControl<			T	1		[T	1	[[r
Altg2740 NT Altg2740 Altg2740 Altg2740 Altg2320 Altg3320	ID	Function	AGI	Family	Subfamily	Name	Control1	Control2	Control3	Control4	Control5
Xyin Atsg61840 Yur Atsg61840 Yur Atsg61840 Yur Atsg61840 Yur Atsg61840											
AlgebbaNumAlgebbaAlgebaNumAlgeba<	At1g27440		At1g27440	GT47	D	IRX10	At3g57630	At1g21480	At5g33290	At2g35100	At5g62220
Arg28110 Vian backhone Arg22940 Arg28110 Arg27000 Arg22400 Arg27001 Arg27000 Arg27001 Arg27000 Arg27000 Arg27000 Arg27000 Arg2700 Arg2700 Arg2700 Arg27000 Arg2700 Arg27000 Arg2700 Arg2700 Arg2700 Arg2700 Arg2700 Arg2700 Arg2700 Arg2700 Arg2700 <td>A+5 a61940</td> <td></td> <td>A+5g61840</td> <td>GT47</td> <td>D</td> <td></td> <td>A+2a57620</td> <td>A+1a21/80</td> <td>A+5a22200</td> <td>A+2a25100</td> <td>A+5a62220</td>	A+5 a61940		A+5g61840	GT47	D		A+2a57620	A+1a21/80	A+5a22200	A+2a25100	A+5a62220
Ardge10 related Ardge210 Ardge240 Graf D HK7 Ardge760 Ardge2400 Ardge2400 <	ALJ201840		Alg01840	0147	D	INATOL	Alsg57030	ALIZ21400	ALJEJJZJU	ALZESSIO	ALJEUZZZU
Name Name <th< td=""><td>At2g28110</td><td></td><td>At2g28110</td><td>GT47</td><td>D</td><td>IRX7</td><td>At3g57630</td><td>At1g21480</td><td>At5g33290</td><td>At2g35100</td><td>At5g62220</td></th<>	At2g28110		At2g28110	GT47	D	IRX7	At3g57630	At1g21480	At5g33290	At2g35100	At5g62220
Argargaron Kiyan backbone Argargaron GT43 GT43A IRX9 AAH71961 - - - - Argaron Kiyan backbone Argaron GT43 GT43A IRX9 AAH71961 - - - - Argaron Kyian backbone Argaron GT43 GT43B IRX14 AAH71961 - - - - Ardgaron Myian backbone Argaron GT43 GT43B IRX14L AAH71961 -<		Xylan backbone									
Alzgaroon related Alzgaroon GT43 GT43 GT43 RR9 AAH71961 - - - - - Alugaron related Allg27600 GT43 GT43 GT43 RR91 AAH71961 - <td>At5g22940</td> <td></td> <td>At5g22940</td> <td>GT47</td> <td>D</td> <td>IRX7L</td> <td>At3g57630</td> <td>At1g21480</td> <td>At5g33290</td> <td>At2g35100</td> <td>At5g62220</td>	At5g22940		At5g22940	GT47	D	IRX7L	At3g57630	At1g21480	At5g33290	At2g35100	At5g62220
Artg2700 Feilade Artg2760 GT43 GT43 RX91 AAH71961 - - - - Artg26800 related Artg26800 GT43 GT43B IRX14 AAH71961 - - - - Artg26807 related Artg67230 GT43 GT43B IRX14L AAH71961 - - - - Artg67230 related Artg67230 GT43 GT43B IRX14L AAH71961 - - - - Artg7000 related Artg7900 GT8 GAU IRX8 Artg26130 Artg220810 - - - - Myan backbore Artg19300 GT8 GUX GUX Artg26130 Artg220810 -	412-27000		412-27000	6743	CT 43 A		44174064				
At122700 related At4236890 related At436890 related At567230 related At567400 related At6733300 GT8 GUX GUX At5618400 related related <thr></thr> related related <	At2g37090		At2g37090	G143	G143A	IRX9	AAH/1961	-	-	-	-
Ardg36890 related Ardg36890 GT43 GT43B IRX14 AAH71961 -<	At1g27600	-	At1g27600	GT43	GT43A	IRX9L	AAH71961	-	-	-	-
Yigha backbone At5g67230 GT43 GT43B IRX14L AAH71961 - -	0	Xylan backbone									
Artg6723 related Atg67230 GT43 GT43B RN14L AAH71961 - - - - Atg654600 related Atg54600 GT8 GAU IRX8 Al3g61130 At220810 - - - - Atg1900 related At1g19300 GT8 GAU PARVUS At3g61130 At2g20810 - - - - At1g1900 related At1g19300 GT8 GUX GUX1 At3g16100 At2g20810 -	At4g36890		At4g36890	GT43	GT43B	IRX14	AAH71961	-	-	-	-
Atsg54690 related Atsg54690 GTB GAUT IRX8 Atsg61130 Atzg20810 - - - Atsg1910 related Attg19300 GTB GAUT PARVUS At3g61130 At2g20810 - - - - At1g1900 related At1g19300 GTB GAU PARVUS At3g1860 -	ALE - C7220		415-67220	6743	CT 43D		44174064				
At5g54690 related At5g54690 GT8 GAUT IRX8 At3g61130 At2g20810 - - - At1g19300 related At1g19300 GT8 GATL PARVUS At3g61130 At2g20810 - - - Xylan alpha-1,2- At3g18660 GT8 GUX GUX1 At5g18480 - - - - At4g33330 GIAT At3g1810-1,2- At4g33330 GT8 GUX GUX2 At5g18480 - - - - Xylan alpha-1,2- At1g77130 GIAT At1g7130 GUX GUX3 At5g18480 - - - - Xylan alpha-1,2- At1g7130 GIAT At1g74940 GT8 GUX GUX4 At5g18480 - </td <td>At5g67230</td> <td></td> <td>At5g67230</td> <td>G143</td> <td>G143B</td> <td>IRX14L</td> <td>AAH/1961</td> <td>-</td> <td>-</td> <td>-</td> <td>-</td>	At5g67230		At5g67230	G143	G143B	IRX14L	AAH/1961	-	-	-	-
Xylan backbone Atlg19300 GTB GATL PARVUS At3g61130 At2g20810 - - - At1g19300 GIAT At3g18660 GTB GUX GUX1 At3g18610 - - - - At1g1330 GIAT At3g18660 GTB GUX GUX1 At5g18480 -	At5g54690		At5g54690	GT8	GAUT	IRX8	At3g61130	At2g20810	-	-	-
Xylan alpha-1,2- GicAT At3g18660 GT8 GUX GUX1 At5g18480 - - - - At4g13330 GicAT At4g13330 GT8 GUX GUX2 At5g18480 - - - - - - At1g17130 GICAT At1g17130 GT8 GUX GUX3 At5g18480 -	0										
At3g18660 CicAT At3g18660 GT8 GUX GUX1 At5g18480 - - - - At4g3330 GicAT At4g3330 GT8 GUX GUX2 At5g18480 - - - - Xylan alpha-1,2- At1g77130 GicAT At1g77130 GT8 GUX GUX3 At5g18480 - <td>At1g19300</td> <td></td> <td>At1g19300</td> <td>GT8</td> <td>GATL</td> <td>PARVUS</td> <td>At3g61130</td> <td>At2g20810</td> <td>-</td> <td>-</td> <td>-</td>	At1g19300		At1g19300	GT8	GATL	PARVUS	At3g61130	At2g20810	-	-	-
Xylan alpha-1,2- GlcAT Atdg33330 GT8 GUX GUX2 At5g18480 - </td <td>112-10660</td> <td></td> <td>412-40550</td> <td>CT0</td> <td>CUN</td> <td>CLIVA</td> <td>415-40400</td> <td></td> <td></td> <td></td> <td></td>	112-10660		412-40550	CT0	CUN	CLIVA	415-40400				
Ar4g33330 GIAT At4g33330 GT8 GUX GUX2 At5g18480 - - - - At1g77130 GIAT At1g77130 GT8 GUX GUX3 At5g18480 - - - - - At1g77130 GIAT At1g574940 GT8 GUX GUX3 At5g18480 -	At3g18660		At3g18660	GI8	GUX	GUX1	At5g18480	-	-	-	-
Xylan alpha-1,2- GicAT Atlg77130 GT8 GUX GUX3 At5g18480 - </td <td>At4g33330</td> <td></td> <td>At4g33330</td> <td>GT8</td> <td>GUX</td> <td>GUX2</td> <td>At5g18480</td> <td>-</td> <td>-</td> <td>-</td> <td>-</td>	At4g33330		At4g33330	GT8	GUX	GUX2	At5g18480	-	-	-	-
Junc Xylan alpha-1,2- G(AT At1g54940 At1g54940 GT8 G(AT GUX GUX At1g518480 At5g18480 -											
At1g54940GicATAt1g54940GT8GUXGUX4At5g18480 <td>At1g77130</td> <td>GlcAT</td> <td>At1g77130</td> <td>GT8</td> <td>GUX</td> <td>GUX3</td> <td>At5g18480</td> <td>-</td> <td>-</td> <td>-</td> <td>-</td>	At1g77130	GlcAT	At1g77130	GT8	GUX	GUX3	At5g18480	-	-	-	-
Xylan alpha-1,2- GicAT Attg08990 GT8 GUX GUX5 Atsg18480 - - - - Attg03800 4-O-MethylT Attg03800 DUF579 DUF579 GXMT1 -				070	O 11/2	0.04					
At1g08990GIcATAt1g08990GT8GUXGUX5At5g18480 <td>At1g54940</td> <td></td> <td>At1g54940</td> <td>GT8</td> <td>GUX</td> <td>GUX4</td> <td>At5g18480</td> <td>-</td> <td>-</td> <td>-</td> <td>-</td>	At1g54940		At1g54940	GT8	GUX	GUX4	At5g18480	-	-	-	-
Glucoronoxylan At1g33800 DUF579 DUF579 GXMT1 -	At1g08990		At1g08990	GT8	GUX	GUX5	At5g18480	-	-	-	-
Xylan backbone At3g50220Xylan backbone relatedAt3g50220DUF579IRX15Xylan backbone relatedAt5g67210DUF579DUF579IRX15LXylan alpha-1,3- Os02g22480Os02g22480GT61GT61CWOs XAT1At5g55500Xylan alpha-1,3- Os03g37010Os03g37010GT61GT61CWOs XAT2At5g55500Xylan alpha-1,3- Vylan alpha-1,3- AraTFR873610GT61GT61CWOs XAT2At5g55500Xylan alpha-1,3- FR846232GT61GT61CWTa XAT1At5g55500Xylan alpha-1,3- FR846232GT61GT61CWTa XAT1At5g55500 </td <td>0</td> <td></td> <td>0</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>	0		0								
At3g50200relatedAt3g50200DUF579DUF579IRX15 <td>At1g33800</td> <td>4-O-MethylT</td> <td>At1g33800</td> <td>DUF579</td> <td>DUF579</td> <td>GXMT1</td> <td>-</td> <td>-</td> <td>-</td> <td>-</td> <td>-</td>	At1g33800	4-O-MethylT	At1g33800	DUF579	DUF579	GXMT1	-	-	-	-	-
Xylan backbone related At5g67210 DUF579 DUF579 IRX15L - </td <td></td> <td></td> <td></td> <td>0</td> <td>0.15570</td> <td>10145</td> <td></td> <td></td> <td></td> <td></td> <td></td>				0	0.15570	10145					
At5g67210relatedAt5g67210DUF579DUF579IRX15L </td <td>At3g50220</td> <td></td> <td>At3g50220</td> <td>DUF579</td> <td>DUF579</td> <td>IRX15</td> <td>-</td> <td>-</td> <td>-</td> <td>-</td> <td>-</td>	At3g50220		At3g50220	DUF579	DUF579	IRX15	-	-	-	-	-
Xylan alpha-1,3- AraTOs02g22480GT61GT61CWOs XAT1At5g55500Mylan alpha-1,3- AraTOs03g37010GT61GT61CWOs XAT2At5g55500Mylan alpha-1,3- FR873610AraTOs03g37010GT61GT61CWOs XAT2At5g55500	At5g67210	-	At5g67210	DUF579	DUF579	IRX15L	-	-	-	-	-
Xylan alpha-1,3- Os03g37010Xylan alpha-1,3- Os03g37010GT61GT61CWOs XAT2At5g55500FR873610AraTFR873610GT61GT61CWTa XAT1At5g55500						-					
Os03g37010AraTOs03g37010GT61GT61CWOs XAT2Atsg55500<	Os02g22480		Os02g22480	GT61	GT61CW	Os XAT1	At5g55500	-	-	-	-
Xylan alpha-1,3- AraT FR873610 GT61 GT61 GT61CW Ta XAT1 At5g55500 - - - - Xylan alpha-1,3- FR846232 AraT FR846232 GT61 GT61CW Ta XAT2 At5g55500 -	0-02-27040		0-02-27040	CTC4	CTC4 ON	0. 1472	A+F + F F F 6 0 0				
FR873610AraTFR873610GT61GT61 WTa XAT1Atsg55500Kylan alpha-1,3- AraTFR846232GT61GT61 WTa XAT2Atsg55500 <td< td=""><td>USU3g37010</td><td></td><td>USU3g37010</td><td>G161</td><td>GIBICW</td><td>US XA12</td><td>At5g55500</td><td>-</td><td>-</td><td>-</td><td>-</td></td<>	USU3g37010		USU3g37010	G161	GIBICW	US XA12	At5g55500	-	-	-	-
Xylan alpha-1,3- AraTFR846232GT61GT61CWTa XAT2At5g55500S02g22380beta-1,2-XylTOs02g22380GT61GT61CWOs XAX1At5g55500Putative xylanClosest AtAt3g18170AraTAt3g18170GT61GT61CWto XAT1At5g55500Putative xylanClosest AtPutative xylan-Closest AtAt3g18170GT61GT61CWto XAT1At5g55500At3g18170AraTAt3g18170GT61GT61CWto XAT1At5g55500At3g18170GT61GT61GT61CWto XAT1At5g55500 </td <td>FR873610</td> <td></td> <td>FR873610</td> <td>GT61</td> <td>GT61CW</td> <td>Ta XAT1</td> <td>At5g55500</td> <td>-</td> <td>-</td> <td>-</td> <td>-</td>	FR873610		FR873610	GT61	GT61CW	Ta XAT1	At5g55500	-	-	-	-
FR846232AraTFR846232GT61GT61CWTa XAT2At5g5500Nylan arabinosyl-No20222380GT61GT61CWOs XAX1At5g5500 <t< td=""><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></t<>											
Os02g22380beta-1,2-XyITOs02g22380GT61GT61CWOs XAX1At5g5500Putative xylanAraTAt3g18170GT61GT61CWto XAT1At5g5500 <td>FR846232</td> <td>AraT</td> <td>FR846232</td> <td>GT61</td> <td>GT61CW</td> <td>Ta XAT2</td> <td>At5g55500</td> <td>-</td> <td>-</td> <td>-</td> <td>-</td>	FR846232	AraT	FR846232	GT61	GT61CW	Ta XAT2	At5g55500	-	-	-	-
Putative xylanPutative xylanAt3g18170GT61GT61CWClosest At to XAT1At5g55500At3g18170AraTAt3g18170GT61GT61CWto XAT1At5g55500At3g18170AraTAt3g18170GT61GT61CWto XAT1At5g55500 <td< td=""><td>0-02-22200</td><td></td><td>0-02-22200</td><td>CTC4</td><td>CTC4 ON</td><td>0- 1414</td><td>A+F + F F F 6 0 0</td><td></td><td></td><td></td><td></td></td<>	0-02-22200		0-02-22200	CTC4	CTC4 ON	0- 1414	A+F + F F F 6 0 0				
At3g18170AraTAt3g18170GT61GT61CWto XAT1At5g5500Putative xylanClosest At <td< td=""><td>USU2822380</td><td></td><td>USU2822380</td><td>6161</td><td>GIDICW</td><td></td><td>At5g55500</td><td>-</td><td>-</td><td>-</td><td>-</td></td<>	USU2822380		USU2822380	6161	GIDICW		At5g55500	-	-	-	-
Putative xylan At3g18170 GT61 GT61W Closest At At5g55500 - - - At3g18170 AraT At3g18170 GT61 GT61CW to XAT1 At5g55500 - - - - Xylan acetyltransferase At3g55990 TBR/TBL TBR/TBL TBL29 At5g55800 - - - - Xylan coumaroyl ransferase Os06g39390 BAHD BAHD Os AT10 At3g48720 - - -	At3g18170		At3g18170	GT61	GT61CW		At5g55500	-	-	-	-
Xylan Xylan acetyltransferase At3g55990 TBR/TBL TBR/TBL TBR/TBL TBR/T	J										
At3g55990 acetyltransferase At3g55990 TBR/TBL TBL29 At5g58600 - - - - - Xylan coumaroyl Xylan coumaroyl BAHD BAHD Os AT10 At3g48720 - - - - - 0s06g39390 transferase Os06g39390 BAHD BAHD Os AT10 At3g48720 - - - -	At3g18170		At3g18170	GT61	GT61CW	to XAT1	At5g55500	-	-	-	-
Xylan coumaroyl transferase Solog39390 BAHD BAHD Os AT10 At3g48720 - - - - Closest At - - - - - - - -	A+2 ~ F F 0 0 0		A+2~55000	דרם / דרי			A+F ~F 0 C 0 C				
Os06g39390 transferase Os06g39390 BAHD BAHD Os AT10 At3g48720 - <th< td=""><td>At3g55990</td><td></td><td>At3g55990</td><td>IBK/IBL</td><td>I BK/ I BL</td><td>I BL29</td><td>At5g58600</td><td>-</td><td>-</td><td>-</td><td>-</td></th<>	At3g55990		At3g55990	IBK/IBL	I BK/ I BL	I BL29	At5g58600	-	-	-	-
Closest At	Os06g39390		Os06g39390	BAHD	BAHD	Os AT10	At3g48720	-	-	-	-
At3g62160 - At3g62160 BAHD BAHD to OsAT10 At3g48720	00		<u> </u>								
	At3g62160	-	At3g62160	BAHD	BAHD	to OsAT10	At3g48720	-	-	-	-

Table S1 Positive/negative list used for identification of xylan synthesis related sequences

Table S2 Protein sequences of *Klebsormidium flaccidum* xylan gene orthologs.

Supplementary Table 1. P	rotein sequences of Klebsormidium flaccidum xylan gene orthologs
IRX10/IRX7 orthologs	
kfl00398_0130 / KflRX10	MPNKLFTGFAALFCFCFFLLGSIIQRSSLFVLDVENKATVQAAERELTERTWDVQEGTTGVVARRALLDRLMLEEEVPEDDEEDKVPSVLRESPEGRLKVFVYDLPSKYNTKLLEKDSRCLNHMFA AEIFVHRNLLNSSIRTMDPSEADWFYTPVYTTCDLTPNGLPLPFKAPQLMRSAIRWVAEHYPYWNRSEGADHFFTVPHDFGACFHYQEEAIKRRILPVLQRATLVQTFGQNNHVCLKEGSIVV PPYAPPQKMRSRFIPPDTPRSIFVYFRGMFYDISNDPEGGYYARGARAAIWENFKDNPLFDLSSDHPSTYYEDMQRAVFCLCPLGWAPWSPRIVESVIFGCIPVIIGDDIVLPFSDAIRWDEISVY VAEKDVQNLDTILTAIKPEEIMAMQRKMWKPEVRTAMLFPQPAEPGDSFHQILNGLARKLPHKPEVFVKPGEHPIDWTAGPQQDQDEWRGDDEN
kfl00254_0250	MPPSDATLGRAGSMSNASGRSVSPATNRRCVCGYCAPSSFCTCASSPAPSTGESLGGLQRSSQMSLTSAFSSKQLRKQASSTRLDQGAFFYFRQRLAFSCTGLMLVMLCVALASRHLQHASV DGLERVPPFRDSRRMMMSADDGISVSTGTTAVIQNGLESTGLALPKEVIPNSPLGGPLLEAPARKYNAIASQMVSITSAFSSKQLRKQASSTRLDQGAFFYFRQRLAFSCTGLMLVMLCVALASRHLQHASV FFIPVYSTGDLNDQGRALPQMAGMMIKQVMHHVVHTWPFFNRTHGSDHVVVVPHDFGACFTFREEEAVTRKVPAVLRKLLMLQTFGQENHTCSRSGRSIIIPPVVNVNWVKSVPGLPEGP GQPPRDKFAYFRGLVHDFGNDPWGNKYSKGVRYQIWKHHKDNPLFDVAEMVPFAYYEDMRRSVFCLCPLGWAPWSPRIVESLQQGCIPVIISDNIILPFREFVDWPAITVHVPEKDVPRLDEI LSAIPVEEVAAMQKRIWSPAVRRAVLFPQPAEEGDAFHVVLRELARKLEQRREAQREKDRKSKSKKKSPTNALRGESQTAALHALR
kfl00144_0240	MGCGRRWIAVNVVGVCLLLLVFGGVVGPDLTQSGLGPGDSKSGRVSGVGHGRRARALLQSRRALQQTDRTFARERRRWDAEALQNVGVPVILGRGVSVEDRVLHFRRRGISDFFEAEPEWP ESQESEEPEDEVGAESSMEGDTDQVLDSKSVGEAGSGKTGHKGLGKGGPGSVPELIIRDSDSEVRTAETFAPVQTLFVNVTANSTGGAANSSNLEKKGPLLGQGLPLNITGGVATNKSSRLNQS EFSSRDPDFDAHPHILQNVTEAERVEQTTAESEPLTKGAPGEVRDIEQEQGLELEUETTAPTQPPEEAEMQNPFAADPETAPMLVLPPQRSIKVYIYDLPPHLNRDFVTRDPRVRQTMFATEVLIH ENFEQSAVRTLDPDEADFYFIPVYTAAVLEDSGRTIPDSSHGTMLEAVMYVAERWPYFNRMRGLDHVVVAPHDYGACFHYRDELASQTKFRMVLRYLTVLQTFGVHNHVCYRPGRSVVIPPYI DVRKLRQLAFAPERDAHEKSVLAFFRGLYIDAGNDATGKVYSRGVRLAIGQTFANHPDFDISDAPSLTYLQDMMRATFCLCPLGWAPWSPRIFESVQIGCIPVIIADDIILPFENYFNWERIAVRV KEADVPRLEEILKGIPPEKVAAMQEAMVWDPQVRQALIFPEPAEPGDAFEMILNELEAKVERIRHKRAEVFRKHGQ
kfl00113_0290	MHGVKLPRLTTSSGNHSKQPHRRGPSYSKRRAPLRRWIRRLGGCTVCCASVILFLLMSAVVFLAVKSWPGAAVSSAFDLGTDESPANVWRRGGERGGEHDAEEWGGGEIAKGLDLGAYTA MVGGDVAPPTSIADLGLPVPLRIVYELPPSFNVDFLRNPRCASHLFAAEVAIHRALLQSPVRTLOPAEADFFFVPVYTACNFNTENGFPNLSGAPEQLRTAVELLAAEYQWNRTNGRDHVFV ATHDFGACFHAMEKAAVQAGVPQFLRNSILQTFGQTNNHPCQDVDHIQIPPYVPSPLVLANWQPPESQNRTIFAYMSGKIEVHPKESVSGDVYGRAVRTGIWKRFHRKRRFLIRRKRPEAYY EEMLHARFCLCPRGWAPWSPRIVESVLFGCIPVIIADHIKLPFDDLLNWPAMGLILQENEVPNLATILESLSEDRVASMQRELWRAEVRRTLLFLEPSVRGDATWQILDALAKKKMERDKLQ
IRX8 ortholog	
kfl00095_0020 / KflRX8	MKLKLSPSMRRITITAQEDFHPMGLKSAPACCSYSCVMVSLLGLALILPMLFITTAIVKLDSIDNCIDGECYTRGRKMSPTDRADELVTGLLDVPVSVGRQLNEERTGVGGDAVVGEREETIDDLI AQVSELNDDSASVILKLRALVEVHEGKARMAKLQRALVRHFASKGVPKGLHCLSLRLTADYSTNPEANRPLAGPEEISRLTDNSLRHFVLLSDNVLATSVVVNSTVTNAKEPEKVVFHVVTDETTY SAMHAWFALFPPEPATIEVRGVHEMQWLTAAKAPVLDLMENSLALQKLNYGDRILEEAANDPPQVLAAKLQARSPKYISIMNHLRMYLPEVFPELDRIIFLDDDVVQKDLTGLWEMDMRG KVNAGVETCRGADRFALSKQFRSYFNFSNPLIAEHFAPETCAWAFGLNVFDLKRWRESNITETYHHWQKQNMMLINMTUWRLGTLPPALIAFVGHTLPLDHSWHMLGLGYQPDSSIEVIKSSA VIHYNGPAKPWLDMAFTDFRSYWTRWVDYGNEFVHQCNIFQMRMMGRKKAPSSSSTPPGQKAKGPRQDSGAADGSLPERRSFVHGSAFLIVCFACFLLGFRLSWEGKAAFLGAYRIGPES GKLVLRDTGTSLDYADSESGGGAAQGGLDLRPGAWESGQGDKYGAGVADGEWERIETRILASAGGVVNASESNRSRRLLPRPPPSPEHPWVGRHNISVRPWPHPDAGETLLAHALIAAVQ REQQNQFGVPLDRRKRMIVVTPTYRRAFQQHLSHLTDTTRLAPEPLSWLVVEAAGFTEETQALISKTGLADIHLLPVPGVTMPDVRAPERPIVEDYLRDVALGYIQKQAMEGVVVFADDSATL RLALFQELQKVRWFAAFSVGMLVQGGGSQVDLPVQGPVCGPEGDVRGWHAADHELLRLKKPAVVPKRALLGAAEKRLPGGEGAQRRLLALEVPERTEGFGAGSDSISQLGAGTVVRQUG VEQSEYFEPLEIFIPKLEWVGFAFNARLLWAPAERPAWLREWKPSTTLGKVDAVVGDVGDSGDEDPTDLLARMVEDESQVETLGRCGKDVLAWWLKVEAAEEGQYPAGWRLAENLPVVAEPQE WSEEQWAEYEARLVEEKKKKKKKSKKKKAPEATKVEEDAAARQAAEQAAADEARRVEEEKQRELEAARQAEEERQTRLAEAQKAEVEDNLDSVQKAKAEVRTDPEKEGEKAESKLESSALA EKEALTNNAKEVLNKLVEGADGGVADDTQAREKQTS
IRX9 ortholog	
>kfl00579_0060 / KflRX9	MPPGGGKPMAKHAPIIGSLLNNRRPLLLGFCFLLGFLLGVFDTRRFEASFPDAVSHSFPGQALEDEQRGLVGAGASRILVQQQLPAQPRLEAPVQLLIVTPTYNRPFQAMYLTRLAHTLRLVRPP VLWIVVETGWQSEESVALLMRTGLMYRHLVVQQFSPQHKDRGVHQRNHALDHIQQHALQGIVYFADDDNNYSVELFEEIRKVKRFGVLPVAMLVEGAEKTTIEGPVCAGGRVIGWHTSEES AAQRRFHSDLCGFAFNSSLMWEPGGEGRRRNAPVRFRDNIKEGFQETTFIEQLVGSEAEMEGLADDCREVLVWHLRAEAARDSVYPTSWMLRASMNITTPARTPESIREP
IRX14 ortholog	
kfl00262_0200 / KfIRX14	MMGKRKAPSSSSTPPGQKAKGPRQDSGAADGSLPERRSFVHGSAFLIVCFACFLLGFRLSWEGKAAFLGAYRIGPESGKLVLRDTGTSLDYADSESGGGAAQGGLDLRPGAWESGQGDKYGA GVADGEWERIETRILASAGGVVNASESNRSRLLPRPPSSPPEHPWVGRHNISVRPWPHPDAGETLLAHALIAAVQREQQNQFGVPLDRRKRMIVVTPTYRRAFQQHLSHLTDTLRLAPEPLS WLVVEAGGPTEETQAILSKTGLADIHLLPVPGVTMPDVRAPERPIVEDYLRDVALGYIQKQAMEGVVFAADSATLRLALFQELQKVRWFAAFSVGMLVQGGSQVDLPVQGPVCCPEGDV RGWHAADHELLRLKKPAVVPKRALLGAAEKRLPGGEGAQRRLLALEVPERTEGFGAGSDSISQLGAGTVVRQLQQVEQSEYFEPELFIPKLEWVGFAFNARLLWAPAERPAWLREWKPSTTLG KVDAVVGDYEGDEDPTDLLARMVEDESQVETLGRCGKDVLAWWLKVEAAEEGQYPAGWRLAENLPVVAEPQEWSEEQWAEYEARLVEEKKKKKKKKKKAEPEATKVEEDAAARQAA EQAAADEARRVEEEKQRELEAARQAEEERQTRLAEAQKAEVEDNLDSVQKAKAEVRTDPEKEGEKAESKLESSALAEKEALTNNAKEVLINKLVEGADGGVADDTQAREKQTS

Table S3 Protein BLAST anal	sis of Klebsormidium flaccidun	n xylan gene orthologs
-----------------------------	--------------------------------	------------------------

Supplementary Table S2. Protein BLAST analysis of <i>Klebsormidium flaccidum</i> xylan gene orthologs							
Protein name	Max score	Total score	Query cover	E value	Ident		
IRX10L	860	860	1	0	1		
Potri012G109600	778	778	0.95	0	0.93		
PoIRX10_5	764	764	1	0	0.88		
IRX10	750	750	1	0	0.86		
Potri003G162000	741	741	0.93	0	0.9		
Potri001G068100	738	738	0.99	0	0.85		
Os01g70200	732	732	0.96	0	0.86		
PoIRX10_6	730	730	0.95	0	0.86		
Bd2g59410	720	720	0.93	0	0.87		
Potri012G109200	719	719	0.95	0	0.88		
Os02g32110	719	719	0.93	0	0.88		
Os01g70190	716	716	0.94	0	0.85		
Sm442111	711	711	0.93	0	0.87		
Bd2g59400	711	711	0.93	0	0.85		
Os04g32670	711	711	0.93	0	0.86		
Bd5g08400	710	710	0.93	0	0.86		
PoIRX10_3	708	708	0.98	0	0.81		
Bd3g44420	703	703	0.93	0	0.85		
Pp1s7_455V6	674	674	0.99	0	0.77		
Bd2g59380	667	667	0.92	0	0.8		
Os01g70180	650	650	0.94	0	0.77		
Os10g10080	644	644	0.98	0	0.74		
kfl00398_0130 / KflRX10	620	620	0.9	0	0.75		
PoIRX10_1	593	593	0.97	0	0.69		
PoIRX10_2	573	573	0.88	0	0.73		
PoIRX10_4	566	566	0.91	0	0.71		
kfl00254_0250	379	379	0.83	4E-130	0.5		
kfl00144_0240	345	345	0.84	2E-115	0.47		
kfl00113_0290	314	314	0.81	4E-106	0.46		
FRA8	299	318	0.87	1E-100	0.42		
XGD1	103	103	0.62	4E-27	0.3		
AtIRX8	1112	1112	1	0	1		
Potri.001G416800	901	901	1	0	0.82		
Potri.011G132600	871	871	0.99	0	0.81		
Potri.004G124500	634	634	1	0	0.55		
Potri.017G090800	622	622	1	0	0.54		
Bradi4g03670	596	596	0.99	0	0.54		
LOC_Os12g38930	583	583	0.95	0	0.55		
Potri.004G124500	574	574	0.87	0	0.57		
Bradi1g70290 Bradi1g70290	571 571	571 571	0.99 0.99	0	0.51 0.51		
U	566	566	0.99	0	0.51		
LOC_Os03g11330 kfl00095_0020_IRX8	478	478	0.99	2E-159	0.43		
AtIRX9-L	823	823	1	0	1		
kfl00579_0060_IRX9	288	288	0.74	8E-99	0.49		
Potri.016G086400	244	244	0.79	9E-82	0.42		
Potri.006G131000	244	244	0.79	1E-80	0.42		
AtIRX9	218	218	0.77	1E-71	0.39		
Bradi2g37970	172	172	0.79	2E-53	0.34		
LOC_Os05g03174	169	169	0.81	2E-52	0.33		
LOC_Os04g58040	50.4	68.9	0.26	1E-10	0.39		
AtIRX14	1082	1082	1	0	1		
AtIRX14-L	652	652	0.96	0	0.65		
Potri.007G047500	642	642	0.88	0	0.7		
Potri.005G141500	640	640	1	0	0.65		
Potri.007G047500	589	589	0.83	0	0.69		
Bradi1g33320	480	480	0.8	2E-169	0.54		
LOC Os06g47340	475	475	0.8	2E-167	0.55		
LOC_Os04g55670	246	246	0.84	2E-79	0.36		
kfl00262_0200_IRX14	206	350	0.8	4E-62	0.35		

Pairwise protein sequence comparison was performed using BLAST (Altschul et al., 1997). For each sub section of the table the proteins are compared pairwise with the top entry maked in bold. *K. flaccidum* proteins are indicated in red.

Notes S1 Newick tree files and sequences used for Fig. 2 and Fig. S2-S9.