

Table 1, supplementary data. Annotation of Arabidopsis LRXs and related proteins in databases.

Results from query with "extensin AND "repeat" as key words at NCBI Map Viewer (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?chr=arabid.inf). Twelve genes labelled with pink boxes were found with with search. The 11 LRX genes annotated by Baumberger et al. (2003) are labelled with violet boxes. Functional domains were found using InterProScan (<http://www.ebi.ac.uk/Tools/InterProScan/>) and ScanProsite (<http://www.expasy.ch/tools/scanprosite/>). Names of domains (IPR: InterProScan; PF: PFAM; PR: SPRINT; PS: PROSITE) are indicated as well as the corresponding scores and positions on the amino acid sequence between square brackets.

Annotations in databases		Interproscan and Prosite results	Annotation by Baumberger et al. (2003)
At1g12040	<p>NCBI</p> <p>LRX1 (LEUCINE-RICH REPEAT/EXTENSIN 1) similar to LRX2 (LEUCINE-RICH REPEAT/EXTENSIN 2), protein binding / structural constituent of cell wall [Arabidopsis thaliana] (TAIR:AT1G62440.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR LEUCINE-RICH REPEAT/EXTENSIN 1, LRX1 encodes a a chimeric leucine-rich repeat/extensin protein that regulates root hair morphogenesis and elongation. Null mutants develop root hairs that frequently abort, swell, or branch. Gene is expressed in root hair cells and protein is specifically localized in the wall of the hair proper.</p> <p>Leucine-rich repeat (N-terminal:IPR013210) Leucine-rich repeat:Molecular Function: protein bi (IR001611)</p> <p>TIGR leucine-rich repeat family protein / extensin family protein (LRX1) similar to extensin-like protein [Lycopersicon esculentum] gjl5917664[gb]AAD55979 contains leucine-rich repeats (PF00560) contains proline rich extensin domains (IPR002965)</p> <p>MIPS leucine-rich repeat/extensin 1 (LRX1)</p>	<p>IPR001611: Leucine-rich repeat 6000.000322506199 [148-170]T, 32.9999983317489 [172-191]T, 2100.0003726265404 [243-265]T, 3500.00043291567 [267-289]T, 39.9999971482335 [291-313]T</p> <p>PF00560: LRR_1</p> <p>IPR013210: Leucine-rich repeat, N-terminal PF08263: LRR_NT 0.04300000526153389 [51-90]T</p> <p>unintegrated 4.8E-14 [383-395]T, 4.8E-14 [406-427]T, 4.8E-14 [428-444]T, 4.8E-14 [457-482]T</p> <p>PR01217: PRICHEXTENSIN</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 104.855 [383 - 741]</p>	AtLRX1
At1g49490	<p>NCBI</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT3G19020.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Pistil-specific extensin-like protein (IPR003882) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT3G19020.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Pistil-specific extensin-like protein (IPR003882) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TIGR leucine-rich repeat family protein / extensin family protein contains similarity to disease resistance protein Gt:3894383 from [Lycopersicon esculentum] contains leucine-rich repeats (PF00560) contains proline rich extensin domains (IPR002965)</p> <p>MIPS hypothetical protein</p>	<p>IPR001611: Leucine-rich repeat 1899.9996417184002 [180-202]T 3399.99972437719 [250-272]T 6000.000322506199 [274-296]T 58.999994304626 [298-320]T</p> <p>PF00560: LRR_1</p> <p>IPR003882: Pistil-specific extensin-like protein PR01218: PSTLEXTENSIN 1.4E-5 [530-553]T 1.4E-5 [565-583]T</p> <p>IPR013210: Leucine-rich repeat, N-terminal PF08263: LRR_NT 2.00000006595059 [59-98]T</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 86.183 [394 - 842]</p>	AtPEX2

At1g62440	NCBI	<p>IPR001611: Leucine-rich repeat PF00560: LRR_1 0.022 [214-233]T, 4.1 [285-307]T, 0.029 [309-331]T, 0.029 [333-355]T</p> <p>IPR003883: Extensin-like protein PF02095: Extensin_1 5.9 [636-645]T, 63 [664-673]T, 2.9 [679-688]T, 22 [693-702]T, 6.6 [722-731]T, 22 [737-746]T</p> <p>Extensins are plant cell-wall proteins; they can account for up to 20% of the dry weight of the cell wall. They are highly-glycosylated, possibly reflecting their interactions with cell-wall carbohydrates. Amongst their functions is cell wall strengthening in response to mechanical stress (e.g., during attack by pests, plant-bending in the wind, etc.). This repeat occurs within extensin-like proteins.</p>	AtLRX2
	<p>LRX2 (LEUCINE-RICH REPEAT/EXTENSIN 2) similar to LRX1 (LEUCINE-RICH REPEAT/EXTENSIN 1), protein binding / structural constituent of cell wall [Arabidopsis thaliana] (TAIR:AT1G12040.1)</p> <p>contains InterPro domain Extensin-like protein (IPR003883) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR LEUCINE-RICH REPEAT/EXTENSIN 2, LRX2 encodes a paralog of LRX1 (LEUCINE-RICH REPEAT/EXTENSIN 1) which acts synergistically with LRX1 in root hair cell morphogenesis. Leucine-rich repeat, N-terminal (IPR013210) Extensin-like protein;Molecular Function: structur (IPR003883) Leucine-rich repeat;Molecular Function: protein bi (IPR001611)</p> <p>TIGR leucine-rich repeat family protein / extensin family protein similar to extensin-like protein [Lycopersicon esculentum] gjl5917664 gb AAD55979 contains leucine-rich repeats (PF00560) contains proline rich extensin domains (IPR002965)</p> <p>MIPS putative extensin-like protein (gnl PID e1310400)</p>	<p>PS50099: PRO_RICH Proline-rich region profile score = 108.140 [425 - 824]</p>	
At2g15880	NCBI	<p>IPR001611: Leucine-rich repeat PF00560: LRR_1 51.9999946268774 [299-321]T</p> <p>IPR003882: Pistil-specific extensin-like protein PR01218: PSTLEXTENSIN 4.1E-10 [514-537]T, 4.1E-10 [549-567]T, 4.1E-10 [663-684]T</p> <p>unintegrated PR01217: PRICHEXTENSIN 2.7E-11 [390-402]T, 2.7E-11 [404-420]T, 2.7E-11 [424-436]T, 2.7E-11 [442-463]T, 2.7E-11 [465-481]T</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 109.004 [394 - 722]</p>	AtPEX3
	<p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G33970.1)</p> <p>contains InterPro domain Pistil-specific extensin-like protein (IPR003882) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR leucine-rich repeat family protein / extensin family protein leucine-rich repeat family protein / extensin family protein; similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G33970.1)</p> <p>contains InterPro domain Pistil-specific extensin-like protein (IPR003882) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TIGR leucine-rich repeat family protein / extensin family protein similar to extensin-like protein [Lycopersicon esculentum] gjl5917664 gb AAD55979 contains leucine-rich repeats (PF00560) contains proline rich extensin domains (IPR002965)</p> <p>MIPS unknown protein</p>		

At2g19780	<p>NCBI</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G29240.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G29240.1) similar to hypothetical protein [Vitis vinifera] (GB:CAN62837.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TIGR leucine-rich repeat family protein / extensin family protein contains leucine rich-repeat (LRR) domains (PF00560, IPR001611) similar to leucine-rich repeat/extensin 1 (GI:13809918) [Arabidopsis thaliana] contains similarity to disease resistance protein [Lycopersicon esculentum] gj13894383 gb AAC78591</p> <p>MIPS putative disease resistance protein</p>	<p>IPR001611: Leucine-rich repeat 4.599999811982681 [134-156]T, 20.0000020319536 [158-178]T, 5.90000024020674 [182-201]T, 6099.99977126217 [205-224]T, 45.000003473935095 [301-323]T</p> <p>PF00560: LRR_1</p> <p>IPR013210: Leucine-rich repeat, N-terminal PF08263: LRR_NT 8.29999581180007E-5 [68-107]T</p>	<p>none</p>
At3g19020	<p>NCBI</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT1G49490.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Pistil-specific extensin-like protein (IPR003882) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT1G49490.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210)</p> <p>contains InterPro domain Pistil-specific extensin-like protein (IPR003882) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TIGR leucine-rich repeat family protein / extensin family protein similar to extensin-like protein [Lycopersicon esculentum] gj15917664 gb AAD55979 contains leucine-rich repeats (PF00560) contains proline rich extensin domains (IPR002965)</p> <p>MIPS hypothetical protein</p>	<p>IPR001611: Leucine-rich repeat 4000.00081278147 [193-215]T, 7300.001901602701 [263-285]T, 3800.00071256818 [287-309]T</p> <p>PF00560: LRR_1</p> <p>IPR003882: Pistil-specific extensin-like protein PR01218: PSTLEXTENSIN 8.7E-6 [675-698]T, 8.7E-6 [710-728]T, 8.7E-6 [811-832]T</p> <p>IPR013210: Leucine-rich repeat, N-terminal PF08263: LRR_NT 5.3000002124871 [72-111]T</p> <p>unintegrated PR01217: PRICHEXTENSIN 6.4E-9 [406-418]T, 6.4E-9 [423-444]T, 6.4E-9 [446-462]T, 6.4E-9 [470-487]T</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 132.344 [407 - 954]</p>	<p>AtPEX1</p>
At3g22800	<p>NCBI</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT3G24480.1) similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G13340.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Pistil-specific extensin-like protein (IPR003882) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G13340.1)</p> <p>contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Pistil-specific extensin-like protein (IPR003882) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TIGR leucine-rich repeat family protein / extensin family protein similar to extensin-like protein [Lycopersicon esculentum] gj15917664 gb AAD55979 contains leucine-rich repeats (PF00560) contains proline rich extensin domains (IPR002965)</p> <p>MIPS hypothetical protein</p>	<p>IPR001611: Leucine-rich repeat 2200.0002175608097 [148-170]T, 38.999996884611306 [172-194]T, 1499.99977583351 [268-290]T, 32.9999983317489 [292-314]T</p> <p>PF00560: LRR_1</p> <p>IPR003882: Pistil-specific extensin-like protein PR01218: PSTLEXTENSIN 1.2E-5 [400-423]T, 1.2E-5 [435-453]T</p> <p>IPR013210: Leucine-rich repeat, N-terminal PF08263: LRR_NT 0.0989999902584295 [51-90]T</p> <p>unintegrated PR01217: PRICHEXTENSIN 3.6E-14 [374-386]T, 3.6E-14 [386-407]T, 3.6E-14 [414-430]T</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 52.470 [376 - 467]</p>	<p>AtLRX6</p>

<p>At3g24480</p> <p>NCBI</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G18670.1) similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G13340.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210)</p> <p>contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G18670.1) similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G13340.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TIGR</p> <p>leucine-rich repeat family protein / extensin family protein similar to extensin-like protein [Lycopersicon esculentum] gjl5917664[gb]AAD55979 contains leucine-rich repeats (PF00560) contains proline rich extensin domains (IPR002965)</p> <p>MIPS</p> <p>disease resistance protein, putative</p>		<p>IPR001611: Leucine-rich repeat</p> <p>2100.0003726265404 [195-214]T, 5099.9988243997805 [289-311]T, 3099.99945899369 [313-335]T</p> <p>PF00560: LRR_1</p> <p>IPR013210: Leucine-rich repeat, N-terminal</p> <p>PF08263: LRR_NT 0.0739999967908919 [74-113]T</p> <p>unintegrated</p> <p>PR01217: PRICHEXTENSN 5.4E-16 [411-432]T, 5.4E-16 [434-450]T, 5.4E-16 [456-473]T</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 34.144 [403 - 492]</p>	<p>AtLRX4</p>
<p>At4g06744</p> <p>NCBI</p> <p>leucine-rich repeat family protein / extensin family protein Identical to Uncharacterized protein At4g06744 precursor [Arabidopsis Thaliana] (GB:Q8W3M4) similar to leucine-rich repeat family protein [Arabidopsis thaliana] (TAIR:AT3G19320.1) similar to leucine-rich repeat family protein [Arabidopsis thaliana] (TAIR:AT1G49750.1) similar to Leucine-rich repeat, plant specific [Medicago truncatula] (GB:ABD32635.1) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR</p> <p>leucine-rich repeat family protein / extensin family protein Identical to Uncharacterized protein At4g06744 precursor [Arabidopsis Thaliana] (GB:Q8W3M4) similar to leucine-rich repeat family protein [Arabidopsis thaliana] (TAIR:AT3G19320.1) similar to Leucine-rich repeat, plant specific [Medicago truncatula] (GB:ABD32635.1) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TIGR</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat/extensin 1 (GI:13809918) (Arabidopsis thaliana) contains PF00560: Leucine Rich Repeat domains</p> <p>MIPS</p> <p>not found</p>		<p>IPR001611: Leucine-rich repeat</p> <p>2700.00035356384 [143-165]T, 4100.000492084209 [167-189]T, 8199.9995664553 [215-238]T, 5200.00089003336 [264-286]T, 48.9999951782359 [314-335]T</p> <p>PF00560: LRR_1</p>	<p>none</p>

At4g13340	<p>NCBI</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT3G24480.1) similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G18670.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Pistil-specific extensin-like protein (IPR003882) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT3G24480.1) similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G18670.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Pistil-specific extensin-like protein (IPR003882) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TIGR</p> <p>leucine-rich repeat family protein / extensin family protein similar to extensin-like protein [Lycopersicon esculentum] gjl5917664 gb AAD55979 contains leucine-rich repeats (PF00560) contains proline rich extensin domains (IPR002965)</p> <p>MIPS</p> <p>extensin-like protein strong similarity to known protein : extensin-like protein, Zea mays, PIR2:S49915</p>	<p>IPR001611: Leucine-rich repeat 2100.0003726265404 [187-206]T, 7199.99952919347 [281-303]T, 1599.99977185865 [305-327]T</p> <p>PF00560: LRR_1</p> <p>IPR003882: Pistil-specific extensin-like protein</p> <p>PR01218: PSTLEXTENSIN 6.2E-5 [416-439]T, 6.2E-5 [449-467]T</p> <p>IPR013210: Leucine-rich repeat, N-terminal PF08263: LRR_NT 0.00230000061930563 [66-105]T unintegrated</p> <p>PR01217: PRICHEXTENSIN 1.0E-14 [392-413]T, 1.0E-14 [425-441]T, 1.0E-14 [443-460]T</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 153.609 [394 - 758]</p>	AtLRX3
At4g18670	<p>NCBI</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT3G24480.1) similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G13340.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611) contains InterPro domain Atrophin (IPR002951)</p> <p>TAIR</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT3G24480.1) similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT4G13340.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611) contains InterPro domain Atrophin (IPR002951)</p> <p>TIGR</p> <p>leucine-rich repeat family protein / extensin family protein similar to extensin-like protein [Lycopersicon esculentum] gjl5917664 gb AAD55979 contains leucine-rich repeats (PF00560) contains proline rich extensin domains (IPR002965)</p> <p>MIPS</p> <p>extensin-like protein strong similarity to known protein : extensin-like protein - maize, PIR2:S49915</p>	<p>IPR001611: Leucine-rich repeat 3200.00074719642 [199-218]T, 3099.99945899369 [293-315]T, 4600.00044330866 [317-339]T</p> <p>PF00560: LRR_1</p> <p>IPR002951: Atrophin</p> <p>PR01222: ATROPHIN 2.1E-5 [410-438]T, 2.1E-5 [759-780]T</p> <p>IPR013210: Leucine-rich repeat, N-terminal PF08263: LRR_NT 0.0739999967908919 [78-117]T unintegrated</p> <p>PR01217: PRICHEXTENSIN 2.4E-11 [402-423]T, 2.4E-11 [426-442]T, 2.4E-11 [446-463]T, 2.4E-11 [465-490]T</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 99.495 [406 - 837]</p> <p>PS50324: SER_RICH Serine-rich region profile score = 13.379 [418 - 596]</p> <p>PS50316: HIS_RICH Histidine-rich region profile score = 13.990 [628 - 674]</p>	AtLRX5

At4g29240	<p>NCBI</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT2G19780.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT2G19780.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TIGR</p> <p>leucine-rich repeat family protein / extensin family protein contains PF00560: Leucine Rich Repeat domains similar to leucine-rich repeat/extensin 1 (GI:13809918) [Arabidopsis thaliana]</p> <p>MIPS</p> <p>extensin-like protein</p>	<p>IPR001611: Leucine-rich repeat 31.0000030991121 [139-161]T, 31.999986882983 [163-183]T, 29.999985646001 [187-206]T, 6899.99963379704 [210-232]T, 3099.99945899369 [306-328]T</p> <p>PF00560: LRR_1</p> <p>IPR013210: Leucine-rich repeat, N-terminal PF08263: LRR_NT 0.016000002110419 [73-112]T</p> <p>PS50315: GLY_RICH Glycine-rich region profile score = 14.253 [21 - 61]</p>	<p>none</p>
At4g33970	<p>NCBI</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT2G15880.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT2G15880.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TIGR</p> <p>leucine-rich repeat family protein / extensin family protein similar to extensin-like protein [Lycopersicon esculentum] gj5917664[gb]AAD55979 contains leucine-rich repeats (PF00560) contains proline rich extensin domains (IPR002965)</p> <p>MIPS</p> <p>extensin-like protein strong similarity to known protein : extensin-like protein, Zea mays, PIR2:S49915</p>	<p>IPR001611: Leucine-rich repeat PF00560: LRR_1 31.0000030991121 [325-347]T</p> <p>IPR013210: Leucine-rich repeat, N-terminal PF08263: LRR_NT 0.52000001763605 [86-125]T</p> <p>unintegrated 4.0E-14 [418-430]T, 4.0E-14 [434-450]T, 4.0E-14 [454-466]T, 4.0E-14 [467-488]T, 4.0E-14 [490-506]T</p> <p>PR01217: PRICHEXTENSIN</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 82.898 [419 - 694]</p>	<p>AtPEX4</p>
At5g25550	<p>NCBI</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT3G24480.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TAIR</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat family protein / extensin family protein [Arabidopsis thaliana] (TAIR:AT3G24480.1) contains InterPro domain Leucine-rich repeat, N-terminal (IPR013210) contains InterPro domain Leucine-rich repeat (IPR001611)</p> <p>TIGR</p> <p>leucine-rich repeat family protein / extensin family protein similar to leucine-rich repeat/extensin 1 (GI:13809918) [Arabidopsis thaliana] contains PF00560: Leucine Rich Repeat domains</p> <p>MIPS</p> <p>extensin-like protein strong similarity to known protein : Pex1 extensin-like protein, Lycopersicon esculentum, EMBL:AF159296</p>	<p>IPR001611: Leucine-rich repeat PF00560: LRR_1 0.034 [172-191]T, 7.8 [218-240]T, 2.2 [267-289]T, 0.011 [291-313]T</p> <p>IPR013210: Leucine-rich repeat, N-terminal PF08263: LRR_NT 6.5e-05 [51-90]T</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 16.337 [382 - 430]</p>	<p>AtLRX7</p>

Table 2, supplementary data. Annotation of some cell wall proteins in databases.

Functional domains were found using InterProScan (<http://www.ebi.ac.uk/Tools/InterProScan/>) and ScanProsite (<http://www.expasy.ch/tools/scanprosite/>). Names of domains are as follows: IPR: InterProScan; PF: PFAM; PR: SPRINT; PTHR: PANTHER; PS: PROSITE; SSF: Superfamily. The corresponding scores are given as well as the positions on the amino acid sequence between square brackets.

Annotations in databases		Interproscan and Prosite results	Annotation by experts
Some extensin genes			
At1g21310	<p>NCBI</p> <p>AtEXT3 (ROOT SHOOT HYPOCOTYL DEFECTIVE) contains InterPro domain Extensin-like region (IPR006706)</p> <p>TAIR</p> <p>AtEXT3 (ROOT SHOOT HYPOCOTYL DEFECTIVE) contains InterPro domain Extensin-like region (IPR006706)</p> <p>TIGR proline-rich extensin-like family protein contains extensin-like region (PF04554)</p> <p>MIPS extensin 3 (atExt3)(AtExt5)</p>	<p>IPR006706: Extensin-like region</p> <p>Extensins are homologous hydroxyproline-rich glycoproteins (HRGPs) found in the plant extracellular matrix. The key to the role of HRGPs in cell wall self-assembly and cell extension lies in their chemistry, which is dependent on extensive post-translational modifications.</p> <p>PF04554: Extensin-like region 1.9e-40 [9-220]T, 1.9e-33 [221-428]T</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 95.519 [34 - 428]</p>	<p>Johnson et al. (2003)</p> <p>extensin (ROOT SHOOT HYPOCOTYL DEFECTIVE) (RSH)</p>
At1g26240	<p>NCBI</p> <p>proline-rich extensin-like family protein contains InterPro domain Extensin-like region (IPR006706)</p> <p>TAIR</p> <p>proline-rich extensin-like family protein contains InterPro domain Extensin-like region (IPR006706)</p> <p>TIGR</p> <p>proline-rich extensin-like family protein similar to hydroxyproline-rich glycoprotein precursor gj 727264 gb AAA87902 contains proline-rich extensin domains (IPR002965)</p> <p>MIPS hypothetical protein</p>	<p>IPR006706: Extensin-like region</p> <p>1.10000150671643E-32 [10-143]T, 2.10000267834025E-23 [144-243]T, 4.0999959904575E-28 [244-362]T, 9.70001793430386E-24 [364-478]T</p> <p>PF04554: Extensin-like region unintegrated</p> <p>PR01217: PRICHEXTENSIN 2.5E-10 [28-44]T, 2.5E-10 [50-62]T, 2.5E-10 [63-84]T, 2.5E-10 [88-104]T</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 150.151 [32 - 476]</p> <p>PS50328: TYR_RICH Tyrosine-rich region profile score = 32.719 [28 - 470]</p>	<p>extensin</p>
At1g26250	<p>NCBI</p> <p>proline-rich extensin, putative contains InterPro domain Extensin-like region (IPR006706)</p> <p>TAIR</p> <p>proline-rich extensin, putative contains InterPro domain Extensin-like region (IPR006706)</p> <p>TIGR</p> <p>proline-rich extensin, putative similar to extensin gj 1165322 gb AAB53156 contains proline-rich extensin domains (IPR002965)</p> <p>MIPS hypothetical protein</p>	<p>IPR006706: Extensin-like region</p> <p>5.89998809405682E-29 [9-163]T, 1.99999636034536E-24 [164-273]T, 7.300005909151711E-16 [274-443]T</p> <p>PF04554: Extensin-like region unintegrated</p> <p>PR01217: PRICHEXTENSIN 5.5E-10 [24-40]T, 5.5E-10 [44-56]T, 5.5E-10 [63-84]T, 5.5E-10 [88-104]T</p> <p>PS50099: PRO_RICH Proline-rich region profile score = 131.825 [29 - 441]</p> <p>PS50328: TYR_RICH Tyrosine-rich region profile score = 37.973 [17 - 435]</p>	<p>extensin</p>

Some pollen Ole e1 allergen and extensin family proteins				
At1g78040	NCBI	IPR006041: Pollen Ole e 1 allergen and extensin	none	
	pollen Ole e 1 allergen and extensin family protein similar to unknown [Populus trichocarpa] (GB:ABK94215.1) contains InterPro domain Pollen Ole e 1 allergen and extensin (IPR006041)	PF01190: Pollen_Ole_e_I 9.70001793430386E-27 [32-153]T		These proteins are most probably secreted and consist of about 145 residues. There are six cysteines which are conserved in the sequence of these proteins. They seem to be involved in disulphide bonds
	TAIR			
	TIGR			
	MIPS			
	phosphoglycerate mutase 1 like protein Prosite motif: 1-6 TonB-dependent receptor proteins signatures	TonB box, conserved site PS00430: TONB_DEPENDENT_REC_1 0.0 [1-35]?		
At3g33790	NCBI	IPR006041: Pollen Ole e 1 allergen and extensin	AtAGP30	
	pollen Ole e 1 allergen and extensin family protein similar to AGP31 (ARABINOGLACTAN-PROTEIN 31) structural constituent of cell wall (AT1G28290.1) contains InterPro domain Pollen Ole e 1 allergen and extensin (IPR006041)	PF01190: Pollen_Ole_e_I 9.09992432541697E-68 [112-239]T		
	TAIR	PS50099: PRO_RICH Proline-rich region profile score = 16.164 [30 - 103]		
	TIGR			
	MIPS			
	putative proline-rich protein			
others				
At2g15770	NCBI	IPR003245: Plastocyanin-like	AtEN23 Nersissian and Shipp (2002)	
	glycine-rich protein similar to glycine-rich protein (AT2G15780.1) contains InterPro domain Cupredoxin (IPR008972) contains InterPro domain Plastocyanin-like (IPR003245)	PF02298: Cu_bind_like 0.026000001808627 [154-184]T		
	TAIR	IPR008972: Cupredoxin SSF4953: Cupredoxins superfamily 4.0E-17 [141-258]T		
	glycine-rich protein contains InterPro domain Cupredoxin (IPR008972) contains InterPro domain Plastocyanin-like (IPR003245)	PS50324: SER_RICH Serine-rich region profile score = 18.782 [21 - 136]		
	TIGR	PS50315: GLY_RICH Glycine-rich region profile score = 8.908 [35 - 96]		
	hypothetical protein			
At1g15825	NCBI	PS50099: PRO_RICH Proline-rich region profile score = 24.463 [2 - 123]	none	
	hydroxyproline-rich glycoprotein family protein similar to leucine-rich repeat family protein / extensin family protein (AT3G19020.1)			
	TAIR			
	TIGR			
	MIPS			
	putative protein			

At1g12090	<p>NCBI</p> <p>ELP, EXTENSIN-LIKE PROTEIN similar to protease inhibitor/seed storage/lipid transfer protein (LTP) family protein (AT1G62510.1) similar to extensin-like protein [Brassica napus] (AAK30571.1) contains InterPro domain Plant lipid transfer protein/seed storage/trypsin-alpha amylase inhibitor (IPR003612) contains InterPro domain Bifunctional inhibitor/plant lipid transfer protein/seed storage (IPR016140) contains InterPro domain Plant lipid transfer protein and hydrophobic protein helical (IPR013770)</p> <p>TAIR</p> <p>ELP, EXTENSIN-LIKE PROTEIN Plant lipid transfer protein/seed storage/trypsin (IPR003612) Bifunctional inhibitor/plant lipid transfer protein (IPR016140) Plant lipid transfer protein and hydrophobic protein (IPR013770)</p> <p>TIGR</p> <p>protease inhibitor/seed storage/lipid transfer protein (LTP) family protein similar to 14 kDa polypeptide [Catharanthus roseus] (407410)</p> <p>contains Pfam protease inhibitor/seed storage/LTP family domain (PF00234)</p> <p>MIPS</p> <p>pEARLI 1-like protein 4Fe-4S ferredoxins, iron-sulfur binding region signature</p>	<p>IPR003612: Plant lipid transfer protein/seed storage/trypsin-alpha amylase inhibitor PF00234: Tryp_alpha_amyl 6.09999977126231E-23 [55-137]T</p> <p>IPR001450: 4Fe-4S ferredoxin, iron-sulphur binding PF00198: 4FE4S_FERREDOXIN 0.0 [23-34]?</p>	none
AAK30571 Brassica napus	<p>NCBI</p> <p>extensin-like protein</p> <p>98% identity with At1g12090</p>	<p>IPR003612: Plant lipid transfer protein/seed storage/trypsin-alpha amylase inhibitor PF00234: Tryp_alpha_amyl 1.3999989204989E-19 [55-137]T</p> <p>IPR001450: 4Fe-4S ferredoxin, iron-sulphur binding PF00198: 4FE4S_FERREDOXIN 0.0 [23-34]?</p>	