Table S1. Plant proteins containing HMG-box domain(s).

Protein[1,2]	Length (aa)	Mass (Da)	Locus[3]	Databases[4]
HMGB				
AtHMGB1	178	20265	At3g51880	
AtHMGB2	144	15982	At1g20693	
AtHMGB3	141	15681	At1g20696	
AtHMGB4	138	15364	At2g17560	
AtHMGB5	125	14203	At4g35570	
AtHMGB6	241	26964	At5g23420	
AtHMGB14	151	17481	At2g34450	
BdHMGB1	160	17347	Bradi1g29730	MIPS
BdHMGB2	127	14781	Bradi2g45927	MIPS
BdHMGB3	217	23410	Bradi5g18460	MIPS
BdHMGB4	144	15737	Bradi3g51140	MIPS
BdHMGB5	218	23816	Bradi3g12920	MIPS
BdHMGB6	128	13997	Bradi4g37247	MIPS
CrHMGB3401	179	19516		
CrHMGB3402[5]	336	37857		
CrHMGB3404[5]	200	22263		
CrHMGB3405[5]	206	22663		
CrHMGB3406	99	11757		
GmHMGB1	169	18863	Glyma01g28820	SUPFAM
GmHMGB2	166	18371	Glyma03g08270	SUPFAM
GmHMGB3	181	20249	Glyma04g01840	SUPFAM
GmHMGB4	177	20610	Glyma05g25900	SUPFAM
GmHMGB5	157	17310	Glyma06g01970	SUPFAM
GmHMGB6	205	24591	Glyma06g36940	SUPFAM
GmHMGB7	148	17318	Glyma08g08840	SUPFAM
GmHMGB8	142	15928	Glyma09g33070	SUPFAM
GmHMGB9	200	22433	Glyma10g43440	SUPFAM
MtHMGB1304	155	16952		
MtHMGB1305	170	19096		
MtHMGB1306	152	17554		
MtHMGB1307	142	15991		
MtHMGB1308	111	12684		
OsHMGB1	157	17100	Os06g51220	
OsHMGB705	202	22328	Os08g01100	

OsHMGB706	145	15765	Os02g44930	
OsHMGB707	131	13873	Os04g47690	
OsHMGB710	146	16349	Os09g37910	
OsHMGB711	133	15566	Os01g47600	
PpHMGB2	165	18684		
PpHMGB3	157	17517		
PpHMGB1504	207	23103		
PpHMGB1505	215	22675		
PtHMGB901	152	16531		
PtHMGB902	152	16709		
PtHMGB907	201	22544		
PtHMGB910	176	19879		
PtHMGB911	179	20033		
PtHMGB914	165	19322		
PtHMGB915	144	15916		
SbHMGB2602	126	13993		
SmHMGB1	156	17965		GENOM
VcHMGB6403	199	21506		
VvHMGB1	153	17168		SUPFAM
VvHMGB2	175	20397		SUPFAM
VvHMGB3	154	17137		SUPFAM
VvHMGB4	128	14506		SUPFAM
VvHMGB5	168	18888		SUPFAM
VvHMGB6	190	21558		SUPFAM
ZmHMGB1	157	17146		
ZmHMGB2	139	15316		
ZmHMGB3	138	15007		
ZmHMGB4	126	14104		
ZmHMGB5	123	13637		
ZmHMGB106	212	23534		
ZmHMGB107	127	14215		
ARID-HMG				
AtARID-HMG1	338	38774	At1g76110	
AtARID-HMG2	319	36296	At3g13350	
AtHMGB11	337	38049	At1g55650	
AtHMGB15	448	50004	At1g04880	
BdARID-HMG1	444	49374	Bradi3g43310	MIPS

GmHMGB1207	342	38529		
GmHMGB1208	313	35456		
MtHMGB1301	443	50132		
MtHMGB1302	424	47766		
OsHMGB702	467	51629	Os02g27060	
OsHMGB709	300	34993	Os09g37250	
PpHMGB1511	753	81975		
PtHMGB903	329	37740		
PtHMGB904	316	36023		
PtHMGB908	389	43825		
PtHMGB909	467	51892		
SbHMGB2601	455	50524		
SmARID-HMG1	446	50134		GENOM
SmARID-HMG2	496	55792		GENOM
StHMGB1401	314	35751		
VvHMGB6901	458	51094		
VvHMGB6902	328	37339		
VvHMGB6904	300	34108		
ZmHMGB115	442	49700		
3xHMG-box				
At3xHMG-box1	456	53231	At4g23800	
At3xHMG-box2	446	52303	At4g11080	
Bd3xHMG-box1	537	60501	Bradi3g09690	MIPS
CpHMGB13004	470	55645		
GmHMGB1201	473	55698		
GmHMGB1206	478	56132		
OsHMGB704	504	56024	Os02g15810	
PpHMGB1506	483	56028		GENOM
PpHMGB1507	368	42921		GENOM
PtHMGB912	498	58004		
PtHMGB913	480	56174		
RcHMGB20901	514	60136		
SbHMGB2603	488	55669		
Sm3xHMG-box1	457	52813		GENOM
Sm3xHMG-box2	413	47880		GENOM
VvHMGB6903	366	43326		
ZmHMGB108	487	55953		

SSRP1				
AtSSRP1	646	71646	At3g28730	
BdSSRP1	643	71427	Bradi2g33580	
BdSSRP2	641	70776	Bradi2g05430	
CrSSRP3403	576	61145		
GmSSRP1201	579	65991		
GmSSRP1202	557	62444		
OsSSRP701	641	71334	Os01g08970	
OsSSRP702	640	71048	Os05g08970	
PtSSRP901	644	72498		
PtSSRP902	688	78278		
PpSSRP1501	674	74420		
PpSSRP	662	73084		GENOM
SbSSRP2601	639	70441		
SbSSRP2602	639	71362		
SmSSRP	647	71678		GENOM
VvSSRP6901	640	71534		
VcSSRP6401	645	69944		
ZmSSRP1	611	68309		

[1]The species of which the protein sequences were retrieved are indicated as follows: *Arabidopsis thaliana* (At), *Brachypodium distachyon* (Bd), *Carica papaya* (Cp), *Chlamydomonas reinhardtii* (Cr), *Glycine max* (Gm), *Medicago truncatula* (Mt), *Oryza sativa* (Os), *Populus trichocarpa* (Pt), *Physcomitrella patens* (Pp), *Ricinus communis* (Rc), *Selaginella moellendorfii* (Sm), *Solanum tuberosum* (St), *Sorghum bicolor* (Sb), *Vitis vinifera* (Vv), *Volvox carteri* (Vc), *Zea mays* (Zm).

[2]For the majority of HMG-box proteins we use the published name or the name given in The Plant Chromatin Database. The remaining proteins were named according to the HMG-box family (HMGB, ARID-HMG, 3xHMG-box, SSRP1) they belong to, since in several cases the nomenclature in the different databases is inconsistent and/or the names are too long to be used in the phylogenetic trees (cf. Figs. 1, S1, S2).

[3]If available, the loci encoding the respective proteins are enclosed.

[4]Unless the amino acid sequence is retrieved from The Plant Chromatin Database (http://www.chromdb.org), the source of the respective sequence is indicated: SUPFAM (http://supfam2.cs.bris.ac.uk), MIPS (http://mips.helmholtz-muenchen.de/plant/brachypodium/), GENOM (http://genome.jgi-psf.org).

[5]The three amino acid sequences from *Chlamydomonas* contain two putative HMG-box domains.

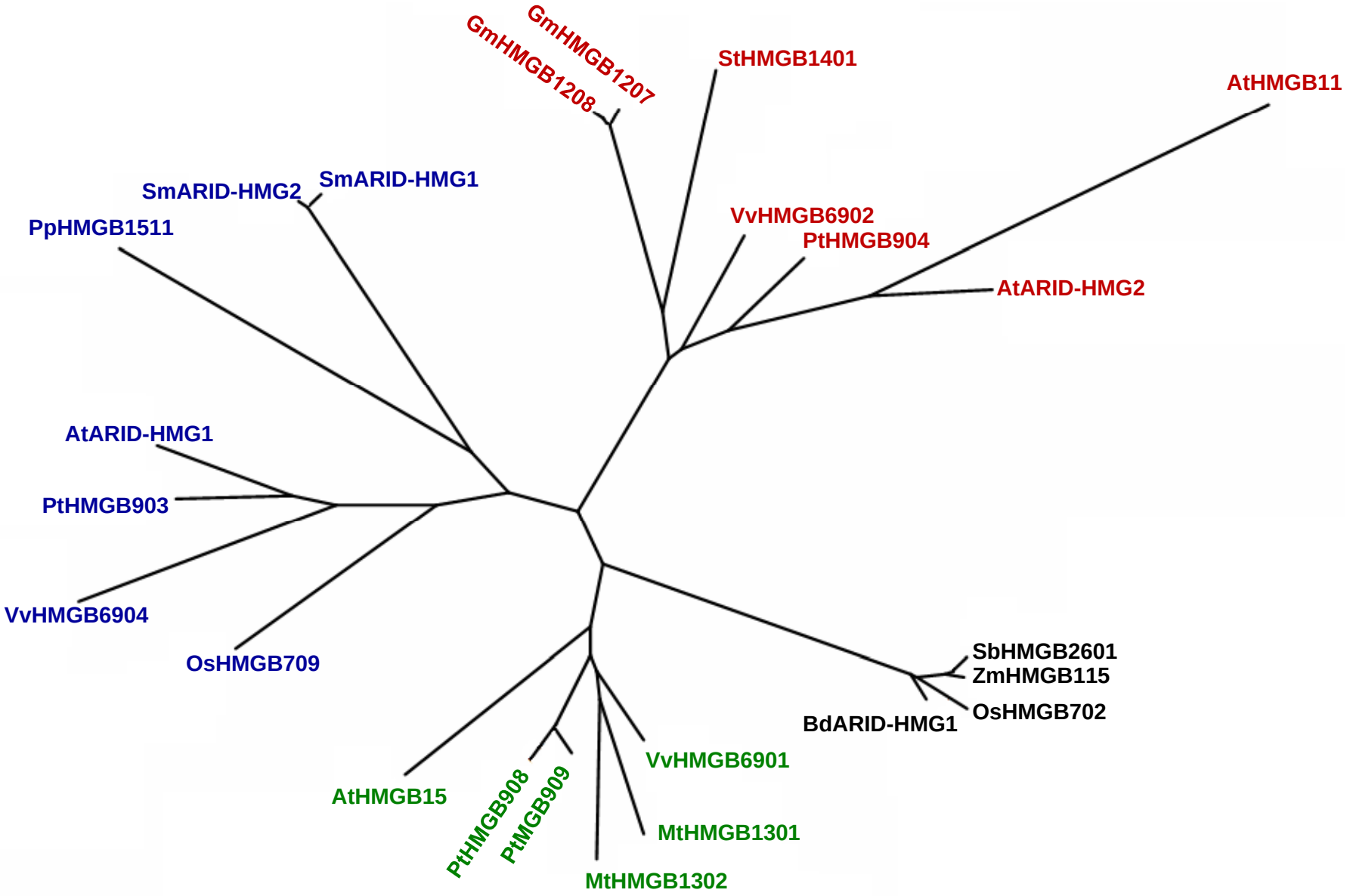

Figure S1. Amino acid sequence similarity of HMG-ARID proteins of different plant species (cf. Table S1 for designation of proteins and species). A multiple sequence alignment was generated and used to construct the shown neighbor-joining tree (http://pbil.univ-lyon1.fr/software/seaview.html). The sequences occur in four subgroups representing related monocot (in black) and dicot (in green) proteins, as well as a more distantly related dicot group (in red) and a group that contains in addition to monocot/dicot members sequences from *Physcomitrella* and *Selaginella* (in blue).

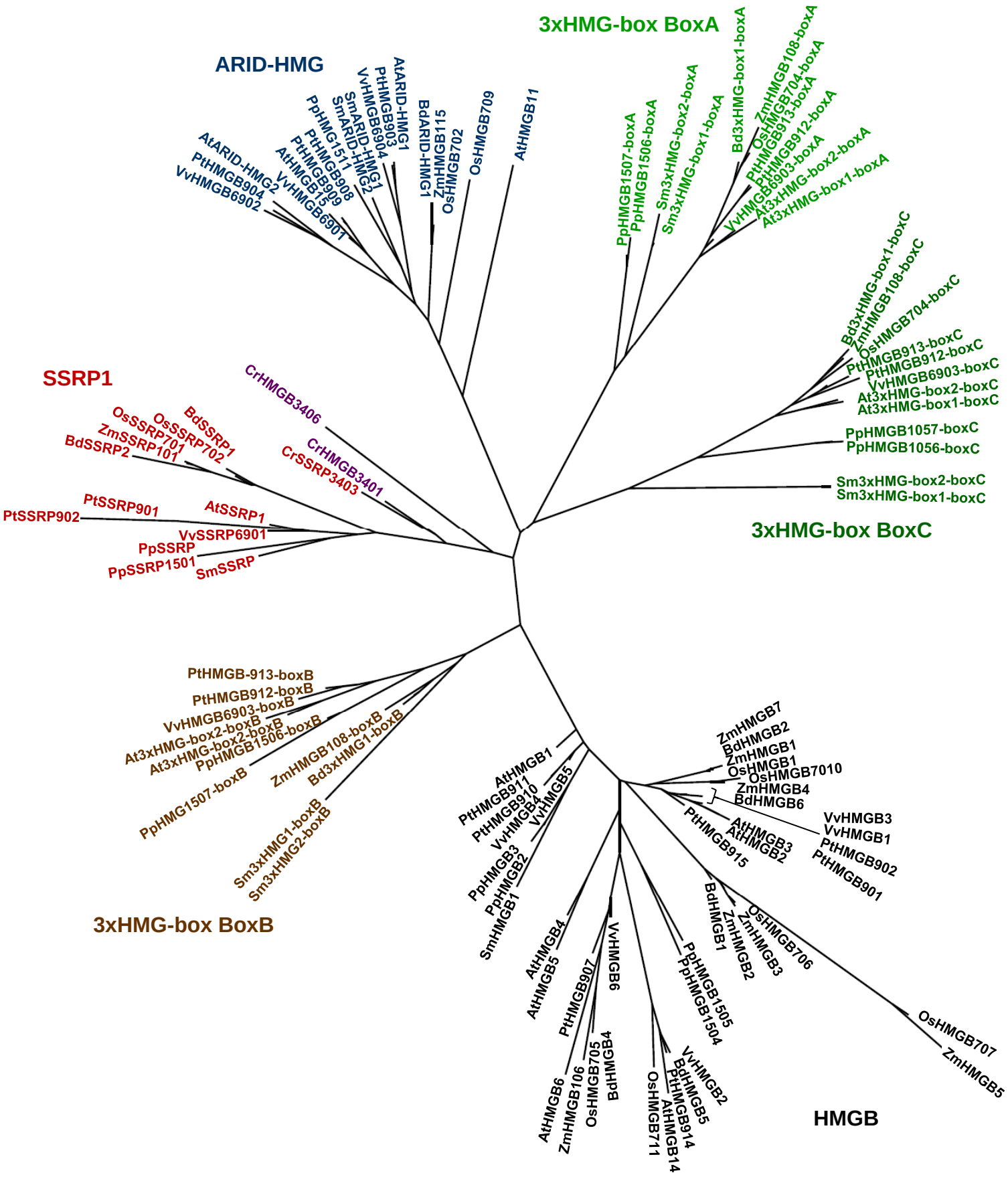

Figure S2. Amino acid sequence similarity of the individual HMG-box domains of the different plant HMG-box protein families (cf. Table S1). A multiple sequence alignment was generated and used to construct a neighbor-joining tree (http://pbil.univ-lyon1.fr/software/seaview.html). The sequences occur in subgroups representing the HMG-box domains originating from HMGB, SSRP1 and ARID-HMG proteins, except that the HMG domains of two putative *Clamydomonas* HMGB-type proteins (in violet) group with the SSRP1 domains. Moreover, the individual HMG-box domains of the 3xHMG-box proteins (termed BoxA, BoxB, BoxC) form separate groups.

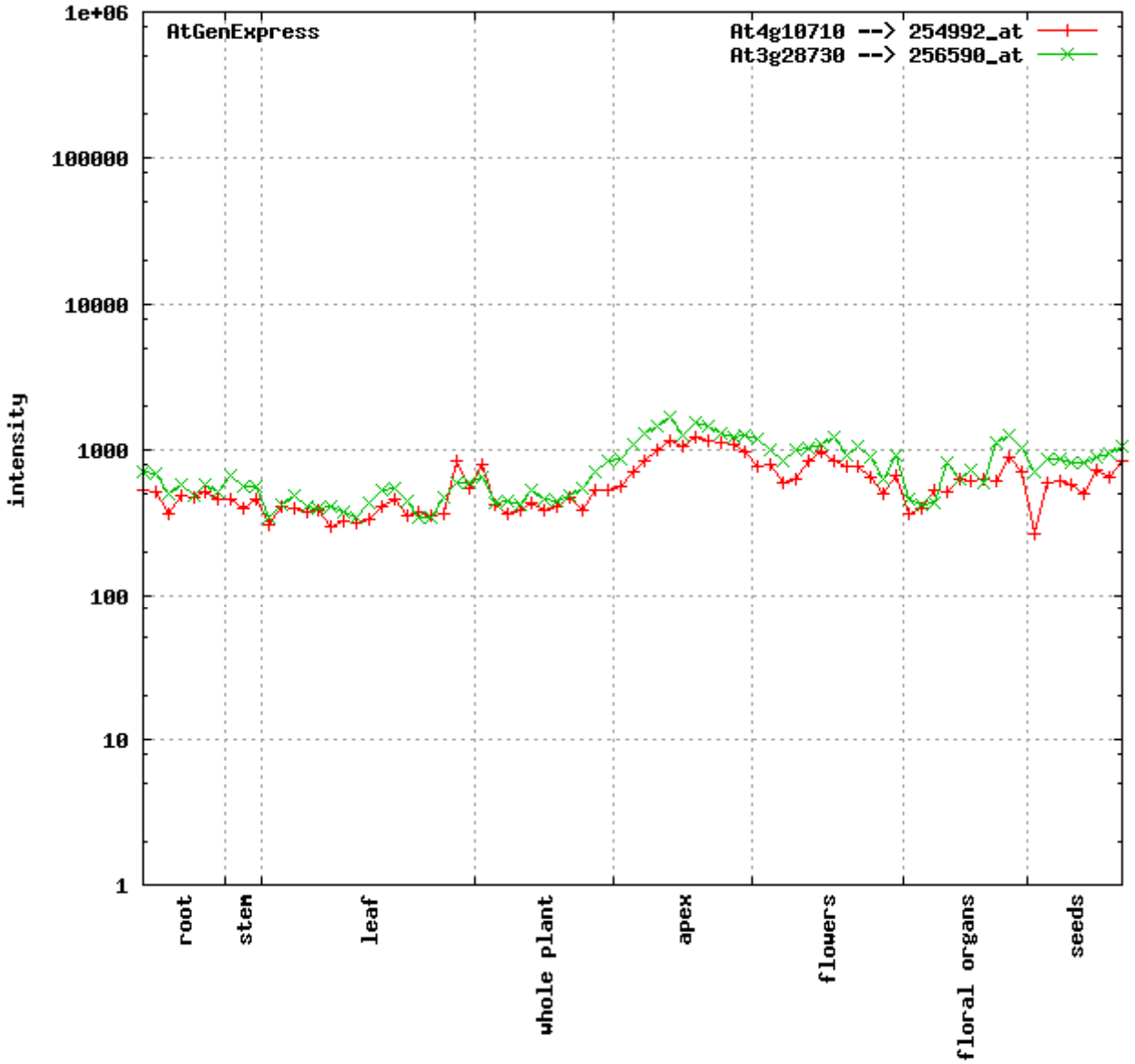

Figure S3. Similar expression of the genes encoding the FACT subunits SSRP1 (At3g28730) and SPT16 (At4g10710) in different *Arabidopsis* tissues. The data reflecting mRNA levels originate from a series of microarray experiments and are displayed using the AtGenExpress visualization tool (http://jsp.weigelworld.org/expviz/expviz.jsp).