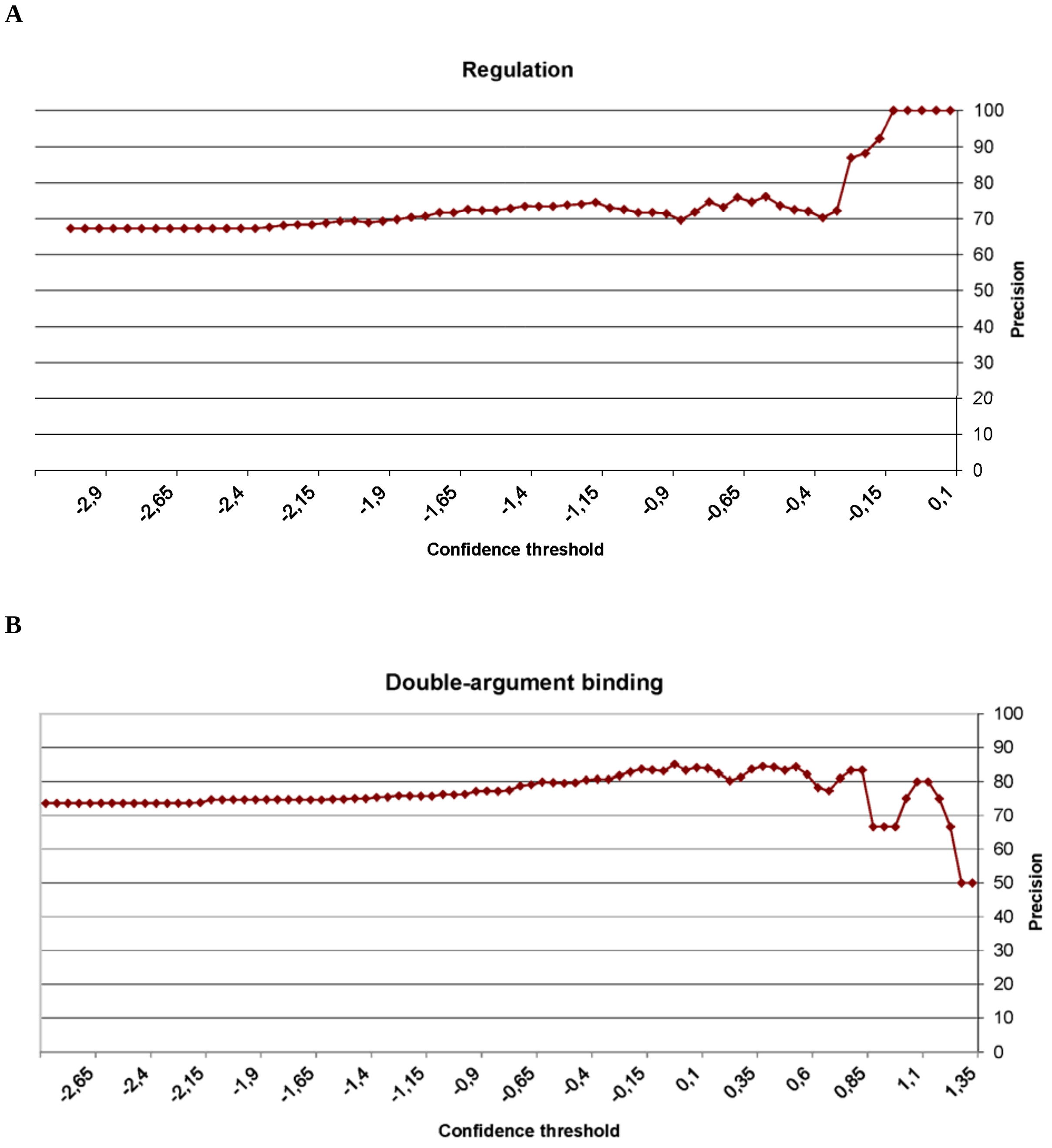
Supplemental Figure 1. Precision rates in the PLEV dataset, plotted against the confidence thresholds of the text mining data. A) Regulatory events. B) Binding events. A more stringent cut-off results in more accurate predictions in general (cf. regulatory events). However, a few high-ranking but wrong predictions can disturb the ranking severely for the most stringent cut-off values (artifact at the end of the curve for double-argument bindings).



Supplemental Table 1. Overview of the mid- and large-scale PPI mapping studies.

PubMed ID	Number of PPIs	Type of proteins	REF
PMID:17426018	205	cell cycle	[1]
PMID:19000166	215	VH1/BRL2 receptor-like kinase	[2]
PMID:19130088	233	calcium-dependent protein kinases	[3]
PMID:17340043	241	TFIID	[4]
PMID:19452453	256	14-3-3 protein complexes	[5]
PMID:14749489	280	Skp1-related genes	[6]
PMID:12795696	443	SCF ubiquitin E3 ligase subunits	[7]
PMID:20407024	514	cell cycle	[8]
PMID:21189294	560	nuclear pore complex	[9]
PMID:21952135	617	G-protein interactome	[10]
PMID:18642946	658	two-component signalling complex	[11]
PMID:15781858	722	TALE homeodomain	[12]
PMID:15805477	989	MADS-box genes	[13]
PMID:17360592	1607	calmodulin-related proteins	[14]
PMID:20706207	1778	cell cycle	[15]

1. Van Leene J, Stals H, Eeckhout D, Persiau G, Van De Slijke E, Van Isterdael G, De Clercq A, Bonnet E, Laukens K, Remmerie N *et al*: **A tandem affinity purification-based technology platform to study the cell cycle interactome in Arabidopsis thaliana**. *Mol Cell Proteomics* 2007, **6**(7):1226-1238.
2. Ceserani T, Trofka A, Gandotra N, Nelson T: **VH1/BRL2 receptor-like kinase interacts with vascular-specific adaptor proteins VIT and VIK to influence leaf venation**. *Plant J* 2009, **57**(6):1000-1014.
3. Uno Y, Rodriguez Milla MA, Maher E, Cushman JC: **Identification of proteins that interact with catalytically active calcium-dependent protein kinases from Arabidopsis**. *Mol Genet Genomics* 2009, **281**(4):375-390.
4. Lawit SJ, O'Grady K, Gurley WB, Czarnecka-Verner E: **Yeast two-hybrid map of Arabidopsis TFIID**. *Plant Mol Biol* 2007, **64**(1-2):73-87.
5. Chang IF, Curran A, Woolsey R, Quilici D, Cushman JC, Mittler R, Harmon A, Harper JF: **Proteomic profiling of tandem affinity purified 14-3-3 protein complexes in Arabidopsis thaliana**. *Proteomics* 2009, **9**(11):2967-2985.
6. Takahashi N, Kuroda H, Kuromori T, Hirayama T, Seki M, Shinozaki K, Shimada H, Matsui M: **Expression and interaction analysis of Arabidopsis Skp1-related genes**. *Plant Cell Physiol* 2004, **45**(1):83-91.
7. Risseeuw EP, Daskalchuk TE, Banks TW, Liu E, Cotelesage J, Hellmann H, Estelle M, Somers DE, Crosby WL: **Protein interaction analysis of SCF ubiquitin E3 ligase subunits from Arabidopsis**. *Plant J* 2003, **34**(6):753-767.
8. Boruc J, Van den Daele H, Hollunder J, Rombauts S, Mylle E, Hilson P, Inze D, De Veylder L, Russinova E: **Functional modules in the Arabidopsis core cell cycle binary protein-protein interaction network**. *Plant Cell* 2010, **22**(4):1264-1280.
9. Tamura K, Fukao Y, Iwamoto M, Haraguchi T, Hara-Nishimura I: **Identification and characterization of nuclear pore complex components in Arabidopsis thaliana**. *Plant Cell* 2010, **22**(12):4084-4097.

10. Klopffleisch K, Phan N, Augustin K, Bayne RS, Booker KS, Botella JR, Carpita NC, Carr T, Chen JG, Cooke TR *et al*: **Arabidopsis G-protein interactome reveals connections to cell wall carbohydrates and morphogenesis**. *Mol Syst Biol* 2011, **7**:532.
11. Dortay H, Gruhn N, Pfeifer A, Schwerdtner M, Schmulling T, Heyl A: **Toward an interaction map of the two-component signaling pathway of Arabidopsis thaliana**. *J Proteome Res* 2008, **7**(9):3649-3660.
12. Hackbusch J, Richter K, Muller J, Salamini F, Uhrig JF: **A central role of Arabidopsis thaliana ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins**. *Proc Natl Acad Sci U S A* 2005, **102**(13):4908-4912.
13. de Folter S, Immink RG, Kieffer M, Parenicova L, Henz SR, Weigel D, Busscher M, Kooiker M, Colombo L, Kater MM *et al*: **Comprehensive interaction map of the Arabidopsis MADS Box transcription factors**. *Plant Cell* 2005, **17**(5):1424-1433.
14. Popescu SC, Popescu GV, Bachan S, Zhang Z, Seay M, Gerstein M, Snyder M, Dinesh-Kumar SP: **Differential binding of calmodulin-related proteins to their targets revealed through high-density Arabidopsis protein microarrays**. *Proc Natl Acad Sci U S A* 2007, **104**(11):4730-4735.
15. Van Leene J, Hollunder J, Eeckhout D, Persiau G, Van De Slijke E, Stals H, Van Isterdael G, Verkest A, Neirynck S, Buffel Y *et al*: **Targeted interactomics reveals a complex core cell cycle machinery in Arabidopsis thaliana**. *Mol Syst Biol* 2010, **6**:397.

Supplemental Table 2. GO enrichment results for the network presented in Figure 3. (FDR: Bonferroni-corrected p-value)

GO ID	GO description	Type	FDR	Enrichment
GO:0009913	epidermal cell differentiation	BP	1.61E-23	46.702
GO:0008544	epidermis development	BP	1.99E-23	46.168
GO:0010026	trichome differentiation	BP	3.77E-19	79.933
GO:0045165	cell fate commitment	BP	8.64E-18	117.558
GO:0009888	tissue development	BP	5.33E-17	20.611
GO:0030154	cell differentiation	BP	1.55E-16	19.422
GO:0001071	nucleic acid binding transcription factor activity	MF	2.63E-14	6.655
GO:0003700	sequence-specific DNA binding transcription factor activity	MF	2.63E-14	6.655
GO:0048869	cellular developmental process	BP	4.94E-14	14.051
GO:0031326	regulation of cellular biosynthetic process	BP	2.55E-12	5.25
GO:0009889	regulation of biosynthetic process	BP	2.73E-12	5.236
GO:0010468	regulation of gene expression	BP	6.87E-12	5.043
GO:0048513	organ development	BP	1.11E-11	9.2
GO:0048731	system development	BP	1.18E-11	9.17
GO:2001141	regulation of RNA biosynthetic process	BP	1.41E-11	5.281
GO:0051252	regulation of RNA metabolic process	BP	1.56E-11	5.258
GO:0080090	regulation of primary metabolic process	BP	1.57E-11	4.877
GO:0003677	DNA binding	MF	2.11E-11	5.423
GO:0010556	regulation of macromolecule biosynthetic process	BP	2.27E-11	5.176
GO:2000112	regulation of cellular macromolecule biosynthetic process	BP	2.27E-11	5.176
GO:0019219	regulation of nucleobase-containing compound metabolic process	BP	3.89E-11	5.059
GO:0006351	transcription, DNA-dependent	BP	4.19E-11	5.043
GO:0051171	regulation of nitrogen compound metabolic process	BP	4.71E-11	5.019
GO:0032774	RNA biosynthetic process	BP	4.86E-11	5.012
GO:0001708	cell fate specification	BP	1.66E-10	120.846
GO:0048856	anatomical structure development	BP	2.01E-07	4.903
GO:0009813	flavonoid biosynthetic process	BP	4.65E-07	41.893
GO:0009962	regulation of flavonoid biosynthetic process	BP	1.12E-06	112.214
GO:0009812	flavonoid metabolic process	BP	1.46E-06	35.705
GO:0010090	trichome morphogenesis	BP	3.79E-06	48.966
GO:0007275	multicellular organismal development	BP	4.00E-06	4.154
GO:0048468	cell development	BP	5.43E-06	16.031
GO:1900376	regulation of secondary metabolite biosynthetic process	BP	5.76E-06	83.122
GO:0032501	multicellular organismal process	BP	9.08E-06	3.967
GO:2000762	regulation of phenylpropanoid metabolic process	BP	1.01E-05	74.81
GO:0009653	anatomical structure morphogenesis	BP	1.27E-05	8.377
GO:0032502	developmental process	BP	1.79E-05	3.817

GO:0009699	phenylpropanoid biosynthetic process	BP	7.03E-05	20.536
GO:0010053	root epidermal cell differentiation	BP	7.19E-05	30.26
GO:0043455	regulation of secondary metabolic process	BP	8.47E-05	49.873
GO:0009718	anthocyanin biosynthetic process	BP	0.0003217	78.062
GO:0009698	phenylpropanoid metabolic process	BP	0.0003971	15.949
GO:0044550	secondary metabolite biosynthetic process	BP	0.0004253	15.789
GO:0010054	trichoblast differentiation	BP	0.0015763	28.054
GO:0000904	cell morphogenesis involved in differentiation	BP	0.0022587	16.938
GO:0010015	root morphogenesis	BP	0.0022587	16.938
GO:0046283	anthocyanin metabolic process	BP	0.002347	48.525
GO:0009753	response to jasmonic acid stimulus	BP	0.003949	15.389
GO:0019438	aromatic compound biosynthetic process	BP	0.0042171	11.221
GO:0007267	cell-cell signaling	BP	0.006239	38.201
GO:0019748	secondary metabolic process	BP	0.010822	9.728
GO:0022603	regulation of anatomical structure morphogenesis	BP	0.0136005	31.499
GO:0031537	regulation of anthocyanin metabolic process	BP	0.0173762	70.872
GO:0010091	trichome branching	BP	0.0408675	53.863

Supplemental Table 3. Illustrative examples of event extraction from text, randomly chosen from the interactions depicted in Figure 3. Each event can have multiple arguments of a specific type, such as T(heme) and C(ause), expressing the subject and the object of the interaction, respectively.

	PMID/PMC	Extracted data	Confidence	Manual evaluation
1	PM 18434419	Binding(T: GL1, T: ttg2)	Very high	Correct
	In vivo binding of the R2R3-MYB protein GLABRA1 (GL1) to the promoters of GLABRA2 (GL2), TRANSPARENT TESTA GLABRA2 (TTG2), CAPRICE (CPC) and ENHANCER OF TRIPTYCHON AND CAPRICE1 (ETC1) establishes that these genes are major transcriptional targets for the TTG1-bHLH-MYB regulatory complex.			
2	PM 18434419	Negative_regulation(T: TTG1) + Regulation(T: GL3)	Average + Low	Correct individual events – complex event not extracted
	The loss of TTG1 and GL1 through mutation, affects the subcellular distribution of GL3.			
3	PMC 2492867	Positive_regulation(C: GL3, T: Gene_expression(T: GL2))	Very high	Correct
	Based on the results of yeast-two-hybrid interaction assays, it has been proposed that TTG1, GL1 or WER, and GL3 or EGL3 form an activator complex to induce GL2 expression [16].			
4	PMC 2929091	Positive_regulation(C: SPL9, T: Transcription(T: TRY))	Very high	Correct
	A significant increase of TRY and TCL1 transcript abundance was observed after 5 h (Figure 4A), suggesting that SPL9 rapidly activated the transcription of TRY and TCL1.			
5	PMC 2492867	Positive_regulation(C: GL1, T: TCL1)	Very high	Negation not extracted
	By using Arabidopsis protoplast transfection assays, we found that all single-repeat R3 MYBs examined interact with GL3, and that GL1 or WER and GL3 or EGL3 are required and sufficient to activate the transcription of TRY, CPC, ETC1 and ETC3, but not TCL1 and ETC2.			
6	PM 15310821	Binding(T: LHY, T: LKP2) + Negation	Very high	Negation not extracted
	The two-hybrid analysis also demonstrated that LKP2 interacted with TOC1, a clock component, but not with CCA1 or LHY, negative regulators of TOC1 gene expression.			
7	PM 16500988	Positive_regulation(C: PIF4, T: Regulation(T: SHB1)) + Speculation	Average	Correct – speculation extracted
	PIF4 appears to specifically mediate SHB1 regulation of hypocotyl elongation and CHLOROPHYLL a/b BINDING PROTEIN3 or CHALCONE SYNTHASE expression under red light.			

Supplemental Table 4. An in-depth analysis of the low confidence text mining events appearing in Figure 6 (dashed edges).

	Event	PMID/PMC	Confidence	Manual evaluation
1	AS2 regulates KAN1	PM18849474	-1.12306	correct
	We also show that the abaxial expression of KAN1 is mediated directly or indirectly by AS2.			
2	KAN1 regulates AS2	PM18849474	-0.86297	correct
	KANADI1 regulates adaxial-abaxial polarity in Arabidopsis by directly repressing the transcription of ASYMMETRIC LEAVES2.			
3	AS1 regulates YAB5 AS2 regulates YAB5	PMC2621207	-1.03652	speculation not extracted
	The expression of YAB5 was recently suggested to be negatively modulated in the adaxial domain by AS1 and AS2, two genes critical for the development of properly expanded leaves [68].			
4	AS2 regulates KNAT2	PMC2993911	-1.01328	corrct
	AS2 is required to prevent expression of the class I KNOX homeobox genes BREVIPEDICELLUS (BP), *KNAT2, and KNAT6 in the leaf (Ori et al., 2000; Semiarti et al., 2001; Lin et al., 2003).*			
5	AS1 – GTE6 binding	PM16166385	-1.43169	correct
	Using chromatin immunoprecipitation (ChIP) assays, we show that GTE6 is associated with the promoter and the start of the transcribed region of AS1 and up-regulates expression of AS1 through acetylation of histones H3 and H4. *The bromodomain protein GTE6 controls leaf development in Arabidopsis by histone acetylation at ASYMMETRIC LEAVES1.*			
6	GTE6 regulates AS1	PM16166385	-1.06718	correct
	GTE6 positively regulates the expression of ASYMMETRIC LEAVES1 (AS1), which encodes a myb-domain protein that controls proximodistal patterning of leaves.			
7	GTE6 regulates AS1	PM16166385	-1.33552	correct
	Genetic studies demonstrated that AS1 is epistatic to GTE6, indicating that GTE6 regulates AS1 during leaf morphogenesis. *GTE6 was shown to positively regulate the myb domain gene ASYMMETRIC LEAF1 (AS1), which is involved in leaf axis specification in mature leaves.*			
8	CUC1 – CUC3 binding	PM12837947	-1.08633	not a binding event
	This gene, CUP-SHAPED COTYLEDON3 (CUC3), encodes a putative NAC-domain transcription factor that is homologous with CUC1 and CUC2.			

Supplemental Table 5. Overview of the 26 candidate genes associated to core CC genes, based solely on text mining. When the claim was wrongly extracted from literature, the note "no valid text mining evidence" is displayed, and such findings should be regarded as false positive predictions.

	Gene	**Gene symbols**	**Connected to core CC gene**	**Literature evidence**
1	AT4G32551	RON2, LEUNIG	HEN3 CDKE;1	The transcription corepressor **LEUNIG** interacts with the histone deacetylase HDA19 and mediator components MED14 (SWP) and CDK8 (**HEN3**) to repress transcription
2	AT1G24590	SOB2	CYCD1;1	**ESR2** directly targeted **CYCLIN D1;1**
3	AT4G02980	ABP1	CYCD3;1 (undirect)	Therefore, both overexpression of **CYCD3.1** and reduction of RBR bypass the cell cycle arrest mediated by inactivation of **ABP1** in root cap stem cells, suggesting that the CYCLIN D/RBR pathway is operating downstream of ABP1
4a	AT4G37650	SHR	CYCD6;1	**SHR** has been demonstrated to directly regulate the expression of genes including (...) the D-type cyclin, **CYCD6;1**
4b	AT4G37650	SHR	RBR1	(...) SCARECROW and **SHR** (...) also regulate the transcription of a set of cell cycle regulatory proteins, including **RETINOBLASTOMA RELATED** (...).
5	AT2G35670	FIS2	RBR1	We thus concluded that **RBR1** is necessary for **FIS2** expression in the central cell.
6	AT4G25530	HDG6, FWA	RBR1	Hence, both MSI1 and **RBR1** are required for **FWA** and FIS2 expression.
7	AT2G19520	FVE, NFC4	RBR1 (unspecific)	It is noteworthy that **FVE** can directly interact with a **plant retinoblastoma protein**
8	AT4G29910	ORC5	NOT to RBR1 (negation)	Therefore, all these data together led us to conclude that expression of all AtORC genes, **except AtORC5**, is regulated by the **RBR**/E2F pathway
9	AT1G02580	SDG5, MEA	NOT to RBR1 (negation)	We thus concluded that **MEA** expression in the central cell is **not** regulated by MSI1 and **RBR1**
10	AT2G26430	RCY1	-	no valid text mining evidence

11	AT4G39400	DWF2	-	no valid text mining evidence
12	AT2G33880	WOX9A	-	no valid text mining evidence
13	AT3G59420	CR4	-	no valid text mining evidence
14	AT3G19590	BUB3.1	-	no valid text mining evidence
15	AT3G15970	NUP50	-	no valid text mining evidence
16	AT4G14110	FUS7, COP9	-	no valid text mining evidence
17	AT5G51600	PLE	-	no valid text mining evidence
18	AT2G46410	CPC	-	no valid text mining evidence
19	AT5G04340	ZAT6	-	no valid text mining evidence
20	AT1G24290		-	no valid text mining evidence
21	AT5G20730	TIR5, ARF7	-	no valid text mining evidence
22	AT1G75950	UIP1, ASK1	-	no valid text mining evidence
23	AT4G23713	MIR319A	-	no valid text mining evidence
24	AT5G10380	RING1	-	no valid text mining evidence
25	AT1G72930	TIR	-	no valid text mining evidence
26	AT2G47980	SCC3	-	no valid text mining evidence