

Supplementary Material

Dutt, S., *et al.* 2010. *In silico* analysis of sequential, structural and functional diversity of wheat cystatins and its implication in plant defense. *Genomics Proteomics Bioinformatics* 8: 42-56.

DOI: 10.1016/S1672-0229(10)60005-8

Tables S1-S5

Table S1 TargetP prediction results of cystatins

Cystatin	Length	cTP	mTP	SP	other	Loc	RC	CS
WC1AB038392	142	0.000	0.356	0.960	0.039	S	2	28
WC2AB038395	78	0.037	0.415	0.179	0.605	O	5	-
WC3AB038394	125	0.089	0.485	0.036	0.475	M	5	-
WC4AB038393	142	0.002	0.559	0.925	0.008	S	4	27
WC5AF364099	128	0.010	0.041	0.983	0.033	S	1	22
WCMDAB223039	243	0.014	0.175	0.897	0.008	S	2	37
OCIOs01g58890	140	0.002	0.603	0.894	0.006	S	4	26
OCHIOs05g41460	151	0.793	0.338	0.083	0.009	C	3	15
OCHIOs05g33880	184	0.006	0.652	0.649	0.004	M	5	35
OCIVOs01g68660	158	0.005	0.115	0.957	0.005	S	1	24
OCVOs01g68670	148	0.001	0.221	0.994	0.006	S	2	25
OCVIOs03g11180	113	0.077	0.353	0.385	0.029	S	5	18
OCVIOs03g11170	117	0.005	0.048	0.990	0.030	S	1	24
OCVIII0s03g31510	123	0.010	0.171	0.862	0.019	S	2	19
OCIXOs03g11160	114	0.004	0.048	0.986	0.110	S	1	23
OCXOs04g28250	151	0.033	0.003	0.983	0.032	S	1	22
OCXIOs09g08100	120	0.008	0.064	0.951	0.016	S	1	24
OCXIOs01g16430	250	0.020	0.534	0.692	0.024	S	5	32
icy1AJ536590	107	0.076	0.591	0.043	0.497	M	5	-
icy2AJ748337	140	0.000	0.693	0.937	0.012	S	4	26
icy3AJ748338	179	0.006	0.198	0.928	0.004	S	2	30
icy4AJ748344	243	0.009	0.190	0.898	0.007	S	2	32
icy5AJ748340	151	0.002	0.149	0.980	0.012	S	1	25
icy6AJ748341	124	0.001	0.347	0.938	0.015	S	3	22
icy7AJ748345	124	0.027	0.149	0.620	0.186	S	3	16
cc1AM05530	135	0.003	0.248	0.974	0.003	S	2	27
cc2AM05531	134	0.006	0.130	0.982	0.007	S	1	27
cc3AM05532	97	0.210	0.119	0.064	0.815	O	2	-
cc4AM05533	226	0.012	0.025	0.988	0.012	S	1	26
cc5AM05534	68	0.044	0.394	0.107	0.431	O	5	-
cc6AM05535	116	0.015	0.368	0.854	0.004	S	3	22
cc7AM05536	97	0.372	0.107	0.029	0.840	O	3	-
cc8AM05537	127	0.043	0.017	0.948	0.065	S	1	29
cc9AM05538	78	0.821	0.321	0.052	0.010	C	3	30
cc10AM05539	120	0.136	0.225	0.043	0.811	O	3	-
SB1X87168	130	0.041	0.267	0.790	0.007	S	3	20

Note: cTP, chloroplast transit peptide; mTP, mitochondrial transit peptide; SP, signal peptide; Location: S, secretory; M, mitochondrial; O, any other location; RC, reliability class from 1 to 5, where 1 indicates the strongest prediction; CS, cleavage site as predicted by SignalP.

Table S2 Conserved motifs as determined by MEME program in wheat, rice and barley cystatins

Motif	Width (residues)	Best possible match
1	15	LYEAKVWEK P WENFK
2	11	Q VVAGTLYYLT
3	15	HVQEL A R F A V AEHNK
4	15	ANALLEFERVVKAKQ
5	15	PLVGGWSDIPDVEDN
6	8	ELQEFKPA
7	15	KSIQERSNSLFPYEL
8	15	GGHEPGWRDVPVHDP
9	15	EDFAKFDILMCLKRG
10	15	FTQTQMQSARDKAAM
11	14	MRKHRVVGLVAALL
12	15	KEEK M KA E VHKN L E G
13	14	GDATSFTIADLGAK
14	8	VKEGGAKK
15	11	LFLVGSASLAI

Note: Bold letters represent the signature sequences of cystatins.

Table S3 Different motifs and their functions in the promoter sequences of OCI, OCII, OCXII and icy1

No.	Motifs found	Sequence	Function	OCI	OCII	OCXII	icy1
1	ABRE	GCCGCGTGGC	cis-acting element involved in abscisic acid responsiveness	+	+	+	-
2	Box W1	TTGACC	fungal elicitor responsive element	-	-	+	-
3	CAAT-box	CAAT	common cis-acting element in promoter and enhancer regions	+	+	+	+
4	CAT-box	GCCACT	cis-acting regulatory element related to meristem expression	-	-	+	-
5	CGTCA motif	CGTCA	cis-acting regulatory element involved in MeJA responsiveness	+	+	+	-
6	G-box	GACATGTGGT	cis-acting regulatory element involved in light responsiveness	+	+	-	+
7	GC-motif	CCCCCG	enhancer-like element involved in anoxic specific inducibility	+	-	-	+
8	GCN4 motif	CAAGCCA	cis-acting regulatory element involved in endosperm expression	-	-	-	+
9	GT1 motif	GGTTAA	light responsive element	+	+	-	-
10	LTR	CCGAAA	cis-acting element involved in low-temperature responsiveness	+	-	-	+
11	MBS	CAACTG	MYB binding site involved in drought inducibility	+	+	-	-
12	Motif II b	CCGCCGCGCT	abscisic acid responsive element	+	-	-	-
13	O2-	GTTGACGTGA	cis-acting regulatory element involved in zein metabolism regulation	+	-	+	-
14	Skn-1 motif	GTCAT	cis-acting regulatory element required for endosperm expression	+	+	+	+
15	Sp1	CC(G/A)CCC	light responsive element	+	+	+	+
16	TATA	TATA	core promoter element around -30 of transcription start	+	+	+	+
17	TCA	GAGAAGAATA	cis-acting element involved in salicylic acid responsiveness	-	-	+	-
18	TC-rich repeats	ATTCTCTAAC	cis-acting element involved in defense and stress responsiveness	-	-	-	+
19	TGACG	TGACG	cis-acting regulatory element involved in MeJA responsiveness	+	+	-	+
20	TGA element	AACGAC	auxin responsive element	-	+	-	-

Note: The promoter sequences cover 1,000 bases upstream of the start site. +/- represents presence or absence of a particular motif in a promoter. Motifs identified only on the positive strand are considered.

Table S4 Protein–protein interactions of <math><4\text{\AA}</math> in complexes of WC1 and WC5 with papain compared with steffin–papain complex

Papain residues	N-terminal residues																
	Stefin B			WC1							WC5						
	M6	M7	S8	G9	A10	P11	V47	G48	G49	I50	E39	N40	G41				
Y61		+															
N64				+	+	+											
G65			+	+													
G66		+	+														
Y67		+															
P68	+																
W69	+																
V133			+														
Q135													+				
A137							+								+		
K156								+							+		
V157								+	+				+	+			
D158			+	+				+	+				+	+	+		
H159									+								
A160			+														
C200													+				
First hairpin loop and adjacent residues making contact																	
	Q53	V54	V55	A56	N59		Q93	T94	V95	A96	G97	M99	K101	L84	V85	S86	G87
N18				+						+							+
Q19				+					+	+					+	+	
G20				+	+				+	+	+				+	+	+
S21				+					+			+			+	+	
C22				+			+		+						+	+	
G23				+													
C63	+												+				
N64													+				
Q94								+									
A136		+															
A137		+															
D158		+												+			
S176									+	+							
W177				+	+				+	+					+		
Second hairpin residues																	
	L102	P103	H104				P123							W115			
G108			+														
Q142							+							+			
L143		+															
W177	+																

Note: Only those N-terminal residues of WC1 and WC5 have been listed, which form interactions with the active site cleft of papain. “+” indicates the presence of interaction of less than 4Å between two residues.

Table S5 Values of Φ , ψ and ω of amino acid residues of stefin B, WC1 and WC5 listed in Table S4

Cystatin	Residue	phi	psi	omega
Stefin B	Met7	-62.0	98.5	167.7
	Ser8	-97.5	135.6	-171.4
	Gly9	96.7	10.5	168.2
	Ala10	-88.6	145.8	176.3
	Pro11	-56.4	156.5	173.1
	Gln53	-142.6	141.0	158.5
	Val54	-87.3	112.1	-179.9
	Val55	-118.7	-148.6	176.8
	Ala56	-80.9	51.2	169.3
	Asn59	-88.6	130.0	169.2
	Leu102	-67.5	140.1	168.4
	Pro103	-46.1	-44.5	-173.8
	His104	-67.2	-2.6	-175.6
	WC1	Val47	-67.0	80.4
Gly48		75.7	179.2	160.5
Gly49		-96.1	-162.6	174.2
Ile50		41.7	41.4	174.0
Gln93		-142.9	117.3	163.8
Thr94		-109.1	115.7	-179.1
Val95		-140.4	-128.3	179.6
Ala96		-79.4	72.0	172.4
Gly97		108.6	174.3	-171.4
Met99		-114.0	138.3	179.7
Tyr101		-119.3	112.1	170.2
Pro123		-66.4	-15.4	-166.0
WC5	Gln39	-60.9	-23.9	177.9
	Asn40	-151.1	164.8	178.2
	Gly41	-108.5	175.6	-173.9
	Leu84	-87.7	139.7	-178.8
	Val85	-142.4	-152.1	-176.4
	Ser86	-78.0	55.3	172.2
	Trp115	-80.7	44.8	175.2

Note: As described in the text, every residue showing interaction of less than 4Å attains a different conformation in their respective complexes with papain.