BIOINFORMATION

Supplementary material:

Table 1: Identification of selected ESTs of each gene family by standalone BLAST

EST IDs	No. of selected ESTs after BLAST analysis	EST IDs	No. of selected ESTs after BLAST analysis
17.3 KDa class II heat shock protein (Hsp17.3)			
gi | 107954700	2	gi | 108954720	113
gi | 108954815	4	gi | 108955108	147
gi | 108955133	15	gi | 108955061	148
gi | 108954762	30	gi | 108954987	367
Desiccation protective protein LEA5			
gi | 108954782	36	gi | 108954830	84
gi | 108954989	82	gi | 108954970	84
DnaJ heat shock N-terminal domain containing protein			
gi | 108954994	17		
Hsp70-60 KDachaperonin			
gi | 108954655	22	gi | 108955170	22
gi | 108955006	22		
Type 2 metallothionein			
gi | 108955024	126	gi | 108954648	500
gi | 108955082	126	gi | 108954669	500
gi | 108955151	126	gi | 108955131	500

Table 2: Selection of putative candidate genes with the help of TAIR and WISE2 result

Family Name	Assembled Contigs	TAIR Result			WISE2 Result	
		Arabidopsis Gene Id	E-Value	Description	Exon position	WISE2 Score (in bits)
17KDa Class II Heat Shock Protein	108954762	AT5G12030.1	5.00E-04	At-Heat Shock Protein 17.6 Kda	96-554	215.62
	108954815	AT5G37670.1	2.00E-08	HSP20-like chaperones superfamily protein	55-489	213.46
	108955061	**AT2G29500.1**	**7.00E-25**	**HSP20-like chaperones superfamily protein**	**238-687**	**283.27**
		AT3G46230.1	2.00E-21	At Heat Shock Protein 17.4 Kda	238-687	266.81
	108955108	**AT2G29500.1**	**7.00E-25**	**HSP20-like chaperones superfamily protein**	**238-687**	**283.27**
		AT3G46230.1	2.00E-21	At-Heat Shock Protein 17.4 Kda	238-687	266.81
DnaJ heat shock protein	**108954994**	**AT5G01390.4**	**1.00E-23**	**DnaJ heat shock family protein**	**205-1023**	**409.42**
Type 2 Metallothionein	**108954648**	**AT3G09390.1**	**7.00E-09**	**At Metallothionein-1**	**115-354**	**146.79**
		AT5G02380.1	0.096	Cysteine-rich protein with copper-binding activity	115-354	130.76
	108954669	**AT3G09390.1**	**7.00E-09**	**At Metallothionein-1**	**115-354**	**146.79**
		AT5G02380.1	0.096	Cysteine-rich protein with copper-binding activity	115-354	130.76
	108955024	AT3G15353.1	1.00E-06	Metallothionein, binds to and detoxifies excess copper and other metals	72-266	89.56
	108955082	AT3G15353.1	1.00E-06	Metallothionein, binds to and detoxifies excess copper and other metals	72-266	89.56
	108955131	**AT3G09390.1**	**7.00E-09**	**At Metallothionein-1**	**115-354**	**146.79**
		AT5G02380.1	0.096	Cysteine-rich protein with copper-binding activity	115-354	130.76
	108955151	AT3G15353.1	1.00E-06	Metallothionein, binds to and detoxifies excess copper and other metals	72-266	89.56

Table 3: Best predicted model with their C-Score, TM Score and RMSD value where C-Score is the confidence score for the predicted model, TM-score is a measure of global structural similarity between query and template protein and Root Mean Square Deviation is the RMSD between residues that are structurally aligned by TM-align.

Best predicted model			
Best Model	**C-Score**	**TM Score**	**RMSD**
HSP best Model 1	0.3	0.75 ± 0.10	4.2 ± 2.8Å
DnaJ best Model 1	-3.18	0.36 ± 0.12	14.2 ± 3.8Å
Type 2 metallothionein 1	-2.98	0.38 ± 0.13	10.0 ± 4.6Å

Table 4: Identified best two structural analogs in PDB where Coverage represents the coverage of global structural alignment and is equal to the number of structurally aligned residues divided by length of the query protein.

Top 2 Identified structural analogs in PDB				
Gene Family	**PDB Hit**	**TM-score**	**RMSD**	**Coverage**
HSP	1gmeA	0.864	1.47	0.908
	2byuI	0.613	0.54	0.621
DnaJ	1nltA	0.62	1.7	0.651
	3agzA	0.484	2.78	0.549
Type 2 metallothionein	1sj8A	0.486	3.97	0.914
	2o8rA	0.467	3.78	0.864

Table 5: Best template protein for similar binding sites.

Template proteins with similar binding site					
Gene Family	**Cscore^{LB}**	**PDB Hit**	**TM-score**	**RMSD**	**BS-score**	**Ligand Name**
HSP	0.04	3l1e0	0.493	2.79	1.03	PEPTIDE
DnaJ	0.18	3agyB	0.461	2.73	1.52	PEPTIDE
Type 2 metallothionein	0.01	3hkeB	0.459	4.07	0.57	T13