

Table 1 Suppl. Primers used for expression analysis of selected unigenes, their amplicon size (A.s.), and melting temperature (T_m).

Contig #	Gene name	Forward and reverse primers (5'-3')	A.s.[bp]	T _m [°C]
Contig1	<i>Cacys</i>	CATTCAAATGGGTCAAAACCTTCAAAG GAAGCAACAGGTTGTTTCTGGAAGT	227	61.8
Contig9	<i>CacdK</i>	CAGAGCAGAGGGTGGGTCCACTAC AAAGCAGTTCCAAAGTTATCCAAACCTC	248	61.0
Contig17	<i>Camet</i>	CCCTAGATAATGTCTTGCTGCGGAG ACATTGCAACCATTCAGCTCC	209	61.5
Contig22	<i>Caacy</i>	ACGAAGGTGAAACAGTTGCTGGAAG CAATAACAATGATCACGTCTTAACTCTAGTAC	220	62.6
Contig26	<i>Casae</i>	GTAAATATGACAACATCAAGGCGTCTTG CATTTGCTACAATATAACAACATTTGCTTAAC	239	60.2
Contig32	<i>Cauae</i>	TGTGGGAGATAATCCTACTTTGAGAGAGC CCACTACATCAAAATGTCGTCGGTATG	208	62.7
Contig34	<i>Cauce</i>	TCCATTCTCGGATCAACATGACTTG CATGAAATCTATGTGAATCCTCCCAACTAG	213	62.9
Contig47	<i>Cauti</i>	CCAAGGATGATTTGAGACTCCCCAC CCGTTAGAACATAAGAGTTGGTTTTTCATG	209	61.8
<i>TeaEST_1440</i>	<i>Calea</i>	GCTAAGCTTCTCTCTGCCTTCATCG GCAACTCAGCCACATCTATATCGTCTG	231	62.8
<i>TeaEST_1499</i>	<i>Cauce2</i>	GGAAGCATTTGTCTCGACATTCTGAAAG GAGTAACATTTTTTCAGAAGAGCCACG	245	62.8
Reference gene	<i>Ca26s RNA</i>	TTGGTGAATTCTGCTTCACAATGATAGG AAAGCTCGTTTGATTCTGATTTCCAGTAC	209	60.0

Table 2 Suppl. The *BLASTx* analysis of unigenes.

Unigene ID	Number of ESTs	Length	<i>BLASTx</i> hits	Acc. ID	E-value
Contig1	3	581	cystatin [<i>Camellia sinensis</i>]	ACN85343.1	1.00E-55
Contig2	4	571	FRIGIDA-like protein isoform 1 [<i>Theobroma cacao</i>]	117	1.00E-27
Contig3	2	500	auxin-repressed protein-like protein ARP2 [<i>Manihot esculenta</i>]	AAX84678.1	3.00E-17
Contig4	3	551	no significant hits		
Contig5	2	399	putative RNA binding protein [<i>Nicotiana tabacum</i>]	AAB92518.1	5.00E-18
Contig6	2	268	no significant hits		
Contig7	2	187	no significant hits		
Contig8	2	217	no significant hits		
Contig9	2	647	cyclin-dependent kinase regulatory subunit 1-like [<i>Setaria italica</i>]	XP_004985721.1	1.00E-45
Contig10	2	502	maturase, partial (mitochondrion) [<i>Ginkgo biloba</i>]	AAF14714.1	2.00E-06
Contig11	3	288	no significant hits		
Contig12	2	210	no significant hits		
Contig13	2	319	no significant hits		
Contig14	2	334	β -amylase [<i>Actinidia arguta</i>]	AFO84078.1	2.00E-11
Contig15	2	232	no significant hits		
Contig16	2	280	hypothetical protein POPTR_0001s18640g [<i>Populus trichocarpa</i>]	XP_002299719.2	2.00E-26
Contig17	2	620	metallothionin 2 [<i>Camellia sinensis</i>]	ABD97258.1	9.00E-26
Contig18	2	398	60S ribosomal protein L40-1 [<i>Arabidopsis thaliana</i>]	NP_565836.1	3.00E-05
Contig19	3	356	no significant hits		
Contig20	3	370	protein transporter, Pam16 [<i>Arabidopsis thaliana</i>]	NP_568943.1	1.00E-07
Contig21	28	663	hypothetical protein OCS_05324 [<i>Ophiocordyceps sinensis</i> CO18]	EQK98963.1	2.00E-13
Contig22	4	417	acyl-CoA-binding protein [<i>Camellia sinensis</i>]	AEC10987.1	3.00E-23
Contig23	2	439	pollen-specific protein SF3-like [<i>Fragaria vesca</i> subsp. <i>vesca</i>]	XP_004299580.1	4.00E-20
Contig24	2	374	no significant hits		

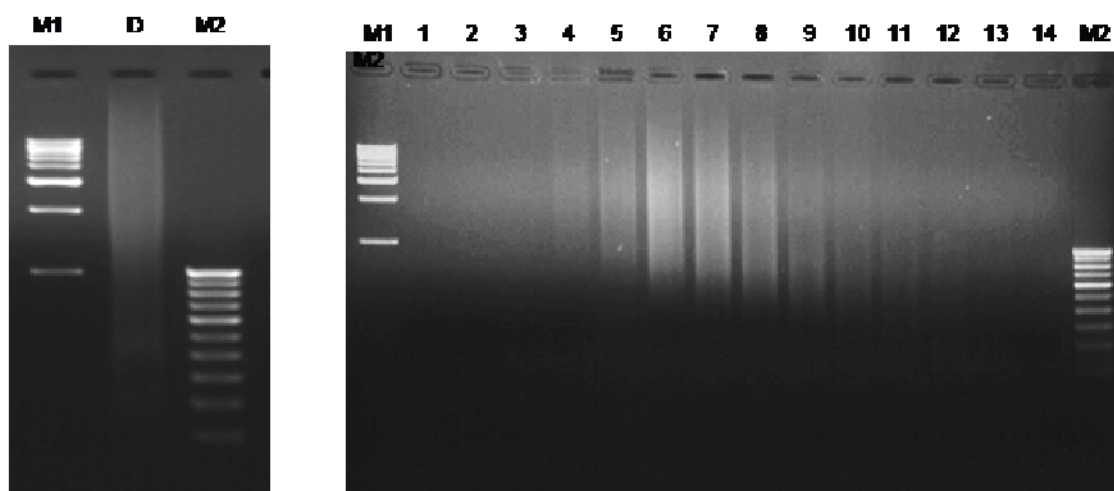


Fig 1 Suppl.. The *left panel*: Size ranges of double stranded cDNA synthesized through long distance (LD)-PCR. The *right panel*: The size fractionations of cDNAs through a chromaspin 400 column (D - LD-PCR products, 1 to 14 - serial numbers of fractions collected in corresponding tubes; M1 - 1 kb DNA marker, M2 - 100 bp DNA marker).

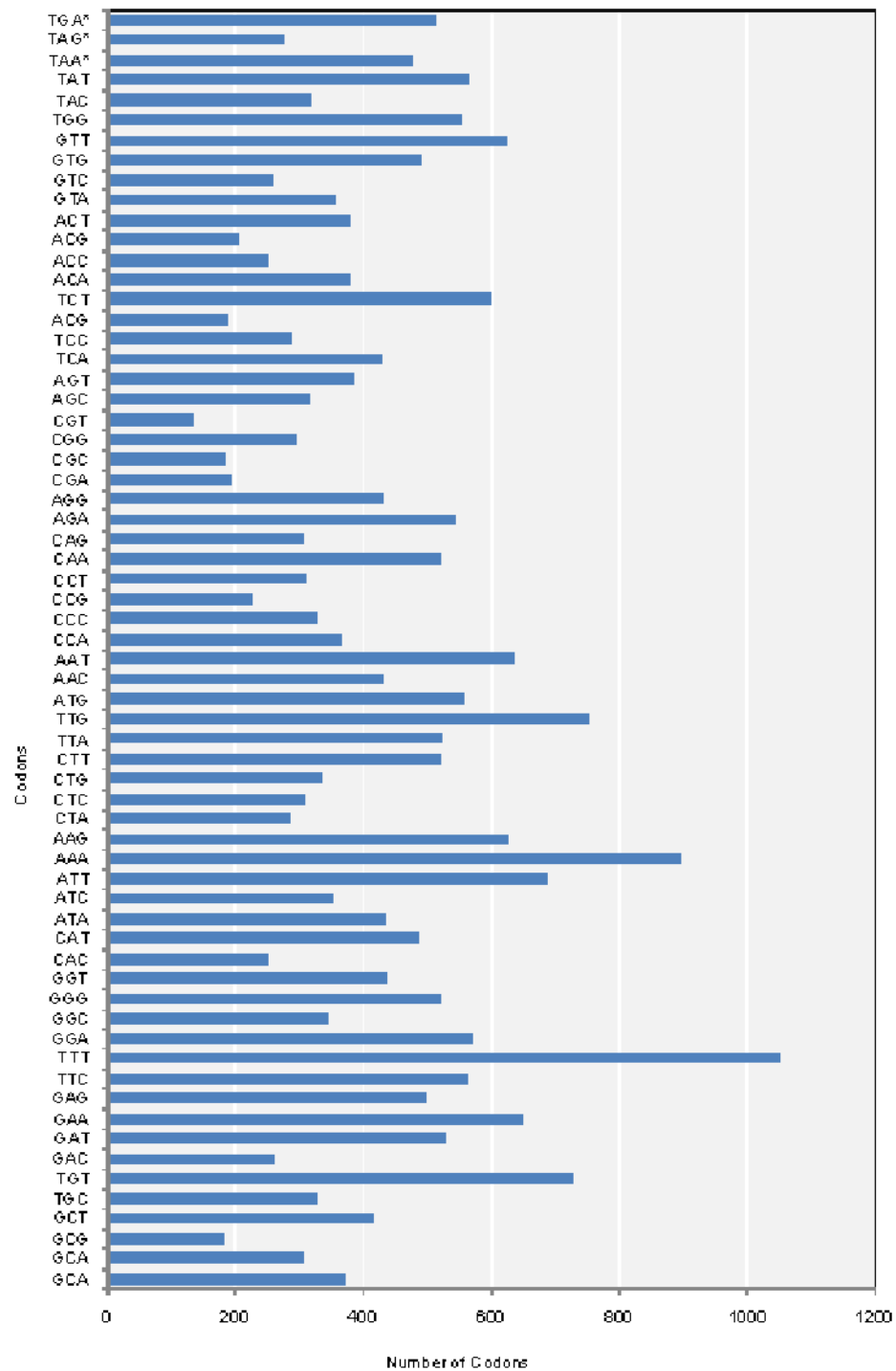


Fig 2 Suppl. The codon usage analysis of standard full-length unigenes.

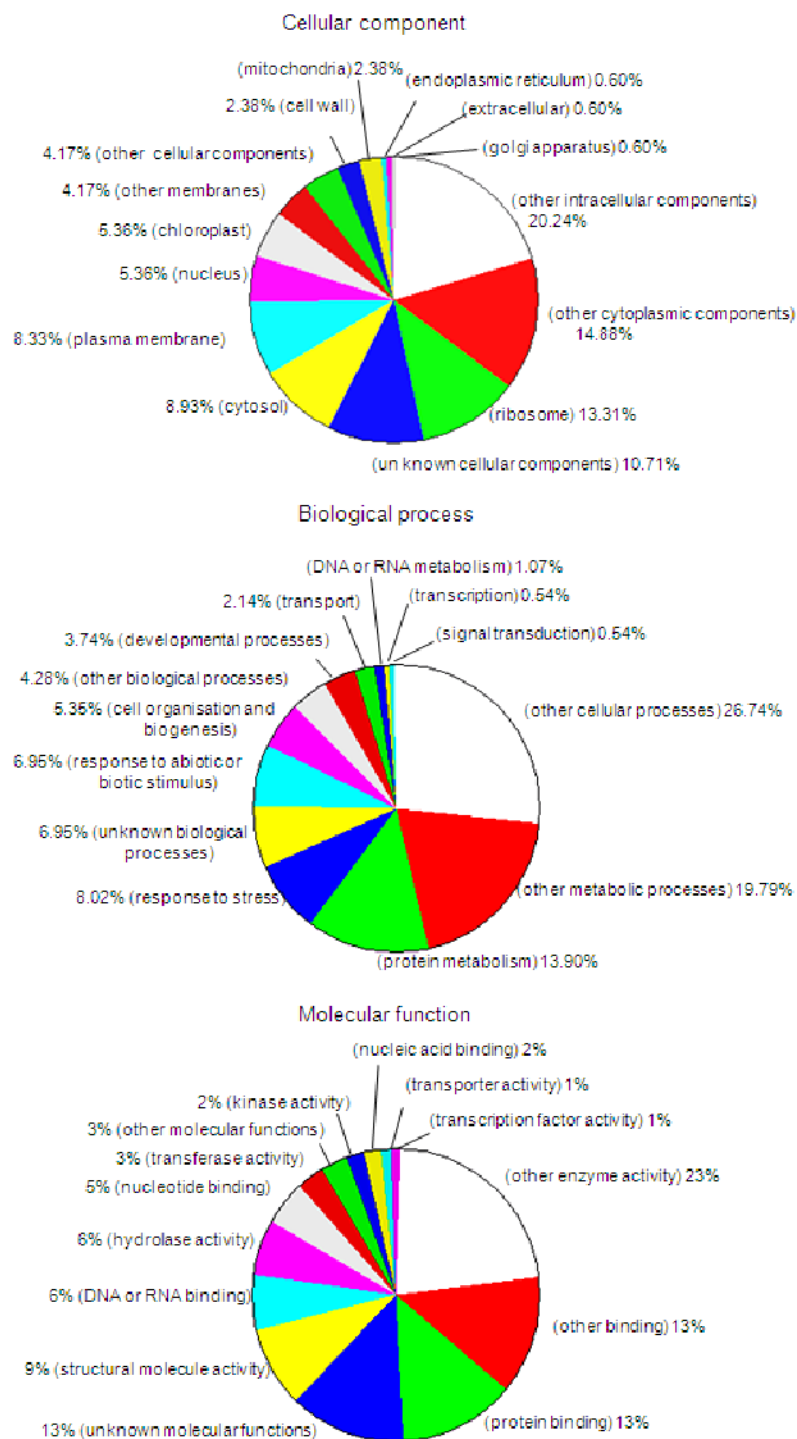


Fig. 3 Suppl. The gene ontology classification of full-length unigenes as defined for the *Arabidopsis* proteome.