OPEN

Transcriptomic and proteomic analysis reveals mechanisms of embryo abortion during chrysanthemum cross breeding

Fengjiao Zhang[1,2], Zhiquan Wang[1], Wen Dong[3], Chunqing Sun[1], Haibin Wang[1], Aiping Song[1], Lizhong He[1], Weimin Fang[1], Fadi Chen[1,2] & Nianjun Teng[1,2]

[1]College of Horticulture, Nanjing Agricultural University, Nanjing 210095, China, [2]Jiangsu Province Engineering Lab for Modern Facility Agriculture Technology & Equipment, Nanjing 210095, China, [3]China Rural Technology Development Center, Beijing 100045, China.

Embryo abortion is the main cause of failure in chrysanthemum cross breeding, and the genes and proteins associated with embryo abortion are poorly understood. Here, we applied RNA sequencing and isobaric tags for relative and absolute quantitation (iTRAQ) to analyse transcriptomic and proteomic profiles of normal and abortive embryos. More than 68,000 annotated unigenes and 700 proteins were obtained from normal and abortive embryos. Functional analysis showed that 140 differentially expressed genes (DEGs) and 41 differentially expressed proteins (DEPs) were involved in embryo abortion. Most DEGs and DEPs associated with cell death, protein degradation, reactive oxygen species scavenging, and stress-response transcriptional factors were significantly up-regulated in abortive embryos relative to normal embryos. In contrast, most genes and proteins related to cell division and expansion, the cytoskeleton, protein synthesis and energy metabolism were significantly down-regulated in abortive embryos. Furthermore, abortive embryos had the highest activity of three executioner caspase-like enzymes. These results indicate that embryo abortion may be related to programmed cell death and the senescence- or death-associated genes or proteins contribute to embryo abortion. This adds to our understanding of embryo abortion and will aid in the cross breeding of chrysanthemum and other crops in the future.

Embryo abortion is very common in plants and has attracted considerable attention in plant biology, as it not only seriously reduces the seed and fruit production of many crops but also negatively affects the efficiency of plant cross breeding[1,2]. To solve this important problem, many researchers have carried out a vast number of studies in the past several decades to investigate factors that cause plant embryo abortion[3,4]. However, most of these studies have mainly focused on morphological, anatomical, and physiological levels due to the limitations of traditional biological techniques. Few studies have been reported that examine the factors contributing to plant embryo abortion at a molecular level. Therefore, the mechanisms underlying plant embryo abortion are poorly understood.

The recent development of modern molecular biological techniques, such as novel high-throughput sequencing technologies including Solexa/Illumina RNA sequencing (RNA-seq) and a new proteomics approach (i.e., isobaric tags for relative and absolute quantitation, iTRAQ) may provide a promising means to address the genes and proteins associated with plant embryo abortion[5–7]. In particular, transcriptome sequencing is a very powerful method that produces functional genomic data for non-model organisms without prior information about the entire genome[8,9]. This novel technique has been extensively applied to many plant species, but surprisingly, few studies on embryo abortion have been reported to date[10,11].

The chrysanthemum (*Chrysanthemum morifolium*) is the second most important flower in the world and plays a vital role in the flower production industry[12]. More importantly, this species is also grown as a medicinal tea and nutritious vegetable[13]. Thus, chrysanthemum is an economically important crop both in China and the rest of the world. However, a large problem exists in chrysanthemum cultivation: most chrysanthemum cultivars are only weakly resistant to biotic and abiotic factors[13,14]. To improve chrysanthemum resistance, breeders often carry out wide crosses between chrysanthemum cultivars and wild chrysanthemum species with high levels of tolerance to biotic and abiotic stresses. However, reproductive barriers often exist in these crosses, and as a consequence, breeding efficiency is considerably reduced[15]. Our previous investigation indicated that embryo abortion is the main reproductive barrier responsible for the failure of these crosses[15,16]. So far, studies have focused on the causes

of embryo abortion, such as maternal genotype, endosperm nutrition[17] and the sensitivity of the endosperm from the parental genome dosage imbalance that enhanced or repressed seed development[18–20], among others. To our knowledge, no study has investigated molecular events occurring in embryo abortion during chrysanthemum cross breeding, or any other crop, and thus little information is available pertaining to the genes and proteins involved. Therefore, it is necessary to examine the components implicated in chrysanthemum embryo abortion at the molecular level, which will increase our understanding of the molecular events involved in embryo abortion of chrysanthemum and other crops.

Transcriptome analysis has been instrumental in identifying differentially expressed genes (DEGs) implicated in various organismal biological processes[21,22]. However, it cannot capture the key proteins implicated in the same processes. Therefore, in addition to transcriptome analysis by RNA-seq, we have also investigated proteins associated with chrysanthemum embryo abortion using iTRAQ, which has mainly been used in proteome analyses of animal and microorganisms[23–25]. The objectives of this study were to identify functional genes and proteins involved in embryo abortion and highlight molecular mechanisms related to chrysanthemum embryo abortion.

Results

Embryo development and abortion. Nearly all ovaries were normal in morphology just before pollination (Fig. 1a, e). At 12 days after pollination (DAP), most chrysanthemum embryos reached the stage of globular embryo and 89.7 ± 3.1% of embryos were normal at this developmental stage (Fig. 1b, f, i). Subsequently, many embryos began to degenerate. At 18 DAP, only 43.3 ± 1.8% of embryos were normal, and most developed into the heart-shaped embryo stage (Fig. 1c, g, j), and the remaining embryos underwent degeneration (Fig. 1d, h, k). These results indicate that the period from globular to heart-shaped embryo stages were the key developmental stage of chrysanthemum embryo abortion in the cross.

High-throughput transcriptome sequencing and read assembly. To study transcriptomic gene expression during embryonic development, three cDNA libraries from normal embryos at 12 DAP (NE12), normal embryos at 18 DAP (NE18) and abortive embryos at 18 DAP (AE18) were subjected to Illumina sequencing (Table 1). After removing invalid reads and data cleaning, we obtained 152,347,304 clean reads containing a total of 13,711,257,360 nucleotides (13.7 Gb) with a mean length of 90 bp. The Q20 percentage (proportion of nucleotides with quality value larger than 20 in reads) and GC (guanine and cytosine) percentage were 97.6% and 43.6%, respectively. After clustering the high-quality reads, we finally obtained 615,312 contigs and 116,697 unigenes. The unigenes with an average length of 743 bp included 52,860 clusters and 63,837 singletons.

In addition, we compared RNA-Seq datasets with the EST datasets produced from vegetative tissues of *Chrysanthemum nankingense*[26]. The results indicate there are 31,674 genes with the same sequence of CDS in both libraries, occupying 27.14% of the 116,697 unigenes in the chrysanthemum embryo transcriptome.

Gene annotation and functional classification. A total of 116,697 unigenes were annotated using Blastx to protein databases and Blastn to nucleotide databases Nt (E-value<0.00001). In the Nr, Nt, SwissProt, KEGG, COG and GO databases, 65,327, 43,427, 43,359, 40,363, 25,297 and 49,535 unigenes were aligned, respectively (Table 2).

To understand the function of unigenes intuitively and effectively, we used the COGs functional annotation system for the assembled unigenes. As a result, 25,297 unigenes have specific protein function definition, accounting for 37.12% of the total annotated unigenes and involving 25 COG functional classes (Fig. S1). The five largest categories were 'general function' (8,167, 16.1%); 'transcription' (4,729, 9.3%); 'Replication, recombination and repair' (4,147, 8.2%); 'Post-translational modification, protein turnover, chaperones' (3,899, 7.7%); 'translation, ribosomal structure, and biogenesis' (3,756, 7.4%). The following two categories, 'nuclear structure' (14, 0.028%) and 'extracellular structures' (35, 0.069%) accounted for the lowest percentages.

We used the GO assignments to classify the functions of the predicted *C. morifolium* genes according to the annotation from Nr database. The results showed that 49,535 unigenes were categorised into 55 functional groups, which could be classified into three main categories (biological process, cellular component and molecular function), and these contain 24, 15 and 16 functional subcategories, respectively (Fig. S2). In the biological processes category, the major subcategories were 'cellular process', 'metabolic process', and 'response to stimulus'. In the two remaining categories, 'cell', 'cell part' and 'catalytic activity', 'binding' were remarkable, respectively. This classification suggests that these biological processes may play a significant role during embryonic development. On the other hand, of these terms, 'channel regulator activity', 'metallochaperone activity', 'translation regulator activity', 'extracellular matrix part' and 'protein tag' were rare (Fig. S2).

To identify biological pathways activated in the embryos of *C. morifolium*, the assembled unigenes were annotated against the KEGG database (E-value<0.00001). A total of 40,363 unigenes were assigned to the 129 KEGG pathways (Table S1). The majority of these pathways were 'metabolic pathways [ko01100]' (9,241 genes), 'biosynthesis of secondary metabolites [ko01110]' (4,736 genes), 'plant-pathogen interaction [ko04626]' (2,346 genes) and 'plant hormone signal transduction [ko04075]' (1,894 genes). In addition, 824 genes were related to 'RNA degradation [ko03018]', 765 genes to 'ubiquitin mediated proteolysis [ko04120]' and 260 genes to 'proteasome [ko03050]', which were closely related to protein degradation and cell death.

DEGs in three samples. In analysis, using FPKM values[27] calculated gene expression levels at the different embryonic developmental stages and DEGs with FDR ≤ 0.001 and |log$_2$(ratio)| larger than 2. To explore gene expression levels in AE18, which is the reference point, variations in gene expression were identified in two comparisons: between AE18 and NE12 and between AE18 and NE18. The number of up- and down-regulated genes is shown in Venn diagrams[28] (http://bioinfogp.cnb.csic.es/tools/venny/index.html). Compared with NE12, there were 4,399 and 3,868 genes up- and down-regulated in AE18, respectively. Compared with NE18, in AE18, there were 1,310 and 1,010 genes were up- and down-regulated, respectively. In AE18, 306 and 178 genes were up- and down-regulated, respectively, compared with the two samples of NE12 and NE18 (Fig. S3; Tables S2 & S3). Changes in differentially expressed genes at different stages indicated that the abundance of transcripts differed between developmental stages of the embryo.

DEGs were selected based on their possible function in the Nr annotation database and the expression level fold-change. In addition, DEGs with similar functions were clustered in the same category. There were a total of 143 DEGs and 10 categories listed in one heat map (Fig. 2; Table S4). The number of genes related to cell senescence and death, energy metabolism and protein synthesis accounted for a high proportion of these categories. Compared with NE12, NE18 and AE18 were significantly changed in the following categories: cytoskeleton, cell division, phytohormone signalling, reactive oxygen species (ROS) scavenger and protein degradation. In addition, transcription regulators, a key factor in regulating gene expression, were changed, including AP2/ERF, WRKY, zinc finger protein and others. Among the three samples of NE12, NE18 and AE 18, AP2/ERF, WRKY, MYB and ethylene-responsive transcription factor trended upwards. Conversely, zinc finger protein and leucine-rich repeat receptor protein kinase trended downward.
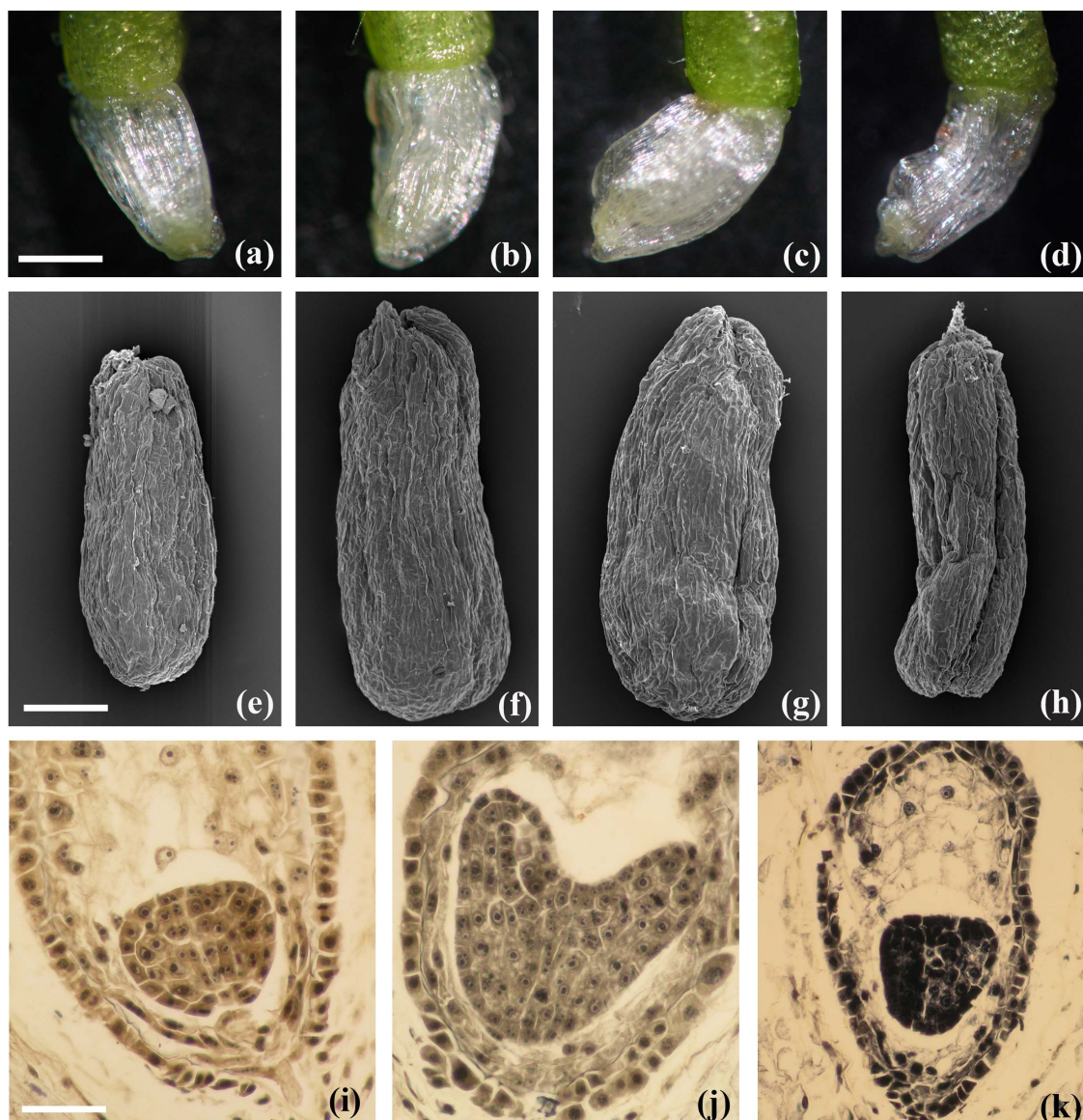
Figure 1 | Morphological and anatomical features of chrysanthemum ovaries and embryos. Chrysanthemum ovaries just before pollination (a, e). Normal chrysanthemum ovaries (b, f) and embryo (i) at 12 days after pollination (DAP). Normal chrysanthemum ovaries (c, g) and embryo (j) at 18 DAP. Abortive chrysanthemum ovaries (d, h) and embryo (k) at 18 DAP. Bars = 400 μm (a–d), 200 μm (e–h), 100 μm (i–k).

Validation of gene expression profiles by qRT-PCR. The quality of the RNA-Seq libraries was further validated experimentally using qRT-PCR. Expression quantification of 14 genes, randomly selected from DEGs, related to cell death and division, phytohormone signalling, late embryogenesis abundant protein, plasma membrane H^+ ATPase and transcription factor, were performed. qRT-PCR results agree with the RNA-Seq data (Fig. 3), confirming the reliability of our sequencing data.

Table 1	Summary for chrysanthemum embryo transcriptome	
Total number of clean reads	152 347 304
Total nucleotide length	13 711 257 360 bp
Average read length	90 bp
Q20 percentage	97.60%
GC percentage	43.60%
Total number of contigs	615 312
Total number of unigenes	116 697
Mean length of unigenes	743 bp

Protein profiles of normal and abortive embryos. Based on iTRAQ-nano-HPLC-MS/MS proteomic investigation, we obtained 715 proteins in protein profiles of the three samples. When AE18 was the references point, there were 41 proteins with significant changes in abundance, 21 of which were up-regulated and 20 of which were down-regulated in NE12 and NE18 (Table 3). Proteins associated with cell senescence and death, embryo sac development arrest, reactive oxygen species scavenger, and protein degradation were significantly up-regulated in AE18 relative to those in NE12 during embryo development. Meanwhile, compared with NE18, levels of these proteins were up-regulated as well, which involved normal and abortive embryos in the same developmental stage. For example, the abundance of two proteins related to cell senescence and death and three proteins associated with protein degradation was significantly higher in AE18 than NE12, and higher than NE18 (Table 3). In contrast, proteins associated with cytoskeleton and ribosome constituent, energy metabolism, as well as protein and amino-acid biosynthesis in AE18 were significantly down-regulated among NE12, NE18, and AE18. For instance, five proteins associated with cytoskeleton and ribosome constituent, eight proteins

Table 2 | Annotation of unigene sequences in chrysanthemum embryos

Sequence database	Number of annotated unigene sequences	Percentage of annotated unigene sequences
Total unigenes	68 157	100
Nr	65 327	95.85
Nt	43 427	63.72
Swiss-Prot	43 359	63.62
KEGG	40 363	59.22
COG	25 297	37.12
GO	49 535	72.68

implicated in energy metabolism (two in tricarboxylic acid cycle and three in glycolysis), and five proteins involved in protein and amino-acid biosynthesis were significantly down-regulated in AE18 (Table 3).

Correlation analysis of transcriptomic and proteomic data. We carried out a correlation analysis between DEPs and transcriptomic results to examine the congruence between transcriptomic and proteomic data. Forty-one DEPs were identified in the proteomic analysis, 39 of which (95.1%) were found to have corresponding genes in the transcriptomic data. The corresponding genes of 'Cysteine synthase-like [*Glycine max*]' and 'Omega-3 fatty acid

desaturase, endoplasmic reticulum [*Arabidopsis thaliana*]' were not found in the transcriptomic data (Table 3), indicating that transcriptomic and proteomic data are of high identity match. The corresponding genes of three DEPs, 'Pathogenesis-related protein [*Zinnia violacea*]', 'Pyruvate kinase [*A. thaliana*]', and 'Eukaryotic initiation factor 4A [*Oryza sativa* Japonica Group]', were also screened in the list of DEGs (Fig. 2; Table S4). However, the corresponding genes of the other 36 DEPs were not listed in Fig. 2. The reason for this is that the change in expression ratios of those genes was less than two fold. In the comparison between AE18 and NE 12, the expression trend of 21 DEPs (53.8%) was consistent with that of the corresponding genes, whereas the expression trend of the remaining 18 DEPs was not. Similar results were found in the comparison between AE18 and NE 18, where 24 of 39 DEPs (61.5%) had an expression trend consistent with that of the corresponding genes, but the expression trend of the remaining 15 DEPs was not inconsistent with that of the corresponding genes. Such results demonstrate only a moderate correlation between the mRNA and protein abundance ratios, as transcriptional and translational control appeared rather uncoupled in this biological system.

Caspase-like activities in normal and abortive embryos. AE18 (abortive embryos) has the highest cleavage activities for Ac-LEVD-AMC, Ac-VEID-AMC and Ac-LEHD-AMC compared with NE12 and NE18 (Fig. 4). The results indicate that chrysanthemum
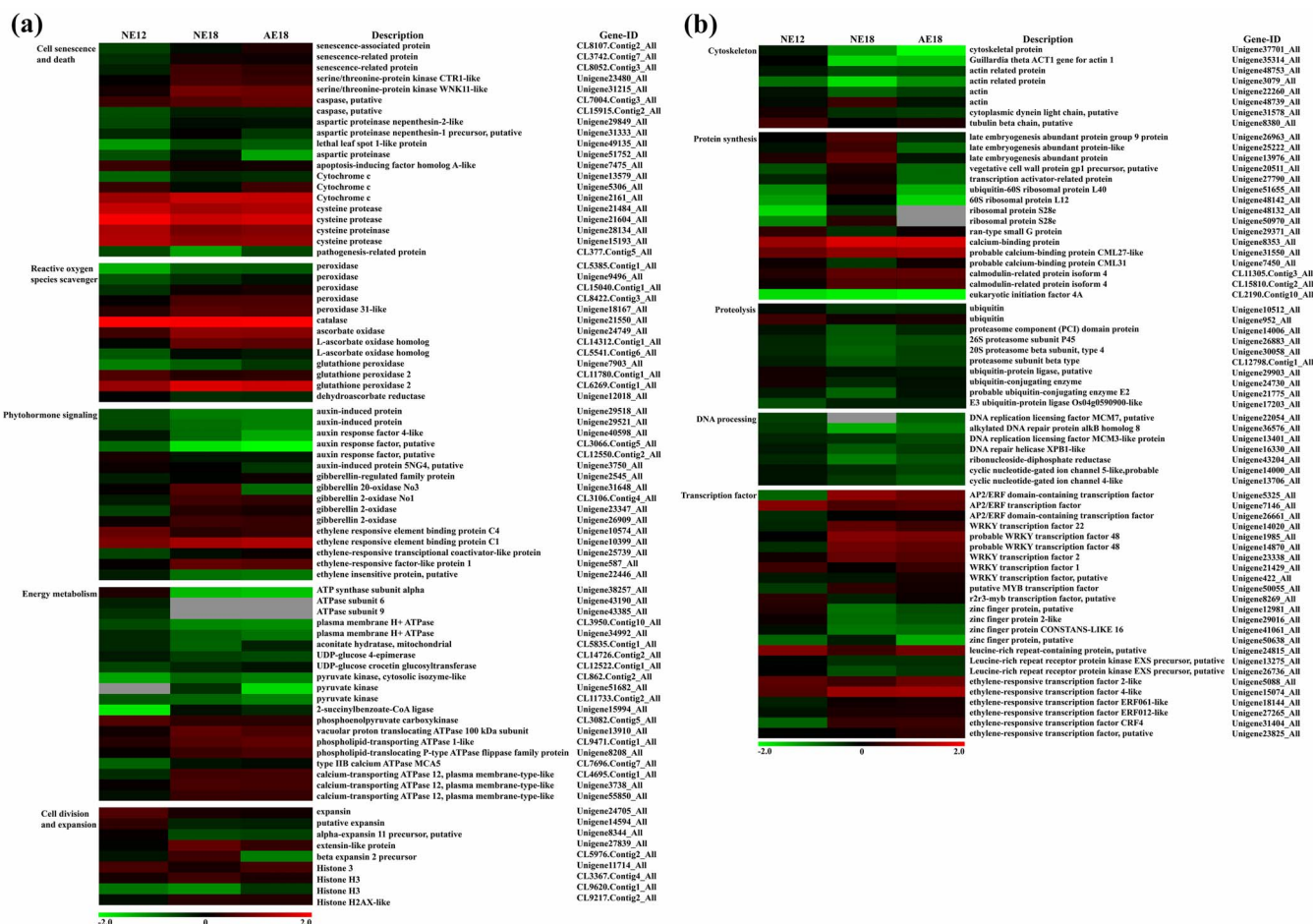

Figure 2 | Heatmap of 143 differentially expressed genes in normal and abortive chrysanthemum embryos. The 143 genes are grouped into 10 main categories and the bar represents the scale of the expression levels for each gene. The bar represents the scale of the expression levels for each gene ($\log_{10}$FPKM (fragments per kb per million fragments)) in the NE12, NE18 and AE18 as indicated by red/green rectangles. Red rectangles mean up-regulation of genes and green means down-regulation. All genes in this list have a P-value for differential expression $< 10^{-5}$ (The P-value was presented in Tables S2 & S3). All information for each gene list can be found in Tables S4.
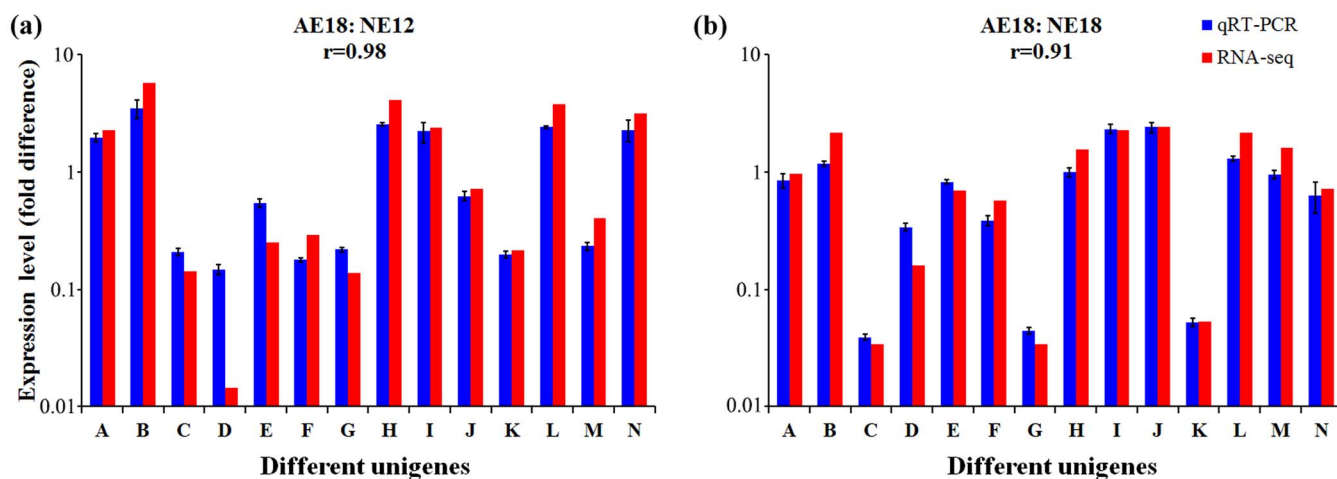
Figure 3 | Validation of RNA-seq results by qRT-PCR. Correlation between the expression profiles of the 14 transcripts was determined by RNA-Seq (red bars) and qRT-PCR (blue bars). The 14 points (A–N) from left to right on the x-axis represent genes encoding caspase putative, senescence-associated protein, beta expansin 2 precursor, cytoskeletal protein, plasma membrane H⁺ ATPase, cyclic nucleotide-gated ion channel 4-like, gibberellin 20-oxidase No3, ethylene-responsive transcriptional coactivator-like protein, ethylene-responsive transcription factor putative, proteasome component (PCI) domain protein, late embryogenesis abundant protein-like, glutathione peroxidase, leucine-rich repeat receptor protein kinase EXS precursor putative and WRKY transcription factor 2, respectively.

embryo abortion is may be attributed to programmed cell death, and the corresponding three executioner caspase (LEVDase, VEIDase and LEHDase) played a vital role in embryo abortion. It is interesting that NE12 has the highest activity of DEVDase among the three samples. A possible reason is that programmed cell death often occurs in developing embryos to eliminate unwanted cells, such as suspensor cells, and DEVDase plays a key role in this process.

Discussion

Genes and proteins associated with programmed cell death during embryo abortion. Plant embryo abortion is a very complex bioprocess during which many events work collectively[11,29]. Among those events, cell senescence and death are often considered the most important that occur during plant embryo abortion. In plants, cell death is a physiological process required to remove unwanted or damaged cells, and usually has two types, programmed cell death (PCD) and necrosis[30]. PCD commonly occurs in flowering and reproduction, seed and embryo development and development of leaf, stem and root[29–31]. It has several key diagnostic features including leakage of cytochrome c from mitochondria into the cytosol, nuclear DNA fragmentation, increase in caspase-like and metacaspase activities, and increase in ROS[30,32–34]. In this study, we found that many genes or proteins associated with PCD features were significantly up-regulated in abortive embryos relative to normal embryos (NE12 and NE18), such as caspase-putative, aspartic proteinase, cysteine protease, cysteine synthase-like and cytochrome c. In addition, the expression levels of some ROS scavengers were much higher in abortive embryos than normal embryos, indicating that higher levels of ROS were produced in abortive embryos. Moreover, several genes encoding senescence-related proteins were also dramatically up-regulated in abortive embryos. Similarly, two senescence-associated genes were once identified and considered to potentially contribute to peanut embryo abortion[11]. Furthermore, among the three samples of NE12, NE18 and AE 18, caspase activity assays showed that AE18 has the highest activities of three executioner caspase (caspase 4, caspase 6, caspase 9). The results presented here demonstrate that chrysanthemum embryo abortion during cross breeding is possibly a result of PCD and the senescence-related genes is a key trigger of PCD.

Genes and proteins related to phytohormone are implicated in embryo abortion. Change in phytohormone concentration is an important event associated with embryo abortion. For example, the authors found that the level of IAA was much lower in abortive watermelon embryos than normal embryos, whereas ABA concentration was significantly higher in abortive embryos[35]. Similar results were also found in Chinese white poplar and *Litchi Chinensis*[36]. Except for IAA and ABA, ethylene production also plays an important role in embryo abortion, as increased ethylene production has been related to accelerated senescence, reduced growth, and responses to environmental stresses. For instance, Hays *et al.* reported that high temperature treatment resulted in kernel abortion of heat-susceptible winter wheat (*Triticum aestivum*). Enhanced ethylene production induced by heat stress was responsible for its kernel abortion, as treatment with the ethylene receptor inhibitor 1-methylcyclopropane (1-MCP) prior to exposure to heat stress blocked kernel abortion[37]. Those previous studies established a link between phytohormone and embryo abortion at a physiological level, but the underlying mechanism was not clear. In the current work, some auxin-related genes or proteins, such as auxin-induced protein and auxin response factor, were significantly down-regulated in abortive chrysanthemum embryos, which is consistent with down-regulation of several genes related to cell division and expansion. In contrast, some ethylene-associated genes or proteins (e.g., ethylene responsive factor and ERF transcription factor) were significantly up-regulated in abortive embryos. Such results suggest that phytohormone was implicated in chrysanthemum embryo abortion and ethylene may also be a factor causing embryo abortion.

Genes and proteins associated with energy metabolism. Energy metabolism is the process of the energy associated with the release, transfer and utilisation occurring in mitochondrion, the key function of which is the generation of adenosine triphosphate (ATP), the major energy source of the cell[38]. The rate of living theory, a classic theory of aging, proposed the existence of a link between energy metabolism and aging. The central thought of the theory is that life span is determined by the rate at which it produces and uses energy at the cellular level[39]. In *Pinus bungeana*, authors found that the abundance of several proteins associated with energy production

5

Table 3 | Differentially expressed proteins in normal and abortive chrysanthemum embryos

Protein (taxonomy)	AE18:NE12	AE18:NE18	NE18:NE12	Function/category	Peptide number
Cysteine synthase-like [Glycine max]	1.69 ± 0.11	1.41 ± 0.03	1.20 ± 0.05	Cell senescence and death	2
Pathogenesis-related protein [Zinnia violacea]	2.56 ± 0.10	3.09 ± 0.15	0.83 ± 0.02	Cell senescence and death	2
Selenium-binding protein 2 [Arabidopsis thaliana]	1.51 ± 0.05	1.37 ± 0.04	1.11 ± 0.01	Embryo sac development arrest	5
Superoxide dismutase, chloroplastic [Cu-Zn] [Solidago canadensis var. scabra]	2.51 ± 0.10	2.76 ± 0.10	0.91 ± 0.02	Reactive oxygen species scavenger	5
Ascorbate peroxidase 2-like protein [Tragopogon porrifolius]	2.09 ± 0.04	2.38 ± 0.06	0.88 ± 0.02	Reactive oxygen species scavenger	3
Putative polyubiquitin 2 [Oryza sativa Japonica Group]	1.34 ± 0.05	1.54 ± 0.04	0.87 ± 0.02	Protein degradation	6
Ubiquitin-fold modifier 1 [Arabidopsis thaliana]	1.75 ± 0.08	1.87 ± 0.05	0.93 ± 0.02	Protein degradation	3
Proteasome subunit alpha type, putative [Ricinus communis]	2.20 ± 0.10	2.05 ± 0.04	1.07 ± 0.03	Protein degradation	3
Calmodulin-7 [Arabidopsis thaliana]	2.55 ± 0.13	2.40 ± 0.09	1.06 ± 0.02	Calcium ion binding	9
Polcalcin [Artemisia vulgaris]	1.74 ± 0.09	1.93 ± 0.04	0.90 ± 0.03	Calcium ion binding	7
Cyclophilin, putative [Ricinus communis]	1.75 ± 0.07	1.79 ± 0.06	0.98 ± 0.01	Protein folding	3
putative DNA repair protein RAD23-4 [Arabidopsis thaliana]	1.69 ± 0.06	1.91 ± 0.06	0.89 ± 0.01	DNA repair	5
Ribonuclease E inhibitor RraA/ Dimethylmenaquinone methyltransferase [Arabidopsis thaliana]	1.82 ± 0.13	1.56 ± 0.05	1.16 ± 0.05	RNA metabolic process	4
glycine-rich RNA binding protein [Oryza sativa Japonica Group]	1.52 ± 0.07	1.27 ± 0.05	1.19 ± 0.01	RNA transcription or processing during stress	1
Omega-3 fatty acid desaturase, endoplasmic reticulum [Arabidopsis thaliana]	1.85 ± 0.09	1.40 ± 0.03	1.32 ± 0.04	Fatty acid biosynthesis	2
Translationally-controlled tumor protein homolog [Arabidopsis thaliana]	1.55 ± 0.06	1.43 ± 0.04	1.09 ± 0.02	Microtubule binding	3
Histone 1 [Malus domestica]	1.40 ± 0.11	1.57 ± 0.04	0.89 ± 0.05	Chromosome constituent	1
Histone H2B.6 [Arabidopsis thaliana]	1.33 ± 0.09	1.55 ± 0.06	0.86 ± 0.03	Chromosome constituent	6
Histone H4 [Glycine max]	1.28 ± 0.06	1.60 ± 0.03	0.80 ± 0.02	Chromosome component	5
Prohibitin-2, mitochondrial [Arabidopsis thaliana]	1.40 ± 0.09	1.51 ± 0.04	0.92 ± 0.04	Chaperone for respiration chain proteins	10
Actin-7 [Arabidopsis thaliana]	0.78 ± 0.06	0.76 ± 0.03	1.01 ± 0.05	Cytoskeleton constituent	8
Tubulin alpha-6 chain [Arabidopsis thaliana]	0.63 ± 0.02	0.78 ± 0.02	0.80 ± 0.01	Cytoskeleton constituent	5
Beta-tubulin [Zinnia violacea]	0.67 ± 0.03	0.74 ± 0.03	0.90 ± 0.00	Cytoskeleton constituent	5
40S ribosomal protein S15a-1 [Arabidopsis thaliana]	0.43 ± 0.01	0.58 ± 0.02	0.73 ± 0.00	Ribosome constituent	6
Ribosomal protein S3, component of cytosolic 80S ribosome and 40S small subunit [Ostreococcus lucimarinus (strain CCE9901)]	0.64 ± 0.03	0.75 ± 0.03	0.86 ± 0.00	Ribosome constituent	5
putative ATP synthase beta chain [Oryza sativa (japonica cultivar-group)]	0.66 ± 0.03	0.72 ± 0.02	0.91 ± 0.01	Energy metabolism	3
V-type proton ATPase catalytic subunit A [Daucus carota]	0.82 ± 0.03	0.68 ± 0.0 1	1.20 ± 0.02	Energy metabolism	9
UDP-glucose 6-dehydrogenase [Oryza glaberrima]	0.69 ± 0.03	0.72 ± 0.02	0.95 ± 0.02	Energy metabolism	4
2-oxoglutarate dehydrogenase, E1 subunit-like protein [Arabidopsis thaliana]	0.80 ± 0.06	0.71 ± 0.02	1.11 ± 0.05	Energy metabolism	3
Aconitate hydratase [Arabidopsis thaliana]	0.64 ± 0.03	0.74 ± 0.02	0.85 ± 0.03	Energy metabolism	7
Pyruvate kinase [Arabidopsis thaliana]	0.75 ± 0.03	0.77 ± 0.03	0.97 ± 0.01	Energy metabolism	4
Fructose-bisphosphate aldolase [Medicago truncatula]	0.77 ± 0.03	0.78 ± 0.02	0.98 ± 0.02	Energy metabolism	5
Glyceraldehyde 3-phosphate dehydrogenase (Fragment) [Salvia elegans]	0.73 ± 0.03	0.60 ± 0.01	1.21 ± 0.03	Energy metabolism	8
Elongation factor EF-2 [Arabidopsis thaliana]	0.71 ± 0.05	0.74 ± 0.02	0.96 ± 0.05	Protein biosynthesis	23
Elongation factor 1-alpha 1[Arabidopsis thaliana]	0.78 ± 0.03	0.78 ± 0.01	1.00 ± 0.03	Protein biosynthesis	11
Eukaryotic initiation factor 4A [Oryza sativa Japonica Group]	0.79 ± 0.02	0.70 ± 0.01	1.13 ± 0.01	Protein biosynthesis	7
Glutamyl-tRNA synthetase, cytoplasmic, putative [Ricinus communis]	0.54 ± 0.02	0.47 ± 0.02	1.13 ± 0.00	Protein biosynthesis	2
Methionine synthase [Helianthus annuus x Helianthus debilis subsp. Debilis]	0.65 ± 0.05	0.69 ± 0.01	0.94 ± 0.06	Amino-acid biosynthesis	6
14-3-3-like protein 2 [Gossypium hirsutum]	0.48 ± 0.02	0.71 ± 0.01	0.67 ± 0.01	ATP binding	5
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase [Solenostemon scutellarioides]	0.55 ± 0.02	0.62 ± 0.02	0.88 ± 0.01	Methyltransferase	12
Putative S-adenosylmethionine synthetase [Capsicum annuum]	0.66 ± 0.02	0.74 ± 0.02	0.90 ± 0.01	Methionine adenosyltransferase activity	3

NE12: Normal embryos at 12 days after pollination; NE18: Normal embryos at 18 days after pollination; AE18: Abortive embryos at 18 days after pollination. Values represent mean ± SE.
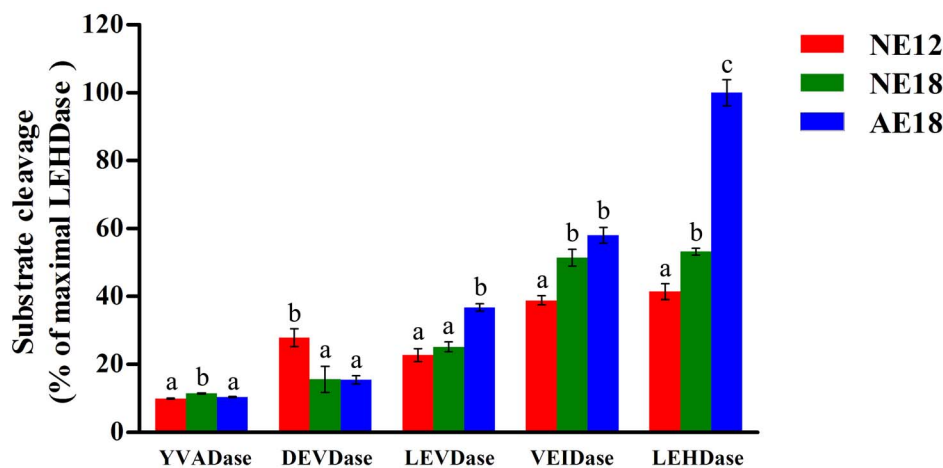
Figure 4 | Caspase-like activities in normal and abortive chrysanthemum embryos. Embryo extracts were measured for substrate specificity using different caspase substrates: Ac-YVAD-AMC (YVADase), Ac-DEVD-AMC (DEVDase), Ac-LEVD-AMC (LEVDase), Ac-VEID-AMC (VEIDase), Ac-LEHD-AMC (LEHDase). Relative fluorescence units were calculated and expressed as a percentage of the LEHDase activity. AE 18 (abortive embryos) has the highest cleavage activities for Ac-LEVD-AMC, Ac-VEID-AMC and Ac-LEHD-AMC compared with NE12 and NE 18, indicating that chrysanthemum embryo abortion is a result of programmed cell death. Different letters indicate significant differences at alpha = 0.05 by the Bonferroni t-test (the bar is standard error).

were down-regulated in the pollen tubes inhibited by nifedipine, and thus concluded that a decrease in energy production may be a factor contributing to inhibition of pollen tubes[40]. In the present study, transcriptomic and proteomic analyses showed that several energy-related genes, such as ATP synthase subunit, ATPase subunit and plasma membrane H^+ ATPase etc., and eight proteins (i.e., ATP synthase beta chain, UDP-glucose 6-dehydrogenase and pyruvate kinase) implicated in tricarboxylic acid cycle or in glycolysis were significantly down-regulated in abortive chrysanthemum embryos compared with normal embryos, indicating that the rate of energy metabolism in abortive embryos was significantly lower than that in normal embryos. In other words, the chrysanthemum abortion process was concomitant with a significant reduction in energy production.

Abnormal protein metabolism was also related to chrysanthemum embryo abortion. Protein is the basis material of life, involved in cell growth and cell cycle progression, a key factor for robust plant growth. Protein synthesis is affected by many factors, for example, initiation, elongation, and release factors. Modification of some factors affects the rate of translation and protein synthesis[41]. In the current study, we found that some genes and proteins associated with protein synthesis, such as initiation factor, elongation factor and ribosomal protein, were down-regulated in abnormal embryos. For example, the abundance of cytoskeleton constituent including actin and tubulin in NE12 and NE18 were higher than that in AE18. Because cytoskeleton is crucial for cell structure, cell division and expansion, and signal transduction[42], the significant down-regulation of actin and tubulin in AE18 may affect the elongation or reduction of protein synthesis, and may be associated with chrysanthemum embryo abortion.

Meanwhile, the mRNA levels of several genes encoding ubiquitin, 26S proteasome, and ubiquitin-conjugating enzyme were significantly higher in abortive embryos relative to normal embryos. The ubiquitin-proteasome system is the main protein-degradation pathway; more than 80% of protein degradation is dependent on this pathway in plants[43–45]. Therefore, the higher abundance of several components in ubiquitin-proteasome system suggests that the process of protein degradation is accelerated in chrysanthemum abortion. Consequently, our findings suggest that reduced protein synthesis and enhanced protein degradation were closely related to chrysanthemum embryo abortion.

Transcriptional factors during embryo development. Transcription factors are essential for the regulation of gene expression during flower development, embryogenesis and seed maturation as well as biotic and abiotic stresses that include pathogen, drought, high temperature and salt[46,47]. In this study, the abundance of many transcription factors, such as AP2/ERF, WRKY, Myb, and zinc fingers, were changed during abortive chrysanthemum embryos compared with normal embryos. Interestingly, most genes encoding stress-responsive transcription factors, including AP2/ERF, WRKY, and Myb, were significantly up-regulated in abortive embryos. Conversely, some others encoding zinc finger protein and leucine-rich repeat receptor protein kinase were significantly down-regulated in abortive embryos relative to normal embryos. Zinc finger proteins and leucine zipper protein have been reported to be required to balance cell proliferation and maintain specification of undifferentiated cells during active growth, including during embryogenesis and seed germination[46,48]. A possible reason is that chrysanthemum embryos may receive incompatibility signalling when they reach the globular embryo stage. This signalling may be similar to high expression of stress-response transcription factors in response to a stimulus of biotic and abiotic stresses. Therefore, transcription factors play an important role in complex network of chrysanthemum embryo abortion.

In conclusion, we performed transcriptomic and proteomic analysis as complementary approaches to investigate embryo abortion during chrysanthemum cross breeding. Many functional genes and proteins were identified that potentially contribute to chrysanthemum embryo abortion, especially those associated with cell senescence and death. In addition, a new and important finding is that abortive embryos have higher caspase-like activities compared with normal embryos. Caspase-like activity is a key factor in PCD, so we presume that embryo abortion is a result of PCD. Taken together, this study highlights our understanding of mechanisms related to chrysanthemum embryo abortion, and may be useful in the future for overcoming embryo abortion during cross breeding of chrysanthemum and other crops.

Methods

Plant material and cross. An interspecific cross was performed as described in the previous report[15]. *C. morifolium* and *C. nankingense* were female and male parents. Because chrysanthemum embryos are too small to be sampled, and endosperm development also influences embryo development, ovules including embryo,

7

endosperm and nucellus were collected by manual dissection with the naked eye. At 12 DAP, when nearly all the embryos reached the stage of globular embryo and were normal, ovules were collected. At 18 DAP when normal embryos developed into the heart-shaped embryo stage and abnormal embryos were undergoing degeneration, normal and abnormal ovules were collected (Fig. 1a–k). We had three planting areas, and 20 chrysanthemum plants were planted in each area. Each chrysanthemum plant had more than 50 inflorescences, and approximately 1000 inflorescences can be used in each planting area. The interspecific cross was performed, and at least 500 inflorescences (each inflorescence with 18–23 female ligulate florets or ray florets containing one ovule) were hand-emasculated and bagged in each planting area. We collected samples at two stages after pollination. At 12 days after pollination, we collected approximately 2300 normal ovules (>0.7 g) from each planting area, and there were nearly 7000 ovules (>2.1 g) in the three planting areas. At 18 days after pollination, we similarly collected >2.1 g normal and abnormal ovules, respectively. Then, 0.5 g, 0.5 g and 0.6 g ovules were used for RNA-Seq, proteomics and caspase activity assays. After collection, the three samples, normal embryos at 12 DAP (NE12), normal embryos at 18 DAP (NE18) and abortive embryos at 18 DAP (AE18), were immediately frozen in liquid nitrogen and stored at −80°C.

Total RNA and protein extraction. Total RNA was extracted using Trizol reagent according to the manufacturer's protocol (Takara Bio Inc., Otsu, Japan). The integrity and purity of the RNA was determined by Agilent 2100 RNA 6000 Kit and electrophoresis on 1% agarose gel.

cDNA preparation and Illumina deep-sequencing. Illumina sequencing was performed at the Beijing Genomics Institute (BGI)-Shenzhen, Shenzhen, China (http://www.genomics.cn/index.php) according to the manufacturer's instructions (Illumina, San Diego, CA, USA)[49].

De novo assembly and sequence clustering. We obtained clean reads from raw data. The application of Trinity[51] completed the de novo assembly of the transcriptome into unigenes. These assembled unigenes can be taken into further processing of sequence splicing and redundancy removing with sequence clustering software to acquire non-redundant unigenes. Finally, the application of Blastx (E-value < 0.00001) revealed alignment between unigenes and protein databases including NCBI nonredundant protein (Nr), Swiss-Prot, Kyoto Encyclopedia of Genes and Genomes (KEGG), and Cluster of Orthologous Groups (COG), and the directions of sequences were determined by the best aligning results. ESTScan[50] was also used to predict the directions of sequences when a unigene was not aligned to those in any of the databases mentioned above.

Unigene function annotation. Unigene sequences were aligned by Blastx to protein databases such as Nr, Swiss-Prot, KEGG and COG and aligned by Blastn to the nucleotide database Nt (E-value < 0.00001). In this way, we achieved the best functional annotations. With Nr annotation, we used the Blast2 GO program[51] to obtain gene ontology (GO) annotation of unigenes. GO has three ontologies describing molecular function, cellular component and biological process[51]. We then used WEGO software[52] to do GO functional classification for all unigenes and to understand the distribution of gene functions at the macro level. KEGG mapping, which helps study complicated biological behaviours of genes, was used to determine metabolic pathway annotation of unigenes.

Analysis of difference in unigenes expression. The expression levels of unigenes were calculated using the method of fragments per kb per million fragments (FPKM)[27]. This method eliminates the influence of different gene length and sequencing level on the calculation of gene expression. Therefore, the calculated gene expression can be directly used for comparing the difference of gene expression between samples. Referring to the statistical method of Audic and Claverie[53] based on sequencing, we analysed the differential expression of unigenes. Firstly, we mapped all differentially expressed genes (DEGs) to each term of GO database and calculated the gene numbers for each GO term, which yields a gene list and gene numbers for every certain GO term. Next, we used a hypergeometric test to find significantly enriched GO terms in DEGs compared to the genome background.

Quantitative real-time PCR (qRT-PCR) validation. To ensure the libraries' reliability, we randomly selected 14 differentially expressed genes, occupying about one-tenth of the 143 DEGs, and validated the data using the three biological replicates samples by qRT-PCR. qRT-PCR assays were conducted as described by Song *et al.*[54]. Special primers for thirty genes were designed using PRIMER3 RELEASE 2.3.4[55] and the reference sequence for quantitative expression analysis was Elongation Factor 1α (EF1α) gene (Table S5). All reactions were performed in biological triplicate. The workflow for transcriptomic analysis is illustrated in Fig. S4.

Protein extraction, digestion and iTRAQ labelling. Protein extraction was performed using trichloroacetic acid (TCA)/acetone method[40]. The protein was redissolved in dissolution buffer (0.5 M triethylammonium bicarbonate, 0.1% (w/v) SDS). Protein concentration was determined according to Bradford assay. From each sample, 100 μg of protein (~30 μL) was denatured and alkylated, and the cysteines blocked according to the manufacturer's instructions of the 8-plex iTRAQ reagent kit (Applied Biosystems, California, USA). Protein was digested with trypsin (Promega, Madison, WI) at a trypsin/protein ratio of 1 : 50 for 16 h at 37°C. For labelling, each iTRAQ reagent was dissolved in 80 μL of isopropanol and added to the respective

peptide mixture. Proteins were labelled with the iTRAQ tags as follows: 1–118 isobaric tag, 2–119 isobaric tag, 3–121 isobaric tag. The labelled samples were combined and dried in vacuo. A SepPac™ C18 cartridge (1 cm³/50 mg, Waters Corporation, Milford, MA) was used to remove the salt buffer. The peptide mixture was re-dissolved in low salt buffer (5 mM K_3PO_4 with 25% (v/v) acetonitrile, pH = 2.5, adjusted with H_3PO_4), and then subjected to strong cation exchange chromatography to remove the hydrolysed unbound iTRAQ reagents using a PolySULFOETHYL A (2.1 × 100 mmi.d.) HPLC column (PolyLC Inc, Columbia, MD, USA). Peptide mixture was collected after high salt buffer (10 mM K_3PO_4, 500 mM KCl, 25% (v/v) acetonitrile, pH = 2.5, adjusted with H_3PO_4) elution. After the acetonitrile was removed by volatilisation, the SepPac™ C18 cartridge was used for the purpose of removing remove excess KCl, and then dried in a vacuum concentrator.

High pH reverse-phase separation. The peptide mixture was re-dissolved in buffer A (20 mM NH_4HCO_2, pH10.0, adjusted with NH_4OH), and then fractionated by high pH separation using a Shimazu UFLC system (Shimazu, Kyoto, Japan) connected to a reverse phase column (PolyRP-300 column, 2.1 mm × 150 mm, 5 μm, 300 Å, The Sepax Technologies, Inc. USA). High pH separation was performed using a four-step linear gradient. Starting from 5% B (20 mM NH_4HCO_2 in 90% acetonitrile, pH 10.0, adjusted with NH_4OH) to 35% B in 30 min, increased to 80% B in 2 min, and then to 5% B in 2 min. The column was re-equilibrated at initial conditions for 15 min. The column flow rate was maintained at 200 μL/min and column temperature was maintained at room temperature. Fifteen fractions were collected and each fraction was dried in a vacuum concentrator for the next step.

Low pH nano-HPLC-MS/MS analysis. The peptides were resuspended with 20 μl solvent A (0.1% (v/v) formic acid in water, pH 2.5), separated by nanoLC and analysed by on-line electrospray tandem mass spectrometry. The experiments were performed on a Nano Aquity UPLC system (Waters Corporation, Milford, MA) connected to an LTQ Orbitrap XL mass spectrometer (Thermo Electron Corp., Bremen, Germany) equipped with an online nanoelectrospray ion source (Michrom Bioresources, Auburn, USA). An 18 μL peptide sample was loaded onto the Captrap Peptide column (Michrom Bioresources, Auburn, USA), with a flow of 20 μL/min for 5 min and subsequently eluted with a three-step linear gradient. Starting from 5% B to 45% solvent B (0.1% (v/v) formic acid in 90% acetonitrile, pH 2.5) in 100 min, increased to 80% B in 1 min, and then hold on 80% B for 4 min. The column was re-equilibrated at initial conditions for 15 min. The column flow rate was maintained at 500 nL/min and column temperature was maintained at 35°C. The electrospray voltage of 1.6 kV versus the inlet of the mass spectrometer was used. LTQ Orbitrap XL mass spectrometer was operated in the data-dependent mode to switch automatically between MS and MS/MS acquisition. Survey full-scan MS spectra (m/z 400–1800) were acquired in the Obitrap with a mass resolution of 30,000 at m/z 400, followed by five sequential HCD-MS/MS scans. The automatic gain control was set to 500 000 ions to prevent over-filling of the ion trap. The minimum MS signal for triggering MS/MS was set to 1000. In all cases, one microscan was recorded. MS/MS scans were acquired in the Obitrap with a mass resolution of 7500. The dissociation mode was HCD (higher energy C-trap dissociation). Dynamic exclusion was used with two repeat counts, 10 s repeat duration, and the m/z values triggering MS/MS were put on an exclusion list for 120 s. For MS/MS, precursor ions were activated using 40% normalised collision energy and an activation time of 30 ms.

Database searching and criteria. The protein data search methods and software were carried out as presented in a previous article[56]. The experiments for proteomics investigation were repeated three times, and the workflow for proteomics investigation is illustrated in Fig. S5.

Caspase activity assays. Embryo protein extraction and caspase activity assays were performed as described by Maurice Bosch and Vernonica E. Franklin-Tong[30]. Caspase-like activities were measured by using AMC substrates: Ac-YVAD-AMC, Ac-DEVD-AMC, Ac-LEVD-AMC, Ac-VEID-AMC, Ac-LEHD-AMC (sigma, USA), which are selective fluorogenic substrate for caspase 1, caspase 3, caspase 4, caspase 6, caspase 9, respectively. Each 200 μL reaction mixture contained 20 μg of embryo protein extract and 100 μM fluorogenic substrate. MicroplateReader (Tecan Infinite M200, Switzerland) was used to measure (excitation 380 nm, emission 460 nm) the release of fluorophore by cleavage at 27°C for 2 h. All analyses were repeated three times.

1. Bawa, K. S. Embryo and seed abortion in plants. *Nature* **342**, 625 (1989).
2. Batygina, R. B. Generative organs of flower *Embryology of flowering plants: Terminology and concepts* Batygina, R. B. (ed.) (Science Publishers Inc, USA, 2002).
3. Arathi, H. S., Ganeshaiah, K. N., Shaanker, R. U. & Hegde, S. G. Factors affecting embryo abortion in *Syzygium cuminii* (L.) Skeels (Myrtaceae). *Int. J. Plant Sci.* **157**, 49–52 (1996).
4. Ndoutoumou, P. N., Toussaint, A. & Baudoin, J. P. Embryo abortion and histological features in the interspecific cross between *Phaseolus vulgaris* L. and *P. coccineus* L. *Plant Cell Tissue Organ Cult.* **88**, 329–332 (2007).
5. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009).

8

6. Palmieri, M. C. *et al.* Proteomic analysis of grapevine resistance induced by *Trichoderma harzianum* T39 reveals specific defence pathways activated against downy mildew. *J. Exp. Bot.* **63**, 6237–6251 (2012).

7. Robertson, F. P. *et al.* Proteome changes in tomato lines transformed with phytoene synthase-1 in the sense and antisense orientations. *J. Exp. Bot.* **63**, 6035–6043 (2012).

8. Raherison, E. *et al.* Transcriptome profiling in conifers and the PiceaGenExpress database show patterns of diversification within gene families and interspecific conservation in vascular gene expression. *BMC Genomics* **13**, 434 (2012).

9. Qiu, Z. B. *et al.* The regulation of cambial activity in Chinese fir (*Cunninghamia lanceolata*) involves extensive transcriptome remodeling. *New Phytol.* **199**, 708–719 (2013).

10. Palovaara, J., Saiga, S. & Weijers, D. Transcriptomics approaches in the early *Arabidopsis* embryo. *Trends Plant Sci.* **18**, 514–521 (2013).

11. Chen, X. P. *et al.* Deep sequencing analysis of the transcriptomes of peanut aerial and subterranean young pods identifies candidate genes related to early embryo abortion. *Plant Biotechnol. J.* **11**, 115–127 (2013).

12. Teixeira da Silva, J. A. *et al.* Chrysanthemum Biotechnology: *Quo vadis*? *Crit. Rev. Plant Sci.* **32**, 21–52 (2013).

13. Anderson, N. O. Chrysanthemum. *Flower Breeding and Genetics* Anderson, N. O. (ed.) 389–437 (Springer, Netherlands, 2006).

14. Sun, C. Q. *et al.* Interspecific hybrids between *Chrysanthemum grandiflorum* (Ramat.) Kitamura and *C. indicum* (L.) Des Moul. and their drought tolerance evaluation. *Euphytica* **174**, 51–60 (2010).

15. Sun, C. Q. *et al.* Factors affecting seed set in the crosses between *Dendranthema grandiflorum* (Ramat.) Kitamura and its wild species. *Euphytica* **171**, 181–192 (2010).

16. Sun, C. Q. *et al.* Overcoming pre-fertilization barriers in the wide cross between *Chrysanthemum grandiflorum* (Ramat.) Kitamura and *C. nankingense* (Nakai) Tzvel. by using special pollination techniques. *Euphytica* **178**, 195–202 (2011).

17. Lafon-Placette, C. & Köhler, C. Embryo and endosperm, partners in seed development. *Curr. Opin. Plant Biol.* **17**, 64–69 (2014).

18. Köhler, C. & Kradolfer, D. Epigenetic mechanisms in the endosperm and their consequences for the evolution of flowering plants. *BBA-Gene Regul. Mech.* **1809**, 438–443 (2011).

19. Costa, L. M. *et al.* Maternal control of nutrient allocation in plant seeds by genomic imprinting. *Curr. Biol.* **22**, 160–165 (2012).

20. Kradolfer, D., Wolff, P., Jiang, H., Siretskiy, A. & Köhler, C. An imprinted gene underlies postzygotic reproductive isolation in *Arabidopsis thaliana*. *Dev. cell* **26**, 525–535 (2013).

21. Bochenek, M. *et al.* Transcriptome analysis of the sulfate deficiency response in the marine microalga *Emiliania huxleyi*. *New Phytol.* **199**, 650–662 (2013).

22. Guggisberg, A., Lai, Z., Huang, J. & Rieseberg, L. H. Transcriptome divergence between introduced and native populations of Canada thistle, Cirsium arvense. *New Phytol.* **199**, 595–608 (2013).

23. Muth, T. *et al.* TraqX: a free, platform independent tool for isobaric tag quantitation at the protein level. *Proteomics* **10**, 1223–1225 (2010).

24. Hou, S. Y., Jones, S. W., Choe, L. H., Papoutsakis, E. T. & Lee, K. H. Workflow for quantitative proteomic analysis of *Clostridia acetobutylicum* ATCC 824 using iTRAQ tags. *Methods* **61**, 269–276 (2013).

25. Tse, W. K. F. *et al.* Transcriptomic and iTRAQ proteomic approaches reveal novel short-term hyperosmotic stress responsive proteins in the gill of the Japanese eel (*Anguilla japonica*). *J. Proteomics* **89**, 81–94 (2013).

26. Wang, H. B. *et al.* Next-generation sequencing of the *Chrysanthemum nankingense* (Asteraceae) transcriptome permits large-scale unigene assembly and SSR marker discovery. *PloS one* **8**, e62293 (2013).

27. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. methods* **5**, 621–628 (2008).

28. Oliveros, J. C. VENNY. An interactive tool for comparing lists with Venn Diagrams.<http://bioinfogp.cnb.csic.es/tools/venny/index.html> (2007) Date of access: 06/06/2014.

29. Zhao, P. *et al.* A bipartite molecular module controls cell death activation in the basal cell lineage of plant embryos. *PLoS Biol.* **11**, e1001655 (2013).

30. Thomas, S. G. & Franklin-Tong, V. E. Self-incompatibility triggers programmed cell death in *Papaver* pollen. *Nature* **429**, 305–309 (2004).

31. Rogers, H. J. Cell death and organ development in plants. *Curr. Top. Develop. Biol.* **71**, 225–261 (2005).

32. Ge, X. C. *et al.* An Arabidopsis aspartic protease functions as an anti-cell-death component in reproduction and embryogenesis. *EMBO Rep.* **6**, 282–288 (2005).

33. Bosch, M. & Franklin-Tong, V. E. Temporal and spatial activation of caspase-like enzymes induced by self-incompatibility in *Papaver* pollen. *Proc. Natl. Acad. Sci. USA* **104**, 18327–18332 (2007).

34. Wilkins, K. A. *et al.* Reactive oxygen species and nitric oxide mediate actin reorganization and programmed cell death in the self-incompatibility response of *papaver*. *Plant Physiol.* **156**, 404–416 (2011).

35. Xue, Z. Q. *et al.* Changes of Endogenous Hormones in Relation to Embryo Abortion in Seedless Watermelon. *Acta Agr. Boreali-Occidentalis Sinica* **3**, 044 (2007).

36. Ma, K. F. *et al.* The low fertility of Chinese white poplar: dynamic changes in anatomical structure, endogenous hormone concentrations, and key gene expression in the reproduction of a naturally occurring hybrid. *Plant Cell Rep.* **32**, 401–414 (2013).

37. Hays, D. B., Do, J. H., Mason, R. E., Morgan, G. & Finlayson, S. A. Heat stress induced ethylene production in developing wheat grains induces kernel abortion and increased maturation in a susceptible cultivar. *Plant Sci.* **172**, 1113–1123 (2007).

38. Facucho-Oliveira, J. M. & John, J. S. The relationship between pluripotency and mitochondrial DNA proliferation during early embryo development and embryonic stem cell differentiation. *Stem Cell Rev. Rep.* **5**, 140–158 (2009).

39. Speakman, J. R. Body size, energy metabolism and lifespan. *J. Exp. Biol.* **208**, 1717–1730 (2005).

40. Wu, X. Q. *et al.* Integrative proteomic and cytological analysis of the effects of extracellular Ca^{2+} influx on *Pinus bungeana* pollen tube development. *J. Proteome Res.* **7**, 4299–4312 (2008).

41. Clark, B. F. C. *et al.* Prokaryotic and eukaryotic translation actors: International union of biochemistry and molecular biology (IUBMB). *Biochimie* **78**, 1119–1122 (1996).

42. Drake, T. & Vavylonis, D. Cytoskeletal dynamics in fission yeast: a review of models for polarization and division. *HFSP J.* **4**, 122–130 (2010).

43. Verma, R. *et al.* Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science* **298**, 611–615 (2002).

44. Guo, X. *et al.* UBLCP1 is a 26S proteasome phosphatase that regulates nuclear proteasome activity. *Proc. Natl. Acad. Sci. USA* **108**, 18649–18654 (2011).

45. Teixeira, L. K. & Reed, S. I. Ubiquitin ligases and cell cycle control. *Annu. Rev. Biochem.* **82**, 387–414 (2013).

46. Kim, D. H. *et al.* SOMNUS, a CCCH-type zinc finger protein in *Arabidopsis*, negatively regulates light-dependent seed germination downstream of PIL5. *Plant Cell* **20**, 1260–1277 (2008).

47. Xu, Z. S., Chen, M., Li, L. C. & Ma, Y. Z. Functions and application of the AP2/ERF transcription factor family in crop improvement. *J. Integr. Plant Biol.* **53**, 570–585 (2011).

48. Schmitz, R. J., Hong, L., Michaels, S. & Amasino, R. M. *FRIGIDA-ESSENTIAL 1* interacts genetically with *FRIGIDA* and *FRIGIDA-LIKE 1* to promote the winter-annual habit of *Arabidopsis thaliana*. *Develpment* **132**, 5471–5478 (2005).

49. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).

50. Iseli, C., Jongeneel, C. V. & Bucher, P. ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *ISMB* **99**, 138–148 (1999).

51. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

52. Ye, J. *et al.* WEGO: a web tool for plotting GO annotations. *Nucl. Acid. Res.* **34**, W293–W297 (2006).

53. Audic, S. & Claverie, J. M. The significance of digital gene expression profiles. *Genome Res.* **7**, 986–995 (1997).

54. Song, A. P. *et al.* Isolation and characterisation of *Chrysanthemum crassum* SOS1, encoding a putative plasma membrane Na^+/H^+ antiporter. *Plant Biol.* **14**, 706–713 (2012).

55. Rozen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Bioinformatics methods and protocols* Misener, S. & Krawetz, S. A. (eds.) 365–386 (Humana Press, New York, 1999).

56. Liu, S. Y. *et al.* iTRAQ-based quantitative proteomic analysis for identification of oligodendroglioma biomarkers related with loss of heterozygosity on chromosomal arm 1p. *J. Proteomics.* **77**, 480–491 (2012).

Author contributions

N.J.T., F.J.Z., F.D.C. and W.M.F. designed the experiments. F.J.Z., Z.Q.W., W.D., C.Q.S., H.B.W., A.P.S. and L.Z.H. performed the experiments. F.J.Z., N.J.T., W.D., H.B.W. and A.P.S. analyzed the data. N.J.T. and F.J.Z. wrote the manuscript. All authors read and approved the final manuscript.

Additional information