Plant Omics Journal POJ

POJ 7(1):1-11 (2014) ISSN:1836-3644

Transcriptomic landscape of chrysanthemums infected by *Chrysanthemum stunt viroid*

Yeonhwa Jo[#], Kyoung-Min Jo[#], Sang-Ho Park, Kook-Hyung Kim[*], and Won Kyong Cho[*]

Department of Agricultural Biotechnology and Research Institute for Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, Seoul, 151-921, Republic of Korea

[*]Corresponding authors: wonkyong@gmail.com; kookkim@snu.ac.kr
[#] These authors contributed equally to this work.

Abstract

CSVd is one of the main pathogens infecting chrysanthemum species and is responsible for serious economic losses. To identify the genes involved in viroid infection, we obtained the ESTs of chrysanthemums infected by CSVd using a high-throughput next generation sequencing method. The cDNA library was prepared from the total RNAs extracted from leaves of chrysanthemum infected with CSVd. The prepared cDNA library was sequenced using the FLX 454 sequencer. As a result, a total of 11,600 ESTs were obtained from 99,750 reads after de novo assembly. A BLAST search was performed to annotate the functions of the 11,600 chrysanthemum ESTs. In addition, GO terms, enzyme codes, and InterPro domains for each chrysanthemum EST were obtained using the Blast2GO program with default parameters. Gene ontology enrichment analysis identified 144 significantly enriched GO terms, which were further divided into biological processes (93 GO terms), cellular components (19 GO terms), and molecular functions (32 GO terms). CSVd infection affected the mRNAs involved in various metabolic pathways, stresses, transcription, translation, transport, chloroplasts, plasmodesmata, actins, and microtubules. Comparative analysis revealed 70% of the chrysanthemum ESTs were orthologous to those of five representative plant species. The current study provides information on a large number of chrysanthemum genes. The obtained information will be valuable for studying plant-viroid interactions.

Keywords: Chrysanthemum, Expressed sequence tag, *Chrysanthemum stunt viroid*, Viroid infection, Next generation sequencing.
Abbreviations: CSVd_*Chrysanthemum stunt viroid*; ESTs_ Expressed sequence tags; GO_Gene ontology.

Introduction

Chrysanthemums (*Chrysanthemum morifolium* Ramat. or *Dendranthema grandiflora* Tzvelev.) belonging to the *Asteraceae* family has been cultivated for more than 3,000 years in Asia and Europe. It is currently one of popular flowers and very important to the worldwide floriculture industry. In general, most of the cultivated chrysanthemums are derived from a hexaploid species that includes five to six species and has 54 chromosomes on average (Dowrick, 1953; Dowrick and El-Bayoumi, 1966).

Several genetic studies have demonstrated that crosses between related or unrelated chrysanthemum species are very often unsuccessful due to their incompatibility system (Drewlow et al., 1973; Ronald and Ascher, 1975). Furthermore, biparental crosses between chrysanthemum species produce many progeny varieties whose morphological characteristics are strongly different from the F1 parent (Dowrick and El-Bayoumi, 1966). Many commercial chrysanthemum cultivars originated from somatic mutants that resulted from the loss and gain of chromosomal material during mitosis (Stewart and Dermen, 1970). Thus, a large number of chrysanthemum varieties and cultivars have been produced by hybridization. Recent studies have generated new chrysanthemum varieties by genetic modification (Courtney-Gutterson et al., 1994; Teixeira da Silva, 2004). They are mainly asexually propagated by vegetative cuttings (Roest and Bokelmann, 1975).

In China, flowers of *Chrysanthemum morifolium* are used for making medicinal tea, which is used to aid in such processes as influenza recovery (Miyazawa and Hisama, 2003). Moreover, the insecticide pyrethrin, which is a neurotoxin that attacks insect nervous systems, can be obtained from *Chrysanthemum cinerariaefolium.* The flower produces a natural organic pyrethrin (Hitmi et al., 2000).

ESTs are known to be economical and efficient methods to gain knowledge on genes in molecular biology (Ohlrogge and Benning, 2000). The recent rise of next generation sequencing technologies has facilitated the study of high-throughput functional genomics (Morozova and Marra, 2008). In the *Asteraceae* family, the ESTs of several species including chicory, chrysanthemums, sunflowers, and gerbera hybrid have been previously reported (Chen et al., 2009; Dauchot et al., 2009; Fernández et al., 2003; Laitinen et al., 2005).

Viroids are the smallest pathogens composed of circular RNAs and cause infectious diseases in their plant hosts (Ding, 2010). So far, two viroids infecting chrysanthemums are known: *Chrysanthemum stunt viroid* (CSVd) and *Chrysanthemum chlorotic mottle viroid* (CChMVd) (Hosokawa et al., 2006). CSVd belongs to the family *Pospiviroidae* and replicates in the nucleus while CChMVd, a member of the family *Avsunviroidae* replicates in chloroplasts. The RNA structures and symptoms of the two viroids are different (Haseloff and Symons, 1981;

Table 1. Summary of the sequenced reads obtained by 454 FLX sequencing.

Number of reads	Number of bases	Assembled	Partial	Singleton	Repeat	Outlier
99,750	37,815,742	71,847	13,142	12,652	131	984

Navarro and Flores, 1997). In particular, CSVd infection leads to serious damage to chrysanthemums by producing symptoms such as bleached leaves, stunting, and reduced levels of anthocyanin (Randles et al., 2003). The chrysanthemum is an ideal model plant for plant-viroid interaction studies due to several advantages. The first is that the chrysanthemum is the host for two different viroids including CSVd and CChMVd. The second is that it is relatively easy to generate transgenic chrysanthemum plants with a high level of efficiency for transformation and propagation, which enables molecular functional gene characterization (Rout and Das, 1997).
Presently, only a few studies have performed to identify the host genes that are differentially expressed upon viroid infection using microarray technology (Owens et al., 2012; Rizza et al., 2012). In the case of chrysanthemums, knowledge of only a limited number of genes is available. Thus, EST analysis is a useful to gain knowledge of the genes and their functions associated with viroid infection.

In this study, we performed transcriptomic analysis of the chrysanthemum in response to CSVd infection using next generation sequencing, which is an inexpensive and effective way for de novo sequencing when compared to Sanger sequencing (Shendure and Ji, 2008). The current study provides information on the genes associated with CSVd infection. This information will provide a good resource for future studies of plant-viroid interaction.

Results and Discussion

Construction of expressed sequence tags for CSVd-infected chrysanthemums

The transcriptome of CSVd-infected chrysanthemum was analyzed by GS FLX 454 sequencing, resulting in 99,750 reads. Of those reads, 71,847 reads were used for de novo assembly, resulting in 13,142 partial and 12,652 singleton sequences (Table 1). After removing 131 repeated and 984 outlier sequences, a total of 11,600 unigenes (ESTs) were obtained (Table S1). The average contig count was 1.23 and the number of ESTs with one contig was 10,047 (Table 2). The average EST length was 582.3 bp and a majority of the identified ESTs had lengths from 400 bp to 600 bp (Fig. 1). The number of ESTs 401-500 bp in length was 2,707. The largest EST was 7,241 bp, while the smallest EST was 39 bp (Fig. 1). With an advantage of FLX 454 sequencing technology, the lengths of ESTs in this study were enough to find novel chrysanthemum genes.

Functional classification of chrysanthemum ESTs

To unravel the putative functions of the identified chrysanthemum ESTs, a total of 11,600 chrysanthemum ESTs were submitted to the Swiss-Prot database to find homology using the blastX function implementation in the Blast2GO program. Among these, 6,964 ESTs (60.03%) were matched to known proteins in the Swiss-Prot database, while approximately

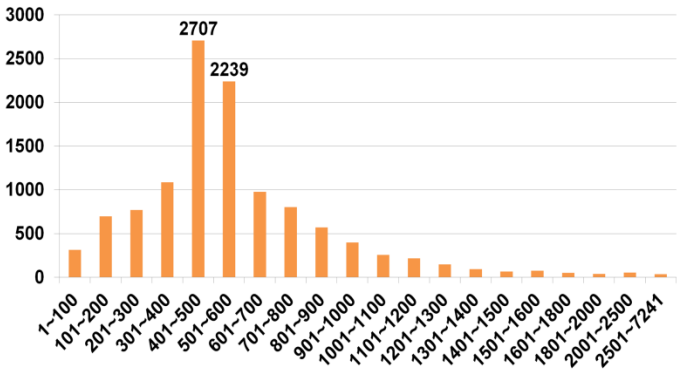

Fig 1. Length distribution of identified ESTs. Left column indicates the number of ESTs. The bottom line indicates the length distributions of ESTs. The lengths of a majority of ESTs were between 401 bp and 600 bp.

40% of the ESTs were not assigned to any known functions. This result suggests that a majority of chrysanthemum genes could be novel genes which are so far identified in other plant species. The obtained E-values ranged from 9.9E-4 to 0. Approximately half of the matched EST sequences (3,202 ESTs) have E-value less than 1E-50. To identify the potential functions of the unknown proteins, the conserved domains in each EST were scanned using the InterPro program. The most frequently identified domains were protein kinase domains, such as the protein kinase-like domain (IPR011009 and IPR000719) and serine/threonine kinases (IPR002290 and IPR008271), followed by a NAD(P)-binding domain (IPR016040), cytochrome P450, and a nucleotide-binding domain (IPR012677). Moreover, an RNA recognition motif domain (IPR000504), a thioredoxin-like fold (IPR012336), WD40 repeats (IPR001680), an armadillo-type fold (IPR016024), and an aldolase-type TIM barrel (IPR013785) were frequently present in the chrysanthemum ESTs (Table 3).
According to the gene ontology (GO) functional classification, 6,779 ESTs had at least one GO term. The total number of GO terms was 78,483, which were further divided into 38,235 GO terms for biological processes, 20,435 GO terms for cellular processes, and 18,829 GO terms for molecular functions. Some ESTs, such as unigene08644, had 130 GO terms. To identify the gene functions strongly related to viroid infection, a GO enrichment analysis was performed using orthologous *Arabidopsis* genes. Applying a stringent p-value (less than 0.01), a total of 144 GO terms associated with CSVd-infected chrysanthemum were identified. These terms could be further divided into three different ontologies: biological processes (93 GO terms), cellular components (19 GO terms), and molecular functions (32 GO terms) (Table S2). The genes responsible for various metabolic pathways, stresses, transcription, translation, transport and organelles were significantly enriched in the transcriptome of the CSVd-infected chrysanthemum.

Table 2. Identified ESTs (Isotigs) from CSVd-infected chrysanthemum using 454 FLX sequencing.

Number of ESTs	Number of bases	Number with one contig	Average contig count	Largest contig count	Average EST length (bp)	maximum EST length (bp)
11,600	6,754,699	10,047	1.23	12	582.3	7,241

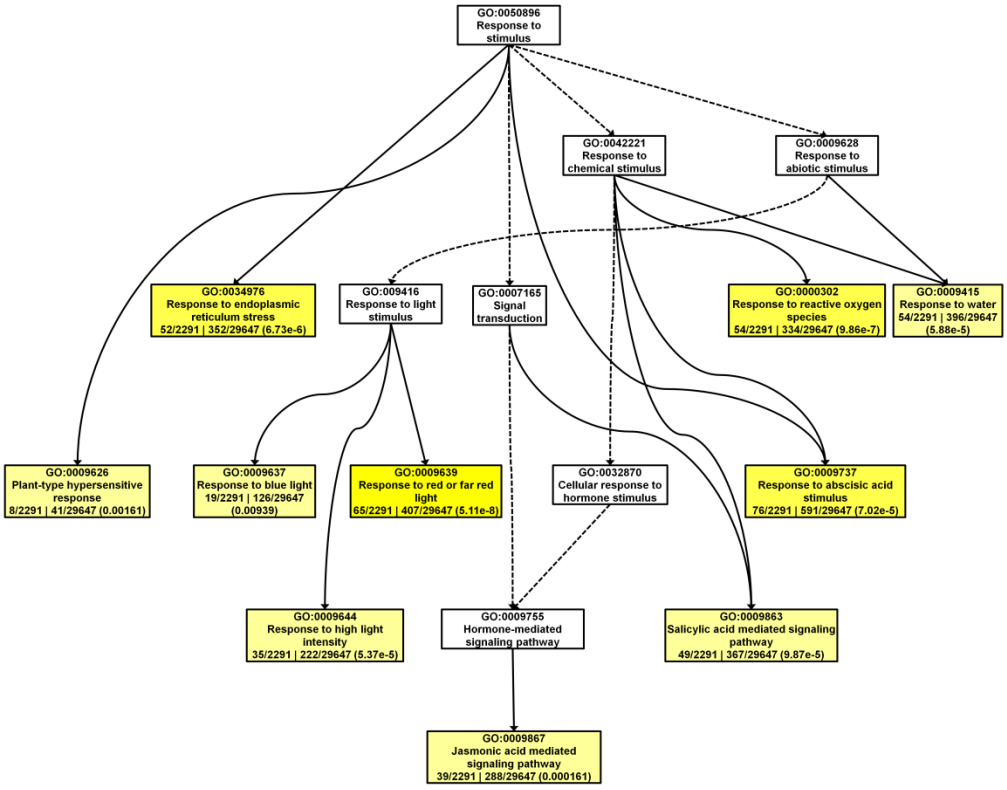

Fig 2. Enriched GO terms associated with stress responses. The directed acyclic graph (DAG) illustrates the relationships of the GO terms over-represented in the CSVd-infected chrysanthemum transcriptome, according to biological processes. The ID and name of the enriched GO terms are described in each rectangle. Detailed information about the number of genes in the test and reference groups, and the FDR value for each GO term can be found in Table S2. The terms at the top are relatively general concepts, while the terms near the bottom indicate more specific concepts. The solid and dotted connecting lines indicate terms that are significantly or not significantly over-represented, respectively. The significance of the over-represented terms corrleates with the darkening color within the rectangle.

Identification of chrysanthemum genes involved in various metabolic pathways

Using the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways implemented in the Blast2GO program (Götz et al., 2008; Kanehisa et al., 2008), 2,451 ESTs were assigned to 3,308 redundant enzymes (779 non-redundant enzymes) involved in various metabolic pathways. The most frequently identified enzymes were protein-serine/threonine kinases (231 ESTs) followed by ubiquitin protein ligases (153 ESTs), ATPases (76 ESTs), and peroxidases (46 ESTs) (Table 4). Furthermore, GO enrichment analysis revealed that many overrepresented GO terms were associated with various metabolic processes such as the carbohydrate biosynthetic process (GO:0016051), glucose metabolic process (GO:0006006), glycerophospholipid biosynthetic process (GO:0046474), isopentenyl diphosphate biosynthetic process mevalonate-independent pathway (GO:0019288), galactolipid biosynthetic process (GO:0019375), carotenoid biosynthetic process (GO:0016117), starch metabolic process (GO:0005982), and fatty acid beta-oxidation (GO:0006635).

Glucose is the most important carbohydrate and is synthesized by photosynthesis. The synthesized energy is very often stored in plants in the form of starch or lipids. In particular, beta-oxidation of fatty acids is important for plant development as well as signaling molecules such as jasmonic acid biosynthesis, which is related to the plant stress response (Baker et al., 2006). In addition, carotenoids are the most important components in plants that furnish flowers and fruits, and they play an important role in photosynthesis (Hirschberg, 2001).

Genes associated with stress and hormone signaling

The GO terms for response to diverse stresses were frequently identified. Such stress-related GO terms involved various

abiotic stresses such as wounding, water, salt stress, heat, reactive oxygen species, and abscisic acid stress (Fig. 2). In addition, some genes were implicated in response to endoplasmic reticulum stress (GO:0034976). Interestingly, GO terms associated with various light stresses such as red or far red light, blue light, and high light intensity were enriched. In addition, 57 genes involved in systemic acquired resistance (GO:0009627) and 8 genes coding for the plant-type hypersensitive response (GO:0009626) were identified (Fig. 2 and Table S2). A previous study demonstrated that viroid infection accumulated transcripts associated with pathogenesis-related (PR) proteins (Gadea et al., 1996). We also identified many ESTs encoding PR proteins such as peroxidase, desaturase-like protein, lipoxygenase, and proteinase inhibitor.
The genes implicated in light stresses were dominantly associated with chloroplasts. Systemic acquired resistance (SAR) is the phenomenon wherein plants develop increased resistance at the whole plant level against pathogen infection. In particular, the eight genes for plant-type hypersensitive response identified, FMO1 (AT1G19250), DND1 (AT5G15410), LSD1 (AT4G20380), CNGC4 (AT5G54250), PBS2 (AT5G51700), RIN4 (AT3G25070), CPR5 (AT5G64930), and CAD1 (AT1G29690).

A recent microarray-based study reported that infection with *Potato spindle tuber viroid* (PSTVd) affected the gene expression involved in diverse hormone signaling, including such hormones as gibberellin and brassinosteroid (Owens et al., 2012). Similarly, genes of the salicylic acid-mediated signaling pathway (GO:0009863), jasmonic acid-mediated signaling pathway (GO:0009867), and regulation of gibberellic acid-mediated signaling pathway (GO:0009937) were frequently identified (Fig. 2 and Table S2). In particular, five genes were identified as being involved in gibberellic acid signaling and are known to be negative regulators of plant growth. For instance, the *RGL2* (AT3G03450) gene is a member of the GRAS transcription factor (TF) family and regulates negatively seed germination and is involved in flower and fruit development (Piskurewicz et al., 2008). *GAI* (AT1G14920) and *RGA1* (AT2G01570) are members of the DELLA protein family and are involved in reduction of ROS (Reactive oxygen species) accumulation upon stresses, as well as the repression of GA-induced vegetative growth and floral initiation (Tyler et al., 2004). *GID1B* (AT3G63010) and *SLY1* (AT4G24210) regulate seed germination by interacting with DELLA proteins (Ariizumi et al., 2011; Griffiths et al., 2006). Recent evidence suggests that jasmonic acid and salicylic acid signaling are important in modulating plant immune responses by DELLA proteins (Navarro et al., 2008). Moreover, a previous study demonstrated the level of gibberellic acid is related to stunting symptoms in chrysanthemums infected by CSVd (Huh et al., 2006). In summary, CSVd infection affected the complex transcriptome and influenced genes involved in stress response and hormone signaling.

Genes associated with the regulation of diverse biological processes and transporting systems

We found several genes involving in many biological processes. For example, 29 genes for trichome differentiation (GO:0010026) and 47 genes for shoot morphogenesis (GO:0010016) were overrepresented. In addition, we also found GO terms for the regulation of the circadian rhythm (GO:0042752) and positive regulation of cellular process (GO:0048522). By contrast, some identified chrysanthemum genes were associated with the negative regulation of some biological processes, such as the negative regulation of seed germination (GO:0010187), negative regulation of programmed cell death (GO:0043069) and negative regulation of the cellular macromolecule biosynthetic process (GO:2000113) (Table S2).

We found many transport-associated GO terms. For instance, three GO terms were related to protein transporting, including protein targeting (GO:0006605), protein secretion (GO:0009306), and protein import (GO:0017038). In addition, ATP transport (GO:0015867), amine transport (GO:0015837), Golgi vesicle transport (GO:0048193), mitochondrial transport (GO:0006839), and nucleotide transport (GO:0006862) (Table S2) terms were identified. Among these terms, a previous study reported that various virus infections led to the significant accumulation of polyamine content (González et al., 1997) and polyamines are regarded as important molecules in plant-microbe interactions (Walters, 2000). Thus, viroid infection might promote cellular transport systems to trigger host defense mechanisms.

Genes associated with transcriptional and translational regulation

Transcription factors (TFs) play an important role in gene expression. To identify potential chrysanthemum transcription factors, we used a database of *Arabidopsis* TFs that includes more than 2,000 TF genes (http://datf.cbi.pku.edu.cn/). Overall, 208 ESTs were homologous to the *Arabidopsis* TFs. They were assigned to 45 different TF families (Table S3). The most frequently identified TF families were C3H (14 TFs) and bZIP (14 TFs) followed by GRAS (13 TFs). In addition, TFs belonging to the MYB, NAC, WRKY, and bHLH families were also identified. Moreover, we identified 45 genes for rRNA processing (GO:0006364) and they were members of photosystem I, photosystem II, sigma factor 2, hcf136, PGRL1A, and PTAC6. Four genes involved in the maturation of SSU-rRNA from tricistronic rRNA transcripts (SSU-rRNA, 5.8S rRNA, LSU-rRNA) (GO:0000462) were also identified.
Interestingly, we found many genes that function in protein modification, such as protein modification by small protein removal (GO:0070646), protein peptidyl-prolyl isomerization (GO:0000413), N-terminal protein myristoylation (GO:0006499), protein maturation (GO:0051604), protein folding (GO:0006457), modification-dependent protein catabolic process (GO:0019941), protein glycosylation (GO:0006486), response to misfolded protein (GO:0051788), and the cullin-RING ubiquitin ligase complex (GO:0031461) (Fig. 3 and Table S2).

Recent studies have provided evidence that translational mechanisms are important for biotic stress tolerance. For instance, the majority of genes assigned to GO terms for misfolded proteins are ubiquitin-conjugating enzymes, which play an important role in post-translational modification and participate in other biological processes such as cell cycle control, transcription, DNA repair, stress response, signaling, immunity, plant growth, embryogenesis, and circadian rhythms (Schulman, 2011). In particular, the overexpression of genes encoding ubiquitin-conjugating enzymes confers enhanced resistance to abiotic stresses such as drought, salt, and water stresses (Lyzenga and Stone, 2012; Wan et al., 2011; Zhou et al., 2010). In addition, the endoplasmic reticulum (ER) in eukaryotes

Table 3. The top 20 InterPro conserved domains frequently identified among 11,600 ESTs.

InterPro No.	No. of hits	Description
IPR011009	200	Protein kinase-like domain
IPR000719	174	Protein kinase, catalytic domain
IPR016040	151	NAD(P)-binding domain
IPR002290	90	Serine/threonine- / dual-specificity protein kinase, Catalytic domain
IPR008271	86	Serine/threonine-protein kinase, active site
IPR001128	69	Cytochrome P450
IPR012677	69	Nucleotide-binding, alpha-beta plait
IPR020635	64	Tyrosine-protein kinase, catalytic domain
IPR000504	59	RNA recognition motif domain
IPR012336	58	Thioredoxin-like fold
IPR015943	57	WD40/YVTN repeat-like-containing domain
IPR016024	56	Armadillo-type fold
IPR013785	52	Aldolase-type TIM barrel
IPR001680	51	WD40 repeat
IPR001245	47	Serine-threonine/tyrosine-protein kinase catalytic domain
IPR011989	47	Armadillo-like helical
IPR002401	44	Cytochrome P450, E-class, group I
IPR013083	44	Zinc finger, RING/FYVE/PHD-type
IPR017853	44	Glycoside hydrolase, superfamily
IPR011046	43	WD40 repeat-like-containing domain

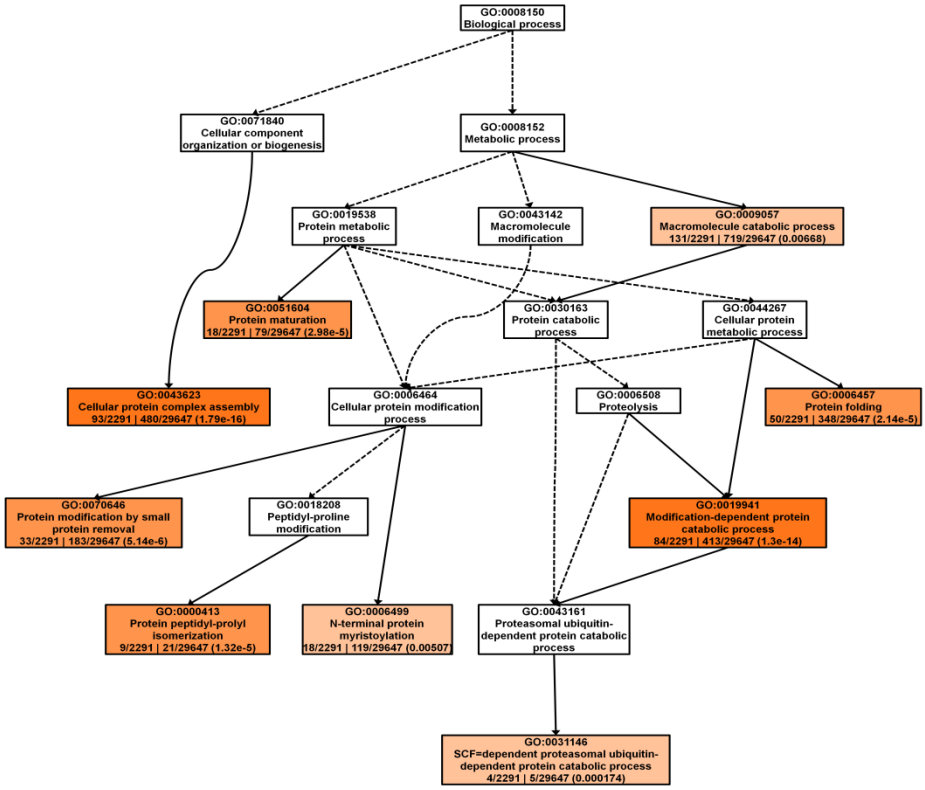

Fig 3. Enriched GO terms associated with protein modification. The directed acyclic graph (DAG) illustrates the relationships of the protein modification GO terms highly over-represented in the CSVd-infected chrysanthemum transcriptome.

functions in protein folding and the maturation of secretory and membrane proteins that are associated with abiotic stress tolerance (Hüttner and Strasser, 2012).

Important organelles associated with CSVd infection

The GO enrichment analysis found that genes for various cellular components such as the nucleus, membrane, cytoplasm, plastid, and mitochondrion were enriched in CSVd-infected chrysanthemum samples. Of them, a majority of ESTs were associated with two important organelles, the plastids and mitochondria (Fig. 4). The GO terms related to plastids such as chloroplasts (GO:0009507), plastid stroma (GO:0009532), thylakoids (GO:0009579), and the photosystem (GO:0009521).

Table 4. The top 20 enzymes frequently identified among the chrysanthemum ESTs.

Index	Enzyme code	Annotation	No. of ESTs
1	EC:2.7.11.0	Protein-serine/threonine kinases	231
2	EC:6.3.2.19	Ubiquitin protein ligase	153
3	EC:3.6.1.3	ATPase	76
4	EC:1.11.1.7	Peroxidase	46
5	EC:3.4.21.0	Serine endopeptidases	30
6	EC:1.1.1.0	Alcohol dehydrogenase	28
7	EC:3.2.1.2	Beta-amylase	27
8	EC:3.4.22.0	Cysteine endopeptidases	26
9	EC:2.5.1.18	Glutathione transferase	25
10	EC:2.3.1.48	Histone acetyltransferase	24
11	EC:3.2.1.0	Glycosidases	24
12	EC:1.14.13.21	Flavonoid 3'-monooxygenase	22
13	EC:3.1.1.1	Carboxylesterase	22
14	EC:2.4.1.203	Trans-zeatin O-beta-D-glucosyltransferase	21
15	EC:2.4.1.215	Transferases	21
16	EC:3.6.3.6	ATP phosphohydrolase	21
17	EC:1.11.1.9	Glutathione peroxidase	20
18	EC:3.1.3.16	Phosphoprotein phosphatase	20
19	EC:3.4.25.0	Threonine endopeptidases	19
20	EC:5.2.1.8	Peptidylprolylisomerase	19

The detailed information for each enzyme can be found in KEGG: Kyoto Encyclopedia of Genes and Genomes (http://www.genome.jp/kegg/) with corresponding enzyme code.

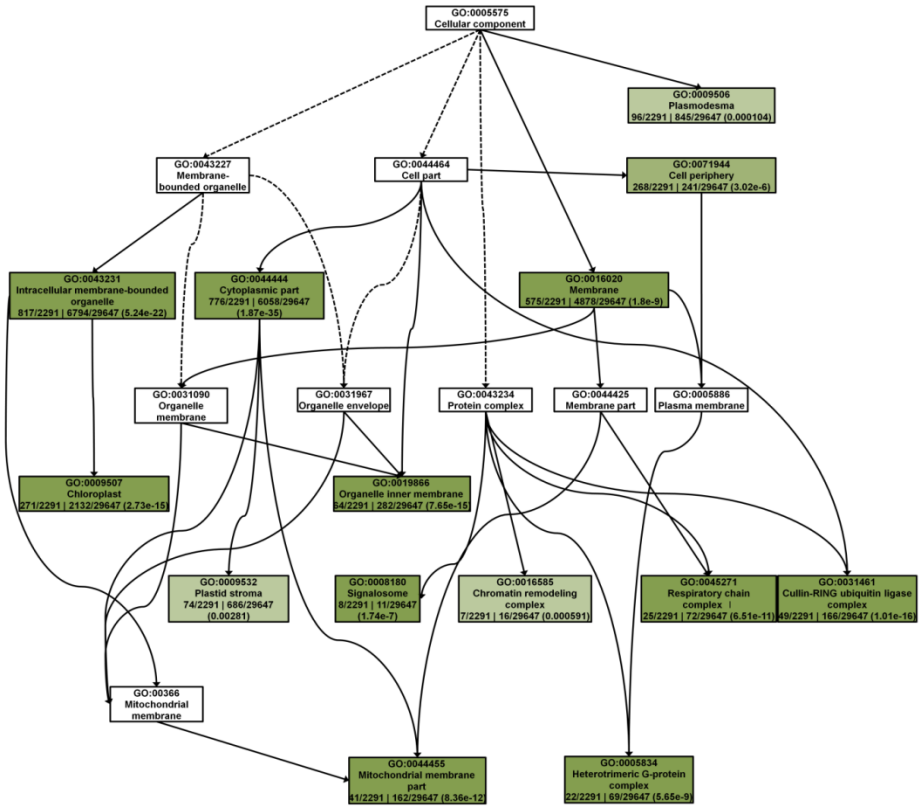

Fig 4. Enriched GO terms associated with cellular component. The directed acyclic graph (DAG) illustrates the relationships of the cellular component GO terms highly over-represented in the CSVd-infected chrysanthemum transcriptome.

Moreover, several GO terms were associated with mitochondrion, including the mitochondrial membrane (GO:0044455) and respiratory chain complex I (GO:0045271) (Fig. 4).

In addition, we found 96 genes associated with the plasmodesmata (GO:0009506) (Fig. 4), which is an intercellular channel establishing cell-to-cell connections in higher plants. Among these genes, PDLP7 (AT5G37660) is a

Table 5. Comparison of the identified chrysanthemum genes with those of five representative plant species for which whole genome sequences are available.

Species	*Populustrichocarpa*	*Arabidopsis thaliana*	*Oryza sativa*	*Physcomitrealla patens*	*Chlamydomon reinhardtii*
No. of orthologous genes	8,267	8,083	7,815	7,124	4,814

member of plasmodesmata-located proteins (PDLP), that are type I membrane receptor-like proteins (Thomas et al., 2008). Furthermore, a recent study demonstrated that such PDLP proteins are positive regulators of plant virus movement by interacting with viral movement proteins (Amari et al., 2010).

Thus, it is likely that the identified PDLP orthologous genes could function in viroid movement. In addition, genes for cytoskeletal features, such as microtubule- (GO:0007020) and actin-related genes (GO:0045010), were identified. The cytoskeletal inclusion of actin filaments and microtubules plays an important role in the cell-to-cell movement of plant viruses (Heinlein et al., 1995; Wright et al., 2007). However, we do not know whether such features are also required for viroid movement.

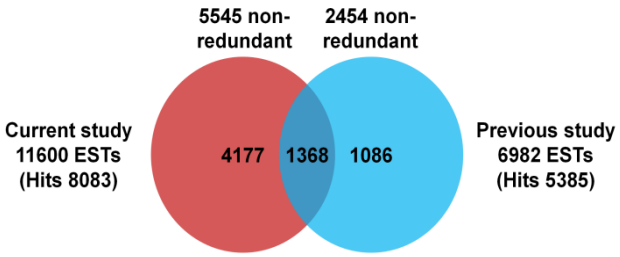

Fig 5. Comparative study identifies chrysanthemum genes expressed preferentially in the flower. ESTs in each data set were converted to corresponding *Arabidopsis* genes. After removal of redundant genes, only non-redundant *Arabidopsis* genes were used for comparison.

Comparison of diverse transcriptome analyses associated with viroid infection

The most comprehensive way in molecular biology to understand changes of the host following viroid infection might be transcriptome analysis. Previously, a few studies have applied microarray systems to monitor the expression patterns of host genes in response to viroid infection. For instance, a macroarray composed of 768 clones derived from a tomato cDNA library was used to examine host gene expression by two different strains of PSTVd and this study found that genes associated with a defense/stress response, cell wall structure, chloroplast function, protein metabolism, and other diverse functions were differentially expressed (Itaya et al., 2002). A cDNA microarray carrying 21,081 unigenes were used to examine the expression patterns of citrus plants in response to *Citrus exocortis viroid* (CEVd) at different time points and unveiled that the genes responsible for chloroplast, cell wall, peroxidase and symporter activities were differentially expressed (Rizza et al., 2012). Furthermore, tomato genome arrays were used to examine transcriptional changes upon PSTVd infection and revealed that the expression of hormone signaling-related genes was significantly changed (Owens et al., 2012). At a small scale, using the differential display technique (DDRT-PCR), several differentially expressed genes of Etrog citron plants in response to *Citrus viroid IIIb* were identified (Tessitori et al., 2007). Our study provides a list of chrysanthemum genes after CSVd infection, the enriched gene functions might be similar to those of previous studies associated with viroid infection. For instance, genes for chloroplast function, stresses, transporting system, and protein metabolism were also differentially expressed in other two studies (Itaya et al., 2002; Rizza et al., 2012). Moreover, we found chrysanthemum genes involved in hormone signaling, plasmodesmata, actins, microtubules, protein modification, protein transportation, trichome differentiation and shoot morphogenesis (Itaya et al., 2002; Rizza et al., 2012b). Currently, we are establishing a cDNA microarray containing approximately 13,000 chrysanthemum genes to carry out a transcriptome analysis following viroid infection with CSVd and CChMVd. Such an analysis will elucidate the differentially expressed gene sets in the same host following infection with two different viroids.

Comparison of two chrysanthemum EST datasets

The number of nucleotide sequences for the chrysanthemum is very limited. Most of them were derived from a previous study performing EST analysis of chrysanthemum flowers in which 6,982 ESTs were identified (Chen et al., 2009). It might be of interest to compare different EST datasets. For that, each EST was blasted against the *Arabidopsis* proteome using tblastX (less than 1E-5 as a cutoff). As a result, 8,083 and 5,385 ESTs from the current and previous studies, respectively, were matched to *Arabidopsis*, producing 5,545 and 2,454 non-redundant *Arabidopsis* orthologous genes, respectively. In fact, 11,600 ESTs were derived from leaves, stems, and roots. A previous study identified only ESTs from the flowers of chrysanthemum. A total of 1,368 genes were commonly present in both studies. In addition, 4,177 and 1,086 genes were only detected in the current and previous studies, respectively. It is likely that 1,086 genes might be chrysanthemum genes preferentially expressed in flower tissues (Fig. 5).

Conservation and divergence of chrysanthemum genes

It might of interest to examine the evolutionary relationship of identified chrysanthemum ESTs with those of five representative plant species for which whole genome sequences are available. For that, a total of identified 11,600 ESTs were BLASTed against five plant proteomes using tblastX (less than 1E-5 as a cutoff). The five plant species included *Arabidopsis thaliana*, *Oryza sativa*, *Populus trichocarpa*, *Physcomitrella patens*, and *Chlamydomonas reinhardtii*. The results indicated that the transcriptome of the chrysanthemum is most close to that of *Populus trichocarpa* (8,267 orthologous proteins), followed by *Arabidopsis thaliana* (8,083 orthologous proteins) and *Oryza sativa* (7,815 orthologous proteins). More than one-fourth of chrysanthemum ESTs might be species specific

Fig 6. Pictures of chrysanthemum cultivar Shinma and CSVd detection by RT-PCR. (A) Shoots of chrysanthemum cultivar Shima grown in medium (B) Flowers of chrysanthemum cultivar Shima (C) Detection of CSVd in chrysanthemum cultivar Shima by RT-PCR. Chrysanthemum *Actin* gene was used as positive control.

(Vandepoele and Van de Peer, 2005). In addition, 7,124 and 4,814 chrysanthemum genes are orthologous to those of *Physcomitrella patens* and *Chlamydomonas reinhardtii*, respectively (Merchant et al., 2007; Rensing et al., 2008), implying that at least half of chrysanthemum genes were derived from algae (Table 5, Table S4)

Materials and Methods

Plant preparation and total RNA isolation

Chrysanthemum morifolium Ramatuelle cultivar Shinma infected by CSVd was obtained from the National Institute of Horticultural and Herbal Science in Korea and propagated in Murashige and Skoog (MS) solid medium containing 3% sucrose and 0.1% activated carbon (Fig. 6A and 6B). CSVd infection was confirmed by reverse-transcription polymerase chain reaction (RT-PCR) (Fig. 6C). Four-week old plants after cutting were used for RNA extraction. Plant tissues, including leaves, stems and roots of the three different CSVd-infected chrysanthemums were harvested and pooled. An RNeasy Plant Mini Kit was used (Qiagen, Valencia, USA) to obtain RNA samples, which were treated with a TURBO DNA-free Kit (Ambion, Austin, USA) to obtain high quality RNAs.

Library preparation

A total of 50 μg total RNA was used to isolate mRNA using the FastTrack™ MAG Micro mRNA Isolation Kit following manufacturer's protocol (Life Technologies, Carlsbad, USA). After quantification of the isolated mRNA using an Agilent RNA 6000 NANO CHIP (Agilent technologies, Santa Clara, USA), 200 ng of mRNA was utilized for cDNA library construction. According to the protocol for the cDNA rapid library preparation, mRNAs fragmented by $ZnCl_2$ and heat treatment were used to synthesize first-strand cDNA using reverse transcriptase and random primers. After removing the RNA template, double-stranded (ds) cDNA was synthesized. Short adapters were ligated onto both ends of ds cDNA, which were then polished. The adapters are useful to prime sequences for both amplification and sequencing, and they can be used as the "sequencing key" for base calling. Finally, the quality of the constructed cDNA fragments was measured with a 2100 BioAnalyzer (Agilent technologies, Santa Clara, USA). The RNA integrity number (RIN) is 8.8.

Emulsion PCR

The cDNA fragments were immobilized on DNA capture beads and then resuspended in amplification solution. The mixture was emulsified and PCR amplification was conducted. After that, bead-immobilized and amplified DNA fragments were obtained.

Sequencing run

The DNA bead incubation mix carrying DNA polymerase was added to the DNA-carrying beads which were further layered onto a PicoTiterPlate device (PTP) with enzyme beads. The PTP was loaded into the Genome Sequencer FLX 454 Instrument, and sequencing reagents were sequentially flowed over the plate. Information from all the wells of the PTP could be recorded simultaneously by the CCD camera and could be processed in real time by the onboard computer. The data were saved in standard flowgram format (SFF) files for further analysis. The FLX 454 sequencing was performed at Macrogen Inc. in Korea (http://www.macrogen.com).

Data processing

The obtained SFF data were subjected to data processing using the Roche GS FLX software (version 2.6). The obtained reads from the SFF file were de novo assembled using a GS de novo assembler (version 2.6) with default parameters. The size of N50 isotig was 634 bp. Various contaminants, low quality and low-complexity sequences were removed using two different sequence cleanup programs: SeqClean (http://sourceforge.net/projects/seqclean/) and Lucy (http://lucy.sourceforge.net/) (Patel and Jain, 2012). The obtained SFF file is accessible as "SRX162109" in the NCBI SRA database (http://www.ncbi.nlm.nih.gov/sra/). The 11,600 EST sequences are available upon request.

Annotation of chrysanthemum ESTs

The Blast2GO program was used to annotate the functions of the 11,600 chrysanthemum ESTs [21]. All nucleotide sequences, in the FASTA format, were blasted against proteins in the

Swiss-Prot database using a blastX search (using 1E-3 as a cutoff) implemented in Blast2GO to identify the potential functions of the genes. In addition, GO terms, enzyme codes, and InterPro domains for each chrysanthemum EST were obtained using the Blast2GO program with default parameters.

Gene ontology enrichment analysis

To decipher the significantly enriched chrysanthemum genes involved in CSVd infection, a GO enrichment analysis was performed using the GOEAST program. As the full chrysanthemum genome is not available, GO annotations for the orthologous *Arabidopsis* proteins were used. For the analysis, all ESTs were mapped onto the *Arabidopsis* proteome, resulting in 8,083 orthologous *Arabidopsis* proteins. Next, 5,546 corresponding non-redundant *Arabidopsis* genes (loci) were used for GO enrichment analysis with the following parameters: "Fisher Exact Test" statistical test and the Adrian Alexa's improved weighted scoring, without multi-test adjustment method, with a 0.01 enrichment significance level and a minimum of five associated genes in reference. Due to frequent updates in the GOEAST database, the results of future analyses might be different. The GO enrichment was performed on June 30th, 2012.

Conclusions

The understanding of plant-viroid interaction is necessary to prevent disease. Unfortunately, few viroid studies have been performed compared to those examining plant viruses. Little is known about the host proteins that regulate viroid replication and movement. To establish fundamental studies on plant-viroid interaction, we performed an EST analysis of samples from CSVd-infected chrysanthemums using next generation sequencing, which led to the identification of 11,600 ESTs. A GO enrichment analysis identified a total of 144 GO terms. These terms provided gene functions that are significantly overrepresented during CSVd infection. The majority of genes were responsible for diverse metabolic pathways, various stress responses, transcription, translation, and transportation. The most striking findings were the genes associated with plant hormone signaling, such as those related with gibberellic acid, a negative regulator of plant growth. Furthermore, the identification of genes coding for several cellular components such as chloroplasts, plasmodesmata, actins, and microtubules implies the possible roles of such organelles in viroid replication and movement. The identified 11,600 ESTs in this study provide a comprehensive overview of the transcriptome of a chrysanthemum infected by CSVd. They will provide a good resource for future plant-viroid interaction studies.

Acknowledgments

This research was supported in part by Grants from the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2011-0010027), Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ009764), the Next-Generation BioGreen 21 Program (No. PJ00798402), and the Agenda Program (PJ00922904) through Rural Development Administration, Republic of Korea.

.

Supplementary data (online only): Table S1. Detailed information for the 11,600 identified chrysanthemum ESTs.
Table S2. Significantly enriched GO terms in the CSVd-infected chrysanthemum transcriptome.
Table S3. A total of 208 chrysanthemum transcription factors present in chrysanthemum ESTs.
Table S4. A comparative analysis of 11,600 chrysanthemum ESTs against five plant genomes.

References

Amari K, Boutant E, Hofmann C, Schmitt-Keichinger C, Fernandez-Calvino L, Didier P, Lerich A, Mutterer J, Thomas CL, Heinlein M (2010) A family of plasmodesmal proteins with receptor-like properties for plant viral movement proteins. PLoS Pathog. 6: e1001119

Ariizumi T, Lawrence PK, Steber CM (2011) The role of two f-box proteins, SLEEPY1 and SNEEZY, in Arabidopsis gibberellin signaling. Plant Physiol. 155: 765-775

Baker A, Graham IA, Holdsworth M, Smith SM, Theodoulou FL (2006) Chewing the fat: β-oxidation in signalling and development. Trends Plant Sci. 11: 124-132

Chen S, Miao H, Chen F, Jiang B, Lu J, Fang W (2009) Analysis of expressed sequence tags (ESTs) collected from the inflorescence of chrysanthemum. Plant Mol Biol Rep. 27: 503-510

Courtney-Gutterson N, Napoli C, Lemieux C, Morgan A, Firoozabady E, Robinson KEP (1994) Modification of flower color in florist's chrysanthemum: production of a white–flowering variety through molecular genetics. Biotechnology (N Y). 12: 268-271

Dauchot N, Mingeot D, Purnelle B, Muys C, Watillon B, Boutry M, Van Cutsem P (2009) Construction of 12 EST libraries and characterization of a 12,226 EST dataset for chicory (*Cichorium intybus*) root, leaves and nodules in the context of carbohydrate metabolism investigation. BMC Plant Biol. 9: 14

Ding B (2010) Viroids: self-replicating, mobile, and fast-evolving noncoding regulatory RNAs. Wiley Interdiscip Rev: RNA. 1: 362-375

Dowrick G (1953) The chromosomes of chrysanthemum. Heredity. 7: 59-72

Dowrick G, El-Bayoumi A (1966) The origin of new forms of the garden chrysanthemum. Euphytica. 15: 32-38

Drewlow L, Ascher P, Widmer R (1973) Genetic studies of self incompatibility in the garden chrysanthemum, *Chrysanthemum morifolium* Ramat. Theor Appl Genet. 43: 1-5

Fernández P, Paniego N, Lew S, Hopp H, Heinz R (2003) Differential representation of sunflower ESTs in enriched organ-specific cDNA libraries in a small scale sequencing project. BMC Genomics. 4: 40

Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A (2008) High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 36: 3420-3435

Gadea J, Mayda ME, Conejero V, Vera P (1996) Characterization of defense-related genes ectopically expressed in viroid-infected tomato plants. Mol Plant Microbe Interact. 9: 409-415

González E, Mosquera M, San José M, Díaz T (1997) Influence of virus on the chlorophyll, carotenoid and polyamine contents in grapevine microcuttings. J Phytopathol. 145: 185-187

Griffiths J, Murase K, Rieu I, Zentella R, Zhang ZL, Powers SJ, Gong F, Phillips AL, Hedden P, Sun T (2006) Genetic characterization and functional analysis of the GID1 gibberellin receptors in *Arabidopsis*. Plant Cell. 18: 3399-3414

Hüttner S, Strasser R (2012) Endoplasmic reticulum-associated degradation of glycoproteins in plants. Front Plant Sci. 3

Haseloff J, Symons RH (1981) Chrysanthemum stunt viroid: primary sequence and secondary structure. Nucleic Acids Res. 9: 2741-2752

Heinlein M, Epel BL, Padgett HS, Beachy RN (1995) Interaction of tobamovirus movement proteins with the plant cytoskeleton. Science. 270: 1983-1985

Hirschberg J (2001) Carotenoid biosynthesis in flowering plants. Curr Opin Plant Biol. 4: 210-218

Hitmi A, Coudret A, Barthomeuf C (2000) The production of pyrethrins by plant cell and tissue cultures of *Chrysanthemum cinerariaefolium* and Tagetes species. Crit Rev Biochem Mol Biol. 35: 317-337

Hosokawa M, Matsushita Y, Uchida H, Yazawa S (2006) Direct RT-PCR method for detecting two chrysanthemum viroids using minimal amounts of plant tissue. J Virol Methods. 131: 28-33

Huh EJ, Lee SK, Chung BN, Lee IJ, Choi SY (2006) Changes of growth and gibberellin contents in chrysanthemum by infection of *Chrysanthemum stunt viroid*. Hort Environ Biotechnol. 47: 366-370

Itaya A, Matsuda Y, Gonzales RA, Nelson RS, Ding B (2002) *Potato spindle tuber viroid* strains of different pathogenicity induces and suppresses expression of common and unique genes in infected tomato. Mol Plant Microbe Interact. 15: 990-999

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T (2008) KEGG for linking genomes to life and the environment. Nucleic Acids Res. 36: D480-D484

Laitinen RAE, Immanen J, Auvinen P, Rudd S, Alatalo E, Paulin L, Ainasoja M, Kotilainen M, Koskela S, Teeri TH (2005) Analysis of the floral transcriptome uncovers new regulators of organ determination and gene families related to flower organ differentiation in *Gerbera hybrida* (Asteraceae). Genome Res. 15: 475-486

Lyzenga WJ, Stone SL (2012) Abiotic stress tolerance mediated by protein ubiquitination. J Exp Bot. 63: 599-616

Merchant SS, Prochnik SE, Vallon O, Harris EH, Karpowicz SJ, Witman GB, Terry A, Salamov A, Fritz-Laylin LK, Maréchal-Drouard L (2007) The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science. 318: 245-250

Miyazawa M, Hisama M (2003) Antimutagenic activity of flavonoids from *Chrysanthemum morifolium*. Biosci Biotechnol Biochem. 67: 2091-2099

Morozova O, Marra MA (2008) Applications of next-generation sequencing technologies in functional genomics. Genomics. 92: 255-264

Navarro B, Flores R (1997) *Chrysanthemum chlorotic mottle viroid*: unusual structural properties of a subgroup of self-cleaving viroids with hammerhead ribozymes. Proc Natl Acad Sci USA. 94: 11262

Navarro L, Bari R, Achard P, Lisón P, Nemri A, Harberd NP, Jones JDG (2008) DELLAs control plant immune responses by modulating the balance of jasmonic acid and salicylic acid signaling. Curr Biol. 18: 650-655

Ohlrogge J, Benning C (2000) Unraveling plant metabolism by EST analysis. Curr Opin Plant Biol. 3: 224-228

Owens RA, Tech K, Shao J, Sano T, Baker CJ (2012) Global analysis in tomato gene expression during *Potato spindle tuber viroid* infection reveals a complex array of changes affecting hormone signaling. Mol Plant Microbe Interact.

Patel RK, Jain M (2012) NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS One. 7: e30619

Piskurewicz U, Jikumaru Y, Kinoshita N, Nambara E, Kamiya Y, Lopez-Molina L (2008) The gibberellic acid signaling repressor RGL2 inhibits *Arabidopsis* seed germination by stimulating abscisic acid synthesis and ABI5 activity. Plant Cell. 20: 2729-2745

Randles JW, Hadidi A, Flores R, Semancik J (2003) Economic impact of viroid diseases. Viroids. 3-11

Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud P-F, Lindquist EA, Kamisugi Y (2008) The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science. 319: 64-69

Rizza S, Conesa A, Juarez J, Catara A, Navarro L, Duran-Vila N, Ancillo G (2012) Microarray analysis of Etrog citron (*Citrus medica* L.) reveals changes in chloroplast, cell wall, peroxidase and symporter activities in response to viroid infection. Mol Plant Pathol. 13: 852–864

Roest S, Bokelmann G (1975) Vegetative propagation of *Chrysanthemum morifolium* Ram. *in vitro*. Scientia Hort. 3: 317-330

Ronald W, Ascher P (1975) Self compatibility in garden chrysanthemum: occurrence, inheritance and breeding potential. Theor Appl Genet. 46: 45-54

Rout G, Das P (1997) Recent trends in the biotechnology of *Chrysanthemum*: a critical review. Scientia Hort. 69: 239-257

Schulman BA (2011) Twists and turns in ubiquitin-like protein conjugation cascades. Protein Sci. 20: 1941-1954

Shendure J, Ji H (2008) Next-generation DNA sequencing. Nat Biotechnol. 26: 1135-1145

Stewart R, Dermen H (1970) Somatic genetic analysis of the apical layers of chimeral sports in chrysanthemum by experimental production of adventitious shoots. Amer J Bot. 1061-1071

Teixeira da Silva JA (2004) Ornamental chrysanthemums: improvement by biotechnology. Plant Cell Tiss Org Cult. 79: 1-18

Tessitori M, Maria G, Capasso C, Catara G, Rizza S, De Luca V, Catara A, Capasso A, Carginale V (2007) Differential display analysis of gene expression in Etrog citron leaves infected by *Citrus viroid III*. Biochim Biophys Acta. 1769: 228-235

Thomas CL, Bayer EM, Ritzenthaler C, Fernandez-Calvino L, Maule AJ (2008) Specific targeting of a plasmodesmal protein affecting cell-to-cell communication. PLoS Biol. 6: e7

Tyler L, Thomas SG, Hu J, Dill A, Alonso JM, Ecker JR, Sun T (2004) DELLA proteins and gibberellin-regulated seed germination and floral development in Arabidopsis. Plant Physiol. 135: 1008-1019

Vandepoele K, Van de Peer Y (2005) Exploring the plant transcriptome through phylogenetic profiling. Plant Physiol. 137: 31-42

Walters D (2000) Polyamines in plant-microbe interactions. Physiol Mol Plant Pathol. 57: 137-146

Wan X, Mo A, Liu S, Yang L, Li L (2011) Constitutive expression of a peanut ubiquitin-conjugating enzyme gene in *Arabidopsis* confers improved water-stress tolerance through regulation of stress-responsive gene expression. J Biosci Bioeng. 111: 478-484

Wright KM, Wood NT, Roberts AG, Chapman S, Boevink P, MacKenzie KM, Oparka KJ (2007) Targeting of TMV movement protein to plasmodesmata requires the actin/ER network; evidence from FRAP. Traffic. 8: 21-31

Zhou GA, Chang RZ, Qiu LJ (2010) Overexpression of soybean ubiquitin-conjugating enzyme gene *GmUBC2* confers enhanced drought and salt tolerance through modulating abiotic stress-responsive gene expression in *Arabidopsis*. Plant Mol Biol. 72: 357-367