BMC Genomics

RESEARCH ARTICLE

Open Access

Transcriptomic analysis of grain amaranth (*Amaranthus hypochondriacus*) using 454 pyrosequencing: comparison with *A. tuberculatus*, expression profiling in stems and in response to biotic and abiotic stress

John P Délano-Frier[1*], Hamlet Avilés-Arnaut[1], Kena Casarrubias-Castillo[1], Gabriela Casique-Arroyo[1], Paula A Castrillón-Arbeláez[1], Luis Herrera-Estrella[2], Julio Massange-Sánchez[1], Norma A Martínez-Gallardo[1], Fannie I Parra-Cota[1], Erandi Vargas-Ortiz[1] and María G Estrada-Hernández[1,3]

Abstract

Background: *Amaranthus hypochondriacus*, a grain amaranth, is a C4 plant noted by its ability to tolerate stressful conditions and produce highly nutritious seeds. These possess an optimal amino acid balance and constitute a rich source of health-promoting peptides. Although several recent studies, mostly involving subtractive hybridization strategies, have contributed to increase the relatively low number of grain amaranth expressed sequence tags (ESTs), transcriptomic information of this species remains limited, particularly regarding tissue-specific and biotic stress-related genes. Thus, a large scale transcriptome analysis was performed to generate stem- and (a)biotic stress-responsive gene expression profiles in grain amaranth.

Results: A total of 2,700,168 raw reads were obtained from six 454 pyrosequencing runs, which were assembled into 21,207 high quality sequences (20,408 isotigs + 799 contigs). The average sequence length was 1,064 bp and 930 bp for isotigs and contigs, respectively. Only 5,113 singletons were recovered after quality control. Contigs/isotigs were further incorporated into 15,667 isogroups. All unique sequences were queried against the nr, TAIR, UniRef100, UniRef50 and Amaranthaceae EST databases for annotation. Functional GO annotation was performed with all contigs/isotigs that produced significant hits with the TAIR database. Only 8,260 sequences were found to be homologous when the transcriptomes of *A. tuberculatus* and *A. hypochondriacus* were compared, most of which were associated with basic house-keeping processes. Digital expression analysis identified 1,971 differentially expressed genes in response to at least one of four stress treatments tested. These included several multiple-stress-inducible genes that could represent potential candidates for use in the engineering of stress-resistant plants. The transcriptomic data generated from pigmented stems shared similarity with findings reported in developing stems of Arabidopsis and black cottonwood (*Populus trichocarpa*).

Conclusions: This study represents the first large-scale transcriptomic analysis of *A. hypochondriacus*, considered to be a highly nutritious and stress-tolerant crop. Numerous genes were found to be induced in response to (a)biotic stress, many of which could further the understanding of the mechanisms that contribute to multiple stress-resistance in plants, a trait that has potential biotechnological applications in agriculture.

* Correspondence: jdelano@ira.cinvestav.mx
[1]Unidad de Biotecnología e Ingeniería Genética de Plantas, (Cinvestav-Unidad Irapuato) Km 9.6 del Libramiento Norte Carretera Irapuato-León. Apartado Postal 629, C.P. 36821, Irapuato, Gto., México
Full list of author information is available at the end of the article

Background

The genus *Amaranthus* L. (Caryophyllales: Amarantha-ceae) comprises C4 dicotyledonous herbaceous plants classified into approximately 70 species. It has a worldwide distribution, although most species are found in the warm temperate and tropical regions of the world [1,2]. Many amaranth species are cultivated as ornamentals or a source of highly nutritious pseudocereals (e.g. *grain amaranths*) and vegetables; others, are notoriously aggressive weeds that affect many agricultural areas of the world [3,4]. The grain amaranths (predominantly *Amaranthus hypochon-driacus* L., *A. cruentus* L., and *A. caudatus* L.) are ancestral crops native to the New World. They are classified along with their putative progenitor species (*A. hybridus* L., *A. quitensis* H.B.K., and *A. powellii* S. Wats.) in what is known as the *A. hybridus* complex [5]. Restricted for cen-turies to a limited cultivation in Meso America as a result of religious intolerance [6], grain amaranths have gradually acquired renewed interest due to their various nutritional [7-12] and health-related traits [13], in addition to their highly desirable agronomic characteristics. These charac-teristics offer a viable alternative to cereals and other crops in many stressful agricultural settings, particularly those where soil moisture conditions vary considerably between growing seasons [14-16]. The increased ability to withstand drought stress that characterizes grain amaranth is closely related to its superior water use efficiency (WUE) [17-20], variously defined as the ratio of economic yield to evapo-transpiration or of the amount CO_2 assimilated to water loss [21,22]. WUE in grain amaranth has been found to be higher than in other C3 and C4 crops, includ-ing wheat, corn, cotton and sorghum [23]. Moreover, the high salt tolerance of grain amaranth has also been asso-ciated with a high WUE [16]. The drought-tolerance of grain amaranth has been attributed to the inherently stress-attenuating physiology of the C4 pathway, an inde-terminate flowering habit and the capacity to grow long taproots and develop an extensive lateral root system in response to water shortage in the soil [20,24,25]. Recently, the results of a combined proteomic/genomic approach suggested that amaranth's root response to drought stress involves a coordinated response that includes osmolyte accumulation and the activation of stress-related genes needed for the scavenging of reactive oxygen species, pro-tein stabilization and transcriptional regulation of plant growth [26,27].

The use of molecular tools for the advanced genomic study of the genus *Amaranthus* has recently increased, with at least six published reports appearing in the last three years. The construction of a bacterial artificial chro-mosome (BAC) library for *A. hypochondiacus* representing a 10.6-X coverage of its haploid genome content was reported in 2008 [28]. Shortly afterwards, this BAC library was utilized to generate a set of microsatellite markers for the grain amaranths, which were used to clarify taxonomic relationships within the *A. hybridus* complex. Additional applicability for these microsatellite markers for the study of other economically important species within the *Amaranthus* genus, including weeds and ornamentals, was proposed [29,30]. The utilization of next-generation 454 pyrosequencing technology was subsequently explored as a tool to obtain genomic data for waterhemp (*A. tuberculatus*), a notorious weed of maize and soybean crops in the USA [31]. The sequence data obtained (43 Mbp), which covered 10% of this spe-cies' genome, included the nearly complete sequence of the chloroplast genome and revealed genomic data per-taining herbicide resistance genes, simple sequence repeat markers, and repeated elements (e.g., transpo-sons). This materialized later with the publication of a deep coverage of waterhemp's transcriptome that yielded a total of 44,469 unigenes, 49% of which displayed highly significant similarities to *Arabidopsis* proteins [32]. Moreover, this study generated preliminary sequence information for all of the major herbicide target-site genes for which waterhemp has documented resistance, in addition to two other herbicide targets not previously reported as having evolved resistance in any plant spe-cies. Similarly impressive results were obtained when more than 500 Mbp sequence data, derived from a single 454-pyrosequencing run, were utilized in combination with novel genomic reduction protocol to discover thou-sands of single nucleotide polymorphisms in different populations of *A. caudatus* [33].

The information regarding resistance responses to insects and pathogens in amaranth is relatively scarce. The limited number of defense-related genes reported includes protease and α-amylase inhibitors, agglutinins, anti-microbial peptides and ribosome-inactivating pro-teins [34-39]. This information, however, was comple-mented by a recent study describing several more insect- and pathogen-induced genes [40]. Similarly limited is the genetic information underlying the mechanisms that con-fer amaranth with its capacity to withstand drought and/ or saline stress, although several abiotic-stress-related genes have been identified in amaranth and in phylogen-etically related species such as spinach, cultivated and wild species of beet root, *Mesembryanthemum crystalli-num* and the halophytes *Suaeda* spp., *Salicornia* spp., and *Atriplex* spp. [26,27,40-50].

In this study, the results derived from a large-scale transcriptomic analysis of *A. hypochondriacus* plants performed by massive parallel pyrosequencing technol-ogy, are described. The data includes genes found to be specifically- or highly-expressed in stems and also in leaves under four different stress conditions (drought

and salt stress, insect herbivory and bacterial infection). This allowed the identification of several stress-responsive genes, including many with unknown function and/ or that are expressed in multiple conditions of stress. These may constitute potentially novel mechanisms utilized by this, and related plant species, to deal with highly unfavorable conditions. A comparison of the *A. hypochondriacus* and *A. tuberculatus*, a weedy amaranth species, transcriptomes yielded low levels of similarity. Annotation of homologous transcripts in both species indicated that the majority was associated with genes required for basic biological processes, although an important fraction of them included abiotic stress-related genes.

Methods

Sample preparation for 454 sequencing

Seeds of *Amaranthus hypochondriacus* cultivar Revancha and of accession 38040 (origin: India) were kindly provided by E. Espitia (INIFAP, México) and D. Brenner (USDA, Iowa State University, Ames, IA), respectively. Seeds were germinated in 60-well germinating trays filled with a sterile soil preparation composed of a general soil mixture (three parts Sunshine Mix 3TM [SunGro Horticulture, Bellevue, WA], one part loam, two parts mulch, one part vermiculite [SunGro Hort] and one part perlite [Termolita S.A., Nuevo León, México] and coconut paste [Hummert de México, Morelos, México] in a 1:1 v/v relation). The trays were maintained in a growth chamber kept at 26°C, ≈75% R.H. and with a 16: 8 h light (at approximately 300 µmol m^{-2} s^{-1}) dark photoperiod. Amaranth plantlets were subsequently transplanted to 1.3-L plastic pots, containing sterile general soil mixture, 21 days after germination. They were fertilized once, one week after transplant, with a 20:10:20 (N: P: K) nutrient soil drench solution according to the manufacturer's instructions (Peters Professional; Scotts-Sierra Horticultural Products, Marysville, OH, USA). Plants having six expanded leaves were employed for experimentation. Total RNA was obtained from leaves (*A. hypochondriacus* cv. Revancha) or pigmented stems (*A. hypochondriacus* India 38040) using the Trizol reagent (Invitrogen Corp., Carlsbad, CA, USA) as instructed, treated with RNAase-free DNAase and re-purified with the RNeasy kit (Qiagen, Valencia, CA, USA) following the manufacturer's protocol. Different sources of RNA were used to generate the six cDNA libraries employed for pyrosequencing runs: i) leaves of intact plants grown under natural greenhouse conditions in the summer of 2009 (Source 1, S1) ; ii) pooled damaged leaf tissue from plants subjected to herbivory for 1, 4 and 12 h (≈20% maximum leaf-tissue loss) by larvae of the salt marsh caterpillar *Estigmene acrea* (S2); iii) leaves of noticeably wilted plants resulting from the drought-stress imposed after withholding

watering for 3 days (S3) (drought-stress was most probably caused by the confinement of the treated plants in pots, which impeded taproot elongation, a known morphological response to drought in amaranth [see above]), and iv) leaves of plants, showing increased thickness and coarser leaf texture as a result of the acute salt-stress produced by watering the plants for three straight days with 100 ml of a 400 mM NaCl solution, (S4). Leaf material was also obtained from leaves of plants infected with *Pseudomonas argentinensis*, a bacterial amaranth pathogen, as described previously [51] (S5) and from pigmented (red) stem tissue of un-stressed 38040 plants (S6). RNA source S1 to S5 were obtained exclusively from plants of the Revancha cultivar.

cDNA library construction for pyrosequencing

Two different methods were employed for the generation of the cDNA libraries. In method 1, cDNA synthesis (S1) was performed by using SMART II™ cDNA Synthesis kit (Clontech Laboratories, Inc., Mountain View, CA, USA) following manufacturer's recommendations. The SMART II oligonucleotide (Clontech), which has extra G nucleotides at its 3' end, was used to create an extended template useful for full-length cDNA enrichment. Double stranded cDNA was quantified with a spectrophotometer (Nano Drop 1000, Thermo Scientific, Wilmington, DE, USA) and then concentrated by speed vacuum to a concentration of 500 ng/ul. The products were run on a 2% agarose gel to verify cDNA quality and fragment length. The main size distribution was within the 500 to 4,000 bp range. Approximately 5 µg of each cDNA sample were sheared via nebulization into small fragments, and then sequenced (runs 1 and 2; see below).

In method 2, cDNA synthesis (S2 - S6; destined for the differential gene expression analysis) was performed following a previously described RNA amplification protocol [52]. This procedure is based on a reverse transcription with an oligo(dT) primer bearing a T7 promoter using ArrayScript™ reverse transcriptase (RT), engineered to produce higher yields of first-strand cDNA than wild-type enzymes. ArrayScript RT catalyzes the synthesis of almost exclusively full-length cDNAs. The cDNAs then undergo a second-strand synthesis and cleanup to get a template suitable for *in vitro* transcription with the T7 RNA polymerase. This methodology generates hundreds to thousands of antisense RNA copies of each mRNA in a sample (also called cRNA) from which a second round of cDNA synthesis is performed. This RNA amplification methodology was originally developed as a method to increase very small amounts RNA samples to produce enough material for microarray hybridization [53]. Moreover, several previous reports have confirmed that no bias is generated by the amplification of RNA [54-56].

Steps from aRNA isolation through to pyrosequencing were performed as a service by the National Laboratory of Genomics for Biodiversity (Langebio) at Cinvestav, Irapuato México. Preliminary titration runs were followed by six micro-bead sequencing runs, using Roche-454 GS FLX (454 Life Sciences/Roche; Branford, CT, USA) (runs 1 and 2) and Roche- 454 GS-FLXTM (runs 3 to 6) instruments, respectively. The first two runs involved cDNAs derived from S1. Runs 3 and 4 were done with S2 and S3. The two final runs (5 and 6) involved equimolar cDNA amounts derived from S2, S3, S4 and S5 and S2, S3, S4 and S6, respectively. In runs 5 and 6, the respective cDNAs were placed in defined sections of the pico-titer plate, which was equally divided into four sectors, to permit identification for subsequent analysis (i.e. the digital expression analysis; see below).

Bioinformatics

The 454-reads were assembled using software version 2.3 Newbler, which has a cDNA option for transcriptome assembly. This option allows the formation of isogroups (a collection of isotigs and/or contigs). In broad terms, isotigs are transcripts, built out of the contigs. Different isotigs within the same isogroup represent alternative splice variants. Thus, an isogroup can be considered the equivalent of a gene.

The resulting sequence set (contigs/isotigs) was annotated using Basic Local Alignment Search Tool (BLASTX) [57] against the non-redundant (nr) database from the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov), the Arabidopsis database from The Arabidopsis Information Resource (TAIR) (http://arabidopsis.org/index.jsp), the UniRef50 and UniRef100 databases (UniProt Reference Clusters; European Bioinformatics Institute) and all the Amaranthaceae sequences (ESTs) downloaded from Gen-Bank. Those sequences that did not produce a significant hit ($E \geq 1 \times 10^{-10}$) with the nr database (3901 sequences; ≈15% of the total) were compared to the PFAM database for annotation. The latter comprises a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains, [58]. Significant BLAST results against TAIR database were used for functional gene ontology (GO) annotation [59].

Transcriptome comparison: *A. tuberculatus* vs. *A. hypochondriacus*

The raw sequence files (SRR039408, SRR039411 and SRR039412) derived from the recently reported *A. tuberculatus* transcriptome pyrosequencing effort [32] were downloaded directly from the NCBI Sequence Read Archive (SRA) at http://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?study=SRP002251. Reads were assembled after quality control, following an identical procedure as that used for *A. hypochondriacus*. Transcript annotation for *A. tuberculatus* was performed by querying the UniRef 100, and Amaranthaceae ESTs databases. Both transcriptomes were then aligned with each other using BLASTN to identify homologous contigs. Sequence homology was defined only at E values ≤ 1×10^{-10} and identity ≥ 90%. Homologous transcripts were quantified and classified into five different categories, i.e. those: i) producing the same hit; ii) different hits; iii) and iv) one hit for one species and no hit for the other, and vice-versa, or v) no hit, when queried against the above databases. Annotated transcripts detected only in *A. hypochondriacus* or *A. tuberculatus* were also quantified.

Digital expression analysis

The number of reads per gene was counted in each of the 454-sequencing outputs derived from the salt stress, water stress, insect herbivory and bacterial infection treatments and also from stem tissue (runs 5 and 6). Genes having read counts lower than 5 were eliminated. To calculate relative expression profiles in each stress treatment, Relative Abundance (RA) values were computed for each gene per treatment sample by dividing its 454-sequence count by the total 454-sequence count in the treatment sample. Differentially expressed genes in one or more treatments were detected by using the R [60] and χ^2 test statistics using a freely available web tool (http://telethon.bio.unipd.it/bioinfo/IDEG6_form/) [61]. A gene was considered to be differentially expressed when at least one statistical test yielded significance values ≤ 0.0001. A similar procedure was employed to identify transcripts that were stem-specific or highly abundant in this tissue.

The following considerations were adopted for the organization of the digital stress-related gene expression data: i) a *minimum* (MIN) or baseline/control expression value for a given gene was assigned to the lowest RA in the four-treatment set examined. The RAs that produced an expression ratio ≤ 2 when divided by MIN were also considered as MINs; ii) a gene was considered to be *significantly expressed* (SE) by a given treatment when its RA yielded a ratio ≥ 2 when divided by MIN, and iii) *maximum expression* (ME) levels for a given gene were assigned to the treatment having the highest SE. Treatments were reported to produce additional MEs when their respective SEs yielded a ratio ≤ 2 when divided by ME. This classification was devised to give an indication of the influence that a given treatment or group of treatments had on the expression levels of a particular gene. Six basic patterns of expression could be generated on the basis of the above definitions: 1) induced expression in only one treatment (only MEs); 2) induced expression in two treatments (ME-ME or ME-SE combinations) and 3) induced expression in three treatments (ME-ME-ME,

ME-ME-SE or ME-SE-SE combinations]). A total of 50 different patterns of expression were produced when all four stress treatments analyzed in this study were accommodated into the above basic patterns.
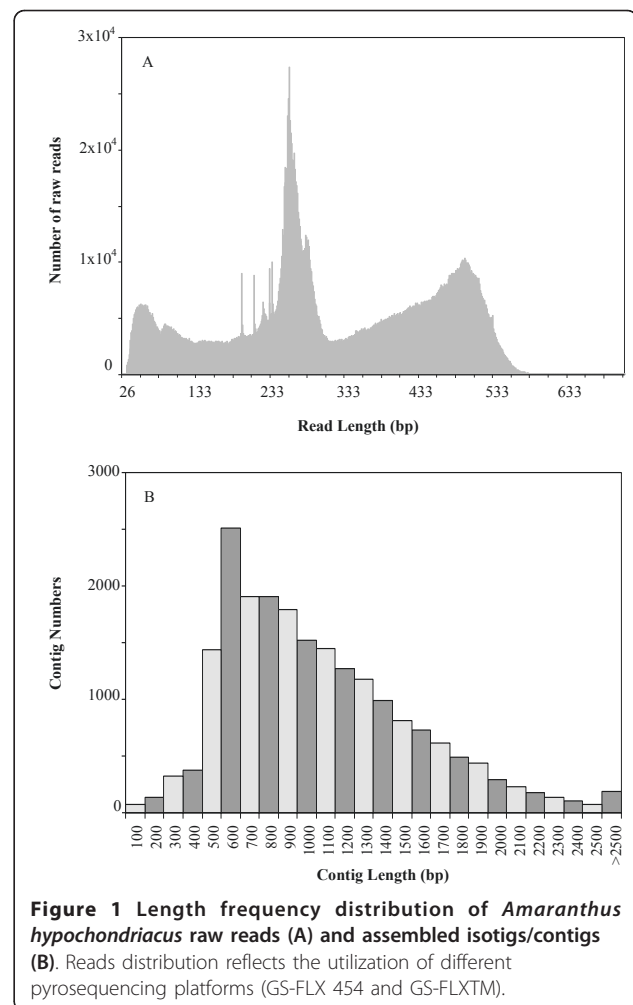
Results and Discussion

Roche GS-FLX and GS-FLXTM sequencing and assembly

Six sequencing runs yielded ≈910 Mb total data size equivalent to 2,913,966 raw reads. The raw sequence files are available from the NCBI Sequence Read Archive (SRA) at http://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?study=SRP006173, as files SRR172675 (S1), SRR172676 and SRR183482 (S2), SRR172677 (S3), SRR172678 and SRR183483 (S4), SRR172679 (S5) and SRR172680 (S6). Length frequency distribution of raw reads clustered around the 200-to-300 bp and 300-to-400 bp range as the result of using two different platforms for sequencing (Figure 1A). A total of 2,700,168 reads (93% of total) entered into the assembly process which yielded 21,207 high quality assembled sequences (20,408 isotigs + 799 contigs, equivalent to 87% of reads entering assembly and ≈82% of all assembled sequences). These ranged in length from 80 to 3,379 bp (Figure 1B) and had an average sequence length of 1,014 bp (isotigs) and 930 bp (contigs). A total of 178,636 reads (≈6% of total) remained as singletons (coverage depth = 1); of these, only 5,113 clean sequences remained after quality control. Isotigs were further incorporated into 15,667 isogroups. A status summary of the sequencing, assembly and annotation (see below) process is presented in Table 1.

Annotation of *A. hypochondriacus* contigs/isotigs

All contigs/isotigs were queried against the nr, TAIR, UniRef100, UniRef50 and Amaranthaceae ESTs and PFAM databases for annotation. Approximately 82% of all entries produced significant hits ($E \leq 1 \times 10^{-10}$) when queried against the nr database (Table 1). The 3,901 sequences with no significant hit versus the nr database were queried against the PFAM protein domain database in order to determine their putative function. Only a small fraction of these sequences (≈2%) produced significant hits (E values $\leq 1 \times 10^{-5}$) to known protein domains. These results are available in Additional file 1. Annotation of the 5,113 clean singletons against the TAIR database yielded approximately 1,000 significant hits.

The best hit for each unigene queried against the TAIR database was utilized to assign functional GO annotation in terms of *biological process* (11,224 sequences), *molecular function* (11,499 sequences) and *cellular component* (11,227 sequences) *groups*. The results are summarized in Figure 2. As expected, the largest percentage in each GO group (12% to 15%) was conformed by contigs/isotigs with an unknown functional annotation. No obvious differences in the number

Figure 1 Length frequency distribution of *Amaranthus hypochondriacus* raw reads (A) and assembled isotigs/contigs (B). Reads distribution reflects the utilization of different pyrosequencing platforms (GS-FLX 454 and GS-FLXTM).
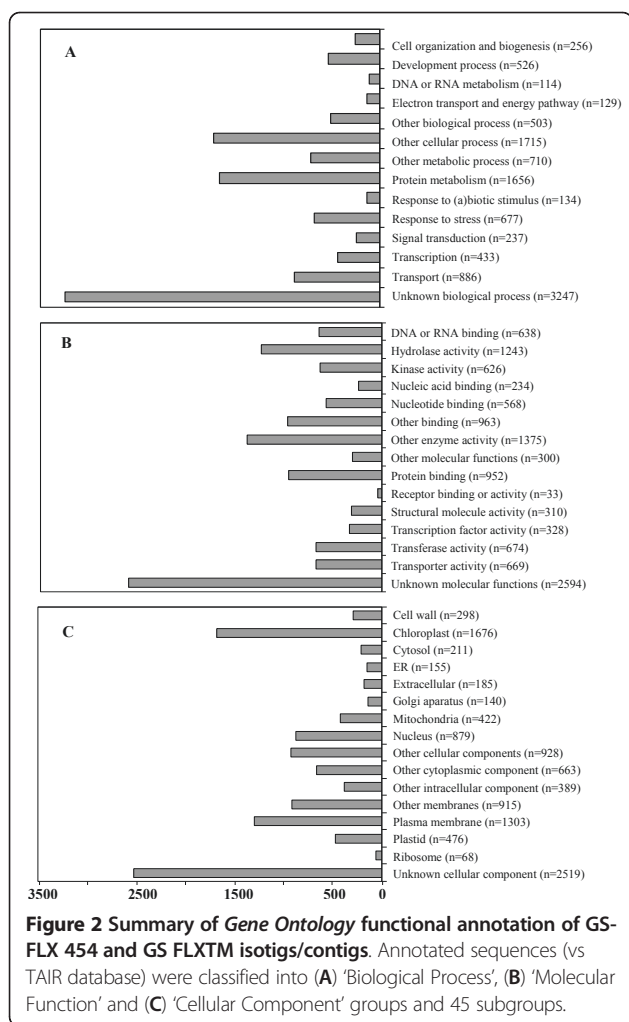
of sequences assigned to each category, including response to (a)biotic stress, were observed between grain amaranth and *Arabidopsis thaliana*. This was probably a reflection of Arabidopsis' known capacity respond strongly to abiotic and biotic stresses at the transcriptional level [62,63]. This outcome also argues against the possibility of grain amaranth possessing a different transcriptomic signature, particularly in the stress and response to stimuli categories, that could explain its characteristic (a)biotic stress tolerance, in contrast to what has been observed in plant species adapted to extreme habitats (e.g. the Arabidopsis-related halophyte *Thellungiella halophila* [64] and extremophile mangroves [65]). Thus, functional GO assignment for *Biological Process* (Figure 2A) indicated that 3% of the contigs/isotigs were grouped under stress/stimuli response, 2% in development processes and an additional 4% in other biological and metabolic processes. These categories were of our particular interest considering that one of the primal objectives of this transcriptome study was to provide information leading to the

Table 1 Summary of *A. hypochondriacus* 454 sequencing data trimming, assembly and annotation

| Category	Metric	Value
Run Metrics	Total raw reads	2,913,966 (100%)
	Total bases	909,631,600
	Reads after quality control and trimming	2,700,168 (92.6%)
	Bases entering assembly	877,153,000 (96.4%)
Assembly	Aligned reads	2,417,008 (89.5%)
	Aligned bases	803,229,499 (88.3%)
	Assembled reads	1,886,081
	Fully assembled	1,422,449
	Partially assembled	463,632
	Singletons (5.9)	178,636
	Repeats	68,980
	Outliers	56,216
	Too short	46,623
Isogroup Metrics	Total isogroups	15,667
	Average contig content	3.0
	Largest contig content	22,172
	Number with one contig	12,739
	Average isotig content	1.3
	Largest isotig content	52
	Number with one isotig	12,950
Isotig metrics	Total isotigs	20,408
	Average contig content	1.7
	Largest contig content	17
	Number with one contig	12,985
	Number of bases	20,710, 069
	Average isotig size	1, 014
	N50 isotig size	1,196
	Largest isotig size	4,762
Large contig metrics	Number of contigs	15,608
	Number of bases	15,170,717
	Average contig size	971
	N50 contig size	1,063
	Largest contig size	3,379
All contig metrics	Number of contigs	25,998
	Number of bases	18,043,010
Annotation (contigs/isotigs)	nr (NCBI)	17,282
	TAIR	16,597
Annotation (singletons)	UniRef 100	17,440
	UniRef 50	4,396
	Amaranthaceae ESTs	10,846
	TAIR	≈1,000

identification of (a)biotic stress-responsive genes (see below). From the number of transcripts to which a defense role was assigned (1% of total), more than half were associated with bacterial infection (41%) and jasmonic acid (JA)-regulation (24%), including many JA biosynthetic (e.g. *LOX13*, *AOS, AOC, OPR3*) and JA-responsive genes (Figure 3A; see also additional files 2, 3 and 4).

The overall perspective obtained from the above information is that grain amaranth possesses a diverse arsenal of genes to resist pathogen infection and insect herbivory, the majority of which are reported for the first time in this species. These include genes potentially involved in oxalate and phytoecdysteroid synthesis (results not shown), which are believed to be effective defensive weapons in amaranth and other species [66-68]. The

Figure 2 Summary of *Gene Ontology* functional annotation of GS-FLX 454 and GS FLXTM isotigs/contigs. Annotated sequences (vs TAIR database) were classified into (**A**) 'Biological Process', (**B**) 'Molecular Function' and (**C**) 'Cellular Component' groups and 45 subgroups.

Figure 3 Number of *A. hypochondriacus* isotigs/contigs categorized within the Response to Biotic Stress (A) and phytohormone function (B). BI, FI, INC, JA and SA represent bacterial and fungal infection, incompatible plant-pathogen interaction, jasmonic acid and salicylic acid, respectively. ABA, AUX, CK, ET and GA represent abscisic acid, auxins, cytokinins, ethylene and gibberellins, respectively.

implementation of a relatively robust defense response was somewhat unexpected, at least against insect herbivory, considering that the unusually high tolerance to defoliation we have observed in *A. hypochondriacus* plants (see below), might be expected to exempt this species from an investment in metabolically costly inducible defense responses (e.g. protease inhibitors and lectins). The nature of the pathogen-resistant genes isolated was also complex, and included a whole gamut of bacterial and fungal elicitor-induced and pathogenesis-related proteins, extracellular receptors similar to those involved in elicitor-induced defense responses, proteases, transcription factors (TFs) and enzymes involved in reactive oxygen species generation-detoxification.

Also important from our perspective were genes potentially involved in compensatory photosynthesis, carbohydrate re-localization (Table 2) and regulation/synthesis of phytohormone levels (Figure 3B), possibly related to the increased ramification observed in grain amaranth plants as a response to defoliation caused by insect herbivory
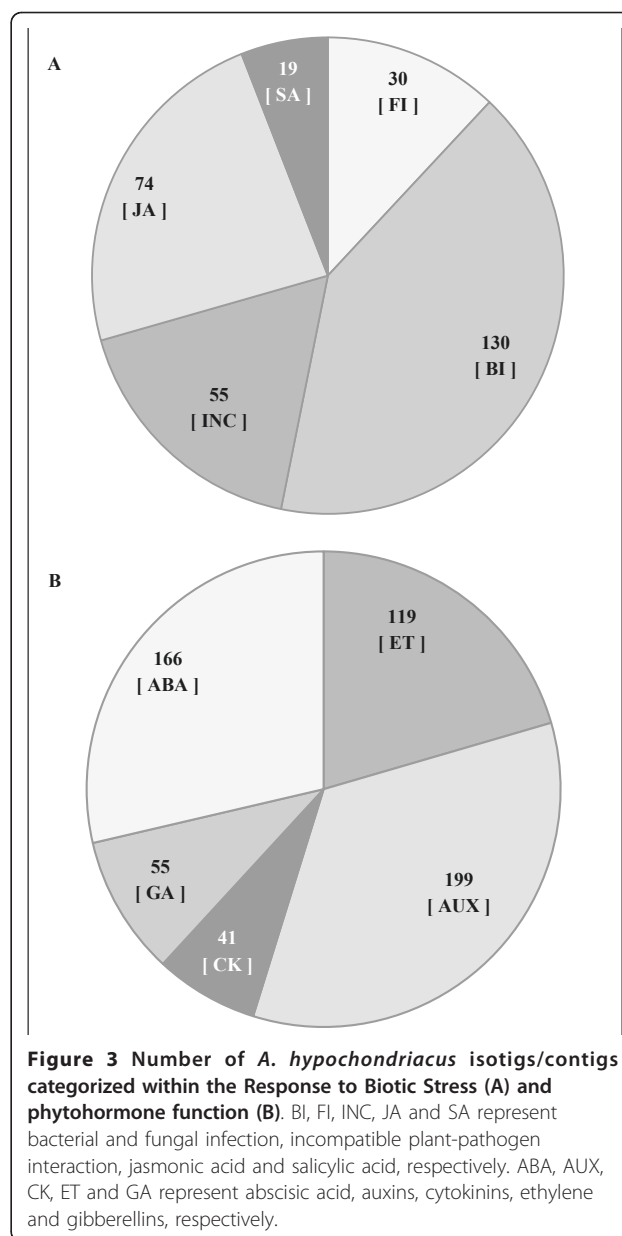
and/or mechanical damage [40,69]. Many of the genes identified can be used for studying unrelated processes. For example, the analysis of phytohormone-related genes, in combination with those showing homology with flowering genes is being pursued to gain an insight of the genetic mechanisms responsible for the several symptoms produced by phytoplasm infection of grain amaranth in the field, including phyllody [70].

Transcriptome comparison between *A. hypochondriacus* and *A. tuberculatus*

The publicly available raw transcriptomic 454 pyrosequencing data generated for *A. tuberculatus* [32] was

Table 2 Selected genes related to carbohydrate (CHO) synthesis metabolism, storage and mobilization

Gene description	No. Isotigs
Starch synthase I	3
Starch synthase II	7
Starch synthase III	1
Starch synthase V	3
Starch synthase VI	1
Granule-bound starch synthase I	2
Starch ramifying enzyme I	5
Starch ramifying enzyme II	1
Starch phosphorylase I	1
Starch phosphorylase H	4
Pullulanase	1
Iso-amylase II	2
Iso-amylase III	1
SnRK1 (SNF1-Related Protein Kinase-1)	2
SNF4 (Sucrose non-fermenting-4)	1
Glucose-6-P/phosphate transporter	5
Phosphoenol pyruvate/phosphate transporter	8
Triose P/phosphate transporter	8
AGPase small subunit	1
AGPase large subunit	3
Sucrose synthase	7
Invertase (vacuolar)	4
Invertase (neutral/alkaline)	14
Invertase (cell wall)	1
Invertase inhibitors/PMEI	7
P-glucomutase	10

The genes listed were identified in the GS-FLX 454 and GS-FLXTM pyrosequencing of *A. hypochondriacus* and could be potentially involved in CHO re-localization associated with tolerance to defoliation by insect herbivory or mechanical damage.

Table 3 Comparison of *A. hypochondriacus* (*Ah*) and *A. tuberculatus* (*At*) transcriptomes (I)

Species	UniRef100	Amaranthaceae ESTs
A. hypochondriacus	17,440	10,846
A. tuberculatus	6,625	7,185

Number of sequences with significant hits ($E \leq 1 \times 10^{-10}$) to the UniRef 100 and Amaranthaceae ESTs databases in each species.

homologous transcripts annotated with the Amaranthaceae EST data base indicated that the majority had an unknown function/provenance (21%). The highest proportion (71%) was found in EST libraries generated from immature seed and floral tissues in *Chenopodium quinoa* [72], inflorescence, germinating tissue, roots in various stages of development, hypocotyls, seed stalks and cotyledons of beet root and chlorenchyma cells of the non-Kranz C4 species *Bienertia sinuspersici* [73]. Stress related genes constituted the smallest fraction (8%), mostly represented by ESTs generated from salt-stress halophyte species (*Salicornia brachiata* [44], *Suaeda salsa* [74], *S. maritima* [46], *Atriplex centralasiatica* [75] and *C. glaucum*, in addition to ESTs from immature tissue of *Salsola tragus*. All the biotic-stress related transcripts identified came from cDNA libraries of beet roots subjected to maggot (*Tetanops myopaeformis*) feeding [76,77]. On the other hand, two thirds of the homologous transcripts annotated with the Uniref100 data base had an unknown function. Subsequent classification of transcripts (33%) having an assigned function in the biological processes category placed the majority of them (16%) within a group consisting of basic house-keeping functions (e.g. cellular component organization and biogenesis, cell cycle, cell death, regulation of gene expression, translation, cellular homeostasis, anatomical structure morphogenesis and growth, carbohydrate, protein and DNA metabolic processes, transport and photosynthesis), primary and secondary metabolism (7%), signal transduction and transcription regulation (4%). The rest included transcripts expressed in response to biotic (2%) and abiotic stress (4%). The majority of the latter were isolated from Amaranthaceae and related halophytes mostly exposed to salt stress, Interesting (a)biotic stress-related genes present in both species include a plastid-lipid associated protein known to be induced in response to multiple stresses in many plant species [78], AtPOB1, a BTB/POZ-domain protein that was found the to positively regulate disease responses in Arabidopsis and tobacco [79], the phloem sap protein AtPP2-A1 whose over-expression in Arabidopsis strongly repressed phloem feeding of the green peach aphid *Myzus persicae* [80], a transcript similar to the non-specific lipid-transfer protein type 2 from *Tamarix hispida*, whose expression was found to be part of an adaptive response to abiotic stresses in this

re-assembled using the same computational methods as for *A. hypochondriacus*. In our hands, however, the assembly yielded a ratio of contigs/singletons (12,216/53,803) that differed from the one reported by the former workers (22,035/22,434), perhaps as a consequence of the use of different assemblers [71]. The discrepancy occurred despite the fact that 83% of the total *A. tuberculatus* raw reads entering the process was assembled. BLASTN alignment of the resulting 12,216 *A. tuberculatus* contigs with the 21,207 *A. hypochondriacus* isotigs/contigs yielded 8,260 homologous sequences ($E \leq 1 \times 10^{-10}$ and $\geq 90\%$ identity). The number of contigs from each species that produced significant hits ($E \leq 1 \times 10^{-10}$) when queried against the Uniref 100 and Amaranthaceae ESTs data bases, is shown in Table 3. Combined use of above information led to the quantification of the number of homologous contigs producing the same hit, different hits, one hit for one species and none for the other, and vice-versa, and no hit. The results obtained are shown in Table 4. The analysis of the

Table 4 Comparison of *A. hypochondriacus* (*Ah*) and *A. tuberculatus* (*At*) transcriptomes (II). Annotation of homologous contigs

Annotation	UniRef100	Amaranthaceae ESTs
Homologous contigs with different hit	1,406	2,394
Homologous contigs with same hit	2,858	2,331
Homologous contigs with no hit	559	1,088
Ah contig with hit but not its *At* homologue	1,406	757
At contig with hit but not its *Ah* homologue	235	1,690

species [81], polyamine oxidase, an H_2O_2 producing enzyme supposedly involved in cell wall differentiation processes and defense responses, which was recently found to be required for wound healing in maize [82], methionine sulfoxide reductase, found to be active in defense against pathogens in pepper plants, via the regulation of cell redox status [83], and the DEAD-box ATP-dependent RNA helicase 7, a type of DNA repair protein recently shown to confer multi-stress resistance when expressed in plants [84-86]. Also remarkable was the identification several genes related to heavy metal ion homeostasis and tolerance, cation detoxification, water transport and stress-related phytohormone (e.g. abscisic acid and JA) biosynthesis and signal transduction (see additional file 5).

The number of annotated transcripts that were detected in only one species was comparatively large (Table 5). An illustrative example of the differences observed between weedy and grain amaranth transcriptomes is given by the analysis of herbicide-target genes that were annotated with the UniRef 100 and Amaranthaceae ESTs databases. It indicated that 29 of these were found in both species, whereas 13 and 8 sequences were found only in *A. hypochondriacus* and *A. tuberculatus*, respectively (Table 6).

The rather stringent parameters employed for the transcriptome comparison could have led to the transcriptome differences herein observed, although the use of lower E-value thresholds (say $E \leq 1 \times 10^{-5}$) might have not contributed much to increase level of transcript homology, as suggested by a previous genome sequencing study in *Eucalyptus grandis* [87]. However, another more plausible possible explanation is that the

Table 5 Comparison of *A. hypochondriacus* (*Ah*) and *A. tuberculatus* (*At*) transcriptomes (III)

Species	UniRef100	Amaranthaceae ESTs
A. hypochondriacus	9,974	5,364
A. tuberculatus	2,222	2,750

Number of annotated transcripts detected exclusively in one species.

above discrepancies were the reflection of fundamental differences in the overall experimental design utilized to generate both transcriptomic data. For instance, many biotic stress-related genes detected in *A. hypochondriacus* were absent in A. *tuberculatus* (results not shown). An alternative hypotheses proposing that the difference observed was due to an important sequence divergence occurred during speciation/domestication will require much further research to be validated.

Digital expression profiling
Stress-responsive transcriptional profile in leaves
This technique, also known as tag sampling or RNA-seq, is considered to be an efficient method for gene expression analysis [88,89]. The digital expression profiling analysis performed for *A. hypochondriacus* identified a total of 1,971 differentially expressed genes in response to at least one of the four stress treatments tested (i.e. water stress, salt stress, insect herbivory and bacterial infection) (Additional file 6). Fifty different gene expression profiles were generated to determine the influence of any given stress treatment on the expression levels of a particular gene. The results are shown in Figure 4. An evident feature of this analysis was the high percentage of un-annotated genes or genes with unknown function that were induced by stress. These represent a potentially rich source of genetic material that could be systematically analyzed for the discovery of genes involved in novel mechanisms of stress resistance.

All the stress-inducible genes with known function that were identified in 41 of the 50 gene expression categories were also tabulated (Additional file 7). These included several TFs known to be regulators of stress responses in other plant species, e.g. AREB-like protein [90], Dof-type zinc finger domain-containing protein [91], BTB/POZ domain-containing protein [92], GRF zinc finger containing protein [93], RAP 2.4-like protein [94], JAZ1 repressor [95], ATEBP/ERF72/RAP2.3 (related to AP2-3) [96], RAV [97], MYB-like transcription factor [98], TINY-like protein 2 [99], Cys2/His2 zinc-finger transcription factor [100], the little known GAGA-motif binding transcriptional activator [101]; SCOF-1 zinc finger proteins, found to be induced by cold or salt stress in Arabidopsis and other plants, apparently to enhance ABRE-dependent gene expression [102], a putative NAC transcription factor [103], and histone-fold/TFIID-TAF/NF-Y [104]. Others have been identified in several xerophytes/halophytes as possible factors that contribute to their ability to colonize extreme habitats, e.g. lycopene synthase [105] water channel proteins [106], myo-inositol-1-phosphate synthase, [107] cystathionine gamma-synthase [108] phosphoenolpyruvate carboxylase [109], Na^+/H^+ antiporter [110], protein phosphatase-2C [111], Ca^{2+}/H^+ antiporter [112], calcineurin B-like protein [113], inositol

Table 6 Comparison of *A. hypochondriacus* (*Ah*) and *A. tuberculatus* (*At*) transcriptomes: number of hits (isotigs/contigs) to herbicide target-site genes in the UniRef 100 and other databases

Herbicide Target-site Gene	UniRef 100 Annotation			Annotation: all databases[a]
	Hit in *Ah* and *At*	Hit in *Ah* only	Hit in *At* only	*Ah*
Tubulin	11	4	4	34
Acetolactate synthase	2	0	1	5
Protoporphyrinogen oxidase	1	1	2	2
Glutamine synthetase	6	1	0	11
1-Deoxy-D-Xylulose-5-phosphate synthase	3	3	1	3
4-Hydroxyphenylpyruvate dioxygenase	3	0	0	3
Acetyl-CoA carboxylase	1	4	0	13
Phytoene desaturase	2	0	0	2
5-Enolpyruvylshikimate-3-phosphate synthase	0	0	0	1
Dihydropteroate synthase	0	0	0	1
D1 protein (plastidic gene)	0	0	0	2

[a]nr, TAIR, UniRef 100, UniRef 50, Amaranthaceae ESTs databases.

monophosphatase [46], and salt-induced hydrophilic protein [114].

Not surprisingly, numerous transcripts coding for reactive oxygen scavengers were found to be strongly induced, many of them by multiple stresses, e.g. [Fe] superoxide dismutase, glutathione S-transferase Z1, germin-like oxidase and several catalases, peroxidases and ascorbate peroxidases. Also, the strong and multiple-stress induction of aspartyl protease, various cysteine proteases, a subtilisin-like protease and a vacuolar
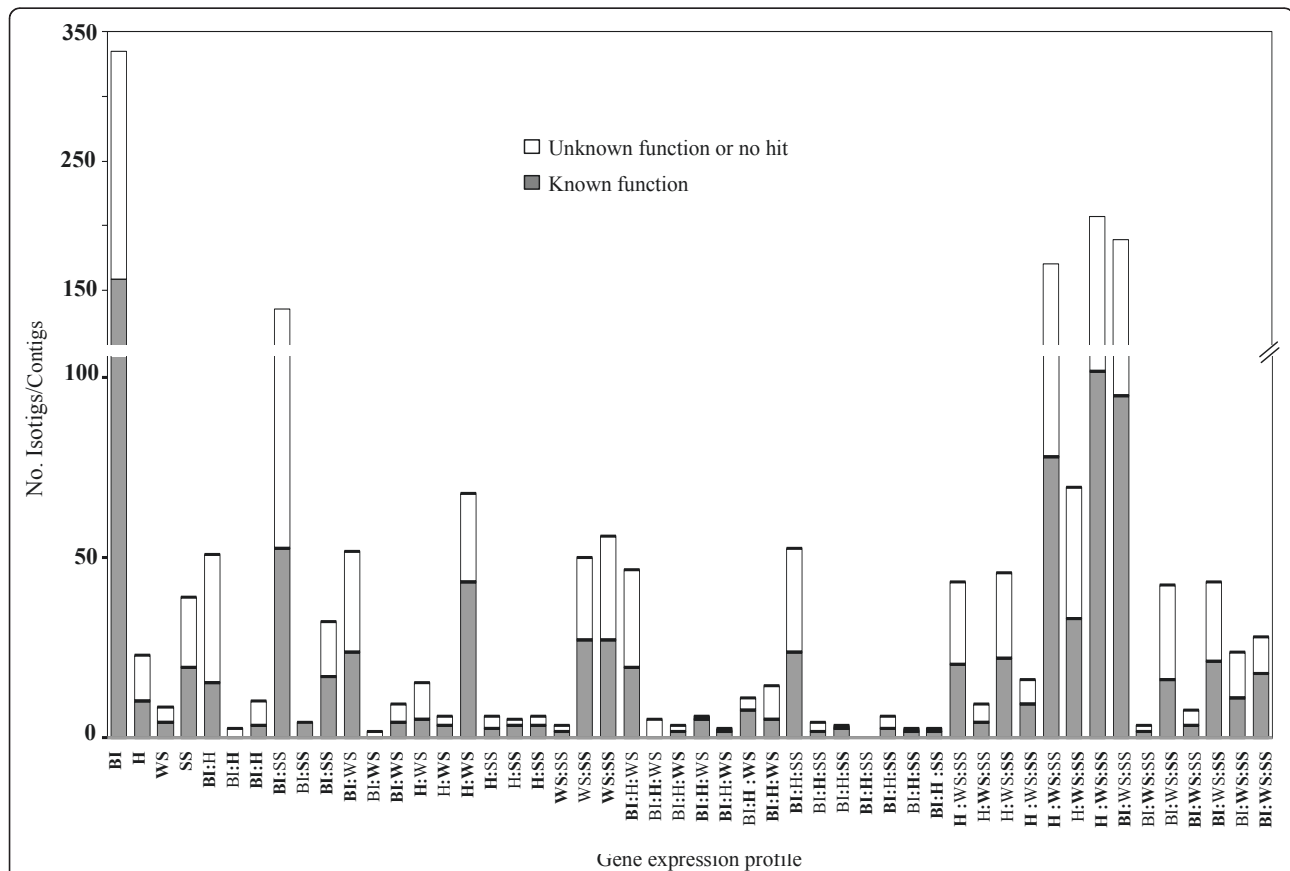

Figure 4 Number of contigs/isotigs within the 50 gene expression combinations generated to categorize digital expression data. Number of significantly expressed genes in response to: salt stress [SS], water stress [WS], insect herbivory [H] and bacterial infection [BI]. Bold letters represent *Maximum Expression* values (see text).

processing enzyme (VPE) supports a role for protein-recycling processes in response to stress, similarly to what was found during the salinity stress adaptation competence process in the extremophile *T. halophila* [115], whereas the expression of expansins, xyloglucan endotransglycosylases, several cellulose synthase subunits, glycine-, proline- and hydroxyproline-rich proteins is supported by the observed capacity to adjust cell wall properties in many plants undergoing stress [116,117]. Many of these carbohydrate-active genes were also highly expressed in stems (see below).

Of particular importance were genes highly expressed by several stress treatments, not previously reported in amaranth or related halophyles/extremophyles. These have obvious potential biotechnological applications and could also contribute to the elucidation of molecular mechanisms leading to resistance to multiple stress conditions. A selection includes the following: *Drm3*, required for *de novo* DNA methylation in *Arabidopsis thaliana* where it is proposed to regulate gene silencing processes [118]; *Enhancer of SOS 3*-1 which encodes a chloroplast-localized protein that interacts with the critical SOS3 and SOS2 regulators of salt stress tolerance in *Arabidopsis* [119,120]; YCF3 and HCF101 (high chlorophyll fluorescence 101) proteins deemed to be essential for assembly and accumulation of the photosystem I (PSI) complex and prevention of photo-oxidative damage [121,122]; translational initiation factor eIF1, found to be a determinant of sodium tolerance in yeast and plants, implying that translation is a salt toxicity target and that its recovery might be a crucial mechanism for cell survival under NaCl stress conditions [123] in addition to its proposed regulation of ion accumulation and the intracellular redox status [124]; ATP-dependent FtsH protease 9, involved in the degradation of the D1 protein of photo-damaged (PSII), a step which is needed to avoid the accumulation of excessive levels of reactive oxygen species [125]; the ACD1-LIKE electron carrier, resembling the *Arabidopsis-accelerated cell death* gene product, involved in the oxygenation of pheophorbide *a* that is required to prevent photooxidative destruction of the cell and also found to be up-regulated during salt stress adaptation process in *T. halophila* [115,126]; the prohibitin gene *PHB1*, family members of which have been found to accumulate in response to different stress conditions in many plants, presumably to act as safeguards of mitochondrial function and integrity, triggers for the retrograde mitochondrion-to-nucleus signaling and/or mediators of the interplay between H_2O_2 and NO, by a still undefined mechanism [127]; the Yellow Stripe Like 6 protein, whose members are hypothesized to participate in the delivery of metal micronutrients to and from vascular tissues and in metal tolerance and hyper-accumulation

[128]; putative linker histone H1 variant protein, expressed by drought stress conditions in tomato, and acting by a mechanism other than chromatin organization that is proposed to involve a negative regulation of stomatal conductance [129]; GASA-1/LtCOR1-like, a gibberellin regulated protein putatively involved in the regulation of fruit ripening [130] or the establishment of the dormant state in cambial meristems of trees [131]; beta and gamma-tubulin chains, whose expression is coincident with the increasingly important role played by the cytoskeleton in the mediation of the plant cell's response to stress [132,133]; translation initiation factor 5A, found to be involved in an apparently isoform-dependent regulation of stress response pathways and resistance through a largely unknown mechanism [134]; *argonaute 4-like* gene, the primary protein involved in methylation of heterochromatin and recently recognized as a critical factor for small RNA mediated systemic signaling required for plant (a)biotic stress responses and nutrient deprivation [135,136]; a putative arginase, highlighting the role of arginine as a precursor for the biosynthesis of polyamines and nitric oxide, employed as messengers for the adaptation of plants to stress [137-139], and pore-forming toxin-like lectin protein Hfr-2, recognized as an important biotic resistance factor in wheat against Hessian fly infestation and fungal (*Puccinia striiformis*) infection, and implicated in the vegetative phase change in maize [140-143], but with no known function in abiotic stress regulation. The functional characterization of a select set of multi-stress-inducible *A. hypochondriacus* genes, in Arabidopsis, tobacco and/or grain amaranth, is now under progress in our laboratory.

Transcriptional profile in stems

Comparison of the stem-derived cDNA library (S6) with those generated from leaves subjected to biotic and abiotic stress (S2 to S5) permitted to identify a small group of transcripts whose expression was exclusively detected in stems. Remarkably, the accumulation of several other transcripts was higher in stems than in foliar tissue of amaranth plants exposed to (a)biotic stress (see additional file 8). The transcript profile observed was consistent with previously data reported for stem transcriptomic analyses in *Arabidopsis thaliana* [144,145]. All annotated transcripts were classified into different categories, similarly to the above studies.

Lignin and cuticule wax biosynthesis was represented by genes coding for proteins presumably involved in mono-lignol biosynthesis (e.g. cytochrome P450 reductases, needed for the activity of several key cytochrome P450 enzymes of the phenylpropanoid pathway [144]), mono-lignol transport (e.g. ABC transporters [146]) and cuticular lipid export (e.g. white-brown-complex ABC transporter family [147]). The modest number of up-regulated lignin

biosynthesis genes that were detected was probably related to the use of young amaranth plants, not yet undergoing active lignification, for experimentation.

The carbohydrate-active enzyme category was highly represented. This was not surprising considering that these proteins play a fundamental role in cell wall biosynthesis and modification and are therefore tightly regulated during stem development. It included a number of glycosyl transferases and several glycosyl hydrolases (GH) representing families having cellulase (GH9), β-1,3-glucanase (GH3), xylanase (GH10), xyloglucan endotransglucosylase-hydrolase (GH16), glucan endo-1,3-beta-D-glucosidase (GH17), invertase (GH31) and β-D-galactosidase (GH35) activity. These enzymes are variously required for cell wall loosening and elongation, formation of the secondary cell walls of vascular tissues, hydrolysis of the xylan backbone, post-translational modifications (as glycosylations) of proteins and mobilization of energy in form of sucrose. Also detected were pectin methylesterases (PME) involved in the modification of the physical, chemical, and biological properties of pectins. The concomitant expression of a PME inhibitor probably represented a need to regulate PME in young amaranth stems in order to avoid the wall rigidification associated with PME activity. In addition, a putative β-expansin protein was detected; these proteins modulate the interaction between hemi-celluloses and cellulose presumably via a disruption of their shared hydrogen bonds [145].

Within the extracellular oxido-reductases group were found two peroxidases, belonging to the peroxidase 25 and 64 families, respectively. Peroxidases have been found to be expressed at moderate to high levels in developing stems, where they are believed to reduce cell wall extensibility due to their role in the formation of covalent links between pectin residues, hydroxyproline-rich proteins like extensins, and lignin precursors. One gene encoding a multicopper oxidase of the SKS family (*SKS5*) was identified. The function of these proteins in stem development is not well known, although the expression of *SKS5* was latterly found to be up-regulated in metal hyper-accumulating ecotypes of *Thlaspi caerulescens* [148]. Another oxido-reductase identified in amaranth stems was an 2-OG-Fe(II) oxygenase protein of unknown function that was recently found to be associated with defense mechanisms against fungal infection in Arabidopsis [149].

Several genes encoding proteins with putative interaction domains with polysaccharides and/or other proteins were identified. Many of the genes classified within this category are kinases, peptide receptors and receptor-like kinases that regulate developmental processes in plants such as the CLAVATA1-like receptor [150], CLAVATA3/ESR-related receptor [151], Abnormal Leaf

Shape 2 receptor-like kinase [152], leucine-rich repeat receptor-like kinase RLK7 [153] and LRR XI-23 kinase [154]. A number of hydroxiproline-rich (glyco) proteins, most probably representing arabinogalactan-proteins (AGPs), structural proteins (e.g. extensins, proline-rich proteins, PRPs) and a related prolyl 4-hydroxylase (catalytic alpha-2 subunit) needed for the hydroxylation of proline residues [155], were also highly expressed in stems. Numerous roles for AGPs in plant development have been suggested by means of their influence on cell fate determination, somatic embryogenesis, and cell proliferation. Also, AGPs have been assumed to be signal molecules and to associate with pectic polysaccharides, whereas extensins, PRPs and others (e.g. glycine-rich proteins) have been shown to be expressed in specific cell types including xylem and phloem tissues [145].

Also present were genes coding for a Rhomboid-like 2 endopeptidase, and two proteins with inhibitor activity: a lipid transfer protein/trypsin-alpha amylase inhibitor and a cysteine proteinase inhibitor. In addition, transcripts for an F-Box protein (SKIP2) and a 26S proteasome non-ATPase regulatory subunit, known to be involved in the targeted degradation of proteins triggered in response to various stimuli during growth and/or diverse stress conditions, were also detected. It has been suggested that proteinase activity and its modulation by proteinase inhibitors is necessary for the processing and/or turnover of cell wall proteins, generation of peptide signals, programmed cell death and/or balancing cell expansion/proliferation rates, which are collectively required for proper stem development [156,157].

Among the miscellaneous protein category were found genes coding for proteins involved in lipid metabolism (GDSL-lipases [158] and a putative glycerophosphoryl diester phosphodiesterase [159]), which are suggested to be important for stem development, a copper-binding plantacyanin (ARPN), assumed to regulate oxido-reduction processes in cell walls, several proteins known to be required for stem cell maintenance in the shoot apical meristem (histone H2A; [160]; Aurora 2 histone kinase [161]), metal tolerance (e.g. selenocysteine methyltransferase [162]) and components of the cytoskeleton, most probably involved in cell division and elongation [163]. The finding of a transcript coding for the catalytic LigB subunit of an aromatic ring-opening dioxygenase family (i.e. a putative dopa dioxygenase) the prominent enzyme in betacyanin biosynthesis, and of biosynthetically related glycosyl transferases (GTs) (e.g. GT from *Phytolacca Americana* and a UDP-GT) [164] was consistent with the highly pigmented phenotype of the stem tissue used to generate the sequenced cDNA library. The determination of the structure and regulation of pigment-related genes, their tissue- and stress-related expression patterns, and their probable role in defense

against insect herbivory in grain amaranth is now being actively pursued in our laboratory. Several TFs were also detected. In accordance with a previous report [144], most of TFs found to be highly expressed in stem tissue of grain amaranth were of the MYB, AP2-EREBP, GRAS, bHLH-domain and homeodomain families (e.g. WOX4 [165]). TFs in stems have been variously associated with the regulation of vascular tissue bio-genesis and differentiation, phenylpropanoid gene expression and fiber development [144]. Finally, a high level of expression was found for several abiotic stress and defense-related genes in stems of *A. hypochondriacus*. The presence of highly expressed defense-related genes was in accordance with a recent report showing that genes involved in plant defense and protective functions were dominant in developing stems of *Populus trichocarpa* [156]. In this respect, the concomitant presence of a putative jasmonate o-methyl transferase and a CXE carboxylesterase gene coding for a protein that can presumably identify methyl jasmonate (MeJA) as its substrate (in addition to methyl salicylate and indol-3-acetate) in *Actinidia arguta*, argues in favor of a possible role for MeJA in signaling, both within and between amaranth plants, during biotic and/or abiotic stress [166,167]. Other interesting genes identified in amaranth stems to which an active role in pathogen defense has been recently ascribed include those coding for an epoxide hydrolase 2 [168] and a VPE-1B [169], respectively. The role of epoxide hydrolase in defense is thought to be associated with its involvement in detoxification, signaling, and/or metabolism of antimicrobial compounds, whereas VPE's importance is believed to derive from its involvement in elicitor-triggered immunity connected with the combined induction of a hypersensitive response (HR) and stomatal closure. As mentioned above, VPE expression has also been associated with responses to abiotic stress.

Conclusions

The work herewith presented describes the first large-scale 454 pyrosequencing transcriptomic analysis of *A. hypochondriacus*, an under-utilized and stress-tolerant crop known to produce highly nutritious seeds and foliage. This study allowed the identification of numerous genes that are presently been analyzed to determine their role in unknown or poorly understood aspects of grain amaranth physiology, such as the mechanisms employed to tolerate defoliation, either by mechanical damage or insect defoliation. Furthermore, a digital expression analysis of transcriptome-derived data allowed the identification of numerous genes that are expressed in response to (a)biotic stress and also in a stem-specific manner. This information greatly complemented the relatively scant knowledge regarding stress-related gene expression in grain amaranth, particularly with regards to insect herbivory and bacterial infection. Furthermore, it uncovered many multiple-stress genes that could contribute to the effective response capacity against several types of environmental insults often reported in grain amaranth. Finally, a comparison with transcriptomic data obtained from an amaranth weedy species produced large differences in the number and types of transcripts detected. Although this outcome most probably resulted from fundamental experimental differences in the way the respective transcriptomic data was obtained, it is tempting to speculate that such a difference reflected a large degree of divergence between wild and cultivated amaranths generated during speciation and/or as a consequence of the domestication of *A. hypochondriacus*.

Additional material

> **Additional file 1: Functional annotation of clean GS-FLX 454 and GS FLXTM contigs/isotigs by comparison to the PFAM database**.
>
> **Additional file 2: Transcripts associated with jasmonic acid-related responses identified in *A. hypochondriacus***.
>
> **Additional file 3: Transcripts associated with responses to bacterial infection identified in *A. hypochondriacus***.
>
> **Additional file 4: Transcripts associated with the incompatible plant-pathogen interaction identified in *A. hypochondriacus***.
>
> **Additional file 5: Annotated (Uniref 100) homologous transcripts in *A. hypochondriacus* and *A. tuberculatus***.
>
> **Additional file 6: Digital expression data; total contigs/isotigs that were differentially expressed in response to (a)biotic stress**.
>
> **Additional file 7: Digital expression data; total contigs/isotigs having a known function that were differentially expressed in response to (a)biotic stress**.
>
> **Additional file 8: Digital expression data; total contigs/isotigs having a known function that were differentially expressed in stems**.

Author details

[1]Unidad de Biotecnología e Ingeniería Genética de Plantas, (Cinvestav-Unidad Irapuato) Km 9.6 del Libramiento Norte Carretera Irapuato-León. Apartado Postal 629, C.P. 36821, Irapuato, Gto., México. [2]Laboratorio Nacional de Génomica para la Biodiversidad, Km 9.6 del Libramiento Norte Carretera Irapuato-León. Apartado Postal 629, C.P. 36821, Irapuato, Gto., México. [3]Department of Entomology, College of Agricultural Sciences. Penn State University, University Park, PA 16802, USA.

Authors' contributions

JPDF and LHE drafted the manuscript. MGE performed the digital expression analysis. MGE performed the statistical analysis of the digital expression data. JPDF analyzed the transcriptional and digital expression data. HAA and NAMG performed the GO functional annotation. GCA and FPC analyzed the

raw transcriptomic and assembly data. KCC, GCA and PACA performed the plant stress treatments. KCC, GCA, PACA, JMS, FPC and EVO analyzed the transcriptomic data pertaining biotic stress, phytohormones and carbohydrate metabolism, as possibly related to tolerance to defoliation by herbivory or mechanical damage. JPDF designed and coordinated the study. All authors read and approved the final manuscript.

References

1. Mosyakin S, Robertson K: *Amaranthus L., in Flora of North America North of Mexico.* NY: New York; 20034.
2. Sauer J: **Grain amaranths.** London: Longman Group; 1976.
3. Holm L, Doll J, Holm E, Pancho J, Herberger J: **World Weeds: Natural Histories and Distribution.** Toronto: JohnWiley & Sons; 1997.
4. Steckel LE: **The dioecious Amaranthus spp.: Here to stay.** *Weed Technol* 2007, **21**:567-570.
5. Mallory MA, Hall RV, McNabb AP, Pratt DB, Jellen EN, Maughan PJ: **Development and characterization of microsatellite markers for the grain amaranths.** *Crop Sci* 2008, **48**:1098-1106.
6. Iturbide G, Gispert M: **Grain amaranths (*Amaranthus spp.*).** Rome, Italy: FAO; 1994.
7. Becker R, Wheeler EL, Lorenz K, Stafford AE, Grosjean OK, Betschart AA, Saunders RM: **A Compositional Study of Amaranth Grain.** *J Food Sci* 1981, **46**:1175-1180.
8. Breene WM: **Food Uses of Grain Amaranth.** *Cereal Food World* 1991, **36**:426-430.
9. Bressani R, Gonzales JM, Zuniga J, Breuner M, Elias LG: **Yield, Selected Chemical-Composition and Nutritive-Value of 14 Selections of Amaranth Grain Representing 4 Species.** *J Sci Food Agr* 1987, **38**:347-356.
10. Bressani R, Sanchezmarroquin A, Morales E: **Chemical-Composition of Grain Amaranth Cultivars and Effects of Processing on Their Nutritional Quality.** *Food Rev Int* 1992, **8**:23-49.
11. Gorinstein S, Pawelzik E, Delgado-Licon E, Haruenkit R, Weisz M, Trakhtenberg S: **Characterisation of pseudocereal and cereal proteins by protein and amino acid analyses.** *J Sci Food Agr* 2002, **82**:886-891.
12. Tucker JB: **Amaranth - the Once and Future Crop.** *Bioscience* 1986, **36**:9-13.
13. Huerta-Ocampo J, Barba de la Rosa A: **Amaranth: a pseudo-cereal with nutraceutical properties.** *Curr Nutr Food Sci* 2011, **7**:1-9.
14. Brenner D, Baltensperger D, Kulakow P, Lehmann J, Myers R, Slabbert M, Sleugh B: **Genetic resources and breeding of Amaranthus.** *Plant Breeding Reviews* 2000, **19**:227-285.
15. Kadereit G, Borsch T, Weising K, Freitag H: **Phylogeny of Amaranthaceae and Chenopodiaceae and the evolution of C-4 photosynthesis.** *Int J Plant Sci* 2003, **164**:959-986.
16. Omami EN, Hammes PS, Robbertse PJ: **Differences in salinity tolerance for growth and water-use efficiency in some amaranth (*Amaranthus spp.*) genotypes.** *New Zeal J Crop Hort* 2006, **34**:11-22.
17. Jiayi L, W S, L X, L X, G J: **An observation of the root system growth of grain amaranth and its drought resistance.** *Agric Res Arid Areas* 1989, **3**:34-41.
18. Kauffman CS, PW H: **Grain amaranth: A crop with low water requirements and high nutritional value.** Praeger Publ., New York: W. Lockeretz; 1983.
19. Myers R: **Amaranth: New crop opportunity.** In *Progress in new crops.* Edited by: Janick J. Alexandria, VA: ASHS Press; 1996:207-220.
20. Putnam DH: **Agronomic practices for grain amaranth.** In *Proc Natl Amaranth Symp, 4th.* Edited by: Serv ME. Minneapolis, MN: Univ. of Minnesota, St. Paul; 1990:151-162.
21. Miller TE, Wing JS, Huete AR: **The Agricultural Potential of Selected C-4 Plants in Arid Environments.** *J Arid Environ* 1984, **7**:275-286.
22. Morison JIL, Gifford RM: **Plant-growth and water-use with limited water-supply in high CO_2 concentrations. 2. Plant dry-weight, partitioning and water-use efficiency.** *Aust J Plant Physiol* 1984, **11**:375-384.
23. Weber LE: **Amaranth grain production guide.** Rodale Press, Emmaus, PA; 1990.
24. Hauptli H: **Agronomic potential and breeding strategy for grain amaranths.** Maxatawny, PA: Rodale Press, Emmaus, PA; 1977.
25. Johnson BL, Henderson TL: **Water use patterns of grain amaranth in the northern Great Plains.** *Agron J* 2002, **94**:1437-1443.
26. Huerta-Ocampo J, Leon-Galvan M, Ortega-Cruz L, Barrera-Pacheco A, De León-Rodríguez A, Mendoza-Hernández G, Barba de la Rosa A: **Water stress induces up-regulation of DOF1 and MIF1 transcription factors and down-regulation of proteins involved in secondary metabolism in amaranth roots (*Amaranthus hypochondriacus* L.).** *Plant Biology* 2010.
27. Huerta-Ocampo JA, Briones-Cerecero EP, Mendoza-Hernandez G, De Leon-Rodriguez A, de la Rosa APB: **Proteomic Analysis of Amaranth (*Amaranthus hypochondriacus* L.) Leaves under Drought Stress.** *Int J Plant Sci* 2009, **170**:990-998.
28. Maughan PJ, Sisneros N, Luo MZ, Kudrna D, Ammiraju JSS, Wing RA: **Construction of an *Amaranthus hypochondriacus* bacterial artificial chromosome library and genomic sequencing of herbicide target genes.** *Crop Sci* 2008, **48**:85-94.
29. Lee JR, Hong GY, Dixit A, Chung JW, Ma KH, Lee JH, Kang HK, Cho YH, Gwag JG, Park YJ: **Characterization of microsatellite loci developed for *Amaranthus hypochondriacus* and their cross-amplifications in wild species.** *Conserv Genet* 2008, **9**:243-246.
30. Mallory MA, Hall RV, McNabb AP, Pratt DB, Jellen EN, Maughan PJ: **Development and characterization of microsatellite markers for the grain amaranths.** *Crop Sci* 2008, **48**:1098-1106.
31. Lee RM, Thimmapuram J, Thinglum KA, Gong G, Hernandez AG, Wright CL, Kim RW, Mikel MA, Tranel PJ: **Sampling the Waterhemp (*Amaranthus tuberculatus*) Genome Using Pyrosequencing Technology.** *Weed Sci* 2009, **57**:463-469.
32. Riggins CW, Peng YH, Stewart CN, Tranel PJ: **Characterization of de novo transcriptome for waterhemp (*Amaranthus tuberculatus*) using GS-FLX 454 pyrosequencing and its application for studies of herbicide target-site genes.** *Pest Manag Sci* 2010, **66**:1042-1052.
33. Maughan P, Yourstone S, Jellen E, Udall J: **SNP discovery via genomic reduction, bar coding, and 454-pyrosequencing in amaranth.** *Plant Genome* 2009, **2**:260-270.
34. Broekaert WF, Marien W, Terras FRG, Debolle MFC, Proost P, Vandamme J, Dillen L, Claeys M, Rees SB, Vanderleyden J, Cammue BPA: **Antimicrobial Peptides from *Amaranthus caudatus* Seeds with Sequence Homology to the Cysteine Glycine-Rich Domain of Chitin-Binding Proteins.** *Biochemistry-Us* 1992, **31**:4308-4314.
35. Chagollalopez A, Blancolabra A, Patthy A, Sanchez R, Pongor S: **A Novel Alpha-Amylase Inhibitor from Amaranth (*Amaranthus hypochondriacus*) Seeds.** *J Biol Chem* 1994, **269**:23675-23680.
36. Sanchez-Hernandez C, Martinez-Gallardo N, Guerrero-Rangel A, Valdes-Rodriguez S, Delano-Frier J: **Trypsin and alpha-amylase inhibitors are differentially induced in leaves of amaranth (*Amaranthus hypochondriacus*) in response to biotic and abiotic stress.** *Physiol Plantarum* 2004, **122**:254-264.
37. Valdes-Rodriguez S, Cedro-Tanda A, Aguilar-Hernandez V, Cortes-Onofre E, Blanco-Labra A, Guerrero-Rangel A: **Recombinant amaranth cystatin (AhCPI) inhibits the growth of phytopathogenic fungi.** *Plant Physiol Bioch* 2010, **48**:469-475.
38. Wu J, Luo X, Guo H, Xiao J, Tian Y: **Transgenic cotton, expressing *Amaranthus caudatus* agglutinin, confers enhanced resistance to aphids.** *Plant Breeding* 2006, **125**:390-394.
39. Roy S, Sadhana P, Begum M, Kumar S, Lodha ML, Kapoor HC: **Purification, characterization and cloning of antiviral/ribosome inactivating protein from *Amaranthus tricolor* leaves.** *Phytochemistry* 2006, **67**:1865-1873.
40. Fomsgaard I, Añon M, Barba de la Rosa A, Christophersen C, Dusek K, Délano-Frier J, Espinoza Pérez J, Fonseca A, Janovská D, Kudsk P, Labouriau R, Lacayo-Romero M, Martínez N, Matus F, Matusová K, Mathiassen S, Noellemeyer E, Pedersen H, Stavelikova H, Steffensen S, de Troiani R, Taberner A: **Adding Value to Holy Grain: Providing the Key Tools for the Exploitation of Amaranth - the Protein-rich Grain of the Aztecs.** *Results from a Joint European - Latin American Research Project* Denmark: Department of Integrated Pest Management, Aarhus University, Faculty of Agricultural Sciences; 2010.
41. Adler G, Blumwald E, Bar-Zvi D: **The sugar beet gene encoding the sodium/proton exchanger 1 (BvNHX1) is regulated by a MYB transcription factor.** *Planta* 2010, **232**:187-195.
42. de Araujo S, Silveira J, Almeida T, Rocha I, Morais D, Viegas R: **Salinity tolerance of halophyte *Atriplex nummularia* L. grown under increasing NaCl levels.** *R Bras Eng Agríc Ambiental* 2006, **10**:848-854.
43. Guo SL, Yin HB, Zhang X, Zhao FY, Li PH, Chen SH, Zhao YX, Zhang H: **Molecular cloning and characterization of a vacuolar H +-pyrophosphatase gene, SsVP, from the halophyte *Suaeda salsa* and its**

overexpression increases salt and drought tolerance of *Arabidopsis*. *Plant Mol Biol* 2006, **60**:41-50.

44. Jha B, Agarwal PK, Reddy PS, Lal S, Sopory SK, Reddy MK: **Identification of salt-induced genes from *Salicornia brachiata*, an extreme halophyte through expressed sequence tags analysis**. *Genes Genet Syst* 2009, **84**:111-120.

45. Kirch HH, Vera-Estrella R, Golldack D, Quigley F, Michalowski CB, Barkla BJ, Bohnert HJ: **Expression of water channel proteins in *Mesembryanthemum crystallinum***. *Plant Physiol* 2000, **123**:111-124.

46. Sahu BB, Shaw BP: **Isolation, identification and expression analysis of salt-induced genes in *Suaeda maritima*, a natural halophyte, using PCR-based suppression subtractive hybridization**. *BMC Plant Biol* 2009, **9**:69-94.

47. Wang LW, Showalter AM: **Cloning and salt-induced, ABA-independent expression of choline mono-oxygenase in *Atriplex prostrata***. *Physiol Plantarum* 2004, **120**:405-412.

48. Wu W, Su Q, Xia XY, Wang Y, Luan YS, An LJ: **The *Suaeda liaotungensis* kitag betaine aldehyde dehydrogenase gene improves salt tolerance of transgenic maize mediated with minimum linear length of DNA fragment**. *Euphytica* 2008, **159**:17-25.

49. Yamada A, Tsutsumi K, Tanimoto S, Ozeki Y: **Plant RelA/SpoT homolog confers salt tolerance in *Escherichia coli* and *Saccharomyces cerevisiae***. *Plant Cell Physiol* 2003, **44**:3-9.

50. Zhang Y, Yin H, Li D, Zhu WW, Li QL: **Functional analysis of BADH gene promoter from *Suaeda liaotungensis* K**. *Plant Cell Rep* 2008, **27**:585-592.

51. Casarrubias-Castillo K: **Resistencia a bacteriosis en genotipos de amaranto (*Amaranthus spp.*) de interés agronómico**. *MSc thesis* CINVESTAV: Departamento de Biotecnología y Bioquímica; 2009.

52. Vega-Arreguin JC, Ibarra-Laclette E, Jimenez-Moraila B, Martinez O, Vielle-Calzada JP, Herrera-Estrella L, Herrera-Estrella A: **Deep sampling of the Palomero maize transcriptome by a high throughput strategy of pyrosequencing**. *BMC Genomics* 2009, **10**.

53. Yue H, Eastman PS, Wang BB, Minor J, Doctolero MH, Nuttall RL, Stack R, Becker JW, Montgomery JR, Vainer M, Johnston R: **An evaluation of the performance of cDNA microarrays for detecting changes in global mRNA expression**. *Nucleic Acids Res* 2001, **29**:E41-41.

54. Feldman AL, Costouros NG, Wang E, Qian M, Marincola FM, Alexander HR, Libutti SK: **Advantages of mRNA amplification for microarray analysis**. *Biotechniques* 2002, **33**:906-914.

55. Li Y, Li T, Liu SZ, Qiu MY, Han ZY, Jiang ZL, Li RY, Ying K, Xie Y, Mao YM: **Systematic comparison of the fidelity of aRNA, mRNA and T-RNA on gene expression profiling using cDNA microarray**. *J Biotechnol* 2004, **107**:19-28.

56. Polacek DC, Passerini AG, Shi CZ, Francesco NM, Manduchi E, Grant GR, Powell S, Bischof H, Winkler H, Stoeckert CJ, Davies PF: **Fidelity and enhanced sensitivity of differential transcription profiles following linear amplification of nanogram amounts of endothelial mRNA**. *Physiol Genomics* 2003, **13**:147-156.

57. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic Local Alignment Search Tool**. *J Mol Biol* 1990, **215**:403-410.

58. Sonnhammer ELL, Eddy SR, Durbin R: **Pfam: A comprehensive database of protein domain families based on seed alignments**. *Proteins* 1997, **28**:405-420.

59. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G, Consortium GO: **Gene Ontology: tool for the unification of biology**. *Nat Genet* 2000, **25**:25-29.

60. Stekel DJ, Git Y, Falciani F: **The comparison of gene expression from multiple cDNA libraries**. *Genome Res* 2000, **10**:2055-2061.

61. Romualdi C, Bortoluzzi S, D'Alessi F, Danieli GA: **IDEG6: a web tool for detection of differentially expressed genes in multiple tag sampling experiments**. *Physiol Genomics* 2003, **12**:159-162.

62. De Vos M, Van Oosten VR, Van Poecke RMP, Van Pelt JA, Pozo MJ, Mueller MJ, Buchala AJ, Metraux JP, Van Loon LC, Dicke M, Pieterse CMJ: **Signal signature and transcriptome changes of *Arabidopsis* during pathogen and insect attack**. *Mol Plant Microbe In* 2005, **18**:923-937.

63. Kreps JA, Wu YJ, Chang HS, Zhu T, Wang X, Harper JF: **Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress**. *Plant Physiol* 2002, **130**:2129-2141.

64. Taji T, Sakurai T, Mochida K, Ishiwata A, Kurotani A, Totoki Y, Toyoda A, Sakaki Y, Seki M, Ono H, Sakata Y, Tanaka S, Shinozaki K: **Large-scale collection and annotation of full-length enriched cDNAs from a model halophyte, *Thellungiella halophila***. *BMC Plant Biol* 2008, **8**.

65. Dassanayake M, Haas JS, Bohnert HJ, Cheeseman JM: **Shedding light on an extremophile lifestyle through transcriptomics**. *New Phytol* 2009, **183**:764-775.

66. Bakrim A, Maria A, Sayah F, Lafont R, Takvorian N: **Ecdysteroids in spinach (*Spinacia oleracea* L.): Biosynthesis, transport and regulation of levels**. *Plant Physiol Bioch* 2008, **46**:844-854.

67. Korth KL, Doege SJ, Park SH, Goggin FL, Wang Q, Gomez SK, Liu G, Jia L, Nakata PA: ***Medicago truncatula* mutants demonstrate the role of plant calcium oxalate crystals as an effective defense against chewing insects**. *Plant Physiol* 2006, **141**:188-195.

68. Schmelz EA, Grebenok RJ, Ohnmeiss TE, Bowers WS: **Interactions between *Spinacia oleracea* and *Bradysia impatiens*: A role for phytoecdysteroids**. *Arch Insect Biochem* 2002, **51**:204-221.

69. Vargas-Ortiz E: **Estudio de la redistribución de carbohidratos y nitrógeno como posible mecanismo de tolerancia a pérdida de tejido foliar por herbivoría y/o daño mecánico en amaranto**. *MSc thesis* CINVESTAV: Departamento de Biotecnología y Bioquímica; 2009.

70. Ochoa-Sanchez JC, Parra-Cota FI, Avina-Padilla K, Delano-Frier J, Martinez-Soriano JP: **Amaranthus spp.: a new host of "Candidatus Phytoplasma aurantifolia"**. *Phytoparasitica* 2009, **37**:381-384.

71. Kumar S, Blaxter ML: **Comparing de novo assemblers for 454 transcriptome data**. *BMC Genomics* 2010, **11**:571.

72. Coles ND, Coleman CE, Christensen SA, Jellen EN, Stevens MR, Bonifacio A, Rojas-Beltran JA, Fairbanks DJ, Maughan PJ: **Development and use of an expressed sequenced tag library in quinoa (*Chenopodium quinoa* Willd.) for the discovery of single nucleotide polymorphisms**. *Plant Sci* 2005, **168**:439-447.

73. Park J, Okita TW, Edwards GE: **Expression profiling and proteomic analysis of isolated photosynthetic cells of the non-Kranz C-4 species *Bienertia sinuspersici***. *Funct Plant Biol* 2010, **37**:1-13.

74. Zhang L, Ma XL, Zhang Q, Ma CL, Wang PP, Sun YF, Zhao YX, Zhang H: **Expressed sequence tags from a NaCl-treated *Suaeda salsa* cDNA library**. *Gene* 2001, **267**:193-200.

75. Xu J, Yin HX, Yang LL, Xie ZX, Liu XJ: **Differential salt tolerance in seedlings derived from dimorphic seeds of *Atriplex centralasiatica*: from physiology to molecular analysis**. *Planta* 2011, **233**:859-871.

76. Puthoff D, Smigocki A: **Sugar beet (*Beta vulgaris* L.) genes regulated by sugar beet root maggot (*Tetanops myopaeformis*) infestation**. *Am Soc Sugar Beet Tech Proc* 2005, **33**:214-219.

77. Puthoff DP, Smigocki AC: **Insect feeding-induced differential expression of *Beta vulgaris* root genes and their regulation by defense-associated signals**. *Plant Cell Rep* 2007, **26**:71-84.

78. Leitner-Dagan Y, Ovadis M, Shklarman E, Elad Y, David DR, Vainstein A: **Expression and functional analyses of the plastid lipid-associated protein CHRC suggest its role in chromoplastogenesis and stress**. *Plant Physiol* 2006, **142**:233-244.

79. Mesmar : **An investigation into the role of ubiquitination in plant immunity**. University of Glasgow; 2009.

80. Zhang CL, Shi HJ, Chen L, Wang XM, Lu BB, Zhang SP, Liang YA, Liu RX, Qian J, Sun WW, You ZZ, Dong HS: **Harpin-induced expression and transgenic overexpression of the phloem protein gene AtPP2-A1 in *Arabidopsis* repress phloem feeding of the green peach aphid *Myzus persicae***. *BMC Plant Biol* 2011, **11**.

81. Wang C, Yang CP, Gao CQ, Wang YC: **Cloning and expression analysis of 14 lipid transfer protein genes from *Tamarix hispida* responding to different abiotic stresses**. *Tree Physiol* 2009, **29**:1607-1619.

82. Angelini R, Tisi A, Rea G, Chen MM, Botta M, Federico R, Cona A: **Involvement of polyamine oxidase in wound healing**. *Plant Physiol* 2008, **146**:162-177.

83. Oh SK, Baek KH, Seong ES, Joung YH, Choi GJ, Park JM, Cho HS, Kim EA, Lee S, Choi D: ***CaMsrB2*, pepper *Methionine Sulfoxide Reductase B2*, is a novel defense regulator against oxidative stress and pathogen attack**. *Plant Physiol* 2010, **154**:245-261.

84. Kim JS, Kim KA, Oh TR, Park CM, Kang H: **Functional characterization of DEAD-Box RNA Helicases in *Arabidopsis thaliana* under abiotic stress conditions**. *Plant Cell Physiol* 2008, **49**:1563-1571.

85. Li DY, Liu HZ, Zhang HJ, Wang XE, Song FM: **OsBIRH1, a DEAD-box RNA helicase with functions in modulating defence responses against pathogen infection and oxidative stress**. *J Exp Bot* 2008, **59**:2133-2146.

86. Vashisht AA, Tuteja N: **Stress responsive DEAD-box helicases: A new pathway to engineer plant stress tolerance.** *J Photoch Photobio B* 2006, **84**:150-160.

87. Novaes E, Drost DR, Farmerie WG, Pappas GJ, Grattapaglia D, Sederoff RR, Kirst M: **High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome.** *Bmc Genomics* 2008, **9**.

88. Velculescu VE, Kinzler KW: **Gene expression analysis goes digital.** *Nat Biotechnol* 2007, **25**:878-880.

89. Wang Z, Gerstein M, Snyder M: **RNA-Seq: a revolutionary tool for transcriptomics.** *Nat Rev Genet* 2009, **10**:57-63.

90. Orellana S, Yanez M, Espinoza A, Verdugo I, Gonzalez E, Ruiz-Lara S, Casaretto JA: **The transcription factor SlAREB1 confers drought, salt stress tolerance and regulates biotic and abiotic stress-related genes in tomato.** *Plant Cell Environ* 2010, **33**:2191-2208.

91. Takatsuji H: **Zinc-finger transcription factors in plants.** *Cell Mol Life Sci* 1998, **54**:582-596.

92. Boyle P, Le Su E, Rochon A, Shearer HL, Murmu J, Chu JY, Fobert PR, Despres C: **The BTB/POZ Domain of the *Arabidopsis* Disease Resistance Protein NPR1 Interacts with the Repression Domain of TGA2 to Negate Its Function.** *Plant Cell* 2009, **21**:3700-3713.

93. Choi D, Kim JH, Kende H: **Whole genome analysis of the OsGRF gene family encoding plant-specific putative transcription activators in rice (*Oryza sativa* L.).** *Plant Cell Physiol* 2004, **45**:897-904.

94. Dong CJ, Liu JY: **The *Arabidopsis* EAR-motif-containing protein RAP2.1 functions as an active transcriptional repressor to keep stress responses under tight control.** *BMC Plant Biol* 2010, **10**:47-62.

95. Staswick PE: **JAZing up jasmonate signaling.** *Trends Plant Sci* 2008, **13**:66-71.

96. Assunção AGL, Herrero E, Lin YF, Huettel B, Talukdar S, Smaczniak C, Immink RGH, van Eldik M, Fiers M, Schat H, Aarts MGM: ***Arabidopsis thaliana* transcription factors bZIP19 and bZIP23 regulate the adaptation to zinc deficiency.** *P Natl Acad Sci USA* 2010, **107**:10296-10301.

97. Chen XF, Wang Z, Wang XM, Dong J, Ren JZ, Gao HW: **Isolation and characterization of GoRAV, a novel gene encoding a RAV-type protein in *Galegae orientalis*.** *Genes Genet Syst* 2009, **84**:101-109.

98. Chen WQ, Provart NJ, Glazebrook J, Katagiri F, Chang HS, Eulgem T, Mauch F, Luan S, Zou GZ, Whitham SA, Budworth PR, Tao Y, Xie ZY, Chen X, Lam S, Kreps JA, Harper JF, Si-Ammour A, Mauch-Mani B, Heinlein M, Kobayashi K, Hohn T, Dangl JL, Wang X, Zhu T: **Expression profile matrix of *Arabidopsis* transcription factor genes suggests their putative functions in response to environmental stresses.** *Plant Cell* 2002, **14**:559-574.

99. Sun W, Xu XN, Zhu HS, Liu AH, Liu L, Li JM, Hua XJ: **Comparative Transcriptomic profiling of a salt-tolerant wild tomato species and a salt-sensitive tomato cultivar.** *Plant Cell Physiol* 2010, **51**:997-1006.

100. Huai JL, Zheng J, Wang GY: **Overexpression of a new Cys(2)/His(2) zinc finger protein ZmZF1 from maize confers salt and drought tolerance in transgenic *Arabidopsis*.** *Plant Cell Tiss Org* 2009, **99**:117-124.

101. Wanke D, Hohenstatt ML, Dynowski M, Bloss U, Hecker A, Elgass K, Hummel S, Hahn A, Caesar K, Schleifenbaum F, Harter K, Berendzen KW: **Alanine Zipper-Like Coiled-Coil Domains Are Necessary for Homotypic Dimerization of Plant GAGA-Factors in the Nucleus and Nucleolus.** *Plos One* 2011, **6**.

102. Kim JC, Lee SH, Cheong YH, Yoo CM, Lee SI, Chun HJ, Yun DJ, Hong JC, Lee SY, Lim CO, Cho MJ: **A novel cold-inducible zinc finger protein from soybean, SCOF-1, enhances cold tolerance in transgenic plants.** *Plant J* 2001, **25**:247-259.

103. Ganesan G, Sankararamasubramanian HM, Narayanan JM, Sivaprakash KR, Parida A: **Transcript level characterization of a cDNA encoding stress regulated NAC transcription factor in the mangrove plant *Avicennia marina*.** *Plant Physiol Bioch* 2008, **46**:928-934.

104. Jia JP, Fu JJ, Zheng J, Zhou X, Huai JL, Wang JH, Wang M, Zhang Y, Chen XP, Zhang JP, Zhao JF, Su Z, Lv YP, Wang GY: **Annotation and expression profile analysis of 2073 full-length cDNAs from stress-induced maize (*Zea mays* L.) seedlings.** *Plant J* 2006, **48**:710-727.

105. Han HP, Li YX, Zhou SF: **Overexpression of phytoene synthase gene from *Salicornia europaea* alters response to reactive oxygen species under salt stress in transgenic *Arabidopsis*.** *Biotechnol Lett* 2008, **30**:1501-1507.

106. Johansson I, Karlsson M, Johanson U, Larsson C, Kjellbom P: **The role of aquaporins in cellular and whole plant water balance.** *Biochim Biophys Acta, Biomembr* 2000, **1465**:324-342.

107. Baisakh N, Subudhi PK, Varadwaj P: **Primary responses to salt stress in a halophyte, smooth cordgrass (*Spartina alterniflora* Loisel.).** *Funct Integr Genomic* 2008, **8**:287-300.

108. Tabuchi T, Kawaguchi Y, Azuma T, Nanmori T, Yasuda T: **Similar regulation patterns of choline monooxygenase, phosphoethanolamine N-methyltransferase and S-adenosyl-L-methionine synthetase in leaves of the halophyte *Atriplex nummularia* L.** *Plant Cell Physiol* 2005, **46**:505-513.

109. Cushman JC, Meyer G, Michalowski CB, Schmitt JM, Bohnert HJ: **Salt Stress Leads to Differential Expression of 2 Isogenes of Phosphoenolpyruvate Carboxylase during Crassulacean Acid Metabolism Induction in the Common Ice Plant.** *Plant Cell* 1989, **1**:715-725.

110. Li W, Wang D, Jin T, Chang Q, Yin D, Xu S, Liu B, Liu L: **The vacuolar Na+/H+ antiporter gene SsNHX1 from the halophyte *Salsola soda* confers salt tolerance in transgenic alfalfa (*Medicago sativa* L.).** *Plant Molecular Biology Reporter* 2011, **29**:278-290.

111. Senadheera P, Maathuis FJ: **Differentially regulated kinases and phosphatases in roots may contribute to inter-cultivar difference in rice salinity tolerance.** *Plant Signal Behav* 2009, **4**.

112. Han N, Shan Q, Bao H, Wang B: **Cloning and characterization of a Ca2+/H+ antiporter from halophyte *Suaeda salsa* L.** *Plant Molecular Biology Reporter* 2011, **29**:449-457.

113. Chen JH, Sun Y, Sun F, Xia XL, Yin WL: **Tobacco plants ectopically expressing the *Ammopiptanthus mongolicus* AmCBL1 gene display enhanced tolerance to multiple abiotic stresses.** *Plant Growth Regulation* 2011, **63**:259-269.

114. Guan B, Jiang GQ, Wang YX, Wang ZC, Haxim Y, Bao QA, Hu YZ, Zhang FC, Wang Y: **Identification of differentially expressed transcripts involved in the salt-stress response of *Salsola ferganica* by suppression subtractive hybridization.** *Plant Cell Tiss Org* 2010, **103**:343-352.

115. Gong QQ, Li PH, Ma SS, Rupassara SI, Bohnert HJ: **Salinity stress adaptation competence in the extremophile *Thellungiella halophila* in comparison with its relative *Arabidopsis thaliana*.** *Plant J* 2005, **44**:826-839.

116. Harb A, Krishnan A, Ambavaram MMR, Pereira A: **Molecular and physiological analysis of drought stress in *Arabidopsis* reveals early responses leading to acclimation in plant growth.** *Plant Physiol* 2010, **154**:1254-1271.

117. Wu YJ, Jeong BR, Fry SC, Boyer JS: **Change in XET activities, cell wall extensibility and hypocotyl elongation of soybean seedlings at low water potential.** *Planta* 2005, **220**:593-601.

118. Cao XF, Jacobsen SE: **Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes.** *P Natl Acad Sci USA* 2002, **99**:16491-16498.

119. Verslues PE, Batelli G, Grillo S, Agius F, Mm YS, Zhu J, Agarwal M, Katiyar-Agarwal S, Zhu JK: **Interaction of SOS2 with nucleoside diphosphate kinase 2 and catalases reveals a point of connection between salt stress and H2O2 signaling in *Arabidopsis thaliana*.** *Mol Cell Biol* 2007, **27**:7771-7780.

120. Zhu JH, Fu XM, Koo YD, Zhu JK, Jenney FE, Adams MWW, Zhu YM, Shi HZ, Yun DJ, Hasegawa PM, Bressan RA: **An enhancer mutant of *Arabidopsis* salt overly sensitive 3 mediates both ion homeostasis and the oxidative stress response.** *Mol Cell Biol* 2007, **27**:5214-5224.

121. Naver H, Boudreau E, Rochaix JD: **Functional studies of Ycf3: its role in assembly of photosystem I and interactions with some of its subunits.** *Plant Cell* 2001, **13**:2731-2745.

122. Schwenkert S, Netz DJA, Frazzon J, Pierik AJ, Bill E, Gross J, Lill R, Meurer J: **Chloroplast HCF101 is a scaffold protein for [4Fe-4S] cluster assembly.** *Biochem J* 2010, **425**:207-214.

123. Rausell A, Kanhonou R, Yenush L, Serrano R, Ros R: **The translation initiation factor eIF1A is an important determinant in the tolerance to NaCl stress in yeast and plants.** *Plant J* 2003, **34**:257-267.

124. Diedhiou CJ, Popova OV, Dietz KJ, Golldack D: **The SUI-homologous translation initiation factor eIF-1 is involved in regulation of ion homeostasis in rice.** *Plant Biology* 2008, **10**:298-309.

125. Kato Y, Miura E, Ido K, Ifuku K, Sakamoto W: **The variegated mutants lacking chloroplastic FtsHs are defective in d1 degradation and accumulate reactive oxygen species.** *Plant Physiol* 2009, **151**:1790-1801.

126. Tanaka R, Hirashima M, Satoh S, Tanaka A: **The *Arabidopsis*-accelerated cell death gene *ACD1* is involved in oxygenation of pheophorbide a: Inhibition of the pheophorbide a oxygenase activity does not lead to the "Stay-Green" phenotype in *Arabidopsis*.** *Plant Cell Physiol* 2003, **44**:1266-1274.

127. Van Aken O, Whelan J, Van Breusegem F: **Prohibitins: mitochondrial partners in development and stress response.** *Trends Plant Sci* 2010, **15**:275-282.

128. Maestri E, Marmiroli M, Visioli G, Marmiroli N: **Metal tolerance and hyperaccumulation: Costs and trade-offs between traits and environment.** *Environ Exp Bot* 2010, **68**:1-13.

129. Scippa GS, Di Michele M, Onelli E, Patrignani G, Chiatante D, Bray EA: **The histone-like protein H1-S and the response of tomato leaves to water deficit.** *J Exp Bot* 2004, **55**:99-109.

130. Pilati S, Perazzolli M, Malossini A, Cestaro A, Dematte L, Fontana P, Dal Ri A, Viola R, Velasco R, Moser C: **Genome-wide transcriptional analysis of grapevine berry ripening reveals a set of genes similarly modulated during three seasons and the occurrence of an oxidative burst at veraison.** *BMC Genomics* 2007, **8**:428-450.

131. Schrader J, Moyle R, Bhalerao R, Hertzberg M, Lundeberg J, Nilsson P, Bhalerao RP: **Cambial meristem dormancy in trees involves extensive remodelling of the transcriptome.** *Plant J* 2004, **40**:173-187.

132. Baskin TI, Meekes HTHM, Liang BM, Sharp RE: **Regulation of growth anisotropy in well-watered and water-stressed maize roots. II. Role of cortical microtubules and cellulose microfibrils.** *Plant Physiol* 1999, **119**:681-692.

133. Wang C, Zhang L, Yuan M, Ge Y, Liu Y, Fan J, Ruan Y, Cui Z, Tong S, Zhang S: **The microfilament cytoskeleton plays a vital role in salt and osmotic stress tolerance in *Arabidopsis*.** *Plant Biology* 2010, **12**:70-78.

134. Xu JY, Zhang BL, Jiang CH, Ming F: **RceIF5A, encoding an eukaryotic translation initiation factor 5A in *Rosa chinensis*, can enhance thermotolerance, oxidative and osmotic stress resistance of *Arabidopsis thaliana*.** *Plant Mol Biol* 2011, **75**:167-178.

135. Ruiz-Ferrer V, Voinnet O: **Roles of Plant Small RNAs in Biotic Stress Responses.** *Annu Rev Plant Biol* 2009, **60**:485-510.

136. Sunkar R, Chinnusamy V, Zhu JH, Zhu JK: **Small RNAs as big players in plant abiotic stress responses and nutrient deprivation.** *Trends Plant Sci* 2007, **12**:301-309.

137. Gao HJ, Yang HQ, Wang JX: **Arginine metabolism in roots and leaves of apple (*Malus domestica* Borkh.): The tissue-specific formation of both nitric oxide and polyamines.** *Sci Hortic-Amsterdam* 2009, **119**:147-152.

138. Jubault M, Hamon C, Gravot A, Lariagon C, Delourme R, Bouchereau A, Manzanares-Dauleux MJ: **Differential regulation of root arginine catabolism and polyamine metabolism in clubroot-susceptible and partially resistant *Arabidopsis* genotypes.** *Plant Physiol* 2008, **146**:2008-2019.

139. Kuznetsov V, Shevyakova N: **Polyamines and stress tolerance in plants.** *Plant Stress* 2007, **1**:50-71.

140. Puthoff DP, Sardesai N, Subramanyam S, Nemacheck JA, Williams CE: **Hfr-2, a wheat cytolytic toxin-like gene, is up-regulated by virulent Hessian fly larval feeding.** *Mol Plant Pathol* 2005, **6**:411-423.

141. Strable J, Borsuk L, Nettleton D, Schnable PS, Irish EE: **Microarray analysis of vegetative phase change in maize.** *Plant J* 2008, **56**:1045-1057.

142. Subramanyam S, Smith DF, Clemens JC, Webb MA, Sardesai N, Williams CE: **Functional characterization of HFR1, a high-mannose N-glycan-specific wheat lectin induced by hessian fly larvae.** *Plant Physiol* 2008, **147**:1412-1426.

143. Yu XM, Wang XJ, Wang CF, Chen XM, Qu ZP, Yu XD, Han QM, Zhao J, Guo J, Huang LL, Kang ZS: **Wheat defense genes in fungal (*Puccinia striiformis*) infection.** *Funct Integr Genomic* 2010, **10**:227-239.

144. Ehlting J, Mattheus N, Aeschliman DS, Li E, Hamberger B, Cullis IF, Zhuang J, Kaneda M, Mansfield SD, Samuels L, Ritland K, Ellis BE, Bohlmann J, Douglas CJ: **Global transcript profiling of primary stems from *Arabidopsis thaliana* identifies candidate genes for missing links in lignin biosynthesis and transcriptional regulators of fiber differentiation.** *Plant J* 2005, **42**:618-640.

145. Minic Z, Jamet E, San-Clemente H, Pelletier S, Renou JP, Rihouey C, Okinyo DP, Proux C, Lerouge P, Jouanin L: **Transcriptomic analysis of *Arabidopsis* developing stems: a close-up on cell wall genes.** *BMC Plant Biol* 2009, **9**:6.

146. Boerjan W, Ralph J, Baucher M: **Lignin biosynthesis.** *Annu Rev Plant Biol* 2003, **54**:519-546.

147. Bird D, Beisson F, Brigham A, Shin J, Greer S, Jetter R, Kunst L, Wu X, Yephremov A, Samuels L: **Characterization of *Arabidopsis* ABCG11/WBC11, an ATP binding cassette (ABC) transporter that is required for cuticular lipid secretion.** *Plant J* 2007, **52**:485-498.

148. Plessl M, Rigola D, Hassinen VH, Tervahauta A, Karenlampi S, Schat H, Aarts MG, Ernst D: **Comparison of two ecotypes of the metal hyperaccumulator *Thlaspi caerulescens* (J. & C. PRESL) at the transcriptional level.** *Protoplasma* 2010, **239**:81-93.

149. van Damme M, Huibers RP, Elberse J, Van den Ackerveken G: ***Arabidopsis* DMR6 encodes a putative 2OG-Fe(II) oxygenase that is defense-associated but required for susceptibility to downy mildew.** *Plant J* 2008, **54**:785-793.

150. Song SK, Lee MM, Clark SE: **POL and PLL1 phosphatases are CLAVATA1 signaling intermediates required for *Arabidopsis* shoot and floral stem cells.** *Development* 2006, **133**:4691-4698.

151. Miwa H, Kinoshita A, Fukuda H, Sawa S: **Plant meristems: *CLAVATA3/ESR*-related signaling in the shoot apical meristem and the root apical meristem.** *J Plant Res* 2009, **122**:31-39.

152. Tanaka H, Watanabe M, Sasabe M, Hiroe T, Tanaka T, Tsukaya H, Ikezaki M, Machida C, Machida Y: **Novel receptor-like kinase *ALE2* controls shoot development by specifying epidermis in *Arabidopsis*.** *Development* 2007, **134**:1643-1652.

153. Pitorre D, Llauro C, Jobet E, Guilleminot J, Brizard JP, Delseny M, Lasserre E: **RLK7, a leucine-rich repeat receptor-like kinase, is required for proper germination speed and tolerance to oxidative stress in *Arabidopsis* thaliana.** *Planta* 2010, **232**:1339-1353.

154. Samuel MA, Mudgil Y, Salt JN, Delmas F, Ramachandran S, Chilelli A, Goring DR: **Interactions between the S-domain receptor kinases and AtPUB-ARM E3 ubiquitin ligases suggest a conserved signaling pathway in *Arabidopsis*.** *Plant Physiol* 2008, **147**:2084-2095.

155. Ellis M, Egelund J, Schultz CJ, Bacic A: **Arabinogalactan-proteins: key regulators at the cell surface? (vol 153, pg 403, 2010).** *Plant Physiol* 2010, **154**:1012-1012.

156. Dharmawardhana P, Brunner AM, Strauss SH: **Genome-wide transcriptome analysis of the transition from primary to secondary stem development in *Populus trichocarpa*.** *BMC Genomics* 2010, **11**.

157. Kurepa J, Wang S, Li Y, Zaitlin D, Pierce AJ, Smalle JA: **Loss of 26S proteasome function leads to increased cell size and decreased cell number in *Arabidopsis* shoot organs.** *Plant Physiol* 2009, **150**:178-189.

158. Irshad M, Canut H, Borderies G, Pont-Lezica R, Jamet E: **A new picture of cell wall protein dynamics in elongating cells of *Arabidopsis thaliana*: Confirmed actors and newcomers.** *BMC Plant Biol* 2008, **8**.

159. Zhao Q, Dixon RA: **Transcriptional networks for lignin biosynthesis: more complex than we thought?** *Trends Plant Sci* 2011, **16**:227-233.

160. Shen WH, Xu L: **Chromatin remodeling in stem cell maintenance in *Arabidopsis thaliana*.** *Mol Plant* 2009, **2**:600-609.

161. Demidov D, Van Damme D, Geelen D, Blattner FR, Houben A: **Identification and dynamics of two classes of aurora-like kinases in *Arabidopsis* and other plants.** *Plant Cell* 2005, **17**:836-848.

162. Sors TG, Martin CP, Salt DE: **Characterization of selenocysteine methyltransferases from *Astragalus* species with contrasting selenium accumulation capacity.** *Plant J* 2009, **59**:110-122.

163. Baluska F, Jasik J, Edelmann HG, Salajova T, Volkmann D: **Latrunculin B-induced plant dwarfism: Plant cell elongation is F-actin-dependent.** *Dev Biol* 2001, **231**:113-124.

164. Strack D, Vogt T, Schliemann W: **Recent advances in betalain research.** *Phytochemistry* 2003, **62**:247-269.

165. Ji JB, Strable J, Shimizu R, Koenig D, Sinha N, Scanlon MJ: **WOX4 promotes procambial development.** *Plant Physiol* 2010, **152**:1346-1356.

166. Islam MM, Hossain MA, Jannat R, Munemasa S, Nakamura Y, Mori IC, Murata Y: **Cytosolic alkalization and cytosolic calcium oscillation in *Arabidopsis* guard cells response to ABA and MeJA.** *Plant Cell Physiol* 2010, **51**:1721-1730.

167. Sasaki Y, Asamizu E, Shibata D, Nakamura Y, Kaneko T, Awai K, Amagai M, Kuwata C, Tsugane T, Masuda T, Shimada H, Takamiya K, Ohta H, Tabata S: **Monitoring of methyl jasmonate-responsive genes in *Arabidopsis* by cDNA macroarray: self-activation of jasmonic acid biosynthesis and crosstalk with other phytohormone signaling pathways.** *DNA Res* 2001, **8**:153-161.

168. Wijekoon CP, Goodwin PH, Hsiang T: **The involvement of two epoxide hydrolase genes, *NbEH1.1* and *NbEH1.2*, of *Nicotiana benthamiana* in the interaction with *Colletotrichum destructivum*, *Colletotrichum orbiculare* or *Pseudomonas syringae pv. tabaci*.** *Funct Plant Biol* 2008, **35**:1112-1122.

169. Zhang HJ, Dong SM, Wang MF, Wang W, Song WW, Dou XY, Zheng XB, Zhang ZG: **The role of vacuolar processing enzyme (VPE) from *Nicotiana***

benthamiana in the elicitor-triggered hypersensitive response and stomatal closure. *J Exp Bot* 2010, **61**:3799-3812.