BMC
Plant Biology

Developmental profiling of gene expression in soybean trifoliate leaves and cotyledons

Anne V. Brown[1] and Karen A. Hudson[2*]

Abstract

Background: Immediately following germination, the developing soybean seedling relies on the nutrient reserves stored in the cotyledons to sustain heterotrophic growth. During the seed filling period, developing seeds rely on the transport of nutrients from the trifoliate leaves. In soybean, both cotyledons and leaves develop the capacity for photosynthesis, and subsequently senesce and abscise once their function has ended. Before this occurs, the nutrients they contain are mobilized and transported to other parts of the plant. These processes are carefully orchestrated by genetic regulation throughout the development of the leaf or cotyledon.

Results: To identify genes involved in the processes of leaf or cotyledon development and senescence in soybean, we used RNA-seq to profile multiple stages of cotyledon and leaf tissues. Differentially expressed genes between stages of leaf or cotyledon development were determined, major patterns of gene expression were defined, and shared genes were identified. Over 38,000 transcripts were expressed during the course of leaf and cotyledon development. Of those transcripts, 5,000 were expressed in a tissue specific pattern. Of the genes that were differentially expressed between both later stage tissues, 90 % had the same direction of change, suggesting that the mechanisms of senescence are conserved between tissues. Analysis of the enrichment of biological functions within genes sharing common expression profiles highlights the main processes occurring within these defined temporal windows of leaf and cotyledon development. Over 1,000 genes were identified with predicted regulatory functions that may have a role in control of leaf or cotyledon senescence.

Conclusions: The process of leaf and cotyledon development can be divided into distinct stages characterized by the expression of specific gene sets. The importance of the WRKY, NAC, and GRAS family transcription factors as major regulators of plant senescence is confirmed for both soybean leaf and cotyledon tissues. These results help validate functional annotation for soybean genes and promoters.

Keywords: Leaf senescence, Cotyledon senescence, *Glycine max*, Leaf development

Background

Soybean (*Glycine max*) is the world's largest oilseed crop with many uses including oil, food, and biofuels. In 2013 over 75 million acres of soybean were grown in the US, for a total crop value of 42 billion dollars [1]. The completion of the soybean genome sequence in 2010 has accelerated progress in understanding the genetic control of many processes, as well as enabling functional genomics research towards crop improvement [2].

A leaf undergoes several important changes throughout its development. As the newly initiated leaf expands and gains photosynthetic competency, it undergoes a transition from nutrient sink to source, and later it undergoes senescence as nutrients are mobilized and exported prior to leaf abscission and a highly controlled process of cell death. During trifoliate leaf senescence in soybean, the nutrients that remain in the are transported to the developing seeds, therefore this process is crucial to yield and seed quality. If leaves senesce before seed development is complete (for example, as a result of a biotic stress in the form of foliar disease, or abiotic stress such as drought), the plant may suffer yield reductions up to 50 % [3], while senescence delay and extension of the seed filling period results in potentially increased yield [4]. Studies have also shown that a suppressed leaf senescence during water-deficient conditions can confer

* Correspondence: kkaczoro@purdue.edu
[2]USDA-ARS Crop Production and Pest Control Research Unit, 915 West State Street, West Lafayette, IN 47907, USA
Full list of author information is available at the end of the article

drought tolerance to tobacco plants [5]. Further understanding of the pattern of gene expression that is triggered during leaf senescence and how the initiation of leaf death is controlled presents an opportunity to maximize the photosynthetic period and potentially improve crop yields.

The cotyledon, a modified leaf that functions to sustain early seedling development follows a similar pattern, at least in soybean which undergoes epigeal phanerocotylar germination. Even among other legume species the cotyledon has a different fate – *Pisum* and *Phaseolus* cotyledons function exclusively as storage organs and do not attain photosynthetic competence, while in *Ricinius communis* the cotyledons lack the fleshy characteristics of soybean or common bean cotyledons and much more closely resemble leaves [6]. Establishment of the photosynthetic machinery in the soybean cotyledon is energetically costly for relatively little gain; it is thought that photosynthesis by soybean seedling cotyledons serves to balance respiratory losses soon after germination while the carbohydrate, lipid, and protein reserves are mobilized [7].

Previous expression profiling studies have identified genes associated with leaf development and senescence in the model plants *Arabidopsis* and maize and have further defined processes that occur during this period [8–10]. Genes that are up-regulated during senescence, termed senescence-associated genes (SAGs) encode proteins required for degradation, but which have roles in a variety of processes including defense responses, detoxification, and signaling. Genes that decrease in expression during senescence, referred to as senescence down-regulated genes (SDGs) primarily encode proteins involved with photosynthesis [11–13]. Along with SAGs and SDGs a number of transcription factors have been identified in various species that are implicated in the regulation of SAGs and SDGs [14–16]. The largest families of transcription factors associated with senescence include NAC, WRKY, MYB, C2H2 zinc finger, bZIP and AP2/EREBP [14, 15]. Some of these transcription factors have also been identified as highly expressed in the early seedling cotyledons [17]. Plant hormones play an important role in plant growth and development and also in the regulation of senescence. The salicylic acid, jasmonic acid and ethylene response pathways are known to be involved with plant defense responses and also positively regulate the senescence process, while cytokinins inhibit senescence [18, 19]. Studies in *Arabidopsis* have shown that mutants defective in salicylic acid signaling exhibit a delayed senescence and all three pathways are required for the expression of SAGs [20].

To characterize the temporal shifts in gene expression that underlie the biochemical and metabolic processes occurring in soybean leaves and cotyledons, and to identify the genes that govern these transitions, we performed RNAseq expression profiling on soybean leaves and cotyledons at multiple stages of the life cycle of each organ, including the stages before and during senescence. To more fully understand the mobilization of nutrient reserves in the cotyledon, we also examined the common and distinct processes between these two photosynthetic tissues.
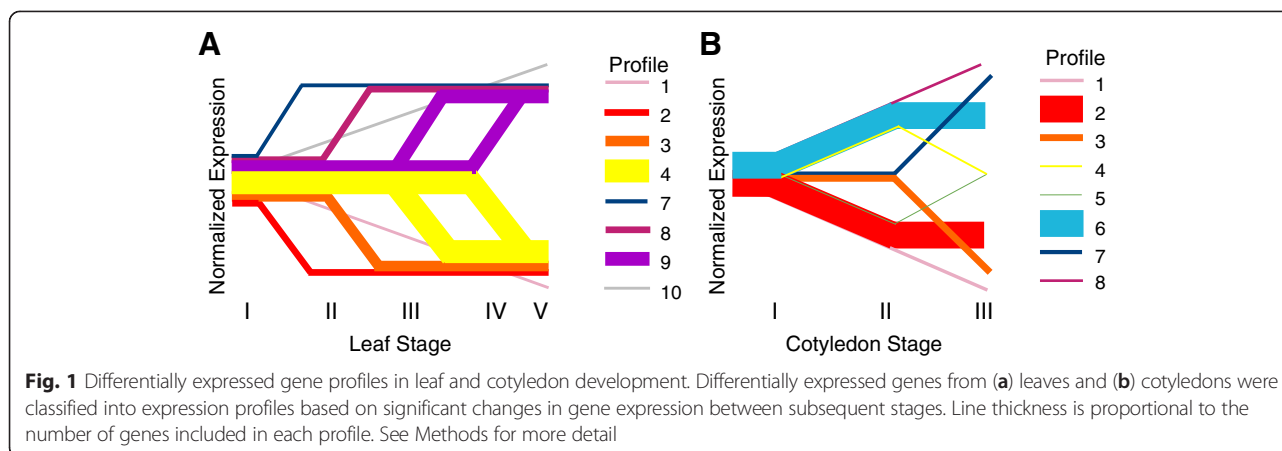
Results and discussion
Stages and genes
Five stages were selected for profiling from the leaves from leaf opening to onset of senescence, which we refer to as L-I through L-V throughout this experiment. Three stages of development were analyzed from soybean cotyledons, which we refer to as C-I through C-III. The stages of leaf and cotyledon samples collected are shown in Additional files 1 and 2, respectively. Principal component analysis plots for leaf and cotyledon samples show consistency between biological replicates and significant differences between the stages selected for comparative analysis (Additional file 3). Stages L-II and L-III as well as stages L-IV and L-V are similar to one another, but very different from stage L-I. Principal component analysis for the cotyledon data suggests that gene expression in stages C-II and C-III is overall more similar than between stages C-I (immediately post-emergence) and C-II. A total of 38,079 (70.2 %) transcripts from the Gmax189 version of the soybean genome annotation were found to be expressed during at least one stage of leaf development, while 39,039 transcripts (72.1 %) were found to be expressed during at least one stage of cotyledon development (Additional files 4 and 5).

Genes expressed differentially during development and senescence
Of the 54,175 genes in the soybean Gmax189 dataset, 12,497 genes (23.1 %) were found to be differentially expressed between at least two stages within the cotyledon timecourse while 9,500 genes (17.5 %) were differentially expressed between two stages of the leaf timecourse. Each differentially expressed gene was assigned to an expression profile based on significant fold change and expression across all of the stages in an organ order to identify major dynamic patterns of gene expression (Fig. 1). Overall, the set of differentially expressed genes through cotyledon and leaf development agree well with results of prior studies of leaf senescence in *Arabidopsis thaliana* and other plants. Close homologs for thirty-one percent (31 %) of genes identified as SAGs in *Arabidopsis* [13] were found to be differentially expressed in the leaf dataset, while 44 % were differentially expressed in the cotyledon dataset (Additional file 6). For example, *Arabidopsis* gene AtNAP (At1g69490) is a NAC transcription factor that has an important role in

Fig. 1 Differentially expressed gene profiles in leaf and cotyledon development. Differentially expressed genes from (**a**) leaves and (**b**) cotyledons were classified into expression profiles based on significant changes in gene expression between subsequent stages. Line thickness is proportional to the number of genes included in each profile. See Methods for more detail

leaf senescence [21]. The top soybean orthologs of this gene are Glyma13g35550, Glyma1g06150, Glyma20g04400, Glyma02g12220, and Glyma07g35630. All of these were found to be significantly upregulated in either cotyledon or leaf in this experiment suggesting conservation of function of this transcription factor in soybean leaf senescence.
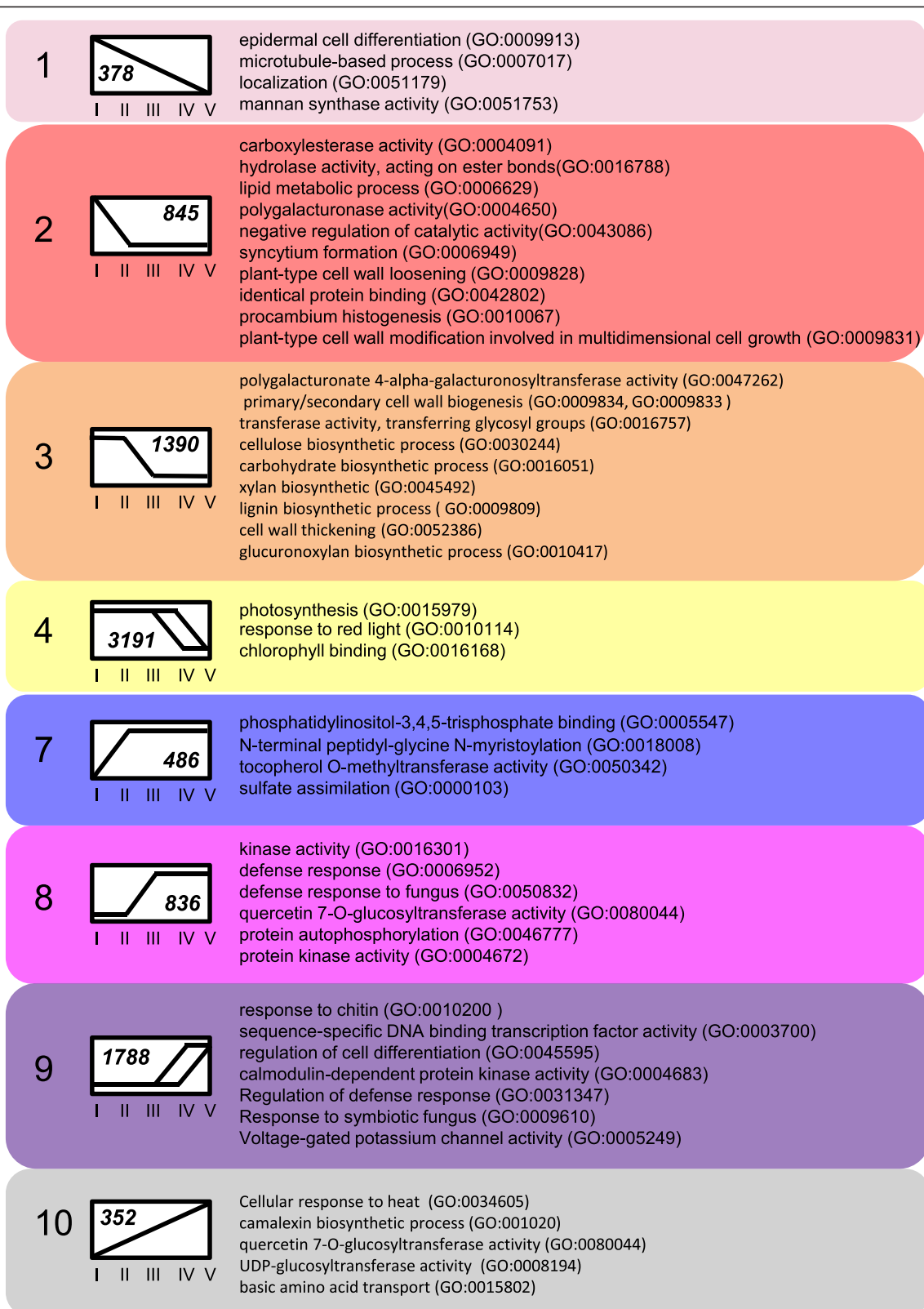
Temporal pattern of leaf gene expression

To characterize the biochemical and physiological processes occurring during different stages of leaf development, we used Gene Ontology Term (GO-term) enrichment analysis [22] to identify groups of genes involved in similar functions that were significantly enriched within the co-regulated gene sets that were identified. A table listing all enriched GO-terms and associated genes identified in each leaf or cotyledon expression profile is given in Additional file 7, and summarized in Figs. 2 & 3. Genes belonging to profile 2 or 3 that decreased significantly between stage L-I and L-II or L-III fell into several classes, predominantly associated with cell growth and expansion, cell wall biosynthesis, cell wall thickening and loosening, as well as the development of the epidermal cell layer and cuticular wax biosynthesis (Fig. 2). This finding is consistent with the L-I stage being critical for leaf expansion and maturation (Leaves are fully expanded by stage L-II). Genes associated with photosynthetic processes and light response decreased in expression during the later stages of L-IV and L-V (profile 4) consistent with the decline of photosynthesis as the tissues began to senesce. Genes expressed at higher levels in the later stages (profile 10) function in arginine and glutamate transport, consistent with remobilization of amino acids (nitrogen) from the senescing tissues. Seven genes with predicted roles in sulfate assimilation are upregulated early (profile 7) which is consistent with the previously observed developmental regulation of the ATP sulfurylase genes [23]. A significant number of genes potentially involved in disease resistance and defense responses were also upregulated after stage L-III, which is consistent with the overlapping functions

of many senescence-associated genes in responses to biotic stress (reviewed in [24]).
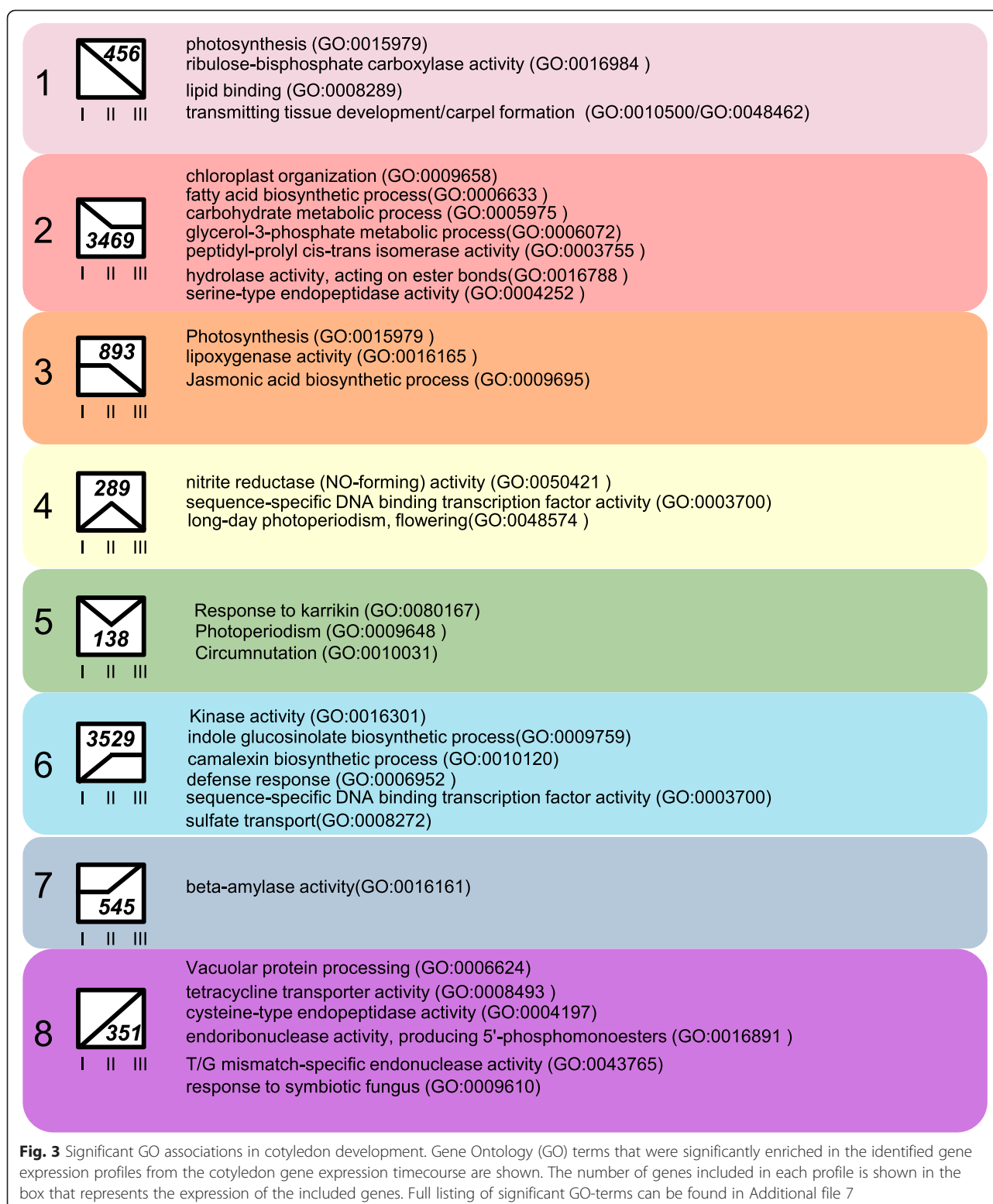
Temporal pattern of cotyledon gene expression

Genes that increased in expression between stages C-I and C-III (profile 8) or C-II and C-III (profile 7) were implicated in protein and starch breakdown (Fig. 3). Genes in profiles 4 and 6 that increase in expression between stage C-I and C-II have roles in defense response, nitrate assimilation, and sulfate transport. A large number (3469, or approximately 10 % of the total number of genes observed to be expressed in cotyledons) decreased in expression between stages C-I and C-II (profile 2) and these genes appear to function in lipid and carbohydrate metabolism and proteolysis, consistent with the role of cotyledons in mobilizing reserves to the developing seedling shortly after germination [17]. Some of the changes in gene expression observed between C-I and C-II may be due to the differences in the growth conditions for these tissues, however we can infer that the genes that respond in a similar pattern in L-I and L-III may be developmentally rather than environmentally induced. Genes that decrease in stage C-III (profiles 1 and 3) are primarily involved in the biosynthesis and maintenance of the photosynthetic complex, and jasmonic acid biosynthesis.

A significant number of genes included in cotyledon profile 6 (upregulated early) are annotated as sulfate transporters. Nitrogen (N), Potassium (K), Phosphorus (P), and Sulfur (S) are major nutrients mobilized from yellowing leaves to new growth or developing seeds. It was shown that N, P, K, and S levels drop dramatically during *Arabidopsis* leaf senescence, with a loss of over 60 % of the initial nutrient content [25]. To understand the role of transcriptional regulation in the process of nutrient mobilization, we examined the response of the all of genes annotated as transporters of nitrogen, sulfur, potassium, or phosphorus. Approximately half of the annotated or BLASTP-identified N transport genes were expressed at levels below the threshold of detection of this experiment, but many of the

1   378
I II III IV V

epidermal cell differentiation (GO:0009913)
microtubule-based process (GO:0007017)
localization (GO:0051179)
mannan synthase activity (GO:0051753)

2   845
I II III IV V

carboxylesterase activity (GO:0004091)
hydrolase activity, acting on ester bonds(GO:0016788)
lipid metabolic process (GO:0006629)
polygalacturonase activity(GO:0004650)
negative regulation of catalytic activity(GO:0043086)
syncytium formation (GO:0006949)
plant-type cell wall loosening (GO:0009828)
identical protein binding (GO:0042802)
procambium histogenesis (GO:0010067)
plant-type cell wall modification involved in multidimensional cell growth (GO:0009831)

3   1390
I II III IV V

polygalacturonate 4-alpha-galacturonosyltransferase activity (GO:0047262)
 primary/secondary cell wall biogenesis (GO:0009834, GO:0009833)
transferase activity, transferring glycosyl groups (GO:0016757)
cellulose biosynthetic process (GO:0030244)
carbohydrate biosynthetic process (GO:0016051)
xylan biosynthetic (GO:0045492)
lignin biosynthetic process (GO:0009809)
cell wall thickening (GO:0052386)
glucuronoxylan biosynthetic process (GO:0010417)

4   3191
I II III IV V

photosynthesis (GO:0015979)
response to red light (GO:0010114)
chlorophyll binding (GO:0016168)

7   486
I II III IV V

phosphatidylinositol-3,4,5-trisphosphate binding (GO:0005547)
N-terminal peptidyl-glycine N-myristoylation (GO:0018008)
tocopherol O-methyltransferase activity (GO:0050342)
sulfate assimilation (GO:0000103)

8   836
I II III IV V

kinase activity (GO:0016301)
defense response (GO:0006952)
defense response to fungus (GO:0050832)
quercetin 7-O-glucosyltransferase activity (GO:0080044)
protein autophosphorylation (GO:0046777)
protein kinase activity (GO:0004672)

9   1788
I II III IV V

response to chitin (GO:0010200)
sequence-specific DNA binding transcription factor activity (GO:0003700)
regulation of cell differentiation (GO:0045595)
calmodulin-dependent protein kinase activity (GO:0004683)
Regulation of defense response (GO:0031347)
Response to symbiotic fungus (GO:0009610)
Voltage-gated potassium channel activity (GO:0005249)

10   352
I II III IV V

Cellular response to heat (GO:0034605)
camalexin biosynthetic process (GO:001020)
quercetin 7-O-glucosyltransferase activity (GO:0080044)
UDP-glucosyltransferase activity (GO:0008194)
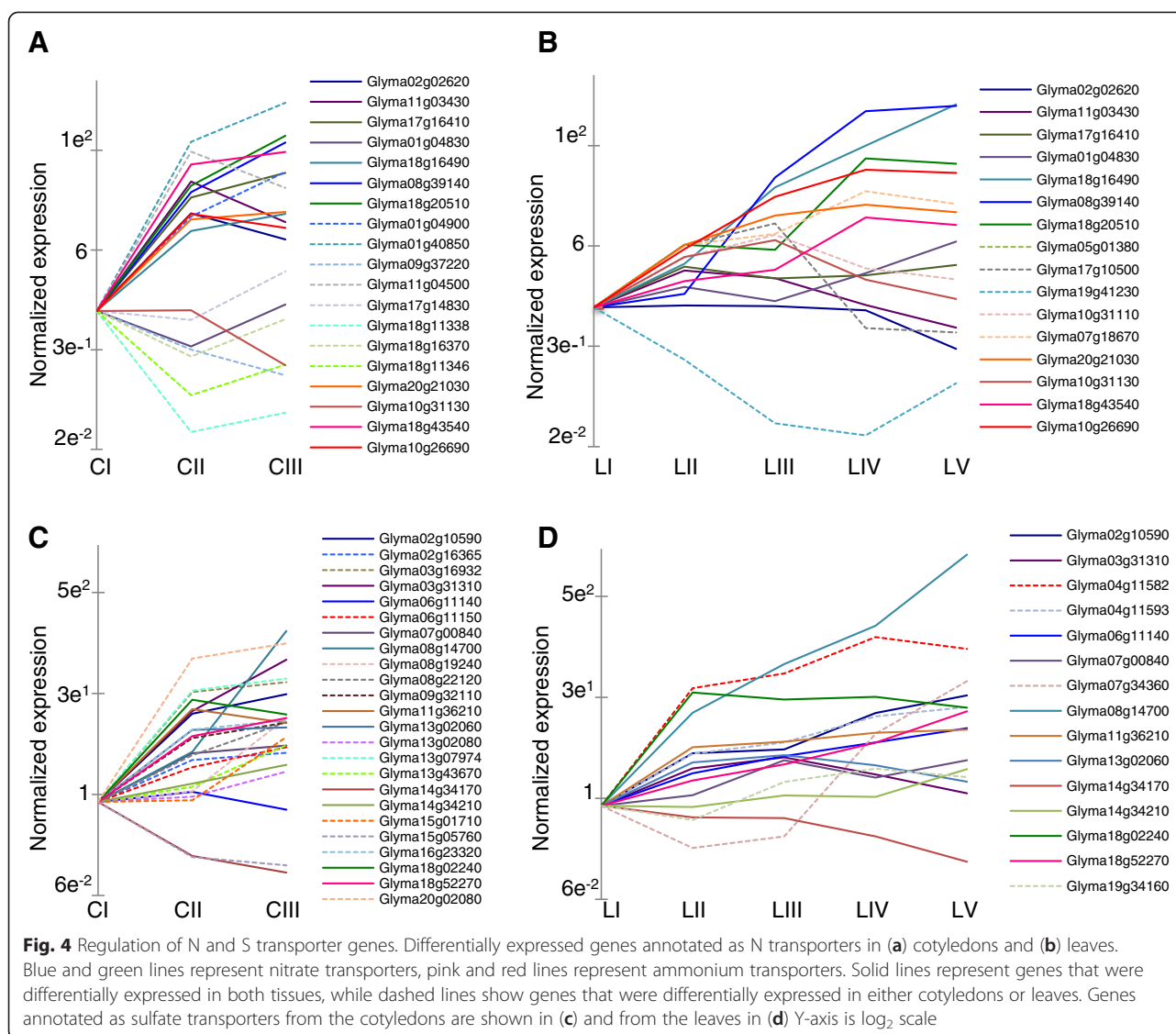basic amino acid transport (GO:0015802)

Fig. 2 Significant GO term associations in leaf developmental patterns. Gene Ontology (GO) terms that were significantly enriched in the identified gene expression profiles from developing leaves are shown. The number of genes included in each profile is shown in the box that represents the expression pattern of the included genes. A full listing of significant GO-terms can be found in Additional file 7

1 — 456
photosynthesis (GO:0015979)
ribulose-bisphosphate carboxylase activity (GO:0016984)
lipid binding (GO:0008289)
transmitting tissue development/carpel formation (GO:0010500/GO:0048462)

2 — 3469
chloroplast organization (GO:0009658)
fatty acid biosynthetic process(GO:0006633)
carbohydrate metabolic process (GO:0005975)
glycerol-3-phosphate metabolic process(GO:0006072)
peptidyl-prolyl cis-trans isomerase activity (GO:0003755)
hydrolase activity, acting on ester bonds(GO:0016788)
serine-type endopeptidase activity (GO:0004252)

3 — 893
Photosynthesis (GO:0015979)
lipoxygenase activity (GO:0016165)
Jasmonic acid biosynthetic process (GO:0009695)

4 — 289
nitrite reductase (NO-forming) activity (GO:0050421)
sequence-specific DNA binding transcription factor activity (GO:0003700)
long-day photoperiodism, flowering(GO:0048574)

5 — 138
Response to karrikin (GO:0080167)
Photoperiodism (GO:0009648)
Circumnutation (GO:0010031)

6 — 3529
Kinase activity (GO:0016301)
indole glucosinolate biosynthetic process(GO:0009759)
camalexin biosynthetic process (GO:0010120)
defense response (GO:0006952)
sequence-specific DNA binding transcription factor activity (GO:0003700)
sulfate transport(GO:0008272)

7 — 545
beta-amylase activity(GO:0016161)

8 — 351
Vacuolar protein processing (GO:0006624)
tetracycline transporter activity (GO:0008493)
cysteine-type endopeptidase activity (GO:0004197)
endoribonuclease activity, producing 5'-phosphomonoesters (GO:0016891)
T/G mismatch-specific endonuclease activity (GO:0043765)
response to symbiotic fungus (GO:0009610)

Fig. 3 Significant GO associations in cotyledon development. Gene Ontology (GO) terms that were significantly enriched in the identified gene expression profiles from the cotyledon gene expression timecourse are shown. The number of genes included in each profile is shown in the box that represents the expression of the included genes. Full listing of significant GO-terms can be found in Additional file 7

transporter genes expressed at high levels were identified as differentially regulated. These genes are listed in Additional file 8. Fig. 4 shows the differentially expressed N (ammonium and nitrate) and S transporters in the cotyledon and leaves. The upregulated nitrate transporters predominantly belong to the NPF2, NPF4, and NPF7 families, with an affinity for nitrate or dipeptide transport, and related to the Arabidopsis NRT1.7 with a role in source-sink

Fig. 4 Regulation of N and S transporter genes. Differentially expressed genes annotated as N transporters in (**a**) cotyledons and (**b**) leaves. Blue and green lines represent nitrate transporters, pink and red lines represent ammonium transporters. Solid lines represent genes that were differentially expressed in both tissues, while dashed lines show genes that were differentially expressed in either cotyledons or leaves. Genes annotated as sulfate transporters from the cotyledons are shown in (**c**) and from the leaves in (**d**) Y-axis is log$_2$ scale

remobilization of N [26, 27]. Consistent with results from *Arabidopsis*, nitrogen and sulfate transporter genes are expressed at higher levels in the later stages of both the leaves and cotyledon. Of 39 transcript models annotated as sulfate transporters, 22 were upregulated between stages C-I and C-II in cotyledons, and 14 were significantly upregulated in leaf tissues, and eleven of these were common to both sets. N transporters were also upregulated in both leaves and cotyledons. Potassium and phosphate transporters were differentially expressed during leaf and cotyledon development and senescence, but demonstrated less coordinated regulation (Additional file 8).

Modifications of the leaf expression blueprint to program a nutrient reserve

We determined that 36,012 transcripts (87.6 %) were expressed in both the cotyledon and the leaf, while

the remainder of the genes showed evidence for specificity to either leaves or cotyledons in this experiment (Additional file 9). A total of 3,027 (7.4 %) genes were only present in cotyledons and 2,067 genes (5.0 %) were only expressed in the leaf (Fig. 5). Over-represented GO-terms associated with the cotyledon-specific genes suggest the mobilization of cotyledon-specific nutrients, including the response to trehalose stimulus and alpha-glucan/water dikinase activity. The chloroplast protein alpha-glucan water dikinase 1 mediates the incorporation of phosphate into starch-like alpha-glucan, mostly at the C-6 position of glucose units in *Arabidopsis.* This acts as an overall regulator of starch mobilization, and is required for starch degradation [28]. The β-amylase enzymes are involved in the hydrolysis of starch and an important step in germinating seedlings, and post germination growth [29, 30]. Over-represented functions

Fig. 5 Leaf and cotyledon specific genes. Of the 41,106 genes identified as expressed in the experiment, 7.4 % of genes were expressed exclusively in the cotyledons, 5 % of the genes were expressed only in the leaf, and more than 87 % were shared between the two tissues. The list complete list of genes that demonstrated tissue specific expression is included in Additional file 9

for the leaf-specific genes include categories associated with photoperiodism and flower development, consistent with these processes occurring long after the cotyledons have senesced.

Genes expressed in both leaf and cotyledon tissues, and which demonstrated significant differential expression between both C-II and C-III and L-III and L-IV (genes that change in expression in anticipation of tissue senescence in both organs) were shown to have a significant overlap, and 99.0 % of these genes shared the same direction of change (up- or down-regulation) in both tissues. This may indicate that the control of the later phases of senescence uses the same molecular machinery in both tissues (Additional file 10). Genes expressed differentially between early stages, as between both C-I and C-II and L-I and L-II also overlapped significantly and over 95 % of these genes shared the same direction of change, which likewise implies a shared mechanism for early growth and developmental patterning common to both tissue types.

The control network of temporal gene expression
One of the largest classes of genes found to be differentially expressed between multiple stages were the transcription factors. To investigate regulatory networks involved with developmental processes we identified transcription factors that were differentially expressed throughout the time course. We assembled a list of 7,251 transcription factors from the soybean genome based on Mapman classification terms [31]. In leaves, a total of 990 genes predicted to encode transcription factors were differentially expressed between two stages. A total of 1415 transcription factors were differentially expressed between at least two stages of cotyledon development.

The transcription factors from both the cotyledon and leaves were placed into bins based on annotated family

[31]. We found representatives of 60 different families of transcription factors in the cotyledon and 57 families in the leaf. The most abundant transcription factor families (containing 20 or more genes) are shown in Fig. 6 for leaves and Fig. 7 for cotyledons. In both tissues, we observe that the transcription factors most commonly found to increase in expression with tissue age are from the WRKY family and these comprise the largest family of transcriptional regulators expressed differentially in our dataset. This is consistent with previous studies that find these transcription factors to function as major regulators of senescence [29, 32–34]. Additionally, members of the GRAS family are also expressed predominantly during leaf senescence and early cotyledon senescence, consistent with studies in other species [35–37]. Twenty-one of the GRAS transcription factors were upregulated during senescence in both cotyledons and leaves. In leaves, the bHLH, C2C2(DOF) zinc finger transcription factors, and G2-like transcription factors share the expression profiles (profiles 2–4) that span the photosynthetic period, consistent with a likely role in chloroplast development and regulation of photosynthetic genes and carbohydrate metabolism [38–41]. Auxin plays an important role in plant growth and development by regulating gene expression [42]. In both the leaf and cotyledon, the Aux/IAA transcription factors decrease in expression following the initial stages of development (leaf profile 3 and cotyledon profile 2). This result is consistent with studies of senescence in cotton, rice, and wheat suggesting these transcription factors may act as negative regulators of senescence [43–45]. Of 7,251 genes identified by MapMan as transcription factors, only six were annotated as part of the NAC family. As these transcription factors are known to be involved in senescence and defense [14, 21, 46], we examined the Gmax_189 genome annotations to identify 236 NAC or NAC-related transcription factors (Additional file 11), 38 of these were found to differentially expressed in the leaves and 54 in the cotyledon. The majority of the differentially expressed NAC transcription factors (23 in leaf and 40 in cotyledon) were upregulated during tissue senescence, with 8 genes common to both organs.

To identify potential *cis*-acting elements that govern the temporal patterns of expression, over-represented promoter motifs were identified from genes in each profile using Elefinder [47, 48]. The Elefinder method identifies statistically over-represented, characterized binding sites for known transcription factors from a variety of plant species within a specified promoter set. This can suggest a likely mechanism for the control of the specific expression pattern of a gene although the incidence of a given binding site only suggests which class of transcription factor could recognize and regulate the promoter. Where gene sets were too large to scan for over-
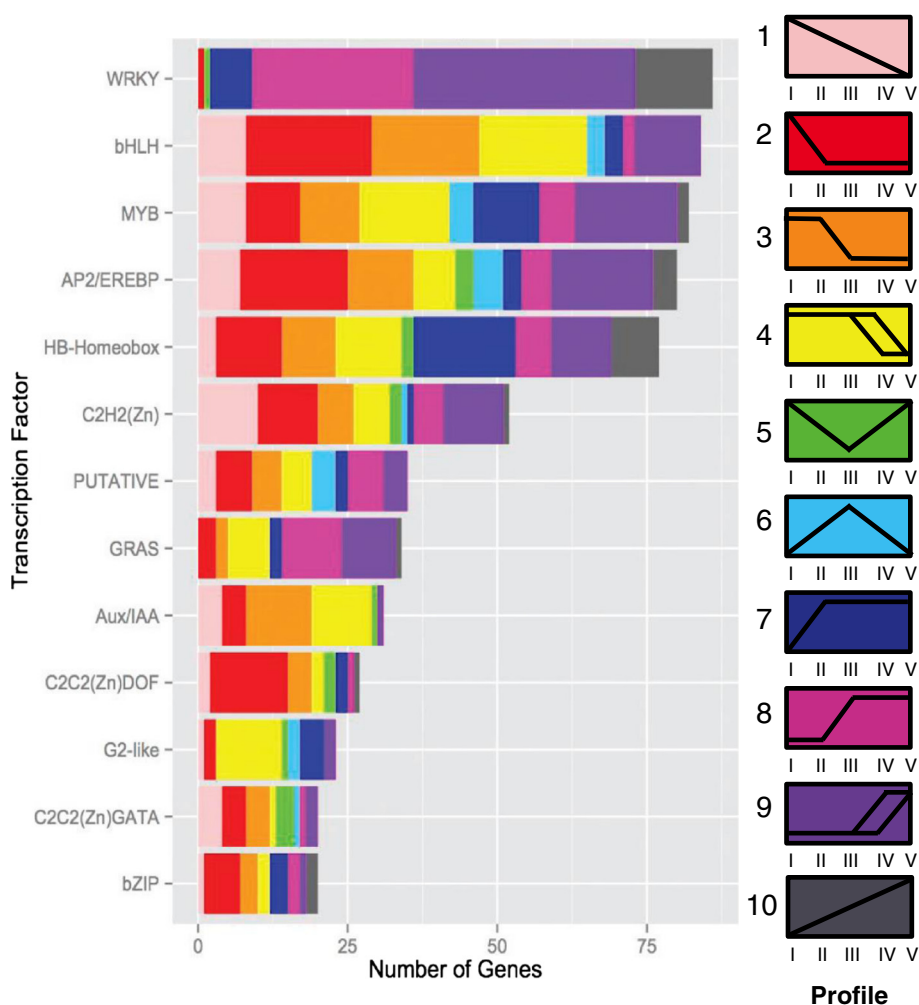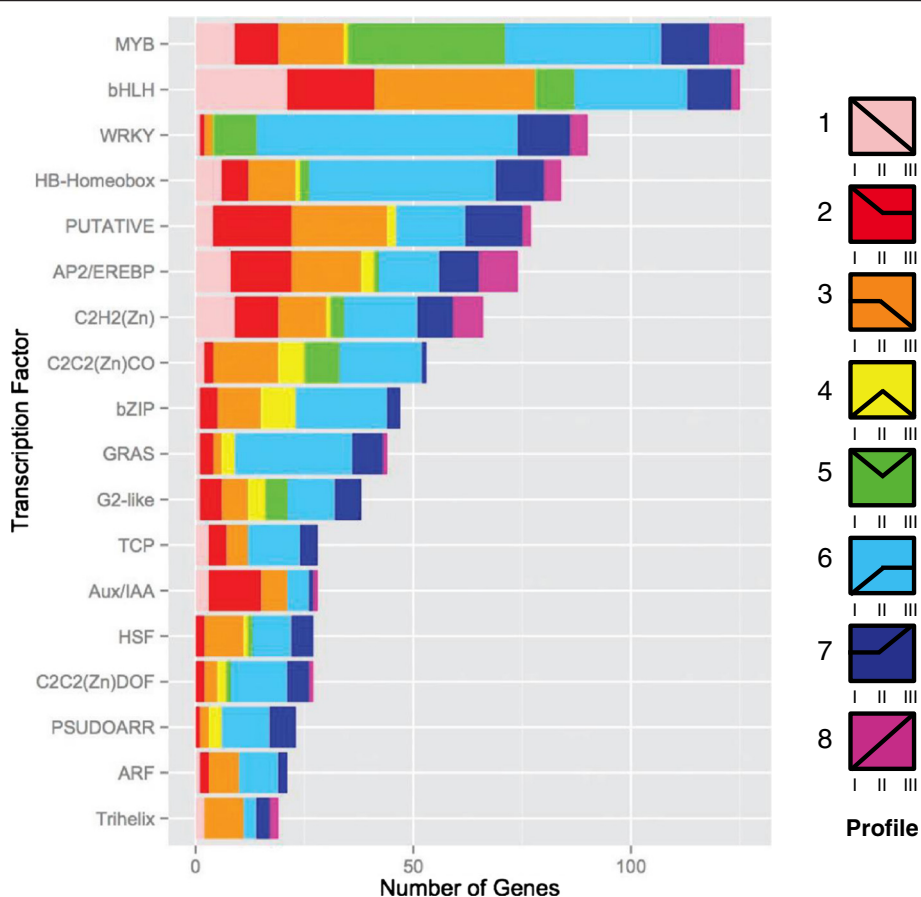
Fig. 6 Specific families of leaf transcription factors are enriched in early leaf development and senescence phases. Differentially expressed transcription factors categorized by expression profile and classified by family. Transcription factor families represented by more than twenty genes are shown [31]

represented elements in their entirety, the top 400,425 or 450 genes (based on largest fold change) were analyzed. Fig. 8 lists all of the significant motifs associated with each expression pattern in leaves, and the associated *e*-value, and the number of occurrences of the element in the data set. In the leaf, elements present in the promoters of genes that decline from stage L-I throughout the timecourse (leaf profiles 1–4) are associated with photosynthetic genes, including the MYC2 binding site, the Ibox, and ATB2/AtbZIP44/AtbZIP53/GBF5 binding site. Similarly, genes included in cotyledon profiles 1 and 4 that decreased in expression between stages C-II and C-III contained the G-box and SORLIP1 element in their promoters, consistent with the known abundance of these motifs upstream of light regulated genes associated with photosynthesis. Photosynthesis genes were observed to be downregulated between these stages (Fig. 8), thus activation via these elements may be reduced during the studied

period [47, 49, 50]. The ABFs element is enriched in promoters of genes declining in expression throughout the cotyledon timecourse, and is a G-box variant implicated in ABA response and targeted by bZIP transcription factors [51] (Fig. 9). The ARF1 binding site, targeted by the Aux/IAA transcription factors is enriched in genes downregulated in later stages of development in both the leaf (profiles 3 and 4) and cotyledon (profile 3) [52, 53]. In both leaves and cotyledons, the DPBF1&2 element is enriched in genes upregulated in later stages (for example, profiles 9 and 10 in the leaf and 5, 7, and 8 in the cotyledon) [54–56]. In both leaves and cotyledons, the AtHB6 motif is over-represented in genes that had a constant increase in expression (cotyledon profile 8 and leaf profile 10). The ATHB6 transcription factor is a member of the homeodomain-leucine zipper family of transcription factors thought to regulate active cell growth and differentiation (reviewed in [57]). The Bellringer/replumless/

Fig. 7 Specific families of transcription factors are enriched in different phases of cotyledon post-germination development. Differentially expressed transcription factors categorized by expression profile and classified by family. Transcription factor families represented by more than twenty genes are shown [31]

pennywise binding site is enriched in promoters of genes from multiple distinct expression profiles in cotyledons and leaves. This AT-rich *cis*-regulatory element is a target for homeodomain transcription factors, however in the context of regulation of the AGAMOUS gene in *Arabidopsis* the element is located within an intron [58]. In both tissues the WRKY-targeted W-box [14, 59] is the most significantly enriched promoter element, and occurs in genes that increase in expression leading up to leaf senescence.

Transcriptional control mechanisms leading to differential gene expression are believed to play important roles in coordination of the senescence process [14]. Our results support previous findings that major transcription factor families (eg, WRKY) are associated with senescence. The WRKY family is one of the largest transcription factor families in higher plants and 133 WRKY members have been identified in the soybean genome [60]. Eighty-six WRKY transcription factors were differentially expressed in the leaf, and 91 in the cotyledon. Of these WRKY genes, 44 were differentially expressed in both leaves and cotyledons, indicating that the WRKY transcription factors play a major

role in regulation in both tissues. Supporting the hypothesis that the co-regulated sets of genes expressed in distinct stages of cotyledon development share a mechanism of regulation, it was found that the W-box (the binding site for WRKY transcription factors) has by far the most significant over-representation in the promoters of genes that increase in expression during senescence in both cotyledon and leaf tissues. These results are consistent with the hypothesis that these master regulators lie at the top of a network driving the expression of genes that drive the physiological changes observable during senescence.

Conclusions

We have profiled the temporal and developmental regulation of approximately 41,000 genes in trifoliate leaves and cotyledons of soybean. We found several commonalities between these two tissues for the establishment of photosynthetic machinery, and in the final processes of mobilization of nutrients, but have also identified specific transcripts that appear to be involved in mobilization of stored nutrients in the cotyledons and are expressed

MOTIF	1	2	3	4	7	8	9	10
DPBF1&2					9.63e-07 *301*		2.03e-05 *262*	6.06e-06 *219*
ATHB6								1.51e-06 *54*
W-Box					2.85e-07 *1399*	1.85e-67 *1505*	0.00 *1380*	1.37e-88 *1251*
T-Box		3.16e-11 *588*		7.39e-12 *592*	1.12e-09 *589*	1.55e-16 *560*	9.75e-05 *496*	1.81e-07 *424*
Bellringer/replumless/ pennywise BS1 In AG		2.24e-05 *746*	1.21e-11 *771*	8.21e-06 *752*	1.68e-15 *734*	7.28e-05 *682*	1.05e-05 *672*	1.83e-16 *608*
ATB2/AtbZIP44/AtbZIP53/GBF5 BS	1.75e-08 *465*	1.97e-05 *493*	4.66e-07 *538*	1.59e-05 *495*	5.26e-05 *496*			
LTRE	1.37e-05 *42*	1.69e-05 *45*						
ARF1			1.52e-06 *338*	1.13e-08 *324*				
Ibox				4.80e-06 *395*		2.88e-06 *328*		7.04e-06 *294*
AtMYC2 BS in RD22	2.64e-05 *367*	7.98e-22 *481*	2.71e-10 *464*	6.47e-11 *440*				
Total number of Promoters	327	363 (Top 425)	389 (Top 450)	364 (Top 450)	368 (Top 450)	333 (Top 450)	324 (Top 400)	265

Fig. 8 Significantly over-represented motifs in promoters of genes that change in expression during leaf development and senescence. The table shows the significant promoter motifs (*e*-value cutoff less than *e*-05) and the number of elements (in italics) found in the set of promoters of genes classified in one of the six expression profiles [47]. For leaf analysis, multiple profiles were combined to identify significant motifs. Leaf profiles 2 and 3 were combined, while profile 4 was pooled with genes decreasing in stage L-V. Profiles 5 and 6 did not include a sufficient number of genes to identify significant motifs. Profiles 7 and 8 were pooled into one profile. Profile 9 included additional genes that increased in expression between stages L-IV and L-V. Where gene sets were too large for analysis, the top 425 or 450 genes (based on magnitude of fold change) were analyzed

earlier in cotyledon development. Significantly, we have identified families of transcriptional regulators and corresponding *cis*-regulatory motifs that appear to co-ordinately regulate the changes in expression observed in several groups of genes during the programmed transition to senescence. Our data help to identify the specific machinery by which the genome is regulated during the later stages of leaf development, and may allow more specific biotechnology approaches targeting developmental stages and tissues with transgene expression.

Methods

Plant growth and tissue collection

Williams-82 soybean seeds were planted in one part Pro-mix (BRK): two parts sand/soil min in 3-gal pots [61]. Plants were grown in Conviron PGR14 chambers in 12 hr light/12 hr dark photoperiod conditions in light provided by a combination of fluorescent and incandescent lighting at an intensity of 350 μmol m^{-2} s^{-1}, constant temperature of 25 C and 75 % relative humidity. Williams-82 plants grown in these conditions flower after 6 weeks and produce seed. Samples were collected at 2:00 pm (6 hours after lights were turned on) to minimize diurnal effects. Cotyledon samples grown for stage C-I were grown in magenta boxes on MS media (without sucrose, MP Biomedicals, LLC, Solon, Ohio) to reduce the risk of soil contamination. Three biological replicate sets of cotyledons (one pair of cotyledons) were collected for each sample. Five distinct stages determined by physical appearance and plant developmental stage were collected from the time of soil emergence to the onset of senescence (see Additional file 2). Stage C-I cotyledons were collected immediately after emergence (4 days after planting) and were green but not yet open. Stage II cotyledons were open, green, and collected after the plant produced the first set of unifoliate leaves. Stage III cotyledons were collected at the first sign of yellowing and shrinking. Three biological replicate sets of leaves (one trifoliate) were collected from different individuals, at the 4th node (3rd set of trifoliate leaves) of each plant. Initial tissue collection of the sets of trifoliate leaves began at the V3 [62] stage and subsequently occurred every seven days. Eight stages were collected Stage L-I was collected after the trifoliates opened but before the leaves were fully expanded. Stages L-II and L-III were collected 21 and 40 days after stage L-I, respectively. Stages L-IV and L-V were collected after 49 and 56 days, and these two samples were marked by the appearance of signs of leaf senescence (Additional file 1). RNA was extracted using procedures described previously [61].

Library preparation and sequencing

5 μg of RNA from each sample was used for library construction using standard protocols. Paired-end libraries were constructed for both leaf and cotyledon samples.

MOTIF	1	2	3	4	5	6	7	8
DPBF1&2		5.97e-05 281			1.56e-07 105		2.16e-22 354	4.16e-28 290
W-Box						2.48e-22 1123	9.20e-06 1239	3.57e-42 1210
BoxII							2.53e-06 383	
T-Box						8.30e-10 435		
Bellringer/replumless/pennywise BS1 In AG		8.87e-06 728	1.47e-07 622	4.00e-18 546		1.81e-11 590	4.00e-07 689	9.49e-10 657
ATHB6 Binding Site Motif								6.20e-05 54
RAV1					9.72e-16 729		8.05e-05 1856	
ARF1			4.86e-05 269					
Evening Element					1.56e-07 41			
ATB2/AtbZIP44/AtbZIP53/GBF5 BS		3.76e-06 484		3.70e-06 329		9.71e-05 363		
Ibox	5.20e-11 468							
AtMYC2 BS	5.98e-07 462	2.96e-17 453					8.18e-05 363	
SORLIP1	3.16e-05 201			9.48e-09 136				
ABFs	1.09e-05 33			1.11e-05 23				
G-Box/RAV1-B	3.14e-07 148			3.62e-06 90			6.38e-06 118	
Total # of Promoters	406	352 (Top 425)	312 (Top 450)	233	116	266 (Top 425)	327 (Top 425)	303

Fig. 9 Significantly over-represented motifs in promoters of genes that change in expression during cotyledon senescence. The table shows the most significant promoter motifs (*e*-value cutoff less than *e*-05) and the number of occurrences (in italics) for each significant element found in promoters of genes classified in one of the eight expression profiles identified for differentially expressed cotyledon genes [47]. Where gene sets were too large for analysis, the top 425 or 450 genes (based on magnitude of fold change) were analyzed

Median insert size was 200 bp. Three biological replicate RNA samples from each stage of the cotyledons were sequenced using the Illumina Hi-Scan with three libraries per lane. The total number of mate-paired reads for each sample ranged from 5,000,000 to 54,000,000. Three biological replicate RNA samples from each stage of the leaves were sequenced in one lane on the Illumina Hi-Seq 2500, and the number of reads ranged from 5,000,000 to 18,000,000. For both leaves and cotyledons, read lengths of 100 bp were collected.

Read alignment and normalization
Reads were aligned to the Glycine max transcriptome (Gmax_189) [2] using the ultrafast Bowtie aligner [63]. The alignment allowed only one mismatch per read. Approximately 87 % of reads matched known soybean transcript models. Of those reads 23 % matched only one transcript and 77 % matched multiple transcripts. (These figures were approximately the same for both leaf and cotyledon tissues.) Reads that hit more than one transcript were filtered from the bowtie output and only the reads that matched only one transcript were used for further analysis [64]. Read counts were obtained from the Bowtie output using a set of Perl scripts. Transcripts with less than 10 total read counts per gene in at least one stage were removed. The remaining transcripts were normalized for each sample by dividing the effective library size for each sample from the raw count of that sample (using the DESeq software package). These normalized counts were used for further analysis. Principal component analysis (PCA) was performed using the DESeq package, on the whole transcriptome set with low-expression genes removed. This method used variance-stabilized data to obtain sample-to-sample distance. To identify differentially expressed genes both DESeq and EdgeR software packages were used to perform pairwise comparisons of stages and for both packages with a corrected p-value of 0.05, and a 1.5 $\log_2$ fold (2.8 fold) change [65, 66]. The set of differentially expressed genes common to both of these two methods make up the set of genes used for further analysis. (Additional file 4). Differentially expressed genes from cotyledons were assigned to a profile using a script in R based on whether expression significantly increased or decreased between two stages, leaf genes were manually evaluated for inclusion into each profile. The majority of genes had either a downward or upward trend. Relatively few genes fell into profiles that initially decreased and then increased, or vice versa, the

genes that exhibited the up-then-down pattern (profile 6, 70 genes) or down-then-up pattern (profile 5, 49 genes) were combined for promoter and GO-enrichment analysis. Genes that showed increased expression between more than two stages in leaves were also combined into the "overall up" (profile 10) and genes that decreased in expression between more than two stages were placed in the "overall down" profile (profile 1). Leaf profiles 4 and 9 also included distinct patterns, as stages L-IV and L-V were similar (see Additional file 3), genes that increased in expression between L-III and L-IV or between L-IV and L-V were pooled into profile 9, and genes whose expression decreased between L-III and L-IV or L-IV and L-V were pooled into profile 4.

A larger number of differentially expressed genes were found in the cotyledons relative to leaves, potentially due to higher sequence depth for the cotyledon samples.

Data analysis
GO-term analysis was performed using the Soybase GO-enrichment tool (www.soybase.org) using GO annotation terms from genome version Gmax1.89. Promoter motif analysis was performed using the Elefinder tool at http://stan.cropsci.uiuc.edu/tools.php [48]. BLASTP was used to identify soybean orthologs of Arabidopsis SAGs from the TAIR10 version of the *Arabidopsis thaliana* genome annotation.

Availability of supporting data
The raw and processed data sets supporting the results of this article are available in the NCBI Gene Expression Omnibus, as set GSE61857 at http://www.ncbi.nlm.nih.gov/geo/.

Additional files

Additional file 1: Stages of leaf development.

Additional file 2: Stages of cotyledon development.

Additional file 3 Principal component analysis of leaf and cotyledon gene expression.

Additional file 4: Normalized count data for all genes in the leaf and cotyledon tissues.

Additional file 5: Differentially expressed genes and normalized expression values for leaf and cotyledon developmental timecourses with annotations.

Additional file 6: Normalized expression values of Homologs for Arabidopsis Senescence Associated Genes in leaf and cotyledon.

Additional file 7: Enriched GO terms for co-expressed genes.

Additional file 8: Normalized expression values of genes involved in nitrogen, sulfur, potassium, and phosphorus transport.

Additional file 9: Tissue specific genes.

Additional file 10: Differentially expressed genes common between leaf and cotyledon.

Additional file 11: Normalized expression values of NAC annotated genes in leaf and cotyledon.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
AVB and KAH designed the experiment. AVB performed the experiments and analyzed the data. Both AVB and KAH interpreted the results. AVB and KAH wrote the manuscript. Both authors read and approved the final manuscript.

Author details
[1]Department of Agronomy, Purdue University, 915 West State Street, West Lafayette, IN 47907, USA. [2]USDA-ARS Crop Production and Pest Control Research Unit, 915 West State Street, West Lafayette, IN 47907, USA.

References
1. ASA. SoyStats 2013. In: Association AS, editor. A reference guide to important soybean facts and figures. 2013.
2. Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, et al. Genome sequence of the palaeopolyploid soybean. Nature. 2010;463(7278):178–83.
3. Hartman GL, Miles MR, Frederick RD. Breeding for resistance to soybean rust. Plant Dis. 2005;89(6):664–6.
4. Smith JR, Nelson RL. Selection for seed-filling period in soybean. Crop Sci. 1986;26(3):466–9.
5. Rivero RM, Kojima M, Gepstein A, Sakakibara H, Mittler R, Gepstein S, et al. Delayed leaf senescence induces extreme drought tolerance in a flowering plant. Proc Natl Acad Sci. 2007;104(49):19631–6.
6. Smith DL. Cotyledon anatomy in the Leguminosae. Bot J Linn Soc. 1983;86(4):325–55.
7. Harris M, Mackender RO, Smith DL. Photosynthesis of cotyledons of soybean seedlings. New Phytol. 1986;104(3):319–29.
8. Breeze E, Harrison E, McHattie S, Hughes L, Hickman R, Hill C, et al. High-resolution temporal profiling of transcripts during *Arabidopsis* leaf senescence reveals a distinct chronology of processes and regulation. Plant Cell. 2011;23(3):873–94.
9. Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, et al. The developmental dynamics of the maize leaf transcriptome. Nat Genet. 2010;42(12):1060–7.
10. Watanabe M, Balazadeh S, Tohge T, Erban A, Giavalisco P, Kopka J, et al. Comprehensive dissection of spatiotemporal metabolic shifts in primary, secondary, and lipid metabolism during developmental senescence in *Arabidopsis*. Plant Physiol. 2013;162(3):1290–310.
11. Lohman KN, Gan S, John MC, Amasino RM. Molecular analysis of natural leaf senescence in *Arabidopsis thaliana*. Physiol Plant. 1994;92(2):322–8.
12. Gan S, Amasino RM. Making sense of senescence (molecular genetic regulation and manipulation of leaf senescence). Plant Physiol. 1997;113(2):313–9.

13. Gepstein S, Sabehi G, Carp MJ, Hajouj T, Nesher MFO, Yariv I, et al. Large-scale identification of leaf senescence-associated genes. The Plant Journal. 2003;36(5):629–42.

14. Balazadeh S, Riaño-Pachón D, Mueller-Roeber B. Transcription factors regulating leaf senescence in *Arabidopsis thaliana*. Plant Biol. 2008;10(s1):63–75.

15. Hinderhofer K, Zentgraf U. Identification of a transcription factor specifically expressed at the onset of leaf senescence. Planta. 2001;213(3):469–73.

16. Breeze E, Harrison E, Page T, Warner N, Shen C, Zhang C, et al. Transcriptional regulation of plant senescence: from functional genomics to systems biology. Plant Biol. 2008;10(s1):99–109.

17. Shamimuzzaman M, Vodkin L. Transcription factors and glyoxylate cycle genes prominent in the transition of soybean cotyledons to the first functional leaves of the seedling. Functional & Integrative Genomics. 2014;4:683–96.

18. Buchanan-Wollaston V, Earl S, Harrison E, Mathas E, Navabpour S, Page T, et al. The molecular analysis of leaf senescence–a genomics approach. Plant Biotechnol J. 2003;1(1):3–22.

19. Fischer AM. The complex regulation of senescence. Crit Rev Plant Sci. 2012;31(2):124–47.

20. Buchanan-Wollaston V, Page T, Harrison E, Breeze E, Lim PO, Nam HG, et al. Comparative transcriptome analysis reveals significant differences in gene expression and signalling pathways between developmental and dark/starvation-induced senescence in *Arabidopsis*. The Plant Journal. 2005;42(4):567–85.

21. Guo Y, Gan S. AtNAP, a NAC family transcription factor, has an important role in leaf senescence. The Plant Journal. 2006;46(4):601–12.

22. Morales AM, O'Rourke JA, Van DE Mortel M, Scheider KT, Bancroft TJ, Borém A, et al. Transcriptome analyses and virus induced gene silencing identify genes in the Rpp4-mediated Asian soybean rust resistance pathway. Funct Plant Biol. 2013;40(10):1029–47.

23. Prioretti L, Gontero B, Hell R, Giordano M. Diversity and regulation of ATP sulfurylase in photosynthetic organisms. Frontiers in Plant Science. 2014;5.

24. Guo Y, Gan S. Leaf senescence: signals, execution, and regulation. Curr Top Dev Biol. 2005;71:83–112.

25. Himelblau E, Amasino RM. Nutrients mobilized from leaves of *Arabidopsis thaliana* during leaf senescence. J Plant Physiol. 2001;158(10):1317–23.

26. Leran S, Varala K, Boyer JC, Chiurazzi M, Crawford N, Daniel-Vedele F, et al. A unified nomenclature of NITRATE TRANSPORTER 1/PEPTIDE TRANSPORTER family members in plants. Trends Plant Sci. 2014;19(1):5–9.

27. Fan SC, Lin CS, Hsu PK, Lin SH, Tsay YF. The *Arabidopsis* nitrate transporter NRT1.7, expressed in phloem, is responsible for source-to-sink remobilization of nitrate. Plant Cell. 2009;21(9):2750–61.

28. Yu T-S, Kofler H, Häusler RE, Hille D, Flügge U-I, Zeeman SC, et al. The *Arabidopsis* sex1 mutant is defective in the R1 protein, a general regulator of starch degradation in plants, and not in the chloroplast hexose transporter. Plant Cell. 2001;13(8):1907–18.

29. Rushton PJ, Somssich IE, Ringler P, Shen QJ. WRKY transcription factors. Trends Plant Sci. 2010;15(5):247–58.

30. Laurière C, Doyen C, Thévenot C, Daussant J. β-amylases in cereals a study of the maize β-amylase system. Plant Physiol. 1992;100(2):887–93.

31. Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, et al. mapman: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. The Plant Journal. 2004;37(6):914–39.

32. Robatzek S, Somssich IE. A new member of the *Arabidopsis* WRKY transcription factor family, AtWRKY6, is associated with both senescence- and defence-related processes. The Plant Journal. 2001;28(2):123–33.

33. Eulgem T, Rushton PJ, Robatzek S, Somssich IE. The WRKY superfamily of plant transcription factors. Trends Plant Sci. 2000;5(5):199–206.

34. Eulgem T, Somssich IE. Networks of WRKY transcription factors in defense signaling. Curr Opin Plant Biol. 2007;10(4):366–71.

35. Pysh LD, Wysocka-Diller JW, Camilleri C, Bouchez D, Benfey PN. The GRAS gene family in *Arabidopsis*: sequence characterization and basic expression analysis of the SCARECROW-LIKE genes. The Plant Journal. 1999;18(1):111–9.

36. Lin JF, Wu SH. Molecular events in senescing *Arabidopsis* leaves. The Plant Journal. 2004;39(4):612–28.

37. Lin M, Pang C, Fan S, Song M, Wei H, Yu S. Global analysis of the Gossypium hirsutum L. transcriptome during leaf senescence by RNA-Seq. BMC Plant Biology. 2015;15:43.

38. Lijavetzky D, Carbonero P, Vicente-Carbajosa J. Genome-wide comparative phylogenetic analysis of the rice and Arabidopsis Dof gene families. BMC Evol Biol. 2003;3(1):17.

39. Yanagisawa S, Sheen J. Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression. The Plant Cell Online. 1998;10(1):75–89.

40. Waters MT, Wang P, Korkaric M, Capper RG, Saunders NJ, Langdale JA. GLK transcription factors coordinate expression of the photosynthetic apparatus in Arabidopsis. Plant Cell. 2009;21(4):1109–28.

41. Jin J, Zhang H, Kong L, Gao G, Luo J. PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. Nucleic Acids Res. 2013;42(D1):D1182–7. gkt1016.

42. Ulmasov T, Murfett J, Hagen G, Guilfoyle TJ. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. The Plant Cell Online. 1997;9(11):1963–71.

43. Liu L, Zhou Y, Zhou G, Ye R, Zhao L, Li X, et al. Identification of early senescence-associated genes in rice flag leaves. Plant Mol Biol. 2008;67(1–2):37–55.

44. Mueller-Roeber B, Balazadeh S. Auxin and its role in plant senescence. J Plant Growth Regul. 2014;33(1):21–33.

45. Gregersen PL, Holm PB. Transcriptome analysis of senescence in the flag leaf of wheat (*Triticum aestivum* L.). Plant Biotechnol J. 2007;5(1):192–206.

46. Hickman R, Hill C, Penfold CA, Breeze E, Bowden L, Moore JD, et al. A local regulatory network around three NAC transcription factors in stress responses and senescence in *Arabidopsis* leaves. The Plant Journal. 2013;75(1):26–39.

47. Hudson ME, Quail PH. Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. Plant Physiol. 2003;133(4):1605–16.

48. Li Y, Swaminathan K, Hudson ME. Rapid, organ-specific transcriptional responses to light regulate photomorphogenic development in dicot seedlings. Plant Physiol. 2011;156(4):2124–40.

49. Alvarez-Canterbury AMR, Flores DJ, Keymanesh K, To K, Brusslan JA. A double SORLIP1 element is required for high light induction of ELIP genes in Arabidopsis thaliana. Plant Mol Biol. 2014;84(3):259–67.

50. Giuliano G, Pichersky E, Malik V, Timko M, Scolnik P, Cashmore AR. An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. Proc Natl Acad Sci. 1988;85(19):7089–93.

51. Guiltinan MJ, Marcotte WR, Quatrano RS. A plant leucine zipper protein that recognizes an abscisic acid response element. Science. 1990;250(4978):267–71.

52. Ulmasov T, Hagen G, Guilfoyle TJ. ARF1, a transcription factor that binds to auxin response elements. Science. 1997;276(5320):1865–8.

53. Hagen G, Guilfoyle T. Auxin-responsive gene expression: genes, promoters and regulatory factors. Plant Mol Biol. 2002;49(3–4):373–85.

54. Izawa T, Foster R, Chua N-H. Plant bZIP protein DNA binding specificity. J Mol Biol. 1993;230(4):1131–44.

55. Kim SY, Ma J, Perret P, Li Z, Thomas TL. Arabidopsis ABI5 subfamily members have distinct DNA-binding and transcriptional activities. Plant Physiol. 2002;130(2):688–97.

56. Jakoby M, Weisshaar B, Dröge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, et al. bZIP transcription factors in *Arabidopsis*. Trends Plant Sci. 2002;7(3):106–11.

57. Harris JC, Hrmova M, Lopato S, Langridge P. Modulation of plant growth by HD-Zip class I and II transcription factors in response to environmental stimuli. New Phytol. 2011;190(4):823–37.

58. Bao X, Franks RG, Levin JZ, Liu Z. Repression of AGAMOUS by BELLRINGER in floral and inflorescence meristems. Plant Cell. 2004;16(6):1478–89.

59. Meng Y, Li H, Wang Q, Liu B, Lin C. Blue light–dependent interaction between Cryptochrome2 and CIB1 regulates transcription and leaf senescence in soybean. Plant Cell. 2013;25(11):4405–20.

60. Yin G, Xu H, Xiao S, Qin Y, Li Y, Yan Y, et al. The large soybean (*Glycine max*) WRKY TF family expanded by segmental duplication events and subsequent divergent selection among subgroups. BMC Plant Biology. 2013;13(1):148.

61. Hudson KA. The circadian clock-controlled transcriptome of developing soybean seeds. The Plant Genome. 2010;3(1):3–13.

62. Fehr WR, Caviness CE. Stages of Soybean Development. Iowa, USA: Iowa State University Cooperative Extension Service; 1977. vol. Special Report 80.

63. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357–9.

64. Severin AJ, Woody JL, Bolon Y-T, Joseph B, Diers BW, Farmer AD, et al. RNA-Seq Atlas of Glycine max: a guide to the soybean transcriptome. BMC Plant Biology. 2010;10(1):160.

65. Anders S, Huber W. Differential expression of RNA-Seq data at the gene level–the DESeq package. Heidelberg, Germany: EMBL; 2012.

66. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40.