**OPEN**

Correspondence and
requests for materials
should be addressed to
X.Y.S. (xshan@bch.
msstate.edu)

# Relating significance and relations of differentially expressed genes in response to *Aspergillus flavus* infection in maize

Matthew C. Asters[1], W. Paul Williams[2], Andy D. Perkins[3], J. Erik Mylroie[2], Gary L. Windham[2] & Xueyan Shan[1]

[1]Department of Biochemistry, Molecular Biology, Entomology, and Plant Pathology, Mississippi State University, Mississippi State, Mississippi, USA, [2]Agricultural Research Service, United States Department of Agriculture, Corn Host Plant Resistance Research Unit, Mississippi State, Mississippi, USA, [3]Department of Computer Science and Engineering, Mississippi State University, Mississippi State, Mississippi, USA.

***Aspergillus flavus* is a pathogenic fungus infecting maize and producing aflatoxins that are health hazards to humans and animals. Characterizing host defense mechanism and prioritizing candidate resistance genes are important to the development of resistant maize germplasm. We investigated methods amenable for the analysis of the significance and relations among maize candidate genes based on the empirical gene expression data obtained by RT-qPCR technique from maize inbred lines. We optimized a pipeline of analysis tools chosen from various programs to provide rigorous statistical analysis and state of the art data visualization. A network-based method was also explored to construct the empirical gene expression relational structures. Maize genes at the centers in the network were considered as important candidate genes for maize DNA marker studies. The methods in this research can be used to analyze large RT-qPCR datasets and establish complex empirical gene relational structures across multiple experimental conditions.**

The fungus *Aspergillus flavus* causes maize ear rot and produces aflatoxins which are potent health hazards to humans and animals[1,2]. Enhancement of maize host plant resistance to *A. flavus* infection is desirable to reduce aflatoxin contamination at the pre-harvest stage of maize production. The host plant resistance in maize to *A. flavus* infection is a quantitative trait involving co-expression of many genes[3–5]. Identification of controlling genes and their empirical network relations is essential to the development of DNA markers and the transfer of maize resistance into elite commercial maize lines.

Plants have developed multiple defense mechanisms against pathogen invasion[6]. An early event in defense responses is triggered by the pathogen molecules that carry pathogen-associated molecular patterns (PAMPs) such as lipopolysaccharides and ssRNA[7,8]. PAMP-triggered immunity (PTI) is activated to control the spread of pathogen at the infection site[9]. A further event of defense responses happens when pathogens release effectors into the host plant cells to overcome the first defense system and enable the parasitic infection. In some cases the pathogen effectors can be recognized by specific host plant resistance proteins (R proteins) and the effector-triggered immunity (ETI) is activated to turn on the systemic defense mechanism for elevated resistance in the whole plant[10–12]. Both the PAMP-triggered immunity (PTI) and the effector-triggered immunity (ETI) in plants are associated with the activation or repression of specific plant defense-related genes. The RNA transport pathway protein complexes are critical in the regulation of gene expression and activation for effective plant defense responses[13–15].

RNA transport pathways comprise various protein complexes that regulate gene expression and nucleocytoplasmic trafficking. RNAs are transcribed in the nucleus and transported across the nuclear membrane with the help of specific protein complexes in RNA transport pathways[16]. Specific RNA molecules are transported through well-defined pathways. The transport of messenger RNAs (mRNAs) is different from that of ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), or small nuclear RNAs (snRNAs). For instance, protein complexes such as cap binding complex (CBC), spliceosome, transcription-export complex (TREX), exon-junction complex (EJC), and translation initiation factors (eIFs) are involved in the serial of events associated with the transport and trans-

| Table 1 | Alfatoxin concentrations in mature kernels of the six maize inbred lines used in this study | | | | |
|---|---|---|---|---|---|
| Pedigree | Host Plant Response | Treatment | Alfatoxin (ng/g) | | Level* |
| Va35 | Susceptible | Inoculated | 1821 | | a |
| | | Uninoculated | ≈1 | | |
| Mp04:85 | Susceptible | Inoculated | 1594 | | ab |
| | | Uninoculated | ≈1 | | |
| Mp04:89 | Susceptible | Inoculated | 1236 | | ab |
| | | Uninoculated | ≈1 | | |
| Mp04:104 | Resistant | Inoculated | 255 | | c |
| | | Uninoculated | ≈1 | | |
| Mp718 | Resistant | Inoculated | 199 | | c |
| | | Uninoculated | ≈1 | | |
| Mp719 | Resistant | Inoculated | 53 | | d |
| | | Uninoculated | ≈1 | | |

*Means followed by the same letter do not differ at p = 0.05 (Fisher's Protected LSD). Test of significance were performed on transformed value (log(Y + 1)).

lation of mRNAs. On the other hand, importins, exportins, Ran-GTP related protein complex, and the survival mortor neuron complex (SMN) are involved in the transport of rRNA, tRNA, and snRNA molecules[17–19]. Nevertheless, all RNAs are transported across the nuclear membrane through interactions with the nucleoporins in nuclear pore complexes (NPCs)[20,21]. In fact, components of RNA transport pathways interlink in functions and overlap with the pathways of nucleocytoplasmic trafficking for all macromolecules including proteins. The nucleocytoplasmic trafficking pathways are fundamental for normal cell functions as well as plant defense responses[22,23].

Studies have demonstrated that RNA transport pathway genes play direct roles in plant defense systems. Several reports have shown that nucleoporins directly regulate the transport of R proteins. Mutations in certain nucleoporins reduce the nuclear accumulation of specific R proteins and hence compromise resistance[24–26]. The expression patterns of maize RNA transport pathway genes and their relations in response to *A. flavus* infection have not yet been reported. Identification of maize defense-related genes, their regulatory roles, and expression relations responding to *A. flavus* infection in the empirical gene expression network is most important for maize resistance breeding. The advance of quantitative real time PCR (RT-qPCR) technique makes it possible to precisely describe gene expression patterns and compare the changes in gene expression levels[27]. In contrast to the comprehensive genome wide microarray and RNA sequencing techniques, RT-qPCR provides a powerful and flexible tool which allows focusing on individual pathways across a wide range of experimental conditions with remarkable sensitivity, specificity, and accuracy[28]. Although a number of RT-qPCR analysis packages are available, they vary widely in terms of algorithms and capacities for data analysis. Appropriate analysis procedures that are tailored to perform comprehensive quantitative analysis of RT-qPCR data are needed to conduct rigorous statistical analysis and make inferences from gene expression data[29]. The objectives of this study were to explore and select appropriate methods for analysis of RT-qPCR gene expression data and investigate the expression of maize RNA transport pathway genes in response to *A. flavus* infection in selected resistant and susceptible maize inbred lines. Particularly, construction of empirical gene expression relational structures was investigated in order to identify candidate genes that play important roles in maize host resistance to *A. flavus* infection.

## Results
**Aflatoxin concentrations in mature kernels of the resistant and susceptible maize inbred lines.** The selected maize inbred lines used in this study were recombinant inbred lines developed from $F_2$ plants of the cross Mp715 × Va35. There were two identical copies for each gene in each maize inbred line. Only up to two different alleles for each gene were present among all the six maize inbred lines because they were offspring lines from a single cross. To determine the resistance or susceptibility of each maize inbred line, aflatoxin accumulation was evaluated using the mean values of aflatoxin concentrations per 50 g ground mature kernels. Aflatoxin concentrations of the six maize inbred lines (Mp718, Mp719, Mp04:104, Mp04:89, Mp04:85, and Va35) were evaluated along with other maize inbred lines planted in the field. Each maize inbred line had three replications and was subjected to two treatments (inoculated and non-inoculated with *A. flavus*).The six maize inbred lines exhibited four levels of aflatoxin accumulation (Table 1). Mp719 exhibited the highest level of resistance among the six maize inbred lines with a significantly low value of aflatoxin concentration (53 ng/g). The susceptible maize inbred line Va35 was highest in aflatoxin concentration with a mean value of 1821 ng/g. The significance levels in the differences of aflatoxin accumulation levels among the tested maize inbred lines have been consistent over field trials for multiple years[30,31].

**Quantitative RT-PCR assays.** Total RNA samples were prepared from developing kernels of the six resistant and susceptible maize inbred lines resulting in a total of 72 samples. Quantitative RT-PCR analysis and standard curve assays were performed and PCR efficiencies were calculated (Table 2). Test of RT-qPCR primers was performed on a total of 66 maize genes initially, including 50 RNA transport pathway genes and 16 differentially expressed candidate genes identified from previous studies[5]. Out of the 66 RT-qPCR primer evaluation assays, 56 gene primers yielded RT-qPCR data of good quality with a PCR efficiency (in r-squared value) >0.9 (Table 2) and therefore were included for the subsequent whole plate data analysis. The primer sequences for the selected RNA transport pathway genes are listed in Table 2. The gene IDs and functions of all the 56 tested genes are listed in Table 3. There were three whole plate assays for the reference gene GAPDH. Missing values in one GAPDH assay were corrected by calculation of the corresponding values from the other GAPDH assays.

**Identification of differentially expressed genes in RNA transport pathways.** The relative delta $C_q$ values obtained from preprocessing the raw RT-qPCR data were used as the gene expression values for the subsequent descriptive statistical analysis, analysis of variance (ANOVA), correlation analysis, and network analysis. Summary of the distributions of the gene expression values were presented by boxplots in Figure 1A with the median, spread and outliers showing for each gene. Large amount of outliers in the expression values from RNA transport pathway genes were observed. Since there were only up to two different alleles for each gene being involved in all 72 samples, the abundant expression variations observed in these genes indicated that different gene regulating patterns existed in these maize recombinant inbred lines. Scatterplots were used to evaluate if there were any trends present in the regulating patterns for each gene related to resistant or susceptible maize inbred lines (Figure 1B–

## Table 2 | Primer sequences and PCR efficiencies in RT-qPCR reactions

| Primer | Forward | Reverse | PCR Efficiency (r-squared values)* |
|---|---|---|---|
| ACIN1 | GTACGCCCATTTACCCTGAA | ATTCGGCGACCAACTTATTG | 0.9837 |
| CBC | GGACGAAGTGGTGGACAAGT | CTCATAGGGAGGTGGTGCAT | 0.9857 |
| CRM1 | GCTTCAGGATTCCAAAACCA | TAGACCATCAACCGCACAAA | 0.9814 |
| eEF1A | AGGCTGACTGTGCTGTCCTT | ACGTGCCTTGGAGTATTTGG | 0.9536 |
| eIF1 | GAAAGAACGAGGCTGACGAC | TCCACACCTTCTTGTGCATC | 0.9454 |
| eIF2 | CGAGGCCTTCAAGCTCATAG | CTTGAGAGGCTGTGGTGTCA | 0.9231 |
| EIF4A3 | GAGGACCTTGCGAACAAGAG | GGTTCCGTCATGAACTTGCT | 0.7357 |
| eIF4E | GGTCGATTTTGGGAGGATTT | CATCACAGCCTGATGGTCTG | 0.9849 |
| eIF5 | GGACCACTCAACCTCTCCAA | CTCCTCAGTGGACAGCATCA | 0.8817 |
| eIF5B | TTGGGTGTTGATGTCCTTGA | GAAGCTCCTCTGCTGCTCAT | 0.9992 |
| Gemin2 | GGGTTCTCCGGTGAGTATGA | AGATCCCCATTTCTGTGTGC | 0.9014 |
| IPOB | TCTATGCGATGCCTATGCTG | ATAATATGCGGAGCGAATGG | 0.9557 |
| MAGOH | CAGGAGGATCATCCAGGAGT | GTCCTTGCTGGACTGGACAT | 0.9774 |
| NMD3 | CACCACCTGGAGGAGAAAAA | TCTTCCCGAAATCTGACACC | 0.9706 |
| Nup53 | GCGAACTGGGTGCATATCTT | TCCCCTGGTAGTTCTCGTTG | 0.9937 |
| Nup62 | TGAAACACAAACAGGCTTGG | TCATGCTCCACAACTTCTGC | 0.9770 |
| Nup85 | CCTGATCATGCAGAATGGTG | CCAAGCAGCCCTTAAACTTG | 0.9788 |
| Nup88 | CTCTGTGCATCCTGTCCTCA | CTTCCAACCCTTCATCTCCA | 0.9855 |
| Nup96 | ACGGCAGCATCAAATTTTTC | AGTCACCTCTGCAGCCTTGT | 0.9811 |
| Nup98 | ACGGCAGCATCAAATTTTTC | AGTCACCTCTGCAGCCTTGT | 0.9774 |
| Nup133 | GGATGCTGTGGTGTATGACG | TTCCTTGCCCCTAATCACTG | 0.9776 |
| Nup160 | GTGCCACTTCTGCCTTCTTC | CAGCAACAATGACTGGCAAC | 0.9693 |
| PABP | TGCTATGATACCCCCTCAGC | TCCAGCTCCACGCCTAATAC | 0.9301 |
| plcln | CTCTACACCTTCGCCGACAT | TCCTGGTGGTGAGGTAAAGG | 0.9282 |
| Pinin | CCCGAATTTGGTGAGAAGAA | TCTGCAATTTGCTCACGTTC | 0.9966 |
| PRMT5 | TTGGGAAAGCCTGGTTACTG | GAAATCCCATCAGGCTTCAA | 0.9819 |
| PYM | ACAACCGGAATCATCTGCTC | GCTCCTTAAGCCATCCATCA | 0.9711 |
| Rae1 | ATTCAAATGCCACAGGGAAG | TACTGCAAGGAATGGGAAGG | 0.9829 |
| Ran | AGAGAAGTTTGGTGGCCTCA | ATTACCGCAGAGGACAATGG | 0.9519 |
| RanGAP | AAGCTGACCTTGGCTGAAAA | TTCAGCTGCACAAAATCTGG | 0.9782 |
| Ref/Aly | CAGAAGTTGGGGAGCTCAAG | TGGTTCGACATTGTTTCCAA | 0.9829 |
| SAP18 | AAGAGGTTGCTCTTCCAGCA | CATGATTGCAACGCTCAAGT | 0.9763 |
| Sec13 | GCTTGTCTCAGGAGGGTTTG | ATGACAACCTTCCCGTCTTG | 0.9965 |
| SPN1 | GCAAACGATGTTTGGTTGTG | AAATGCAATCGAGGATGGAG | 0.9660 |
| SUMO | GATGGCAACGAGGTGTTCTT | TCCTCCATCTCCAGCTCATC | 0.9480 |
| TEX1 | GACCTGGTGTCGAGGGAATA | CTGTGTGGCCTTTCAGTTCA | 0.9654 |
| TGS1 | CTATGGAGCAAACGCACTCA | TGCCAGCAGAGTAGATGGTG | 0.9461 |
| THOC5 | GATGTCTCTGCACCTCGTCA | CCCGGATACTTCGTCTGAAA | 0.9160 |
| THOC7 | AACTGATTTGAAGCGGGAGA | TCCCTTCTTCCATGCTCATC | 0.9504 |
| UAP56 | CTACGTCGGGATCCACAGTT | GTTTTCCCCATCCCAGATTT | 0.9750 |
| UBC9 | GTTCCCTCAGGGTTTCTTCC | AGGTGATAGCCATCCGTTTG | 0.9221 |
| Urn | CATGGGTGGAGTAAGGCTGT | TTCGCGTCCCAGAATTTTAC | 0.9521 |
| GAPDH | CGACTTACTTGGTGACAGCAG | CGCCATCCACATTTATTCTCG | 0.9713 |
| AI065864 | AGAATCGATCCGCCAAGTTA | AGGTTGCAACGCTATTGGTC | 0.9942 |
| AI065909 | TACCACAGCAGAGCAACCAC | ATCTCCGGCTGAAGAAGACA | 0.9858 |
| AI664980 | CTGACACAAAGCGACCTTCA | ATCCTGTTCGCTACCGTGTT | 0.9989 |
| BE050050 | CCGTGGAAATGTGGTAATCC | ATCCACGTCAACCATCTTCC | 0.9589 |
| BG266083 | CTTTGCATCACAAAGCTCCA | GGTGAGGAAGAGCAAATGGT | 0.9984 |
| BM078796 | TTTTCTCCACCTCGGTCTTG | AGCGTGAGCTCCTACGACAT | 0.9958 |
| BM379345 | TTCACACACACCACACAATACC | CTGCAACTGTTGATCCCATC | 0.996 |
| BM498943 | CTCTGTATTGGCCCACGACT | AATTGTCGAGGTCGGAGATG | 0.9957 |
| CA399536 | GGCTGATGCAATAAGGTGGT | TTGTTGCCATTCTACCCACA | 0.9893 |
| CD443591 | ATAGCAGCCATCCTCCATTG | GGGAAGAACATCCCCTTGAT | 0.993 |
| TC207503 | AAACGCCATTGCACATTACA | TCTTGAAGGATCGTGTGCTG | 0.98 |
| TC223736 | AACGGTCAGAATTGGAGTGC | GACGACGCAACAGATCTCAA | 0.9873 |
| TC231674 | GGGCTTCTTGTTGTGCTCTC | TTAAAGCGCTGCCTTATTCC | 0.9469 |
| TC237311 | TGAGGATCATGGAGGAGGAC | CCACATTCACGGGCTTATCT | 0.9932 |
| TC238832 | AGACATGGGATACCGAGACG | AGCTCCATCAGCTCCTTGAA | 0.9995 |
| TC247683 | ATGATGGGAGGCTGACTTTG | TCTCAGCGAAATTCATCGTG | 0.9765 |

*To calculate RT-qPCR efficiency (E), use formula E = (1 + rsquared value).

E). Four examples are shown in Figure 1B–E regarding the regulating trends in gene expression patterns discovered in this study. A translation initiation factor gene eIF5B appeared to express consistently among samples (Figure 1B). The nucleoporin Nup133 gene had significant variations in gene expression values with down-regulation patterns showing in resistant maize inbred lines (Figure 1C). AI664980 was another example showing significant variations among samples with down-regulation patterns in resistant maize inbred lines (Figure 1E). TC231674 was found highly expressed in the resistant maize inbred line Mp718. It showed significant variations among samples with up-regulation patterns in resistant maize inbred lines (Figure 1D).

**Table 3 | Grouping by functions of the analyzed maize genes obtained from database searching**

| Grouping | Gene ID | Function |
|---|---|---|
| **Nuclear Pore Complex(NPC)** | Nup53* | nuclear pore complex protein Nup53 |
| | Nup62* | nuclear pore complex protein Nup62 |
| | Nup85 | nuclear pore complex protein Nup85 |
| | Nup88* | nuclear pore complex protein Nup88 |
| | Nup96 | nuclear pore complex protein Nup98-Nup96 |
| | Nup98 | nuclear pore complex protein Nup98-Nup96 |
| | Nup133* | nuclear pore complex protein Nup133 |
| | Nup160* | nuclear pore complex protein Nup160 |
| | Rae1 | mRNA export factor |
| | Ran* | GTP-binding nuclear protein |
| | RanGAP | Ran GTPase-activating protein 1 |
| | Sec13* | protein transport protein SEC13 |
| | SUMO* | small ubiquitin-related modifier |
| | UBC9* | ubiquitin-conjugating enzyme E2 I |
| **Survival Mortor Neuron Complex(SMN)** | Gemin2 | survival of motor neuron protein-interacting protein 1 |
| | IPOB* | importin subunit beta-1 |
| | plcln* | chloride channel, nucleotide-sensitive, 1A |
| | PRMT5* | protein arginine N-methyltransferase 5 |
| | SPN1* | snurportin-1 |
| | TGS1* | trimethylguanosine synthase |
| | Urn | serine-threonine kinase receptor-associated protein |
| **Translation Initiation Factors(EIFs)** | CBC | nuclear cap-binding protein subunit 2 |
| | CRM1 | exportin-1 |
| | eEF1A | elongation factor 1-alpha |
| | eIF1 | translation initiation factor 1A |
| | eIF2 | translation initiation factor 2 subunit 1 |
| | eIF4E | translation initiation factor 4E |
| | eIF5 | translation initiation factor 5 |
| | eIF5B | translation initiation factor 5B |
| | PABP | polyadenylate-binding protein |
| **Exon-Junction Complex(EJC)** | ACIN1 | apoptotic chromatin condensation inducer in the nucleus |
| | EIF4A3 | ATP-dependent RNA helicase |
| | MAGOH* | protein mago nashi |
| | Pinin | pinin_SDK_memA domain family protein |
| | Ref/Aly* | THO complex subunit 4 |
| | SAP18* | histone deacetylase complex subunit |
| | UAP56 | ATP-dependent RNA helicase UAP56/SUB2 |
| **Trasncription-Export Complex(TREX)** | NMD3 | nonsense-mediated mRNA decay protein 3 |
| | PYM | partner of Y14 and mago |
| | Tex1 | THO complex subunit 3 |
| | THOC5 | THO complex subunit 5 |
| | THOC7* | THO complex subunit 7 |
| **Genes Associated with Mp313E(RES)** | TC231674* | NPCs-NUP85 like, RNA Transport |
| | AI065909 | |
| | BE050050* | |
| | BM078796 | Heat Shock Protein 26 (HSP26) |
| | BM379345* | Metallothionein-Like Protein (MTLP) |
| | BM498943 | Ethylene Responsive Protein (ETHRP) |
| | CD443591* | |
| | TC207503 | Prenylated Rab Acceptor (PRA1) |
| | TC223736 | Heat Shock Protein 90 (HSP90) |
| | TC237311 | Heat Shock Protein 101 (HSP101) |
| | TC238832* | Lecithin choesterol acyltransferase (LCAT) |
| | TC247683* | |
| **Genes Associated with Va35(SUS)** | AI065864* | Exonuclease-Endonuclease-Phosphatase (EEP) |
| | AI664980* | Glycine Rich RNA Binding Protein2 (GRBP2) |
| | BG266083* | Heat Shock Protein 18a(HSP18a) |
| | CA399536* | |

*P values were obtained by performing ANOVA on the RT-qPCR data. The significance level is p value $<0.05$.

Analysis of variance (ANOVA) was used to determine the significant levels of the differentially expressed genes among different groups and to identify the sources of the variations associated with maize resistance to aflatoxin accumulation. Contrasts for ANOVA analysis were constructed among maize inbred lines (by pedigree) and between resistant and susceptible groups (by RES). Table 4 shows the p values for each gene obtained from ANOVA analysis based on the general linear models by RES, pedigree, RES*INOC, or pedigree*INOC. Of the 56 genes analyzed, 23 were differentially expressed among the tested maize inbred lines at a significance level of $p < 0.05$ (Table 4, column Pedigree), and 17 were found significant in expression differences between the resistant group and the susceptible group at $p < 0.05$ (Table 4, column Res). These significant genes included the nucleoporins Nup133, Nup62, Nup160, Nup85,
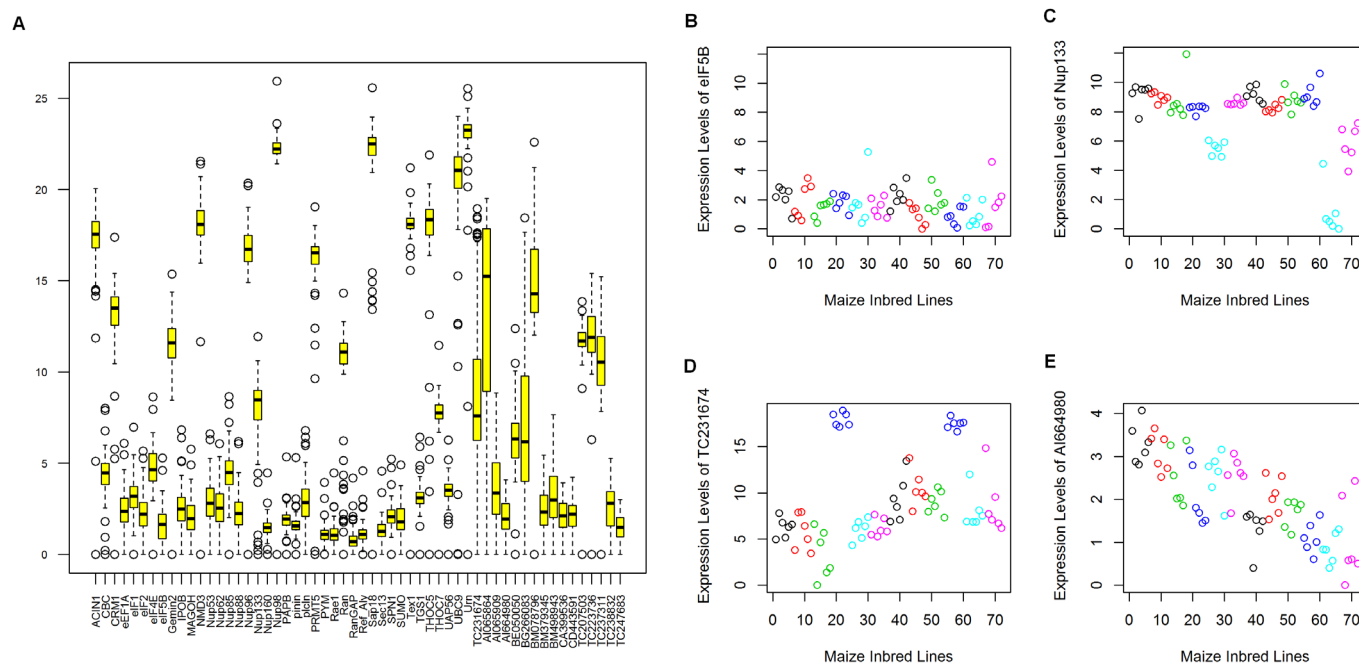
**Figure 1 | An overview on the RT-qPCR gene expression data for the candidate genes after normalized with the reference gene GAPDH.** (A) Boxplots showing the distributions (median, spread and outliers) of the gene expression values for each candidate gene. The horizontal axis represents the gene IDs. The vertical axis represents the relative delta $C_q$ values. (B–E) Examples of scatterplots showing trends in the expression values over the 72 samples for four selected genes. The horizontal axis represents the 72 samples with different colors coded for the six maize inbred lines. Samples 1–36 were collected at 2 DAI and samples 37–72 were at 7 DAI. The vertical axis represents the relative delta $C_q$ values.

Nup88, Nup53, UBC9, SUMO, and Sec13; the Survival Motor Neuron complex genes Ran, TGS1, SPN1, IPOB, plcln, and PRMT5; the Exon-Junction Complex genes MAGOH, Sap18, Ref_Aly. Most of the significant genes identified among the RNA transport pathway genes were from the NPC and SMN protein complexes. Some of the previously identified candidate genes BG266083, CD443591, BE050050, TC231674, CA399536, AI065864, AI664980, TC238832, BM379345, and TC247683 were again found differentially expressed in this research that had a set of germplasm different from the previous studies[5]. None of the translation initiation factors (EIFs) were found differentially expressed among the resistant maize lines and susceptible lines (Table 3 and 4).

**Correlations in gene expression between the tested maize genes.** To determine if there were co-expression patterns in gene expression between all pairs of the tested genes, correlation analysis was performed on gene expression values and the correlation matrices were visualized by using R package "Corrgram". Figures 2A–B are correlograms displaying the correlation matrices of Pearson's correlation coefficients between selected pairs of genes. The genes displayed in Figure 2A–B were selected from the significant genes identified from the ANOVA analysis presented in different functional groups. Correlations displayed in a correlogram were organized in the order that genes have similar expression patterns were grouped together. The signs and values of the Pearson's coefficients were reflected schematically with the correlation coefficients and the 95% confidence intervals displayed at the lower triangle, whereas the color-coded pie graphs in the upper triangle. Figure 2A showed expression correlations among genes in the NPC and SMN protein complexes of maize RNA pathways along with other previously identified candidate genes. Resistance related genes BE050050, TC231674, and BM498943 were found positively correlated to each other with a coefficient being at least 0.66. Resistance related gene TC238832 was highly correlated with the SUMO gene, an ubiquitin related disease defense gene. Susceptibility related gene AI664980 was found positively correlated to a nucleoporin gene Nup62 (0.88) and

negatively correlated with a resistance related gene BE050050 (−0.97). Figure 2B showed the expression correlations among selected RNA pathway genes in the EIFs, EJC, and TREX protein complexes. Susceptibility related gene BG266083 was highly correlated with Sap18 which was an Exon-Junction Complex gene. A TREX gene THOC7 was found negatively correlated with the resistance related genes TC231674, BE050050, and BM379345, which suggested that down-regulation of the THOC7 gene was likely involved in maize defense responses.

Figure 3 is an eigenvector plot showing the results from a Principal Component analysis (PCA) analysis on the correlation coefficients of the selected significant genes. The distance between genes in the correlation coefficients was illustrated by the angle formed between the gene eigenvectors. The length of an eigenvector represents the largest variance for each gene in the correlation coefficients. Gene eigenvectors placed close to each other were more similar in the expression patterns and hence were more positively correlated. For instance, THOC7 and Nup62 were found to be highly correlated with the expression of the susceptibility related gene AI664980. On the other hand, UBC9 was found to be positively correlated with the resistance related genes TC231674, BE050050, and BM498943 genes.

The inclusion of previously identified candidate genes in this research provided a way to make some of the observations found in this study experimentally verifiable. Based on the directions of the gene eigenvectors, the expression of resistance related gene TC231674 (on chromosome 5, highly expressed in resistant maize inbred line Mp718) was shown positively correlated with the expression of resistance related gene BE050050 (close to maize resistance SSR marker bnlg2291 on chromosome 4) and negatively correlated with the expression of the susceptibility related gene AI664980 (GRBP2, highly expressed in susceptible maize inbred line Va35) across all the tested maize inbred lines in this study. This observation was consistent with the previous findings on the expression patterns of these defense related genes in a different set of germplasm where TC231674 was found highly expressed in a different resistant maize inbred line Mp313E[5].

**Table 4 | P values obtained by using analysis of variance (ANOVA) for all the tested genes between different contrasting groups**

| GENE ID | RES | PEDIGREE | RES by INOC | PEDIGREE by INOC |
|---|---|---|---|---|
| Nup133#* | 6.85E-06# | 1.07E-15* | 9.06E-06 | 1.71E-14 |
| Nup62#* | 1.96E-05# | 0.000934241* | 2.07E-05 | 0.001574375 |
| Ran#* | 7.84E-05# | 4.68E-05* | 9.79E-05 | 0.000124295 |
| BG266083#* | 0.000760222# | 5.48E-08* | 0.000900728 | 1.95E-07 |
| CD443591#* | 0.001011978# | 2.32E-06* | 0.001049726 | 5.69E-06 |
| BE050050#* | 0.001477429# | 0.020556229* | 0.001616464 | 0.022687508 |
| TC231674#* | 0.001629489# | 4.34E-17* | 0.001891075 | 1.14E-15 |
| CA399536#* | 0.001821379# | 6.43E-10* | 0.002083541 | 3.01E-09 |
| SPN1#* | 0.002466566# | 0.035539412* | 0.002710653 | 0.040688655 |
| Sap18#* | 0.002781819# | 0.000638323* | 0.00289673 | 0.000910661 |
| AI065864#* | 0.002854662# | 4.16E-30* | 0.003046812 | 2.36E-30 |
| AI664980#* | 0.003382439# | 0.06263276* | 0.003848577 | 0.080973849 |
| UBC9#* | 0.004676791# | 2.77E-05* | 0.005275719 | 7.91E-05 |
| THOC7# | 0.025062924# | 0.053049132 | 0.022244284 | 0.050063859 |
| Nup160# | 0.038681499# | 0.061780109 | 0.039487014 | 0.075999717 |
| Ref_Aly# | 0.04002563# | 0.305000259 | 0.039914685 | 0.296859356 |
| BM379345#* | 0.041254015# | 0.00305644* | 0.042569074 | 0.004359345 |
| Rae1 | 0.074822354 | 0.189985307 | 0.071777903 | 0.199925282 |
| MAGOH* | 0.083705888 | 0.012263152* | 0.086877894 | 0.01828929 |
| Gemin2 | 0.109430056 | 0.278073942 | 0.109846956 | 0.285886467 |
| TC247683* | 0.110272735 | 1.69E-05* | 0.109958835 | 2.82E-05 |
| THOC5 | 0.123659983 | 0.561576115 | 0.126237595 | 0.564615945 |
| TGS1* | 0.153359964 | 0.000933695* | 0.157352443 | 0.001206472 |
| RanGAP | 0.165012818 | 0.323225827 | 0.165980776 | 0.350264571 |
| BM078796 | 0.176571285 | 0.161989688 | 0.170022083 | 0.13998881 |
| pinin | 0.177138676 | 0.265426044 | 0.17581723 | 0.278794776 |
| plcln* | 0.187276309 | 4.29E-08* | 0.191235837 | 8.35E-08 |
| eIF5B | 0.200763378 | 0.339198949 | 0.204002426 | 0.381158455 |
| AI065909 | 0.208411015 | 0.244545466 | 0.205745798 | 0.252342955 |
| BM498943 | 0.215685896 | 0.059314695 | 0.216795808 | 0.071322972 |
| NMD3 | 0.267670961 | 0.263074967 | 0.268940316 | 0.275968665 |
| eEF1A | 0.280758999 | 0.069161183 | 0.283664394 | 0.088458125 |
| PYM | 0.329579917 | 0.441751167 | 0.327269603 | 0.460562856 |
| eIF2 | 0.338916479 | 0.113125646 | 0.34270682 | 0.140936244 |
| ACIN1 | 0.397054749 | 0.841492917 | 0.398470605 | 0.841896453 |
| SUMO* | 0.421678638 | 0.027676888* | 0.42466696 | 0.036632958 |
| eIF4E | 0.430938336 | 0.114396136 | 0.428871994 | 0.122125126 |
| CBC | 0.444664816 | 0.083222119 | 0.445024629 | 0.087939214 |
| CRM1 | 0.451961026 | 0.20491634 | 0.456955632 | 0.222832837 |
| PAPB | 0.474446141 | 0.785186063 | 0.47377122 | 0.789371648 |
| Nup98 | 0.520779041 | 0.491404919 | 0.523641684 | 0.515304266 |
| Nup53* | 0.555509261 | 5.65E-05* | 0.558261927 | 0.000127474 |
| UAP56 | 0.557866549 | 0.344854255 | 0.559285991 | 0.364086021 |
| TC207503 | 0.602504748 | 0.065869398 | 0.602795853 | 0.065693085 |
| Nup85 | 0.649107717 | 0.15302179 | 0.653127094 | 0.176059677 |
| Nup88* | 0.666415513 | 0.008156039* | 0.669772544 | 0.012650164 |
| TC238832* | 0.693399134 | 0.000301083* | 0.694436938 | 0.00050016 |
| Sec13* | 0.700390081 | 0.039807357* | 0.695923088 | 0.040694796 |
| TC223736 | 0.70870598 | 0.704724564 | 0.705340869 | 0.693326914 |
| Nup96 | 0.782725932 | 0.156130738 | 0.780095262 | 0.146805988 |
| Tex1 | 0.81266049 | 0.649474045 | 0.810910727 | 0.650319502 |
| IPOB* | 0.820575413 | 0.008944226* | 0.822316491 | 0.013708982 |
| Urn | 0.879871826 | 0.712300347 | 0.880040238 | 0.712395487 |
| eIF1 | 0.881820391 | 0.223508933 | 0.882667121 | 0.243135231 |
| TC237311 | 0.934340407 | 0.078564895 | 0.934658645 | 0.085891538 |
| PRMT5* | 0.993641272 | 0.044085505* | 0.993661343 | 0.048024231 |

#*P values were obtained by performing ANOVA on the RT-qPCR data. The significance level is p value <0.05.

**Determination of the roles and relations among the tested genes in the empirical gene expression network.** The genes selected from the RNA transport pathways were considered as elements in a static gene network in terms of potential biological processes. In order to determine the dynamic roles and relations in expression of these genes responsive to *A. flavus* infection, we wanted to explore methods to construct empirical gene relational networks that were based on the variations in the actual gene expression levels. To achieve this goal, we conducted PCA on the gene expression data.

The scores of the first two principal components (pc1 and pc2) associated with each gene were used to calculate a Euclidean distance matrix between all pairs of genes for the network construction. Figures 4A–B are network graphs constructed based on the Euclidean distance matrices. The vertices in the network represented genes. The edges represented the Euclidean distance between each pair of genes on the pc1 × pc2 plane. To highlight genes by protein complexes or groups, the vertices were color-coded for the seven different groups where the genes were chosen from. Five
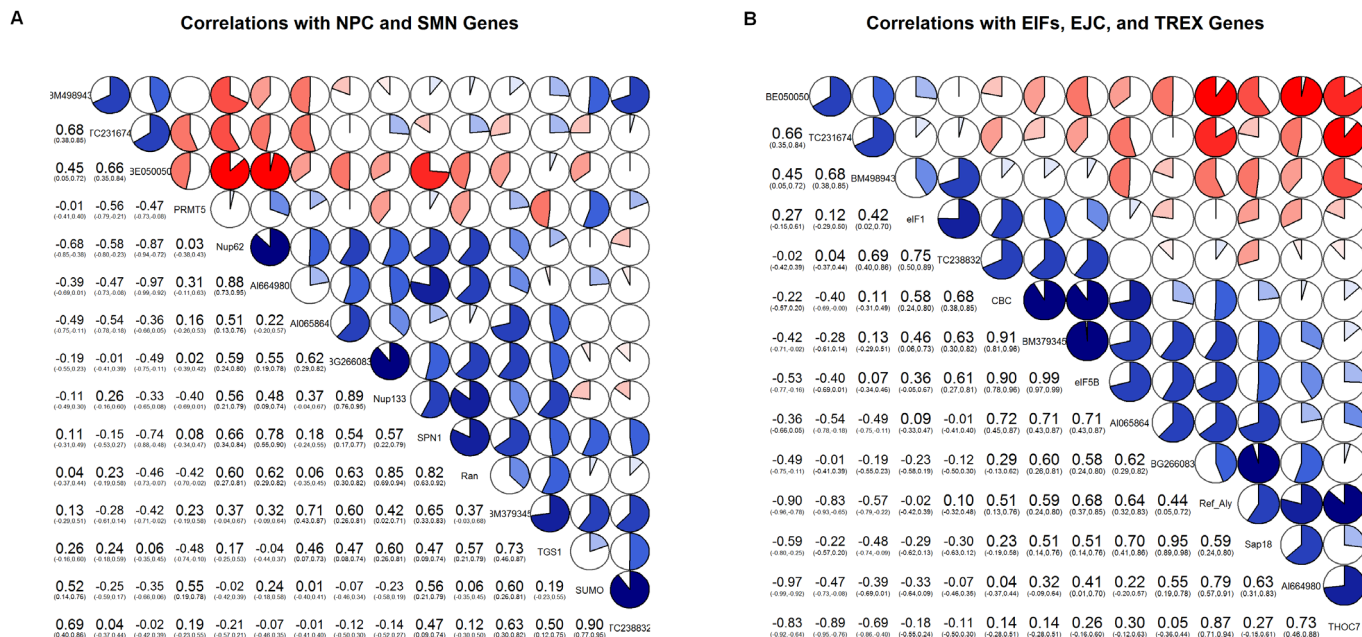
**A** Correlations with NPC and SMN Genes

**B** Correlations with EIFs, EJC, and TREX Genes



**Figure 2 | Correlogram displays of correlation matrices for gene expression data.** 2(A) Pearson's coefficients in genes from the NPC, SMN, RES, and SUS groups. 2(B) Pearson's coefficients in genes from the EIFs, EJCs, TREX, RES, and SUS groups. Correlations between genes are displayed in the order that genes have similar expression patterns are grouped together. The pie graphs are filled in proportion to the Pearson's coefficient values, clockwise for positive correlations (in blue) and anti-clockwise for negative correlations (in red). The numbers are Pearson's coefficients with 95% confidence intervals.

of the groups (EIFs, EJC, NPC, SMN, and TREX) represented the genes from different protein complexes in the RNA transport pathways and two groups (RES and SUS) represented the candidate genes selected from previous studies which were included in this research. Figure 4A is a network constructed for all the 56 genes in this study. The connectivity threshold was set arbitrarily as the Euclidean distance value being 2 for an exploratory criterion. Twenty-four genes were connected in the
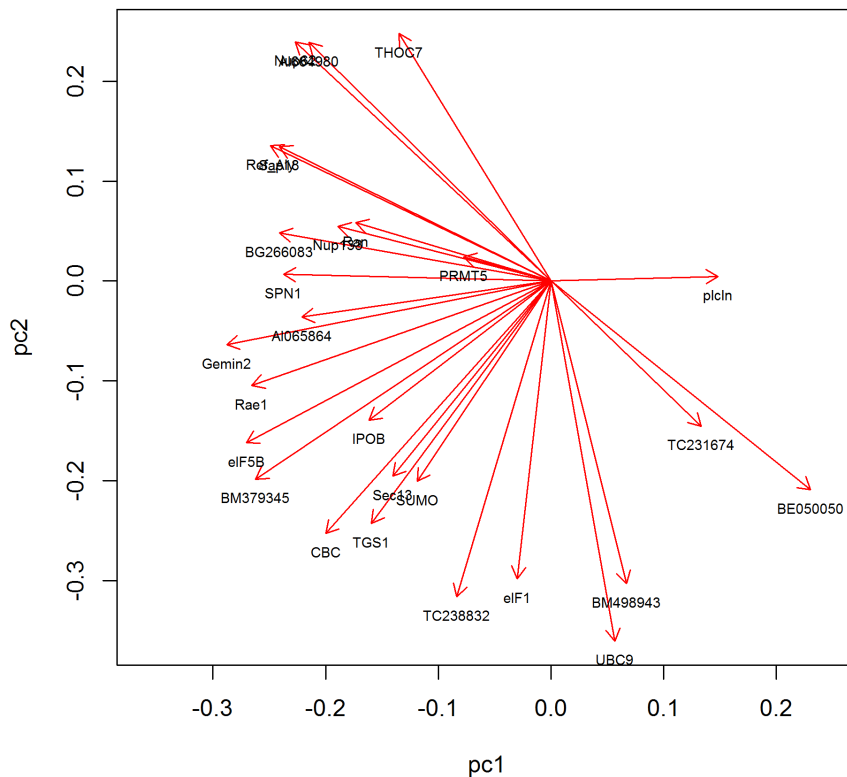


**Figure 3 | An eigenvector plot displaying the correlations on gene expression values among the significant candidate genes.** The length of each eigenvectors represents the largest variance in the Pearson's coefficients for each gene. The ordering of the eigenvectors is based on the distance between genes in terms of Pearson's coefficients. Gene eigenvectors close to each other are more positively correlated and hence the genes are more similar in the expression patterns.
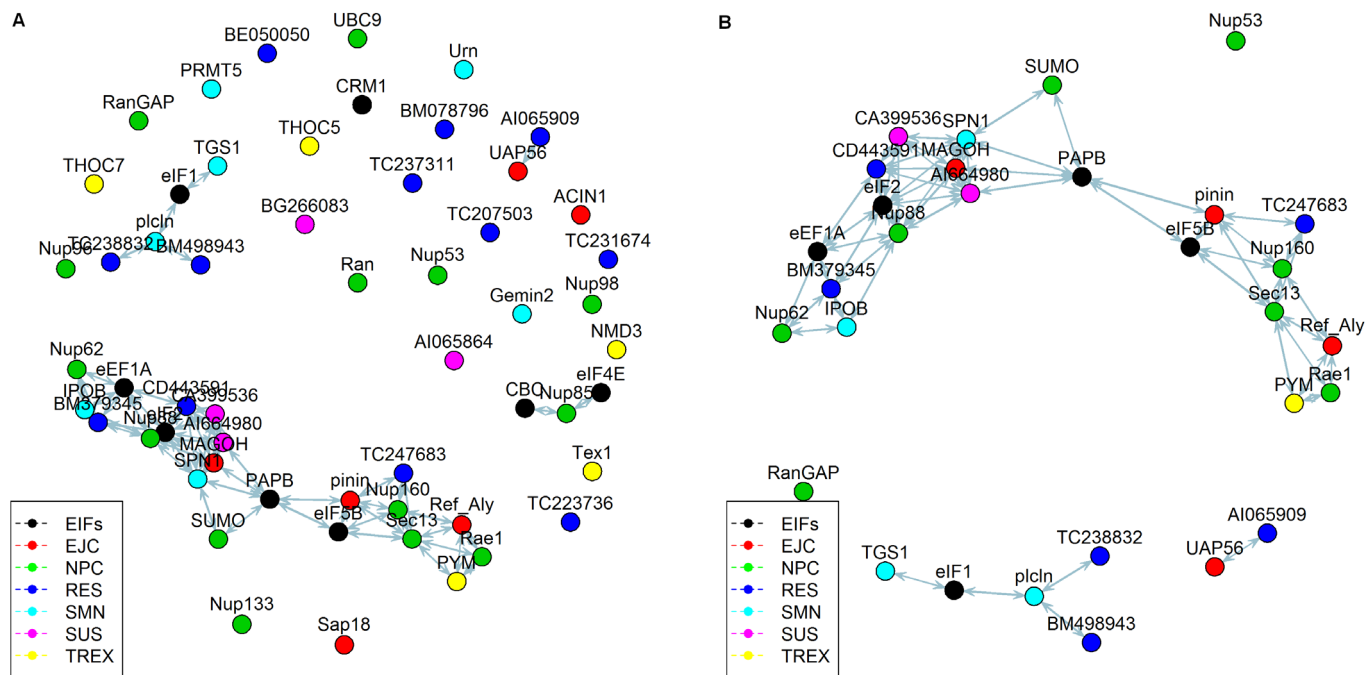
**Figure 4 | Network graphs showing the empirical relational structures revealed from the gene expression data.** 4(A) A network built at a threshold of the Euclidean distance < 2 on all the tested 56 genes. 4(B) A subgraph built at a threshold of the Euclidean distance < 1.6. The vertices were color-coded to highlight genes in seven different subgroups. Five of the subgroups (EIFs, EJC, NPC, SMN, and TREX) were from RNA transport pathways and two subgroups (RES and SUS) were selected from previous studies. The hubs with multiple connections indicate genes with important roles in the network of cellular functions.

network. Six resistance related genes, including TC231674 and BE050050, appeared as isolates in the network. Two susceptibility related genes BG266083 and AI065864 were not connected to any of the tested genes either. Further experiment with more genes will be required to reveal genes closely related in expression patterns with these genes in the empirical expression relational networks regarding maize defense. Figure 4B is a subgraph extracted from the same network dataset to show the genes that were connected in the network at a threshold of 1.6 in the Euclidean distance value. Seven genes (Nup88, eIF2, CD443591, CA399536, SPN1, AI664980, and MAGOH) had the highest vertex degrees and were clustered closely together, suggesting a co-expression pattern of these genes in response to *A. flavus* infecton among the tested maize inbred lines. Two other centers (hubs) were also revealed in the subgraph in Figure 4B. The resistance related gene BM379345 was found at the center of co-expression with five tested genes (IPOB, Nup62, eEF1A, eIF2, and Nup88). Nucleoporin genes Nup160, Sec13, and Rae1 were hubs for co-expression with genes PYM, Ref_Aly, eIF5B, pinin, and TC247683. The hubs with multiple connections were considered as genes of important roles in the network of cellular functions. The susceptibility related gene AI664980 was found adjacent to multiple RNA transport pathway genes including the Nup88, MAGOH, PAPB, and SPN1 genes and appeared to play an important role in the defense related nucleocytoplasmic trafficking activities based on the statistical inferences from network analysis. Maize genes at the centers in the network will be considered as important candidate genes and will be used in priority for further maize DNA marker studies.

## Discussion

Numerous studies have shown that the resistant maize inbred lines exhibited significantly low levels of aflatoxin accumulation. Determination of the mechanisms underlying such maize host resistance to aflatoxin accumulation has been proven difficult due to the complex nature of this quantitative trait. Many genes were found to be involved in the maize host plant resistance[3–5]. The exploration for methods to describe functional roles and relations of genes statistically using effective experimental designs and empirical gene expression data will expedite the discovery of DNA markers and uncover the mechanism of maize host resistance. In this study, we conducted RT-qPCR gene expression analysis on 56 genes including genes from RNA transport pathways that comprise the potential components of maize host resistance. The functions and relations of the genes were examined from three aspects: 1) statistical analysis (ANOVA) on gene expression data for the identification of differentially expressed genes and the determination of the significance levels; 2) correlation analysis for delineation of genes positively correlated or negatively correlated in response to *A. flavus* infection; and 3) network analysis for depiction of relations of genes in the empirical functional network. Significant genes related to maize defense to *A. flavus* infection were identified. Through the application of multidisciplinary methods, a wealth of data was generated for data mining and experimental validation. Evidence and supportive data have already been found through complementary research projects. For example, two differentially expressed genes, AI664980 and BG266083, which were found significant in the susceptible maize inbred line Va35 from previous reports, showed high significance again from this study in susceptible maize inbred lines (Mp04 : 85, Mp04 : 89, Va35). These genes are known to be involved in plant responses toward various stress and pathogens[32–34]. Statistical inferences drawn from our RT-qPCR gene expression analysis indicated that genes in RNA transport pathways, especially in NPC and SMN complexes, were highly significant and involved in maize resistance. One supporting example was that the resistance related gene TC231674 found previously in a different maize resistant inbred line Mp313E was highlighted again in the resistant maize inbred line Mp718. Interestingly, the highly expressed gene TC231674 found in two resistant inbred lines was homologous to the human nucleoporin Nup85 but it was not the same gene as the maize Nup85. The role of TC231674 gene in terms of interactions with maize nuclear pore complexes (NPCs) is yet to be determined.

Numerous additional examples can be used to show the richness of data yielded through the combination of the multidisciplinary methods in this study. One resistance related gene TC238832 was found positively correlated to the SUMO gene in the nuclear pore complex of the RNA transport pathway (Figure 2A). Figure 3 showed the relationship between these two genes according to the respective eigenvectors. The small angle between the two vectors represented a comparison of the Pearson's coefficients and a positive correlation between the TC238832 and SUMO genes. A similar but stronger relationship was noticed between the resistance related gene BM379345 and the eIF5B gene (Figure 2B), a translation initiation factor in the RNA transport pathway. The nuclear cap-binding protein subunit 2 gene (CBC gene) was another gene found positively correlated to both BM379345 and eIF5B. The positive correlations among CBC, eIF5B, and the BM379345 genes showed relationships that hinted at the possibility of these genes being related to resistance. By comparing these genes to the differentially expressed genes associated with susceptibility or resistance, we can gain new insights on the functions of the genes involved in the RNA transport and the plant defense mechanisms.

A network was constructed showing maize genes closely related in terms of the magnitudes and directions of their largest variances in the expression values among the resistant and susceptible maize inbred lines. Applying network-based methods to describe empirical gene expression data was an exploratory strategy we investigated to reveal genes potentially important in the regulation of host-fungus defense responses. It provided new strategies on prioritizing candidate genes. Information revealed by network analysis also provided more insights into the roles of highly expressed resistance related genes whose functions were yet to be characterized. While this study resulted in promising results, more research and analysis, such as testing of DNA markers associated with the resistance related genes, are required to verify the results and determine the mechanisms of maize host plant resistance to *Aspergillus flavus* infection and alfatoxin reduction.

# Methods

**Plant materials and experimental design.** Six maize inbred lines (Mp718, Mp719, Mp04 : 104, Mp04 : 85, Mp04 : 89, and Va35) were used in this experiment. Five of them (Mp718, Mp719, Mp04 : 104, Mp04 : 85, and Mp04 : 89) were recombinant maize inbred lines obtained by eight generation selfing from $F_2$ plants of a cross of Mp715 × Va35 and were selected against aflatoxin accumulation under *Aspergillus flavus* inoculation in field conditions. The maize inbred line seeds were maintained by the United States Department of Agriculture, Agricultural Research Service, Corn Host Plant Resistance Research Unit (USDA-ARS-CHPRRU) at Mississippi State University. Mp718, Mp719, and Mp04 : 104 were maize inbred lines showing resistance to *Aspergillus flavus* infection and aflatoxin accumulation. Mp04 : 85, Mp04 : 89, and Va35 were susceptible to *Aspergillus flavus*. All maize lines were planted at the R. R. Foil Plant Science Farm at Mississippi State University. The experimental design was a randomized complete block design including three replications and two treatments (inoculated and un-inoculated with *A. flavus*) for each maize inbred line and two sample collection time points (2 and 7 days after inoculation). All primary ears were self-pollinated. Fourteen days after pollination, the inoculation of *A. flavus* was performed using the *A. flavus* strain NRRL 3357 (ATCC # 200026; SRRC 167). The procedure of fungal culture preparation and the fungal inoculation with side-needle technique were the same as described previously[35]. Two and seven days after inoculation, which was 16 and 21 days after self-pollination, developing kernels from inoculated and uninoculated primary ears were collected for RNA preparation. All remaining primary ears from each plot were harvested at maturity and processed for measurement of aflatoxin concentrations as previously described[35-38].

**RNA extraction.** Developing kernels were collected from the resistant maize inbred lines (Mp718, Mp719, Mp04 : 104) and the susceptible maize inbred lines (Va35, Mp04 : 85, and Mp04 : 89), flash frozen in liquid nitrogen in the field, and stored at −80°C for further analysis. Total RNAs were isolated from the kernels using the BioRad Aurum™ Total RNA Fatty and Fibrous Tissue kit. Frozen kernels were ground into powder under liquid nitrogen and combined with PureZOL for disruption. Chloroform was added to the sample for extraction of the aqueous phase containing the RNA. The sample was subjected to DNase I treatment and followed by a series of washes and centrifugation steps with solutions provided with the kit. Upon completion, total RNA concentrations were determined using a NanoDrop® ND-1000 Spectrophotometer. The Quality control assessments of total RNA was performed with an Agilent 2100 Bioanalyzer. RNA samples with a QC value RIN >8 were used for cDNA synthesis.

**Quantitative real time RT-PCR.** ThermoScript RT-PCR system (Invitrogen, #11146-024) was used for cDNA synthesis. RNA was combined with Oligo(dT) primer, 10 mM dNTP Mix, and DEPC-treated water and incubated for 5 minutes at 65°C. A master mix was created using 5× cDNA Synthesis Buffer, 0.1 M DTT, RNaseOUT™, DEPC-treated water, and ThermoScript™ RT, and added to the reaction mixture. This mixture was incubated at 50°C for 45 minutes for the completion of successful cDNA synthesis. RT-qPCR analysis was conducted using a Roche LightCycler 480 instrument (Roche Applied Science) with the standard 96-well block. LightCycler 480 SYBR Green I Master kit (Roche Applied Science, #04 707 516 001) was used for the RT-qPCR reactions. Fifty genes were selected from the RNA transport pathways, and primers were designed using the Primer3 software[39,40] (Table 2). Sixteen previously identified candidate genes were also included in this study and the sequences of these primers (Table 2) were the same as described before[5]. The housekeeping gene, *Zea mays* glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as the reference gene in this study. The choice of the reference gene has been determined in a previous study[5]. Standard curve assays were performed for each pair of primers by creating a three-fold dilution scheme for each sample to calculate the PCR efficiency. A total of 72 cDNA samples including six maize inbred lines(pedigree), three replications(rep), two treatments(inoc), and two time points(DAI) were loaded onto each 96-well PCR plates for whole plate assays with one plate for each tested gene. A negative check with a ddH2O sample was included on each plate. The RT-qPCR program was as the following: 1) 1 cycle of 95°C for 5 min; 2) 45 cycles of 95°C for 10 sec, 60°C for 15 sec, 72°C for 15 sec; 3) 1 cycle of 95°C for 5 sec, 65°C for 1 min, 97°C at continuous; 4) 1 cycle of 40°C for 10 sec. The mixture used for the RT-qPCR reactions was as follows: 0.5 μl forward primer (10 uM), 0.5 μl reverse primer (10 uM), 3 μl SYBR Green I Master Kit Enzyme mix, 5 μl DEPC-treated water, and 1 μl of cDNA. This working recipe was adapted from the manufacturer's manual with the following changes: The recipe to prepare for a 10 μl reaction from the manufacturer's manual was 2.5 μl DNA template, 5 μl Master Mix, 1 μl PCR primer, and 1.5 μl water. Our working recipe was 1 μl cDNA template, 3 μl Master Mix, 1 μl PCR primer, and 5 μl water. The ratio of DNA template to Master Mix by volume from the manufacturer's manual was 1 : 2. In our recipe, the ratio of cDNA template to Master Mix was 1 : 3. The ratio of Master Mix (ready to use kit mixture containing Taq DNA polymerase, dNTP mix, SYBR Green I dye, and MgCl2) to the DNA template was increased in our recipe. The PCR efficiencies were in the optimal range (1.9–2.0) for PCR (Table 2).

**Preprocessing of raw RT-qPCR data.** The R statistical programming language[41] was used to develop scripts for both the preprocessing of raw RT-qPCR data and the subsequent ANOVA analysis in this study. To enable the high-throughput processing and analysis of RT-qPCR data, we developed R scripts specific for the acquisition and preprocessing of raw RT-qPCR data from Roche output files which were in tab-separated plain text format. The preprocessing of raw data included the following steps: 1) The RT-qPCR cycle threshold values (designated as CP value in output files from Roche instruments and represented in this manuscript as $C_q$ in line with the MIQE guidelines[42]) were batch-extracted from all output data files for the whole plate assays and standard curve assays, 2) R-squared values from the linear regression analysis of the standard curve data for each gene were calculated, 3) The r-squared values were used as the PCR efficiencies for calculation of the gene expression Cq values, 4) The Cq values of the reference gene were subtracted from the Cq values of the targeted genes for the normalization with the reference gene, and 5) the maximum Cq value of each gene was subtracted from all the Cq values for that gene to get the relative delta Cq values which were used as gene expression values for the subsequent statistical analysis.

**Statistical analysis of RT-qPCR data.** Statistical analysis was performed with R scripts. Boxplots were used to visualize the summary of the descriptive statistics for gene expression data, including the median, spread and outliers. Scatterplots were used to display the gene expression data for each gene over 72 samples. Analysis of variance (ANOVA) [http://www.statmethods.net/stats/anova.html] was used for the test of null hypothesis on gene expression data among different treatments including resistance versus susceptibility (res), pedigree, inoculation status (inoc), res by inoc, and pedigree by inoc in order to identify the sources of variation. The test of null hypothesis was based on the F-ratio which was the Mean Square between the groups (MSG) to Mean Square within the group (MSE). The significance level was determined at p < 0.05.

The ANOVA analysis was performed using the corresponding functions in R as listed in the following:

```
# One Way Anova (Completely Randomized Design)
fit <- aov(y ~ A, data = mydataframe)
# Randomized Block Design (B is the blocking factor)
fit <- aov(y ~ A + B, data = mydataframe)
# Two Way Factorial Design
fit <- aov(y ~ A + B + A:B, data = mydataframe)
fit <- aov(y ~ A*B, data = mydataframe)
[http://www.statmethods.net/stats/anova.html]
```

**Correlation analysis and the illustration of Pearson's coefficients.** Pearson's correlation coefficients were calculated on the gene expression data between all pairs of genes for the correlation analysis. The R package "Corrgram" was used to display the correlations between the selected pairs of genes either by using a "correlogram" or a coefficient eigenvector plot[43]. A correlogram was a direct visual display of the matrix of Pearson's coefficients that were calculated from the gene expression data. By this method, correlations between genes were displayed by grouping genes that have similar expression patterns, and the values and signs of the correlations were visualized schematically in numbers and color-coded pie graphs. The pie graphs were filled in proportion to the Pearson's coefficient values, clockwise for positive correlations (in blue) and counter clockwise for negative correlations (in red). The eigenvector plot showed the ordering of variables according to the angles formed by the two gene eigenvectors. The magnitude (length) of the eigenvectors represents the largest variance in the coefficient values for each gene. The angles between the eigenvectors represents the closeness between genes in terms of expression correlations. Gene eigenvectors placed close to each other are more similar in the expression patterns and hence are more positively correlated.

**Network analysis.** Network analysis was performed following the manual of the R packages "sna" and "network" [Butts, C.T. Statnet Project http://statnetproject.org]. Principal component analysis was performed on the mean values (by replication) of gene expression data. The scores of the first two principal components (designated as pc1 and pc2, respectively) were used to calculate the Euclidean distance values between all pairs of genes. The resulting Euclidean distance matrix was used to construct the network. Network graphs were generated to display the empirical relations and roles among the RNA transport pathway genes and other candidate genes. The vertices were color-coded to represent different gene groups. The genes at the centers (namely hubs) and the genes connecting the centers in the network were considered as potentially important candidate genes for future DNA marker studies.

1. Payne, G. A. Aflatoxins in maize. *Critical Rev Plant Sci.* **10**, 423–440 (1992).
2. Gourma, H. & Bullerman, L. B. Aspergillus flavus and Aspergillus parasiticus: Aflatoxigenic fungi of concern in foods and feeds. *J Prot Ecol.* **58**, 1395–1404 (1995).
3. Brooks, T. D., Williams, W. P., Windham, G. L., Willcox, M. C. & Abbas, H. K. Quantitative trait loci contributing resistance to aflatoxin accumulation in the maize inbred Mp313E. *Crop Sci.* **45**, 171–174 (2005).
4. Williams, W. P., Windham, G. L. & Buckley, P. M. [Enhancing maize with resistance to Aspergillus flavus infection and aflatoxin accumulation] *Aflatoxin and Food Safety* [379–394] (CRC Press, New York, 2005).
5. Kelley, R. Y. *et al.* Identification of maize genes associated with host plant resistance or susceptibility to Aspergillus flavus infection and aflatoxin accumulation. *PLoS ONE* **5**, e36892 (2012).
6. Moore, J. W., Loake, G. J. & Spoel, S. H. Transcription dynamics in plant immunity. *The Plant Cell* **23**, 2809–2820 (2011).
7. Akira, S. & Takeda, K. Toll-like receptor signalling. *Nat Rev Immunol.* **4**, 499–511 (2004).
8. Ali, G. S., Prasad, K. V., Day, I. & Reddy, A. S. Ligand-dependent reduction in the membrane mobility of FLAGELLIN SENSITIVE2, an Arabidopsis receptor-like kinase. *Plant Cell Physiol.* **48**, 1601–11 (2007).
9. Thomma, B. P., Nurnberger, T. & Joosten, M. H. Of PAMPs and effectors: The blurred PTI-ETI dichotomy. *The plant cell* **23**, 4–15 (2011).
10. Tsuda, K., Sato, M., Stoddard, T., Glazebrook, J. & Katagiri, F. Network properties of robust immunity in plants. *PLoS Genetics* **5**, e1000772 (2009).
11. Shah, J. & Zeier, J. Long-distance communication and signal amplification in systemic acquired resistance. *Frontiers in Plant Sci.* **4**, 30 (2013).
12. Marin, M., Uversky, V. N. & Otta, T. Intrinsic Disorder in Pathogen Effectors: Protein Flexibility as an Evolutionary Hallmark in a Molecular Arms Race. *The Plant Cell* **25**, 3153–3157 (2013).
13. Piffanelli, P., Devoto, A. & Schulze-Lefert, P. Defence signaling pathways in cereals. *Curr Opin Plant Biol* **2**, 295–300 (1999).
14. Chisholm, S. T., Coaker, G., Day, B. & Staskawicz, B. J. Host-microbe interactions: Shaping the evolution of the plant immune response. *Cell.* **124**, 803–814 (2006).
15. Walley, J. W. *et al.* Mechanical stress induces biotic and abiotic stress responses via a novel cis-element. *PLoS Genetics* **3**, e172 (2007).
16. Nakielny, S., Fischer, U., Michael, W. M. & Dreyfuss, G. RNA transport. *An. Rev Neurosci* **20**, 269–301 (1997).
17. Kindler, S., Wang, H. D., Richter, D. & Tiedge, H. RNA transport and local control of translation. *Annu Rev Cell Dev Biol.* **21**, 223–245 (2005).
18. Jambhekar, A. & Derist, J. L. Cis-acting determinants of asymmetric, cytoplasmic RNA transport. *RNA.* **13**, 625–642 (2007).
19. Vazquez-Pianzola, P. & Suter, B. Conservation of the RNA transport machineries and their coupling to translation control across eukaryotes comparative and functional genomics. *Compar Funct genom.* **2012**, 287852 (2012).
20. Köhler, A. & Hurt, E. Exporting RNA from the nucleus to the cytoplasm. *Nature Reviews Mol Cell Biol.* **8**, 761–773 (2007).
21. Meier, I. & Brkljacic, J. The nuclear pore and plant development. *Cur Opin Plant Biol.* **12**, 87–95 (2009).
22. Deslandes, L. & Rivas, S. The plant cell nucleus, a true arena for the fight between plants and pathogens. *Plant Signal & Behav.* **6**, 42–48 (2011).
23. Rivas, S. Nuclear dynamics during plant innate immunity. *Plant Physiol.* **158**, 87–94 (2012).
24. Cheng, Y. T. *et al.* Nuclear pore complex component MOS7/Nup88 is required for innate immunity and nuclear accumulation of defense regulators in Arabidopsis. *Plant Cell.* **21**, 2503–2516 (2009).
25. Garcia, A. V. & Parker, J. E. Heaven's gate: nuclear accessibility and activities of plant immune regulators. *Trends in Plant Sci* **14**, 479–487 (2009).
26. Faris, J. D. *et al.* A unique wheat disease resistance-like gene governs effector-triggered susceptibility to necrotrophic pathogens. *Proc Natl Acad Sci USA.* **107**, 13544–13549 (2010).
27. Heid, C. A., Stevens, J., Livak, K. J. & Williams, M. P. Real time quantitative PCR. *Genome Res.* **6**, 986–994 (1996).
28. Kubista, M. *et al.* The real-time polymerase chain reaction. *Mol Aspects of Med.* **27**, 95–125 (2006).
29. Rieu, I. & Powers, S. J. Real-time quantitative RT-PCR: Design, Calculations, and Statistics. *Plant Cell* **21**, 1031–1033 (2009).
30. Williams, W. P. Breeding for resistance to aflatoxin accumulation in maize. *Mycotoxin Res* **22**, 27–32 (2006).
31. Williams, W. P. & Windham, G. L. Registration of Mp718 and Mp719 Germplasm lines of Maize. *J Plant Registr.* **6**, No. 2 (2012).
32. Naqvi, S. M. S. *et al.* A glycine-rich RNA-binding protein gene is differentially expressed during acute hypersensitive response following Tobacco Mosaic Virus infection in tobacco. *Plant Mol Biol* **37**, 571–576 (1998).
33. Govrin, E. M. & Levine, A. The hypersensitive response facilitates plant infection by the necrotrophic pathogen Botrytis cinerea. *Curr Biol.* **10**, 751–757 (2000).
34. Singh, U., Deb, D., Singh, A. & Grover, A. Glycine-rich RNA binding protein of Oryza sativa inhibits growth of M15 E. coli cells. *BMC res notes.* **4**, 18 (2011).
35. Windham, G. L. & Williams, W. P. Aspergillus flavus infection and accumulation in resistant and susceptible maize hybrids. *Plant Dis* **82**, 281–284 (1998).
36. Windham, G. L. & Williams, W. P. Evaluation of corn inbreds and advanced breeding lines for resistance to aflatoxin contamination in the field. *Plant Dis* **86**, 232–234 (2002).
37. Windham, G. L., Williams, W. P., Buckley, P. M., Abbas, H. K. & Hawkins, L. K. [Techniques used to identify alfatoxin-resistant corn] *Aflatoxin and Food Safety* [379–394] (CRC Press, New York, 2005).
38. Zummo, N. & Scott, G. E. Evaluation of field inoulation teqhniques for screening maize genotypes against kernel infection by Aspergillus flavus in Mississippi. *Plant Dis* **73**, 313–316 (1989).
39. Koressaar, T. & Remm, M. Enhancements and modifications of primer design program Primer3. *Bioinformatics.* **23**, 1289–1291 (2007).
40. Untergrasser, A. *et al.* Primer3 - new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).
41. R Development Core Team. [R: A language and environment for statistical computing] *R Foundation for Statistical Computing* (Vienna, Austria, 2008).
42. Bustin, S. A. *et al.* The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin Chem.* **55**, 611–622 (2009).
43. Friendly, M. Corrgrams: Exploratory Displays for Correlation Matrices. *Am Stat.* **56**, 316–324 (2002).

## Acknowledgments

## Author contributions

X.S. and W.P.W. conceived and designed the experiments. M.C.A., X.S., W.P.W., J.E.M. and G.L.W. performed the experiments. X.S., M.C.A. and A.D.P. analyzed the data. M.C.A. and X.S. wrote the manuscript. W.P.W., A.D.P., J.E.M. and G.L.W. contributed to and edited the manuscript. All authors have reviewed and approved the final manuscript.

## Additional information