

# SCIENTIFIC REPORTS



OPEN

## Hemolymph protein profiles of subterranean termite *Reticulitermes flavipes* challenged with methicillin resistant *Staphylococcus aureus* or *Pseudomonas aeruginosa*

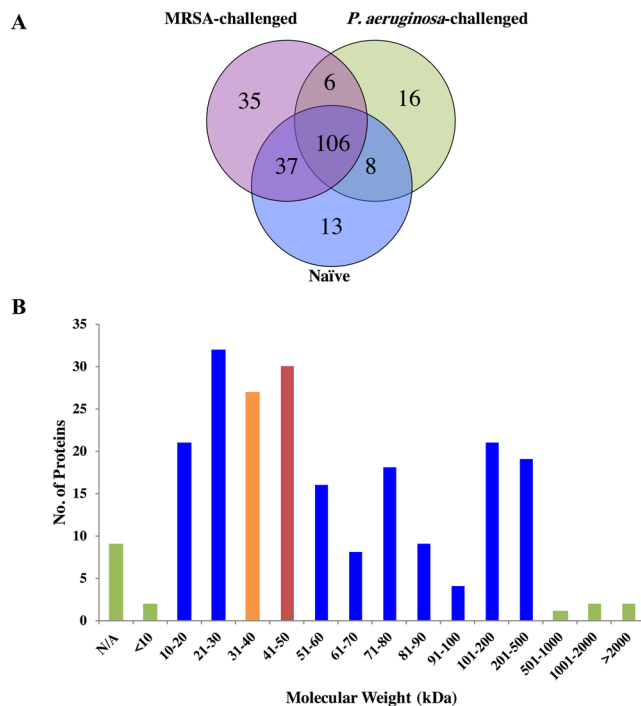
Yuan Zeng<sup>1,4</sup>, Xing Ping Hu<sup>1</sup>, Guanqun Cao<sup>2</sup> & Sang-Jin Suh<sup>3</sup>

When the subterranean termite *Reticulitermes flavipes* is fed heat-killed methicillin resistant *Staphylococcus aureus* (MRSA) or *Pseudomonas aeruginosa*, the termite produces proteins with antibacterial activity against the inducer pathogen in its hemolymph. We used a proteomic approach to characterize the alterations in protein profiles caused by the inducer bacterium in the hemolymph of the termite. Nano-liquid chromatography-tandem mass spectrometry analysis identified a total of 221 proteins and approximately 70% of these proteins could be associated with biological processes and molecular functions. Challenges with these human pathogens induced a total of 57 proteins (35 in MRSA-challenged, 16 in *P. aeruginosa*-challenged, and 6 shared by both treatments) and suppressed 13 proteins by both pathogens. Quasi-Poisson likelihood modeling with false discovery rate adjustment identified a total of 18 and 40 proteins that were differentially expressed at least 2.5-fold in response to MRSA and *P. aeruginosa*-challenge, respectively. We selected 7 differentially expressed proteins and verified their gene expression levels via quantitative real-time RT-PCR. Our findings provide an initial insight into a putative termite immune response against MRSA and *P. aeruginosa*-challenge.

Insect hemolymph plays key roles in insect innate immunity<sup>1</sup>. Although many of the hemolymph components have yet to be characterized, some of the proteins have been identified and their functions elucidated. For example, hexamerins and arylphorins have been determined to be the source of amino acids for components of insect cuticles; lipophorins and other related enzymes (e.g. esterases, lipases) function in lipid transportation and hydrolysis; vitellogenins play an important role in embryo development; and cytokines are involved in intercellular communications. In addition, the hemolymph contains components of the insect immune system, including effector molecules such as cytokines, enzyme cascades, and antioxidant proteins, demonstrating that it is vital for defense against pathogens or tissue damage<sup>2-4</sup>.

Proteomics is an important tool for studying changes in insect hemolymph proteome during development and in response to environmental effects<sup>5-7</sup>. Previous studies have demonstrated changes in the hemolymph proteomes of fruit fly, silkworm, white butterfly, and tobacco hornworm in response to immune challenges<sup>6,8-10</sup>. In termites, Liu *et al.*<sup>11</sup> identified differential expression of 20 hemolymph proteins and upregulated expression of 3 immune genes in the subterranean termite *Reticulitermes chinensis* when challenged with an entomopathogenic fungus *Metarhizium anisopliae*. However, very little is known about the effect of bacterial challenge on the termite hemolymph proteome.

<sup>1</sup>Department of Entomology and Plant Pathology, Auburn University, Auburn, AL, USA. <sup>2</sup>Department of Mathematics and Statistics, Auburn University, Auburn, AL, USA. <sup>3</sup>Department of Biological Sciences, Auburn University, Auburn, AL, USA. <sup>4</sup>Present address: Department of Bioagricultural Sciences and Pest Management, Colorado State University, Fort Collins, CO, USA. Correspondence and requests for materials should be addressed to S.-J.S. (email: [suhsang@auburn.edu](mailto:suhsang@auburn.edu))



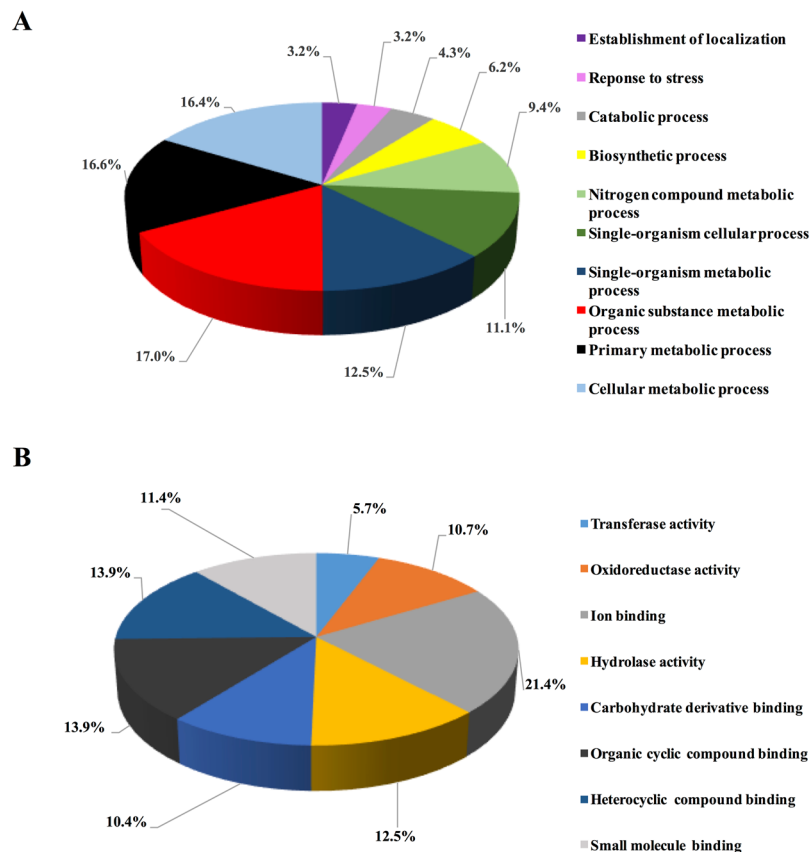
**Figure 1.** Summary of the identified hemolymph proteins in bacterial pathogen-challenged and naïve *R. flavipes* workers. (A) Comparison of proteins identified in naïve and bacterial pathogen-challenged *R. flavipes*. (B) Size distribution of proteins.

The eastern subterranean termite *Reticulitermes flavipes* (Kollar) is widely distributed in the United States and causes serious damage to structures and plants to result in significant economic losses<sup>12</sup>. Subterranean termites have evolved various defense strategies against pathogens as they nest and forage in microbe-rich soil<sup>13</sup>. We have previously reported the existence of a broad-spectrum of constitutive activity against a variety of Gram-positive and Gram-negative bacteria, including human pathogens, in the hemolymph of *R. flavipes*<sup>14,15</sup>. Interestingly, although the naïve termite hemolymph was active against the human pathogen *S. aureus*, it was inactive against MRSA, the methicillin resistant derivative of *S. aureus*, that has emerged as one of the most dangerous Gram-positive pathogens. The hemolymph from naïve termite was also inactive against *Pseudomonas aeruginosa*, one of the most dangerous Gram-negative nosocomial pathogens. However, the termite developed inducible anti-MRSA or anti-*P. aeruginosa* activity in its hemolymph when it was fed heat-killed MRSA or *P. aeruginosa*, respectively<sup>14</sup>. The induction of anti-MRSA or anti-*P. aeruginosa* activity affected the overall antimicrobial spectrum of the *R. flavipes* hemolymph. MRSA-challenged termite hemolymph gained anti-MRSA activity while maintaining activity against two other Gram-positive bacteria, including *Streptococcus pyogenes*, but completely lost activity against three Gram-negative bacteria tested. *P. aeruginosa* and anti-MRSA activity while maintaining activity against other bacteria tested. As an initial step towards elucidating the inducible antimicrobial defense mechanisms of *R. flavipes*, we determined the hemolymph protein profiles of termites challenged with either MRSA or *P. aeruginosa*.

## Results

**Analysis of hemolymph proteome.** NCBI protein database search showed that 22,338 spectra matched those of trypsin digested termite hemolymph proteins at  $\geq 98\%$  probabilities to achieve 0.10% Decoy False Discovery Rate (FDR). The matching spectra corresponded to 181 clusters containing a total of 221 proteins with at least two peptide matches at  $\geq 95\%$  confidence. Of the 221 proteins, 106 proteins were present in all three termite groups, 37 proteins were present in the naïve and the MRSA-challenged termites, 8 proteins were present in the naïve and the *P. aeruginosa*-challenged termites, and 6 proteins were present only in the two bacteria-challenged insects. In summary, two bacteria induced the production of 57 proteins but reduced the production of 13 proteins. Of the 57 induced proteins, 35 were detected only in the MRSA-challenged, 16 in the *P. aeruginosa*-challenged, and 6 in both bacteria-challenged termite groups, respectively (Fig. 1A). Approximately 68% of the proteins had molecular weight (MW) between 10 to 80 kDa (Fig. 1B). The smallest protein was 8 kDa (antioxidant enzyme; AGM32333.1) and the largest protein was 2,068 kDa (uncharacterized protein; XP\_014096235.1).

Blast2Go analysis associated many of the hemolymph protein sequences (70%) with a variety of biological processes (GO:0008150) and molecular functions (GO:0003674). The identified proteins were annotated into 10 categories based on biological processes (Fig. 2A) or 8 categories based on molecular functions (Fig. 2B). A comparison among relative abundances of sub-levels of annotated categories on biological process showed that hemolymph proteins that participate in organonitrogen compound metabolism (GO:1901564), macromolecule metabolism (GO:0043170), carbohydrate metabolism (GO:0005975), phosphorus metabolism (GO:0006793),



**Figure 2.** Protein categorization by gene ontology (GO). **(A)** GO based on biological processes. **(B)** GO based on molecular functions. The percentages indicate relative abundances.

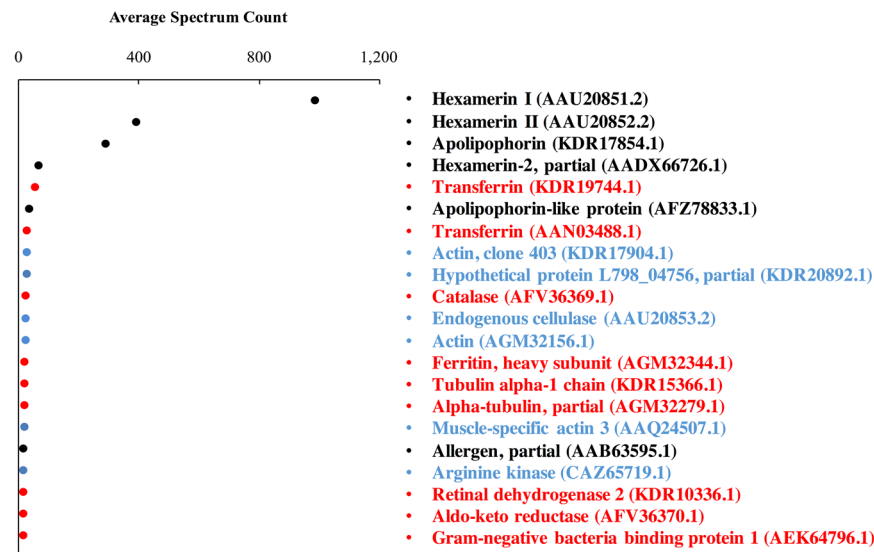
redox reactions (GO:0055114), small molecule metabolism (GO:0044281), and transport (GO:0006810) had a higher abundance. A comparison among relative abundances of sub-levels of annotated molecular functions showed termite hemolymph proteins involved in nucleotide binding (GO:0000166), purine nucleoside binding (GO:0001883), ribonucleoside binding (GO:0032549), and metal ion binding (GO:0046872) had a higher abundance than those annotated with catalytic activities.

According to the total spectral count, the 21 most highly expressed hemolymph proteins were divided into three categories of 6 storage proteins (hexamerin I, hexamerin II, hexamerin-2, apolipoporphins, apolipoporphin-like protein, and allergen), 9 immune-related proteins (2 transferrins, catalase, ferritin, 2 alpha-tubulin, retinal dehydrogenase 2, aldo-keto reductase, and gram-negative bacteria binding protein), and 6 other proteins (3 actins, hypothetical protein L798\_04756, endogenous cellulase, and arginine kinase) (Fig. 3).

**Effect of bacterial pathogen-challenge on hemolymph proteome.** *Alterations in the hemolymph proteome following MRSA-challenge.* MRSA challenge induced production of 41 proteins and reduced production of 21 proteins (Fig. 1A). Of these, 4 proteins were induced at least 2.5-fold and 14 proteins were reduced at least 2.5-fold (Table 1). Of the induced (upregulated) proteins, transferrin, has been associated with antimicrobial activity as a part of the immune response. Among the reduced (downregulated) proteins, 6 are associated with immune response (i.e., response to stress, cytoskeletal modeling, detoxification, and immune effectors). These molecules include beta-glucosidase, papilin, c-type lysozyme, apolipoporphin, peroxiredoxin-6, and cathepsin L-like protein. Thus, exposure to heat-killed MRSA up-regulated one but down-regulated 6 known immune proteins in *R. flavipes*.

Hierarchical clustering based on the expression profile of the 18 discriminatory proteins (4 upregulated and 14 downregulated) separated samples into two main clusters corresponding to the MRSA-challenge (Fig. 4A). Similarity in protein expression profiles clustering the 6 samples was summarized in a biplot by the principle component analysis (PCA) (Fig. 4B). Both principal components (PC1 and PC2) explained a total variation of 95.2%. The first principal component (PC1) differentiated MRSA-challenged and naïve termite hemolymph samples, while the second principal component (PC2) was a measure of antibacterial peptide and detoxification enzyme (e.g. c-type lysozyme 2, peroxiredoxin-6) across all 18 discriminatory proteins.

*Alterations in the hemolymph proteome following P. aeruginosa-challenge.* The *P. aeruginosa*-challenge resulted in differential accumulation of 40 proteins with at least 2.5-fold changes (Table 2). These proteins were annotated to metabolism, development, stress response, immune signaling, immune effectors and other functions. Of the



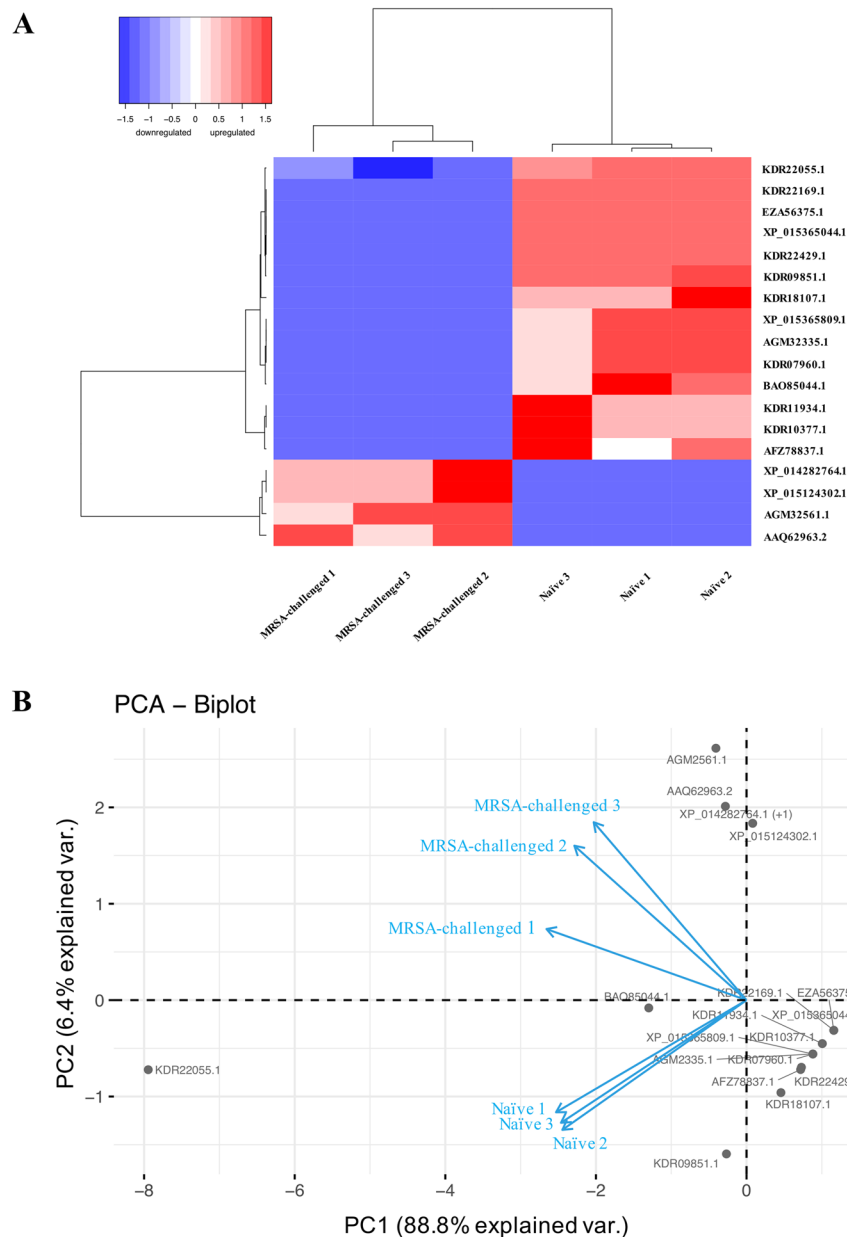
**Figure 3.** Twenty-one most abundant proteins in three *R. flavipes* hemolymph samples. The abundance value of each protein was estimated as average spectrum count. Colors show the protein category: storage protein, black; immune-related protein, red; other proteins, blue.

Protein	Accession Number	Poisson.FDR $p$ -value	Quasi.FDR $p$ -value	Rate Ratio*
transferrin	AAQ62963.2	0.0909	0.0193	32.91
calponin-like domain containing protein	AGM32561.1	0.0909	0.0193	32.91
predicted: myosin heavy chain, muscle isoform X26	XP_014282764.1	0.1375	0.0308	32.59
predicted: ATP synthase subunit alpha	XP_015124302.1	0.1375	0.0308	32.59
protein yellow	KDR22429.1	0.0696	<0.0001	-34.62
multifunctional protein ADE2, partial	KDR09851.1	0.0018	0.0008	-34.29
predicted: alpha, alpha-trehalose-phosphate synthase [UDP-forming]-like isoform X1	XP_015365044.1	0.1938	<0.0001	-33.62
cytosolic carboxypeptidase-like protein 5, partial	KDR22169.1	0.1938	<0.0001	-33.62
multifunctional protein ADE2	EZA56375.1	0.1938	<0.0001	-33.62
predicted: bifunctional purine biosynthesis protein PURH	XP_015365809.1	0.0909	0.0193	-32.91
hypothetical protein L798_01615	KDR07960.1	0.0909	0.0193	-32.91
cathepsin L-like protein	AGM32335.1	0.0909	0.0193	-32.91
lambda-crystallin-like protein	KDR11934.1	0.1375	0.0308	-32.59
peroxiredoxin-6	KDR10377.1	0.1375	0.0308	-32.59
apolipoprotein	KDR18107.1	0.0272	0.0308	-32.15
c-type lysozyme-2	AFZ78837.1	0.0696	0.0469	-31.73
beta-glucosidase	BAO85044.1	0.1375	0.0308	-2.00
papilin	KDR22055.1	0.0049	0.0308	-1.79

**Table 1.** Differentially expressed hemolymph proteins of *R. flavipes* following MRSA-challenge. \*Rate ratio:  $-1.32 \leq \log_2(\text{rate1}/\text{rate2}) \leq 1.32$ .

4 upregulated proteins, 2 immune response related proteins (actin<sup>16,17</sup> and alpha-tubulin, partial<sup>18,19</sup>) levels were increased at least 2.5-fold. In contrast, *P. aeruginosa*-challenge downregulated 36 proteins and 24 of them completely disappeared from the proteome. Most of the downregulated proteins are involved in metabolic processes and stress response. Of the downregulated proteins, 9 have been reported as immune-related proteins including ferritin, alcohol dehydrogenase, glutathione S-transferase, aldo-keto reductase, regucalcin, c-type lysozyme, Cu/Zn superoxide dismutase, hemocytin, and prostaglandin reductase<sup>20–25</sup>.

Hierarchical clustering (Fig. 5A) of these 40 discriminatory proteins in *P. aeruginosa*-challenged and naïve termites essentially separated samples into two main clusters. PCA analysis by the first two principal components explained 96.3% of the variations (Fig. 5B). The first principal component (PC1) differentiated *P. aeruginosa*-challenged and naïve termite hemolymph samples, and the second principal component (PC2) is a measure of enzymes involved in metabolism (glutamine synthetase 2)<sup>26</sup>, detoxification (glutathione S-transferase)<sup>20</sup>, cell migration (filamin A)<sup>27</sup>, and antibacterial defenses (c-type lysozyme 2)<sup>28</sup>.



**Figure 4.** Hierarchical clustering analysis (A) and principal component analysis (B) based on 18 proteins significantly changed in abundances between MRSA-challenged and naïve termites within the dataset. (A) Both samples and proteins were clustered using Ward's method, and with Pearson correlation as similarity metric. The samples are shown horizontally (columns) and the proteins are shown vertically (rows). The dendrograms represent the distances between clusters. Protein expression levels are represented in the color scale of blue (downregulated) to red (upregulated). (B) Grey dots and blue arrows represent hemolymph proteins and variables, respectively. Each axis represents a principal component (PC1 and PC2) with the percentage of the total variance it explains.

**RT-qPCR verification of selected protein levels.** Alterations in protein levels can be due to various factors including gene expression and changes in protein turn-over. We verified our proteome data by determining the transcript levels of seven selected genes via RT-qPCR (Supplementary Table 1) with the gene encoding alpha-tubulin 2 as the internal control. These genes were selected because they were expected to be involved in the insect's immune responses or hormone regulation. As determined from the proteome analysis, MRSA-challenge significantly downregulated the expression level of the beta-glucosidase gene ( $t = 4.2846$ ,  $df = 8$ ,  $p = 0.0011$ ) and induced calponin-like domain containing protein gene ( $t = -3.4864$ ,  $df = 8$ ,  $p = 0.0043$ ) expression when compared with those in the naïve termites (Fig. 6). *P. aeruginosa*-challenge increased the expression level of actin gene ( $t = -7.3798$ ,  $df = 8$ ,  $p < 0.0001$ ) while reduced the expression levels of ferritin encoding gene ( $t = 3.2076$ ,  $df = 8$ ,  $p = 0.0067$ ) and beta-glucosidase gene ( $t = 2.3772$ ,  $df = 8$ ,  $p = 0.0313$ ) (Fig. 6). Expression of the other genes remained similar in the naïve versus MDR-challenged termites to support the proteome data (Fig. 6; Tables 1 and 2).

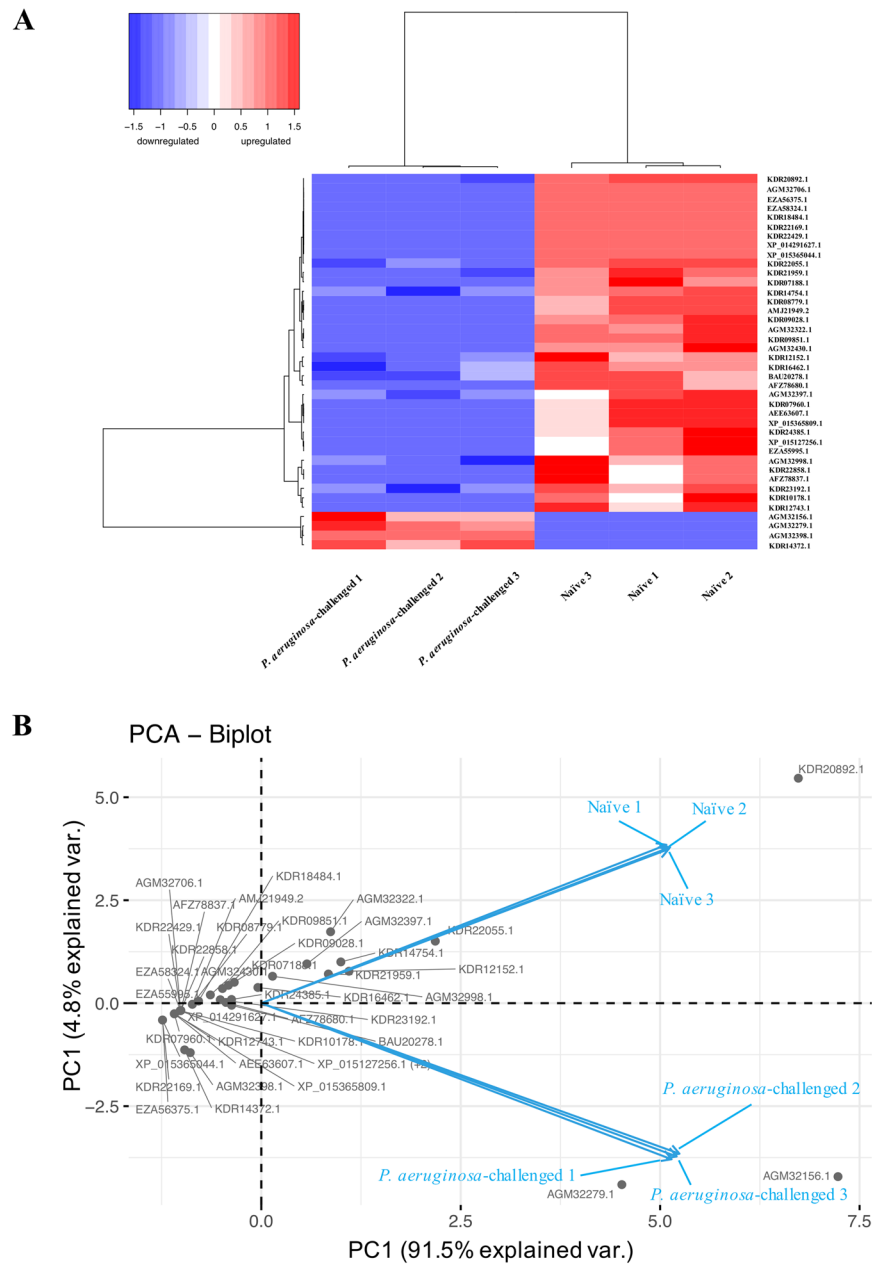
Protein	Accession Number	Poisson.FDR p-value	Quasi.FDR p-value	Rate Ratio*
enolase	AGM32398.1	0.0034	<0.0001	35.20
malate dehydrogenase	KDR14372.1	0.0061	0.0020	33.59
alpha tubulin, partial	AGM32279.1	<0.0001	0.0045	2.50
actin	AGM32156.1	0.0053	0.0218	1.61
ferritin	AGM32322.1	<0.0001	0.0007	-35.55
glutamine synthetase 2 cytoplasmic	KDR18484.1	0.0034	<0.0001	-35.20
uncharacterized protein	AGM32706.1	0.0198	<0.0001	-34.62
translationally-controlled tumor protein-like protein	EZA58324.1	0.0198	<0.0001	-34.62
protein yellow	KDR22429.1	0.0198	<0.0001	-34.62
alcohol dehydrogenase	XP_014291627.1	0.0198	<0.0001	-34.62
selenium-binding protein 1-A	KDR09028.1	0.0001	0.0017	-34.50
multifunctional protein ADE2, partial	KDR09851.1	0.0003	0.0005	-34.29
multifunctional protein ADE2	EZA56375.1	0.1132	<0.0001	-33.62
cytosolic carboxypeptidase-like protein 5, partial	KDR22169.1	0.1132	<0.0001	-33.62
predicted: alpha,alpha-trehalose-phosphate synthase [UDP-forming]-like isoform X1	XP_015365044.1	0.1132	<0.0001	-33.62
aldo-keto reductase 1	AMJ21949.2	0.0061	0.0020	-33.59
beta-glucuronidase	KDR08779.1	0.0061	0.0020	-33.59
teneurin-3	KDR07188.1	0.0002	0.0030	-32.96
unknown	AEE63607.1	0.0351	0.0077	-32.91
hypothetical protein L798_01615	KDR07960.1	0.0351	0.0077	-32.91
regucalcin	KDR12743.1	0.0351	0.0077	-32.91
predicted: bifunctional purine biosynthesis protein PURH	XP_015365809.1	0.0351	0.0077	-32.91
SCP-like extracellular domain containing protein 2	AGM32430.1	0.0011	0.0069	-32.61
c-type lysozyme-2	AFZ78837.1	0.0198	0.0223	-31.73
filamin-B	EZA55995.1	0.0198	0.0223	-31.73
Ribose-phosphate pyrophosphokinase 2, partial	KDR10178.1	0.0198	0.0223	-31.73
neurotrypsin	KDR22858.1	0.0198	0.0223	-31.73
predicted: filamin-A isoform X1	XP_015127256.1	0.0198	0.0223	-31.73
Cu/Zn superoxide dismutase	AGM32998.1	0.0121	0.0405	-2.66
enolase	AGM32397.1	0.0147	0.0456	-2.58
hemocytin, partial	KDR23192.1	0.0361	0.0384	-2.46
synaptic vesicle membrane protein VAT-1-like protein-like, partial	KDR16462.1	0.0128	0.0411	-2.42
hypothetical protein L798_04756, partial	KDR20892.1	0.0000	0.0004	-2.38
prostaglandin reductase 1	KDR24385.1	0.0627	0.0150	-2.00
hypothetical protein L798_11509	KDR14754.1	0.0034	0.0217	-1.95
glutathione S-transferase	AFZ78680.1	0.0927	0.0045	-1.87
papilin	KDR22055.1	0.0019	0.0047	-1.66
pasma alpha-L-fucosidase	KDR21959.1	0.0198	0.0079	-1.58
beta-ureidopropionase	KDR12152.1	0.0198	0.0450	-1.53
putative chemosensory protein	BAU20278.1	0.1696	0.0473	-1.46

**Table 2.** Differentially expressed hemolymph proteins of *R. flavipes* following *P. aeruginosa*-challenge. \*Rate ratio:  $-1.32 \leq \log_2(\text{rate1}/\text{rate2}) \leq 1$ .

## Discussion

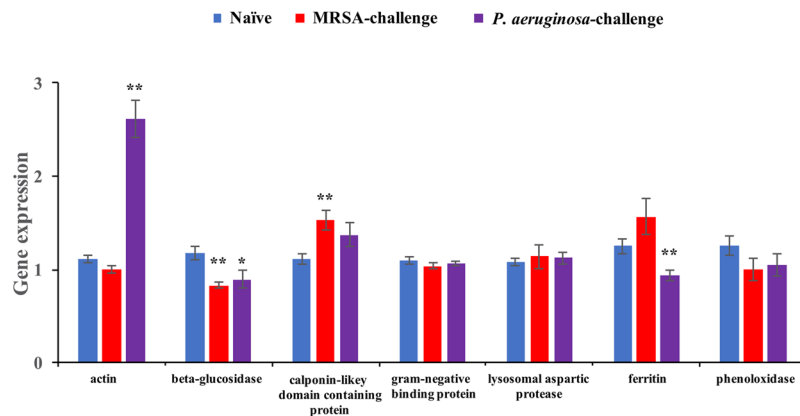
In order to further understand the inducible anti-MRSA and anti-*P. aeruginosa* activity in *R. flavipes*, we analyzed the hemolymph proteomes of the naïve and the bacterial pathogen-challenged *R. flavipes* workers in this study. We used nano-LC-MS/MS to analyze the hemolymph proteome and identified a total of 221 proteins. Of these, 57 proteins were induced by bacterial pathogen-challenges and 164 were constitutively present in naïve termites. The abundance of hexamerins, apolipoporphin, and actins in termite hemolymph suggests that hemolymph is a source for nutrient and ion storage and transportation.





**Figure 5.** Hierarchical clustering analysis (**A**) and principle component analysis (**B**) based on 40 proteins significantly changed in abundances between *P. aeruginosa*-challenged and naïve termites within the dataset. (**A**) Both samples and proteins were clustered using Ward's method, and with Pearson correlation as similarity metric. The samples are shown horizontally (columns), the proteins vertically (rows). The dendrograms represent the distances between clusters. Protein expression levels are represented in the color scale of blue (downregulated) to red (upregulated). (**B**) Grey dots and blue arrows represent hemolymph proteins and variables, respectively. Each axis represents a principal component (PC1 and PC2) with the percentage of the total variance it explains.

Three immune effector proteins, phenoloxidase (PO), Gram-negative binding protein (GNBP) and lysosomal aspartic protease were constitutively present in all three samples and their expression levels remained constant after bacterial pathogen-challenges (data not shown). POs play a crucial role in formation of melanin and reactive oxygen species (ROS) to defend against microbes and other parasites in arthropods<sup>29–31</sup>. GNBP's possess antimicrobial activity with the  $\beta$ -1,3-glucanase activity<sup>32</sup>. Lysosomal aspartic protease, which has been documented in human and various insects<sup>33–35</sup>, may contribute to the overall bactericidal activity<sup>36</sup>. Therefore, it is likely that these molecules are part of the constitutive antibacterial activities we previously reported for *R. flavipes*<sup>14,15</sup>. However, unlike in other studies in which transcript levels of the genes that encode for POs and GNBP's were overexpressed after immune challenge<sup>11,37,38</sup>, these transcripts and the proteins remained unchanged following



**Figure 6.** Transcript levels of genes encoding for seven selected proteins (mean  $\pm$  SE) between bacteria-challenged and naïve termites. Asterisks denote significant differences between bacteria-challenged and naïve termites. (\* $p < 0.05$ ; \*\* $p < 0.01$ ; paired t-test,  $n = 9$ ,  $\alpha = 0.05$ ).

MRSA or *P. aeruginosa*-challenge in our study. This suggests that the anti-MRSA and anti-*P. aeruginosa* activities in the hemolymphs of challenged *R. flavipes* are not due to these previously characterized enzymes.

Some of the differentially expressed proteins in response to bacteria-challenges are involved in immune-related processes including cytoskeletal modeling, iron metabolism, antioxidant-related immune response, stress response, and immune effectors. In addition, enzymes such as enolase and malate dehydrogenase of central metabolic pathways<sup>39,40</sup> were upregulated in *P. aeruginosa*-challenged termites to suggest a higher energetic demand of the host to suppress *P. aeruginosa* infection. This is supported by recent studies in mosquitos (*Anopheles gambiae*)<sup>41</sup> after a parasite infection and moths (*Galleria mellonella*)<sup>42</sup> to bacterial and fungal infections. Interestingly, in termites challenged with heat-killed MRSA, neither the enolase nor the malate dehydrogenase was overexpressed compared to the naïve termites. In contrast, a lipid transporter and an energy storage protein (i.e., apolipoprotein) were decreased suggesting downregulation of the lipid metabolism in response to MRSA-challenge<sup>43</sup>.

Successfully competing for iron is vital for preventing pathogen survival in the hosts<sup>44,45</sup>. In support of this strategy, transferrin was upregulated in MRSA-challenged termites<sup>46–48</sup>. In contrast to transferrin, a c-type lysozyme with hydrolase activity was severely downregulated by MRSA-challenge. Given the activity of lysozyme in hydrolyzing the  $\beta$ -1,4 linkage of peptidoglycan in bacterial cell wall and its importance as an antibacterial enzyme, it is curious that a c-type lysozyme disappeared in MRSA-challenged termites.

In *P. aeruginosa*-challenged termites, two proteins that had previously been associated with immune response were induced. Actin and alpha-tubulin are cytoskeletal elements that help maintain cell shape, participate in cellular division, and intracellular transport of molecules<sup>17,20,49–51</sup>. Expression of genes encoding for actin and alpha-tubulin have been reported to be upregulated in various insects in response to bacterial challenge<sup>7,52</sup>. In comparison to the MRSA-challenged termites, the abundance of actin and tubulin was significantly higher in *P. aeruginosa*-challenged termites (data not shown). A recent study demonstrated that insect actin identified in mosquito can mediate bacterial cell killing through phagocytosis or direct antibacterial action when it binds to the surface of bacterial cells<sup>16</sup>. Thus, it is possible that actin is an antibacterial molecule that inhibits *P. aeruginosa* and other susceptible bacteria<sup>14</sup> in *R. flavipes* hemolymph.

We hypothesized more proteins associated with insect immune response should be upregulated after a 24 h bacterial pathogen-challenge. However, more proteins, including those that have been previously reported to play roles in stress response and immunity, were downregulated in our study. Our results are similar to the study in *D. melanogaster* larvae hemolymph<sup>40</sup> in which enzymes/proteins involved in metabolism (carbohydrate, energy, lipid, and protein metabolisms) and the overall stress response (heat shock, immune response, and detoxification) were down-regulated 24 h following a bacterial-challenge. In addition, a previous study in *R. chinensis* reported an equal number of stress response proteins were either upregulated or downregulated following an entomopathogenic fungus induction<sup>11</sup>.

We selected 7 proteins to verify the LC-MS/MS data by RT-qPCR. As described in the results, our RT-qPCR data validated the upregulated level of actin and downregulated levels of ferritin and beta-glucosidase in *P. aeruginosa*-challenged termites. We also verified that the upregulated level of calponin-like domain containing protein and the reduction of beta-glucosidase in MRSA-challenged termites were due to differential gene expression at the transcriptional level. Although ferritin, an iron storage and transport protein, is important for oxidative stress, a significantly lower level of this protein was observed in termites after *P. aeruginosa* challenge in contrast to several previous studies<sup>20,53,54</sup>. However, in support of our data, a recent study reported a negative correlation between an increase of ferritin with the survivability of a lepidopteran insect against bacterial infection<sup>42</sup>. The expression of a digestive enzyme, beta-glucosidase, was also decreased after the 24 h bacterial pathogen-challenge. This suggests that termites may reduce secretion of proteins involved in cellulose degradation to couple with immune defense although cellulase abundance remained similar among treatments to support the metabolic demand (data not shown). Our data support a model in which the host responds to bacteria not only by producing proteins that can fight the potential infection directly but also by engaging its metabolic pathways and energy resources required to support the effort<sup>18</sup>.



The present study demonstrated that heat-killed MRSA or *P. aeruginosa* alters the proteome profile in the eastern subterranean termite hemolymph 24 h post bacterial feeding. The differential expression of proteins involved in metabolic activity and in insect immunity among naïve, MRSA-challenged and *P. aeruginosa* suggest that termite host defense against bacterial challenge is a concerted response involving proteins that directly kill the microbes as well as proteins in stress response and metabolism that indirectly support the host's defensive effort. Moreover, the identified immune effectors provide insights for the potential new targets of antimicrobial discovery against *P. aeruginosa* and MRSA. In conclusion, by using a proteome-wide scale study on the eastern subterranean termites where immunity and protein energy flux are tightly coupled, we provide an initial insight into a putative termite immune response against two human bacterial pathogens.

## Methods

**Termite collection and hemolymph extraction.** *R. flavipes* workers were collected on the Auburn University campus as previously described<sup>14</sup> and workers were reared with filter papers (Whatman #1) in Urban Entomology Laboratory at  $25 \pm 2^\circ\text{C}$  for at least 20 days. Three groups of 4-g termite workers were introduced into Petri plates (15 cm  $\times$  2.5 cm), surface sterilized with 70% ethanol immediately before being subjected to experiments. One group was fed with a sterile filter paper (Whatman #1) moistened with 3 ml of milli-Q (MQ) water as control. The other two groups were challenged by feeding the termites with the same filter paper moistened with 3 ml of heat-killed *P. aeruginosa* or MRSA suspension, respectively. Each bacterial suspension contained approximately  $1.8 \times 10^9$  cells. Termites were harvested after 24 h feeding and the cell-free hemolymph was extracted using the method described in Zeng *et al.*<sup>14</sup>. For each treatment, termite immunization and hemolymph extraction were performed in triplicates.

**Proteomic analysis.** Nano LC-MS/MS analyses were performed by the Mass Spectrometry & Proteomics Resource of the W. M. Keck Foundation Biotechnology Resource Laboratory of Yale School of Medicine. Cell-free hemolymph extracts were processed by reduction, alkylation and trypsin digestion for nano LC-MS/MS analysis. Briefly, the hemolymph extracts were dried and reconstituted in 40  $\mu\text{l}$  8 M urea, 0.4 M  $\text{NH}_4\text{HCO}_3$ , reduced in 4.0  $\mu\text{l}$  45 mM dithiothreitol (DTT), and incubated 30 min at  $37^\circ\text{C}$  before alkylation in 4.0  $\mu\text{l}$  100 mM iodoacetamide. 10  $\mu\text{l}$  of 0.5 mg/ml trypsin was added to each alkylated sample and incubated at  $37^\circ\text{C}$  for 16 h. After desalination, the eluted samples were dried, resuspended with 10  $\mu\text{l}$  70% formic acid and 340  $\mu\text{l}$  0.1% trifluoroacetic acid, and the protein concentrations (A260/280) were determined by Nanodrop measurements (Thermo Scientific Nanodrop 2000 UV-Vis Spectrophotometer) for mass spectrometric analysis.

Nano LC-MS/MS analysis was performed on a Thermo Scientific Q Exactive Plus mass spectrometer equipped with a Waters nanoAcquity UPLC system utilizing a binary solvent system. Trapping was performed using a Waters Symmetry<sup>®</sup> C18 (180  $\mu\text{m}$   $\times$  20 mm) column and peptides were separated using an ACQUITY UPLC PST (BEH) C18 nanoACQUITY Column. MS was acquired in profile mode over the 300–1,500  $m/z$  range and data dependent MS/MS was acquired in centroid mode. Up to 20 MS/MS were collected per MS scan on species with an intensity threshold of  $10^4$ , charge states 2–6, peptide match preferred, and dynamic exclusion set to 20 seconds.

**Protein identification and compilation of search results.** Databases of Sequest (Thermo Fisher Scientific, San Jose, CA, USA; version 2.1.0.81) and X!Tandem (The GPM, thegpm.org; version CYCLONE, 2010.12.01.1) were used for MS/MS samples' protein identification. Sequest was set up to search against a proteome database including *R. flavipes*, *P. aeruginosa*, and MRSA from NCBI assuming trypsin digestion (<https://www.ncbi.nlm.nih.gov/guide/proteins/>). X!Tandem was set up to search a subset of the database, also assuming trypsin digestion. Scaffold software (v4.6.1, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide probabilities from X!Tandem were assigned by the Peptide Prophet algorithm<sup>55</sup> with Scaffold delta-mass correction. Peptide Probabilities from X!Tandem and Sequest were assigned by the Scaffold Local FDR algorithm. Protein probabilities were assigned by the Protein Prophet algorithm<sup>55</sup>. Protein probabilities and identifications were accepted if they could be established at greater than 95% peptide probability and contained at least two identified peptides to achieve a false discovery rate (FDR; the ratio between the false peptide-spectrum matches (PSM) and the total number of PSMs above the score threshold) less than 5.0%. Proteins sharing significant peptide evidence were grouped into clusters.

**Protein function prediction.** We used Blast2GO PRO to search insect protein collections at NCBI and to predict functions of those accepted protein identifications because of the advanced functional analysis to the genomics research of non-model species<sup>56,57</sup>. BLASTp-fast searches were done with an expectation value maximum of  $10^{-3}$ .

**Statistical analysis.** Generalized linear models (GLM) was used to identify quantities of the accepted proteins across different treatments based on spectral counts. The original protein dataset was first reduced to eliminate proteins that were only present in one replicate with too few spectral counts. The acquired subsets of the protein database consisted of 168 proteins in MRSA-challenged and naïve termites and 177 proteins in *P. aeruginosa*-challenged and naïve termites.

In statistics, GLM is a popular parametric model to fit a given set of empirical observations, i.e., counts data. For our spectral count data, the GLM is expressed as:

$$\text{Log}(Y) = \beta_0 + \beta_1 X + \varepsilon,$$

where  $Y$  denotes spectral count,  $X$  stands for treatment and  $\varepsilon$  is the measurement error.

Since our dataset has been verified as overdispersion (i.e., observed variance is higher than the variance of a theoretical model) which is a common characteristic of LC-MS/MS dataset, the quasi-Poisson likelihood method<sup>58</sup> was used to deal with the overdispersion problem and to increase accuracy. In the following, we give the definition of quasi-Poisson likelihood. For the  $i$ th response  $Y_i$ , let  $E(Y_i) = u_i$ ,  $\text{Var}(Y_i) = \varphi V(u_i)$ , and  $U_i$  as  $U_i = \frac{Y_i - u_i}{\varphi V(u_i)}$ ,  $\varphi > 0$  is the unknown overdispersion parameter. Then a log quasi-likelihood ( $Q_i$ ) for  $Y_i$  is defined by the integration of  $U_i$ :  $Q_i = \int_{y_i}^{u_i} \frac{y_i - t}{\varphi V(t)} dt(t)$ , and the log quasi-likelihood for all  $n$  observations is:  $Q = \sum_{i=1}^n Q_i$ , where  $n = 6$  for each comparison. The unknown parameters  $\beta_0$ ,  $\beta_1$ , and  $\varphi$  can be estimated by many standard statistics software.

As there are over a hundred hypotheses testing between MDR-challenged and naïve termites simultaneously occurred during quasi-Poisson modeling, it is highly likely to get at least one false significant result when running hundreds of hypotheses testing simultaneously. Therefore, finding statistical methods of controlling the rate of false discovery is required. We applied a False Discovery Rate (FDR) method<sup>59</sup> which was proven to be able to control the FDR by correcting  $p$ -values when computing from quasi-Poisson likelihood. Specifically, given the problem of testing  $m$  null hypotheses simultaneously, denote  $H_{(i)}$  ( $i = 1, 2, \dots, m$ ) the null hypothesis, and  $P_{(i)}$  the corresponding ordered  $p$ -value satisfying  $P_{(1)} \leq P_{(2)} \leq \dots \leq P_{(m)}$ . We rejected all  $H_{(i)}$ ,  $i = 1, 2, \dots, k$  and  $k = \max\{i: P_i \leq \frac{i}{m}\alpha\}$ , where  $\alpha$  is the pre-specified significance level.

Significantly expressed proteins were determined by a FDR corrected quasi  $p$ -value less than 0.05 and at least 2.5-fold difference in spectral counts. To evaluate variations of hemolymph proteins between *P. aeruginosa*-challenged, MRSA-challenged and naïve termites, and to visualize strong patterns in our dataset, unsupervised hierarchical clustering was performed using Ward's method with Pearson correlation as similarity metric. This clustering technique organized all data elements into a dendrogram representing the discovered classes. In addition, a principal component analysis (PCA) was applied to the protein expression data to better visualize the dataset after class prediction analysis and the top components were used to illustrate the similarity in protein expression profiles among samples. Hierarchical clustering analysis and PCA was performed on proteins identified as significantly expressed.

All statistical analyses were performed at a significant level of  $\alpha = 0.05$ , using statistical software R (<https://www.r-project.org/>).

**RT-qPCR verification.** Total termite RNA from each treatment was obtained with TRIzol Reagent (Thermo Fisher Scientific, San Jose, CA, USA) according to the manufacturer's instructions and the mRNA was purified with the Oligotex mRNA mini Kit (Qiagen, Valencia, CA, USA). 1  $\mu$ g of mRNA per sample was reverse transcribed using TaqMan<sup>®</sup> reverse transcription reagent (Thermo Fisher Scientific, San Jose, CA, USA) to obtain cDNA products and were amplified and the resulting PCR products containing the genes of interest were sequenced by the Laragen Sequencing & Genotyping (Culver City, CA). Using the acquired sequences, specific primers for qPCR were designed using DNASTAR (Madison, WI; Supplementary Table 2). The qPCR reactions were performed with iTaq<sup>™</sup> Universal SYBR<sup>®</sup> Green Supermix on Bio-Rad CFX96 Touch<sup>™</sup> Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA). For each sample, three replicates were performed. The qPCR reactions for *phenoloxidase* and *ferritin* were performed at following condition: 95 °C for 3 min, followed by 40 cycles of 95 °C for 10 s and 54 °C for 30 s. qPCR for the other genes were performed at following condition: 95 °C for 3 min, followed by 40 cycles of 95 °C for 10 s and 58 °C for 30 s. The expressions of 7 genes of MDR-challenged termites were normalized to those of naïve termites and *alpha-tubulin* was used as an internal control. The relative gene expressions were calculated by the method of  $2^{-\Delta\Delta Ct^2}$ . The gene expression levels between MDR-challenged termite treatments and naïve termites were analyzed using paired t-tests ( $n = 9$ ,  $\alpha = 0.05$ ).

## References

- Kanost, M. R. *et al.* Insect haemolymph proteins. *Adv. Insect Physio.* **22**, 299–396 (1990).
- Kim, T. & Kim, Y. Overview of innate immunity in *Drosophila*. *J. Biochem. Mol. Biol.* **38**, 121 (2005).
- Bulet, P., Stöcklin, R. & Menin, L. Antimicrobial peptides: from invertebrates to vertebrates. *Immunol. Rev.* **198**, 169–184 (2004).
- Evans, J. D. *et al.* Immune pathways and defence mechanisms in honey bees *Apis mellifera*. *Insect Mol. Biol.* **15**, 645–656 (2006).
- Chan, Q. W., Howes, C. G. & Foster, L. J. Quantitative comparison of caste differences in honeybee hemolymph. *Mol. Cell. Proteomics.* **5**, 2252–2262 (2006).
- de Moraes Guedes, S. *et al.* *Drosophila melanogaster* larval hemolymph protein mapping. *Biochem. Biophys. Res. Commun.* **312**, 545–554 (2003).
- Wolte, D. *et al.* Proteome analysis of hemolymph changes during the larval to pupal development stages of honeybee workers (*Apis mellifera ligustica*). *J. Proteome Res.* **12**, 5189–5198 (2013).
- He, Y. *et al.* Changes in the plasma proteome of *Manduca sexta* larvae in relation to the transcriptome variations after an immune challenge: evidence for high molecular weight immune complex formation. *Mol. Cell. Proteomics.* **15**, 1176–1187 (2016).
- Zhang, Y. *et al.* Proteomics of larval hemolymph in *Bombyx mori* reveals various nutrient-storage and immunity-related proteins. *Amino Acids.* **46**, 1021–1031 (2014).
- Karlsson, C. *et al.* Proteomic analysis of the *Drosophila* larval hemolymph clot. *J. Biol. Chem.* **279**, 52033–52041 (2004).
- Liu, L., Li, G., Sun, P., Lei, C. & Huang, Q. Experimental verification and molecular basis of active immunization against fungal pathogens in termites. *Sci. Rep.* **5**, 15106, <https://doi.org/10.1038/srep15106> (2015).
- Rust, M. K. & Su, N. Y. Managing social insects of urban importance. *Annu. Rev. Entomol.* **57**, 355–375 (2012).
- Chouvenec, T., Efstathiou, C. A., Elliott, M. L. & Su, N. Y. Extended disease resistance emerging from the faecal nest of a subterranean termite. *Proc. R. Soc. Lond. [Biol.]* **280**, 20131885, <https://doi.org/10.1098/rspb.2013.1885> (2012).
- Zeng, Y., Hu, X. P. & Suh, S. J. Characterization of Antibacterial Activities of Eastern Subterranean Termite, *Reticulitermes flavipes*, against Human Pathogens. *PLoS ONE.* **11**, e0162249, <https://doi.org/10.1371/journal.pone.0162249> (2016).
- Zeng, Y., Hu, X. P., Yu, X. Q. & Suh, S. J. Multiple antibacterial activities of proteinaceous compounds in crude extract from the eastern subterranean termite, *Reticulitermes flavipes* Kollar (Blattodea: Isoptera: Rhinotermitidae). *Adv. Res.* **2**, 455–461 (2014).
- Sandiford, S. L. *et al.* Cytoplasmic actin is an extracellular insect immune factor which is secreted upon immune challenge and mediates phagocytosis and direct killing of bacteria, and is a *Plasmodium* antagonist. *PLoS Pathog.* **11**, e1004631, <https://doi.org/10.1371/journal.ppat.1004631> (2015).

17. Bartholomay, L. C. *et al.* Description of the transcriptomes of immune response-activated hemocytes from the mosquito vectors *Aedes aegypti* and *Armigeres subalbatus*. *Infect. Immun.* **72**, 4114–4126 (2004).
18. Foley, E. & O'Farrell, P. H. Functional dissection of an innate immune response by a genome-wide RNAi screen. *PLoS Biol.* **2**, e203, <https://doi.org/10.1371/journal.pbio.0020203> (2004).
19. Kallio, J. *et al.* Functional analysis of immune response genes in *Drosophila* identifies JNK pathway as a regulator of antimicrobial peptide gene expression in S2 cells. *Microbes Infect.* **7**, 811–819 (2005).
20. Vierstraete, E. *et al.* A proteomic approach for the analysis of instantly released wound and immune proteins in *Drosophila melanogaster* hemolymph. *Proc. Natl. Acad. Sci.* **101**, 470–475 (2004).
21. Freitak, D., Heckel, D. G. & Vogel, H. Bacterial feeding induces changes in immune-related gene expression and has trans-generational impacts in the cabbage looper (*Trichoplusia ni*). *Front. Zool.* **6**, 7 (2009).
22. Zhang, J., Lu, A., Kong, L., Zhang, Q. & Ling, E. Functional analysis of insect molting fluid proteins on the protection and regulation of ecdysis. *J. Biol. Chem.* **289**, 35891–35906 (2014).
23. Marikovsky, M. *et al.* Cu/Zn superoxide dismutase plays important role in immune response. *J. Immun.* **170**, 2993–3001 (2003).
24. Arai, I. *et al.* Immunohistochemical analysis of the role of hemocytin in nodule formation in the larvae of the silkworm, *Bombyx mori*. *J. Insect Sci.* **13**, (2013).
25. Sparks, M. E., Shelby, K. S., Kuhar, D. & Gundersen-Rindal, D. E. Transcriptome of the invasive brown marmorated stink bug, *Halyomorpha halys* (Stål) (Heteroptera: Pentatomidae). *PLoS ONE*. **9**, e111646, <https://doi.org/10.1371/journal.pone.0111646> (2014).
26. Zhai, Y. *et al.* Activation of the TOR signaling pathway by glutamine regulates insect fecundity. *Sci. Rep.* **5**, 10694, <https://doi.org/10.1038/srep10694> (2015).
27. Sokol, N. S. & Cooley, L. *Drosophila* filamin is required for follicle cell motility during oogenesis. *Dev. Biol.* **260**, 260–272 (2003).
28. Yu, K. H. *et al.* Comparative study on characteristics of lysozymes from the hemolymph of three lepidopteran larvae, *Galleria mellonella*, *Bombyx mori*, *Agrius convolvuli*. *Dev. Comp. Immunol.* **26**, 707–713 (2002).
29. Taft, A. S., Chen, C. C., Li, J. & Christensen, B. M. Molecular cloning of two prophenoloxidase genes from the mosquito *Aedes aegypti*. *Insect Mol. Biol.* **10**, 97–103 (2001).
30. Lai, S. C., Chen, C. C. & Hou, R. F. Immunolocalization of prophenoloxidase in the process of wound healing in the mosquito *Armigeres subalbatus* (Diptera: Culicidae). *J. Med. Entomol.* **39**, 266–274 (2002).
31. Liu, H. *et al.* Phenoloxidase is an important component of the defense against *Aeromonas hydrophila* infection in a crustacean, *Pacifastacus leniusculus*. *J. Biol. Chem.* **282**, 33593–33598 (2007).
32. Bulmer, M. S., Bachelet, I., Raman, R., Rosengaus, R. B. & Sasisekharan, R. Targeting an antimicrobial effector function in insect immunity as a pest control strategy. *Proc. Natl. Acad. Sci.* **106**, 12652–12657 (2009).
33. Hamilton, C., Lejeune, B. T. & Rosengaus, R. B. Trophallaxis and prophylaxis: social immunity in the carpenter ant *Camponotus pennsylvanicus*. *Biol. Lett.* **7**, 89–92 (2011).
34. Cho, W. L., Dhadialla, T. S. & Raikhel, A. S. Purification and characterization of a lysosomal aspartic protease with cathepsin D activity from the mosquito. *Insect Biochem.* **21**, 165–176 (1991).
35. Lemos, F. J. & Terra, W. R. Properties and intracellular distribution of a cathepsin D-like proteinase active at the acid region of *Musca domestica* midgut. *Insect Biochem.* **21**, 457–465 (1991).
36. Thorne, K. J., Oliver, R. C. & Barrett, A. J. Lysis and killing of bacteria by lysosomal proteinases. *Infect. Immun.* **14**, 555–563 (1976).
37. Rodriguez-Andres, J. *et al.* Phenoloxidase activity acts as a mosquito innate immune response against infection with Semliki Forest virus. *PLoS Pathog.* **8**, e1002977, <https://doi.org/10.1371/journal.ppat.1002977> (2012).
38. González-Santoyo, I. & Córdoba-Aguilar, A. Phenoloxidase: a key component of the insect immune system. *Entomol. Exp. Appl.* **142**, 1–16 (2012).
39. Xing, L. *et al.* Dynamics of the interaction between cotton bollworm *Helicoverpa armigera* and nucleopolyhedrovirus as revealed by integrated transcriptomic and proteomic analyses. *Mol. Cell. Proteomics.* **16**, 1009–1028 (2017).
40. de Morais Guedes, S. *et al.* Proteomics of immune-challenged *Drosophila melanogaster* larvae hemolymph. *Biochem. Biophys. Res. Commun.* **328**, 106–115 (2005).
41. Alvarado-Delgado, A. *et al.* Infection with *Plasmodium berghei* ookinetes alters protein expression in the brain of *Anopheles albimanus* mosquitoes. *Parasit. Vectors.* **9**, 542 (2016).
42. Browne, N. *et al.* Prolonged pre-incubation increases the susceptibility of *Galleria mellonella* larvae to bacterial and fungal infection. *Virulence.* **6**, 458–465 (2015).
43. Chan, Q. W., Melathopoulos, A. P., Pernal, S. F. & Foster, L. J. The innate immune and systemic response in honey bees to a bacterial pathogen, *Paenibacillus larvae*. *BMC genomics.* **10**, 387 (2009).
44. Yoshiga, T. *et al.* *Drosophila melanogaster* transferrin. *Eur. J. Biochem.* **260**, 414–420 (1999).
45. Yun, E. Y. *et al.* Molecular cloning and characterization of a cDNA encoding a transferrin homolog from *Bombyx mori*. *Biol. Chem.* **380**, 1455–1459 (1999).
46. Thompson, G. J., Crozier, Y. C. & Crozier, R. H. Isolation and characterization of a termite transferrin gene up-regulated on infection. *Insect Mol. Biol.* **12**, 1–7 (2003).
47. Altincicek, B., Knorr, E. & Vilcinskis, A. Beetle immunity: Identification of immune-inducible genes from the model insect *Tribolium castaneum*. *Dev. Comp. Immunol.* **32**, 585–595 (2008).
48. Wang, D. *et al.* Molecular characterization of iron binding proteins, transferrin and ferritin heavy chain subunit, from the bumblebee *Bombus ignitus*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **152**, 20–27 (2009).
49. Scharlaken, B. *et al.* Differential protein expression in the honey bee head after a bacterial challenge. *Arch. Insect Biochem. Physiol.* **65**, 223–237 (2007).
50. Li, J., Fang, Y., Zhang, L. & Begna, D. Honeybee (*Apis mellifera ligustica*) drone embryo proteomes. *J. Insect Physiol.* **57**, 372–384 (2011).
51. Randolt, K. *et al.* Immune-related proteins induced in the hemolymph after aseptic and septic injury differ in honey bee worker larvae and adults. *Arch. Insect Biochem. Physiol.* **69**, 155–167 (2008).
52. Loseva, O. & Engström, Y. Analysis of signal-dependent changes in the proteome of *Drosophila* blood cells during an immune response. *Mol. Cell. Proteomics.* **3**, 796–808 (2004).
53. Vierstraete, E. *et al.* Proteomics in *Drosophila melanogaster*: first 2D database of larval hemolymph proteins. *Biochem. Biophys. Res. Commun.* **304**, 831–838 (2003).
54. Geiser, D. L., Zhou, G., Mayo, J. J. & Winzerling, J. J. The effect of bacterial challenge on ferritin regulation in the yellow fever mosquito. *Aedes aegypti*. *Insect Sci.* **20**, 601–619 (2013).
55. Keller, A., Nesvizhskii, A. I., Kolker, E. & Aebersold, R. An explanation of the Peptide Prophet algorithm developed. *Anal. Chem.* **74**, 5383–5392 (2002).
56. Conesa, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics.* **21**, 3674–3676 (2005).
57. Götz, S. *et al.* High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic. Acids. Res.* **36**, 3420–3435 (2008).
58. Li, M. *et al.* Comparative shotgun proteomics using spectral count data and quasi-likelihood modeling. *J. Proteome Res.* **9**, 4295–4305 (2010).
59. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* **289**–300 (1995).

## Acknowledgements

The study was supported by the Auburn University Intramural Grant Program (#160211 Suh) and the Hatch program of the National Institute of Food and Agriculture, U.S. Department of Agriculture (#ALA0SUH for S.-J.S. and #ALA015-1-15005 for X.P.H.). Cao's research was supported in part by the Simons Foundation grant #354917 and the National Science Foundation grant DMS 1736470. We thank Dr. Leonardo De La Fuente of Auburn University for access to his equipment.

## Author Contributions

Conceived and designed the study: Y.Z., X.P.H. and S.-J.S. Performed the experiments: Y.Z. Analyzed the data: Y.Z. and G.C. Wrote the paper: Y.Z., G.C., X.P.H. and S.-J.S. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-31681-2>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018