



Transcriptome analysis of host-associated differentiation in *Bemisia tabaci* (Hemiptera: Aleyrodidae)

Wen Xie¹, Qingjun Wu¹, Shaoli Wang¹, Xiaoguo Jiao¹, Litao Guo¹, Xuguo Zhou^{2*} and Youjun Zhang^{1*}

¹ Department of Plant Protection, Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China

² Department of Entomology, S-225 Agricultural Science Center North, University of Kentucky, Lexington, KY, USA

Edited by:

Raman Chandrasekar, Kansas State University, USA

Reviewed by:

Murugan E. Kadarkarai, Bharathiar University, India

Fang-Hao Wan, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, China

*Correspondence:

Xuguo Zhou, Department of Entomology, S-225 Agricultural Science Center North, University of Kentucky, Lexington, KY 40546-0091, USA

e-mail: xuguo Zhou@uky.edu;

Youjun Zhang, Department of Entomology, Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, 12 Zhongguancun Nandajie, Haidian District, Beijing 100081, China
e-mail: zhangyoujun@caas.cn

Host-associated differentiation is one of the driving forces behind the diversification of phytophagous insects. In this study, host induced transcriptomic differences were investigated in the sweetpotato whitefly *Bemisia tabaci*, an invasive agricultural pest worldwide. Comparative transcriptomic analyses using coding sequence (CDS), 5' and 3' untranslated regions (UTR) showed that sequence divergences between the original host plant, cabbage, and the derived hosts, including cotton, cucumber and tomato, were 0.11–0.14%, 0.19–0.26%, and 0.15–0.21%, respectively. In comparison to the derived hosts, 418 female and 303 male transcripts, respectively, were up-regulated in the original cabbage strain. Among them, 17 transcripts were consistently up-regulated in both female and male whiteflies originated from the cabbage host. Specifically, two ESTs annotated as *Cathepsin B* or *Cathepsin B-like* genes were significantly up-regulated in the original cabbage strain, representing a transcriptomic response to the dietary challenges imposed by the host shifting. Results from our transcriptome analysis, in conjunction with previous reports documenting the minor changes in their reproductive capacity, insecticide susceptibility, symbiotic composition and feeding behavior, suggest that the impact of host-associated differentiation in whiteflies is limited. Furthermore, it is unlikely the major factor contributing to their rapid range expansion/invasiveness.

Keywords: *Bemisia tabaci*, comparative transcriptomic analysis, host-associated differentiation, qRT-PCR, *Cathepsin*

INTRODUCTION

Host-associated differentiation (HAD) in insects, occurring mainly in sympatric populations, is generally considered to be an adaptation to different habitats and resources. This adaptive response plays a crucial role in the early stage of speciation and may lead to reproductive isolation (Nosil et al., 2002; Rundle and Nosil, 2005; Stireman et al., 2005; Schluter, 2009). The advent of Genomics Era, especially the development of the next generation sequencing technology (NGS) (Mardis, 2008; Schuster, 2008), provides an unprecedented opportunity for us to understand HAD at the omics level and its potential association with the host-derived genetic divergence.

The sweetpotato whitefly *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae), an invasive agricultural pest, causes severe damages to agricultural and horticultural crops worldwide through direct feeding and indirect vectoring of plant viruses (Jones, 2003). More than 34 genetically distinct, yet morphologically indistinguishable *B. tabaci* species have been identified (Perring, 2001; Dinsdale et al., 2010; Elbaz et al., 2010; De Barro et al., 2011; De Barro, 2012; Liu et al., 2012). Among them, *B. tabaci* B (also known as the Middle East-Asia Minor 1) has emerged as a major insect pest due to their global invasion for the past 20 years (Brown et al., 1995; Liu et al., 2007). In China, native *B. tabaci* was first recorded in the late 1940s. However,

B. tabaci has not been a pest for vegetable, fruit and ornamental crops until the introduction and establishment of the invasive *B. tabaci* B in the mid-1990s (Zhang et al., 2005; Chu et al., 2006).

In recent years, several studies have been focused on the comparative analysis of host adaptation among *B. tabaci* species complex (Iida et al., 2009; Tsueda and Tsuchida, 2011; Jiao et al., 2012, 2013; Saleh et al., 2012). Species-specific divergence can be attributed to the host plant's response to feeding, and further mediated by the interactions among *B. tabaci*, endosymbionts, plant viruses and natural enemies (Lapidot et al., 2001; Colvin et al., 2006; Inbar and Gerling, 2008; Pan et al., 2012). The performance of *B. tabaci* B on various host plants, including cotton, tobacco, cabbage, squash, kidney beans, garden beans and soybeans, were evidently different (Zang et al., 2006; Mansaray and Sundufu, 2009). The greenhouse whitefly showed substantially different feeding behavior in association with host-plant acceptance (Lei et al., 1998). Whitefly cryptic species maintained on different host plants exhibited differential susceptibilities to insecticides (Liang et al., 2007; Castle et al., 2009; Xie et al., 2011). Host plants themselves may play an important role in explaining why morphologically- indistinguishable *B. tabaci* sibling species have different host range (Brown et al., 1995).

Originated from a *B. tabaci* B population collected in 2004, we established several laboratory whitefly strains feeding exclusively on their original host, cabbage, and the derived hosts, cucumber, tomato, cotton and poinsettia. After maintaining on their respective hosts for 3–5 years, these whitefly strains displayed few variations in their life history traits (Lou et al., 2005; Fu et al., 2008), insecticide susceptibility (Xie et al., 2011), symbiotic contents (Pan et al., 2013) and feeding behavior (Liu et al., 2013). Specifically, *B. tabaci* strains maintained on cabbage and poinsettia had a significantly longer developmental time from egg to adult than the tomato strain (Lou et al., 2005; Fu et al., 2008). In terms of insecticide susceptibility, a whitefly strain derived from poinsettia was, in general, less sensitive to insecticide treatments (Xie et al., 2011). As for the symbiont composition, the cucumber strain harbored more *Portiera* than strains maintained on cabbage and cotton, while cabbage strain harbored more *Rickettsia* and *Cardinium* (Pan et al., 2013). Using electrical penetration graph technology to quantitatively measure the feeding behavior, we showed that *B. tabaci* maintained on their original cabbage host outperformed whiteflies derived from cucumber and tomato (Liu et al., 2013). Despite these phenotypic differences between the original and derived hosts, genetic variations among whiteflies has yet been investigated, especially at the transcriptome level. In this study, we examined the sequence divergence between a *B. tabaci* population originated from a cabbage host and three other whitefly populations which have been subjected to the long-term isolation in cotton, cucumber and tomato, respectively, for multiple generations. Results from the comparative transcriptomic profiling will shed light on the molecular mechanisms underlying the HAD, and contribute to our understanding of the host-induced evolution in this emerging global invasive pest.

MATERIALS AND METHODS

SAMPLE PREPARATION

The parental whitefly population was established in 2004 by releasing 10 pairs of *B. tabaci* adults, which had been collected from cabbage in Haidian District of Beijing, into one screen cage containing one cabbage plant in a glasshouse (Xie et al., 2011, 2014; Liu et al., 2013; Pan et al., 2013). This original population was identified as a B-type (Middle East-Asia Minor 1) based on the mitochondria cytochrome oxidase 1 (mtCO1) marker (Zhang et al., 2005). After two generations, more cabbage plants (i.e., three plants in total) were placed into the parental population cage. At the fourth generation, each cabbage plant was randomly transferred from the parental population to a cage with one derived hosts, cotton, cucumber and tomato, respectively. These four *B. tabaci* populations, including one original and three derived, have the same genetic background and have been maintained in the same glasshouse without exposure to any insecticides for at least 4 years (approximately 60 generations) (Xie et al., 2011, 2014; Liu et al., 2013; Pan et al., 2013). Only 2-day-old adults of these different host strain female and male samples were collected respectively (about 1000 female or male from each sample), pooled into a plastic tube using an aspirator, and then snap frozen in liquid nitrogen, and transferred to a -80°C for freezer for the long-term storage.

RNA ISOLATION, cDNA LIBRARY CONSTRUCTION, AND ILLUMINA SEQUENCING

RNA from each female or male sample group was extracted, RNA purity and degradation were checked on 1% agarose gels, and further integrity was also confirmed using the 2100 Bioanalyzer (Agilent Technologies) with a minimum RNA integrated number value of 8. Poly (A)-containing RNA was then obtained from total RNA with the Dynabeads® mRNA purification kit (Invitrogen). Then, the mRNA was fragmented into small pieces using divalent cations under elevated temperature and the cleaved RNA fragments were used for first strand cDNA synthesis using reverse transcriptase and random primers. This was followed by second strand cDNA synthesis using DNA polymerase I and RNaseH. These cDNA fragments were subjected to end repair process and ligation of adapters. These products were purified and enriched with PCR to create the final cDNA library. All eight paired-end cDNA libraries (500 bp size), including female *B. tabaci* from cabbage (Caf), cucumber (Cuf), tomato (Tof), cotton (Cof), and male whiteflies from cabbage (Cam), cucumber (Cum), tomato (Tom), and cotton (Com), were prepared following the manufacturer's recommendations and sequenced on an Illumina GAII platform.

SEQUENCE DIVERGENCE ANALYSIS

Sequence divergence analysis pipeline was modified from Wang et al. (2011) and Xie et al. (2014). Raw reads of eight cDNA libraries were all filtered to remove adaptor sequences and low quality sequences (containing reads with unknown sequences "N" and lower than 10 of mean phred scores). Then, the reads obtained from each library were assembled separately, and all pooled reads of 8 libraries were also assembled into a reference transcriptome using Trinity with default parameters. Pooled reference transcriptome was annotated using a BLAST search against the non-redundant (NR) database in NCBI, SWISS-PROT, KEGG and COG with an e -value cut-off of $1e^{-5}$.

A customized informatics pipeline was developed to calculate the divergence ratio between the original cabbage population and three derived host populations (cotton, cucumber and tomato). A bidirectional best hit method was used in BLAST search (blastn, e -value cutoff is $1e^{-10}$) to identify genes that are putatively pairs of reciprocal best matched sequences (RBMs) in the eight separately assembled cDNA libraries. Only pairs of sequences that corresponded unambiguously to the same ESTs in the pooled reference transcriptome, and were annotated to the same protein in Swissprot database with an $E < 1 \times 10^{-5}$, were selected as orthologs. Coding sequences (CDS) of the orthologous genes corresponding to the ESTs from the pooled reference transcriptome were determined by BLASTx against all known proteins in Swissprot database using a threshold of 1×10^{-5} . CDS with unexpected stop codon in the Blast hit region and/or shorter than 150 bp was removed. Boundaries defining the primary structure of sequences, including the start and codon, 3' and 5'UTR of each pair of orthologs, were predicted following Wang et al. (2011). Based on the extracted CDS, 5'UTR and 3'UTR of each pair of orthologs from each comparing group, sequence divergence was calculated by dividing the number of substitutions by the number of base pairs compared, while substitution rates were

estimated separately for synonymous (Ks) and non-synonymous sites (Ka) using an approximate method implemented in the KaKs Calculator Version 1.2 with pair-wise approximate analyses (Yang and Nielsen, 2000; Tiffin and Hahn, 2002; Zhang et al., 2006; Wang et al., 2011).

IDENTIFICATION OF DIFFERENTIALLY EXPRESSED GENES

Differentially expressed genes were determined according to Xie et al. (2014). Briefly, based on the number of assembled genes, trinity pooled transcripts and raw reads for each sample were analyzed using RSEM (v1.1.15) (Li and Dewey, 2011). To calculate the normalized read counts for each library, RNA composition bias and normalization factors were taken into account and calculated by edgeR (Robinson et al., 2010). The spearman correlation coefficient for all samples was generated with *heatmap 2* in R package (Durinck et al., 2009). The normalized read counts were subjected to the DEGseq package to identify differentially expressed genes using MARS model ($Q < 0.05$ and Fold Change > 2) (Storey and Tibshirani, 2003; Wang et al., 2010).

QUANTITATIVE REAL TIME PCR (qRT-PCR) ANALYSIS

A total of 300 newly emerged *B. tabaci* adults from the original cabbage (Ca) and derived cucumber (Cu), cotton (Co) and tomato (To) hosts were, respectively, collected and snap frozen in liquid nitrogen for the subsequent qRT-PCR analysis (representing three biological replicates, each using 100 adults). Total RNA was extracted using Trizol (Invitrogen) following the manufacturer's protocol. The resultant total RNA was re-suspended in nuclease-free water and the concentration was measured using Nanodrop (Thermo Scientific Nanodrop 2000). Approximately 0.5 μg of total RNA was used to synthesize the first-strand cDNA using a PrimerScript RT reagent Kit (TaKaRa). The resultant cDNA was diluted to 0.1 $\mu\text{g}/\mu\text{l}$ for qRT-PCR analysis (ABI 7500) using a SYBR Green Realtime PCR Master Mix (TaKaRa). qRT-PCR primers were designed using the Primer Express 2.0 software. The cycling parameter was 95°C for 30 s, followed by 40 cycles of 95°C for 5 s and 62°C for 34 s, and concluded with the melting curve analysis (65–95°C in increments of 0.5°C every 5 s) to check for non-specific product amplification. Relative gene expression was calculated by the ΔCT method using *RPL29* as a reference (Li et al., 2013) to eliminate sample-to-sample variations in the initial cDNA samples.

RESULT AND DISCUSSION

GENETIC DIVERGENCE AMONG DIFFERENT HOST STRAINS

To compare the sequence divergence between *B. tabaci* maintained on their original host, cabbage, and their derived hosts, cotton, cucumber and tomato, respectively, we analyzed the orthologous gene pairs between their transcriptomes using a bidirectional best hit approach. As a result, 26020, 32193, 31222, 17916, 18678, and 39541 pairs of reciprocal best matched sequences corresponding, respectively, to Caf/Cof, Caf/Cuf, Caf/Tof, Cam/Com, Cam/Cum, and Cam/Tom pairing groups were identified (Table 1). Among them, 6718 (Caf/Cof), 7914 (Caf/Cuf), 7756 (Caf/Tof), 3642 (Cam/Com), 4920 (Cam/Cum), and 8031 (Cam/Tom) orthologous genes were annotated to the same protein in Swissprot database and corresponded to the same EST in the pooled reference transcriptome. In addition, 651 pairs of 5'UTR and 757 pairs of 3'UTR were identified in Caf/Cof, followed by 965 pairs of 5'UTR and 1323 pairs of 3'UTR in Caf/Cuf, 900 pairs of 5'UTR and 1162 pairs of 3'UTR in Caf/Tof, 200 pairs of 5'UTR and 266 pairs of 3'UTR in Cam/Com, 305 pairs of 5'UTR and 372 pairs of 3'UTR in Cam/Cum, 705 pairs of 5'UTR and 971 pairs of 3'UTR in Cam/Tom. The mean value of 5'UTR, CDS, 3'UTR of orthologous genes was 0.23% (0.19–0.26%), 0.13% (0.11–0.14%) and 0.19% (0.15–0.21%), respectively (Table 1). To identify genes undergoing positive and purifying selections due to the host induction, we estimated rates of non-synonymous (Ka) and synonymous (Ks) substitutions of cotton, cucumber and tomato with their original cabbage strain. If $Ka/Ks > 1$, it is considered to be a positive selection, whereas when $Ka/Ks < 1$, it is a purifying selection. We identified 24 (Caf/Cof), 17 (Caf/Cuf), 23 (Caf/Tof), 9 (Cam/Com), 13 (Cam/Cum), and 36 genes (Cam/Tom) with Ka/Ks ratio greater than one (Table 1; Tables S3–S8; Figure 1).

Bemisia tabaci B has more than 600 known host plants (Oliveira et al., 2001); however, few studies have explored the host induced genetic variation using a next generation sequencing approach. Previous research associated with the genetic divergence in *B. tabaci* were mainly focused on the comparisons of *B. tabaci* samples from different geographical regions using a suite molecular markers, such as mitochondrial 16S DNA, cytochrome oxidase 1 and the nuclear ribosomal intergenic spacer (Dinsdale et al., 2010; De Barro et al., 2011). Moreover, the differences among three cryptic species *B. tabaci* B, Q and Asia II 3 were

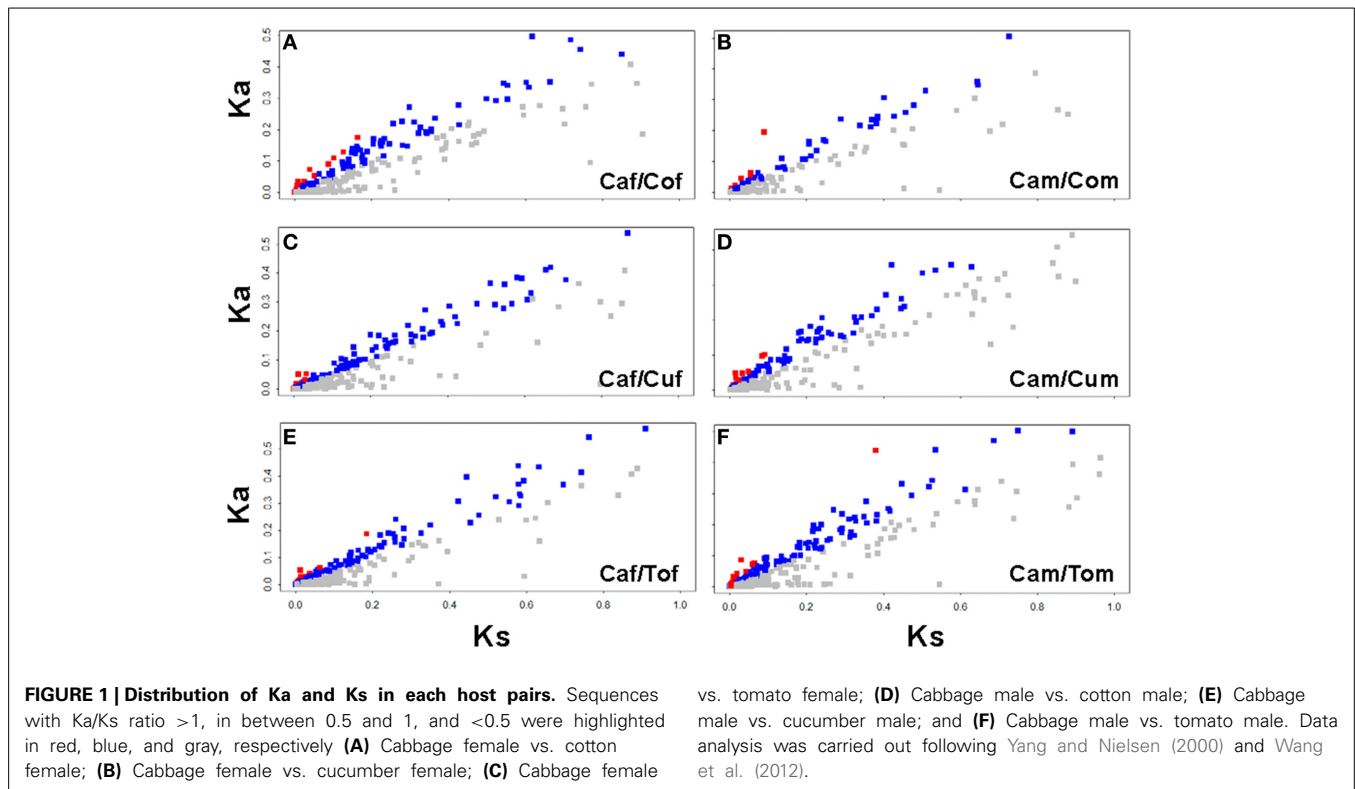
Table 1 | Divergence analysis between the original host and the derived hosts.

Host Pairs*	Matched sequences	Orthologs	5'UTR (%)**	CDS (%)	3'UTR (%)	Ka/Ks > 1 ***
Caf/Cof	26020	6718	651 (0.19)	6673 (0.13)	757 (0.20)	24
Caf/Cuf	32193	7914	965 (0.22)	7873 (0.11)	1323 (0.21)	17
Caf/Tof	31222	7756	900 (0.24)	7697 (0.11)	1162 (0.20)	23
Cam/Com	17916	3642	201 (0.26)	3608 (0.14)	266 (0.19)	9
Cam/Cum	18678	4920	305 (0.23)	4878 (0.14)	372 (0.15)	13
Cam/Tom	39541	8031	705 (0.22)	7958 (0.12)	971 (0.17)	36

*Ca, Co, Cu, and To denote the original host plant cabbage, and the derived host plants cotton, cucumber, and tomato, respectively. Comparative analysis was carried out in both male (m) and female (f) whiteflies.

**The number of transcripts has 5' UTR (the average divergence rate of all 5' UTR).

***The number of orthologs if Ka (non-synonymous substitution rate)/ Ks (synonymous substitution rate) > 1 .



revealed by the whole transcriptome divergence analysis (Wang et al., 2011, 2012). In this study, we indentified thousands of ortholog gene pairs and determined the whole sequence divergence between the original cabbage strain and the derived host strains (cotton, cucumber and tomato; >4 years) to be 0.11–0.14%, 0.19–0.26%, and 0.15–0.21%, respectively, for the CDS, 5′UTR and 3′UTR (Table 1). This value is lower than the reported mean divergence between human and chimpanzee (0.45, 1.12, and 0.86%, respectively, for CDS, 5′UTR and 3′UTR divergence) (Hellmann et al., 2003; Shi et al., 2003). It also lower than the ratio between *B. tabaci* B and Q (0.83, 1.66, and 1.43% respectively for CDS, 5′UTR and 3′UTR divergence) (Wang et al., 2011), between *B. tabaci* B and Asia II 3 (1.73% for CDS) and between *B. tabaci* Q and Asia II 3 (1.84% for CDS) (Wang et al., 2012). Above results indicate that the divergence between the original cabbage strain and derived host strains (cotton, cucumber and tomato) was limited between their transcriptomes, although these whiteflies have been subjected to the long-term isolation in different host plants (>4 years).

GENE EXPRESSION PROFILE AMONG DIFFERENT HOST STRAINS

To identify cabbage strain-biased genes, we compared gene expression profiles between Caf/Cof, Caf/Cuf, Caf/Tof, Cam/Com, Cam/Cum, and Cam/Tom, respectively. Differentially expressed transcripts (DETs) with $p < 0.05$ and \log_2 (Fold change) >1 or \log_2 (Fold change) <−1] were identified bioinformatically among each original/derived host pairings. In summary, 6294 up-regulated and 1139 down-regulated transcripts were found in Caf/Cof, 916 and 593 in Caf/Cuf, 2104 and 577 in Caf/Tof, 4257 and 1070 in Cam/Com, 5055 and 725 in

Table 2 | Differentially expressed genes between the original host and the derived hosts.

Host pairs	Differential expression genes	Up-regulated*	Down-regulated**
Caf/Cof	7433	6294	1139
Caf/Cuf	1509	916	593
Caf/Tof	2681	2104	577
Cam/Com	5327	4257	1070
Cam/Cum	5780	5055	725
Cam/Tom	1214	749	465

* Q -value < 0.05 and \log_2 (Fold change) > 1.

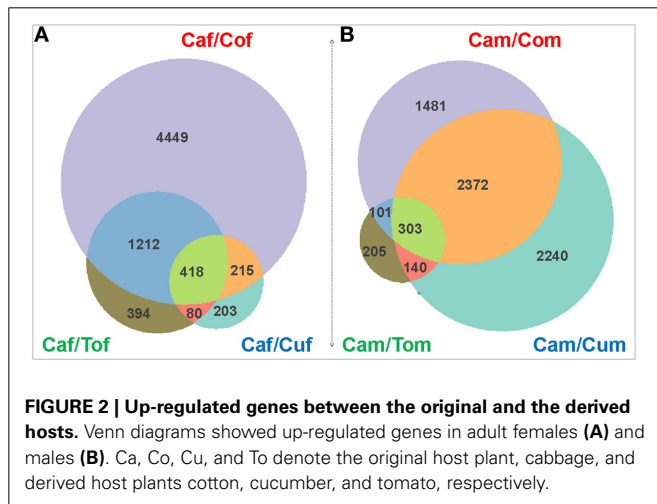
** Q -value < 0.05 and \log_2 (Fold change) < −1.

Cam/Cum, 749 and 465 in Cam/Tom (Figures 2, 3; Figure S1; Table 2). Remarkably, in comparison to the derived hosts, 418 female and 303 male transcripts, respectively, were up-regulated in the original cabbage strain (Tables S1, S2). Among them, 17 transcripts were consistently up-regulated in both female and male whiteflies originated from the cabbage host (Table S9). Annotated transcripts (10) were subjected to the quantitative real-time PCR analysis (Tables S10, S11), in which 50% of the genes (5/10) were consistent with the RNAseq results. Specifically, two ESTs annotated as *Cathepsin B* or *Cathepsin B-like* genes were significantly up-regulated in the cabbage strain in comparison to the other three derive host strains ($P < 0.01$).

Cathepsin B and Cathepsin B-like cysteine protease genes represent large multigene families. Due to an unique “occluding loop,” these cysteine proteases can block the access of substrates

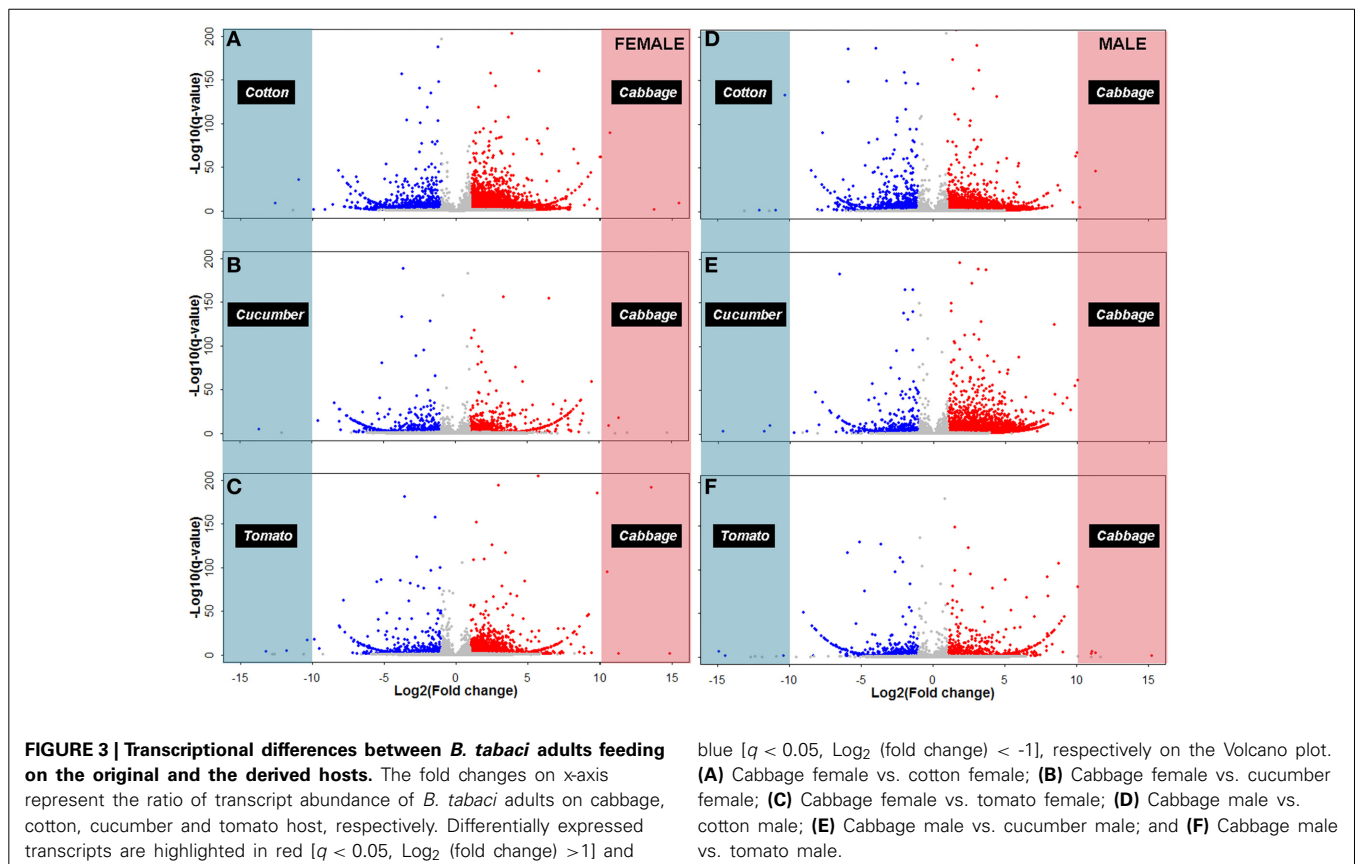
and inhibitors in human (Musil et al., 1991; Illy et al., 1997). Recently, these cysteine proteases were found to play a key role in the biology of trematodes and to be important virulence factors in a picomplexan parasite *Eimeria tenella* (Rieux et al., 2012), liver fluke parasites *Fasciola hepatica* (Beckham et al., 2009) and *Fasciola gigantica* (Siricoon et al., 2012; Chantree et al., 2013). Previous study also shown that cowpea bruchids could fend off plant defensive efforts, such as dietary soybean cysteine protease inhibitor (scN), by over-expression of major digestive cathepsin

L-like proteases as well as by activating scN-insensitive cathepsin B-like proteases (Moon et al., 2004; Koo et al., 2008; Jongasma and Beekwilder, 2011). When *B. tabaci* faces dietary challenges, such as host shift, Cathepsin B and Cathepsin B-like enzymes may play a similar role to compromise the chemical attack initiated by the host plants. Over-expression of these genes represents a transcriptomic response in whiteflies to xenobiotic compounds in cabbage phloem. Future studies involving cloning and functional characterization of these *Cathepsin B* and *Cathepsin B-like* genes are warranted to elucidate their involvement in the *B. tabaci* host adaptation mechanism.



CONCLUSIONS

HAD, one of the driving forces for the speciation and diversification, has fascinated scientists for decades. In this study, host induced transcriptomic responses were investigated using four sweetpotato whitefly strains derived from a single cabbage host. *Bemisia tabaci*, a global invasive agricultural pest, has been notorious for its adaptation capability to different host plants and rapid range expansion/invasion for the past 50 years. Comparative transcriptomic analyses exhibited low level of transcriptional divergences between whiteflies maintained on the original host, cabbage, and the three derived hosts, cotton, cucumber and tomato. Although *B. tabaci* exhibits strong host plasticity in the field, this investigation suggests that the impact of HAD in whiteflies is limited at the transcription level. Therefore, host-associated adaptation is unlikely the primary



factor contributing to the rapid range expansion/invasiveness in whiteflies.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: Wen Xie, Xuguo Zhou, Youjun Zhang. Performed the experiments: Wen Xie. Analyzed the data: Wen Xie, Xuguo Zhou. Contributed reagents/materials/analysis tools: Wen Xie, Qingjun Wu, Shaoli Wang, Xuguo Zhou, Youjun Zhang. Wrote the paper: Wen Xie, Qingjun Wu, Shaoli Wang, Xiaoguo Jiao Litao Guo, Xuguo Zhou, Youjun Zhang.

ACKNOWLEDGMENTS

The authors are grateful to the two anonymous reviewers for their critical comments and suggestions. This research was supported by the National Science Fund for Distinguished Young Scholars (31025020), the Special Fund for Agro-scientific Research in the Public Interest (201303028), the National Science & Technology pillar program (2012BAD19B01), and Key Laboratory of Biology and Genetic Improvement of Horticultural Crops, Ministry of Agriculture, P.R. China.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fphys.2014.00487/abstract>

REFERENCES

- Beckham, S. A., Piedrafita, D., Phillips, C. I., Samarawickrema, N., Law, R. H., Smooker, P. M., et al. (2009). A major cathepsin B protease from the liver fluke *Fasciola hepatica* has atypical active site features and a potential role in the digestive tract of newly excysted juvenile parasites. *Int. J. Biochem. Cell Biol.* 41, 1601–1612. doi: 10.1016/j.biocel.2009.02.003
- Brown, J. K., Frohlich, D. R., and Rosell, R. C. (1995). The sweetpotato or silverleaf whiteflies: biotypes of *Bemisia tabaci* or a species complex? *Ann. Rev. Entomol.* 40, 511–534. doi: 10.1146/annurev.ent.40.010195.002455
- Castle, S. J., Prabhaker, N., Henneberry, T. J., and Toscano, N. C. (2009). Host plant influence on susceptibility of *Bemisia tabaci* (Hemiptera: Aleyrodidae) to insecticides. *Bull. Entomol. Res.* 99, 263–273. doi: 10.1017/S0007485308006329
- Chantree, P., Phatsara, M., Meemon, K., Chaichanasak, P., Changklungmoa, N., Kueakhai, P., et al. (2013). Vaccine potential of recombinant cathepsin B against *Fasciola gigantica*. *Exp. Parasitol.* 135, 102–109. doi: 10.1016/j.exppara.2013.06.010
- Chu, D., Zhang, Y. J., Judith, K. B., Cong, B., Xu, B. Y., Wu, Q. J., et al. (2006). The introduction of the exotic Q biotype of *Bemisia tabaci* from the Mediterranean region China on ornamental crops. *Fla. Entomol.* 89, 168–174. doi: 10.1653/0015-4040(2006)89[168:TIOTEQ]2.0.CO;2
- Colvin, J., Omongo, C. A., Govindappa, M. R., Stevenson, P. C., Maruthi, M. N., Gibson, G., et al. (2006). Host - plant viral infection effects on arthropod-vector population growth, development and behaviour: management and epidemiological implications. *Adv. Virus Res.* 67, 419–452. doi: 10.1016/S0065-3527(06)67011-5
- De Barro, P. J. (2012). The *Bemisia tabaci* Species Complex: questions to guide future research. *J. Integr. Agric.* 11, 187–196. doi: 10.1016/S2095-3119(12)60003-3
- De Barro, P. J., Liu, S. S., Boykin, L. M., and Dinsdale, A. B. (2011). *Bemisia tabaci*: a statement of species status. *Ann. Rev. Entomol.* 56, 1–19. doi: 10.1146/annurev-ento-112408-085504
- Dinsdale, A. B., Cook, L., Riginos, C., Buckley, Y. M., and De Barro, P. J. (2010). Refined global analysis of *Bemisia tabaci* (Hemiptera: Sternorrhyncha: Aleyrodoidea: Aleyrodidae) mitochondrial cytochrome oxidase 1 to identify species level genetic boundaries. *Ann. Entomol. Soc. Am.* 103, 196–208. doi: 10.1603/AN09061
- Durinck, S., Spellman, P. T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package *biomaRt*. *Nat. Protoc.* 4, 1184–1191. doi: 10.1038/nprot.2009.97
- Elbaz, M., Lahav, N., and Morin, S. (2010). Evidence for pre-zygotic reproductive barrier between the B and Q biotypes of *Bemisia tabaci* (Hemiptera: Aleyrodidae). *Bull. Entomol. Res.* 100, 581–590. doi: 10.1017/S0007485309990630
- Fu, W., Xu, B. Y., Wu, Q. J., Wang, S. L., Zhu, G. R., and Zhang, Y. J. (2008). Effects of host plant on development and reproduction of B biotype *Bemisia tabaci*. *Plant Prot.* 34, 63–66. (Chinese)
- Hellmann, I., Zollner, S., Enard, W., Ebersberger, I., Nickel, B., and Paabo, S. (2003). Selection on human genes as revealed by comparisons to chimpanzee cDNA. *Genome Res.* 13, 831–837. doi: 10.1101/gr.944903
- Iida, H., Kitamura, T., and Honda, K. (2009). Comparison of egg-hatch-ingrate, survival rate and development time of the immature stage between B- and Q-biotypes of *Bemisia tabaci* (Gennadius) (Homoptera: Aleyrodidae) on various agricultural crops. *Appl. Entomol. Zool.* 44, 267–273. doi: 10.1303/aez.2009.267
- Illy, C., Quraishi, O., Wang, J., Purisima, E., Vernet, T., and Mort, J. S. (1997). Role of the occluding loop in cathepsin B activity. *J. Biol. Chem.* 272, 1197–1202. doi: 10.1074/jbc.272.2.1197
- Inbar, M., and Gerling, D. (2008). Plant-mediated interactions between whiteflies, herbivores, and natural enemies. *Annu. Rev. Entomol.* 53, 431–448. doi: 10.1146/annurev.ento.53.032107.122456
- Jiao, X. G., Xie, W., Guo, L. T., Liu, B. M., Wang, S. L., Wu, Q. J., et al. (2013). Differences in host selection and performance between B and Q putative species of *Bemisia tabaci* on three host plants. *Entomol. Exp. Appl.* 147, 1–8. doi: 10.1111/eea.12040
- Jiao, X. G., Xie, W., Wang, S. L., Wu, Q. J., Zhou, L., Pan, H. P., et al. (2012). Host preference and nymph performance of B and Q putative species of *Bemisia tabaci* on three host plants. *J. Pest Sci.* 85, 423–430. doi: 10.1007/s10340-012-0441-2
- Jones, D. R. (2003). Plant viruses transmitted by whiteflies. *Eur. J. Plant Pathol.* 109, 195–219. doi: 10.1023/A:1022846630513
- Jongsma, M. A., and Beekwilder, J. (2011). Co-evolution of insect proteases and plant protease inhibitors. *Curr. Protein Pept. Sci.* 12, 437–447. doi: 10.2174/138920311796391115
- Koo, Y. D., Ahn, J. E., Salzman, R. A., Moon, J., Chi, Y. H., Yun, D. J., et al. (2008). Functional expression of an insect cathepsin B-like counter-defence protein. *Insect Mol. Biol.* 17, 235–245. doi: 10.1111/j.1365-2583.2008.00799.x
- Lapidot, M., Friedmann, M., Pilowsky, M., Ben-Joseph, R., and Cohen, S. (2001). Effect of host plant resistance to tomato yellow leaf curl virus (TYLCV) on virus acquisition and transmission by its whitefly vector. *Phytopathology* 91, 1209–1213. doi: 10.1094/PHYTO.2001.91.12.1209
- Lei, H., Tjallingii, W. F., and van Lenteren, J. C. (1998). Probing and feeding characteristics of the greenhouse whitefly in association with host-plant acceptance and whitefly strains. *Entomol. Exp. Appl.* 88, 73–80. doi: 10.1046/j.1570-7458.1998.00348.x
- Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323. doi: 10.1186/1471-2105-12-323
- Li, R. M., Xie, W., Wang, S. L., Wu, Q. J., Yang, N. N., Yang, X., et al. (2013). Reference gene selection for qRT-PCR analysis in the sweetpotato whitefly, *Bemisia tabaci* (Hemiptera: Aleyrodidae). *PLoS ONE* 8:e53006. doi: 10.1371/journal.pone.0053006
- Liang, P., Cui, J. Z., Yang, X. Q., and Gao, X. W. (2007). Effects of host plants on insecticide susceptibility and carboxylesterase activity in *Bemisia tabaci* biotype B and greenhouse whitefly, *Trialeurodes vaporariorum*. *Pest Manag. Sci.* 63, 365–371. doi: 10.1002/ps.1346
- Liu, B. M., Preisser, E. L., Jiao, X. G., Pan, H. P., Xie, W., Wang, S. L., et al. (2013). Plant-mediated changes in the feeding behavior of an invasive whitefly. *Environ. Entomol.* 42, 980–986. doi: 10.1603/EN13071
- Liu, S. S., Colvin, J., and De Barro, P. J. (2012). Species concepts as applied to the whitefly *Bemisia tabaci* systematics: how many species are there? *J. Integr. Agric.* 11, 176–186. doi: 10.1016/S2095-3119(12)60002-1
- Liu, S. S., De Barro, P. J., Xu, J., Luan, J. B., Zang, L. S., Ruan, Y. M., et al. (2007). Asymmetric mating interactions drive widespread invasion and displacement in a whitefly. *Science* 318, 1769–1772. doi: 10.1126/science.1149887

- Lou, Y. P., Zhang, Y. J., Wu, Q. J., Xu, B. Y., and Zhu, G. R. (2005). The analysis on the genetic differentiation of B-biotype *Bemisia tabaci* Populations in different host races. *Agric. Biol. Disaster Prev. Control* 441–447. (Chinese)
- Mansaray, A., and Sundufu, A. J. (2009). Oviposition, development and survivorship of the sweetpotato whitefly *Bemisia tabaci* on soybean, Glycine max, and the garden bean, *Phaseolus vulgaris*. *J. Insect Sci.* 9, 1–6. doi: 10.1023/A:1003215227403
- Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends Genet.* 24, 133–141. doi: 10.1016/j.tig.2007.12.007
- Moon, J., Salzman, R. A., Ahn, J. E., Koiwa, H., and Zhu-Salzman, K. (2004). Transcriptional regulation in cowpea bruchid guts during adaptation to a plant defense protease inhibitor. *Insect Mol. Biol.* 13, 283–291. doi: 10.1111/j.0962-1075.2004.00485.x
- Musil, D., Zucic, D., Turk, D., Engh, R. A., Mayr, I., Huber, R., et al. (1991). The refined 2.15 Å X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity. *EMBO J.* 10, 2321–2330.
- Nosil, P., Crespi, B. J., and Sandoval, C. P. (2002). Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature* 417, 440–443. doi: 10.1038/417440a
- Oliveira, M. R. V., Henneberry, T. J., and Anderson, P. (2001). History, current status and collaborative research projects for *Bemisia tabaci*. *Crop Prot.* 20, 709–723. doi: 10.1016/S0261-2194(01)00108-9
- Pan, H. P., Chu, D., Liu, B. M., Xie, W., Wang, S. L., Wu, Q. J., et al. (2013). Relative amount of symbionts in insect hosts changes with host-plant adaptation and insecticide resistance. *Environ. Entomol.* 42, 74–78. doi: 10.1603/EN12114
- Pan, H. P., Li, X. C., and Zhang, Y. J. (2012). Sex affects the infection frequencies of symbionts in *Bemisia tabaci*. *Commun. Integr. Biol.* 5, 337–339. doi: 10.4161/cib.20398
- Perring, T. M. (2001). The *Bemisia tabaci* species complex. *Crop Prot.* 20, 725–737. doi: 10.1016/S0261-2194(01)00109-0
- Rieux, A., Gras, S., Lecaillon, F., Niepceon, A., Katrib, M., Smith, N. C., et al. (2012). Eimeripain, a cathepsin B-like cysteine protease, expressed throughout sporulation of the apicomplexan parasite *Eimeria tenella*. *PLoS ONE* 7:e31914. doi: 10.1371/journal.pone.0031914
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). EdgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Rundle, H. D., and Nosil, P. (2005). Ecological speciation. *Ecol. Lett.* 8, 336–352. doi: 10.1111/j.1461-0248.2004.00715.x
- Saleh, D., Laarif, A., Clouet, C., and Gauthier, N. (2012). Spatial and host-plant partitioning between coexisting *Bemisia tabaci* cryptic species in Tunisia. *Popul. Ecol.* 54, 261–274. doi: 10.1007/s10144-012-0303-z
- Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science* 323, 737–741. doi: 10.1126/science.1160006
- Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nat. Methods* 5, 16–18. doi: 10.1038/NMETH1156
- Shi, J., Xi, H., Wang, Y., Zhang, C., Jiang, Z., Zhang, K., et al. (2003). Divergence of the genes on human chromosome 21 between human and other hominoids and variation of substitution rates among transcription units. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8331–8336. doi: 10.1073/pnas.1332748100
- Siricoon, S., Grams, S. V., and Grams, R. (2012). Efficient inhibition of cathepsin B by a secreted type 1 cystatin of *Fasciola gigantica*. *Mol. Biochem. Parasitol.* 186, 126–133. doi: 10.1016/j.molbiopara.2012.10.003
- Stireman, J. O., Nason, J. D., and Heard, S. B. (2005). Host-associated genetic differentiation in phytophagous insects: general phenomenon or isolated exceptions? Evidence from a goldenrod insect community. *Evolution* 59, 2573–2587. doi: 10.1111/j.0014-3820.2005.tb00970.x
- Storey, J. D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U.S.A.* 100, 9440–9445. doi: 10.1073/pnas.1530509100
- Tiffin, P., and Hahn, M. W. (2002). Coding sequence divergence between two closely related plant species: *Arabidopsis thaliana* and *Brassica rapa* ssp. *pekinensis*. *J. Mol. Evol.* 54, 746–753. doi: 10.1007/s00239-001-0074-1
- Tsueda, H., and Tsuchida, K. (2011). Reproductive differences between Q and B whiteflies, *Bemisia tabaci*, on three host plants and negative interactions in mixed cohorts. *Entomol. Exp. Appl.* 141, 197–207. doi: 10.1111/j.1570-7458.2011.01189.x
- Wang, L., Feng, Z., Wang, X., Wang, X., and Zhang, X. (2010). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26, 136–138. doi: 10.1093/bioinformatics/btp612
- Wang, X. W., Luan, J. B., Li, J. M., Su, Y. L., Xia, J., and Liu, S. S. (2011). Transcriptome analysis and comparison reveal divergence between two invasive whitefly cryptic species. *BMC Genomics* 12:458. doi: 10.1186/1471-2164-12-458
- Wang, X. W., Zhao, Q. Y., Luan, J. B., Wang, Y. J., Yan, G. H., and Liu, S. S. (2012). Analysis of a native whitefly transcriptome and its sequence divergence with two invasive whitefly species. *BMC Genomics* 13:529. doi: 10.1186/1471-2164-13-529
- Xie, W., Guo, L. T., Jiao, X. G., Yang, N. N., Yang, X., Wu, Q. J., et al. (2014). Transcriptomic dissection of sexual differences in *Bemisia tabaci*, an invasive agricultural pest worldwide. *Sci. Rep.* 4:4088. doi: 10.1038/srep04088
- Xie, W., Wang, S. L., Wu, Q. J., Feng, Y. T., Pan, P. H., Jiao, X. G., et al. (2011). Induction effects of host plants on insecticide susceptibility and detoxification enzymes of *Bemisia tabaci* (Hemiptera: Aleyrodidae). *Pest Manag. Sci.* 67, 87–93. doi: 10.1002/ps.2037
- Yang, Z., and Nielsen, R. (2000). Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.* 17, 32–43. doi: 10.1093/oxfordjournals.molbev.a026236
- Zang, L. S., Chen, W. Q., and Liu, S. S. (2006). Comparison of performance on different host plants between the B biotype and a non-B biotype of *Bemisia tabaci* from Zhejiang, China. *Entomol. Exp. Appl.* 121, 221–227. doi: 10.1111/j.1570-8703.2006.00482.x
- Zhang, L. P., Zhang, Y. J., Zhang, W. J., Wu, Q. J., Xu, B. Y., and Chu, D. (2005). Analysis of genetic diversity among different geographical populations and determination of biotypes of *Bemisia tabaci* in China. *J. Appl. Entomol.* 129, 121–128. doi: 10.1111/j.1439-0418.2005.00950.x
- Zhang, Z., Li, J., Zhao, X. Q., Wang, J., Wang, G. K., and Yu, J. (2006). KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. *Genomics Proteomics Bioinformatics* 4, 259–263. doi: 10.1016/S1672-0229(07)60007-2

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 October 2014; accepted: 26 November 2014; published online: 10 December 2014.

Citation: Xie W, Wu Q, Wang S, Jiao X, Guo L, Zhou X and Zhang Y (2014) Transcriptome analysis of host-associated differentiation in *Bemisia tabaci* (Hemiptera: Aleyrodidae). *Front. Physiol.* 5:487. doi: 10.3389/fphys.2014.00487
This article was submitted to *Integrative Physiology*, a section of the journal *Frontiers in Physiology*.

Copyright © 2014 Xie, Wu, Wang, Jiao, Guo, Zhou and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.