Contents lists available at ScienceDirect

Gene



journal homepage: www.elsevier.com/locate/gene



Research paper

Insect and plant-derived miRNAs in greenbug (*Schizaphis graminum*) and yellow sugarcane aphid (*Sipha flava*) revealed by deep sequencing



Haichuan Wang ^{a,*}, Chi Zhang ^b, Yongchao Dou ^b, Bin Yu ^b, Yunfeng Liu ^b, Tiffany M. Heng-Moss ^c, Guoqing Lu ^d, Michael Wachholtz ^d, Jeffery D. Bradshaw ^c, Paul Twigg ^e, Erin Scully ^f, Nathan Palmer ^g, Gautam Sarath ^g

^a Department of Agronomy and Horticulture, University of Nebraska Lincoln, Lincoln, NE 68583, United States

^b School of Biological Sciences, University of Nebraska Lincoln, Lincoln, NE 68503, United States

^c Department of Entomology, University of Nebraska Lincoln, Lincoln, NE 68583, United States

^d Department of Biology, University of Nebraska Omaha, Omaha, NE 68182, United States

^e Department of Biology, University of Nebraska Kearney, Kearney, NE 68849, United States

^f USDA-ARS, Stored Product Inset and Engineering Research Unit, Center for Grain and Animal Health Research, Manhattan, KS 66502, United States

^g USDA-ARS, Wheat, Sorghum, and Forage Research Unit, Lincoln, NE 68583, United States

ARTICLE INFO

Article history: Received 7 September 2016 Received in revised form 2 November 2016 Accepted 7 November 2016 Available online 9 November 2016

Keywords: Next generation sequencing mirDeep2 miRNAs Cereal-aphids Sorghum Barley

ABSTRACT

Schizaphis graminum (green bug; GB) and *Sipha flava* (yellow sugarcane aphid; YSA) are two cereal aphid species with broad host ranges capable of establishing on sorghum (*Sorghum bicolor*) and several switchgrass (*Panicum virgatum*) cultivars. Switchgrass and sorghum are staple renewable bioenergy crops that are vulnerable to damage by aphids, therefore, identifying novel targets to control aphids has the potential to drastically improve yields and reduce losses in these bioenergy crops. Despite the wealth of genomic and transcriptomic information available from a closely related model aphid species, the pea aphid (*Acyrthosiphon pisum*), similar genomic information, including the identification of small RNAs, is still limited for GB and YSA. Deep sequencing of miRNAs expressed in GB and YSA was conducted and 72 and 56 miRNA candidates (including 14 and eight novel) were identified, respectively. Of the identified miRNAs, 45 were commonly expressed in both aphid species. Further, plant derived miRNAs were also detected in both aphid samples, including 13 (eight known and five novel) sorghum miRNAs and three (novel) barley miRNAs. In addition, potential aphid gene targets for the host plant-derived miRNAs in aphids will ultimately lead to a better understanding of the role of miRNAs in regulating gene expression networks in these two aphids and the potential roles of plant miRNAs in mediating plant-insect interactions.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

MicroRNAs (miRNAs) are a class of small (~22 nucleotide) non-coding RNAs that regulate gene expression in animals, plants, and viruses. MiRNAs were originally discovered in *Caenorhabditis elegans* in 1993 (Lee et al., 1993), and thereafter, many more miRNAs were identified in other organisms, such as insects (Asgari and Sullivan, 2010; Asgari, 2013; Ge et al., 2013; Lucas and Raikhel, 2013) and plants (Jones-Rhoades et al., 2006; Kettles et al., 2013). MiRNAs are known to play important roles in the regulation of gene expression networks (Shukla et al., 2008; Cai et al., 2009) and are involved in a variety of developmental processes (Aravin et al., 2003; Chawla and Sokol, 2011;

 Corresponding author at: Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, 68583-0915, United States.

vedraska-Lincoln, Lincoln, NE, 68583-0915, United S

Mead et al., 2012). Oral feeding of synthetic miRNA mimics in Helicoverpa armigera induced a 40% increase in larval mortality and a 70% reduction in fecundity, indicating the potential use of miRNAs as agents in integrated pest insect management programs (Jayachandran et al., 2013). Further, ingestion of plant-derived miRNAs by aphids and other phytophagous insects during feeding has the potential to modulate expression of insect genes, indicating that plants may produce miRNAs that negatively impact insect fitness. While the precise targets of these plant-produced miRNAs have yet to be identified in many cases, evidence for the presence of plant miRNAs in animal small RNA datasets, including those of the pea aphid (Acyrthosiphon pisum), have been previously reported (Zhang et al., 2012b), confirming that they are ingested during feeding. Additionally, other studies have revealed that exogenous plant-derived miRNA can impact expression of genes and proteins involved herbivore metabolism, revealing the importance of exogenous miRNAs in modulating herbivore-plant interactions and co-evolution (Vaucheret and Chupeau, 2012; Zhang et al., 2012a). A

Abbreviations: GB, green bug; YSA, yellow sugar cane aphid; miRNAs, MicroRNAs.

recent study (Guo et al., 2014) demonstrated that plant-generated artificial small RNAs produced by transgenic *Nicotiana tabacum* cv. Xanti (tobacco) could significantly reduce the expression of the central nervous system enzyme acetylcholinesterase 2 (AchE2) in *Myzus persicae* (green peach aphid). Although these data suggest a putative role for miRNAs in mediating insect-plant interactions, the impacts of endogenous plant miRNAs expressed at physiological levels on herbivore gene expression networks and the precise roles of many plant-derived miRNAs in regulating herbivore gene expression still remain unclear at a functional level.

The inventory of miRNAs expressed by insects is aided by the availability of annotated reference genomes (Asgari and Sullivan, 2010; Asgari, 2013; Etebari and Asgari, 2014; Rebijith et al., 2014; Calla and Geib, 2015). However, miRNAs are generally conserved among closely related insects (Willmann and Poethig, 2007; Nozawa et al., 2010), suggesting that reference genomes from close relatives could be exploited to annotate and predict miRNAs in insects lacking reference genomes. For example, Ge et al. (2013) were able to successfully identify 97 and 91 miRNAs in H. armigera and Spodoptera litura, respectively, using genomic information from a closely related model insect species, Bombyx mori (Ge et al., 2013). The greenbug, Schizaphis graminum (Rondani) (GB), and the yellow sugarcane aphid, Sipha flava (Forbes) (YSA), are major insect pests of many cereal crops. More than 70 graminaceous plant species have been reported as hosts for the GB worldwide and large scale agricultural losses have been reported (Blackman and Eastop, 1984). More recently, Koch et al. (2014) reported that among the four important cereal aphid species, both GB and YSA were able to establish on tetraploid switchgrass (Panicum virgatum L.) populations, indicating their potential as a threat to the sustainable production of switchgrass as a bioenergy crop.

Deep sequencing of small RNAs (for species without draft genome sequences) and computational prediction (for species with sequenced genomes) are two major approaches that have been previously used to identify miRNAs in a number of insects (Mendes et al., 2009), for example, S. litura and H. armigera (Ge et al., 2013), B. mori (Jagadeeswaran et al., 2010; Liu et al., 2010), Mayetiola destructor (Khajuria et al., 2013), and A. pisum (Sathyamurthy and Swamy, 2009). In this study, small RNAs (sRNAs) of 17-25 nt in size expressed in both GB feeding on sorghum and YSA feeding on barley were targeted for next-generation sequencing analysis. High quality reads obtained from GB and YSA were mapped to the reference genomes of pea aphid (A. pisum) and their respective plant hosts to identify potential conserved and novel miRNA candidates. The inventory of miRNAs expressed in these two aphid species while feeding on preferred host plants, the detection of plant-derived miRNAs in aphid samples, and the prediction of potential insect target genes of plant-derived miRNAs can be useful in the future to develop new control strategies, such as RNAi.

2. Material and methods

2.1. Aphid species

GBs and YSAs were maintained on preferred hosts, sorghum (*Sorghum bicolor*) and barley (*Hordeum vulgare* L.), respectively. Plants were maintained in growth chambers set at 26 ± 1 °C and $60 \pm 10\%$ relative humidity with photoperiods of 12:12 h (L:D). Approximately 150 living nymphs of each aphid species were collected using a soft eyebrow brush in a 1.5 ml centrifuge tube. Samples were snap frozen in liquid nitrogen and stored at -80 °C until RNA extraction. Two replicates of 150 nymphs per aphid species were extracted, sequenced, and used for miRNA identification.

2.2. RNA extraction for miRNA preparation

Total RNA was extracted from each pool of GBs and YSAs (~150 aphids per replicate) with Trizol (Invitrogen, Cat. 15596-026) following

the manufacturer's instructions. The quality and quantity of the RNA samples were evaluated on 1% agarose gels and using a NanoDrop-1000 spectrophotometer (Thermo Fisher Scientific).

Small RNAs were fractionated from the total RNAs on a 16% TBU gel (16% acrylamide and 42% urea in 0.5X TBE buffer) as described by Ren et al. (2012). Briefly, 10 µg total RNA (in ~10 µl nuclease free water) was mixed with 10 µl 2x formamide (98%) loading dye. Samples were heated at 65 °C for 5 min, placed on ice, and briefly centrifuged prior to loading on the gel. An RNA ladder (Zymo Research small-RNA ladder Cat. R1090) was used to determine approximate sizes of the gel-separated RNA. Gels were stained with 1x TBE containing ethidium bromide for 2 min or longer until bands of interest were visible. Gel regions corresponding to an approximate size range between 17 and 30 nt were excised. Small RNAs were extracted from the gel by crushing the gel pieces in 200 µl 0.3 M NaCl followed by incubation overnight at 4 °C. The gel mix was passed through Costar Spin-X 0.25 µm filters (Costa Cat.8163), and small RNAs in the supernatant were ethanol precipitated and finally redissolved in 11 µl nuclease free water (Ambion Cat. AM9932). Purified small RNAs were stored at -80 °C prior to sequencing.

2.3. Small RNA sequencing

The small RNAs recovered from the 16% acrylamide gel were ligated to Illumina TruSeq Small RNA library adaptors and converted to cDNA via RT-PCR according to the protocol provided in TruSeq Small RNA library preparation kits (Illumina, Cat. RS-200-0012). PCR products between 145 and 160 nt, consistent with the lengths of miRNAs (~22 nt) ligated to library adapters and barcodes (~125 nt), were gel purified prior to sequencing. The gel portion containing miRNAs was cut and purified as described by Ren et al. (2012).

All small RNA samples were sequenced on an Illumina/Solexa system (Genome Analyzer II) at the Center for Biotechnology, University of Nebraska at Lincoln, with four samples per lane and 75 bp singleend reads.

2.4. Identification of small RNA sequences

Short reads less than 16 nt were discarded using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html). The miRDeep2 package (Friedlander et al., 2012) was used to identify miRNAs. Briefly, identical/redundant reads were collapsed and quantified based on read abundances, and then the unique reads were mapped to the pea aphid genome downloaded from AphidBase (Legeai et al., 2010) using the miRDeep2 mapper module. The remaining reads that did not map to the aphid genome were mapped to the genome of either barley (http://plants.ensembl.org/Hordeum_vulgare/Info/Index) for YSA or sorghum (http://plants.ensembl.org/Sorghum_bicolor/Info/Index) for GB to identify any plant-derived miRNAs that may have been ingested by the aphid. All known miRNAs and precursor miRNAs used as references in the miRDeep2 quantifier module were downloaded from miRBase (version 21, http://www.mirbase.org/). The miRDeep2 module was used to identify novel miRNAs.

2.5. Potential target gene identification

To identify potential targets of plant-derived miRNAs, the 3' UTRs sequences of pea aphid genes were downloaded from the Ensembl database (http://metazoa.ensembl.org/index.html) and compared to the plant-derived miRNAs identified in GBs and YSAs using miRanda (Enright et al., 2004; John et al., 2004) under default parameters. The entire sequence of aphid transcripts with target matches in miRanda were downloaded from Ensembl and annotations and gene ontology (GO) terms were assigned using Blast2GO (Conesa and Gotz, 2008). GO terms enriched in the predicted aphid targets of the plant-derived miRNAs were identified using the functional enrichment feature (Fisher's exact analysis) of Blast2GO with the complete inventory of annotated genes from the pea aphid (http://www.aphidbase. com) as a reference. In addition, KEGG (Kanehisa et al., 2014) pathway analysis was also conducted using Blast2GO to reveal which aphid metabolic pathways were potentially targeted by host plant miRNAs.

2.6. Validation of miRNA

Validation experiments for miRNAs were conducted with GB and sorghum. Approximately 200 μ g of total RNA from GB and sorghum were prepared using Trizol reagent (Invitrogen) following the manufacturer's protocol. The quality and quantity of total RNAs were evaluated on the 1% agarose gel and using a NanoDrop 1000 (Thermo Scientific) respectively, prior to the enrichment of small RNA. For sRNA enrichment, 50 μ l of 50% PEG8000 and 50 μ l of 5 M NaCl were added to 400 μ l (200 μ g) RNA, mixed well, and incubated on ice for 2 h followed by centrifugation at 14,000 rpm for 10 min. The supernatant was carefully transferred into a new centrifuge tube while 1/10 volume of 3 M sodium acetate and 2.5 volumes of chilled 100% ethanol were then added for precipitation. Samples were stored at -20 °C for 2 h and centrifuged for 20 min at maximum speed. The pellets were washed with 75% ethanol and resuspended in DEPC water for cDNA synthesis.

Four host plant-derived miRNAs, including two novel miRNAs (sbi-miR2-3p and sbi-miR3-5p) and two known miRNAs (sbi-miR2927a-5p and osa-miR390-5p), were selected for validation. PCR primers for each miRNA are shown in Table S1. cDNA synthesis and PCR were conducted as described by Wu et al. (2007). The small RNA of U6 from sorghum was used an endogenous control. PCR products were evaluated on 10% PAGE gels.

3. Results

3.1. Sequencing results

Approximately 3.8, 5.1, 8.9, and 9.6 million reads were generated for the GB1, GB2, YSA1, and YSA2 samples, respectively (SRA accession number: SRP058100) (Table 1). After the identical/redundant reads were quantified and collapsed with miRDeep2, the number of reads was reduced by ~75% in both GB and YSA. The remaining 2,244,213 and 4,738,344 reads from the two GB and two YSA samples were used as input for miRNA identification.

3.2. The length distribution and nucleotide bias analysis

As indicated in Fig. 1, 22 nt miRNAs were dominant in both species. MiRNAs of 20 and 21 nt were slightly greater in abundance in the GB samples compared to the YSA samples which had greater abundances of 23 and 24 nt miRNAs. The frequency of the four nucleotides appearing at each position in the 22 nt miRNAs were also analyzed (Fig. 2). Five positions, 1st, 9th, 10th, 12th, and 20th, were dominated by uracil (U) in GB (Fig. 2A), while a bias for U in the 1st, 9th, 12th,

Table 1

The number of miRNA-sequencing reads generated from GB and YSA samples and mapped to pea aphid hairpin RNAs (SRA accession number: SRP058100).

Species	Raw reads	Reads mapped to pea aphid hairpins
S. graminum (GB)	GB1: 3,809,958	952,489
	GB2: 5,166,897	1,291,724
S. flava (YSA)	YSA1: 8,933,780	2,233,445
	YSA2: 9,619,597	2,504,899



Fig. 1. Length distribution of nucleotides in predicted miRNAs in GB (green) and YSA (yellow).

13th, and 17th positions was observed in YSA (Fig. 2B). In both aphids, the first nucleotide was dominated by U (63% U in GB and 73% in YSA respectively).







3.3. Identification of known and novel miRNA

A total of 72 and 56 reads from GB and YSA were identified as potential miRNAs, respectively (Table 2). Among them, 59 in GB and 53 in YSA had matches to the pea aphid genome and 48 of the miRNAs were common to both aphid species. The remaining 13 sRNAs from GB and three sRNAs from YSA had no matches to the pea aphid genome.

Among the 48 miRNAs common to both GB and YSA, 45 miRNAs had complete "seed" matches with known miRNAs from 43 miRNA families in miRBase (Table 3). Of these 45 miRNAs, only seven (13.7%) had "seed" hits with previously characterized miRNAs from pea aphid (*A. pisum*), while the majority (58.8%) had matches to miRNAs that have been previously characterized in fruit flies (*Drosophila melanogaster*). Additionally, the corresponding star strand (miRNA*) was detected for 31 out of the 48 miRNAs. In addition, fourteen and eight sequences from GB and YSA respectively which mapped to the pea aphid genome had no matches to miRNAs in miRBase, suggesting that they could be novel aphid-derived miRNAs.

Additional miRBase queries for the miRNA sequences that did not map to the pea aphid genome showed (Table 4) that 13 sequences from GB samples mapped to the sorghum genome while three sequences from YSA mapped to barely. Eight of the 13 sorghum sequences (from GB) (Table 2) had "seed" matches with previously identified miRNAs from plants, such as *S. bicolor* (1), *Brachypodium distachyon* (1), and *Oryza sativa* (6). The other five putative miRNA sequences from GB and all three miRNA sequences from YSA did not have matches to any known plant miRNAs.

To support the authenticities and confirm the identities of the sequences that did not have matches in miRBase as novel miRNAs, visual examination of the predicted hairpin was performed. A 100 bp region of each potential miRNA precursor sequence obtained from the pea aphid, sorghum, and barley genomes was folded in silico into secondary structures for miRNAs unique to GB (Fig. 3A), unique to YSA (Fig. 3B), common to both GB and YSA (Fig. 3C), unique to sorghum (Fig. 4A), and unique to barley (Fig. 4B). The hairpin structure was formed based on the complementarity of two regions from a single stranded nucleotide sequence. The region highlighted in red corresponds to the mature miRNA while the region in purple or blue signifies the miRNA*. The orange portion within the loop is formed by the unpaired region.

In total, 19 putative novel miRNAs from GB samples (Fig. 3A and C, Fig. 4A) and 11 from YSA samples (Fig. 3B and C, Fig. 4B) were identified (Table 5). Fourteen of 19 and eight of 11 miRNAs were predicted to originate from GB and YSA, respectively (Table 2), while five miRNAs from GB and three miRNAs from YSA were predicted to originate from the host plants, sorghum and barely, accordingly (Table 4). Of the novel aphid derived miRNAs, three (sgr-miR-12-5p to sfl-miR-6-5p, sgr-miR-13-3p to sfl-miR-7-3p, and sgr-miR-14-5 to sfl-miR-8-5p) were shared between the two aphid species. The miRNA* was also detected for two of the three shared miRNAs. For GB, seven out of 14 novel miRNAs had few to no reads derived from the miRNA*s, while miRNA* was detected in low to no abundance in five of the eight novel miRNAs identified from the YSA samples. The low abundances of the star strand for these sequences support their classification as novel miRNAs (Calla

Table 2

miRNAs identified in GB and YSA.



* represents miRNAs which had homologyto miRBase mature miRNAs (mature.fa).

and Geib, 2015). Additionally, miRNA* was not detected for any of the known or novel plant-derived miRNAs from the GB and YSA samples. The higher expression levels of the mature miRNA compared to the star strand provide further evidence that these novel miRNAs are likely to be genuine.

3.4. Prediction of possible host plant miRNAs targets in aphids

Genome sequences are not currently available for YSA and GB; therefore, the genome of a model aphid species, the pea aphid, was used to search for potential target genes of the plant-derived miRNAs identified in this study. In case of the 13 sorghum-derived candidate miRNAs identified in GB, a total of 16,549 potential targets were identified in the pea aphid genome. Of these targets, 11,172 were from eight of the known miRNAs and the remaining 5377 targets were from the five novel miRNAs. The number of potential targets for each miRNAs varied considerably, ranging from 314 to 4225. Similar comparisons in barley identified a total of 8943 targets for the three novel miRNAs identified in YSA, with at least 2450 genes potentially targeted by each barley miRNA. The large number of putative targets identified is because perfect complementarity of only six nucleotides in the seed region has been previously shown to be sufficient to induce RNA knockdown and/or silencing in mammals (Lewis et al., 2005).

In total, 27 GO terms were enriched in the putative targets of two barley miRNAs (Tables S2–S3) and 185 GO terms were enriched in the putative targets of the four known (Tables S4–S7) and three novel (Tables S8–S10) sorghum miRNAs. No enriched GO terms were detected in the remaining six miRNAs from sorghum and one miRNA from barley. GO term enrichment analysis demonstrated that each sorghum and barely miRNA potentially targeted genes involved in different networks in GB and YSA, respectively, (Tables S2–S10). In contrast, only three GO terms were found to be under-represented in the putative GB and YSA targets of the sorghum and barley-derived miRNAs: GO: 0005198 (related to structural molecule activity) targeted by a novel miRNA (hvumiR-3-3p) from barley, and GO: 0003824 (related to catalytic activity) and GO: 0006259 (related to DNA metabolic process) targeted by known miRNAs (sbi-miR6230-3p and sbi-miR2927b-5p) from sorghum.

Additionally, KEGG metabolic pathway analysis revealed that genes targeted by some host plant-driven miRNAs, such as sbi-miR1-3p, sbi-miR2-3p, sbi-miR3-5p, sbi-miR396-5p, sbi-miR3905p, and hvu-miR2-5p, were likely involved in detoxification, such as metabolism of xenobiotics by P450s (Table 6). Genes involved in digestive physiology were also potentially targeted by plant-derived miRNAs, including several genes participating in starch and sucrose metabolism, which were putative targets of sbi-miR3-5p, sbi-miR390-5p, sbi-miR5076-5p, sbi-miR6230-3p, sbi-miR516-3p3, sbi-miR390-5p, sbi-miR166-3p, hvu-miR2-5p, and hvu-miR3-3p.

3.5. Validation of novel and known miRNAs

The RT-PCR results for the validation of miRNAs are shown in Fig. 5. All four miRNAs tested, two novel miRNAs (sbi-miR1-3p and sbi-miR2-3p) and two known miRNAs (sbi-miR2927a-5p and osa-miR390-5p), were successfully amplified from samples of sorghum and GB separately, essentially confirming their expression in sorghum and their presence in GB samples.

4. Discussion

4.1. miRNAs from GB and YSA

In this study, deep sequencing was applied to identify miRNA candidates in two aphid species, GB and YSA, which both lack genomic information but are important pests of many cultivated grass species (e.g., the bioenergy crops sorghum and switchgrass) (Koch et al., 2014).

Table 3

Known miRNAs sequence identified in GB and YSA.

Matching seed in miRBase	miRNA sequence	Length (nt)	Number of reads (miRNAs/miR	NA*)
bmo-miR-2765	ugguaacuccaaaccauugccgg	23	443/17 ^a	1788/36 ^b
dme-miR-2b-3p	uaucacagccagcuuuguaugagc	24	201/50	11383/22
dme-miR-8-3p	uaauacugucagguaaugaug	21	16886/29	105400/122
dme-miR-iab-4-5p	acguauacuaaauguauccug	21	40/1	49/2
api-miR-184b	uggacggagaacugauaagggc	22	111390/1	52788/2
bmo-miR-278-5p	acggacgaagguuucaucaauc	22	558/133	249/134
dme-miR-281 -2-5p	aagagagcuauccgucgacagu	22	11233/18	4242/63
dme-bantam-3p	ugagaucauugugaaagcuaau	22	101265/213	1160273/49
aga-let-7	ugagguaguugguuguauagu	21	228/2	557/1
ame-miR-190	agauauguuugauauucuugguug	24	197/41	497/22
tca-miR-10-5p	uacccuguagauccgaauuugu	22	451408/2605	595126/1479
dme-miR-277-3p	uaaaugcacuaucugguacgac	22	454/5	1485/6
api-miR-3024	ucuuugggauuuaauagagcc	21	20/0	1486/0
dme-miR-14-3p	ucagucuuuuucucucucu	20	1991/1627	1404/50
dme-miR-9a-5p	ucuuugguacuuuagcuguagg	22	202/8	2827/0
api-miR-3042	gagggcagauuauuucugauacg	23	38/0	52/6
api-miR-3050	ugagaucuugauaaacucgccu	22	1159/41	2844776/7
dme-miR-2b-3p	uaucacagccguuuuugacaauu	23	5232/18	11383/22
ame-miR-1000	auauuguccugucacagcag	20	46/0	441/0
dme-miR-965-3p	uaagcguauagcuuuuccccuu	22	35/7	51/21
dme-miR-929-5p	aaauugacucuaguagggagu	21	52/0	382/1
dme-miR-276a-3p	uaggaacuucauaccgugcucu	22	18615/16	74373/72
dme-miR-285-3p	uagcaccauuugaaaccaguac	22	2242/0	353/0
dme-miR-252-5p	cuaaguacuagugccgcggga	21	479/0	1044/0
dme-miR-9a-5p	ucuuugguauucuagcuguagg	22	27/0	12/1
dme-miR-1-3p	uggaauguaaagaagua	17	112/0	128/0
dme-miR-279-3p	ugacuagauccacacucaucc	21	2038/320	4195/12
tca-miR-71 -3p	ucucacuaucuugucuuucaug	22	370/38	564/77
dme-miR-124-3p	uaaggcacgcggugaaugcc	20	7/3	33/0
dme-miR-275-3p	ucagguaccugaaguagcgcg	21	20/0	24/0
dme-miR-275-3p	ucagguaccaagugauuucuga	22	250/0	326/4
dme-miR-2b-3p	uaucacagccacuuugaugagc	22	1066/1	1145/3
tca-miR-7-5p	uggaagacuagugauuuuguuguu	24	23/0	74/0
api-miR-3032	uguuaguauaacucuuaguaaca	23	60/98	27/10
api-miR-3049	ucgggaaggcagcugcggcggacu	24	422/1	347/1
tca-miR-927b-5p	uuuagaauuccuacgcuuuacc	22	330/286	5210/278
dme-miR-263a-5p	aauggcacugaaagaauucacggg	24	223026/12	529704/15
dme-miR-1002-5p	cuaaguaguagcgccaacgguga	23	16/0	87/1
dme-miR-993-3p	gaagcucgucucuacagguaucu	23	1231/1416	1501/2014
dme-miK-9a-5p	ucuuugguuaucuagcuguauga	23	1137/0	2455/0
dme-miR-92a-3p	uauugcacuugacccagccu	20	187/0	614/5
apı-mıR-3040	cagccgguggugacuguuuccaca	24	438/48	732/310
dme-miR-137-3p	uauugcuugagaauacacguag	22	189/12	249/14
dvi-miR-315b	uuuugauuguugcucagaaagccu	24	6349/0	54754/30
dme-miR-87-3p	gugagcaaaguuucaggugugc	22	197/4	230/63

^a Shown in GB samples.

^b Shown in YSA samples.

Table 4	
Host plant small RNAs (miRNAs) identified in aphid.	

Plant	Aphid	miRNA ID	No. reads*	Match in miRBase	miRNAs No. aphic	l target genes
	GB	sbi-miR5163-3p	3	bdi-miR5163b-3p	cagauauucugaagaag	4225
	GB	sbi-miR166-3p	38	osa-miR166a-3p	ucggaccaggcuucauucccc	314
	GB	sbi-miR390-5p	124	osa-miR390-5p	uagcucaguugguuaga	354
	GB	sbi-miR396-5p	7	osa-miR396e-3p	augguucaacagaagacc	1264
_	GB	sbi-miR5076-5p	5	osa-miR5076	aaaaugggauggucugauu	1460
olor	GB	sbi-miR6230-3p	12	sbi-miR6230-3p	caacaagugacuguauu	2430
GB :		sbi-miR2927a-5p	161	osa-miR2927	cgucgucgucucgccguc	312
(S.	GB	sbi-miR2927b-5p	270	osa-miR2927	cgucgucgcgcucgccagc	813
	GB	sbi-miR1-3p	81	N/A	uuuagguucuggcgccc	1825
	GB	sbi-miR2-3p	2292	N/A	uuuccaccacguucccg	450
	GB	sbi-miR3-5p	19	N/A	auuggaauugcuggauc	2107
	GB	sbi-miR4-5p	14	N/A	cgggagggggggggguc	676
÷.	GB	sbi-miR5-5p	10	N/A	gagccgccgucggugcag	319
gare	YSA	hvu-miR1-5p	24	N/A	cgugggucacagguucgc	2454
,vul	YSA	hvu-miR2-5p	1664	N/A	aaaauagcucaguugguaga	2864
H)	YSA	hvu-miR3-3p	717	N/A	guagauuaauuccugcagc	3625

*total reads mapped.

Novel miRNAs were highlighted in blue box

MiRNA reads were mapped to the pea aphid genome (A. pisum), leading to the identification of 72 and 56 miRNAs in GB and YSA, respectively. High levels of sequence similarity between miRNAs identified in GB and YSA to previously identified miRNAs from other insect species, such as Drosophila melanogastar, Anopheles gambiae, Apis mellifera, B. mori, and Tribolium castaneum (http://www.mirbase.org/), strongly suggests that these could be functional miRNA molecules. The identification of the miRNA* strand for over 70% of the miRNAs detected in this study and the lower abundances of the miRNA* strand compared to the guide strand would further confirm that the discovery pipeline was robust in identifying miRNAs in the GB and YSA datasets. During miRNA processing, a miRNA:miRNA* duplex is produced first, and ultimately the miRNA* star strand is degraded while the miRNA strand (guide strand) guides the RNA induced silencing complex (RISC) to interact with the target RNA (Asgari, 2013).

More candidate miRNAs were identified in GB than in YSA. This difference does not seem to be related to differences in sequencing depth or quality because YSA had more high quality reads (defined as reads with mean quality scores above 30) than GB. However, a slight difference in the proportion of mapped reads was observed between GB and YSA. For instance, nearly 10% more reads from GB could be mapped



Fig. 3. Predicted structures of novel miRNAs detected in GB (A), YSA (B), and both aphids (C). The region highlighted in red corresponds to the mature miRNA while the region in purple or blue signifies the miRNA*. The orange portion within the loop is formed by the unpaired region.

to the pea aphid genome as compared to YSA, indicating that sequence variations between miRNA reads and the reference genome used for mapping can impede miRNA identification. This finding also indicates that certain miRNAs are likely species-specific, even among closely related species.

4.2. Length distribution of miRNAs and miRNA*

The length distribution of miRNAs identified in this study further revealed the diversity and heterogeneity of small RNAs in both aphid species. In miRNAs, the region between the 2nd and 7th nucleotides is known as the "seed" region and plays key roles in the interaction of miRNAs with their targets (Lewis et al., 2003, 2005). A previous analysis of miRNAs identified in *H. armigera* and *S. litura* (Ge et al., 2013) revealed no major over- or under-representation of the U nucleotide in the 'seed' region between positions 2 and 7. No major biases were observed in our study, confirming the conservation of this region in GB and YSA and among other insect species in general.

Except for the mature miRNAs identified, the identification of corresponding star strand (miRNA*) was detected (Table 3), indicating that these miRNA* might be afford potential opportunities contributing to the regulation of networks of genes (Guo and Lu, 2010).



Fig. 4. Predicted structures of novel miRNAs detected from sorghum (A) and barley (B) host-plants. Other details are as described in Fig. 3.

Table 5

Sequences and related information for novel miRNA from greenbug (GB) and yellow sugarcane aphid (YSA).

Species	miRNA ID	Sequence Length (nt) Number of reads (miRNAs/miRNAs*)			
S. graminum (GB)	sgr-miR-1-5p ^a	augggugugugucuagugacaug	23	661/698 ^a	
	sgr-miR-2-5p ^a	ggggauaauuguaggacuuacu	22	4558/27	
	sgr-miR-3-5p ^{a,b}	aagguaacuuauucacugugc	21	32/0	
	sgr-miR-4-5p ^a	aggauugaguagggacgucaac uccaaacgcaauucuug	22	11/5	
	sgr-miR-5-5p ^a	gguuucuugaugacugggacc uucguugucgaugaaaccuugaua	21	104/45	
	sgr-miR-6-5p ^{a,b}	uucggcguaacaugaugugugc acacacguguuaugccgcaag	22	1297/0	
	sgr-miR-7-3p ^{a,b}	ccagacgaacucuguuacccc gguaacagguuuguuuugu	21	181/0	
	sgr-miR-8-3p ^a	ugaguugauguucuuguugcu ggaauaaggacauuaguugcacu	21	15/1	
	sgr-miR-9-3p ^{a,b}	acagcaaagugaaagagacuga ugucuuuuuccgcuuugcug	22	181/79	
	sgr-miR-10-3p ^{a,b}	uuuacccaauauauggcuauau guauccauguuuuggguauaca	22	10/0	
	sgr-miR-11-3p ^a	cggagagagauaacaggcuuuggcu ucagucuuuuucucucucu	25	1425/877	
S. flava (YSA)	sfl-miR-1-5p ^b	uuuggagcuuguauuac gauacaguucggaaau	17	22/0 ^b	
	sfl-miR-2-3p ^b	ucacuggugacaauucaugaag ccaugaauuguuucuaguuggu	22	28/0	
	sfl-miR-3-5p ^{a,b}	acggguguagucuuuagugcacg ugacuagaguuacacucguca	23	211/139	
	sfl-miR-4-5p ^{a,b}	gaggaaaaguuauaacgguuaug uaagcguauagcuuuuccccuu	23	48/49	
	sfl-miR-5-5p ^{a,b}	uuacucacucaauuuggaugug caaccuccuugagugagcgac	22	237/4	
S. graminum & S. flava	sfl-miR-6-5p ^a (sgr-miR-12-5p ^b)	cgccugaaacuugcuuugaccu	22	357/64 ^a	937/84 ^b
	$(sgr-miR-13-3n^{a,b})$	uaggccggcggaaacuacuagc	22	654/0	12/0
	(sgr-miR-14-5p ^a)	gcacuggaagaauucacaga gugagucuccuggugucaacua	20	2515/4	3916/7

^a Shown in GB samples.

^b For YSA samples.

4.3. Plant-derived miRNAs

Plant-derived miRNAs have been previously detected in several organisms including insects (Griffiths-Jones et al., 2008; Zhang et al., 2012b). Although these could be "contamination" derived from the plants consumed by these organisms, exogenous dietary plant miRNAs have been found in herbivores (Witwer et al., 2013) that are capable of regulating gene expression in mammals (Vaucheret and Chupeau, 2012; Zhang et al., 2012a). In this study we identified host plant-derived miRNAs in aphids, which provides additional evidence for the existence of consumed exogenous miRNAs in herbivores. Several GO terms were commonly enriched in the putative target genes of barley and sorghum miRNAs, suggesting that they may also target common genes and metabolic pathways. For example, the putative common aphid gene targets of the sorghum and barely miRNAs consist of genes involved in cell signaling/cell-cell communications and central nervous system (CNS) development, indicating that plants may be capable of disrupting or modifying signal transduction cascades and interfering with the CNS to negatively impact aphid fitness. Other common predicted targets of plant-derived miRNAs included genes involved in cell development, morphogenesis, and reproduction.

Both aphids also feed predominately on phloem sap, and similarities and dissimilarities in phloem contents of different plants can be expected. Potentially, this could explain the findings of similar and unique

Table 6

The metabolic pathways of detoxification and nutrient utilization involving the target genes of host plant miRNAs.

miRNA ID	KEGG pathway
sbi-miR1-3p, sbi-miR2-3p and sbi-miR3-5p, sbi-miR390-5p, sbi-miR-396-5p and hvu-miR-3-3p	Drug-metabolism-other enzymes (00983*)
sbi-miR1-3p, sbi-miR2-3p and sbi-miR3-5p and sbi-miR390-5p	Drug metabolism-P450 (00982*)
sbi-miR1-3p, sbi-miR2-3p and sbi-miR3-5p and sbi-miR396-5p	Metabolism of xenobiotics by P450 (00980*)
hvu-miR2-5p	Streptomycin biosynthesis (00521*)
sbi-miR5-5p, sbi-miR6230-3p, sbi-miR5163-3p, sbi-miR396-5p, sbi-miR2927-5p, hvu-miR2-5p and hvu-miR3-3p	Nicotinate & nicotinamide metabolism (00760*)
sbi-miR5-5p, sbi-miR6230-3p, sbi-miR50765p, sbi-miR6230-3p, sbi-miR51633p, sbi-miR390-5p, sbi-miR166-3p,	Butanoate (Butyrate) metabolism (00650*)
hvu-miR2-5p and hvu-miR3-3p	Starch & sucrose metabolism (00500*)



Fig. 5. Validation of plant-derived miRNAs in sorghum and GB samples by RT-PCR.

plant-derived miRNA in GB and YSA samples. However, these initial findings will need to be validated in broader trials to more fully characterize aphid-plant interactions at the miRNA level.

4.4. Validation of miRNAs

Validation of a subset of the miRNAs via RT-PCR provided evidence that the miRNA identified through deep sequencing were expressed by sorghum and present in GBs, confirming that these plant-derived miRNAs are likely encountered and ingested during feedings.

4.5. Potential interaction between aphids and plant-derived miRNAs at molecular level

In this study, about 16,549 and 8943 genes with miRNAs (novel plant miRNA only) target sites in GB in YSA respectively were detected. These genes appeared to be involved in many important phases of insect growth and development, including drug-metabolism (Table 6). However, the impacts of exogenous plant miRNAs expressed at physiological levels on herbivore gene expression networks and precise roles of many plant-derived miRNAs in regulating of herbivore gene expression still needs to be clarified at a functional level. Nevertheless, finding of these plant miRNAs suggest that it could be possible to exploit plant-derived miRNAs to defend against these two aphid pests in the future. As examples, it has been demonstrated that target genes such as AchE2 in aphid (*M. persicaeaphids*) (Guo et al., 2014) and a trypsin-like gene in cotton bollworm (Jayachandran et al., 2013) can be negatively regulated once exposed to *Nicotiana tabacum* cv. Xanti (tobacco) expressing artificial-small RNAs and miRNA mimics respectively.

Whereas many reads could not be aligned to the pea aphid genome, suggesting the possibility that additional miRNAs remains undiscovered and could be revealed once genome and transcriptome information becomes available for these two aphid species. The identification of novel miRNAs unique to either GB or YSA or shared between these two aphid species could lead to the future identification of miRNAs that could serve as targets to control cereal aphids (Witkos et al., 2011).

Supplementary data to this article can be found online at doi:10. 1016/j.gene.2016.11.014.

Role of the funding source

This research project was funded by the United States Department of Agriculture-National Institute of Food and Agriculture [2011-67009-30096 and USDA-ARS CRIS project 5440-21000-030-00D]. The U.S. Department of Agriculture, Agricultural Research Service, is an equal opportunity/affirmative action employer and all agency services are available without discrimination. Mention of commercial products and organizations in this manuscript is solely to provide specific information. It does not constitute endorsement by USDA-ARS over other products and organizations not mentioned.

Contributors

H. Wang conducted experiment design and performed experiment. H. Wang, M. Wachholtz, Y. Dou and Y. Liu analyzed data. H. Wang, C. Zhang, B. Yu, Tiffany M. Heng-Moss, G. Lu, J. Bradshaw, P. Twigg, E. Scully, N. Palmer and G Sarath contributed to the manuscript preparation.

Declaration of interest

The authors state they have no conflict of interest. All authors certify that they have no conflicts of research interests based on the subject matter or materials discussed in this manuscript.

Acknowledgments

We thank Natalie Matz for her help in RNA and miRNA preparations and Kyle Koch for his assistance in collecting aphid samples.

References

- Aravin, A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., Gaasterland, T., Meyer, J., Tuschl, T., 2003. The small RNA profile during *Drosophila melanogaster* development. Dev. Cell 5, 337–350.
- Asgari, S., 2013. MicroRNA functions in insects. Insect Biochem. Mol. Biol. 43, 388–397.
- Asgari, S., Sullivan, C.S., 2010. Role of microRNAs as regulators of host-virus interactions. Insect Virology 347–365.
- Blackman, R.L., Eastop, V.F., 1984. Aphids on the World's Crops. John Wiley & Sons, New York .
- Cai, Q.N., Han, Y., Cao, Y.Z., Hu, Y., Zhao, X., Bi, J.L., 2009. Detoxification of gramine by the cereal aphid Sitobion avenae. J. Chem. Ecol. 35, 320–325.
- Calla, B., Geib, S.M., 2015. MicroRNAs in the oriental fruit fly, Bactrocera dorsalis: extending Drosophilid miRNA conservation to the Tephritidae. BMC Genomics 16.
- Chawla, G., Sokol, N.S., 2011. Micrornas in Drosophila development. Int. Rev. Cell Mol. Biol. 286 (286), 1–65.
- Conesa, A., Gotz, S., 2008. Blast2GO: a comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics 2008, 619832.
- Enright, A.J., John, B., Gaul, U., Tuschl, T., Sander, C., Marks, D.S., 2004. MicroRNA targets in Drosophila. Genome Biol. 5.
- Etebari, K., Asgari, S., 2014. Accuracy of microRNA discovery pipelines in non-model organisms using closely related species genomes. PLoS One 9.
- Friedlander, M.R., Mackowiak, S.D., Li, N., Chen, W., Rajewsky, N., 2012. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res. 40, 37–52.
- Ge, X., Zhang, Y., Jiang, J.H., Zhong, Y., Yang, X.N., Li, Z.Q., Huang, Y.P., Tan, A.J., 2013. Identification of microRNAs in *Helicoverpa armigera* and *Spodoptera litura* based on deep sequencing and homology analysis. Int. J. Biol. Sci. 9, 1–15.
- Griffiths-Jones, S., Saini, H.K., van Dongen, S., Enright, A.J., 2008. miRBase: tools for microRNA genomics. Nucleic Acids Res. 36, D154–D158.
- Guo, L., Lu, Z.H., 2010. The fate of miRNA* strand through evolutionary analysis: implication for degradation as merely carrier strand or potential regulatory molecule? PLoS One 5.
- Guo, H., Song, X., Wang, G., Yang, K., Wang, Y., Niu, L., Chen, X., Fang, R., 2014. Plant-generated artificial small RNAs mediated aphid resistance. PLoS One 9, e97410.
- Jagadeeswaran, G., Zheng, Y., Sumathipala, N., Jiang, H.B., Arrese, E.L., Soulages, J.L., Zhang, W.X., Sunkar, R., 2010. Deep sequencing of small RNA libraries reveals dynamic regulation of conserved and novel microRNAs and microRNA-stars during silkworm development. BMC Genomics 11.
- Jayachandran, B., Hussain, M., Asgari, S., 2013. An insect trypsin-like serine protease as a target of microRNA: utilization of microRNA mimics and inhibitors by oral feeding. Insect Biochem. Mol. Biol. 43, 398–406.
- John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., Marks, D.S., 2004. Human microRNA targets. PLoS Biol. 2, 1862–1879.
- Jones-Rhoades, M.W., Bartel, D.P., Bartel, B., 2006. MicroRNAs and their regulatory roles in plants. Annu. Rev. Plant Biol. 57, 19–53.
- Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M., 2014. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 42, D199–D205.
- Kettles, G.J., Drurey, C., Schoonbeek, H.J., Maule, A.J., Hogenhout, S.A., 2013. Resistance of *Arabidopsis thaliana* to the green peach aphid, *Myzus persicae*, involves camalexin and is regulated by microRNAs. New Phytol. 198, 1178–1190.
- Khajuria, C., Williams, C.E., El Bouhssini, M., Whitworth, R.J., Richards, S., Stuart, J.J., Chen, M.S., 2013. Deep sequencing and genome-wide analysis reveals the expansion of MicroRNA genes in the gall midge *Mayetiola destructor*. BMC Genomics 14, 187.

- Koch, K.G., Fithian, R., Heng-Moss, T.M., Bradshaw, J.D., Sarath, G., Spilker, C., 2014. Evaluation of tetraploid switchgrass (Poales: Poaceae) populations for host suitability and differential resistance to four cereal aphids. J. Econ. Entomol. 107, 424–431.
- Lee, R.C., Feinbaum, R.L., Ambros, V., 1993. The C-elegans heterochronic gene Lin-4 encodes small Rnas with antisense complementarity to Lin-14. Cell 75, 843–854.
- Legeai, F., Shigenobu, S., Gauthier, J.P., Colbourne, J., Rispe, C., Collin, O., Richards, S., Wilson, A.C.C., Murphy, T., Tagu, D., 2010. AphidBase: a centralized bioinformatic resource for annotation of the pea aphid genome. Insect Mol. Biol. 19, 5–12.
- Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., Burge, C.B., 2003. Prediction of mammalian microRNA targets. Cell 115, 787–798.
- Lewis, B.P., Burge, C.B., Bartel, D.P., 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120, 15–20.
- Liu, S., Li, D., Li, Q., Zhao, P., Xiang, Z., Xia, Q., 2010. MicroRNAs of *Bombyx mori* identified by Solexa sequencing. BMC Genomics 11, 148.
- Lucas, K., Raikhel, A.S., 2013. Insect MicroRNAs: biogenesis, expression profiling and biological functions. Insect Biochem. Mol. Biol. 43, 24–38.
- Mead, E.A., Li, M., Tu, Z.J., Zhu, J.S., 2012. Translational regulation of Anopheles gambiae mRNAs in the midgut during Plasmodium falciparum infection. BMC Genomics 13.
- Mendes, N.D., Freitas, A.T., Sagot, M.F., 2009. Current tools for the identification of miRNA genes and their targets. Nucleic Acids Res. 37, 2419–2433.
- Nozawa, M., Miura, S., Nei, M., 2010. Origins and evolution of microRNA genes in Drosophila species. Genome Biol Evol 2, 180–189.
- Rebijith, K.B., Asokan, R., Krishna, V., Hande, R.H., Kumar, N.K.K., Ramamurthy, V.V., 2014. In silico prediction and characterization of microRNAs from *Aphis gossypii* (Hemiptera: Aphididae). Ann. Entomol. Soc. Am. 107, 521–531.
- Ren, G.D., Xie, M., Dou, Y.C., Zhang, S.X., Zhang, C., Yu, B., 2012. Regulation of miRNA abundance by RNA binding protein TOUGH in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. 109, 12817–12821.

- Sathyamurthy, G., Swamy, N.R., 2009. Computational identification of microRNA homologs from Acyrthosiphon pisum (pea aphid). J. Comput. Intell. Bioinforma. 2, 109.
- Shukla, L.I., Chinnusamy, V., Sunkar, R., 2008. The role of microRNAs and other endogenous small RNAs in plant stress responses. Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms 1779, 743–748.
- Vaucheret, H., Chupeau, Y., 2012. Ingested plant miRNAs regulate gene expression in animals. Cell Res. 22, 3–5.
- Willmann, M.R., Poethig, R.S., 2007. Conservation and evolution of miRNA regulatory programs in plant development. Curr. Opin. Plant Biol. 10, 503–511.
- Witkos, T.M., Koscianska, E., Krzyzosiak, W.J., 2011. Practical aspects of microRNA target prediction. Curr. Mol. Med. 11, 93–109.
- Witwer, K.W., McAlexander, M.A., Queen, S.E., Adams, R.J., 2013. Real-time quantitative PCR and droplet digital PCR for plant miRNAs in mammalian blood provide little evidence for general uptake of dietary miRNAs limited evidence for general uptake of dietary plant xenomiRs. RNA Biol. 10, 1080–1086.
- Wu, R.M., Wood, M., Thrush, A., Walton, E., Varkonyi-Gasic, E., 2007. Real-Time PCR Quantification of Plant miRNAs Using. Universal ProbeLibrary Technology. Biochmica .
- Zhang, L., Hou, D.X., Chen, X., Li, D.H., Zhu, L.Y., Zhang, Y.J., Li, J., Bian, Z., Liang, X.Y., Cai, X., Yin, Y., Wang, C., Zhang, T.F., Zhu, D.H., Zhang, D.M., Xu, J., Chen, Q., Ba, Y., Liu, J., Wang, Q., Chen, J.Q., Wang, J., Wang, M., Zhang, Q.P., Zhang, J.F., Zen, K., Zhang, C.Y., 2012a. Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA (vol 22, pg 273, 2012). Cell Res. 22, 273–274.
- Zhang, Y., Wiggins, B.E., Lawrence, C., Petrick, J., Ivashuta, S., Heck, G., 2012b. Analysis of plant-derived miRNAs in animal small RNA datasets. BMC Genomics 13, 381.