# UC Riverside
## UC Riverside Previously Published Works

**Title**

Key Transport and Ammonia Recycling Genes Involved in Aphid Symbiosis Respond to Host-Plant Specialization

**Permalink**

https://escholarship.org/uc/item/6x83q9br

**Journal**

G3: Genes, Genomes, Genetics, 8(7)

**ISSN**

2160-1836

**Authors**

Kim, Dohyup
Minhas, Bushra F
Li-Byarlay, Hongmei
et al.

**Publication Date**

2018-07-01

**DOI**

10.1534/g3.118.200297

Peer reviewed

# Key Transport and Ammonia Recycling Genes Involved in Aphid Symbiosis Respond to Host-Plant Specialization

Dohyup Kim,* Bushra F. Minhas,[†] Hongmei Li-Byarlay,[‡] and Allison K. Hansen*,[1]

*Department of Entomology, University of California, Riverside, Riverside, CA 92501, [†]Illinois Informatics Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801, and [‡]Department of Natural Sciences, Central State University, Wilberforce, OH 45384

ORCID ID: 0000-0003-0718-4247 (D.K.)

**ABSTRACT** Microbes are known to influence insect-plant interactions; however, it is unclear if host-plant diet influences the regulation of nutritional insect symbioses. The pea aphid, *Acyrthosiphon pisum*, requires its nutritional endosymbiont, *Buchnera*, for the production of essential amino acids. We hypothesize that key aphid genes that regulate the nutritional symbioses respond to host-plant diet when aphids feed on a specialized (alfalfa) compared to a universal host-plant diet (fava), which vary in amino acid profiles. Using RNA-Seq and whole genome bisulfite sequencing, we measured gene expression and DNA methylation profiles for such genes when aphids fed on either their specialized or universal host-plant diets. Our results reveal that when aphids feed on their specialized host-plant they significantly up-regulate and/or hypo-methylate key aphid genes in bacteriocytes related to the amino acid metabolism, including glutamine synthetase in the GOGAT cycle that recycles ammonia into glutamine and the glutamine transporter *ApGLNT1*. Moreover, regardless of what host-plant aphids feed on we observed significant up-regulation and differential methylation of key genes involved in the amino acid metabolism and the glycine/serine metabolism, a metabolic program observed in proliferating cancer cells potentially to combat oxidative stress. Based on our results, we suggest that this regulatory response of key symbiosis genes in bacteriocytes allows aphids to feed on a suboptimal host-plant that they specialize on.

When an organism symbiotically lives inside another organism's cells its cellular metabolic processes often become integrated with its hosts'. Archetypes of these ancient cellular integration events are readily observed in eukaryotic cells as mitochondria and plastids (Dyall *et al.* 2004). Such organelles exhibit complex regulatory mechanisms that control all aspects of cellular processes such as cell division, transport,

and metabolism. Similar to the regulation of organelles, simultaneous inter-domain crosstalk of animal host and bacteria exists between host cells and microbial endosymbionts (Zientz *et al.* 2004). This crosstalk is essential to orchestrate the metabolic needs of both players in the symbiosis. One of the clearest examples of these metabolic integration events can be found within intracellular insect-microbe symbioses (Hansen and Moran 2014). More than 10% of insect species possess long-term, mutualistic bacteria that provision nutrients to their insect host, and are housed inside of specialized host cells, referred to as bacteriocytes (Sudakaran *et al.* 2017). Bacteriocytes are adapted to facilitate inter-domain molecular interactions; however, the mechanisms that the host cell uses to regulate, respond to, and control this integrated, symbiotic metabolism is still largely unexplored.

The mutualistic interaction between the pea aphid (*Acyrthosiphon pisum*) and its bacterial endosymbiont, *Buchnera aphidicola*, is one of the best-studied models on nutritional symbioses. In this symbiosis, amino acid pathways of both players are integrated together for the production of essential amino acids (Nakabachi *et al.* 2005; Wilson

et al. 2010; Hansen and Moran 2011; Poliakov et al. 2011). This integrated mixed-domain metabolism ultimately enables aphids to utilize nutrient deficient plant sap as food because like most animals, essential amino acid pathways are not encoded in the aphid's genome (International Aphid Genomics Consortium 2010). For example, the aphid provides nonessential amino acid inputs to Buchnera's essential amino acid pathways, and then Buchnera provides essential amino acids and vitamins to its host (Shigenobu et al. 2000; Nakabachi et al. 2005; International Aphid Genomics Consortium 2010). Several aphid genes, including genes that recycle ammonia into glutamate, complement Buchnera's essential amino acid pathways, and transport nonessential amino acid inputs into bacteriocytes, are predicted to be key aphid genes involved in the regulation of this nutritional symbiosis (Nakabachi et al. 2005; Hansen and Moran 2011; Poliakov et al. 2011; Price et al. 2014).

Previous research on this system suggests that when dietary amino acid contents vary, aphids and Buchnera collectively adjust amino acid biosynthesis based on the aphid's nutritional requirements (Liadouze et al. 1995; Febvay et al. 1999). Moreover, when aphids feed on an artificial diet that varies only in nonessential amino acid profiles, bacteriocytes rebuild distinct profiles of amino acids that depend on the initial nonessential amino acid input(s) (Haribal and Jander 2015). Together, these results suggest that bacteriocytes respond to amino acid variation in the aphid's diet. Host-plants that pea aphids feed upon in the family Fabaceae vary dramatically in free amino acid profiles (Sandström and Pettersson 1994). In turn, it is unclear how aphid bacteriocytes regulate their key symbiotic genes within the amino acid metabolism and in other cellular processes when a polyphagous pea aphid line feeds on their specialized host-plant compared to other host-plants (Hansen and Moran 2014), which vary in free amino acid profiles (Sandström and Pettersson 1994).

Eukaryotic regulons are complex and are orchestrated through a combination of multiple mechanisms, including transcription factors, noncoding RNAs, and epigenetic factors. Among these different layers of gene regulation, the importance of epigenetic factors in influencing gene expression and alternative splicing has only recently begun to be elucidated (Luco et al. 2011; Romanoski et al. 2015). Previously, it has been shown that signals from the environment such as anxiety, stress, and diet can modify DNA methylation, which can subsequently alter gene expression profiles between different tissue types and throughout an organism's development (Feil and Fraga 2012; Tammen et al. 2013). For example, diet can modify DNA methylation patterns in a diversity of animals including insects, which in turn affects gene expression and subsequently influences organismal phenotypes (Niculescu and Zeisel 2002; Kucharski et al. 2008; Anderson et al. 2012). The pea aphid is an ideal insect to observe methylation patterns in because it possesses an asexual, clonal, parthenogenetic life stage (Dixon 1977) with a functional DNA methylation system (Walsh et al. 2010; Pasquier et al. 2014; Mukherjee and Baudach 2016). Such variation at the epigenomic level but not DNA level within a clonal aphid population may be advantageous if it leads to transient phenotypes that are associated with dynamic ecological factors such as host-plant nutrition. Currently it is unknown if the pea aphid has differential methylation in different tissue types and if host-plant environment influences methylation patterns in bacteriocytes. Therefore, DNA methylation may play an important role in the regulation of the aphid-Buchnera integrated metabolism, especially in response to different nutritional environments.

Here, using RNA-Seq and whole genome bisulfite sequencing we investigate if key aphid genes involved in the regulation of the aphid-Buchnera symbiosis are differentially expressed and methylated between a specialized and a universal host-plant diet (i.e., a host-plant diet all aphid biotypes can perform well on). If aphid bacteriocytes can alter these key symbiotic genes in response to feeding on different host-plant environments, aphids can potentially optimize and/or compensate for specialized plant diets that are otherwise unsuitable in nutrient profiles.

## MATERIALS AND METHODS

See SI Materials and Methods for details.

Acyrthosiphon pisum strain (LSR1) was divided into 6 independent sub-lines on either fava (F1, F2, and F3) or alfalfa (A1, A2, and A3) and reared at the same conditions as described in Hansen and Moran (2011). For surrogate aphid fitness trials, weights were measured similar to Vogel and Moran (2011) (N = 20 adult individuals per sub-line). Buchnera abundance was measured per aphid individual with Real Time quantitative PCR (RT-qPCR) (N= 6 individuals per life stage per sub-line) along with transmission electron microscopy (TEM) (N = 7 bacteriocytes for fava; N = 5 bacteriocytes for alfalfa). Bacteriocytes of each sub-line were dissected and counted using a light microscope.

For RNA-Seq trials, for the same 6 sub-lines pea aphid bacteriocytes and body cells were dissected, and RNA extractions were conducted similar to Hansen and Moran (2011), except RNA > 200 bp was retained for sequencing using HiSeq2500 (Illumina, San Diego, CA). For whole genome bisulfite sequencing, the same 6 sub-lines were dissected, extracted, bisulfite converted, and sequenced in a strand-specific manner on HiSeq2500 (Illumina, San Diego, CA) using a TruSeq SBS sequencing kit (Illumina).

For both host-plant treatments bacteriocyte gene expression was compared to body cell gene expression in this study similar to Nakabachi et al. (2005), Hansen and Moran (2011), and Poliakov et al. (2011) to identify unique bacteriocyte-specific signatures of gene expression relative to other cell types, regardless of host-plant treatment (BAC vs. BODY). To identify host-plant specific differences in bacteriocyte gene expression we compared bacteriocytes of alfalfa feeding aphids (ABAC) to bacteriocytes of fava (FBAC) feeding aphids (ABAC vs. FBAC). Specifically, RNA-Seq paired-end data were mapped and analyzed using the HISAT2 (Pertea et al. 2016) and StringTie (Pertea et al. 2015). Differentially expressed genes for BAC vs. BODY were identified using the likelihood ratio tests (LRT) based on generalized linear models with DESeq2 (Love et al. 2014) (see SI Materials and Methods for details). Differentially expressed genes for ABAC vs. FBAC were identified using LRT (see SI Materials and Methods for details). Statistical significance of differentially expressed genes were determined with FDR adjusted $P \leq$ 0.05 and $\geq$1.5X fold change of the normalized expression values.

The HISAT2 (Pertea et al. 2015) and DEXSeq (Anders et al. 2012) pipelines were used to find differentially spliced genes. Genes with differential exon usages between bacteriocytes and body cells (BAC vs. BODY) as well as between bacteriocytes of alfalfa feeding aphids and fava feeding aphids (ABAC vs. FBAC) were identified using a generalized linear model-based approach with FDR adjusted $P \leq$ 0.05 (see SI Materials and Methods for details).

Similar to above for expression data, bacteriocyte methylation data were also compared to body tissues regardless of host-plant type (BAC vs. BODY), and between bacteriocytes from aphids feeding on alfalfa and fava (ABAC vs. FBAC). Methylation read data were aligned with Bowtie2 and Bismark, as suggested in Krueger and Andrews (2011). Site-specific CpG methylation data for each sample were calculated using the methylKit package in R with the minimum coverage of 10 reads per site (see SI Materials and Methods for details) (Akalin et al. 2012). Differentially methylated CpG sites were obtained from methylSig (Park et al. 2014). Statistical significance of differentially

methylated CpG sites was determined with q ≤ 0.01 and ≥10% methylation differences. GSEA (Subramanian *et al.*2005) was used to determine which KEGG pathways were differentially regulated at the normalized $P < 0.05$ and $<0.1$ (See *SI Materials and Methods* for details).

### Data availability

### RESULTS

In this study, the pea aphid strain (LSR1), which originated as an alfalfa (*Medicago sativa*) specialist in the field (International Aphid Genomics Consortium 2010), was used for all trials. Here, the LSR1 strain was divided into six independent sub-lines for all host-plant trials; three sub-lines fed on its specialized host-plant, alfalfa, and the remaining three sub-lines fed on its 'universal' host-plant, fava. The universal host-plant fava was chosen because previous studies have indicated that most pea aphid biotypes favor and display higher fitness on their universal host-plant 'fava' (*Vicia faba*) compared to the host-plant they specialize on in the field (Ferrari *et al.* 2008; Ferrari *et al.* 2012; Peccoud *et al.* 2014). During all trials host-plants were of a particular developmental stage where amino acid profiles in sap vary significantly in alfalfa, compared to fava (Sandström and Pettersson 1994) (see *SI Materials and Methods*).

### Effects of host-plant diet on aphid and Buchnera phenotype

To investigate if pea aphid LSR1 fitness was significantly greater in aphids feeding on fava compared to their specialist host-plant alfalfa we measured adult aphid mass, a surrogate for aphid fitness (Vogel and Moran 2011).

Results indicated that all aphid sub-lines feeding on fava were of similar mass to one another, but aphid mass was significantly greater in aphid sub-lines feeding on fava compared to alfalfa based on Tukey's post-hoc tests ($\chi^2 = 567.017$, d.f. = 1, $P < 0.0005$, Figure 1A). A significant aphid line effect was found for host-plant treatment ($\chi^2 = 48.005$, d.f. = 1, $P < 0.0005$). One aphid sub-line feeding on alfalfa (A2) was significantly greater in mass compared to the other aphid sub-lines (Figure 1A).

Nutritional endosymbionts, such as *Buchnera*, can be regulated in insects at the bacteriocyte and/or symbiont titer level between different insect lifestages and morphs (Humphreys and Douglas 1997; Mira and Moran 2002; Kono *et al.* 2008; Nishikori *et al.* 2009; Stoll *et al.* 2010; Vigneron *et al.* 2014; Parkinson *et al.* 2016; Simonet *et al.* 2016). To determine if host-plant diet affects the number of aphid bacteriocytes and/or *Buchnera* titer we counted bacteriocytes and *Buchnera* cells and then compared them between host-plant treatments. The numbers of bacteriocytes in 4th instar aphids were not significantly different between host-plant treatments or aphid lines ($\chi^2 = 3.522$, d.f.=1, $P = 0.061$; $\chi^2 = 2.065$, d.f.=4, $P = 0.724$, respectively). Average bacteriocyte number per one aphid individual was 68 (95% C.I. 63.8-71.1, N = 45) and 63 (95% C.I. 58.9-66.2, N = 46) for the fava and alfalfa treatments, respectively. Moreover, *Buchnera* abundance did not significantly differ between host-plant treatments based on RT-qPCR ($\chi^2 = 2.674$, d.f.=1, $P = 0.102$, Figure 1B). Normalized *Buchnera* abundance was higher in first day nymphs compared to first day adults regardless of host-plant treatment ($\chi^2 = 55.486$, d.f.=1, $P < 0.0005$, Figure 1B). Consistent with these findings, the number of *Buchnera* cells within bacteriocytes did not differ significantly between aphids from representative host-plant treatments based on TEM images (T = 1.275, d.f.=9, $P = 0.234$). Within a unit area (100 μm²), an average of 21 (N = 7) and 24 (N = 5) *Buchnera* cells were identified within bacteriocytes of 4th instar fava and alfalfa feeding aphids, respectively (Figure 1C).
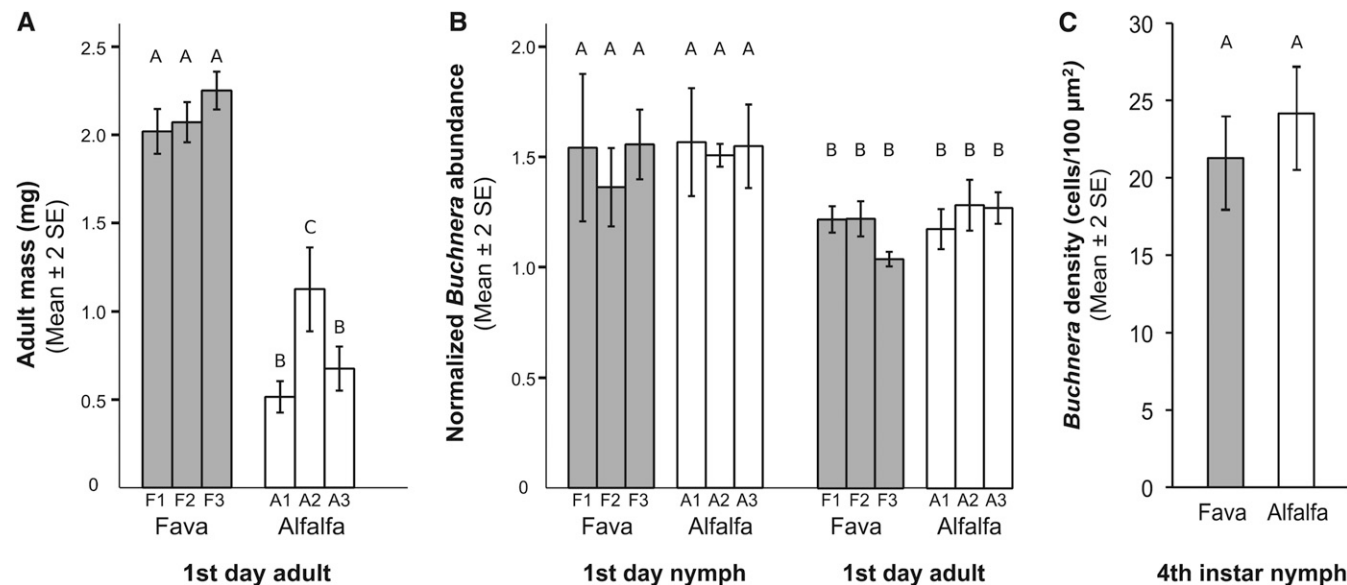


**Figure 1** Effects of a specialized host-plant diet on aphid and *Buchnera* phenotype. Different letters above bars indicate significant differences between aphid sub-lines within each sub-figure (A, B, C) (Tukey's multiple comparison post-hoc test $P < 0.05$) (A) Aphid mass of 1st day adults. N = 20 aphid individuals per aphid sub-line. (B) *Buchnera* abundance of 1st day nymphs and 1st day adults measured by a single copy *Buchnera* gene with RT-qPCR and normalized by a single copy aphid gene. N = 6 aphid individuals per aphid sub-line (C) *Buchnera* cell density of 4th instar nymphs measured by the number of *Buchnera* cells per unit area (100 μm²) using TEM.

## Effects of host-plant diet on the expression of key aphid symbiosis genes

To investigate if host-plant diet affects the expression of key symbiotic genes of aphids at the mRNA level in bacteriocytes, we first conducted RNA-Seq on bacteriocytes and other aphid body tissues (BAC *vs.* BODY) (Table S1). For the BAC *vs.* BODY comparison we identified 1,904 genes that were significantly up-regulated between bacteriocytes and body cells for both host-plant treatments (Dataset S1, Table S2). We further identified 4,211 genes that were significantly down-regulated between bacteriocytes and body cells for both host-plant treatments (Dataset S2, Table S2). To determine host-plant differences between bacteriocytes (ABAC *vs.* FBAC) we identified 54 genes that were up-regulated and 101 genes that were down-regulated in bacteriocytes of alfalfa feeding aphids compared to bacteriocytes of fava feeding aphids (Dataset S3, Table S2).

To characterize the functions of genes differentially expressed in bacteriocytes for each comparison we used GSEA (Subramanian *et al.* 2005). For the BAC *vs.* BODY comparison we found 10 KEGG pathways significantly enriched in bacteriocytes compared to body cells for both host-plant treatments (Table S3a). The top five KEGG pathways in descending order based on the GSEA enrichment score were glycine, serine and threonine metabolism, glyoxylate and dicarboxylate metabolism, phenylalanine metabolism, pentose phosphate pathway, and the nicotinate and nicotinamide metabolism (Table S3a). We also identified five KEGG pathways that were significantly down-regulated in bacteriocytes compared to body cells for both host-plant treatments (Table S3a). Such KEGG pathways in descending order based on the GSEA enrichment score were hippo signaling pathway, notch signaling pathway, other glycan degradation, phototransduction, and neuroactive ligand-receptor interaction.

To identify the pathways that were differentially expressed in bacteriocytes between host-plant treatments (ABAC *vs.* FBAC) we performed GSEA and identified seven KEGG pathways that were significantly enriched in bacteriocytes of alfalfa compared to fava treated aphids. The seven KEGG pathways in descending order based on the GSEA enrichment score were synthesis and degradation of ketone bodies, vitamin B6 metabolism, aminoacyl-tRNA biosynthesis, pyruvate metabolism, Jak-STAT signaling pathway, lipoic acid metabolism, and the butanoate metabolism (Table S3b).

Aphid and *Buchnera* metabolisms are integrated for the production of amino acids within bacteriocytes. This shared amino acid metabolism is hypothesized to be regulated by the aphid host via transporters, the GS/GOGAT cycle, and genes that complement *Buchnera*'s essential amino acid pathways (Nakabachi *et al.* 2005; Hansen and Moran 2011; Poliakov *et al.* 2011; Price *et al.* 2014). We examined these aphid genes and found that 22 out of 27 genes were significantly enriched in bacteriocytes relative to body cells for both host-plant treatments (BAC *vs.* BODY; Figure 2). We further identified that two out of eight of these genes (Glutamine synthetase and *ApGLNT1*) were significantly enriched in bacteriocytes of alfalfa feeding aphids when compared to bacteriocytes of fava feeding aphids (ABAC *vs.* FBAC; Figure 2; Dataset S4). Glutamine synthetase (*GS*) is a key enzyme of the GS/GOGAT cycle and recycles ammonia into glutamine (Hansen and Moran 2011). The transporter *ApGLNT1* imports glutamine into bacteriocytes (Price *et al.* 2014). Collectively these results suggest that aphid genes that synthesize and transport glutamine, an important amino donor for *Buchnera*'s essential amino acid pathways, is enriched in bacteriocytes
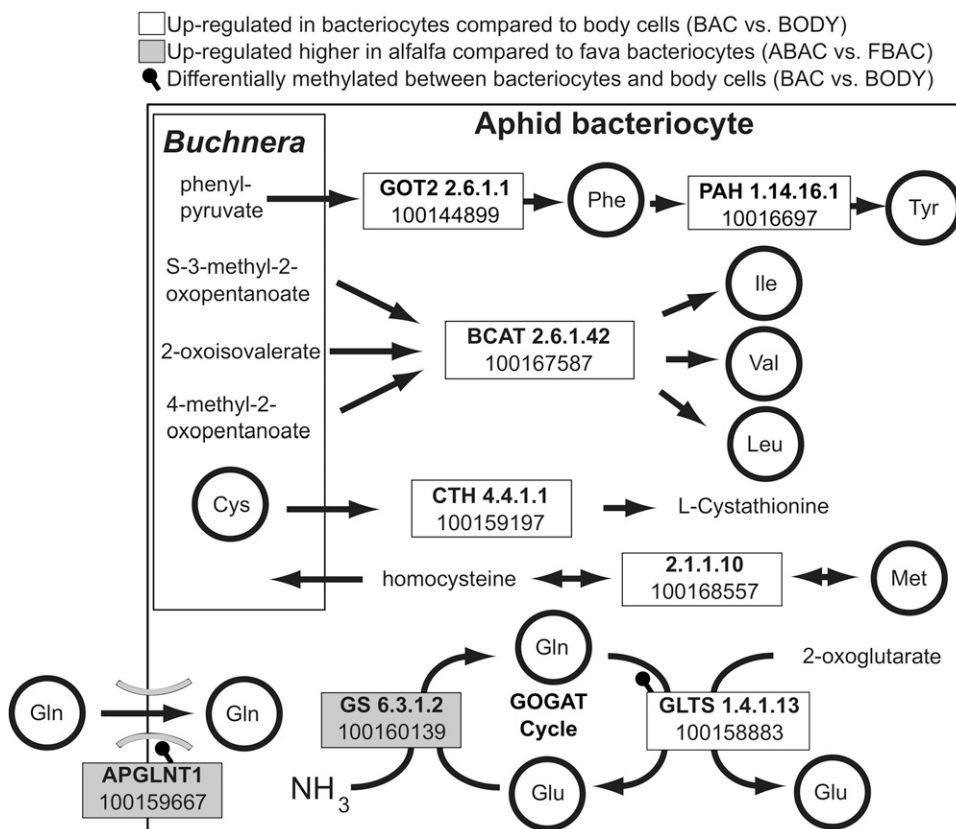


Figure 2 Host-plant effects on differential expression and methylation of key aphid genes that complement and regulate the integrated aphid-*Buchnera* amino acid metabolism. Gene boxes are annotated with either E.C. numbers, Genbank LOC numbers, and/or gene names. Genes are significantly up-regulated in bacteriocytes compared to body cells (BAC *vs.* BODY) and enriched in alfalfa compared to fava bacteriocytes (ABAC *vs.* FBAC) if adjusted p-value ≤ 0.05 and normalized read counts are 50% higher. Genes are differentially methylated significantly if there was ≥10% difference in percent methylation and FDR corrected p-value ≤ 0.01.
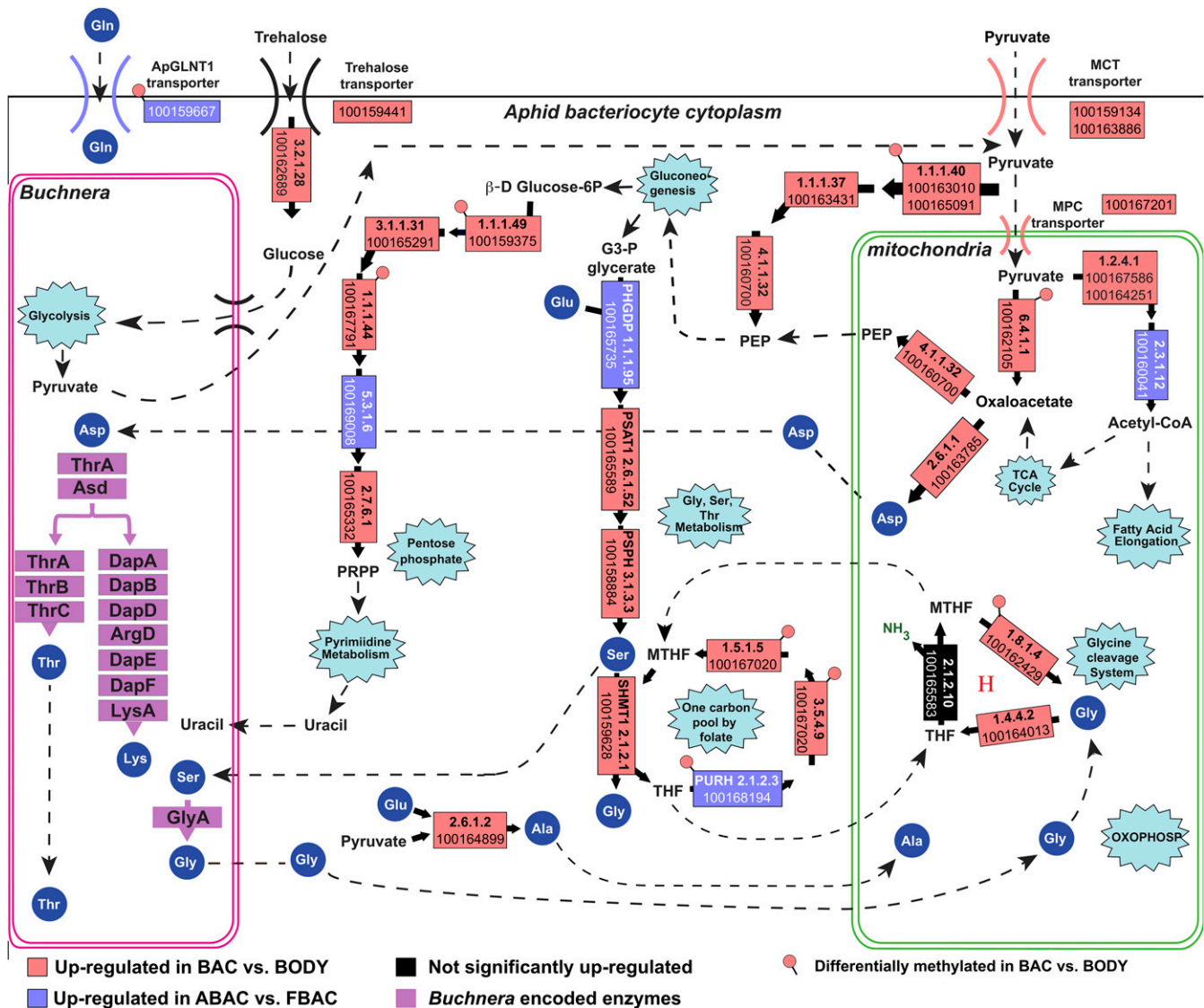
**Figure 3** Differential gene expression and methylation of the Glycine/Serine metabolism in aphid bacteriocytes. Gene boxes are annotated with either E.C. numbers, Genbank LOC numbers, and/or gene names. 'H' denotes the glycine cleavage system H protein (LOC100169052). Genes are significantly up-regulated in bacteriocytes compared to body cells (BAC *vs.* BODY) and in alfalfa compared to fava bacteriocytes (ABAC *vs.* FBAC) if adjusted p-value ≤0.05 and normalized read counts are 50% higher. Genes are differentially methylated significantly if there was ≥10% difference in percent methylation and FDR corrected p-value ≤ 0.01.

of aphids feeding on their specialized host-plant, alfalfa, compared to their universal host-plant, fava (Figure 2).

Another major subset of genes that were enriched significantly in bacteriocytes compared to other body cells for both host-plant treatments (BAC *vs.* BODY) belonged to the glycine/serine metabolism (Figure 3, Dataset S4). Specifically, the genes for serine biosynthesis (D-3-phosphoglycerate dehydrogenase; *PHGDP*, phosphoserine aminotransferase 1; *PSAT1*, phosphoserine phosphatase; *PSPH*) were up-regulated significantly in bacteriocytes compared to body cells. Also, serine hydroxymethyltransferase (*SHMT*), which converts serine to glycine, and the bifunctional purine biosynthesis protein (*PURH*) were up-regulated significantly higher in bacteriocytes compared to body cells. The glycine/serine metabolism also relies on the maintenance of the cofactor tetrahydrofolate (THF). In the bacteriocytes of both treatments THF was maintained by the significant

up-regulation of genes in the one carbon pool by folate metabolism, and the production of 5,10-methenyl-THF through the Glycine Cleavage System (Figure 3, Dataset S4). Among such genes, *PHGDP* and *PURH* were significantly higher in bacteriocytes of alfalfa compared to fava feeding aphids (ABAC *vs.* FBAC; Figure 3).

For both host-plant treatments (BAC *vs.* BODY), gluconeogenesis instead of glycolysis appears to be occurring in bacteriocytes compared to body cells as indicated by the up-regulation of malate dehydrogenase (*MDH*) and phosphoenolpyruvate carboxykinase (*PEPCK*) (Figure 3). In turn, instead of glucose, alternative energy substrates such as extracellular and *Buchnera* derived pyruvate may provide the carbon backbone to fuel the glycine/serine metabolism. For example, pyruvate transporters were significantly up-regulated in bacteriocyte cells compared to body cells (monocarboxylate transporter; *MCT* and mitochondrial pyruvate carrier; *MPC*) (Figure 3).
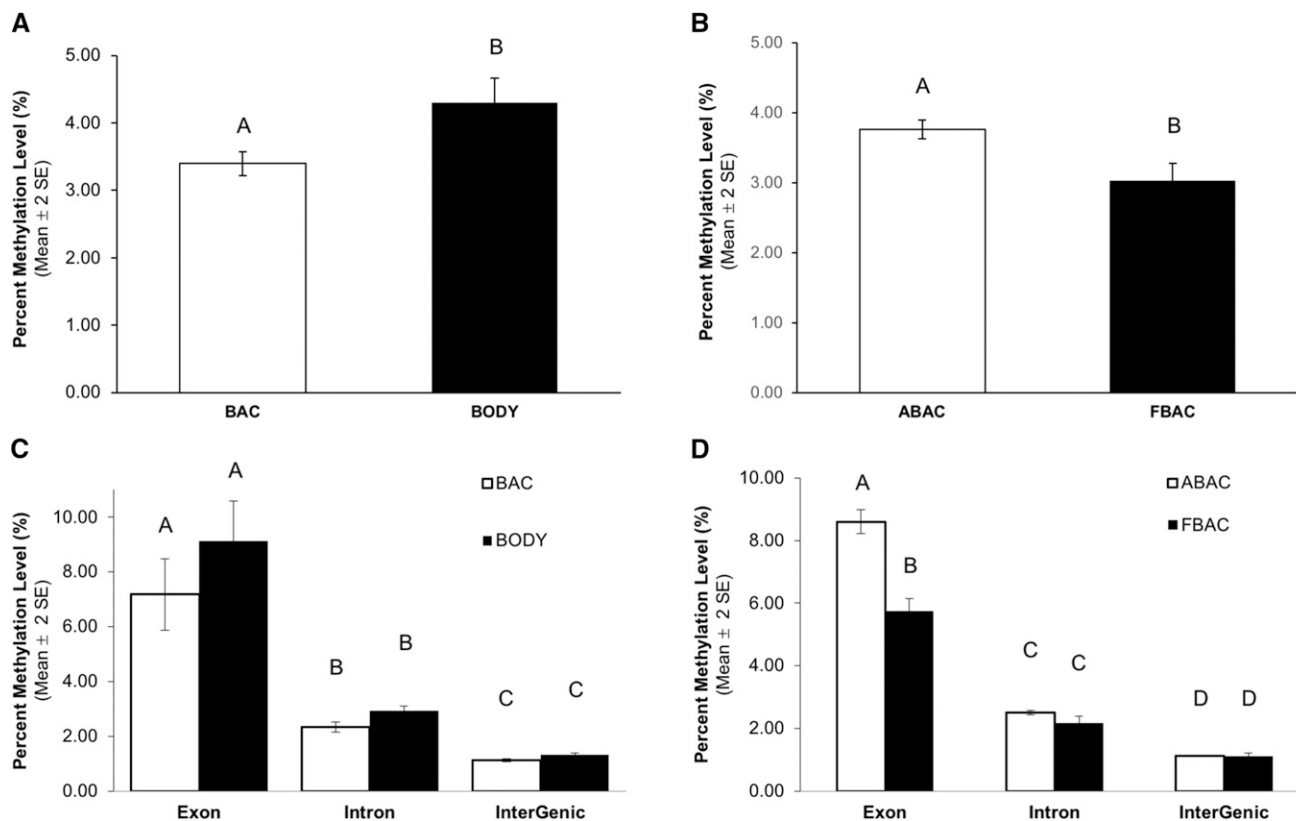
**Figure 4** Percent CpG methylation level for bacteriocyte and body samples from each host-plant treatment. Different letters on bars indicate significant difference between each group and within each subfigure (A-D) (paired *t*-test *P* < 0.05). BAC and BODY denote bacteriocytes and body cells, respectively. Each sample has 6 biological replicates from both alfalfa and fava feeding aphids. ABAC and FBAC denote bacteriocytes of alfalfa feeding aphids and fava feeding aphids, respectively. Each sample has 3 biological replicates. (A) Average methylation levels of BAC and BODY. (B) Average methylation levels of ABAC and FBAC. (C) Average methylation levels of genic regions of BAC and BODY. (D) Average methylation levels of genic regions of ABAC and FBAC.

Another key pathway that is important in the aphid-*Buchnera* symbiosis involves the production of uracil. *Buchnera* is unable to produce its own uracil (Shigenobu *et al.* 2000) and therefore it depends on the host for uracil biosynthesis in this integrated metabolism. The uracil salvage pathway, especially pseudouridine kinase, was found to be significantly enriched in bacteriocytes of aphids feeding on alfalfa compared to fava (ABAC *vs.* FBAC; Dataset S4).

### Effects of host-plant diet on DNA methylation profiles

Diet cues can alter DNA methylation patterns within and adjacent to invertebrate genes (Kucharski *et al.* 2008). Moreover, these DNA methylation marks have been associated with active genes and alternative splicing (Yan *et al.* 2015). To determine if host-plant diet affects the methylation of key symbiotic genes in aphid bacteriocytes we used whole-genome bisulfite sequencing (Table S4). The average percentages of CpG methylation were 3.4% and 4.3% for bacteriocytes and body cells, respectively (Figure 4A). Bacteriocytes had a significantly lower percent of CpG methylation compared to body cells (BAC *vs.* BODY; paired *t*-test, t = 13.47, df = 5, *P* < 0.0001) (Figure 4A). The average percentages of CpG methylation were 3.8% and 3.0% for bacteriocytes of alfalfa compared to fava feeding aphids, respectively (Figure 4B). Bacteriocytes of alfalfa feeding aphids had a significantly higher percent of CpG methylation compared to bacteriocytes of fava feeding aphids (ABAC *vs.* FBAC; t = 5.18, df = 4, *P* = 0.0066) (Figure 4B).

To investigate if there was a difference in percent methylation within and outside genic regions percent methylation within the exon, intron, and the intergenic regions were determined. For both host-plant diets there was not a significant difference in percent methylation between bacteriocyte and body cells for the exon, intron, or the intergenic regions (BAC *vs.* BODY; paired *t*-test, t = 1.76, df = 2, *P* = 0.22) (Figure 4C). In contrast, when comparing between host-plant diets percent methylation in the exon region was significantly higher in bacteriocytes of alfalfa feeding aphids compared to bacteriocytes of fava feeding aphids (ABAC *vs.* FBAC; paired *t*-test, t = 10.34, df = 4, *P* = 0.0005) (Figure 4D). The percent of methylation in the intron and intergenic regions were not significantly different in bacteriocytes of alfalfa feeding aphids compared to bacteriocytes of fava feeding aphids (ABAC *vs.* FBAC; paired *t*-test, t = 2.40, df = 2, *P* = 0.14 for introns; t = 0.57, df = 2, *P* = 0.63 for intergenic) (Figure 4D).

Collectively, for all bacteriocyte and body tissue samples percent methylation within the exon regions was significantly higher compared to the intron regions ((paired *t*-test, t = 12.05, df = 11, *P* < 0.001) (Figure 4C), t = 8.50, df = 5, *P* < 0.001) (Figure 4D)). Also, percent methylation was significantly higher in the intron regions compared to the intergenic regions for all samples ((paired *t*-test, t = 19.03, df = 11, *P* < 0.001), (Figure 4C), (t = 17.02; df = 5; *P* < 0.001) (Figure 4D)).

In order to visually determine how shared CpG sites differ between all samples in percent methylation when aphids feed on their specialized

compared to universal host-plant diet we conducted a Principal Component Analysis (Figure 5). To test if methylated CpG profiles were significantly different between bacteriocyte, body cell, and host-plant treatments we used multi-response permutation procedure (MRPP) (Mielke and Berry 2003). We found that the four *a priori* groups: bacteriocytes of alfalfa feeding aphids (ABAC), body cells of alfalfa feeding aphids (ABODY), bacteriocytes of fava feeding aphids (FBAC), and body cells of fava feeding aphids (FBODY) were significantly different from one another in CpG methylation profiles ($P < 0.001$; A = 0.09356). In the PCA ordination, all body samples clustered tightly together in ordination space away from bacteriocyte samples. Using MRPP, we found that methylated CpG profiles of bacteriocyte samples were significantly different compared to body cell samples (BAC *vs.* BODY; $P < 0.005$; A = 0.08245), with a significantly higher dispersion of within-group differences for bacteriocyte samples (delta = 0.2759) compared to body cell samples (delta = 0.1772). In accord to the PCA ordination MRPP results indicate that bacteriocyte cells have significantly different methylation profiles compared to body cells, which have more similar distributions to one another, regardless of host-plant treatment (Figure 5). Nevertheless, body cells are a mixture of different aphid cell types and therefore we cannot exclude the possibility that more abundant host cell types mask host-plant differences of less abundant cell types. In contrast, among bacteriocyte samples (ABAC *vs.* FBAC), fava samples are more heterogeneous in CpG profiles compared to alfalfa samples, with a higher dispersion of within-group differences for FBAC (delta = 0.2810) compared to ABAC (delta = 0.2645) (Figure 5). These results indicate that methylation profiles are specific to aphid cell type and host-plant diet, especially for bacteriocytes in the aphid's specialized host-plant treatment, alfalfa.

A total of 3,474 CpG sites were significantly differentially methylated between bacteriocytes and body cells of both host-plant treatments (BAC *vs.* BODY) (Table S5a). Between bacteriocytes of alfalfa feeding aphids compared to fava feeding aphids a total of 294 CpG sites were differentially methylated significantly (ABAC *vs.* FBAC). For both comparisons, differential CpG methylation was primarily confined to the gene body regions (82% for BAC *vs.* BODY; 78% for ABAC *vs.* FBAC) (Table S5b) as revealed in other non-mammal animals, which may contribute to gene activation and/or alternative splicing (Hunt *et al.* 2013).

To link the patterns of differential DNA methylation with differential gene expression, we identified 441 genes that were both differentially methylated and differentially expressed significantly between bacteriocyte and body samples for both alfalfa and fava feeding aphids (BAC *vs.* BODY) (Table S5). All 441 genes were up-regulated and hypomethylated in bacteriocytes relative to body cells. We also identified 702 genes that are both differentially methylated and differentially spliced between bacteriocyte and body samples of both alfalfa and fava feeding aphids (BAC *vs.* BODY) (Table S5; Dataset S4). Furthermore, we identified three genes that were both differentially methylated and differentially expressed significantly between bacteriocytes of alfalfa compared to fava feeding aphids (ABAC *vs.* FBAC; Table S5). All three genes were down-regulated and hyper-methylated in alfalfa compared to fava feeding aphid bacteriocytes. These genes were the serine/arginine repetitive matrix 1 gene (LOC100160294), the proton-coupled amino acid transporter 4-like (LOC100159667), and the broad-complex core protein isoforms 1/2/3/4/5-like gene (LOC100167015). We also identified three genes that were both differentially methylated and differentially spliced between bacteriocyte samples from alfalfa and fava feeding aphids (ABAC *vs.* FBAC; Table S5): anoctamin-1-like (LOC100167803), formin-binding protein 1-like (LOC100166693), and tumor protein D54-like (LOC100164449).
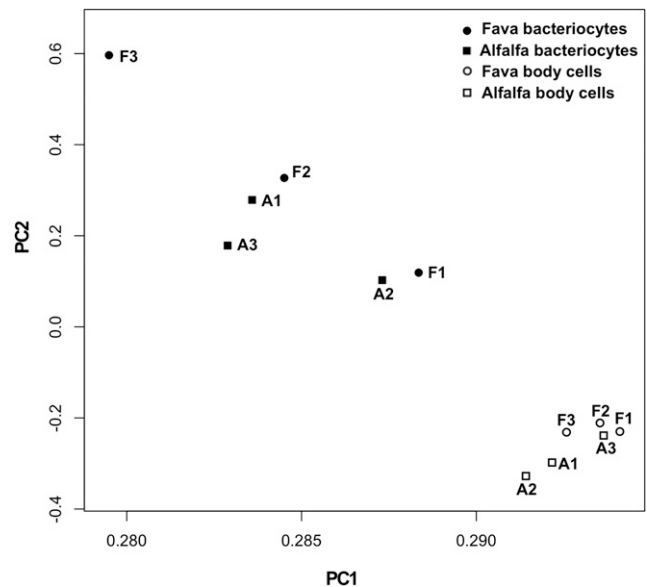


**Figure 5** Principal Component Analysis of CpG methylation profiles for each bacteriocytes and body sample. F1, F2, and F3 denote 3 biological replicates of aphid sub-lines within the fava treatment. A1, A2, A3 denote 3 biological replicates of aphid sub-lines within the alfalfa treatment. PC1 explains 93.1% of total variance with standard deviation of 3.34. PC2 explains 1.5% of total variance with standard deviation of 0.43.

Using GSEA (Subramanian *et al.* 2005), we found five pathways that were both significantly differentially methylated (hypo-methylated) and up-regulated in bacteriocytes compared to body cells. These five pathways were the metabolic pathways, lysosome, protein processing in endoplasmic reticulum, selenocompound metabolism, and tryptophan metabolism (Table S3c). No pathways were differentially methylated and significantly down-regulated in bacteriocytes compared to body cells (BAC *vs.* BODY; Table S3c). Genes that were both differentially methylated and differentially spliced between bacteriocyte and body samples (BAC *vs.* BODY) belong to 16 KEGG pathways (Table S6). These pathways include the bulk movement into cells and digestion (15 genes), the degradation, processing, and transport of RNAs (47 genes), four different signaling pathways including one involved in the immune response (Jak-Stat) (37 genes), protein processing and degradation (20 genes), and the biosynthesis of amino acids (6 genes) (Figure 2).

Fourteen key aphid genes associated with the integrated aphid-*Buchnera* symbiosis and the glycine/serine metabolism were both differentially expressed and differentially methylated between the bacteriocytes and body cells (BAC *vs.* BODY) (Figures 2, 3; Dataset S4). For example, glutamate synthase, an important enzyme of the GS/GOGAT cycle that converts glutamine to glutamate was both hypo-methylated, differentially spliced, and up-regulated in bacteriocytes of both fava and alfalfa feeding aphids compared to body cells (Figure 3). In addition, an active glutamine transporter (ApGLNT1) that was previously characterized to be important for the regulation of *Buchnera*'s essential amino acid biosynthesis pathways (Price *et al.* 2014) was significantly hypo-methylated and up-regulated in bacteriocytes compared to body cells (Figure 2). Also, two enzymes (PURH and cytoplasmic C-1-tetrahydrofolate synthase) in the one carbon pool by folate pathway were hypo-methylated and up-regulated in the bacteriocytes compared to the body cells (Figure 3).

## DISCUSSION

In this study, we demonstrated for the first time that key aphid genes involved in the regulation of the aphid-*Buchnera* symbiosis are differentially expressed, spliced, and methylated when aphids feed on a specialized compared to a universal host-plant diet. Our data indicate that this regulatory and epigenetic response to distinct host-plant types that vary in amino acid profiles may play a significant role in modulating the aphid-*Buchnera* amino acid metabolism when aphids feed on their specialized compared to universal host-plant diet. Moreover, this regulatory response in combination with lower aphid fitness when aphids feed on their specialized compared to universal host-plant diet is consistent with the pea aphid engaging in a compensatory metabolic response when it specializes on a less suitable host-plant. We also identified key aphid genes and pathways involved in the aphid-*Buchnera* symbiosis that are differentially expressed, spliced, and methylated in both host-plant diets. These results collectively suggest that DNA methylation may play both a conserved (maintenance methylation) and an environmentally induced (*de novo* methylation) regulatory role in bacteriocytes when aphids feed on host-plant diets that vary in amino acid profiles.

Here we reveal that instead of regulating bacteriocyte or *Buchnera* cell number the aphid-*Buchnera* integrative metabolism modulates patterns of bacteriocyte DNA methylation and gene expression in response to its specialized host-plant diet, alfalfa, when compared to its universal host-plant diet, fava. Results from our study that provide evidence for this finding include the following: *1)* One aphid enzyme (Glutamine synthetase; *GS*) in the GS/GOGAT cycle was enriched in bacteriocytes of alfalfa feeding aphids compared to fava feeding aphids (Figure 2). The GS/GOGAT cycle is hypothesized to play a key role with *Buchnera* in sustaining aphids on a nitrogen-limited diet, because *GS* recycles waste ammonia for the production of glutamine (Hansen and Moran 2011). *2)* The transporter *ApGLNT1* is significantly up-regulated in the bacteriocytes of alfalfa feeding aphids compared to fava feeding aphids (Figure 2). Interestingly this transporter was significantly hypo-methylated only in bacteriocytes compared to body cells of both alfalfa and fava feeding aphids. This transporter imports glutamine into bacteriocytes and is inhibited by arginine produced by *Buchnera*. In turn, this transporter may play a key role in regulating *Buchnera*'s essential amino acid metabolism (Price *et al.* 2014), by increasing essential amino acid biosynthesis in alfalfa feeding aphids. *3)* The vitamin B6 pathway was significantly enriched in bacteriocytes of alfalfa feeding aphids compared to fava feeding aphids (Table S3b). Vitamin B6 is an essential cofactor in animals and microbes and plays an important role in the amino acid and carbohydrate metabolism and singlet oxygen resistance (John 1995; Daub and Ehrenshaft 2000). Collectively, both the aphid and *Buchnera* do not encode the entire vitamin B6 biosynthesis I or II pathways, however *Buchnera* still encodes serC and thrC and the aphid encodes the enzymes 2.6.1.52, 2.7.1.35, and 1.4.3.5., which make up the majority of the pathway. Both enzymes 2.6.1.52 and 2.7.1.35 are up-regulated significantly in bacteriocytes compared to body tissues 15X and 8X respectively suggesting that there is a demand for vitamin B6 biosynthesis in aphid bacteriocytes. *4)* The uracil salvage pathway was found to be significantly enriched in bacteriocytes of aphids feeding on alfalfa compared to fava. These results suggest that more uracil potentially for both *Buchnera* and/or aphid mRNA biosynthesis is needed for aphid bacteriocytes when they feed on alfalfa compared to fava. *5)* Also six genes were both differentially expressed/spliced and methylated between bacteriocytes of alfalfa compared to fava bacteriocytes and are involved in amino acid transport, protein kinase activity, calcium activated chloride channel activity, gene regulation via epigenomic interactions utilizing the POZ zinc finger domain, and unknown function. More information on how these genes may help regulate the integrated aphid-*Buchnera* metabolism when aphids feed on different host-plants is needed. Collectively, these results suggest that when aphids feed on their suboptimal, specialized host-plant, alfalfa, key aphid genes involved in the regulation of the integrative amino acid metabolism are enriched in bacteriocytes.

The observed regulatory changes in bacteriocytes are indicative of a compensatory response of aphids feeding on alfalfa that require more essential amino acids compared to aphids on fava. For example, previous amino acid concentration data shows a limitation of essential amino acids in alfalfa sap compared to fava sap (Sandström and Pettersson 1994). Specifically, Sandström and Pettersson (1994) revealed that total amino acid concentrations in sap obtained from aphid stylets are relatively similar between alfalfa and fava, however concentrations of nine essential amino acids (arginine, isoleucine, leucine, lysine, phenylalanine, tryptophan, tyrosine, valine, and histidine) are lower in alfalfa sap compared to fava sap. Our data coincides with these amino acid concentration data (Sandström and Pettersson 1994), because we revealed an up-regulation of key aphid genes and pathways that are important in provisioning amino donors, co-factors, and energy to fuel the integrative amino acid metabolism with *Buchnera* for the production of essential amino acids.

In this study, we also observed that regardless of host-plant treatment 1,143 genes were differentially expressed/spliced and methylated in aphid bacteriocytes compared to body cells. These genes included glutamate synthase, the other gene involved in the GS/GOGAT cycle (Figure 2), as well as several pathways involved in cell signaling, immune function, and the regulation of RNA and protein biosynthesis and degradation. One of the most distinctive and highly enriched pathways identified within bacteriocytes compared to body cells was the glycine/serine metabolic profile (Figure 3). Interestingly, to the best of our knowledge this metabolic profile has only been identified previously as a metabolic hallmark in cancer cells (Perroud *et al.* 2006; Jain *et al.* 2012; Amelio *et al.* 2014; Locasale 2013). Unlike some cancer cells, however, expression profiles of bacteriocytes in our study reveal that gluconeogenesis instead of glycolysis is occurring (Figure 3); *i.e.*, the Warburg effect, where cellular energy production relies heavily on aerobic glycolysis (Vander Heiden *et al.* 2009), is not occurring. This pattern of carbohydrate utilization in pea aphid bacteriocytes was identified previously (Poliakov *et al.* 2011). An alternative energy source instead of glucose, such as pyruvate, may be metabolized to fuel the glycine/serine metabolism. This model of using pyruvate as the carbon skeleton for the glycine/serine metabolism has been proposed previously in breast cancer cells (Diers *et al.* 2012). Similar to breast cancer cells we also observed the up-regulation of the pyruvate transporters MCT and MPC and the enzymes that utilize the carbon skeleton of pyruvate into gluconeogenesis in bacteriocyte cells (Figure 3). Currently, it is unclear why cancer cells display this distinctive glycine/serine enriched metabolic profile, however several hypotheses have been proposed, such as reducing oxidative stress from cells that are metabolically very active (Jain *et al.* 2012; di Salvo *et al.* 2013; Maddocks *et al.* 2013; Labuschagne *et al.* 2014). Investigating the regulatory mechanisms behind the differences between metabolically active and prolific cancer cells and metabolically active, yet not prolific aphid bacteriocytes can become a key factor in elucidating the significance of how and why the glycine metabolism is adopted in both malignant cancer cells and symbiotic cells.

The genome-wide CpG methylation patterns of the pea aphid showed higher levels of methylation in gene bodies, especially in alfalfa feeding bacteriocytes (Figure 4B), which is consistent with methylation patterns of other insect species with functional methylation systems (Suzuki *et al.* 2007; Feng *et al.* 2010; Zemach *et al.* 2010; Provataris *et al.* 2018). Bacteriocytes from both host-plant treatments revealed significantly lower percent CpG methylation levels compared to the body cells of pea aphid (Figure 4A). These results suggest that more exon and intron regions within bacteriocytes are hypo-methylated compared to body cells. Host-plant treatments also influenced hypo- *vs.* hyper-methylation patterns in our study, as hundreds of genes were differentially methylated between bacteriocytes and body cells and between bacteriocytes depending on what host-plant they fed upon. In one study (Huh *et al.* 2013) it has been hypothesized that methylation inside of gene bodies plays an important role in reducing transcriptional noise (Huh *et al.* 2013). Differential methylation within gene bodies has also been associated with alternative gene splicing (Shukla *et al.* 2011). Here we observed hundreds of genes that were both differentially spliced and methylated within bacteriocytes compared to body cells and between bacteriocytes of alfalfa compared to fava feeding aphids. Genes that are hyper-methylated may also result in higher gene expression levels in insects. For example, a recent study suggested that gene body methylation of the lysosomal alpha-mannosidase (*LAM*) gene of *Apis melifera* increased the expression level of *LAM* (Wedd *et al.* 2016). In contrast to this latter study, all genes that were both differentially expressed and methylated in this current study were hypo-methylated and up-regulated, which is more in line with what is observed in vertebrates but within the promoter regions (Jones 2012).

The effect of host-plant treatment on DNA methylation patterns within pea aphid bacteriocytes was evident in our study. For example, our results reveal that the CpG methylation profiles of fava feeding aphid bacteriocytes were more heterogeneous between biological replicates compared to alfalfa feeding bacteriocytes (Figure 5). Such heterogeneity may have come from the relaxed nutritional constraints of the 'universal' host-plant, fava, compared to the specialized host-plant, alfalfa. Alternatively, environmental cues from its specialized host-plant, alfalfa, may induce host-plant specific methylation profiles. In this study we identified gene candidates to examine further that were both differentially expressed/spliced and methylated in bacteriocytes compared to body cells and between bacteriocytes when aphids fed on different host plant diets. If host-plant specific patterns of DNA methylation induce tissue and host-plant specific gene expression profiles in bacteriocytes this may be a key regulatory factor that induces phenotypic variation of the integrative metabolism in response to host-plant diet.

In summary, our findings indicate that when aphids feed on their specialized host-plant, alfalfa, key aphid genes that are involved in the integrative metabolism with *Buchnera* are differentially expressed/spliced, and some of these are differentially methylated. Future studies are needed to investigate how and if these methylated sites can influence the regulation of the bacteriocyte, be inherited, and ultimately drive host-plant specialization broadly in sap-feeding insect-nutritional symbioses. We hypothesize that this host-plant induced metabolic modification to the aphid's integrative metabolism may ultimately allow aphids to utilize host-plant diets that were once unsuitable.

## AUTHOR CONTRIBUTIONS

A.K.H. and D.K. conceived and designed the study; A.K.H. prepared samples for RNA-Seq, diet bioassays; H.L.-B. and D. K. prepared samples for whole genome bisulfite sequencing, D. K., A.K.H., and B.F.M analyzed data; and D. K. and A.K.H. wrote the paper with input from all authors. The authors declare no conflict of interest.

## LITERATURE CITED

Akalin, A., M. Kormaksson, S. Li, F. E. Garrett-Bakelman, M. E. Figueroa *et al.*, 2012 methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol. 13: R87. https://doi.org/10.1186/gb-2012-13-10-r87

Amelio, I., F. Cutruzzola, A. Antonov, M. Agostini, and G. Melino, 2014 Serine and glycine metabolism in cancer. Trends Biochem. Sci. 39: 191–198. https://doi.org/10.1016/j.tibs.2014.02.004

Anders, S., A. Reyes, and W. Huber, 2012 Detecting differential usage of exons from RNA-seq data. Genome Res. 22: 2008–2017. https://doi.org/10.1101/gr.133744.111

Anderson, O. S., K. E. Sant, and D. C. Dolinoy, 2012 Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. J. Nutr. Biochem. 23: 853–859. https://doi.org/10.1016/j.jnutbio.2012.03.003

Daub, M. E., and M. Ehrenshaft, 2000 The photoactivated *Cercospora* toxin cercosporin: contributions to plant disease and fundamental biology. Annu. Rev. Phytopathol. 38: 461–490. https://doi.org/10.1146/annurev.phyto.38.1.461

Diers, A. R., K. A. Broniowska, C. F. Chang, and N. Hogg, 2012 Pyruvate fuels mitochondrial respiration and proliferation of breast cancer cells: effect of monocarboxylate transporter inhibition. Biochem. J. 444: 561–571. https://doi.org/10.1042/BJ20120294

di Salvo, M. L., R. Contestabile, A. Paiardini, and B. Maras, 2013 Glycine consumption and mitochondrial serine hydroxymethyltransferase in cancer cells: the heme connection. Med. Hypotheses 80: 633–636. https://doi.org/10.1016/j.mehy.2013.02.008

Dixon, A. F. G., 1977 Aphid ecology: life cycles, polymorphism, and population regulation. Annu. Rev. Ecol. Syst. 8: 329–353. https://doi.org/10.1146/annurev.es.08.110177.001553

Dyall, S. D., M. T. Brown, and P. J. Johnson, 2004 Ancient invasions: from endosymbionts to organelles. Science 304: 253–257. https://doi.org/10.1126/science.1094884

Febvay, G., Y. Rahbé, M. Rynkiewicz, J. Guillaud, and G. Bonnot, 1999 Fate of dietary sucrose and neosynthesis of amino acids in the pea aphid, *Acyrthosiphon pisum*, reared on different diets. J. Exp. Biol. 202: 2639–2652.

Feil, R., and M. F. Fraga, 2012 Epigenetics and the environment: emerging patterns and implications. Nat. Rev. Genet. 13: 97–109. https://doi.org/10.1038/nrg3142

Feng, S., S. J. Cokus, X. Zhang, P. Chen, M. Bostick *et al.*, 2010 Conservation and divergence of methylation patterning in plants and animals. Proc. Natl. Acad. Sci. USA 107: 8689–8694. https://doi.org/10.1073/pnas.1002720107

Ferrari, J., S. Via, and H. C. Godfray, 2008 Population differentiation and genetic variation in performance on eight hosts in the pea aphid complex. Evolution 62: 2508–2524. https://doi.org/10.1111/j.1558-5646.2008.00468.x

Ferrari, J., J. A. West, S. Via, and H. C. Godfray, 2012 Population genetic structure and secondary symbionts in host-associated populations of the pea aphid complex. Evolution 66: 375–390. https://doi.org/10.1111/j.1558-5646.2011.01436.x

Hansen, A. K., and N. A. Moran, 2011 Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. Proc. Natl. Acad. Sci. USA 108: 2849–2854. https://doi.org/10.1073/pnas.1013465108

Hansen, A. K., and N. A. Moran, 2014 The impact of microbial symbionts on host plant utilization by herbivorous insects. Mol. Ecol. 23: 1473–1496. https://doi.org/10.1111/mec.12421

Haribal, M., and G. Jander, 2015    Stable isotope studies reveal pathways for the incorporation of non-essential amino acids in *Acyrthosiphon pisum* (pea aphids). J. Exp. Biol. 218: 3797–3806. https://doi.org/10.1242/jeb.129189

Huh, I., J. Zeng, T. Park, and S. V. Yi, 2013    DNA methylation and transcriptional noise. Epigenetics Chromatin 6: 9 (corrigenda: Epigenetics Chromatin 7: 13). https://doi.org/10.1186/1756-8935-6-9

Humphreys, N. J., and A. E. Douglas, 1997    Partitioning of symbiotic bacteria between generations of an insect: a quantitative study of a *Buchnera sp.* in the pea aphid (*Acyrthosiphon pisum*) reared at different temperatures. Appl. Environ. Microbiol. 63: 3294–3296.

Hunt, B. G., K. M. Glastad, S. V. Yi, and M. A. Goodisman, 2013    The function of intragenic DNA methylation: Insights from insect epigenomes. Integr. Comp. Biol. 53: 319–328. https://doi.org/10.1093/icb/ict003

International Aphid Genomics Consortium, 2010    Genome sequence of the pea aphid *Acyrthosiphon pisum*. PLoS Biol. 8: e1000313. https://doi.org/10.1371/journal.pbio.1000313

Jain, M., R. Nilsson, S. Sharma, N. Madhusudhan, T. Kitami *et al.*, 2012    Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation. Science 336: 1040–1044. https://doi.org/10.1126/science.1218595

John, R. A., 1995    Pyridoxal phosphate-dependent enzymes. Biochim. Biophys. Acta 1248: 81–96. https://doi.org/10.1016/0167-4838(95)00025-P

Jones, P. A., 2012    Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. 13: 484–492. https://doi.org/10.1038/nrg3230

Kono, M., R. Koga, M. Shimada, and T. Fukatsu, 2008    Infection dynamics of coexisting beta- and gammaproteobacteria in the nested endosymbiotic system of mealybugs. Appl. Environ. Microbiol. 74: 4175–4184. https://doi.org/10.1128/AEM.00250-08

Krueger, F., and S. R. Andrews, 2011    Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics 27: 1571–1572. https://doi.org/10.1093/bioinformatics/btr167

Kucharski, R., J. Maleszka, S. Foret, and R. Maleszka, 2008    Nutritional control of reproductive status in honeybees via DNA methylation. Science 319: 1827–1830. https://doi.org/10.1126/science.1153069

Labuschagne, C. F., N. J. Van Den Broek, G. M. Mackay, K. H. Vousden, and O. D. Maddocks, 2014    Serine, but not glycine, supports one-carbon metabolism and proliferation of cancer cells. Cell Reports 7: 1248–1258. https://doi.org/10.1016/j.celrep.2014.04.045

Liadouze, I., G. Febvay, J. Guillaud, and G. Bonnot, 1995    Effect of diet on the free amino acid pools of symbiotic and aposymbiotic pea aphids, *Acyrthosiphon pisum*. J. Insect Physiol. 41: 33–40. https://doi.org/10.1016/0022-1910(94)00085-U

Locasale, J. W., 2013    Serine, glycine and one-carbon units: cancer metabolism in full circle. Nat. Rev. Cancer 13: 572–583. https://doi.org/10.1038/nrc3557

Love, M. I., W. Huber, and S. Anders, 2014    Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15: 550. https://doi.org/10.1186/s13059-014-0550-8

Luco, R. F., M. Allo, I. E. Schor, A. R. Kornblihtt, and T. Misteli, 2011    Epigenetics in alternative pre-mRNA splicing. Cell 144: 16–26. https://doi.org/10.1016/j.cell.2010.11.056

Maddocks, O. D., C. R. Berkers, S. M. Mason, L. Zheng, K. Blyth *et al.*, 2013    Serine starvation induces stress and p53-dependent metabolic remodeling in cancer cells. Nature 24: 542–546. https://doi.org/10.1038/nature11743

Mielke, P. W., and K. J. Berry, 2003    Multivariate multiple regression analyses: A permutation method for linear models. Physiol. Rep. 91: 3–9.

Mira, A., and N. A. Moran, 2002    Estimating population size and transmission bottlenecks in maternally transmitted endosymbiotic bacteria. Microb. Ecol. 44: 137–143. https://doi.org/10.1007/s00248-002-0012-9

Mukherjee, K., and A. F. Baudach, 2016    Epigenetic Control of Polyphenism in Aphids, pp. 89–99 in *Biology and Ecology of Aphids*, edited by Vilcinskas, A., CRC Press, Boca Raton, Florida. https://doi.org/10.1201/b19967-5

Nakabachi, A., S. Shigenobu, N. Sakazume, T. Shiraki, Y. Hayashizaki *et al.*, 2005    Transcriptome analysis of the aphid bacteriocyte, the symbiotic

host cell that harbors an endocellular mutualistic bacterium. Buchnera. Proc. Nat. Acad. Sci. 102: 5477–5482. https://doi.org/10.1073/pnas.0409034102

Niculescu, M. D., and S. H. Zeisel, 2002    Diet, methyl donors and DNA methylation: interactions between dietary folate, methionine and choline. J. Nutr. 132: 2333S–2335S. https://doi.org/10.1093/jn/132.8.2333S

Nishikori, K., K. Morioka, T. Kubo, and M. Morioka, 2009    Age- and morph-dependent activation of the lysosomal system and *Buchnera* degradation in aphid endosymbiosis. J. Insect Physiol. 55: 351–357. https://doi.org/10.1016/j.jinsphys.2009.01.001

Park, Y., M. E. Figueroa, L. S. Rozek, and M. A. Sartor, 2014    MethylSig: a whole genome DNA methylation analysis pipeline. Bioinformatics 30: 2414–2422. https://doi.org/10.1093/bioinformatics/btu339

Parkinson, J. F., B. Gobin, and W. O. Hughes, 2016    Heritability of symbiont density reveals distinct regulatory mechanisms in a tripartite symbiosis. Ecol. Evol. 6: 2053–2060. https://doi.org/10.1002/ece3.2005

Pasquier, C., M. Clément, A. Dombrovsky, S. Penaud, M. Da Rocha *et al.*, 2014    Environmentally selected aphid variants in clonality context display differential patterns of methylation in the genome. PLoS One 9: e115022. https://doi.org/10.1371/journal.pone.0115022

Peccoud, J., M. de la Huerta, J. Bonhomme, C. Laurence, Y. Outreman *et al.*, 2014    Widespread host-dependent hybrid unfitness in the pea aphid species complex. Evolution 68: 2983–2995. https://doi.org/10.1111/evo.12478

Perroud, B., J. Lee, N. Valkova, A. Dhirapong, P. Lin *et al.*, 2006    Pathway analysis of kidney cancer using proteomics and metabolic profiling. Mol. Cancer 5: 64. https://doi.org/10.1186/1476-4598-5-64

Pertea, M., D. Kim, G. M. Pertea, J. T. Leek, and S. L. Salzberg, 2016    Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. Nat. Protoc. 11: 1650–1667. https://doi.org/10.1038/nprot.2016.095

Pertea, M., G. M. Pertea, C. M. Antonescu, T. C. Chang, J. T. Mendell *et al.*, 2015    StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. 33: 290–295. https://doi.org/10.1038/nbt.3122

Poliakov, A., C. W. Russell, L. Ponnala, H. J. Hoops, Q. Sun *et al.*, 2011    Large-scale label-free quantitative proteomics of the pea aphid-*Buchnera* symbiosis. Mol. Cell. Proteomics 10: M110.007039. https://doi.org/10.1074/mcp.M110.007039

Price, D. R., H. Feng, J. D. Baker, S. Bavan, C. W. Leutje *et al.*, 2014    Aphid amino acid transporter regulates glutamine supply to intracellular bacterial symbionts. Proc. Natl. Acad. Sci. USA 111: 320–325. https://doi.org/10.1073/pnas.1306068111

Provataris, P., K. Meusemann, O. Niehuis, S. Grath, and B. Misof, 2018    Signatures of DNA methylation across insects suggest reduced DNA methylation levels in Holometabola. Genome Biol. Evol. 10: 1185–1197. https://doi.org/10.1093/gbe/evy066

Romanoski, C. E., C. K. Glass, H. G. Stunnenberg, L. Wilson, and G. Almouzni, 2015    Epigenomics: Roadmap for regulation. Nature 518: 314–316. https://doi.org/10.1038/518314a

Sandström, J., and J. Pettersson, 1994    Amino acid composition of phloem sap and the relation to intraspecific variation in pea aphid (*Acyrthosiphon pisum*) performance. J. Insect Physiol. 40: 947–955. https://doi.org/10.1016/0022-1910(94)90133-3

Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki, and H. Ishikawa, 2000    Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera sp.* APS. Nature 407: 81–86. https://doi.org/10.1038/35024074

Shukla, S., E. Kavak, M. Gregory, M. Imashimizu, B. Shutinoski *et al.*, 2011    CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature 479: 74–79. https://doi.org/10.1038/nature10442

Simonet, P., G. Duport, K. Gaget, M. Weiss-Gayet, S. Colella *et al.*, 2016    Direct flow cytometry measurements reveal a fine-tuning of symbiotic cell dynamics according to the host developmental needs in aphid symbiosis. Sci. Rep. 6: 19967. https://doi.org/10.1038/srep19967

Stoll, S., H. Feldhaar, M. J. Fraunholz, and R. Gross, 2010    Bacteriocyte dynamics during development of a holometabolous insect, the carpenter

ant *Camponotus floridanus*. BMC Microbiol. 10: 308. https://doi.org/10.1186/1471-2180-10-308

Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert *et al.*, 2005 Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA 102: 15545–15550. https://doi.org/10.1073/pnas.0506580102

Sudakaran, S., C. Kost, and M. Kaltenpoth, 2017 Symbiont acquisition and replancement as a source of ecological innovation. Trends Microbiol. 25: 375–390. https://doi.org/10.1016/j.tim.2017.02.014

Suzuki, M. M., A. R. W. Kerr, D. De Sousa, and A. Bird, 2007 CpG methylation is targeted to transcription units in an invertebrate genome. Genome Res. 17: 625–631. https://doi.org/10.1101/gr.6163007

Tammen, S. A., S. Friso, and S. W. Choi, 2013 Epigenetics: The link between nature and nurture. Mol. Aspects Med. 34: 753–764. https://doi.org/10.1016/j.mam.2012.07.018

Vander Heiden, M. G., L. C. Cantley, and C. B. Thompson, 2009 Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science 324: 1029–1033. https://doi.org/10.1126/science.1160809

Vigneron, A., F. Masson, A. Vallier, S. Balmand, M. Rey *et al.*, 2014 Insects recycle endosymbionts when the benefit is over. Curr. Biol. 24: 2267–2273. https://doi.org/10.1016/j.cub.2014.07.065

Vogel, K. J., and N. A. Moran, 2011 Sources of variation in dietary requirements in an obligate nutritional symbiosis. Proc. R. Soc. Lond. B Biol. Sci. 278: 115–121. https://doi.org/10.1098/rspb.2010.1304

Walsh, T. K., J. A. Brisson, H. M. Robertson, K. Gordon, S. Jaubert-Possamai *et al.*, 2010 A functional DNA methylation system in the pea aphid, *Acyrthosiphon pisum*. Insect Mol. Biol. 19: 215–228. https://doi.org/10.1111/j.1365-2583.2009.00974.x

Wedd, L., R. Kucharski, and R. Maleszka, 2016 Differentially methylated obligatory epialleles modulate context-dependent *LAM* gene expression in the honeybee *Apis mellifera*. Epigenetics 11: 1–10. https://doi.org/10.1080/15592294.2015.1107695

Wilson, A. C., P. D. Ashton, F. Calevro, H. Charles, S. Colella *et al.*, 2010 Genomic insight into the amino acid relations of the pea aphid, *Acyrthosiphon pisum*, with its symbiotic bacterium *Buchnera aphidicola*. Insect Mol. Biol. 19: 249–258. https://doi.org/10.1111/j.1365-2583.2009.00942.x

Yan, H., R. Bonasio, D. F. Simola, J. Liebig, S. L. Berger *et al.*, 2015 DNA methylation in social insects: how epigenetics can control behavior and longevity. Annu. Rev. Entomol. 60: 435–452. https://doi.org/10.1146/annurev-ento-010814-020803

Zemach, A., I. E. McDaniel, P. Silva, and D. Zilberman, 2010 Genome-wide evolutionary analysis of eukaryotic DNA methylation. Scince 328: 916–919. https://doi.org/10.1126/science.1186366

Zientz, E., T. Dandekar, and R. Gross, 2004 Metabolic interdependence of obligate intracellular bacteria and their insect hosts. Microbiol. Mol. Biol. Rev. 68: 745–770. https://doi.org/10.1128/MMBR.68.4.745-770.2004

*Communicating editor: S. Scofield*