



**HAL**  
open science

# Adaptive variation of buchnera endosymbiont density in aphid host acyrthosiphon pisum controlled by environmental conditions

F. Neiers, Jean-Michel Saliou, Loïc Briand, Alain Robichon

## ► To cite this version:

F. Neiers, Jean-Michel Saliou, Loïc Briand, Alain Robichon. Adaptive variation of buchnera endosymbiont density in aphid host acyrthosiphon pisum controlled by environmental conditions. *ACS Omega*, 2021, 6 (28), pp.17902-17914. 10.1021/acsomega.1c01465 . hal-03296049

**HAL Id: hal-03296049**

**<https://hal.inrae.fr/hal-03296049v1>**

Submitted on 22 Jul 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Adaptive Variation of *Buchnera* Endosymbiont Density in Aphid Host *Acyrtosiphon pisum* Controlled by Environmental Conditions

Fabrice Neiers, Jean-Michel Saliou, Loïc Briand, and Alain Robichon\*

Cite This: <https://doi.org/10.1021/acsomega.1c01465>

Read Online

ACCESS |



Metrics &amp; More

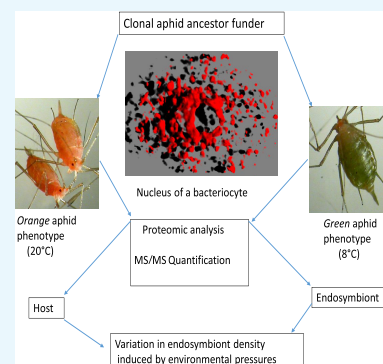


Article Recommendations



Supporting Information

**ABSTRACT:** The scarcity of transcriptional regulatory genes in *Buchnera aphidicola*, an obligate endosymbiont in aphids, suggests the stability of expressed gene patterns and metabolic pathways. This observation argues in favor of the hypothesis that this endosymbiont bacteria might contribute little to the host adaptation when aphid hosts are facing challenging fluctuating environment. Finding evidence for the increased expression or silenced genes involved in metabolic pathways under the pressure of stress conditions and/or a given environment has been challenging for experimenters with this bacterial symbiotic model. Transcriptomic data have shown that *Buchnera* gene expression changes are confined to a narrow range when the aphids face brutal environmental variations. In this report, we demonstrate that instead of manipulating individual genes, the conditions may act on the relative mass of endosymbiont corresponding to the needs of the host. The control of the fluctuating number of endosymbiont cells per individual host appears to be an unexpected regulatory modality that contributes to the adaptation of aphids to their environment. This feature may account for the success of the symbiotic advantages in overcoming the drastic changes in temperature and food supplies during evolution.



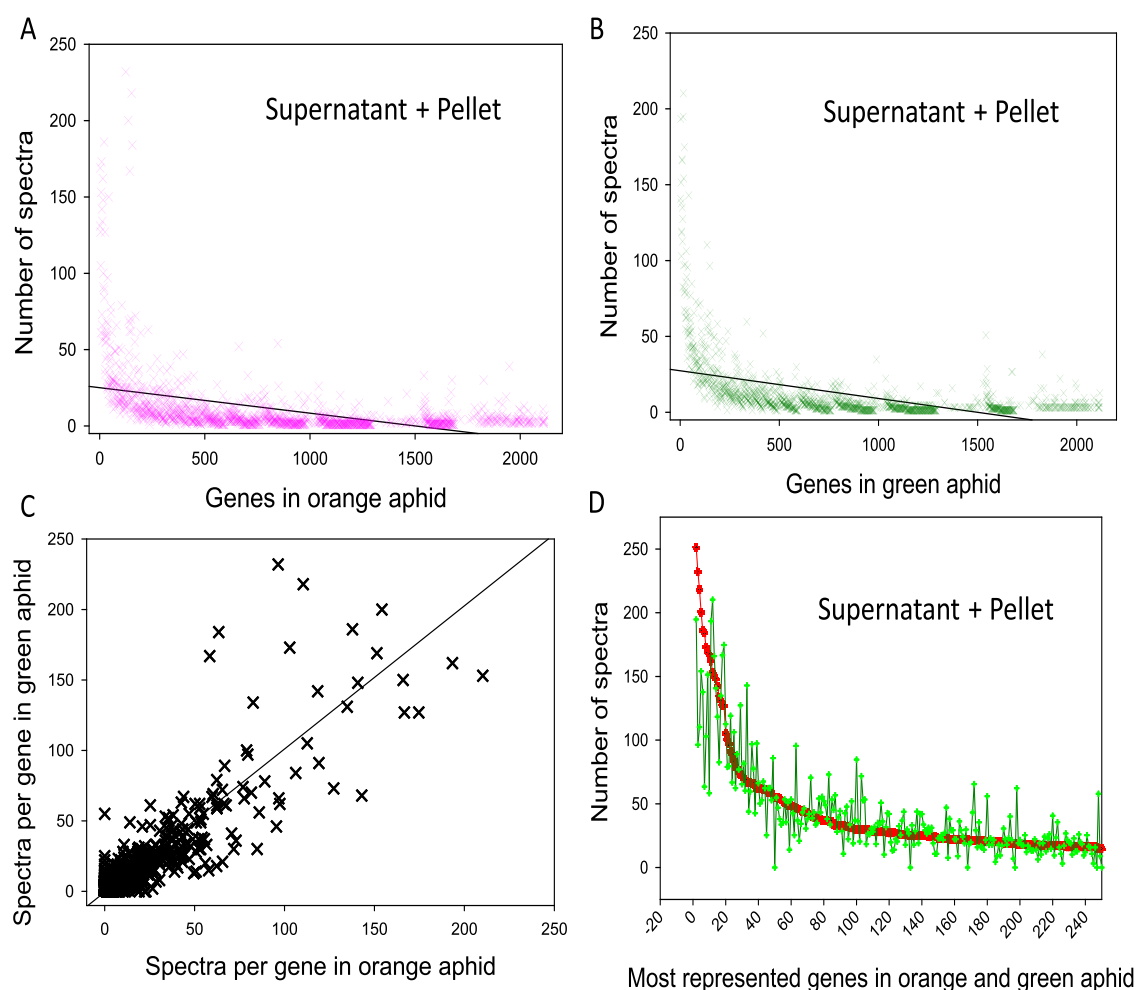
## INTRODUCTION

Certain insect species such as aphid shelter obligate symbiotic bacteria with highly reduced genomes.<sup>1,2</sup> This mutually beneficial symbiosis is particularly spectacular in aphids; an obligate aphid endosymbiont, *Buchnera aphidicola*, is a proteobacterium phylogenetically very close to *Escherichia coli*.<sup>1,2</sup> This association started more than 150 million years ago and is an evolutionary achievement that provides aphids with the essential enzymes that they lack. As a consequence, aphids can successfully live on chemical compositions with reduced complexity that is characteristic for the phloem sap of plants.<sup>3–5</sup> The host species studied in this report is the pea aphid *Acyrtosiphon pisum* whose genome has been sequenced and includes more than 34 000 computationally predicted genes for a total genome size of approximately 464 Mb.<sup>6</sup> The genome of its obligate symbiont *Buchnera* has also been sequenced; the sequencing indicates that the genome underwent a massive reduction in size through co-evolution down to one-seventh of the *E. coli* genome size (0.6 versus 4.2 Mb).<sup>1</sup> The loss of DNA repair, homologous recombination machinery, and synthesis of lipopolysaccharides, which are major components of bacterial membrane, are remarkable.<sup>1</sup> Another remarkable feature of the *Buchnera* endosymbionts is that they are enveloped by a synaptosomal membrane synthesized by the host.<sup>7</sup> These bacteria are mostly housed in giant polynucleated host cells called bacteriocytes that are transmitted vertically from a parthenogenetic mother to the embryos by membrane fusion.<sup>7–11</sup> Aphid bacteriocyte and *Buchnera* gene expression show complex coordination and an

intertwined metabolic pathway; a segment of the pathway is accomplished in the bacteriocyte cytosol, and another segment takes place inside the endosymbiont.<sup>7,12,13</sup> This coordination results in an integrated network of the genes acting in the host or symbiont cytosols. Symbiosis resulting from co-evolution of the aphid and *Buchnera* genomes explains the evolutionary success of the aphid branch that has faced diverse ecological and climate challenges and/or catastrophes.<sup>9–11</sup> Little is known about the coordinated growth/proliferation of the host/symbiont cells. Thus, it remains unknown how the symbiont proliferation is stopped.<sup>14,15</sup> The turnover of the symbionts along their average life span and the mechanism of their lysosomal degradation is surprisingly understudied.<sup>14,15</sup> In this report, we explored whether variations in *Buchnera* cell density are induced by the outcome of environmental stress. If the relative mass of the endosymbionts varies within individual host, which is presently unknown, we assume that these quantitative changes may diverge depending on the nutrient needs and stress types. In this report, we aimed to detect and measure these changes, which may reflect the physiological responses of the host by eventually reducing or increasing the

Received: March 18, 2021

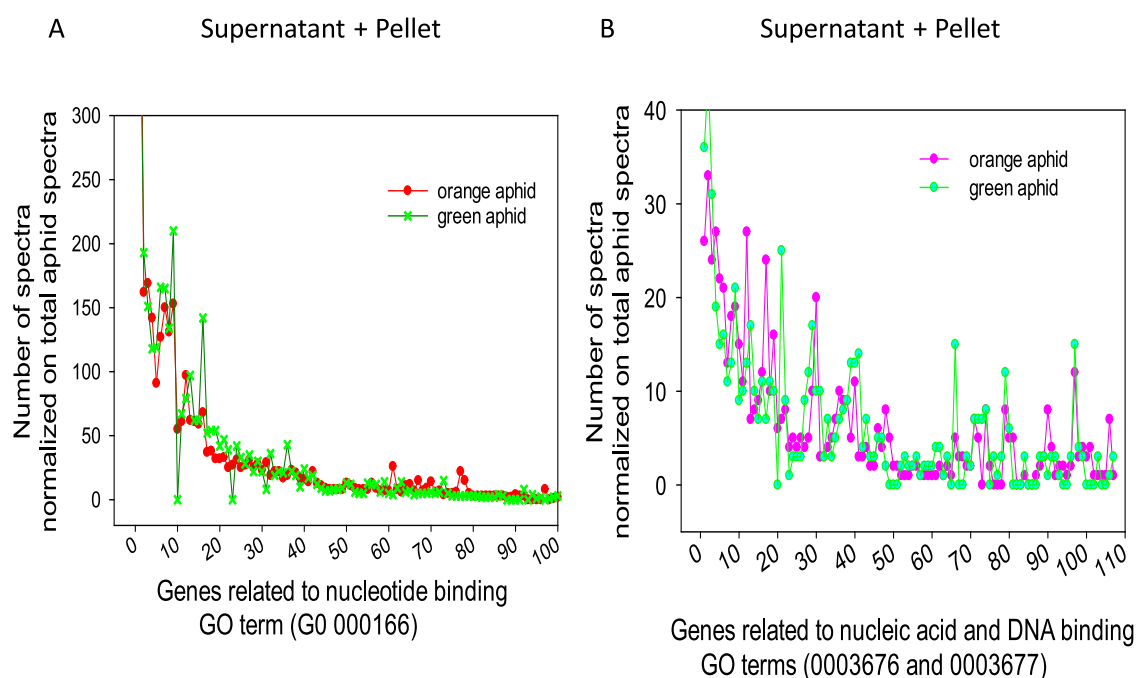
Accepted: June 28, 2021



**Figure 1.** Analysis of the number of spectra per gene in the green and orange aphid host phenotypes. The supernatants and pellets obtained from the green and orange aphid extracts were used for tandem mass spectrometry (MS/MS) peptide identification and relative quantification. After normalization based on the total number of aphid spectra performed separately in the supernatant and pellet lists, each identified protein was assigned a number of spectra calculated as a sum of adding two extract components (pellet and supernatant). Graphs refer to one unique experiment. (A, B) Number of spectra in the orange and green aphid, respectively, related to more than 3000 proteins for a total potential of 34 000. (C) Linear regression plot of the data shown in (A) and (B) ( $r = 0.98$ ,  $n = 2138$ ). (D) Representation of 240 most represented proteins of the green and orange phenotypes. The orange plot has been organized in the descending order of numbers per protein, and the corresponding green plot values are shown for comparison.

growth rates of the bacterial population. If confirmed by experimental evidence, quantitative variations of *Buchnera* cells can be viewed as a systemic mechanism of ecological adaptation in the context of the loss of the transcriptional regulatory genes. The quantitative variation in the number of endosymbiont in their host under environmental stress appears to be unreported or understudied. We aimed to document whether symbiont-driven adaptive traits exist and can account for the observed phenotypic changes of the host produced by a shift by 12 °C temperature over multiple generations (20 to 8 °C). Our aphid host model is parthenogenetic; we excluded meiotic recombination in the gametes and the resulting mixing of the chromosome that can influence the endosymbiont density variations. Two clones originated from a unique parthenogenetic founder mother were selected and maintained under spring cold conditions (8 °C) and at summer temperature (20 °C) for 5 years under stable conditions in the laboratory (photoperiodicity: 16 h light and 8 h dark; fed on the *Sativa pisum* plant). The cold-adapted clone (8 °C) has selected a green pigment, and a moderate temperature (20 °C)

resulted in the maintenance of an orange-colored phenotype.<sup>16</sup> The aphid phenotypic plasticity producing color changes in response to temperature and in the clonality context has been reported by our laboratory.<sup>16</sup> Therefore, we aimed to measure the variations of the symbiont abundance in these two clonal aphid hosts (named orange and green phenotypes) sharing the same ancestor by the metrics of the proteomic wide-scale analysis. To measure the cell density of endosymbiont in aphid or in other insect species turned out to be challenging. The reason resides in the fragility of those cells during the isolation process when they are deprived of the environment and architecture within their host. Moreover, the *Buchnera* cells are polyploid from 50 to 200 chromosomes per cell, which is unique in bacteria era.<sup>17–19</sup> The polyploidy may be derived from the loss of essential loci involved in regulating and assisting the chromosome replication during the cell cycle, through an evolutionary drift process of genome reduction.<sup>20–22</sup> This massive genome shrinkage was selected to obtain an optimized efficiency of symbiotic mutualism to the detriment of proper chromosome segregation, which might



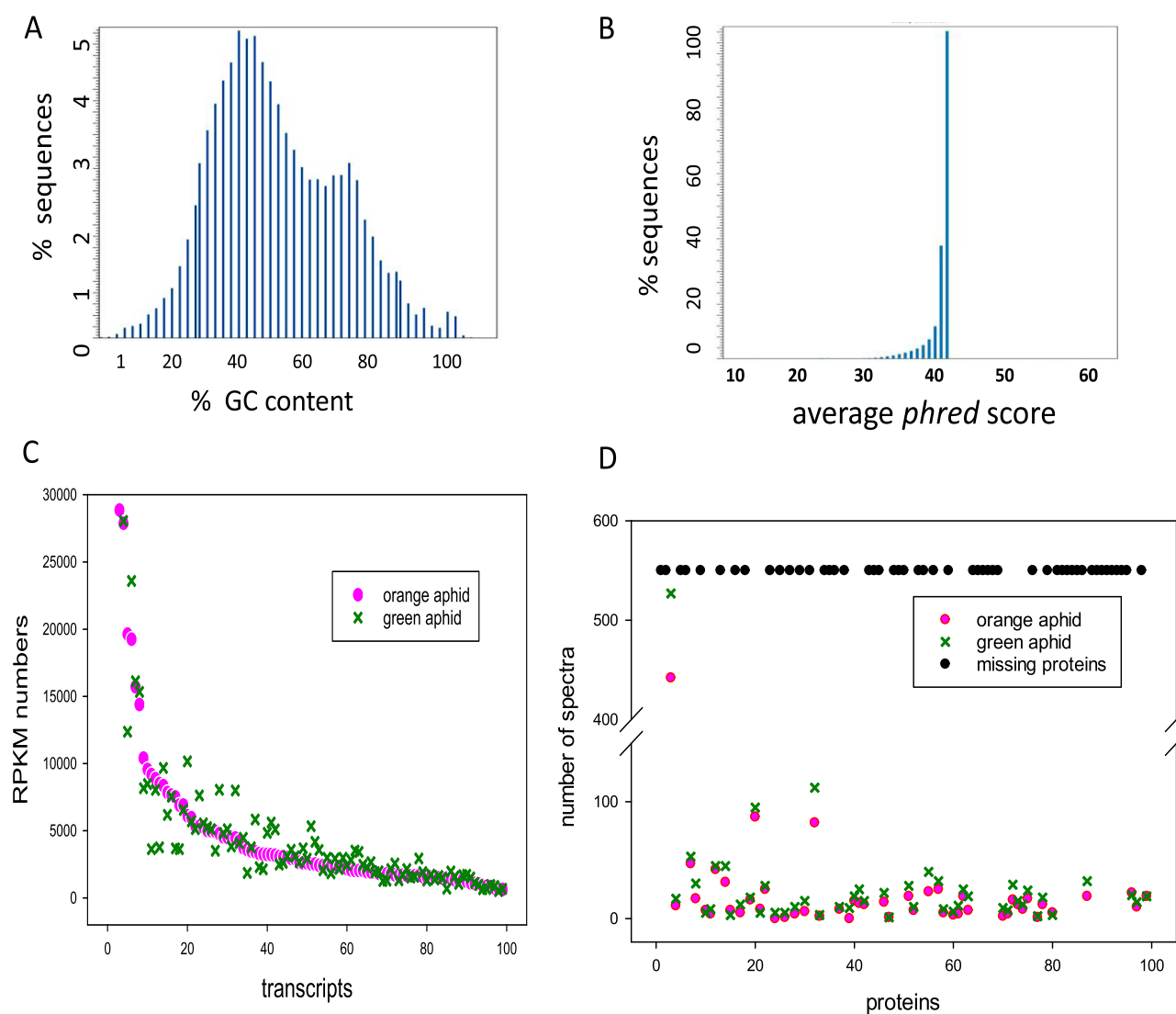
**Figure 2.** Analysis of the total number of spectra per protein depending on GO annotation referring to nucleotide binding and nucleic acid/DNA binding. Two sublists corresponding to the indicated GO terms retrieved from one unique experiment were analyzed for comparison between the green and orange phenotypes. The counting was processed with the combined supernatant and pellet fractions. Graphs show the quantification of the spectra for each gene in the orange and green aphid after normalization of the counts by the total number of aphid spectra (see methods and Table S2 for normalized counts). (A) Genes Ontology annotation GO:000166 (nucleotide-binding function) and (B) GO:0003676/GO:0003677 (DNA- and nucleic acid-binding function).

have led to asymmetry/asynchrony between cell division and replication.<sup>20–22</sup> This might explain the fact that cells do not divide at the same speed that the DNA replicates, and this asynchrony is more pronounced when aphids are aging.<sup>7</sup> The consequence is that the level of messengers should reflect the number of chromosomes (these cells do not have regulatory genes) but not the protein amounts. Protein levels likely would not match the number of chromosomes either (cells, regardless of the number of chromosomes they harbor, have more or less the same size except for the old cells that tend to fuse). On the other hand, the high fragility of *Buchnera* cell external membrane renders those cells highly vulnerable during the isolation process. We have tried a few years ago to isolate *Buchnera* cells: these cells are huge (5  $\mu\text{m}$  diameter), and their membranes are deprived of lipopolysaccharides that makes them easy to brake and/or fuse during the process of grinding the aphid. The difficulties to isolate living *Buchnera* cells after extraction/isolation are aggravated by the fact the host reticulum makes a synaptosomal membrane that envelops them.<sup>7</sup> The breaking down of this membrane weakens *Buchnera* cells that tend to lyse during the isolation process. If the grinding of aphid is soft, many *Buchnera* cells precipitate with the Golgi and endoplasmic reticulum. Indeed, the cell cytometry procedure after extraction might lead to questionable results, although rigorous experimental procedures might minimize variability. The same way, quantitative polymerase chain reaction (qPCR) determination will not approach the cell density but more likely the polyploidy status. This is for this reason that the proteomic approach turned out to be an alternative and the most reliable index for measuring endosymbiont cell density. We therefore have compared the total protein amount relevant of the endosymbiont in two aphid contexts without preliminary isolation of the cells. The

determinations were performed in the total extracts of aphids, which allowed us to use the protein component related to the aphid as an internal control to normalize the corresponding protein quantity of the endosymbiont. In parallel, this proteomic wide-scale analysis allowed us to investigate individual protein variations to eventually unmask certain up- or downregulated functional networks. Moreover, parallel RNAseq determinations were conducted and showed the poor correlation of messenger RNA (mRNA)/protein in aphid *Buchnera*, which can be explained by the highly variable polyploidy of *Buchnera* cells.<sup>17–19</sup> This proteomic approach allowed us to unambiguously establish a correlation between natural phenotypic of the host and the dynamics of the symbiont communities.

## RESULTS

**Comparative Global Analysis of Green and Orange Aphid Host Spectra.** The full lists of proteins along with their assigned number of peptides and spectra are reported for the green and orange aphid and for two separate determinations (Table S1). The analysis of proteome of *Buchnera* endosymbionts and their corresponding hosts was conducted in the global extract without previous fractionation of the two entities. Briefly, 3100 proteins in the aphid host and 183 proteins in the endosymbiont were identified by mass spectrometry (MS) analysis (Table S1). Figure 1 shows the results of one unique experiment because of the highly similar numbers between the orange and green samples in the two separate analyses. Figure 1A,B shows the normalized profiles of each protein found in the lists for the green and orange aphid host phenotypes. The 3100 proteins were detected from a total of 34 000 proteins and isoforms. The linear regression plot value ( $r = 0.908$ ) is in agreement with the stability of the



**Figure 3.** Comparison between RNAseq RPKM and proteomic spectra in *Buchnera* cells. The graphs of the %GC content and average Phred score were provided by the company that performed the RNAseq analysis and emphasize the quality control of the datasets (A, B). The refined analysis of RPKMs related to *Buchnera* for the orange and green aphid is presented in Table S3. The 100 highest numbers of RPKM obtained with the average determinations on three independent RNA samples are listed for the orange and green aphids. The orange plot has been organized in the descending order of numbers for which the corresponding green plot values are shown (C). The protein spectra for the corresponding RPKM messengers shown in (C) are displayed in (D). Protein spectra numbers are the sum of two separate determinations. The row of black spots represents the messengers listed in (C) that were not found in any determination of proteomic analysis (total four determinations: two for the green and two for the orange aphids).

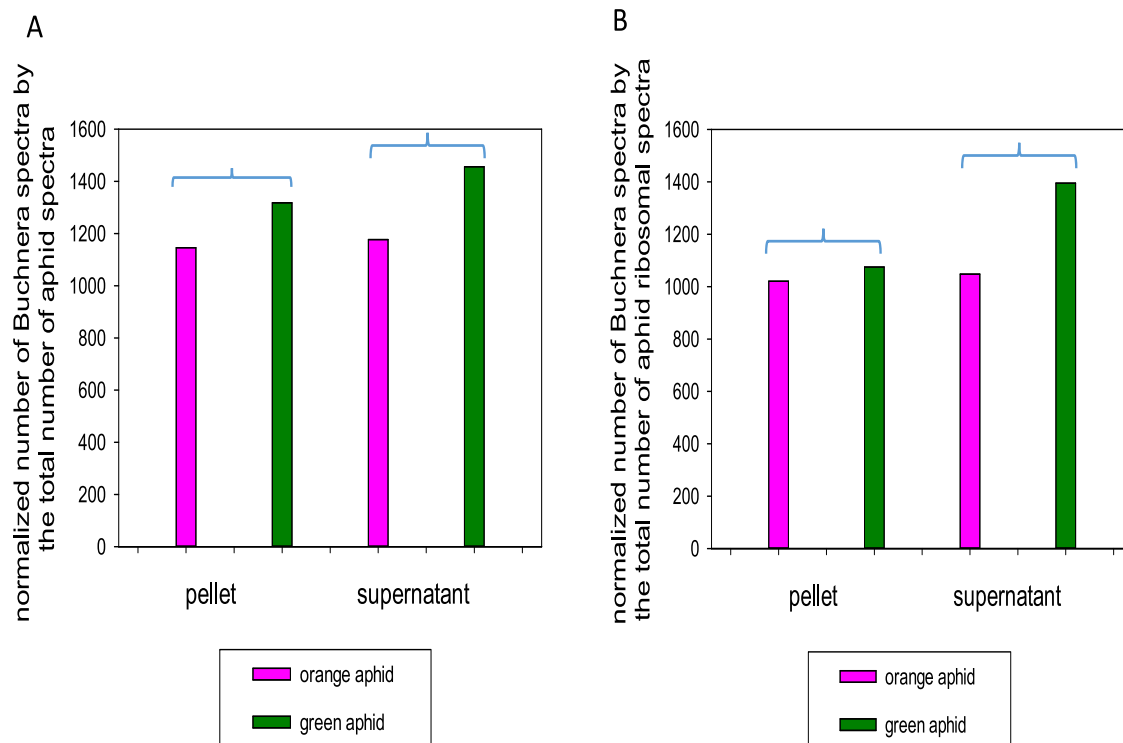
proteome between the phenotypes although numerous individual proteins show certain variability (Figure 1C). Figure 1D represents the variations in the 240 most abundant identified proteins in the orange phenotype linearly organized in a descending order and the corresponding values in the green phenotype. Despite the up and down variations in this limited panel, the global trend confirms the relative stability of the protein levels between the two environmentally induced phenotypes. Figure 2A,B shows the comparative spectral numbers restricted to the most preminent Gene Ontology (GO) terms (GO:0000166 related to nucleotide binding; GO:00003676 for nucleic acid binding and GO:00003677 for DNA binding) retrieved from the nonordered MS lists after normalization based on the total number of aphid spectra (Table S2 for detailed protein datasets). In both cases, the analysis showed little variations within the two groups of genes between orange and green aphid phenotypes and confirmed

the overall stability of the transcriptome when the two aphid phenotypes are compared. In conclusion, the same pattern prevails in different panels of Figures 1 and 2, thus confirming the stability of the proteome between the two clonal temperature-selected aphid phenotypes that originate from a unique parthenogenetic ancestor.

**Comparative Quantification of the *Buchnera* Endosymbiont Cell Density in the Green and Orange *A. pisum*.** Standard RNAseq analysis was performed with the full RNA extract of the aphids (host + endosymbionts). The representative guanine–cytosine (GC) profile content across all samples and the *Q* score distribution in all samples (average Phred score) referring to the six aphid extracts (triplicate for the green and orange aphids) qualify the assay to assess the reliability of comparison studies (Figure 3A,B). Additionally, the global RNAseq reads referring to aphid host or *Buchnera* endosymbiont were counted to quantify the percentage of the



### *Buchnera*: Normalized spectra by total aphid spectra and aphid ribosomal spectra in pellet and supernatant fraction

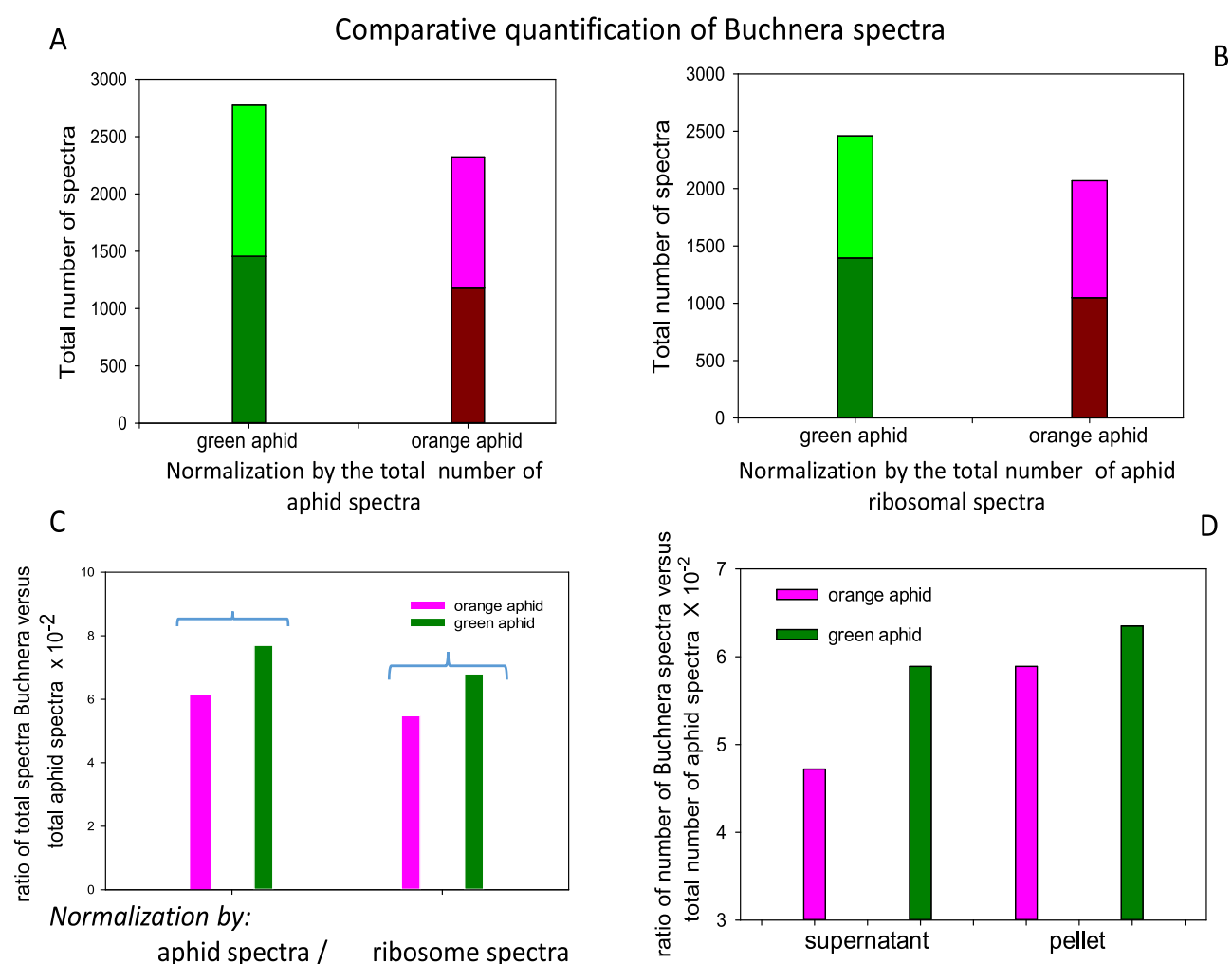


**Figure 4.** Comparative number of spectra related to *Buchnera* endosymbiont in the green and orange aphid. The graphs represent the comparison between the *Buchnera* total spectra in the two aphid phenotypic contexts. All of the individual spectra of each protein were added separately for the supernatant and pellet, and the number were compared between the two phenotypes. The normalized determinations were performed using total aphid spectra (A) and aphid ribosomal proteins spectra (B) (see Table S1 and the Experimental Section for details). Numbers represent the sum of two separate experiments. The two modes of normalization highlight the increase of *Buchnera* spectra in the green aphid phenotype.

mapped fragments relevant to the endosymbiont or the aphid for each sample. The *Buchnera* endosymbiont RNA component was 8.31 and 8.92% of the total, respectively, for the orange and green aphid based on three repeats. The percentage of reads that matched to the *Buchnera* mRNAs was not significantly divergent between the two phenotypes. Figure 3C shows the comparative RNAseq plots corresponding to the 100 most abundant RNAs for the orange and green *Buchnera* in which little differences emerge. Moreover, a correlation study between the RNAseq RPKMs and the protein spectra showed a very weak correspondence between mRNA and protein amounts (Figure 3C,D; Tables S1 and S3). Basically by considering the 100 highest number of RPKMs of both the orange and green aphids (based on the means of three determinations), about half of the corresponding proteins were missing in the mass spectrometry datasets (Figure 3D; Tables S1 and S3). A very weak correspondence was found between the quantification of the RPKMs (average of three repeats) and the protein spectra (average of two repeats) (Figure 3D).

Then, we tried to address the question whether the endosymbiont protein amount varies or is stable in the two ecological contexts that shape the orange and green phenotypes. Considering that the transcriptome and proteome are globally stable between the two aphid host phenotypes, we analyzed the total number of the spectra referring to *Buchnera* in the two contexts. The sum of each spectrum related to individual protein and corrected by two different indexes of normalization allowed us to explore a robust comparison of the

*Buchnera* endosymbiont component between these two phenotypes (Figure 4). The spectra of the two separate determinations were added for each compartment (supernatant + pellet) after normalization of numbers by total aphid spectra (Figure 4A). In summary, the increase in the total number of spectra rose to 13% (1146/1318) in the pellet component and to 20% (1177/1456) in the supernatant component for *Buchnera* in the green phenotype. A confirmatory analysis using the aphid ribosomal protein normalization instead of the total aphid spectra number gave an increase of *Buchnera* proteins up to 25% (1048/1395) for the supernatant component and only a 5% increase (1021/1065) for the pellet component in the *Buchnera* of green phenotype (Figure 4B). An alternative graph is shown to illustrate the sum of the two separate experiments (pellet + supernatant) to highlight the differences using both methods of normalization (total aphid and ribosomal spectra) (Figure 5A,B). The two proteome determinations allowed us to calculate a percentage of *Buchnera* spectra versus the total (host + endosymbiont) in the two phenotype contexts. These percentages gave 6.1 and 7.7% for the *Buchnera* component, respectively, for the orange and green phenotypes normalized by total aphid spectra and 5.5 and 6.8%, respectively, normalized by ribosomal spectra (Figure 5C). Another procedure included calculation of the ratio *Buchnera*/total aphid counts without previous normalization (Figure 5D). The ratios were 4.79 and 5.89% in the supernatant component and 5.89 and 6.38% in the pellet component for the orange and



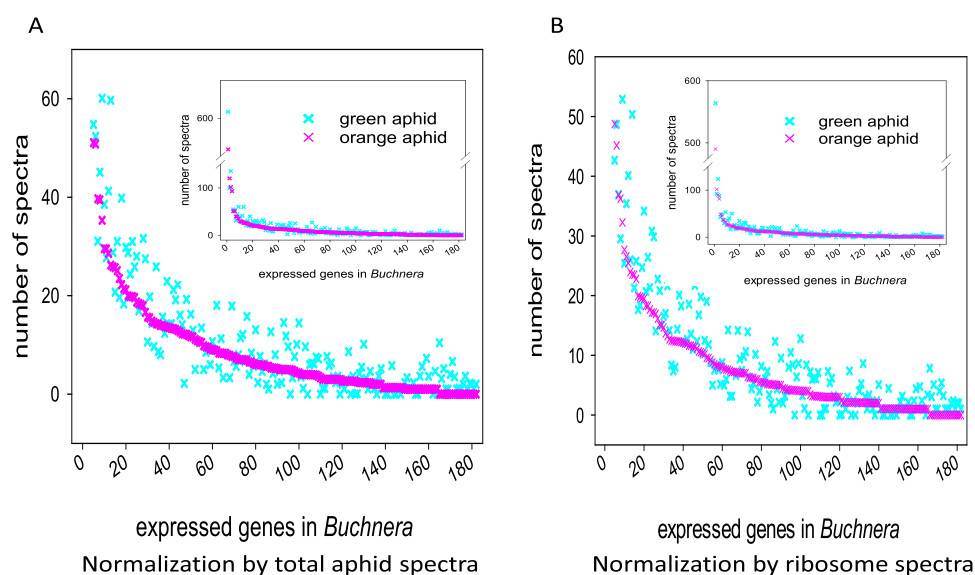
**Figure 5.** Ratio of total spectra of endosymbiont with and without normalization in the two phenotypic contexts. The global spectra counts related to the endosymbiont in each compartment (supernatant and pellet) were added to compare values in both phenotypic aphid contexts. Values are the sum of two separate experiments. The dark and light green or orange colors indicate the pellet and supernatant components, respectively (A) and (B). Then, mean values of the two experiments corresponding to total aphid and endosymbiont spectra were calculated to assess the ratio endosymbiont/aphid. A comparison of the corresponding numbers obtained with the two modes of normalization is shown in (C). Then, the percentage of the global counts related to the endosymbiont versus the total counts in the aphid for each compartment (supernatant and pellet) were determined without normalization. The values represent the mean of two experiments (D).

green phenotypes, respectively. These percentages gave 19 and 8.5% increases of *Buchnera* proteins for the green phenotype in the supernatant and pellet components, respectively.

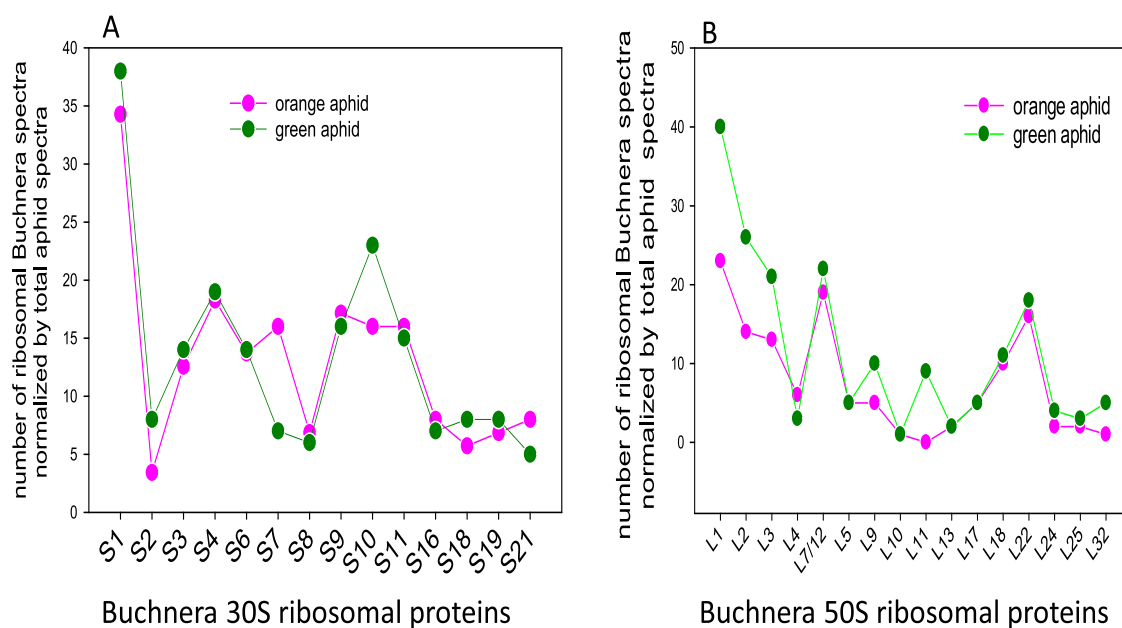
To graphically illustrate these observations, the comparative full spectrum of *Buchnera* proteome (183 proteins) detected in the two contexts is presented after normalization by the two procedures (the orange plot has been organized in the descending order of spectra numbers per protein, and the corresponding protein in green plot has been adjusted) (Figure 6A,B). The represented plots show the average values between the two separate proteome determinations. We observed a significant increase in the number of spectra per protein in the green phenotype although some proteins showed an inverse trend. Regarding the statistical analysis, the two proteomic lists indicating the spectra number for each *Buchnera* protein detected in the two phenotype contexts were compared to assess the significance of the increase in the green host. Using the Spearman rank order correlation test, the correlation coefficient was  $>0.77$  with  $P$ -values  $<0.001$  for all of the determinations (proteomic assays 1 and 2, sum of both assays,

normalization by total aphid spectra or by ribosomal aphid spectra), which indicates a strong positive co-evolution between the two lists. On the other hand, Wilcoxon signed rank test (a nonparametric statistical test that compares two ranked paired groups) and paired  $t$ -test (compare the means of the two lists) gave  $P$ -values  $<0.001$  for all of the same determinations, which indicates a significant increase of the full protein spectrum for *Buchnera* in green context (see detailed quantified indexes in the Experimental Section). Finally, two sublists, ribosomal proteins 50S and 30S, matching the *Buchnera* sequences were examined after normalization with the full aphid spectra leading to the conclusion that the spectral numbers of *Buchnera* green phenotype are slightly higher than in the range counterpart (Figure 7A,B).

**Analysis of Proteome of the Green and Orange Aphid Hosts.** The 3100 proteins identified in aphid are reported in Table S1. Due to the low number of biochemical studies and the lack of protein conservation motifs that were bio-computationally searched, a large number of the identified proteins are uncharacterized. Arbitrarily, 10 proteins present-



**Figure 6.** Comparison of *Buchnera* protein spectra in two aphid phenotypic context. The list of 180 *Buchnera* proteins and their corresponding spectra are shown for comparison after normalization based on the total spectra of aphid (A) or by aphid ribosomal spectra (B). For graphical constraints, the values exceeding 60 spectra are omitted to highlight detailed mainstream proteins. The full spectrum is shown in boxes (see Table S1 for the full spectrum quantification by the two methods). The orange plot has been organized in the descending order of numbers per protein and the corresponding green plot values were, respectively, added for comparison.



**Figure 7.** Comparative number of spectra related to a subset of proteins (30S and 50S ribosomal proteins) in *Buchnera* in the two aphid phenotypes. The numbers of spectra were normalized versus total aphid spectra separately for the pellet and supernatant. Total numbers of spectra of the 30S ribosomal proteins (supernatant + pellet) in the green and orange phenotypes are comparatively represented in the graphs (A). The numbers are the sum of two separate experiments. The same analysis for the 50S proteins was conducted (B).

ing the most elevated spectra in the two aphid host phenotypes are presented in Table 1. As expected, these latter proteins have diverse functions including energy metabolism, translation machinery, or cytoskeletal proteins such as actin (Tables 1 and S1). Few proteins presenting a slight up- or down-regulation in either host phenotypes were observed, whereas no protein with a large spectra number was drastically reduced in one or the other phenotype (Table S1). Interestingly, a chitin-binding protein (J9K245) that harbors a chitin-binding domain known to enhance the chitinase activity was identified.<sup>23</sup> Chitinase regulates the cellular growth and

proliferation and plays a role in defending the host cells against pathogens.<sup>23</sup> One role of this protein in aphid might be to modify the *Buchnera* proliferation depending on ecological conditions, and this role could have been acquired during the co-evolution of host/symbiont and/or drift of an original function.

**Analysis of Proteome of the Green and Orange *Buchnera* Symbiont.** We identified 183 proteins belonging to *B. aphidicola* in the aphid extracts in the supernatant and pellet fractions (see Table S1). Genes ontology annotation and analysis allowed us to classify the proteins according to various

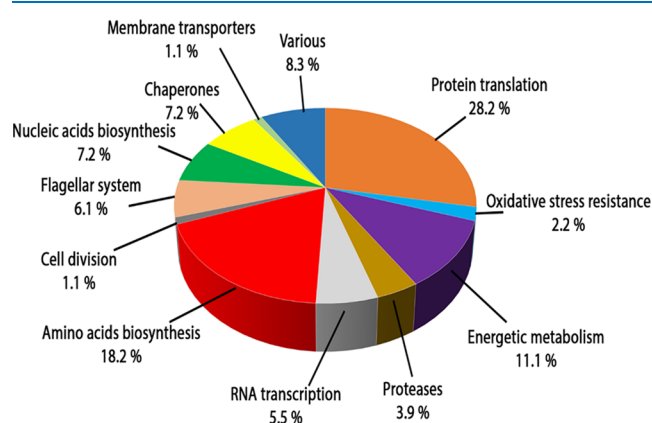


**Table 1. Ten Most Abundant Proteins Identified in Aphid by the Number of Spectra<sup>a</sup>**

protein name	hypothetical function	accession number	theoretical mass (Da)	peptide (number of spectra)
myosin heavy chain	motor proteins	J9K5P5	224 557	183 (423, 374, 199, 104, 87, 286, 131, 400)
paramyosin	motor proteins	J9JLC9	102 700	94 (93, 241, 30, 65, 39, 159, 31, 234)
uncharacterized protein	unknown	J9JKL8	965 258	286 (157, 194, 64, 55, 23, 121, 54, 165)
uncharacterized protein	unknown	J9K5E2	258 943	103 (170, 42, 133, 44, 27, 21, 75, 43, 144)
uncharacterized protein	unknown	X1X113	754 520	221 (125, 138, 44, 27, 21, 75, 43, 144)
ATP synthase subunit $\beta$	energy metabolism	X1WI47	61 799	44 (67, 83, 56, 44, 48, 107, 65, 129)
actin	motor proteins	X2FB37	41 573	24 (55, 52, 47, 7, 60, 81, 118, 145)
molecular chaperone (Hsp70)	chaperone	J9K309	72 924	44 (95, 68, 94, 17, 42, 54, 72, 73)
uncharacterized protein	unknown	J9K2H5	48 022	24 (64, 62, 72, 22, 37, 70, 61, 90)
uncharacterized protein	unknown	J9JYX3	59 795	42 (54, 64, 69, 31, 29, 77, 55, 92)

<sup>a</sup>Different characteristics of the 10 most abundant proteins identified in *A. pisum* are indicated. The number of unique peptides corresponding to the protein in the two experiments are indicated. The values in parentheses correspond to the number of spectra in orange (supernatant, pellet) and green (supernatant, pellet) samples for experiment 1 and orange (supernatant, pellet) and green (supernatant, pellet) samples for experiment 2.

categories as a percentage of the total (Figure 8). A considerable number of proteins (18.2%) are involved in



**Figure 8.** Functional classification of the total *B. aphidicola* proteins. The 183 different proteins identified in *B. aphidicola* were classified into 12 different functions according to Genes Ontology annotation. This includes a category representing miscellaneous proteins (various). For each function, the number of proteins is presented as a percentage of the global number of proteins.

amino acid biosynthesis, confirming the role of the endosymbiont in providing amino acids to the host. The major result resides in the observed stability of most of *Buchnera* individual protein spectra between the two phenotype contexts. However, as expected, strong differences in spectra numbers were found from highly represented to rare proteins, and this was observed in both phenotypes in a similar way. To summarize, we selected arbitrarily 10 highest number of protein spectra referring to the most abundant proteins in both phenotypes for which we noticed the presence of chaperones (GroEL, DnaK, and GroES) and porins (ompF and ompA) (Table 2). Porins are known to perform a variety of functions including mainly transport and diffusion of ions and small molecules.<sup>24,25</sup> The impressive amount of the chaperones and porins suggests that their derived functions were selected by evolution for the needs of symbiotic life. The GroEL proteins related to the flagellar basal bodies have been reported to be intensively present in *Buchnera* endosymbiont, although their role as membrane transporter is still debated.<sup>24,25</sup>

## DISCUSSION

**Lack of Regulatory Modules in *Buchnera* Cell Genome Associated with Aphid Host.** Bacterial symbionts provide essential molecules including certain amino acids,

**Table 2. Ten Most Abundant Proteins Identified in *Buchnera* by Their Number of Spectra<sup>a</sup>**

protein name	hypothetical function	accession number	theoretical mass (Da)	peptide (number of spectra)
molecular chaperone GroEL	chaperone	NP_239860.1	57 912	52 (192, 80, 223, 21, 117, 51, 216, 57)
molecular chaperone DnaK	chaperone	NP_239985.1	70 045	34 (30, 20, 40, 7, 17, 15, 43, 22)
elongation factor Tu	translation	NP_240333.2	43 466	32 (32, 10, 25, 4, 29, 18, 44, 12)
co-chaperonin GroES	chaperone	NP_239859.1	10 244	11 (37, 8, 45, 2, 20, 8, 24, 10)
G-aldehyde 3-phosphate D - hydrogenase A	energy metabolism (glycolysis)	NP_240121.1	36 949	18 (6, 17, 14, 4, 3, 17, 7, 17)
OmpF-like porin	porin, diffusion of low-molecular-weight compounds	NP_240177.1	42 701	17 (12, 29, 12, 7, 1, 6, 0, 13)
alkyl hydroperoxide reductase	scavenger of endogenous hydrogen peroxide	NP_240013.1	22 363	13 (10, 11, 12, 5, 2, 8, 8, 20)
OmpA protein	porin	NP_240151.1	39 306	11 (10, 17, 8, 6, 1, 8, 3, 11)
ketol-acid reductoisomerase	amino acids biosynthesis	NP_240398.1	55 828	21 (7, 6, 8, 7, 0, 10, 5, 20)
5-methyltetrahydropteroyltriglutamate	amino acids biosynthesis	NP_239871.1	87 916	18 (23, 2, 19, 0, 5, 4, 7, 2)

<sup>a</sup>The most abundant proteins identified in our MS lists are shown for the full extract (supernatant + pellet) of *Buchnera* in the combined orange and green aphids. Numbers refer to the peptides found in the two separate determinations, and the values in parentheses correspond to the number of spectra in orange (supernatant, pellet) and green (supernatant, pellet) samples for experiment 1 and orange (supernatant, pellet) and green (supernatant, pellet) samples for experiment 2.

vitamins, and lipids that the aphid host is not capable of synthesizing by itself.<sup>1,6,7,12,13,26</sup> On the other hand, similarly, remarkable facultative symbionts have been described to play a crucial role in aphid pigmentation and resistance against biotic and abiotic stresses such as enhanced tolerance to extreme temperature and protection against pathogens and parasitoids.<sup>16,27–29</sup> Moreover, horizontal gene transfer between the host and symbiont and the transport of proteins inside the endosymbiont encoded by the host genes (or vice versa, i.e., export of proteins encoded by the endosymbiont genes) have not been found, which excludes trans activator mechanism of regulation.<sup>30,31</sup> The effects of stress on insect symbiont dynamics have been scarcely reported, although basic transcriptomic data of aphid endosymbiont *Buchnera* have shown the lack of drastic up- or down-regulation of most of its genes when the hosts were placed under different conditions.<sup>32,33</sup> Computational genome annotations have highlighted that this reduced symbiont genome does not have regulatory modules to control gene expression. Transcript levels of the majority of the *Buchnera* genes seem to react little to the temperature stress in agreement with the finding that the heat shock gene transcripts are constitutively overexpressed.<sup>15</sup> In support of this observation, 20 orthologues of *Buchnera* transcripts corresponding to the heat shock genes were found constitutively overexpressed under various circumstances including nonstress conditions.<sup>15</sup> The transcriptional dynamics of *Buchnera* are surprisingly stable, which does not eliminate the possibility that a few individual genes may be specifically induced by stimuli.<sup>5,12,15</sup> Therefore, we assumed that the full-scale analysis of the proteome of the endosymbiont using the host as an internal control will document certain reliable changes in cell density under different ecological conditions, whereas the significance of transcriptome is biased by polyploidy.

**mRNA and Protein Levels Have Poor Correlation for Many Genes in Various Species.** The number of proteins and their activity are regulated at different levels. First, transcription is governed by epigenetic states guided by chromatin changes, histone chemical modification, DNA methylation, and binding of miRNAs and partner proteins. In parallel, translation is governed by post-transcriptional modification, polyA tail stability, miRNA interference at levels of transcript, rRNAs, circularized RNAs, protein complexes that bind RNAs, and binding proteins selecting specific transcripts. Finally, the third level of regulation involves protein function that is governed by post-translation modifications and targeted location within the cell. This overall knowledge has been the subject of numerous reviews and textbooks. In symbiotic organisms, another layer of regulations resides in the synchronization of cell density of symbiont in a context where modules of gene regulators have disappeared, making those bacterial cells insensitive to external stimuli. Measurement of cell density by qPCR turned out to be questionable when cells present variable polyploidy, leading to conflicting information with proteomic data. The mRNA and protein levels often do not correlate because of the specific regulatory properties inherent to these two molecular categories.<sup>34,35</sup> Consequently, if a subset of mRNA levels strongly correlates with protein levels, numerous examples demonstrated a distorted link between these two determinations.<sup>34,35</sup> For instance, the half-life of certain proteins can vary from minutes to days due to post-translational modifications such as phosphorylation and protein association, thus avoiding the proteases, whereas RNA degradation occurs

within a tight framework usually from 2 to 5 h. Thus, low mRNA expression and high levels of protein can coexist due to the combined factors such as robustness of proteins resistant to proteases and strong mRNA negative feedback inducing rapid degradation.<sup>34,35</sup> If changes in gene expression are not frequently reflecting the variations in the protein levels, qPCR analysis of the transcript markers may fail to document indirectly the corresponding abundance/density of cells. The median Spearman correlation between RNA and protein turned out to be very modest in bacteria, yeast, and human cells.<sup>34,35</sup> The RNA/protein relationships are usually not documented and unavailable at the full-scale transcriptome/proteome. These facts urge us to explore the proteomic approach as a solid alternative tool to assess reliable cell density measures in polyploidy context.<sup>17–19</sup>

**Quantification of *Buchnera* Cell Density by the Proteomic Approach.** A few approaches have been reported to quantify the endosymbiont cell density in aphid, and, to our knowledge, only one study uses direct flow cytometry measurements.<sup>36</sup> Although this attempt provides comparative and relative levels of symbiont, the fragility of the *Buchnera* cells that present huge size compared to most of bacteria species (about 5 versus 1  $\mu\text{m}$  for *E. coli*) and lack membrane lipopolysaccharide component limits the reliability of this procedure. We found that during the extraction process, a substantial amount of *Buchnera* is lysed (this is monitored by the measurement of the released genomic material of *Buchnera* in the supernatant after centrifugation). Another variable part of *Buchnera* precipitates with Golgi and reticulum endoplasmic membranes depending on the aphid grinding and vortexing intensity (the synaptosomal membrane that prolongs the host reticulum envelops *Buchnera* cells). Finally, another study reports the measure of titer of endosymbiont based on qPCR.<sup>37</sup> This method likely might be biased by the degree of polyploidy, which is higher in old aphids.<sup>17–19</sup> The drastically reduced *Buchnera* genome lacks regulatory genes, and, as a consequence, the levels of mRNA should reflect only the number of chromosome copies.

Protein determination methods are complex and less sensitive than RNA quantification; however, the data on relative abundance of global proteins can be currently produced by multiplexed approaches. The difficulties in attributing accurately small peptides to initial proteins in proteomic analysis using mass spectrometry (MS) are now circumvented by the powerful analysis methods of bioinformatics. Quantification of proteins relies on mass spectrometer readout that includes the counts of the MS/MS spectrum that identify a given peptide. The number of identified peptides defines the procedure to assess abundance of each protein within the full proteome in the same run. We took advantage of this technology to compare relative mass/density of *Buchnera* cells. The exact numbering of *Buchnera* cells in an aphid is challenging. The multiplication of chromosomes up to 200 copies per *Buchnera* cell seems exceptional in the bacterial world;<sup>17–19</sup> remarkably, this amplification varies depending on the life cycle stage of the aphid host. The number of genome copies is low during the nymph stages and is drastically increased in adulthood.<sup>17,18</sup> As the aphid is aging, the number of genomic copies proportionally decreases, which makes this quantification a useful marker of age.<sup>17,18</sup> Moreover, in the alatae (winged animals), the number of chromosomal copies was 2-fold higher than that found in *Buchnera* cells of apterae (wingless animals).<sup>17,18</sup> Finally, 4',6'-diamidino-2-phenylindole

(DAPI) staining showed an increase in the number of aggregated cells in aging aphid, which likely suggests cell fusion process combined with a huge amount of chromosome copies.<sup>17,18</sup> All of these elements explain the difficult quantification of the endosymbiont mass and the complex relationship between mRNA and proteins levels. However, our data argue for one main regulatory mode that involves variations in cell density, relegating the discrete up- and down-regulation of a few proteins as accessory mode (see ref 38 for a review of an extensive description of these processes). Our data unambiguously show that the global amount of *Buchnera* spectra varies depending on the conditions and that changes in the endosymbiont density are not affected by the absence of regulatory genes evolutionary trimmed within the genome.

## CONCLUSIONS

The ratio of total *Buchnera* spectra versus total aphid host spectra appears to be an excellent index to measure the relative endosymbiont density. In the past, the metrics used to normalize symbiont abundance (e.g., levels of certain markers, cytometry, and qPCR) turned out to be questionable approaches to estimate reliably the number of *Buchnera* cells per aphid, thus potentially obscuring the meaning and significance of the observations. Another important problem is that large and variable number of chromosome copies per *Buchnera* cell depends on whether the cells come from embryos or nymphs or different ages of adulthood, which has been a hurdle for reliable counting of the cells with DNA fluorescent probes. These elements explain why the symbiont abundance analysis has been understudied by the authors that have mainly explored the endosymbiont mRNAs levels in diverse contexts. Thus, we suggest that the best way to accurately quantify the mass of *Buchnera* cells in aphid subjected to contrasted environments is to perform mass spectrometry analysis of the proteome on a wide-scale host/symbiont without prior extraction. Our results unambiguously show that the number of the endosymbiont cells can be adjusted in a context of environmental challenge. Up- or down-regulation of individual genes based on RNA analysis shows low variability; however, our data on the proteome document the very weak relationship between mRNA and proteins levels and that the proteomic approach constitutes the best probe to assess *Buchnera* cell density. The growth rate and relative density of the endosymbiont likely constitute a regulatory mode of the aphid adaptation bypassing the costly up- and down-regulation of numerous genes.

## EXPERIMENTAL SECTION

**Maintenance of Aphids and Processing of Samples for Proteomic Analysis.** Aphids *A. pisum* were maintained on *Vicia faba* plants in an incubation room at 20 °C with a light/dark photoperiodicity of 16/8 h, respectively. To generate the green phenotype, aphids were reared at 8 °C under the same photoperiodicity conditions.<sup>16</sup> The cold and moderate temperature lines resulted in green and orange phenotypes respectively. Three independent protein extractions were performed at 10-day intervals using 20 parthenogenetic 5- to 7-day-old adults. The extracts of the orange and green aphids were separately pooled, and the proteomic analysis was carried out in four samples: the green and orange aphid supernatants and pellets. The process was repeated twice

to duplicate the determinations leading to eight separate MS analyses. Aphids were flash-frozen in liquid nitrogen, homogenized by a mortar and a pestle, then broken by sonication, and centrifuged (10 000 rpm for 10 min) to separate soluble and insoluble proteins. The two fractions were resuspended in sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer (125 mM Tris-HCl, pH 6.8, 10% v/v mercaptoethanol, 20% v/v glycerol, 4% w/v SDS, and bromophenol blue), and the proteins were separated by gel electrophoresis.

**Proteome Analysis by NanoLC-MS/MS.** The general protocol is inspired by canonical articles that have contributed to standardize the proteomic approaches.<sup>39–41</sup> Proteins were digested with trypsin by filter-aided sample preparation (FASP) method as described in a seminal article.<sup>42</sup> An UltiMate 3000 RSLCnano System (Thermo Fisher Scientific, Waltham, Massachusetts) was used for separation of the protein digests. Peptides were automatically fractionated on a commercial C18 reversed-phase column (75  $\mu\text{m}$   $\times$  250 mm, 2- $\mu\text{m}$  particle, PepMap100 RSLC column, Thermo Fisher Scientific). Trapping was performed for 4 min at 5  $\mu\text{L}/\text{min}$  with solvent A (98% H<sub>2</sub>O, 2% acetonitrile (ACN), and 0.1% formic acid (FA)). Elution was performed using two solvents, A (0.1% FA in water) and B (0.1% FA in ACN), at a flow rate of 300 nL/min and 35 °C. Gradient separation was 3 min at 3% B, 170 min from 3 to 20% B, 10 min from 20 to 80% B, and 80% B maintained for 15 min. The column was equilibrated for 6 min with 3% solvent B prior to the next sample analysis. The eluted peptides were analyzed by Orbitrap Q-Exactive instruments (Thermo Fisher Scientific). The electrospray voltage was 1.9 kV, and the capillary temperature was 275 °C. Full MS scans were acquired in the mass analyzer over  $m/z$  300–1200 range with a resolution of 35 000 ( $m/z$  200). The target value was  $3.00 \times 10^6$ . Fifteen most intense peaks with charge states between 2 and 5 were fragmented in an HCD collision cell with a normalized collision energy of 27%; tandem mass spectrum was acquired in the mass analyzer with a resolution of 17 500 at  $m/z$  200. The target value was  $1.00 \times 10^5$ . The ion selection threshold was  $5.0 \times 10^4$  counts, and the maximum allowed ion accumulation times were 250 ms for full MS scans and 100 ms for tandem mass spectrum. Dynamic exclusion was set to 30 s.

**Processing of the MS Data and Database Searches.** Raw data collected during nanoLC-MS/MS analyses were processed and converted into the \*.mgf peak list format with Proteome Discoverer 1.4 (Thermo Fisher Scientific). MS/MS data were interpreted using Mascot search engine (version 2.4.0, Matrix Science, London, U.K.) installed on a local server. Searches were performed with a tolerance on mass measurement of 0.2 Da for the precursor and 0.2 Da for the fragment ions against a composite target decoy database (58 990 total entries) built with Aphidoidea UniProt database (TaxID 33385—May 22, 2018—58 990 entries) and a *Buchnera* database (574 entries) fused with the sequences of recombinant trypsin and a list of classical contaminants (118 entries). Cysteine carbamidomethylation, methionine oxidation, N-terminal protein acetylation, and cysteine propionamidation were searched as variable modifications. Up to one missed trypsin cleavage was allowed. For each sample, peptides were filtered out with Proline 2.0 according to the cutoff set for proteins hits with one or more peptides larger than nine residues, ion score >15, and 2% protein false-positive rates. Spectral counting analyses were performed with Proline 2.0.



The predicted *Buchnera* proteome is approximately 60-fold smaller than the aphid proteome according to bioinformatics prediction. Because the *Buchnera*/aphid datasets were combined and measured as a ratio of the total spectra, a potential risk of scoring of *Buchnera* proteins lower than expected was possible due to aphid host protein masking. However, we have focused our study on the global evaluation of the symbiont proteins as a comparative ratio to the total symbiont/host proteins. The experimental protocol was repeated twice to calculate mean values based on two separate determinations (see figure legends). A unique protein model was not obtained in some cases due to sequence similarities of MS peptides shared by several genes. This was observed with gene family in which genes are sharing identical domains. In these cases, specific peptides for a particular gene within a family were used to adjust the total number of spectra attributed to this gene. For quantification, each protein accession was scored for total spectral counts.

**GO Ontology for Ordering of the Proteins by Function and Molecular Process.** Aphid proteins were classified by Gene Ontology (GO) ([www.geneontology.org/](http://www.geneontology.org/)) using GO terms within “molecular function”. Metabolism, signaling, structural, homeostasis, transport, and information transfer were the major GO terms that were retrieved during our analysis.

**Presentation of the Analysis and Normalization.** The MS analysis was conducted separately in the supernatant and pellet components after grinding frozen samples and centrifugation of the extracts (full aphid, green, and orange phenotypes). Table S1 represents the number of peptides and spectra for each individual protein found in our analysis that matches the *Buchnera* and *A. pisum* proteomes. The extract preparation and spectra analysis have been carried out twice and values are reported corresponding to experiments 1 and 2 in Table S1. An accession number for each protein is also provided. Two procedures were used to normalize the spectral counts. First, the total spectra for the aphid orange supernatants and pellets were used to normalize their counterparts found in the aphid green phenotype. The ratio of normalization determines an index that was multiplied by each spectral number referring to the individual proteins found in the *Buchnera* within the green phenotype. The adjusted *Buchnera* green values were used for quantitative comparison studies to assess mainly the relative quantity of the endosymbiont proteins in the two phenotype contexts. An all way around normalization (using the green aphid to adjust the orange aphid) was used for the second experiment. The index values are reported in Table S1 for the two separate experiments and for the supernatant and pellet components. Second, instead of total spectra, we used the total spectra of 40S and 60S ribosomal proteins in orange aphid to normalize the green aphid values according to the same principle (the ratio of normalization leads to an index that was multiplied by individual protein spectra of *Buchnera* in the green phenotype). The *Buchnera* protein normalized values by the 40S and 60S ribosomal host proteins are reported for the supernatant and pellet and for the two separate experiments in Table S1. Therefore, the measurements of the *Buchnera* proteins amounts are based on two separate factors of normalization and two separate experiments. The full datasets of the original analysis along with the normalized values are shown in Table S1 (original lists and *Buchnera* and aphid matches). Finally, three entries (component, function, and process) of the Gene

Ontology terms are provided in Excel format for only aphid host thus allowing the reader to select short lists based on the three entry terms and to retrieve gene candidates according to a classification number of GO. We used GO:000166, GO:0003676, and GO:0003677 for supplemental comparison because these terms were overrepresented in our counts (see Table S2). Because of strong similarities between spectra values obtained in experiments 1 and 2 for aphid host, the GO term analysis is provided based on one unique experiment (Table S2).

**Global RNAseq Counting.** Total RNA was extracted according to the recommendation of the sequencing company Genewiz. The quality of the RNA was determined via Agilent bioAnalyzer using an Agilent RNA 6000 Nanokit (Agilent). The sequencing was performed by Genewiz according to the Illumina RNAseq sample preparation workflow. The successive steps included rRNA depletion, RNA fragmentation, random primer RT, 2'-deoxyuridine-5'-triphosphate (dUTP) incorporation, adenosine tailing, Y-shaped adaptor ligation, dUTP strand degradation, and, finally, polymerase chain reaction (PCR) with barcodes linked to the adapters. Sequencing was performed using Illumina TrueSeq chemistry (2 × 50 nt) on an Illumina HiSeq. 2500 platform. All steps were performed by the company Genewiz (South Plainfield, New Jersey). Sequence reads were trimmed from the generated Fastq files to remove possible adapter sequences and low-quality bases at the ends with an error rate <0.05. Sequence reads shorter than 30 nucleotides were discarded. The remaining reads were mapped to two reference genomes using the CLC genomics workbench (*B. aphidicola* str. APS (*A. pisum*) genomic DNA, GCA\_000009605.1\_ASM960v1, and *A. pisum* (Acyr\_2.0, Assembly ID: GCA\_000142985.2)). Total fragments that matched with either the host aphid or *Buchnera* endosymbiont genomes were counted for each of the triplicate samples of the green and orange phenotypes. QIAseq FastSelect 5S/16S/23S Kit was used for the depletion of ribosomal RNA in bacteria and the Illumina Ribo-Zero Plus rRNA Depletion Kit for aphid. The refined analysis based on the numbers of reads and RPKM for each gene in *Buchnera* in the two phenotypes is presented in Tables S3 and S4.

**Statistical Analysis.** For statistical analysis, the Spearman rank order correlation test was used to evaluate the strength of the monotonic relationship between the two sets of variables: lists of protein spectra of the orange [X1] and green phenotypes [X2]. The Wilcoxon signed rank test, a non-parametric statistical test, was used to assess whether the mean ranks of the population [X1] and [X2] diverge in a statistically significant manner. Finally, a parametric procedure, the paired *t*-Student test, was used to evaluate the mean difference between the [X1] and [X2] lists of values.

**Endosymbiont Spectra Quantification: Statistical Comparison between Two Host Phenotypes.** After normalization of the number of spectra between the two host phenotypes by two independent ways, the modified lists of endosymbiont spectra were submitted to the following tests:

**Normalization by Total Ribosomal Aphid Spectra.** For Experiment 1:

Spearman rank order correlation: correlation coefficient: 0.766; *P*-value = 0.0000002; number of samples: 182.

Wilcoxon signed rank test: for 132 protein, *Z*-statistic (based on positive ranks) = 1.836; *P* = 0.067.

For Experiment 2:

Spearman rank order correlation: correlation coefficient: 0.784;  $P$ -value = 0.0000002; number of samples: 182.

Wilcoxon signed rank test:  $Z$ -statistic (based on positive ranks) = 5.202,  $P < 0.001$ .

Paired  $t$ -test:  $t = -4.476$  with 181 degrees of freedom; two-tailed  $P$ -value = 0.0000134; one-tailed  $P$ -value = 0.00000671.

For the Sum Experiment 1 + 2:

Wilcoxon signed rank test:  $Z$ -statistic (based on positive ranks) = 4.246;  $P < 0.001$ .

Spearman rank order correlation: correlation coefficient: 0.817;  $P$ -value = 0.0000002; number of samples: 182.

Paired  $t$ -test:  $t = -3.920$  with 181 degrees of freedom; two-tailed  $P$ -value = 0.000125; one-tailed  $P$ -value = 0.0000626.

**Normalization by Total Aphid Spectra.** For Experiment 1:

Wilcoxon signed rank test:  $Z$ -statistic (based on positive ranks) = 1.981;  $P = 0.048$ .

Spearman rank order correlation: correlation coefficient: 0.784;  $P$ -value = 0.0000002; number of samples: 182.

Paired  $t$ -test: not significant for the full list but with the first 50 most represented  $t = -3.017$  with 49 degrees of freedom; two-tailed  $P$ -value = 0.00404; one-tailed  $P$ -value = 0.00202.

For Experiment 2:

Wilcoxon signed rank test:  $Z$ -statistic (based on positive ranks) = 4.897,  $P < 0.001$ .

Spearman rank order correlation: correlation coefficient: 0.757;  $P$ -value = 0.0000002; number of samples: 182.

Paired  $t$ -test:  $t = -4.412$  with 181 degrees of freedom, two-tailed  $P$ -value = 0.0000175; one-tailed  $P$ -value = 0.00000877.

For the Sum Experiment 1 + 2:

Spearman rank order correlation: correlation coefficient: 0.809;  $P$ -value = 0.0000002; number of samples: 182.

Wilcoxon signed rank test:  $Z$ -statistic (based on positive ranks) = 4.533,  $P < 0.001$ .

Paired  $t$ -test:  $t = -4.649$  with 181 degrees of freedom; two-tailed  $P$ -value = 0.00000643; one-tailed  $P$ -value = 0.00000321.

**Aphid Host Spectra Quantification: Statistical Comparison between the Two Phenotypes.** The lists of spectra were compared after two types of normalization: by the total aphid spectra in one or the other phenotypes or alternatively by the total ribosomal spectra in one or the other phenotype. Using the same statistical test as above, Spearman rank order correlation, Wilcoxon signed rank test, and paired  $t$ -test, no statistically significant differences were obtained when aphid host spectra lists were compared contrary to the endosymbiont spectra lists in the two corresponding host contexts.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsomega.1c01465>.

Normalized spectra counts (Table S1) (XLSX)

GO terms (Table S2) (XLSX)

*Buchnera* read count and RPKM (Table S3) (XLSX)

Spectra and peptide number for each detected protein in the MS determination regarding the aphid host and endosymbiont for the two phenotypes; GO enrichment analysis based on the molecular function of detected proteins in MS determination for aphid host; transcriptomic analysis of endosymbiont in the two phenotypes context; and detailed protocol for the RNA treatment in RNAseq determination (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

Alain Robichon – ISA, Université Côte d'Azur, INRA, CNRS, 06903 Sophia Antipolis, France; [orcid.org/0000-0003-1075-9646](https://orcid.org/0000-0003-1075-9646); Email: [alain.robichon@inra.fr](mailto:alain.robichon@inra.fr)

### Authors

Fabrice Neiers – Centre des Sciences du Goût et de l'Alimentation (CSGA), Université de Bourgogne-Franche Comté, CNRS, INRA, 21000 Dijon, France

Jean-Michel Saliou – Institut Pasteur de Lille, Univ. Lille, CNRS, Inserm, CHU Lille, F-59000 Lille, France

Loïc Briand – Centre des Sciences du Goût et de l'Alimentation (CSGA), Université de Bourgogne-Franche Comté, CNRS, INRA, 21000 Dijon, France

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsomega.1c01465>

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors thank Jean Jacques Rémy for sharing ideas and insights into the topic of RNA regulation and for helpful discussions and expertise. They thank Aviv Dombrovsky that initiated the selection of the two phenotypes (so-called green and orange aphid phenotypes) during his studies in their laboratory on the phenotypic plasticity under environmental pressure. This work was supported by an ANR grant "methylclonome" ANR-12-BSV6-006-01 to Alain Robichon. This work was also supported by the French National Research Agency (ANR) through the LABEX SIGNALIFE program (reference # ANR-11-LABX-0028-01).

## ■ REFERENCES

- (1) Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H. Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. *APS. Nature* **2000**, *407*, 81–86.
- (2) Douglas, A. E. 1998. Nutritional interactions in insect-microbial symbioses: Aphids and their symbiotic bacteria *Buchnera*. *Annu. Rev. Entomol.* **1998**, *43*, 17–37.
- (3) Dixon, A. F. G. *Biology of Aphids*. *Studies in Biology*; E. Arnold: London, 1973.
- (4) Dixon, A. F. G. *Aphid Ecology*; Chapman & Hall: London, 1998.
- (5) Shigenobu, S.; Wilson, A. C. Genomic revelations of a mutualism: the pea aphid and its obligate bacterial symbiont. *Cell. Mol. Life Sci.* **2011**, *68*, 1297–1309.
- (6) The International Aphid Genomics Consortium. Genome sequence of the pea aphid *Acyrtosiphon pisum*. *PLoS Biol.* **2010**, *8*, No. e1000313.
- (7) Charles, H.; Balmand, S.; Lamelas, A.; Cottret, L.; Perez-Brocal, V.; Burdin, B.; Latorre, A.; Febvay, G.; Colella, S.; Calevro, F.; Rahbe, Y. A Genomic Reappraisal of Symbiotic Function in the Aphid/*Buchnera* Symbiosis: Reduced Transporter Sets and Variable Membrane Organizations. *PLoS One* **2011**, No. e29096.
- (8) Dale, C.; Moran, N. A. Molecular interactions between bacterial symbionts and their hosts. *Cell* **2006**, *126*, 453–465.
- (9) Braendle, C.; Miura, T.; Bickel, R.; Shingleton, A. W.; Kambhampati, S.; Stern, D. L. Developmental origin and evolution of bacteriocytes in the aphid-*Buchnera* symbiosis. *PLoS Biol.* **2003**, *1*, No. e21.
- (10) Moran, N.; Wernegreen, J. Lifestyle evolution in symbiotic bacteria: insights from genomics. *Trends Ecol. Evol.* **2000**, *15*, 321–326.



- (11) Miura, T.; Braendle, C.; Shingleton, A.; Sisk, G.; Kambhampati, S.; Stern, D. L. A comparison of parthenogenetic and sexual embryogenesis of the pea aphid *Acyrtosiphon pisum* (Hemiptera: Aphidoidea). *J. Exp. Zool., Part B* **2003**, *295B*, 59–81.
- (12) Nakabachi, A.; Shigenobu, S.; Sakazume, N.; Shiraki, T.; Hayashizaki, Y.; Carninci, P.; Ishikawa, H.; Kudo, T.; Fukatsu, T. Transcriptome analysis of the aphid bacteriocyte, the symbiotic host cell that harbors an endocellular mutualistic bacterium, *Buchnera*. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5477–5482.
- (13) Hansen, A. K.; Moran, N. A. Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 2849–2854.
- (14) Nishikori, K.; Morioka, K.; Kubo, T.; Morioka, M. Age- and morph-dependent activation of the lysosomal system and *Buchnera* degradation in aphid endosymbiosis. *J. Insect Physiol.* **2009**, *55*, 351–357.
- (15) Wilcox, J. L.; Dunbar, H. E.; Wolfinger, R.; Moran, N. A. Consequences of reductive evolution for gene expression in an obligate endosymbiont. *Mol. Microbiol.* **2003**, *48*, 1491–1500.
- (16) Dombrovsky, A.; Arthaud, L.; Ledger, T. N.; Tares, S.; Robichon, A. Profiling the repertoire of phenotypes influenced by environmental cues that occur during asexual reproduction. *Genome Res.* **2009**, *19*, 2052–2063.
- (17) Komaki, K.; Ishikawa, H. Intracellular bacterial symbionts of aphids possess many genomic copies per bacterium. *J. Mol. Evol.* **1999**, *48*, 717–722.
- (18) Komaki, K.; Ishikawa, H. Genomic copy number of intracellular bacterial symbionts of aphids varies in response to developmental stage and morph of their host. *Insect Biochem. Mol. Biol.* **2000**, *30*, 253–258.
- (19) Tamas, I.; Klasson, L.; Canback, B.; Naslund, A. K.; Eriksson, A. S.; Wernegreen, J. J.; Sandstrom, J. P.; Moran, N. A.; Andersson, S. G. E. 50 million years of genomic stasis in endosymbiotic bacteria. *Science* **2002**, *296*, 2376–2379.
- (20) Latorre, A.; Gil, R.; Silva, F. J.; Moya, A. Chromosomal stasis versus plasmid plasticity in aphid endosymbiont *Buchnera aphidicola*. *Heredity* **2005**, *95*, 339–347.
- (21) Plague, G.; Dale, C.; Moran, N. A. Low and homogeneous copy number of plasmid-borne symbiont genes affecting host nutrition in *Buchnera aphidicola* of the aphid *Uroleucon ambrosiae*. *Mol. Ecol.* **2003**, *12*, 1095–1100.
- (22) Gil, R.; Sabater-Muñoz, B.; Latorre, A.; Silva, F. J.; Moya, A. Extreme genome reduction in *Buchnera* spp.: Toward the minimal genome needed for symbiotic life. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 4454–4458.
- (23) Tran, H. T.; Barnich, N.; Mizoguchi, E. Potential role of chitinases and chitin-binding proteins in host-microbial interactions during the development of intestinal inflammation. *Histol. Histopathol.* **2011**, *26*, 1453–1464.
- (24) Fares, M. A.; Moya, A.; Barrio, E. GroEL and the maintenance of bacterial endosymbiosis. *Trends Genet.* **2004**, *20*, 413–416.
- (25) Maezawa, K.; Shigenobu, S.; Taniguchi, H.; Kubo, T.; Aizawa, S.; Morioka, M. Hundreds of flagellar basal bodies cover the cell surface of the endosymbiotic bacterium *Buchnera aphidicola* sp. Strain APS. *J. Bacteriol.* **2006**, *188*, 6539–6543.
- (26) Thomas, G. H.; Zucker, J.; Macdonald, S. J.; Sorokin, A.; Goryanin, I.; Douglas, A. E. A fragile metabolic network adapted for cooperation in the symbiotic bacterium *Buchnera aphidicola*. *BMC Syst. Biol.* **2009**, *3*, No. 24.
- (27) Scarborough, C. L.; Ferrari, J.; Godfray, H. C. J. Aphid protected from pathogen by endosymbiont. *Science* **2005**, *310*, 1781.
- (28) Tsuchida, T.; Koga, R.; Horikawa, M.; Tsunoda, T.; Maoka, T.; Matsumoto, S.; Simon, J. C.; Fukatsu, T. Symbiotic bacterium modifies aphid body color. *Science* **2010**, *330*, 1102–1104.
- (29) Heyworth, E. R.; Ferrari, J. A. facultative endosymbiont in aphids can provide diverse ecological benefits. *J. Evol. Biol.* **2015**, *28*, 1753–1760.
- (30) Poliakov, A.; Russell, C. W.; Ponnala, L.; Hoops, H. J.; Qi Sun, Q.; Douglas, A. E.; Van Wijk, K. J. Large-scale label-free quantitative proteomics of the pea aphid-*Buchnera* symbiosis. *Mol. Cell. Proteomics* **2011**, No. M110.007039.
- (31) Nikoh, N.; McCutcheon, J. P.; Kudo, T.; Miyagishima, S. Y.; Moran, N. A.; Nakabachi, A. Bacterial genes in the aphid genome: absence of functional gene transfer from *Buchnera* to its host. *PLoS Genet.* **2010**, *6*, No. e1000827.
- (32) Humphreys, N. J.; Douglas, A. E. Partitioning of symbiotic bacteria between generations of insect: a quantitative study of a *Buchnera* sp in the pea aphid (*Acyrtosiphon pisum*) reared at different temperatures. *Appl. Environ. Microbiol.* **1997**, *63*, 3294–3296.
- (33) Enders, L. S.; Miller, N. J. Stress-induced changes in abundance differ among obligate and facultative endosymbionts of the soybean aphid. *Ecol. Evol.* **2016**, *6*, 818–829.
- (34) Peshkin, L.; Wühr, M.; Pearl, E.; Haas, W.; Freeman, R. M.; Gerhart, J. C.; Klein, A. M.; Horb, M.; Gygi, S. P.; Kirschner, M. W. On the relationship of protein and mRNA dynamics in vertebrate embryonic development. *Dev. Cell* **2015**, *35*, 383–394.
- (35) Maier, T.; Güell, M.; Serrano, L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett.* **2009**, *583*, 3966–3973.
- (36) Simonet, P.; Dupont, G.; Gaget, K.; Weiss-Gayet, M.; Colella, S.; Febvay, G.; Charles, H.; Viñuelas, J.; Heddi, A.; Calevro, F. Direct flow cytometry measurements reveal a fine-tuning of symbiotic cell dynamics according to the host developmental needs in aphid symbiosis. *Sci. Rep.* **2016**, No. 19967.
- (37) Chong, R. A.; Moran, N. A. Intraspecific genetic variation in hosts affects regulation of obligate heritable symbionts. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 13114–13119.
- (38) Toft, C.; Andersson, S. G. Evolutionary microbial genomics: insights into bacterial host adaptation. *Nat. Rev. Genet.* **2010**, *11*, 465–475.
- (39) Scigelova, M.; Makarov, A. Orbitrap mass analyzer – overview and applications in proteomics. *Proteomics* **2006**, *6*, 16–21.
- (40) Old, W. M.; Meyer-Arendt, K.; Aveline-Wolf, L.; Pierce, K. G.; Mendoza, A.; Sevinsky, J. R.; Resing, K. A.; Ahn, N. G. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol. Cell. Proteomics* **2005**, *4*, 1487–1502.
- (41) Chen, C.; Hou, J.; John, J.; Tanner, J. J.; Cheng, J. Bioinformatics Methods for Mass Spectrometry-Based Proteomics Data Analysis. *Int. J. Mol. Sci.* **2020**, *21*, No. 2873.
- (42) Wiśniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **2009**, *6*, 359–362.