



**HAL**  
open science

## Molecular and cellular profiles of insect bacteriocytes: mutualism and harm at the initial evolutionary step of symbiogenesis

A. Heddi, A. Vallier, C. Anselme, H. Xin, Yvan Rahbé, F. Wäckers

► **To cite this version:**

A. Heddi, A. Vallier, C. Anselme, H. Xin, Yvan Rahbé, et al.. Molecular and cellular profiles of insect bacteriocytes: mutualism and harm at the initial evolutionary step of symbiogenesis. Cellular Microbiology, 2005, 7 (2), pp.293-305. 10.1111/j.1462-5822.2004.00461.x . hal-00391340

**HAL Id: hal-00391340**

**<https://hal.science/hal-00391340v1>**

Submitted on 9 May 2019

**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.

# Molecular and cellular profiles of insect bacteriocytes: mutualism and harm at the initial evolutionary step of symbiogenesis

Abdelaziz Heddi,<sup>1\*</sup> Agnès Vallier,<sup>1</sup> Caroline Anselme,<sup>1</sup> Huawei Xin,<sup>1†</sup> Yvan Rahbe<sup>1</sup> and Felix Wäckers<sup>2</sup>

<sup>1</sup>Laboratoire de Biologie Fonctionnelle Insectes et Interactions, UMR INRA/INSA de Lyon, Bât. Louis Pasteur, 20 Avenue Albert Einstein, 69621 Villeurbanne Cedex, France.

<sup>2</sup>Netherlands Institute of Ecology (NIOO KNAW) Boterhoeksestraat 48, 6666 GA Heteren, the Netherlands.

## Summary

Intracellular symbiosis is considered to be a driving force in eukaryotic cell evolution. In insects, little is known about the molecular bases of the bacteria-bearing host cells (bacteriocytes), particularly in the initial steps of symbiosis, where the bacterial genome has not experienced severe gene deletions because of evolutionary constraints associated with intracellular and vertical transmission. Here, we have applied polymerase chain reaction (PCR)-subtracted cDNA and reverse Northern analysis on the bacteriocytes of a recently established endosymbiosis, the weevil *Sitophilus zeamais*, to discover genes of potential relevance to bacteriocyte genetics. We provide a broad characterization of bacteriocyte transcriptional responses to intracellular bacteria, including pathways covering metabolism-transport-stress (MTS), cell signalling and trafficking, growth and apoptosis, as well as innate immunity. MTS genes show an intriguing diabetes-like pathogenic profile associated with increased stress, as indicated by high levels of upregulations of carbohydrate transporters, aldose reductases and stress-related genes. A high-performance liquid chromatography (HPLC) analysis of tissue carbohydrate contents highlighted an increased carbohydrate assimilation in symbiotic insects and the prevalence of a polyol biosynthetic pathway, as

indicated by the accumulation of sorbitol, mannitol and fructose in the bacteriocytes. These findings provide the first genetic perspectives on the nature of the interaction between insect and cooperative bacteria. They unravel the profound insect bacteriocyte stress associated with increased metabolism and cell trafficking, and they shed light on the potential role of the innate immunity during the pathogeny-mutualism transition at the initial stage of insect symbiogenesis.

## Introduction

Symbiogenesis was defined first by Merezhkowsky (1920) as 'the origin of organisms through combination or association of two or more beings, entering into symbiosis'. This evolutionary process is common in nature and is assumed to play an important role in biodiversity and life evolution (Maynard Smith, 1989; Margulis, 1993; Bordenstein *et al.*, 2001; Heddi, 2003; Lefèvre *et al.*, 2004). In insects, intracellular bacterial symbiosis (endosymbiosis) is often required for reproduction and development of the host (Heddi *et al.*, 1999a). Endosymbionts supplement their host's diet with nutrients of limited availability, thereby improving their invasive power and their ability to settle on new and deficient diet sources and habitats (Douglas, 1998; Heddi *et al.*, 1999a; Wernegreen, 2002).

Endosymbiotic bacteria are an integral part of insect host cell biology and development. They are transmitted maternally through host lineages and, during insect embryogenesis, invade specialized host cells, called the bacteriocytes. Sometimes these form a specific organ, the bacteriome (Heddi *et al.*, 1999a; Braendle *et al.*, 2003). Unlike plant-root symbiosis, where molecular processes and coordinated events triggering the development of nodules are well documented (Schauser *et al.*, 1999; Stracke *et al.*, 2002), the mechanisms leading to insect bacteriocyte cell differentiation and homeostasis remain unknown. The best-studied model in the invertebrate field is the marine squid–*Vibrio* extracellular symbiosis, in which the development of the squid light organ requires specific gene signalling between partners (Foster *et al.*, 2000; Nyholm *et al.*, 2000). Several molecular and cellular changes, such as apoptosis, cell swelling and oxidative stress, occur when *Vibrio fischeri* is recruited from the environment (Foster *et al.*, 2000; Visick *et al.*, 2000).

Received 25 June, 2004; revised 2 September, 2004; accepted 2 September, 2004. \*For correspondence. E-mail Abdelaziz.Heddi@insa-lyon.fr; Tel. (+33) 4 72 43 88 68; Fax (+33) 4 72 43 85 34. †Present address: Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA.

In both root and squid symbioses, each generation of the host recruits specific free-living bacteria from the environment. Insect intracellular symbioses differ from this in two aspects: (i) insect symbionts are strictly intracellular, so that the environmental re-infection process is *a priori* absent, and (ii) insect intracellular endosymbionts are transmitted strictly vertically, which prevents endosymbiont recombination with free-living bacteria. This restricted way of symbiont transmission generates a relatively high rate of endosymbiont DNA evolution, which seems to result in the accumulation of deleterious mutations and a severe genome size reduction (Moran, 1996; Itoh *et al.*, 2002; van Ham *et al.*, 2003). The bacterial DNA degeneration process removes pathways that are redundant with those of the host (Shigenobu *et al.*, 2000), DNA repair mechanisms (Dale *et al.*, 2003), gene regulators and pathogenic genes (Wernegreen, 2002; Dale *et al.*, 2003; Wilcox *et al.*, 2003). Therefore, the complexity of molecular interactions between partners may evolve according to the level of bacterial genomic degeneration and, hence, according to the age of the association. In this evolutionary context, the study of recently established symbiotic associations, particularly those in relation to the insect defence system and bacterial virulence, may provide an insight into the establishment of insect intracellular symbiosis and its maintenance.

Recently, we have identified three  $\gamma$ 3-Proteobacterial endosymbiotic clades within the insect family Dryophthoridae (superfamily Curculionioidea; Lefèvre *et al.*, 2004). Among them the more recent S clade comprises *Sitophilus* spp. primary endosymbionts (SPEs), which are phylogenetically close to pathogenic bacteria (*Proteus*, *Yersinia*, *Salmonella*) and which encode the type three secretion system (TTSS) (Dale *et al.*, 2002). To investigate host molecular responses to their potentially pathogenic endosymbionts, and to gain an insight into bacteriocyte genetics, we have screened genes upregulated in the bacteriocyte cells of the weevil *Sitophilus zeamais* using RT-PCR subtracted cDNA and we confirmed their differential expression by reverse Northern analysis. We have identified a large number of bacteriocyte-upregulated genes that contribute to the understanding of insect bacteriocyte genetic bases. Furthermore, we show that symbiotic cellular homeostasis evolves from a pathologic-like profile, underlined by the innate immunity induction and cellular stress accumulation.

## Results and discussion

Whereas insect endosymbiont physiological and evolutionary features have been well elucidated over the past decade through genome sequencing and comparative genomics, molecular mechanisms and the partners' dialogue that trigger the establishment and the homeostasis

of the bacteriocyte tissue remain unknown. Here, we have provided a broad panel of bacteriocyte-activated functions in response to the bacteria, comprising metabolism-transport-stress (MTS), cell signalling and trafficking, growth and apoptosis, and innate immunity. Expressed sequence tags (EST) annotation and analysis (relative expression, differential ratios) are presented in Table 1. Only the most relevant functions are discussed here.

### *Increased bacteriocyte metabolism*

Functional EST gene grouping has provided molecular insights into how intracellular bacteria improve nutrient absorption and what are the prevalent pathways into the bacteriocytes. Around 30% of the subtracted ESTs were shown to encode genes that are related to MTS (Table 1). MTS gene analysis has roughly indicated two levels of bacteriocyte-enhanced activity: (i), fundamental cell processes (ribosomal proteins, translation and transcription factors) and protein turn-over (proteasome subunits, ubiquitin enzymes, chymotrypsin B), and (ii), a wide range of enzymatic pathways, including NAD<sup>+</sup> and porphobilinogen cofactors synthesis (glutamine-dependent DNA<sup>+</sup> synthase, porphobilinogen synthase), lipid metabolism [triacylglycerol lipase, fatty-acid binding protein (*FABP*)], energetic metabolism (malate dehydrogenase, cytochrome *c* oxidase), and carbohydrate metabolism (Table 1).

*Carbohydrate transport and assimilation.* Carbohydrate metabolism and import/export through the bacteriocyte membranes were highlighted by the induction of 29 MTS genes including sugar transporters, amylases, glucosidases and aldose reductases (Table 1). The identified sugar transporters belong to the major facilitator superfamily (MFS) that are capable of transporting small solutes into the cells in response to chemio-osmotic gradients. Knowing the intimate attachment of the bacteriome tissue to the larval foregut and the high abundance of carbohydrates in the wheat grains (70% of starch), in which the weevil larvae develop, the upregulation of MFS and MTS genes are viewed as driving an increased sugar flow and absorption in symbiotic insects.

To gain insight into this aspect, we have analysed the global insect carbohydrate assimilation by comparing the content of mono-, di- and trisaccharides between aposymbiotic and symbiotic insects (Table 2). Data analysis showed that the overall contents of plant-derived di- and trisaccharides are lower in symbiotic than in aposymbiotic insects, while sucrose was not detected in whole symbiotic insects. Trehalose, the common insect haemolymph sugar, on the other hand, was higher in symbiotic insects. These findings indicate that the assimilation of plant seed oligosaccharides, such as raffinose, melezitose and sucrose, is more efficient in symbiotic insects and that sugar assim-

**Table 1.** Annotation and expression data of the clones sequenced from the bacteriome-subtracted cDNA.

GenBank Acc. No.	Protein description	CN	DR	RE	Target organism	E-value (hit)	SP-AC
<i>Biosynthesis and catabolism</i>							
CN612459	Ribosomal protein 60S, L13 subunit	1	2.64	0.54	<i>Spodoptera</i>	3.7e-36	Q962U1
CN612431	Ribosomal protein S18	1	3.39	0.93	<i>Spodoptera</i>	1.2e-30	Q962R1
CN612426	Transcriptional coactivator kohtalo (RNA polymerase II transcription mediator)	1	nd	nd	<i>Homo sapiens</i>	4.9e-18	Q96PC8
CN612415	Eukaryotic translation initiation factor 3, subunit 4	1	<b>5.79</b>	0.28	<i>Drosophila</i>	3.2e-26	Q9VDM6
CN612444	Proteasome (26S) associated pad1 homologue, RPN11 protein	1	4.99	0.15	<i>Drosophila</i>	7.1e-52	Q9V3H2
CN612419	Proteasome (26S) subunit Mov34-like protein	1	20.70	0.44	<i>Drosophila</i>	3.4e-83	Q9W178
CN612414	Proteasome Beta subunit C7-I	1	9.71	0.26	<i>Homo sapiens</i>	7.5e-39	P49721
CN612409	Ubiquitin carboxyl-terminal hydrolase 16 (Deubiquitinating enzyme 16)	1	<b>5.59</b>	0.27	<i>Drosophila</i>	4.8e-34	Q9W4C3
CN612470	Ubiquitin-protein ligase (Ubiquitin-conjugating enzyme E2)	1	<b>146.40</b>	7.44	<i>Drosophila</i>	1.2e-59	O96840
CN612453	Chymotrypsin B	1	<b>11.04</b>	3.18	<i>Penaeus</i>	1.8e-18	Q00871
CN612433	Glutamine-dependent NAD <sup>+</sup> synthetase	1	nd	nd	<i>Drosophila</i>	2.5e-36	Q9VYA0
CN612473	Homologue to human proline synthetase co-transcribed bacterial homologue protein	1	24.48	1.94	<i>Homo sapiens</i>	1.1e-41	O94903
CN612411	Triacylglycerol lipase	1	<b>3.75</b>	0.12	<i>Bos taurus</i>	9.3e-34	Q29458
CN612449	Myo-inositol-1-phosphate synthase	1	<b>7.26</b>	0.40	<i>Drosophila melanogaster</i>	9.7e-73	Q8MR42
CN612405	Uridine 5'-monophosphate synthase	1	<b>10.05</b>	0.61	<i>Drosophila</i>	1.0e-22	Q01637
CN612450	Malate dehydrogenase	1	7.86	0.81	<i>Drosophila</i>	1.7e-75	Q9U1J0
CN612412	Cytochrome c oxidase II	1	1.06	94	<i>Sitophilus</i>	1.7e-136	Q9B2B8
CN612418	Alpha-amylase 1	2	2.19	2.13	<i>Drosophila</i>	3.3e-37	O15953
CN612420	Alpha-amylase 2	2	1.77	1.53	<i>Tribolium</i>	1.1e-54	Q26854
CN612432	Alpha-amylase 3	1	1.97	25.03	<i>Tenebrio</i>	2.8e-40	P56634
CN612475	Alpha-glucosidase (Maltase I)	1	nd	nd	<i>Drosophila</i>	1.2e-45	O16098
CN612474	Beta-glucosidase	1	<b>25.51</b>	0.11	<i>Tenebrio</i>	2.4e-36	Q9GSE6
CN612446	Aldehyde dehydrogenase (Retinaldehyde-specific dehydrogenase type 2) (RALDH-2)	1	9.70	6.38	<i>Homo sapiens</i>	4.1e-60	O94788
CN612469	Aldose reductase 1	5	13.99 <sup>a</sup>	19.31	<i>Drosophila</i>	5.1e-74	Q9VTK9
CN612416	Aldose reductase 2	1	<b>8.78</b>	0.52	<i>Drosophila</i>	1.2e-29	Q9Y112
CN612417	Aldose reductase 3	1	<b>7.98</b>	0.46	<i>Drosophila</i>	1.2e-22	Q9VTY2
CN612455	Delta-aminolevulinic acid dehydratase (Porphobilinogensynthase)	1	<b>2.83</b>	0.26	<i>Drosophila</i>	1.0e-40	Q9VTV9
<i>Transport</i>							
CN612440	Sugar transporter	1	<b>9.72</b>	0.66	<i>Drosophila</i>	4.1e-28	Q9VNJ2
CN612425	Possible Sugar transporter	1	<b>8.37</b>	0.49	<i>Drosophila</i>	4.6e-19	Q9V7F4
CN612442	Solute carrier family 17 (Anion/sugar transporter, sodium/phosphate symporter)	1	<b>7.01</b>	0.38	<i>Drosophila</i>	9.1e-43	Q9VR44
CN612406	Possible Tricarboxylate/iron carrier (Sideroflexin 1)	1	<b>2.81</b>	0.29	<i>Homo sapiens</i>	2.7e-15	Q9H9B4
CN612429	Hypothetical Fatty acid binding protein	2	<b>155.82<sup>a</sup></b>	79.13	<i>Caenorhabditis</i>	0.00022	Q965W1
CN612439	Hypothetical Fatty acid binding protein	1	<b>189.66</b>	100	<i>Manduca</i>	5.1e-06	P31417
CN612457	Phosphatidylethanolamine-binding protein	1	<b>4.96</b>	0.37	<i>Drosophila</i>	2.8e-27	Q9VK60
CN612476	Cellular retinaldehyde-binding protein	1	nd	nd	<i>Aedes</i>	2.9e-68	Q9BJM8
CN612407	Aspartyl-tRNA synthase	1	14.09	0.94	<i>Drosophila</i>	1.1e-73	Q9XYM1
<i>Stress</i>							
CN612422	Glutathione S-transferase 1	2	<b>26.14</b>	10.15	<i>Musca</i>	2.8e-12	Q9U795
CN612452	Glutathione S-transferase 2	1	<b>13.72</b>	9.38	<i>Drosophila</i>	3.3e-28	Q9V8H2
CN612466	Glutathione S-transferase 3	3	<b>16.88</b>	23.79	<i>Anopheles</i>	3.1e-24	Q9GPL8
CN612461	Lactoylglutathione lyase (Glyoxalase I)	1	<b>13.55</b>	10.46	<i>Drosophila</i>	1.0e-48	Q9V4K3

Table 1. cont.

GenBank Acc. No.	Protein description	CN	DR	RE	Target organism	E-value (hit)	SP-AC
CN612427	Phospholipid hydroperoxide glutathione peroxidase	1	nd	nd	<i>Bos taurus</i>	6.9e-43	Q9N2J2
CN612435	Thioredoxin peroxidase	1	12.58	12.71	<i>Aedes</i>	2.5e-46	Q8WSF6
CN612468	Cytochrome P450	1	<b>12.11</b>	0.78	<i>Drosophila</i>	9.3e-28	Q9VW43
<i>Cell structure and signalling</i>							
CN612428	Alpha-catenin	1	4.68	0.31	<i>Drosophila</i>	8.6e-28	P35220
CN612403	Hypothetical fasciclin	2	<b>6.16<sup>a</sup></b>	0.31	<i>Drosophila</i>	6.3e-24	Q9VJA3
CN612472	Hypothetical Dm-CG7231 like protein bearing actinin-type actin-binding domain	1			<i>Drosophila</i>	5.7e-26	Q9VLU9
CN612447	Tubulin alpha-1 chain	1	5.19	12.71	<i>Drosophila</i>	1.9e-73	P06603
CN612430	EF-hand calcium-binding, calpain	1	<b>6.63</b>	0.35	<i>Drosophila</i>	2.5e-50	Q9VT65
CN612408	EF-Hand calcium-binding protein Dm-CG17661 like, homologue to human-programmed cell death protein 6 (ALG-2)	1	<b>6.51<sup>a</sup></b>	0.34	<i>Homo sapiens</i>	6.2e-17	O75340
CN612464	EF-Hand calcium-binding protein Dm-CG17661 like, homologue to human-programmed cell death protein 6 (ALG-2)	1	<b>11.37</b>	0.432	<i>Drosophila</i>	4.3e-52	Q9W5N8
CN612413	Translationally controlled tumour protein homologue (TCTP)	1	2.27	0.43	<i>Drosophila</i>	1.8e-64	Q9VGS2
CN612441	Ras small GTPase, Rab type	1	<b>7.17<sup>a</sup></b>	0.39	<i>Drosophila</i>	2.1e-84	Q9VC31
CN612445	Vacuolar ATPase subunit	1	2.06	2.10	<i>Manduca</i>	1.1e-06	O76959
CN612392	Hypothetical Dm-VTU2 like protein, vitelline membrane/structural eggshell protein with AP repeats	7	7.41	18.32	<i>Drosophila</i>	1.5e-09	P13238
CN612424	Tetratricopeptide repeats, Homologue to mousse	1	6.21	1.73	<i>Drosophila</i>	9.8e-17	O46085
CN612423	Peroxisome receptor 1 Peptidoglycan-recognition protein	6	<b>28.32<sup>a</sup></b>	9.81	<i>Drosophila</i>	1.6e-59	Q9VGN3
<i>Unknown function</i>							
CN612396	Hypothetical Dm-CG9269 like protein, small conserved glycine-rich eggshell related protein	2	<b>10.93</b>	0.69	<i>Drosophila</i>	1.2e-05	Q9VFT9
CN612448	Hypothetical Dm-CG30336 like protein, with tropomyosin motif	1	<b>7.98</b>	0.45	<i>Drosophila</i>	6.9e-13	AAM50100
CN612378	Hypothetical Dm-CG6005 like protein, homologue to human G-protein (WD-40 repeats)	1	<b>1.76</b>	0.89	<i>Drosophila</i>	2.1e-50	AAM49908
CN612443	Hypothetical Dm-CG5842 like protein, ankyrin signature cation-channel, crystalline-like protein	1	<b>29.87</b>	10.19	<i>Drosophila</i>	5.0e-29	Q9VUD5
CN612376	Hypothetical Dm-BcDNALD22567 like protein	2	<b>6.05</b>	1.41	<i>Drosophila</i>	2.9e-16	Q961C9
CN612379	Hypothetical Dm-CG10825 like protein, Src homology 2 (SH2) domain	1	<b>10.80</b>	2.93	<i>Mus musculus</i>	1.7e-12	Q8VEI2
CN612387	Hypothetical(DC024) Dm-CG11847 like protein, highly homologous to Hs-Q9NWX1 protein	2	<b>70.22</b>	11.88	<i>Homo sapiens</i>	1.0e-52	Q9NWX1
CN612388	Hypothetical Dm-CG4230 like protein	1	11.39	9.45	<i>Mus musculus</i>	6.2e-07	Q99KX1
CN612380	Hypothetical Dm-CG7519 like protein	1	<b>3.90</b>	3.26	<i>Mus musculus</i>	1.2e-15	Q9D8M8
CN612374	Hypothetical Dm-CG8679 like protein, with ankyrin repeat and Lem domain	1	<b>3.07</b>	0.49	<i>Homo sapiens</i>	5.0e-14	BAC04840
CN612372	Hypothetical protein	1	<b>3.65</b>	4.31		>e-02	
CN612373	Hypothetical protein	1	<b>70.32</b>	49.20	<i>Manduca</i>	0.0088	P31417
CN612375	Hypothetical protein	1	<b>3.29</b>	0.86	<i>Dictyostelium</i>	0.00027	Q8T254



Table 1. cont.

GenBank Acc. No.	Protein description	CN	DR	RE	Target organism	E-value (hit)	SP-AC
CN612377	Hypothetical protein	1	<b>7.78</b>	1.34	<i>Acanthamoeba</i>	0.0024	P46761
CN612381	Hypothetical protein	1	<b>22.82</b>	2.70		>e-02	
CN612382	Hypothetical protein	1	<b>8.60</b>	0.82	<i>Mycoplasma</i>	0.00033	Q50274
CN612383	Hypothetical protein	1	<b>29.20</b>	14.25	<i>Pyrococcus</i>	0.00041	O57996
CN612386	Hypothetical protein	1	<b>122.72</b>	16.04	<i>Arabidopsis</i>	0.00029	Q9FKN6
CN612391	Hypothetical protein KID repeats	1	<b>7.76</b>	0.34	<i>Borrelia</i>	5.4e-06	Q9RG81
CN612393	Hypothetical protein	1	<b>5.90</b>	1.64	<i>Homo sapiens</i>	0.0033	Q9H9L7
CN612395	Hypothetical protein	1	<b>23.54</b>	11.95	<i>Arabidopsis</i>	0.00016	O23092
CN612399	Hypothetical protein	1	<b>26.22</b>	7.26	<i>Anaplasma</i>	3.2e-05	Q8RPL3
CN612401	Hypothetical protein	1	<b>80.62</b>	22.08	<i>Mus musculus</i>	0.00087	Q9D3Q7
<i>Non-eukaryotic genes</i>							
CN612481	DNA-binding protein Ner	1	nd	nd	<i>Bacteriophage Mu</i>	3.9e-08	P06020
CN612479	Pol-like protein, RNA-directed DNA polymerase (Reverse transcriptase)	1	nd	nd	<i>Baculovirus</i>	1.1e-15	Q65353
CN612482	Rep-like Pi-like protein	4	nd	nd	<i>Escherichia coli</i>	1.3e-17	Q47523
CN612484	Replication associated protein (pZM2)	1	nd	nd	<i>Zymomonas</i>	4.1e-21	P15256
	<i>Without or with small ORFs</i>	42	nd	nd			

**a.** gene differential tested by real-time RT-PCR.

EST sequences were analysed as described in *Experimental procedures*. CN, number of clones found in the subtracted cDNA; DR, differential expression ratio (bacteriome/aposymbiotic larvae) obtained from reverse Northern analysis; RE, relative expression values (percentage of EST expression in the bacteriome as compared with the most expressed clone, which is the FABP (CN612439); SP-AC, SWISSPROT accession number of hits; nd, non-determined. DR and RE values were calculated from at least two independent reverse Northern analyses and values were shown to be very reliable. Values in bold are ESTs with no detectable band with the aposymbiotic larval probe; their DRs are therefore underestimated.

cDNA subtraction yielded 284 clones. Reverse Northern analysis was sensitive to 162 ESTs, including 153 (~95%) bacteriocyte upregulated ESTs with DR ranging from 1.8 for to 190 and only nine ESTs were shown to be non-differential. Among the 153 differential ESTs, 63 failed to show any signal on blots hybridized with the aposymbiotic cDNA probe regardless their RE. Their DR values are reported as a rough indication because they are underestimated.

Arbitrary classification based on gene function separated them as follows: 33 ESTs (22%) encode metabolic functions, 10 ESTs (7%) encode cellular transport genes, 10 ESTs exhibit stress-related functions, 28 ESTs (18%) are cell structure and signalling genes and 26 ESTs (17%) are of unknown function.

Most of cellular transport genes (8/10), stress-related genes (8/10) and hypothetical genes (25/30) are detectable only in the bacteriome.

ilation may enhance trehalose synthesis and accumulation, leading to significantly increased pools of easily accessible energy stores. It remains to be clarified, though, whether the bacteriome is directly involved in such sugar assimilation or whether SPE modulates the expression of genes involved in intestinal functions, as was recently demonstrated with mammalian resident microorganisms (Hooper *et al.*, 2001). The former case would imply translocation of di- and trisaccharides into the bacteriome, as may be hypothesized by the intriguing presence of sucrose in the bacteriocyte tissue (Table 2). Although previous reports on different insect species have discussed similar observations (Turunen, 1985), neither sucrose synthesis nor sucrose transport have been unambiguously demonstrated in insect tissues. However, by using a heterologous array hybridization to determine the genomic compositions of SPE and *Sodalis* (the closest relative to SPE, that has also recently established a secondary symbiotic association with the tsetse fly), we have demonstrated that among the greatest differences between these bacteria is that SPE has retained genes encoding for plant-derived-car-

bohydrates catabolism while *Sodalis* has lost them (Rio *et al.*, 2003). This finding was interpreted as an adaptive response to the insect nutritional niche as tsetse flies (the host of *Sodalis*) do not feed on plant material but on blood, which is low in complex carbohydrates and rich in simple sugars such as glucose.

*Induction of the polyol pathway in the bacteriocyte.* The increased sugar assimilation and the upregulation of MST genes point to high carbohydrate uptake and metabolism by the bacteriocytes. High glucose levels are known to generate in mammalian diabetic tissues hyperosmotic stress, tissue injury and cellular toxicity, mainly by sorbitol and fructose accumulation through aldose reductase (AR) expression in the polyol pathway. Here, three AR genes (represented by seven clones) were shown to be highly expressed in the bacteriocytes [differential ratio (DR) > 8, Table 1]. To confirm the induction of the polyol pathway, we have compared polyol contents between aposymbiotic insects and the bacteriome tissue. Results show that both sorbitol (×8) and fructose (×5) are highly accumulated in

**Table 2.** Sugar and polyol content (mean  $\pm$  SE) of symbiotic and aposymbiotic insect/tissues for the weevil *S. zeamais*.

Sample/tissue	Sorbitol	Mannitol	Trehalose	Glucose	Fructose	Sucrose	Melezitose	Raffinose	Maltose	Erflose
Sugar structure <sup>a</sup>	hexitol	hexitol	glc (1,1) glc	hexose	hexose	glc (1,2) fru	glc (1,2) fru	gal (1,6) glc (3,1)	glc (1,4) glc glc (1,2) fru	glc (1,4) glc glc (1,2) fru
<i>Whole insect</i>										
( $\mu$ g fresh weight)										
Symbiotic insect	69 $\pm$ 10	3 $\pm$ 0.4	8700 $\pm$ 600	6060 $\pm$ 430	57 $\pm$ 1	nd	13 $\pm$ 2	371 $\pm$ 9	159 $\pm$ 5	nd
Aposymbiotic insect	62 $\pm$ 11	3 $\pm$ 1	6800 $\pm$ 240	4610 $\pm$ 850	69 $\pm$ 13	38 $\pm$ 13	68 $\pm$ 9	410 $\pm$ 19	242 $\pm$ 35	nd
<i>P</i> (t-test sym/apo)	ns	ns	0.05	ns	ns	0.007	0.037	ns	ns	–
% in digestive tract <sup>b</sup>	5%	–	2%	4%	7%	40% <sup>a</sup>	–	2%	3%	–
<i>Tissue content</i>										
(nmoles/mg protein)										
Bacteriome	16.3 $\pm$ 1.6	2.0 $\pm$ 0.79	265 $\pm$ 21	36.2 $\pm$ 5.1	12.3 $\pm$ 1.4	5.4 $\pm$ 2.0	nd	nd	2.93 $\pm$ 0.48	2.0 $\pm$ 0.08
Aposymbiotic tissues	2.4 $\pm$ 0.24	0.12 $\pm$ 0.02	138 $\pm$ 4.9	178 $\pm$ 33	2.67 $\pm$ 0.5	0.77 $\pm$ 0.15	0.94 $\pm$ 0.12	5.71 $\pm$ 0.26	4.91 $\pm$ 0.71	nd
<i>P</i> (t-test bact./apo)	0.0007	ns	0.004	0.004	0.002	ns	–	–	ns	–

a. Overview of polyside structures after their hexose components: glc (glucose), fru (fructose) and gal (galatose).

b. A separate estimation of sugars in digestive tract (versus whole body) was performed; we found no significant difference in tissue distribution in symbiotic and aposymbiotic insects, except for sucrose; reported values are the fraction of each sugar measured in the digestive tract (aposymbiotic insects; '–' is below detection threshold in digestive tracts). Biochemical characterization of sugars and polyols in relevant tissues of symbiotic and aposymbiotic weevils showed: (i) polyols, and mainly sorbitol, were found in internal tissues of both weevil strains, but accumulated 8–12 times the aposymbiotic levels within the bacteriocyte tissue, thereby validating the effectiveness of the biosynthetic pathway revealed by aldose reductases overexpressed in the bacteriocytes (Table 1); (ii) the general sugar load of the symbiotic insects was significantly higher than their aposymbiotic controls, emphasizing the strength of energetic metabolism in symbiotic insects (Heddi *et al.*, 1999a); also, non-reducing sugar (e.g. trehalose versus glucose) accumulation occurred specifically in the bacteriocyte tissue, exemplifying the need for catabolizing enzymes in this tissue (maltase and amylase family enzymes; Table 1); (iii) sugars originating exclusively from plants – the sucrose series glc (1,2) fru adducts (Table 2) were readily detected within the insect tissues, including through specific accumulation of sucrose – not raffinose – in the bacteriocytes; this strongly argues for a di-tri saccharide transport system in the symbiotic tissue, and fits the specific expression of several sugar transporter genes (Table 1), although none of them could be ascribed to a putative sucrose transporter, as the *Drosophila* gene (FlyBase: CG4484, SWISSPROT: Q8VSV) annotated as such by orthology to the plant SUT family of phloem-expressed sucrose transporters. Such sucrose transport has already been speculated in insect cells (Turunen, 1985); (iv) the hexose balance in the 'sink' symbiotic tissue is likely to be drained by the trophic flux towards the harboured bacteria, accompanied by a diversion towards storage forms with specific, yet unknown, functions; the glu/fru ratio is particularly low in the bacteriocytes, while polyols, trehalose, oligosaccharides from the maltose series (erflose) and glycogen all accumulate in the bacteriocytes. nd, below detection threshold; ns, no significant difference for t-test between tested modalities ( $P > 0.05$ ).

the bacteriome (Table 2). However, bacteriocyte examination did not show any obvious cell abnormality or cell swelling (data not shown), despite the induction of the polyol pathway. The insect bacteriocyte system may have selected an appropriate regulation of high sugar uptake to avoid mammalian-like diseases caused by sorbitol and fructose accumulation. This regulation may involve the endosymbiont metabolism, as SPE was shown to encode polyol transporters such as the mannitol carrier (Rio *et al.*, 2003). Intriguingly, the genome of the aphid endosymbiont *Buchnera*, which was drastically reduced, also encodes a polyol transporter among the rare transporters retained by this bacterium (less than 20 transporters, Shigenobu *et al.*, 2000). This is a strong indication of a general function of polyols and cellular stress generation in insect intracellular symbioses. Whether polyol synthesis and accumulation also exist in the aphid bacteriocytes would provide interesting parallels between endosymbioses that are unrelated in evolutionary terms, but have common trophic functions characterized by the processing of plant-derived carbohydrate-rich diet.

*Induction of cellular stress in the bacteriocyte.* The most critical consequence of intense glucose uptake and metabolism, and polyol pathway activation, is the production of reactive oxygen species (ROS) such as H<sub>2</sub>O<sub>2</sub> and HO. (Wolff *et al.*, 1989). Excessive ROS production impairs cellular physiology by altering DNA, membrane structure, lipoproteins and membrane receptors. In weevils, the bacteriocytes seem to express several antioxidant defence system against ROS and other stress molecules, such as advanced glycation end-products (AGE). These include glutathione peroxidases and glutathione-S-transferases that catalyse the reduction of hydrogen peroxide and many lipidic hydroperoxides, and, the glyoxalase I, that reduces, in combination with reduced glutathione, the AGE (such as methylglyoxal) produced by the accumulation of fructose (Suarez *et al.*, 1988). Furthermore, an additional antioxidant enzyme was also screened from the bacteriome-subtracted cDNA, which is a thioredoxin peroxidase that uses thioredoxin as an immediate hydrogen donor and constitutes an enzymatic defence against sulphur-containing radicals.

The induction of variety of antistress systems points to an adaptation of insect bacteriocytes to high metabolic levels and to the ability of these cells to reduce the amount of ROS and AGE and hence to decrease cell alteration and damage.

#### *Innate immunity and bacterial control: permanent induction of a PGRP gene*

In contrast to all insect endosymbionts sequenced so far, SPE encodes functional genes that promote cell invasion,

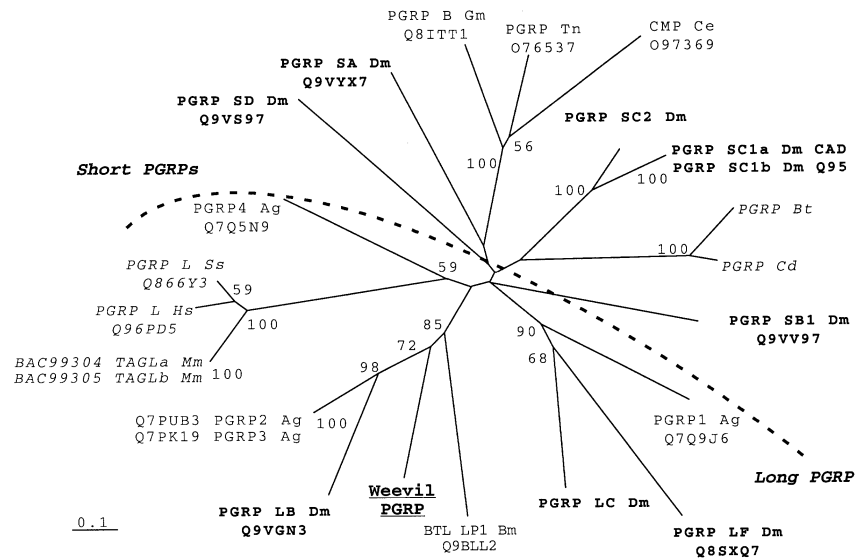
which supports the idea that insect-cooperative bacteria use a pathogenic-like mechanism to enter the host cells (Dale *et al.*, 2002; Lefèvre *et al.*, 2004). Insect molecular responses towards endosymbiont invasion had not been previously reported. In this study, as many as six ESTs, encoding a single peptidoglycan recognition protein (*PGRP*) gene, were found in the subtracted cDNA, with a high differential ratio (DR = 28, Table 1).

The *PGRP* gene family is known to trigger the innate immune defence system via Toll (fungi and Gram<sup>+</sup> bacteria) and Imd (Gram<sup>-</sup> bacteria) pathways that consist of bacterial recognition and of the activation of proteolytic cascades leading to the induction of antibacterial peptide synthesis (Hoffmann and Reichhart, 2002). Bacterial recognition by host *PGRPs* relies on the existence in the bacterial wall of peptidoglycans and lipopolysaccharides (LPS) molecules (Werner *et al.*, 2003). In weevils, the high expression of the *PGRP* gene in the bacteriocyte is regarded as an insect immune response to SPE, eventually via an Imd-like pathway and antimicrobial molecule production. Nevertheless, the fact that the bacteria managed to survive could indicate that either the bacteriocytes produce specific peptides that control the growth of endosymbionts but do not eliminate them, or SPE immunodepresses the bacteriocyte cells downstream from an Imd-like receptor, perhaps by a mechanism similar to those already described in pathogenic bacteria, such as *Yersinia enterocolitica* (Orth, 2002).

Moreover, in addition to the Toll and the Imd pathways, some *PGRP* genes, such as the *PGRP-LB*, are known to interact enzymatically with the bacterial wall. The structure of the *Drosophila* *PGRP-LB* protein was shown to be similar to that of the bacteriophage T7 lysozyme, a zinc-dependent amidase (*N*-acetylmuramyl-L-alanine amidase) (Kim *et al.*, 2003). The amidase activity of the *PGRP-LB* protein was confirmed (Kim *et al.*, 2003) and was shown to rely on the conservation of five amino acid residues H17, Y<sup>46</sup>, H<sup>122</sup>, K<sup>128</sup> (or T) and C<sup>130</sup> (T7 lysozyme numbering) (Cheng *et al.*, 1994; Mellroth *et al.*, 2003). Interestingly, phylogenetic analysis based on the amino acids of the *PGRP* domain indicates that the weevil *PGRP* is closely related, and therefore orthologous, to the *Bombyx mori* bacteriophage T7 lysozyme-like protein as well as to the *Drosophila* *PGRP-LB* gene (Fig. 1). Furthermore, the protein sequence analysis has identified the above five amino acid residues within the weevil gene (data not shown).

Taken together, these findings suggest that the weevil *PGRP* gene might interact with endosymbiont growth and control, either through an Imd-like signalling pathway or by acting directly on the bacterial cell wall through the expression of the amidase activity. The study of the *PGRP* function with respect to the weevil endosymbiont growth will help understanding the interacting mechanisms involved in the early steps of symbiosis. Moreover, what





**Fig. 1.** Phylogenetic tree analysis of selected polypeptides from the PGRP family. Genetic distances were computed by Point Accepted Mutation (PAM) distance and the tree construction was performed by the Neighbour Joining method. For branches supported by bootstrap analysis (500 replications, bootstrap value higher than 50%), the percentage is indicated. Species abbreviation: Ag, *Anopheles gambiae*; Bm, *Bombyx mori*; Bt, *Bos taurus*; Cd, *Camelus dromedarius*; Ce, *Calpodes ethlius*; Dm, *Drosophila melanogaster*; Gm, *Galleria mellonella*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Ss, *Sus scrofa*; Tn, *Trichoplusia ni*; BTL-LP, Bacteriophage T7 lysozyme-like protein; CMP, cuticular molt protein; PGRP, Peptidoglycan recognition protein. For *Drosophila*, the standard nomenclature was: S, short; L, long. *Drosophila* proteins are boldfaced, mammalian proteins are italicized. Amino acid sequences homologous to the weevil PGRP protein were obtained using the BLASTP program in the SPTrEMBL bank. Only sequences with an *e*-value lower than  $10^{-30}$  were conserved. Sequences were aligned using CLUSTALW and alignment was corrected manually using SEAVIEW, and the phylogenetic tree was constructed using the phylo\_win program (<http://pbil.Univ-lyon1.fr>). The analysis demonstrates that the weevil bacteriocyte-expressed PGRP belongs to the subgroup of long PGRPs, as first defined in *Drosophila* (Werner *et al.*, 2000), and is the probable orthologue of the PGRP-LB, a conserved lineage of PGRP genes within insects.

remains to be clarified is how do bacterial recognition and bacterial control by the host evolve in parallel with bacterial genome disintegration during symbiont evolution? For instance, all insect endosymbiont genomes sequenced so far were shown to lack the genes encoding secretion system, and some genes involved in the synthesis of bacterial wall molecules, such as peptidoglycans (Shigenobu *et al.*, 2000; Akman *et al.*, 2002; Gil *et al.*, 2003). In this evolutionary context, we speculate that both bacterial recognition by the host's PGRP (wall structure) and bacterial virulence (secretion system) may attenuate as the symbiosis ages. This assumption is supported by the absence of PGRP sequences in the ESTs obtained so far from aphid bacteriocytes (A. Nakabachi, pers. comm.). This, along with the fact that most of ancient symbionts are surrounded by a third M3 membrane (but not SPE), presumably of eukaryotic origin, suggests that endosymbiont control by the host may switch from innate immunity to an unknown cellular mechanism.

#### *Bacteriocyte cell processes: cell signalling and vesicle trafficking*

Based on the assigned putative gene identity and function, our data predict that the weevil intracellular bacteria modulate several protein profiles related to cytoskeleton

rearrangement, cell signalling and attachment, vesicle trafficking, cell death and tumour processes (Table 1).

Both plant and animal cells undergo extensive cellular remodelling while under attack by pathogens, symbiotic bacteria and fungi. These transformations often induce membrane ruffling and extension, thereby facilitating bacterial pathogen cell internalization and attaching/effacing lesions on target cells. In the weevils, three cytoskeleton-related genes (i.e. alpha-catenin, actinin and tubulin) and one fasciclin I gene were upregulated in the bacteriome tissue.

Fasciclin I gene, an extrinsic membrane protein encoding for the beta-Ig-H3 extracellular domain, is involved in cell adhesion, cell signalling and cell proliferation during morphogenesis. With regard to symbiosis, fasciclin I-like proteins were shown to be upregulated in two extracellular symbiotic models, namely cnidarian-algae (Reynolds *et al.*, 2000) and lichen associations (Paulsrud and Lindblad, 2002). In *Rhizobium*-legume symbiosis, a fasciclin-like gene mutation does not disturb normal nodulation but drastically decreases the percentage of N<sub>2</sub>-fixing nodules (Oke and Long, 1999). The upregulation of fasciclin I in the weevil intracellular symbiosis emphasizes its potential role in both extra- and intracellular symbioses and points to an exciting way of understanding the regulation of mutualistic associations.

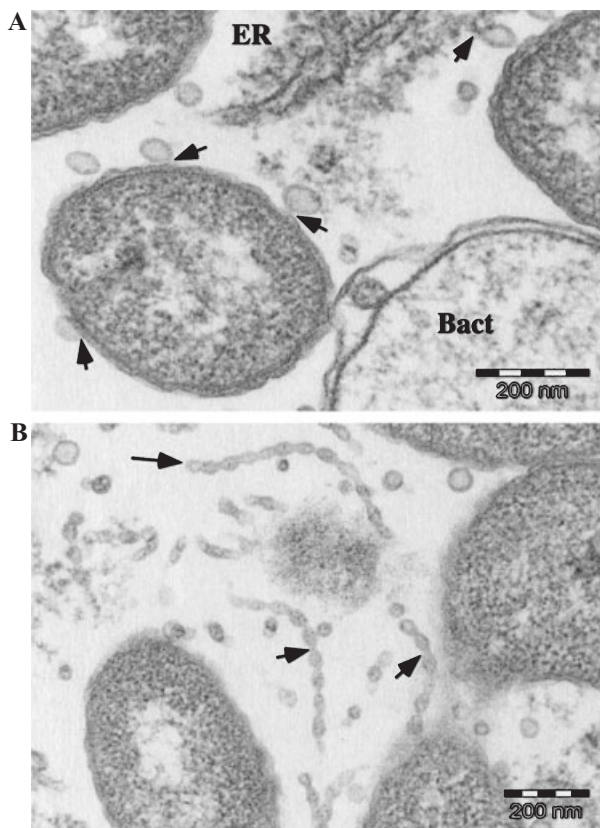
Bacteriocytes also seem to undergo significant vesicle synthesis and trafficking, as shown by the important upregulation of a Rab GTPase and a vacuolar ATPase subunit genes. Similarly, many ATPase subunits and Rab GTPase cDNAs were screened from the bacteriocyte of aphids (A. Nakabachi, pers. comm.). The Rab family is part of the Ras superfamily of small GTPases. Through their effectors, they regulate vesicle formation, actin- and tubulin-dependent vesicle movement, membrane fusion and they may mediate polarized membrane traffic to dynamic cell surface structures. We have investigated this feature in the weevil's bacteriome with electron microscopy (Fig. 2). This yielded intriguing results on vesicle density and location. Some vesicles were tethered at the SPE outer membranes and at the endoplasmic reticulum (ER) membranes (Fig. 2A), and many other vesicles consisted of string structures that are presumably directed to fuse with organelle and cellular membranes. Vesicles were also detected in between the bacteriocyte membrane spaces (data not shown). High vesicle synthesis within the bacteriocytes argues in favour of enhanced polarized cellular traffic, which may help the bacteriome

to fulfil one major complex role, that is the synthesis and transfer of metabolic compounds to the insect body. This function may involve bacterial-specific activities (Miller *et al.*, 2003) as well as host cell pathways working together.

Finally, one of the critical cellular features in insect intracellular symbiosis is the differentiation of the bacteriome tissue (Braendle *et al.*, 2003). Differentiation is likely to result from host-bacterial molecular signalling that recruits pathways related to growth, differentiation and apoptosis. In this study, at least three putative genes exhibiting tumour-apoptosis interplay were identified, namely a calpain, two ESTs homologous to the human-programmed cell death protein 6 (ALG-2) and one translationally controlled protein (TCTP) (Table 1). Calpain is a calcium-activated non-lysosomal neutral thiol protease, which plays crucial roles in various cell functions, such as signal transduction, cell growth and differentiation, apoptosis and necrosis. Interestingly, the ALG-2 gene is also a calcium-binding protein and its role in apoptosis is well documented (Lo *et al.*, 1999). Moreover, among the TCTP proteins that were initially characterized in the growth phase of tumour cells, some were shown to be calcium-binding proteins (Rao *et al.*, 2002). These data suggest that the cellular pathways controlling insect bacteriome tissue homeostasis could be calcium-regulated, as reported recently for plant nodule development (Levy *et al.*, 2004).

In conclusion, the studies described here provide a broad, *in vivo* characterization of bacteriocyte transcriptional responses to intracellular bacteria. The data reveal that intracellular bacteria are able to modulate the expression of insect genes that participate in a variety of fundamental physiological functions, including cell growth, immunity and metabolism. Although data were analysed in terms of simple functional gene clustering, many interconnections between pathways and genes are possible, such as glucose transporter-AR, glutathione-AR, polyol-apoptosis, innate immunity-apoptosis, vesicle synthesis, and these will all need specific investigation. Basically, the usual mutualistic features and beneficial pathways described in insect symbioses totally overlook the profound cellular stress involved in endosymbiosis and the induction of the innate immune system, as attested by the induction of the *PGRP* gene in this work.

From an evolutionary point of view, the bacteriocyte cells could have been selected to skirt around bacterial pathogenicity and insect immunity, probably through the manipulation of Imd-like pathways or by the production of specific peptides. The gradual genome reduction exhibited by intracellular symbionts probably results in a concerted loss of interacting partner's genes, starting from alterations in bacterial virulence gene organization and regulatory features, such as displayed by the heat-shock



**Fig. 2.** Electron microscopy analysis of the *S. zeamais* bacteriocytes. Bact: Bacteria; arrows indicate individual vesicles associated with bacteria and with endoplasmic reticulum (ER) (A) and vesicles grouped as strings (B).

regulon of the older aphid symbiosis (Wilcox *et al.*, 2003). Such progressive decay is probably tightly controlled by adaptive processes acting at the transcriptional regulation level but also shaping out some symbiont and host genes, as it occurs in the GroEL operon and molecules from the *Buchnera* symbiosis (Fares *et al.*, 2004). MTS and cellular genes provide evidence of genetic traits that confer selective advantages to the association. Notably sugar transport and metabolic pathways, adaptation to cellular stress and the molecular signalling network are intriguing physiological mechanisms that will contribute to our understanding of how prokaryotes and eukaryotes interact and coevolve, thereby providing new tools in the study of animal and plant evolution.

## Experimental procedures

### Insect strain selection

*Sitophilus* spp. are naturally infected with two types of bacteria: SPE, the bacteriocyte inducing  $\gamma$ -Proteobacterium, and *Wolbachia*, a rickettsia-like bacterium known to alter invertebrate reproduction (Heddi *et al.*, 1999a). To avoid any *Wolbachia* side-effects on the bacteriocyte gene expression, and to focus on SPE impact only, several *Sitophilus* spp. strains were tested for the presence/absence of *Wolbachia* by polymerase chain reaction (PCR) and fluorescence *in situ* hybridization (FISH), using specific *Wolbachia* primers and probes, as described previously (Heddi *et al.*, 1999a). *S. zeamais* Lagoa strain was found to be monosymbiotic with SPE only, and was therefore used in this work. SPE-aposymbiotic insects were obtained by keeping wild-type symbiotic adults at 35°C and 90% relative humidity for 1 month (Nardon, 1973) followed by female oviposition and insect rearing at 27.5°C and 90% relative humidity on wheat grains. The resulting aposymbiotic insects are less fertile than the wild type, and they develop slowly during the larval stages and they are unable to fly as adults (Nardon, 1988).

### Total RNA isolation

A total of 200 larval bacteriomes were dissected from fourth instar larvae (after larvae beheading) in a drop of buffer A (KCl 25 mM, MgCl<sub>2</sub> 10 mM, sucrose 250 mM, Tris-HCl 35 mM, pH 7.5) and 10 aposymbiotic fourth instar larvae were collected from inside the wheat grains. Total RNA was extracted with the TRIzol reagent (Gibco BRL) and treated with RNase-free DNase I (Promega) to remove any DNA contamination. RNA was then purified through Rneasy mini kit columns (Qiagen), as described in the manufacturer's procedure. After quantification with the spectrophotometer at 260/280 nm, total RNA quality was checked on a 1.2% phosphate-buffer-agarose gel and stored in DEPC-treated H<sub>2</sub>O at -80°C, prior to cDNA synthesis.

### cDNA subtraction

Bacteriome-specific or dominantly expressed genes were screened using a suppressive subtraction hybridization technique with the PCR-select cDNA subtraction kit (Clontech Labo-

ratories), according to the manufacturer's instructions. As bacteriomes do not differentiate in the absence of bacteria in weevils, and as the object of this study was to screen bacteriocyte upregulated genes (but not downregulated ones), a subtracted library was obtained by using the larval bacteriome cDNA as a tester and the whole aposymbiotic larval cDNA as a driver. Moreover, subtraction was performed at a ratio of 1:40 of bacteriome cDNA (tester) to larval cDNA (driver) in order to suppress house-keeping genes and larval non-symbiotic related genes from the bacteriocyte cDNA library as much as possible.

Briefly, cDNA synthesis was performed from 1  $\mu$ g of total RNA, as described in the SMART PCR cDNA synthesis kit (Clontech). The bacteriome cDNA was digested with RsaI, ligated to two adaptors separately (1 and 2R), then hybridized to RsaI-digested aposymbiotic larvae cDNA. After hybridization, the subtracted cDNAs were amplified by two rounds of PCR (12 cycles each) using two different primer sets from the two adaptors (Clontech, cDNA subtraction kit). The amplified products were directly cloned into *pCR 2.1-TOPO* plasmid (Invitrogen). After transformation by electroporation, colonies were randomly selected and subjected to reverse Northern blot to confirm the differential gene expression status. Some genes were additionally checked either with Northern blotting or with real-time RT-PCR.

### Reverse Northern blot analysis

Clones from the subtracted cDNA library were amplified by PCR using Nested 1 and Nested 2R as forward primer and reverse primer respectively (Clontech, cDNA subtraction kit). To confirm the gene differential status, PCR products were run in equal amounts on two 1% TAE-agarose gels and transferred (after denaturation in 0.4 M NaOH and 1 M NaCl, and neutralization in 1.5 M NaCl, Tris-HCl 0.5 M, pH 7.5) onto nylon membranes (Hybond-N, Amersham) in order to obtain two identical blots. Blots were then pre-hybridized (48 h at 42°C), hybridized (48 h at 42°C, as described previously (Heddi *et al.*, 1999b) separately with the aposymbiotic larval cDNA probe and with the larval bacteriome cDNA probe, and then exposed to a Hyperfilm (MP, Amersham) for Phosphorimager analysis (STORM, Amersham). A glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) gene fragment was loaded onto the reverse Northern blot membranes as an internal control. The *GAPDH* fragment was amplified from the total weevil DNA by PCR using degenerated primers *GAPDH*-for (5'-CCWTCYGCYATGCSCCCATG-3') and *GAPDH*-Rev (5'-ACRCGGAARGCCATRCRCGTCA-3') and its identity was confirmed by the sequencing of the PCR products. The *GAPDH* fragments were cut from the agarose gel and sequenced (Genome express, data not shown).

### cDNA probe labelling

cDNA probes were labelled with <sup>32</sup>P-dCTP (3000 Ci mmol<sup>-1</sup>, Amersham) by PCR using Nested primers 1 and 2R (Clontech). Reaction cocktails consisted of 1 U of Taq-Long-Expand polymerase (Roche), 1.5 mM MgCl<sub>2</sub>, 0.2  $\mu$ M primers, 0.2 mM deoxynucleoside triphosphate (dATP, dTTP, dGTP), 2  $\mu$ M dCTP, 0.66  $\mu$ M (100  $\mu$ Ci) <sup>32</sup>P-dCTP, and 50 ng of single-stranded cDNA template in a final volume of 50  $\mu$ l. The PCR parameters were 94°C for 1 min, followed by 18 cycles at 94°C for 15 s, 65°C for 30 s, and 68°C for 6 min. In parallel to the hot PCR, a cold PCR



was made in the same conditions, except that  $^{32}\text{P}$ -dCTP was omitted and replaced by the same amount of cold dCTP (i.e. 0.66  $\mu\text{M}$ ). Cold PCR products were separated through a TAE-agarose gel to check the cDNA smear profile, the size of which should be between 0.3 and 1 kb. In this case, the hot PCR product was purified through a Sephadex G-50 column (Roche) and radioactivity was counted on a Packard apparatus (Camberra company). Several independent probe couples (i.e. bacteriome and aposymbiotic larvae) were used in this study and their activities were very similar, at around  $1.25 \times 10^7$  c.p.m.  $\mu\text{g}^{-1}$  of cDNA.

#### Clone sequencing, blast homology searching and sequence analyses

A set of 153 differential clones, which had been detected by the reverse Northern analysis and showed significant differential ratios, was sequenced by Genome Express (Grenoble, France) on one or both strands using the nested PCR primer 1 or 2R (Clontech).

Sequences were screened individually for open reading frames (ORFs) and translated coding sequences (CDS) were blasted in batch against a non-redundant small protein database at EBI (<http://www.ebi.ac.uk>; WU-Blastp parametered with BLOSUM45 and 10-score recovery). The resulting http file was run for any relevant information retrieval: the SWISSPROT accession number of best hits, links through SWISSPROT to complete hit descriptors, links to protein domains, and cross-links to the Fly-Base *Drosophila* genome database, on which the final annotation check relied, as well as the functional classification of genes (*Drosophila* Gene Ontology). When the expectation value of the protein hit was  $<10^{-25}$  and the hit organism was an insect, the fullest standardized annotation was reported (over the 10 listed). When no insect hit was available, the term 'homologous to' was added in front of the annotation report. Hits with  $e < 10^{-15}$  were qualified as 'possible', whereas hits with  $e < 10^{-2}$  on *Drosophila* were qualified as 'Hypothetical Dm-CGn-like' proteins, with potential protein motif/signatures added for information. Sequences without ORFs were blasted against the European Molecular Biology Laboratory (EMBL) and yielded a few more non-eukaryotic genes, mainly ribosomal genes. A small database of annotated genes was maintained and helped with the automated generation of EMBL submission forms, and with the presented tables.

#### Electron microscopic analysis of the bacteriome tissue

Bacteriome tissues were isolated from fourth instar larvae and fixed in 3% glutaraldehyde in 0.1 M cacodylate buffer (pH 7.4) for 2 h at 4°C. The bacteriomes were postfixated in 1% osmium tetroxide in the cacodylate buffer and in-block staining was performed in 1% aqueous uranyl acetate for 16–18 h. After dehydration, the tissue was embedded in Epon resin, cut into ultrathin sections, stained with lead citrate, and examined with a Zeiss EM 900 electron microscope.

#### Tissue polyol and sugar analysis

Whole L4 larvae ( $n = 5$ – $15$ ) from symbiotic or aposymbiotic *S. zeamais* Laga strains were ground twice in a final volume of

1 ml 70% ethanol in potter/ependorf 1.5 ml tubes. Insect tissues, either bacteriocytes or digestive tracts, were dissected out at 4°C in sucrose-free buffer A (KCl 25 mM,  $\text{MgCl}_2$  10 mM, Tris 35 mM, NaCl 125 mM, pH 7.5). Tissues were pooled on weighed aluminium foil ( $n = 20$  for bacteriocytes,  $n = 5$  for digestive tracts), watered-out with a paper drain, and weighted prior to sugar extraction in 1 ml 70% ethanol. All samples were sonicated in a water bath for  $2 \times 30$  s (resulting in aluminium micronization and tissue grinding for dissected samples), and extracted for 1 h at room temperature. After centrifugation (10 000 g, 5 min, 4°C), the supernatant was collected and vacuum-dried for sugar analysis. Pellets were collected for protein analysis through acid hydrolysis and amino-acid autoanalysis. Three to five repeats were performed per modality.

For sugar analysis, dried samples were dissolved in 50  $\mu\text{l}$  Milli-Q water and injected in a Dionex DX 500 HPLC-system (Dionex, Sunnyvale, CA), equipped with a CarboPac PA1 guard ( $4 \times 50$  mm) and analytical column ( $4 \times 250$  mm) and an ED 40 Electrochemical Detector for Pulsed Amperometric Detection (PAD). The column was eluted with 1 M NaOH and Milli-Q water (10%:90%, 1 ml  $\text{min}^{-1}$ ) and kept at 20°C during analysis. Sugar and polyol peaks were identified by their retention times and quantified based on calibration standards at 2.5 p.p.m., 5 p.p.m., 7.5 p.p.m. and 10 p.p.m. respectively. The concentrations of the individual sugars and polyols were analysed using the program PEAKNET Software Release 5.1 (DX-LAN module). The statistical analysis of sugar and polyol concentrations in weevil tissues was performed by *t*-tests using the JMP software (SAS Institute).

#### Acknowledgements

Authors gratefully acknowledge A. Nakabachi for providing information on the ESTs obtained from the aphid bacteriocytes, P. Nardon and C. Khatchadourian for technical assistance, V. James for English correction, M. McFall-Ngai for the critical reading of the manuscript. This work was supported by the Institut National de la Recherche Agronomique (INRA) and the Institut National des Sciences Appliquées (INSA) Institutions.

#### References

- Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shibata, T., Hattori, M., *et al.* (2002) Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nat Genet* **32**: 402–407.
- Bordenstein, S.R., O'Hara, F.P., and Werren, J.H. (2001) *Wolbachia*-induced incompatibility precedes other hybrid incompatibility in *Nasonia*. *Nature* **409**: 707–710.
- Braendle, C., Miura, T., Bickel, R., Shingleton, A.W., Kambhampati, S., and Stern, D. (2003) Developmental origin and evolution of bacteriocytes in the Aphid-*Buchnera* symbiosis. *PLoS Biology* **1**: 70–76.
- Cheng, X., Zhang, X., Pflugrath, J.W., and Studier, F.W. (1994) The structure of bacteriophage T7 lysozyme, a zing amidase and an inhibitor of T7 RNA polymerase. *Proc Natl Acad Sci USA* **91**: 4034–4038.
- Dale, C., Plague, G.R., Wang, B., Ochman, H., and Moran, N.A. (2002) Type III secretion system and the evolution of mutualistic endosymbiosis. *Proc Natl Acad Sci USA* **99**: 12397–12402.

- Dale, C., Wang, B., Moran, N., and Ochman, H. (2003) Loss of DNA recombinational repair enzymes in the initial stages of genome degeneration. *Mol Biol Evol* **20**: 1188–1194.
- Douglas, A.E. (1998) Host benefit and the evolution of specialization in symbiosis. *Heredity* **81**: 599–603.
- Fares, M.A., Moya, A., and Barrio, E. (2004) GroEL and the maintenance of bacterial endosymbionts. *Trends Genet* **20**: 413–416.
- Foster, J.S., Apicella, M.A., and McFall-Ngai, M. (2000) *Vibrio fischeri* Lipopolysaccharide induces developmental apoptosis, but not complete morphogenesis, of the *Euprymna scolopes* symbiotic light organ. *Dev Biol* **226**: 242–254.
- Gil, A., Silva, F.J., Zientz, E., Delmotte, F., Gonzalez-Candelas, F., Latorre, A., et al. (2003) The genome sequence of *Blochmannia floridanus*. Comparative analysis of reduced genomes. *Proc Natl Acad Sci USA* **100**: 9388–9393.
- van Ham, R.C., Kamerbeek, J., Palacios, C., Rausell, C., Abascal, F., Bastolla, U., et al. (2003) Reductive genome evolution in *Buchnera aphidicola*. *Proc Natl Acad Sci USA* **100**: 581–586.
- Heddi, A. (2003) Endosymbiosis in the weevil of the genus *Sitophilus*: Genetic, physiological, and molecular interactions among associated genomes. In *Insect Symbiosis*. Bourtzis, K. and Miller, T. (eds). Boca Raton: CRC Press, pp. 67–82.
- Heddi, A., Grenier, A.M., Khatchadourian, C., Charles, H., and Nardon, P. (1999a) Four intracellular genomes direct weevil biology: Nuclear, mitochondrial, principal endosymbionts, and *Wolbachia*. *Proc Natl Acad Sci USA* **96**: 6814–6819.
- Heddi, A., Stepien, G., Benke, P.J., and Wallace, D.C. (1999b) Coordinate induction of energy gene expression in tissues of mitochondrial disease patients. *J Biol Chem* **274**: 22968–22976.
- Hoffmann, J.A., and Reichhart, J.M. (2002) *Drosophila* innate immunity: an evolutionary perspective. *Nat Immunol* **3**: 121–126.
- Hooper, L.V., Wong, M.H., Thelin, A., Hansson, L., Falk, P.G., and Gordon, J.I. (2001) Molecular analysis of commensal host-microbial relationships in the intestine. *Science* **291**: 881–884.
- Itoh, T., Martin, W., and Nei, M. (2002) Acceleration of genomic evolution caused by enhanced mutation rate in endocellular symbionts. *Proc Nat Acad Sci USA* **99**: 12944–12948.
- Kim, M.S., Byun, M., and Oh, B.H. (2003) Crystal structure of peptidoglycan recognition protein LB from *Drosophila melanogaster*. *Nat Immunol* **4**: 787–793.
- Lefèvre, C., Charles, H., Vallier, A., Delobel, B., Farrell, B., and Heddi, A. (2004) Endosymbiont phylogenesis in the Dryophthoridae weevils: evidence for bacterial replacement. *Mol Biol Evol* **21**: 965–973.
- Levy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., et al. (2004) A putative Ca<sup>2+</sup> and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science* **303**: 1361–1364.
- Lo, K.W.H., Zhang, Q., Li, M., and Zhang, M. (1999) Apoptosis-linked gene product ALG-2 is a new member of the calpain small subunit subfamily of Ca<sup>2+</sup>-binding proteins. *Biochemistry* **38**: 7498–7508.
- Margulis, L. (1993) *Symbiosis in Cell Evolution: Microbial Communities in the Archean and Proterozoic Eons*, 2nd edn. New York: Freeman, p. 452.
- Maynard Smith, J. (1989) Generating novelty by symbiosis. *Nature* **341**: 284–285.
- Mellroth, P., Karlsson, J., and Steiner, H. (2003) A scavenger function for a *Drosophila* peptidoglycan recognition protein. *J Biol Chem* **278**: 7059–7064.
- Merezhkowsky, C. (1920) La plante considérée comme un complexe symbiotique. *Bull Soc Sc Nat Ouest, 3ème Série, t. VI*: 17–98.
- Miller, S.I., Bader, M., and Guina, T. (2003) Bacterial vesicle formation as a mechanism of protein transfer to animal. *Cell* **115**: 2–3.
- Moran, N.A. (1996) Accelerated evolution and Muller's ratchet in endosymbiotic bacteria. *Proc Natl Acad Sci USA* **93**: 2873–2878.
- Nardon, P. (1973) Obtention d'une souche asymbiotique chez le charançon *Sitophilus sasakii* Tak: différentes méthodes d'obtention et comparaison avec la souche symbiotique d'origine. *C R Acad Sci Paris* **277D**: 981–984.
- Nardon, P. (1988) Cell to cell interactions in insect endocytobiosis. In *Cell to Cell Signals in Plant, Animal and Microbial Symbiosis*. Scannerini, S., Smith, D., Bonfante-Fasolo, P. and Gianinazzi-Pearson, V. (eds). Berlin: Springer-Verlag, pp. 85–100.
- Nyholm, S.V., Stabb, E.V., Ruby, E.G., and McFall-Ngai, M.J. (2000) Establishment of an animal-bacterial association: Recruiting symbiotic vibrios from the environment. *Proc Nat Acad Sci USA* **97**: 10231–10235.
- Oke, V., and Long, S.R. (1999) Bacteroid formation in the rhizobium-legume symbiosis. *Curr Opin Microbiol* **2**: 641–646.
- Orth, K. (2002) Function of the *Yersinia* effector YopJ. *Curr Opin Microbiol* **5**: 38–45.
- Paulsrud, P., and Lindblad, P. (2002) Fasciclin domain proteins are present in *Nostoc* Symbionts of lichens. *Appl Environ Microbiol* **68**: 2036–2039.
- Rao, K.V.N., Chen, L., and Gnanasekar, M. (2002) Cloning and characterization of a calcium-binding, histamine-releasing protein from *Schistosoma mansoni*. *J Biol Chem* **277**: 31207–31213.
- Reynolds, W.S., Schwarz, J.A., and Weis, M. (2000) Symbiosis-enhanced gene expression in cnidarian-algal associations: cloning and characterization of a cDNA, sym32, encoding a possible cell adhesion protein. *Comp Biochem Physiol* **126**: 33–44.
- Rio, R.V.M., Lefevre, C., Heddi, A., and Aksoy, S. (2003) Comparative genomics of insect-symbiotic bacteria: Influence on microbial genome composition. *Appl Environ Microbiol* **69**: 6825–6832.
- Schauser, L., Roussis, A., Stiller, J., and Stougaard, J. (1999) A plant regulator controlling development of symbiotic root nodules. *Nature* **402**: 191–195.
- Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y., and Ishikawa, H. (2000) Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera sp.* APS. *Nature* **407**: 81–86.
- Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., et al. (2002) A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* **417**: 959962.
- Suarez, G., Rajaram, R., Bhuyan, K.C., Oronsky, A.L., and Goidl, J.A. (1988) Administration of an aldose reductase inhibitor induces a decrease of collagen fluorescence in diabetic rats. *J Clin Invest* **82**: 624–627.
- Turunen, S. (1985) Absorption. In *Comprehensive Insect*



- Physiology, Biochemistry and Pharmacology*, Vol. IV. Kerkut, G.A. and Gilbert, L.T.L. (eds). Oxford: Pergamon Press, pp. 241–273.
- Visick, K.L., Foster, J., Doino, J., McFall-Ngai, M., and Ruby, E. (2000) *Vibrio fischeri lux* genes play an important role in colonization and development of the host light organ. *J Bacteriol* **182**: 4578–4586.
- Wernegreen, J. (2002) Genome evolution in bacterial endosymbionts of insects. *Nat Rev* **3**: 850–861.
- Werner, T., Borge-Renberg, K., Mellroth, P., Steiner, H., and Hultmark, D. (2000) A family of peptidoglycan recognition proteins in the fruit fly *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **97**: 13772–13777.
- Werner, T., Borge-Renberg, K., Mellroth, P., Steiner, H., and Hultmark, D. (2003) Functional diversity of the *Drosophila PGRP-LC* gene cluster in the response to lipopolysaccharide and peptidoglycan. *J Biol Chem* **278**: 26319–26322.
- Wilcox, J.L., Dunbar, H.E., Wolfinger, R.D., and Moran, N. (2003) Consequences of reductive evolution for gene expression in an obligate endosymbiont. *Mol Microbiol* **48**: 1491–1500.
- Wolff, S.P., Bascal, Z.A., and Hunt, J.V. (1989) 'Autoxidative glycosylation': free radicals and glycation theory. *Prog Clin Biol Res* **304**: 259–275.