



Article

Identification of the Gene Repertoire of the IMD Pathway and Expression of Antimicrobial Peptide Genes in Several Tissues and Hemolymph of the Cockroach *Blattella germanica*

Leo Zuber ¹, Rebeca Domínguez-Santos ¹, Carlos García-Ferris ^{1,2,*} and Francisco J. Silva ^{1,3,*}

¹ Institute for Integrative Systems Biology (I2SysBio), University of Valencia and CSIC, 46980 Paterna, Spain; leoz1998@gmail.com (L.Z.); rebeca.dominguez@uv.es (R.D.-S.)

² Department of Biochemistry and Molecular Biology, University of Valencia, 46100 Burjassot, Spain

³ Genomics and Health Area, Foundation for the Promotion of Sanitary and Biomedical Research, 46020 Valencia, Spain

* Correspondence: carlos.garcia.ferris@uv.es (C.G.-F.); francisco.silva@uv.es (F.J.S.)

Abstract: Antimicrobial peptide (AMP) genes, triggered by Toll and IMD pathways, are essential components of the innate immune system in the German cockroach *Blattella germanica*. Besides their role in killing pathogenic bacteria, AMPs could be involved in controlling its symbiotic systems (endosymbiont and microbiota). We found that the IMD pathway was active in the adult female transcriptomes of six tissues (salivary glands, foregut, midgut, hindgut, Malpighian tubules and fat body) and hemolymph. Total expression of AMP genes was high in hemolymph and salivary glands and much lower in the other sample types. The expression of specific AMP genes was very heterogeneous among sample types. Two genes, *defensin_g10* and *drosomycin_g5*, displayed relevant expression in the seven sample types, although higher in hemolymph. Other genes only displayed high expression in one tissue. Almost no expression of attacin-like and blattelicin genes was observed in any sample type, although some of them were among the genes with the highest expression in adult female whole bodies. The expression of AMP genes in salivary glands could help control pathogens ingested with food and even determine gut microbiota composition. The low expression levels in midgut and hindgut are probably related to the presence of beneficial microbiota. Furthermore, a reduction in the expression of AMP genes in fat body could be the way to prevent damage to the population of the endosymbiont *Blattabacterium cuenoti* within bacteriocytes.

Keywords: antimicrobial peptides (AMPs); IMD pathway; innate immune response; symbiosis; transcriptome; *Blattella germanica*



Citation: Zuber, L.; Domínguez-Santos, R.; García-Ferris, C.; Silva, F.J. Identification of the Gene Repertoire of the IMD Pathway and Expression of Antimicrobial Peptide Genes in Several Tissues and Hemolymph of the Cockroach *Blattella germanica*. *Int. J. Mol. Sci.* **2022**, *23*, 8444. <https://doi.org/10.3390/ijms23158444>

Academic Editors: Alberto Cuesta and Elena Chaves-Pozo

Received: 22 June 2022

Accepted: 28 July 2022

Published: 30 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Most species of cockroaches (Blattodea) live in tropical/subtropical regions, and only a few interact with humans and may be considered pests [1]. The German cockroach *Blattella germanica* (Linnaeus) is a pest that affects human environments around the world. It impacts human health in several ways, including the transmission of pathogens or by disseminating antibiotic-resistant bacteria. How *B. germanica*, like other cockroaches, has evolved to become a human pest, adapted to unhealthy habitats, is of great interest. Genomics resources are providing molecular understanding of adaptation and immune response. The genomes of some species of cockroaches such as *Periplaneta americana* (Linnaeus) [2] or *B. germanica* [3], in combination with the analysis of transcriptomes [4], have made it possible to identify the expansion of several gene families encoding antimicrobial peptide (AMP) genes that would act on pathogenic bacteria.

However, cockroaches also interact with other non-pathogenic or beneficial bacteria. In fact, cockroaches are paradigmatic, as two symbiotic systems coexist in each individual: a bacterial endosymbiont (*Blattabacterium cuenoti*, hereinafter *Blattabacterium*) located in

bacteriocytes (specialized host cells) in the fat body and ectosymbionts (a rich and varied microbiota) located in the hindgut [5]. The endosymbiont is a member of the phylum Bacteroidetes [6], and its wide distribution among Blattodea species (except most termites) [7] suggests that a single symbiotic association event took place more than 150 million years ago, followed by the coevolution of insect hosts and bacterial symbionts [8–12]. Analysis of the gene repertoires of *Blattabacterium* genomes revealed that the endosymbiont is involved in the nitrogen metabolism of the host, supplying urease, which is responsible for producing ammonia in the final step of the degradation of uric acid (the nitrogen storage compound of cockroaches). Ammonia can be incorporated into glutamate and glutamine by the endosymbiont and the host, respectively, participating in the production of essential amino acids for the host [9,13,14]. The combined metabolism of the insect host and the bacterial symbiont has allowed cockroaches to return to the ancestral system of ammonotelism, frequent in aquatic invertebrates, instead of uricotelism, frequent in most insect species [8].

Moreover, many insect species are associated with a commensal/mutualistic microbiota in the gut, which has an important role in the digestion process of nutrients and helps the absorption of many compounds [15], and while insects have to kill ingested pathogens, they can also maintain the homeostasis of the beneficial gut community.

Pathogens induce the activation of one or two signaling pathways, namely the immune deficiency (IMD) pathway, which is preferentially activated by Gram-negative bacteria, and the Toll pathway, which is preferentially activated by Gram-positive bacteria and fungi, triggering the expression of multiple AMP genes [16,17]. In *B. germanica*, the identification of multiple AMP genes based on whole-body transcriptomic analysis and genome data revealed the expansion of five AMP gene families (defensin, drosomycin, termicin, attacin-like and blattelicin genes), which provide a better defense against pathogens [4]. Blattellicins are a new AMP gene family, evolved from an attacin-like gene in a recent ancestor of *B. germanica*, that encode proteins with an N-terminal signal peptide, a long middle Glx-rich region and an attacin-C domain at the C-terminus.

Furthermore, it has been proposed that AMPs not only have a defensive role against pathogens but also play a part in regulating the interaction of the eukaryotic host with its mutualistic symbionts [18,19]. In this context, it has been recently demonstrated in *Drosophila* that AMPs, and lysozymes to a lesser degree, determine the bacterial community composition and abundance of the microbiota in the gut [20]. Given that *B. germanica* has a rich and varied microbiota, this would imply not having a strong immune response in the midgut or hindgut, or that most species of the microbiota were selected to resist the effect of AMPs more efficiently than pathogens. In addition, it has also been found that AMPs participate in the maintenance and control of endosymbiosis in plants and animals. This was demonstrated in the insect *Sitophilus* with the endosymbiont *Sodalis pierantonius* [21–23], and in plants in the endosymbiosis of legumes with *Rhizobium* [24] and of *Alnus* with *Frankia* [25]. In legumes, AMPs are involved in terminal endosymbiont differentiation, while in *Alnus*, AMPs are involved in the permeabilization of the bacterial membranes, presumably allowing the metabolic complementation with the host.

In this study, we report the levels of expression of AMP genes in six types of tissue and hemolymph of *B. germanica* adult females, with the aim of understanding the way in which the expression of AMPs kills pathogens, maintaining a healthy gut microbiota and regulating the bacterial endosymbiont located in the fat body. We also sought to identify the genes involved in the IMD pathway and analyze their expression, because IMD is probably the main signaling pathway required for the production of those AMPs and may be required for the control of the endosymbiont and for the homeostasis of gut microbiota.

2. Results

2.1. Genes of the IMD Pathway in *B. germanica*

Twenty-seven out of the 30 coding genes described in the 4IN database for the IMD pathway were detected in *B. germanica* (Table 1). Because the *Dredd* gene encoding Death related ced-3/Nedd2-like caspase was duplicated, the complete gene repertoire

included 28 genes (see sequences in Supplementary Materials CDS_IMD_path.fasta and protein_IMD_path.fasta). The sequences of these genes were obtained from adult and nymph transcriptomes. To produce the complete CDS in *cad* and *Dredd_2*, it was necessary to obtain the last 3'-codons from the genome sequence [3]. Because of its length, a complete assembled CDS could not be obtained for *Bruce*, so the sequences included in Supplementary Materials derived from the genome annotation. The reported CDS of only three genes (*akirin*, *ben* and *Npc2*) were identical to the annotated CDS of the genome; in other cases, there were differences of at least one segment, due to the prediction of a different gene structure. The locus tag qualifiers of CDS with complete or partial coverage are indicated (Table 1). The *IKKbeta*, *IKKg* and *key* genes were not identified in either the transcriptomes or the genome.

Table 1. Gene list of the IMD pathway in *B. germanica* based on transcriptomic analysis.

Gene Name (Based on 4IN Database)	Presence in <i>B. germanica</i>	<i>B. germanica</i> Product Name	Protein Length	Located at Locus_Tag †
<i>akirin</i>	YES	Akirin	180	C0J52_14274
<i>ben</i>	YES	Bendless	151	C0J52_14002
<i>Bruce</i>	YES	BIR repeat containing Ubiquitin-conjugating enzyme	4175 ‡	C0J52_00682
<i>cad</i>	YES	Caudal	281 §	
<i>casp</i>	YES	Caspar	669	C0J52_10382
<i>CYLD</i>	YES	Cylindromatosis	1147	C0J52_05308
<i>CASP</i>	YES	Caspase	307	C0J52_04212
<i>Diap2</i>	YES	Death-associated inhibitor of apoptosis 2	578	C0J52_01770
<i>dnr1</i>	YES	Defense repressor 1	544	C0J52_09439
<i>Dredd_1</i>	YES	Death related ced-3/Nedd2-like caspase	592	C0J52_14092
<i>Dredd_2</i>	YES	Death related ced-3/Nedd2-like caspase	572 §	C0J52_22862
<i>dsp1</i>	YES	Dorsal switch protein 1	205	C0J52_00951
<i>Duox</i>		Dual oxidase	1544	C0J52_05752
<i>eff</i>	YES	Effete	147	
<i>FADD</i>	YES	FADD	245	
<i>imd</i>	YES	Immune deficiency	252	C0J52_19439
<i>nemo</i>	YES	NF-kappa-B essential modulator	435	
<i>Npc2</i>	YES	Niemann-Pick type C-2	148	C0J52_22057
<i>Ntf2</i>	YES	Nuclear transport factor 2	130	
<i>PGRP-LC</i>	YES	Peptidoglycan recognition protein	365	C0J52_21009
<i>rel</i>	YES	Relish	1010	C0J52_13050
<i>RYBP</i>	YES	Ring and YY1 Binding Protein	182	C0J52_24359 # C0J52_26497 #
<i>scny</i>	YES	Scrawny	740	C0J52_00345
<i>Skp2</i>	YES	S-phase kinase-associated protein 2 Skp2	466	C0J52_01896
<i>SkpA</i>	YES	S-phase kinase-associated protein 1-related A	162	C0J52_18710
<i>Tab2</i>	YES	TAK1-associated Binding Protein 2	456	C0J52_07566
<i>Tak1</i>	YES	TGF-beta activated kinase 1	471	C0J52_05441
<i>Uev1A</i>	YES	Ubiquitin-conjugating enzyme variant 1A	144	C0J52_05949
<i>IKKbeta</i>	NOT	I-kappaB kinase beta		
<i>IKKg</i>	NOT	IKK gamma		
<i>key</i>	NOT	Kenny		

† The annotated CDS in the genome and those obtained from the transcriptomes differ in more than one exon in most cases. ‡ Because a complete transcript could not be obtained, the CDS was taken from the genome. § The last codons of the CDS were recovered from the genome sequence. # The transcript is located in two genome places with 100% identity and 100% coverage.

2.2. Expression of IMD Pathway Genes in Six Tissues and Hemolymph of *B. germanica*

We analyzed the levels of expression of 28 IMD pathway genes in female adult tissues (fat body, foregut, hindgut, Malpighian tubules, midgut and salivary glands) and hemolymph by estimating the means of the normalized expression (GeTMM) (Figure 1).

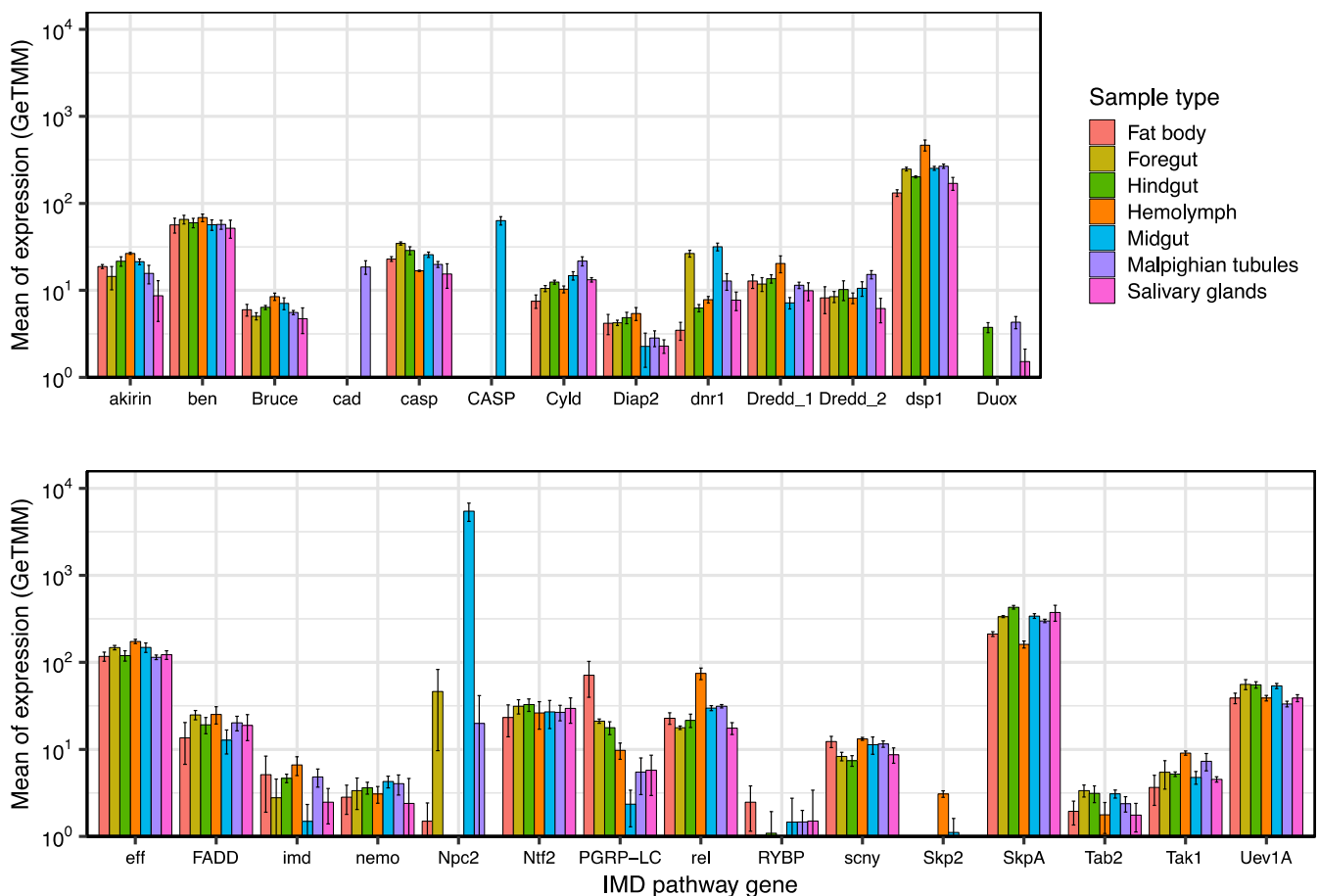


Figure 1. Expression of genes of the IMD pathway in adult females of *B. germanica*. Values are the means plus standard deviations of normalized expression (GeTMM) of each gene in each sample type. Graph bars are standard deviations. See Table 1 for product names. Y-axis is in log scale. The lower ends of the error bars are not shown below 10^0 . Genes are arranged in alphabetical order from left to right and top to bottom.

Most of the genes were expressed in the seven sample types. For example, the *rel* gene, encoding Relish, a key downstream transcription factor involved in the expression of AMP genes [26], is transcribed at a comparable level in the six tissues and slightly higher in hemolymph. Other genes, such as *dsp1*, *eff* and *SkpA*, are highly transcribed in all of the analyzed samples over 10^2 GeTMM.

Furthermore, some genes were differentially expressed in one or more sample types (*cad*, *CASP*, *Duox*, *Npc2*, *RYBP* and *Skp2*). For example, the *cad* gene (encoding Caudal), which in *Drosophila* functions as a gut-specific transcriptional repressor of NF- κ B-dependent induction of AMP genes [27], was almost undetectable in most tissues, except in Malpighian tubules. The *CASP* gene was only substantially expressed in midgut. It encodes Caspase-3, a member of the caspase family of cysteine proteases, previously described as inducing apoptosis in the midgut of *P. americana* under starvation conditions [28]. The most highly expressed IMD pathway gene was *Npc2* (*Niemann-Pick type C-2*), but its expression took place mainly in midgut. The most probable reason for such a high expression is its role in the availability of sterol substrate, as is the case for some *Drosophila Npc2* genes [29]. In fact, *Npc2d* was one of the differentially metabolic overexpressed genes detected in *Drosophila* when comparing normal (with microbiota) to germ-free adult midguts [30].

2.3. Expression of AMP Genes in *B. germanica* Tissues and Hemolymph

The whole expression of AMP genes in six tissues and hemolymph of adult females (Figure 2) was determined by estimating the means of the total normalized expression (GeTMM). Comparatively, AMP genes were highly expressed in hemolymph and salivary glands, and around two/three orders of magnitude lower in the digestive tract, Malpighian tubules and fat body.

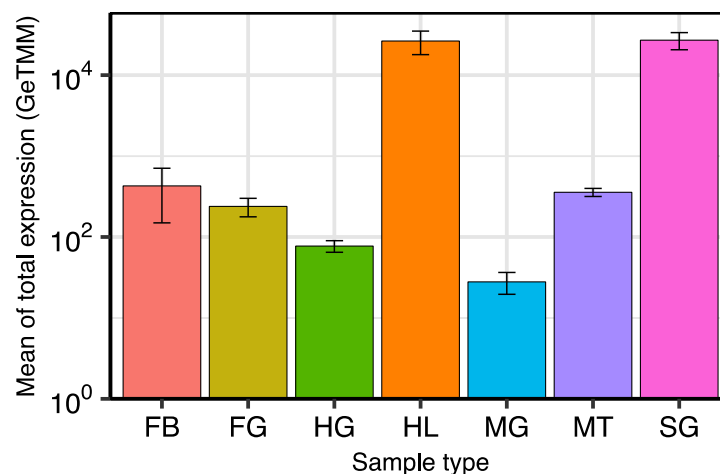


Figure 2. Total normalized expression (GeTMM) of AMP genes in seven sample types of *B. germanica*. Values are the means of expression of all AMP genes in each sample type. Graph bars are standard deviations. Abbreviations: FB (fat body), FG (foregut), HG (hindgut), HL (hemolymph), MG (midgut), MT (Malpighian tubules) and SG (salivary glands). Y-axis is in log scale.

The average levels of expression of individual AMP genes in each sample type revealed large differences among samples and among genes (Figure 3). In general, the expression of most AMP genes was very heterogeneous among sample types, although two genes, *defensin_g10* and *drosomyacin_g5*, were expressed in the seven samples with GeTMM values $> 10^1$. Other genes with expression in most sample types were *defensin_g9*, *defensin_g11-g12*, *drosomyacin_g2* and *drosomyacin_g6*. Some genes displayed very high expression ($>10^3$ GeTMM) in only one sample (hemolymph or salivary glands). However, it was relevant that genes that are relatively highly expressed in whole bodies, such as *blattellicin_g4* and attacin-like genes (Figure 3), were expressed hardly at all in any of the analyzed tissues. The expression pattern of each AMP gene in the seven analyzed samples is shown in Figure S1.

Because hemolymph and salivary glands were the sample types with the highest total AMP gene expression, we analyzed defensin and drosomyacin genes (Figure 4). Most of the expressed genes showed significant differential expression, higher in hemolymph (*defensin_g9*, *defensin_g10*, *defensin_g11-g12*, *defensin_g13*, *defensin_g14*, *drosomyacin_g5* and *drosomyacin_g6*) or in salivary glands (*defensin_g3-g5*, *defensin_g6*, *defensin_g15-16*, *drosomyacin_g2* and *drosomyacin_g3*). Some genes displayed very high expression ($>10^3$ GeTMM), others high expression ($>10^2$ GeTMM), and others moderate expression ($>10^1$ GeTMM). The expression of the other two types of AMP genes, attacin-like and blattellicin genes, is shown in Figure S1.

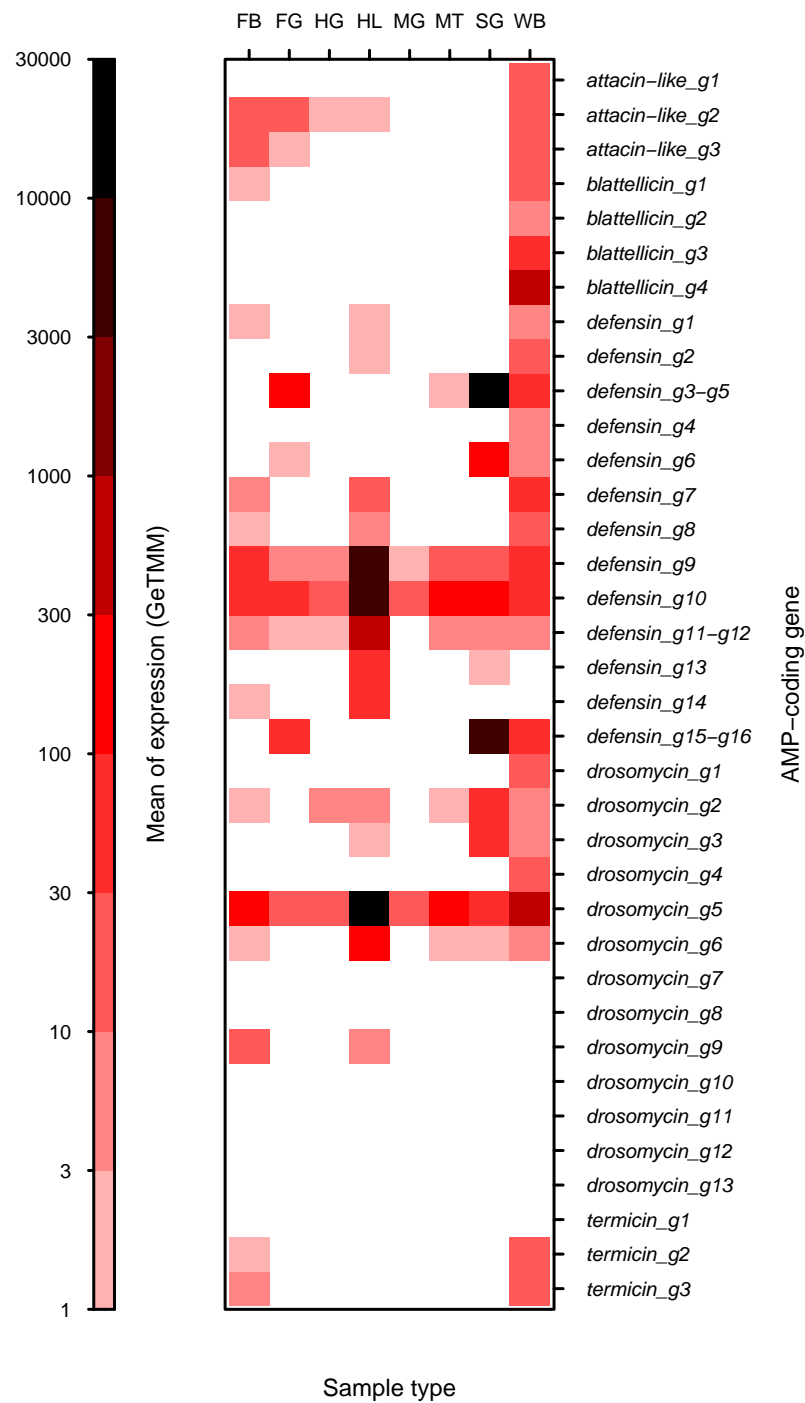


Figure 3. Gene expression of AMP genes in seven sample types and whole bodies of *B. germanica*. Heatmap analysis of normalized expression (GeTMM) of AMP genes. Genes with identical CDS were grouped. The levels of expression in whole bodies (WB) are the means of three female adult SRA projects (SRR5458592.1, SRR5458593.1 and SRR6784711.1). See Figure 2 for abbreviations of sample types. GeTMM values < 1 are in white.

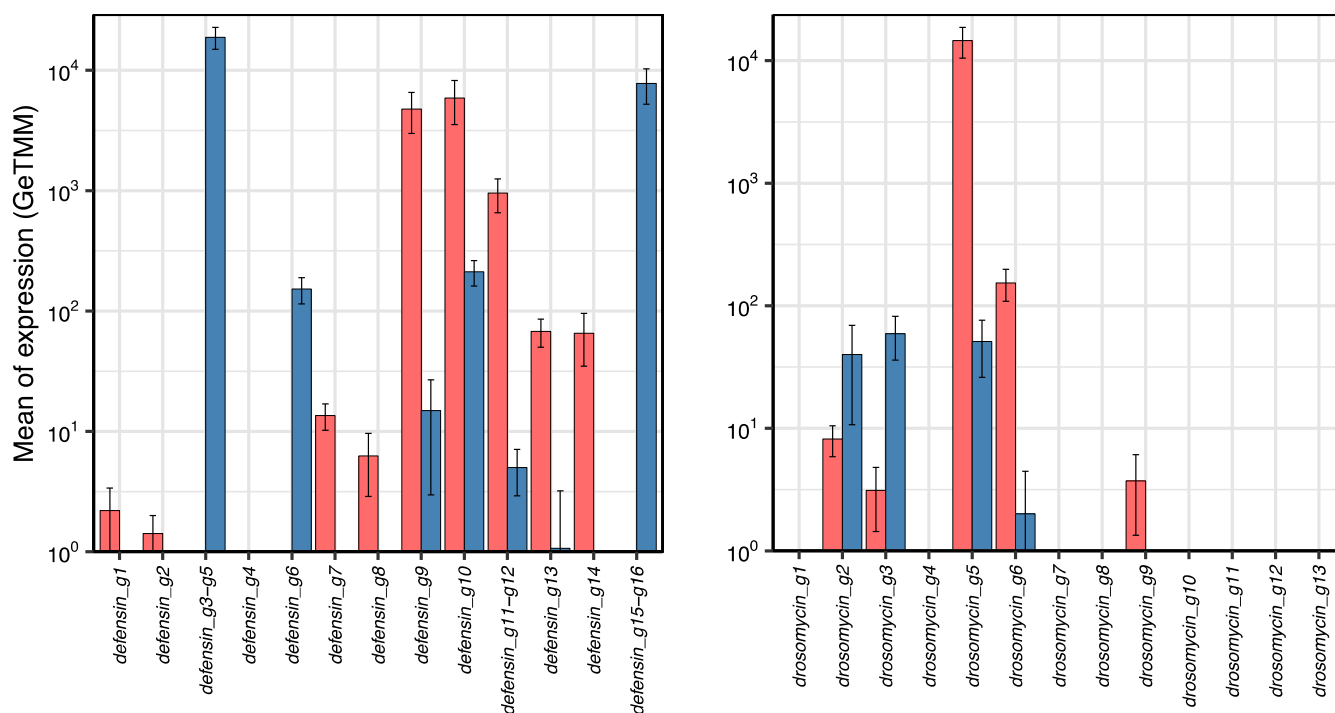


Figure 4. Normalized expression (GeTMM) of defensin and drosomycin genes in hemolymph and salivary glands of *B. germanica*. Values are the means plus standard deviations of gene expression of each sample type. Graph bars are standard deviations. Code colors are pink (hemolymph) and blue (salivary glands). Attacins, blattellicins and termicins are not shown because of their very low mean values. Y-axis is in log scale. The lower ends of the error bars are not shown below 10^0 . All expressed genes, except *defensin_g1* and *defensin_g2*, show a differential expression between the two sample types ($|\log_2\text{FoldChange}| > 2$). Differential expression was significant (adjusted p -value < 0.001) except for *defensin_g7*, *defensin_g8* and *drosomycin_g9*.

Even though the expression of AMP genes in fat body was relatively low (Figure 5), they were analyzed because it is an important organ that harbors bacteriocytes, specialized cells containing the ancient bacterial endosymbiont *Blattabacterium*, and in some insects the fat body is an important organ involved in the production of AMPs for the hemolymph [31]. The average expression of *attacin-like_g2* and *attacin-like_g3* was due to one out of the four replicas. In this, the GeTMM value was higher than 10^2 , while in the three others, the value was smaller than 10^1 . Although low, fat body was the sample type with the highest expression level of some genes such as *drosomycin_g9*, *termicin_g2* and *termicin_g3* (Figure S1).

Most AMP genes were not expressed in foregut, midgut, hindgut and Malpighian tubules (Figure S2). Hindgut and midgut were the tissues with the lowest total AMP gene expression (Figure 2). In both cases, only the widely expressed *drosomycin_g5* and *defensin_g10* displayed an expression level higher than 10^1 (Figure S2).

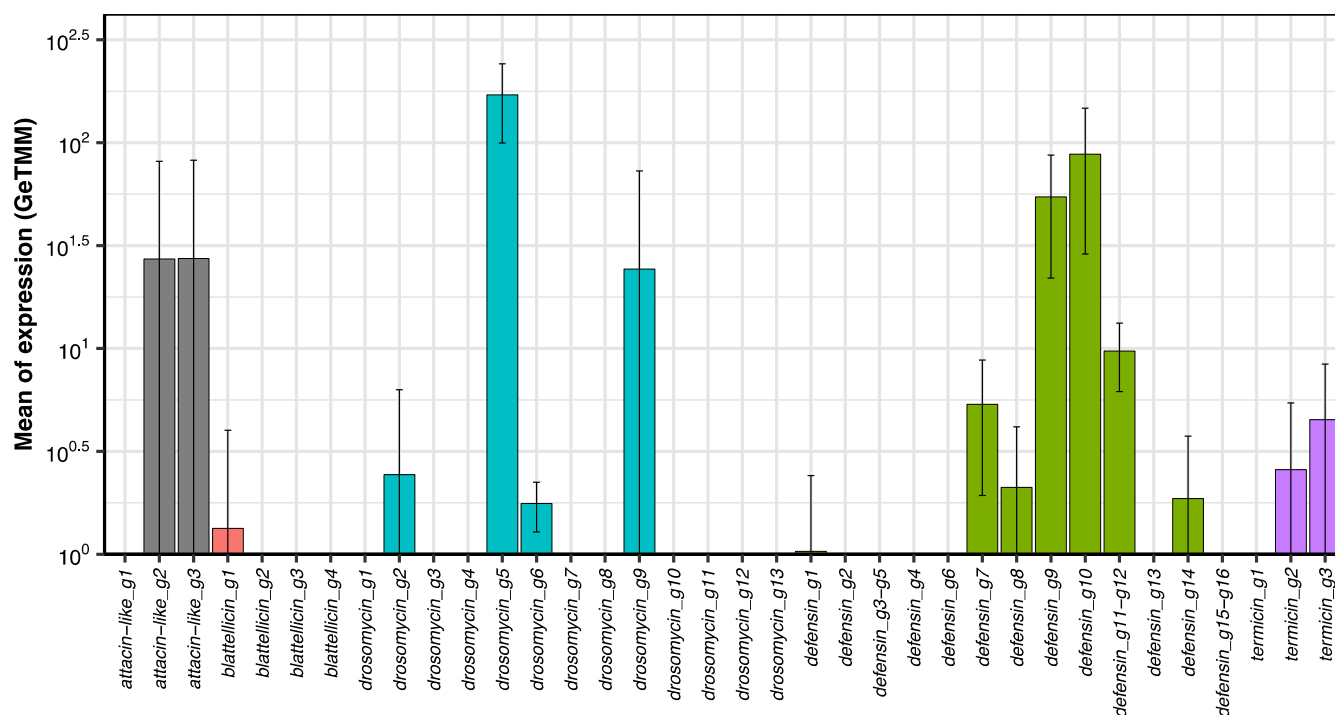


Figure 5. Expression of AMP genes in fat body of *B. germanica*. Means and standard deviations of normalized expression (GeTMM) of AMP genes. Graph bars are standard deviations. Y-axis is in log scale. The lower ends of the error bars are not shown below 10^0 . Color codes for gene families are: attacin-like (gray), blatttellicin (pink), drosomycin (cyan), defensin (green) and termicin (violet).

3. Discussion

AMPs are the main effector molecules of the innate immune system in insects and enable them to fight microbial infections (bacteria, fungi and viruses). The expression of the different AMP genes is triggered as a consequence of the detection of the invasive microorganisms (infecting microbes) by two major NF- κ B-mediated signaling pathways: IMD (which provides protection mainly against Gram-negative bacteria and some Gram-positive bacteria) and Toll (which provides protection mainly against fungi and other Gram-positive bacteria) [32]. Two main modes of producing AMPs appear to exist in insects. While in holometabolous insects (with complete metamorphosis), AMPs are mainly synthesized in the fat body and secreted to the hemolymph, in hemimetabolous insects (with incomplete metamorphosis), AMP genes expression takes place mainly in the hemocytes upon infection [33,34].

B. germanica lives in a hostile environment in the presence of (potentially) pathogenic bacteria and in consequence displays a large arsenal of AMPs. The AMP gene repertoire of *B. germanica* (39 genes) has derived from the expansion of the five types of AMP gene families [4,35]. Comparatively, this species is among those with the largest number of AMP genes, such as the 44 of *Nasonia vitripennis* (Walker) [36] or the 31 of *Musca domestica* (Linnaeus) [37] (summarized by [38]), and much more than those reported in *P. americana* [2]. The phylogenies in *B. germanica* of the most expanded defensin and drosomycin gene families [4] revealed the presence of both identical (or almost identical) and very divergent genes (at CDS level) that may have arisen by old or recent duplications or by events of gene conversion homogenizing the CDS sequences.

The expression of all the genes involved in the IMD signaling pathway has been detected in adult *B. germanica* transcriptomic projects, with the exception of the *IKKbeta*, *IKKg* and *key* genes (Table 1). While *IKKg* and *key* genes are not found in several insect genomes, homologues of *IKKbeta* are detected in all insect genomes described in the 4IN database, except *Acyrtosiphon pisum* (Harris).

We observed that the highest levels of total AMP gene expression, among the analyzed tissues, were detected in salivary glands and hemolymph. Most genes display tissue specificity. Those with the highest expression in hemolymph were *defensin_g9*, *defensin_g10*, and *drosomycin_g5*, and those in salivary glands were *defensin_g3-g5* and *defensin_g15-16*.

Although we did not detect the expression of all AMP genes in *B. germanica* hemolymph, it must be taken into consideration that the samples come from individuals grown under controlled conditions, and that in a challenge situation with diverse pathogenic bacteria, both the levels and the diversity of AMPs produced could be increased.

The large amounts of several AMPs produced by the salivary glands imply that the control of ingested pathogens in the digestive system starts after mixing the saliva with the ingested food. This expression of AMP genes agrees with the idea that anterior parts of the digestive system encounter ingested bacteria first and might need to invoke a stronger response than distal parts [39]. In fact, our results show that in *B. germanica*, this response is not observed in the different parts of the gut (foregut, midgut or hindgut), where total AMP gene expression is very low. Moreover, some AMPs (symbiotic AMPs) are also produced by eukaryotic hosts during their symbiotic interaction with bacteria, counterintuitively not to kill but to control the symbiotic bacterial population [18,19]. Furthermore, it has been shown that AMPs could act by controlling the microbiota present in the intestine, made up mainly of mutualistic bacteria, determining its composition. In fact, it has been proposed that the immune system may have evolved initially to control interaction with this type of mutualistic bacteria in contact with epithelia, such as the intestine, and was later specialized in the fight against pathogenic bacteria [40].

In *Hydra*, the composition of the microbiota is determined by AMPs produced by the epithelial cells in contact with the microbiota [41]. In some insects, there is evidence that AMPs determine the composition and structure of the gut microbiota. This is true of *Drosophila*, where if the expression of some AMP genes is canceled, a change in the microbiota is triggered [20,42]. As compared to *Drosophila*, where some production of AMPs takes place in the midgut [17], our results demonstrate that in *B. germanica*, the majority of AMP production in the digestive system is detected in the salivary glands. Similarly, high expression of an AMP gene has also been observed in the salivary glands and midgut of *Bemisia tabaci* (Gennadius) [43], and some defensin genes are expressed in both salivary glands and midgut, and others are specific to one of these tissues in the tick *Ornithodoros turicata* (Dugès) [44]. In insects, as a general rule, AMP gene expression is silent in the absence of an immune challenge [45], although RNAseq or qRT-PCR methodologies can detect miniscule levels of gene expression. If the AMPs released by salivary glands at the top of the digestive tract were not able to control most of the ingested microbes, they would induce an immune response in other parts of the system, such as foregut, midgut or hindgut. For that reason, we cannot rule out the possibility that many genes with null or basal expression in our samples could experience a strong increase in expression.

In most insects, the fat body is an organ that can produce large amounts of AMPs during infection that are released to hemolymph [45,46]. Alternatively, in other insects, the largest production of AMPs occurs in the hemocytes, which are secreted to the hemolymph [34]. The low AMP expression levels in *B. germanica* fat bodies may increase if different types of microbes reach their cells, since in insects it is an immune-responsive tissue [45]. However, we cannot rule out the existence of mechanisms that protect the population of the endosymbiont *Blattabacterium* in bacteriocytes.

It has been proposed that AMPs are also involved in endosymbiotic associations, controlling the bacterial population and allowing metabolic complementation between the eukaryotic host and the endosymbiont. Symbiotic AMPs can affect endosymbionts through various mechanisms, including interaction with the bacterial membrane (causing disruption or permeation and allowing metabolite leakage) or targeting intracellular machinery and interfering with bacterial metabolism [19]. In the symbiosis between legume plants and nitrogen-fixing *Rhizobium* bacteria, several hundred AMPs (nodule-specific cysteine-rich peptides) allow the plant to manipulate the physiology of the bacterial symbiont [24]. In

insects, the relationship between coleopteracin A (an attacin-related AMP) and the control of a bacterial endosymbiont has been reported in weevils of the genus *Sitophilus* [21], while in the aphid *A. pisum*, some cysteine-rich AMP-like polypeptides are specifically expressed in bacteriocytes, cells containing the bacterial endosymbiont *Buchnera aphidicola* [47], and some of them display antimicrobial activities [48]. In *B. germanica*, the massive production of AMPs that could carry out this function has not been detected, although the existence of unidentified AMP-like genes, previously undetected [4], cannot be ruled out. The use of various machine-learning algorithms to predict genes with antimicrobial activity in fat body transcriptomes may serve to identify non-canonical AMP-like gene candidates involved in this process.

Finally, in adult whole bodies, attacin-like and blattelicin genes are among the most highly expressed (Figure 3). However, almost no expression of these genes was detected in our study, except a moderate expression of *attacin-like_g2* and *attacin-like_g3* in one out of the four samples of fat body, perhaps due to a contamination of surrounding tissues. The most probable place for their expression in adult bodies is the abdomen close to the heart, where the periostial hemocytes are attached and induce, after infection, the aggregation of additional hemocytes. This integrated immune and circulatory system was recently described as a general trend in insects [49]. One of the analyzed insects was *B. germanica*, in which it was detected that hemocytes were located in the dorsal tergum and almost absent in the ventral sternum [49]. However, it has recently been reported that the expression of *attacin-like_g1*, *attacin-like_g2* and *blattelicin_g1* genes in *B. germanica* gut is significantly increased after feeding with live *Salmonella enterica* serovar Typhimurium [50]. In this work, the expression of these three genes as well as *defensin_g1-g2* was also detected at very low levels in control individuals by qRT-PCR (approximate range 10^{-4} to 10^{-5}) [50]. The levels of detection by this method compared to our RNAseq analysis were similar (in both cases relative to the expression of *EF1A*), except for *attacin-like_g1*, which was not detected at all in our samples (*attacin-like_g2* in the three gut sections: range 2×10^{-3} to 1×10^{-4} ; *blattelicin_g1* in hindgut: 3×10^{-4} ; *defensin_g1-g2* in midgut: 8×10^{-5}).

In summary, as we have described, the expression of AMP genes in adult *B. germanica* is mainly located in hemolymph and salivary glands. We have also shown that although some genes are expressed in most tissues, others are tissue-specific. This differential expression in the digestive system and fat body may be related to avoiding negative effects on the beneficial microbiota and the endosymbiont, while maintaining the removal of pathogens.

4. Materials and Methods

4.1. IMD Signaling Pathway Characterization

In order to identify the genes that constitute the IMD immune signaling pathway in *B. germanica*, we performed a search for homologous sequences, using as starting data the Innate Immunity Genes in Insects (4IN) database of the University of Lyon, which is a comprehensive database of innate immunity genes in a diverse range of available insect genomes. The multifasta files obtained after the 4IN search were subjected to multiple alignment with MAFFT v7.475 [51] using the L-INS-i alignment method. Subsequently, three HMMER package 3.3.2 [52] programs were used to perform the homologous sequence search: *hmmbuild* to create HMM alignment profiles from the multiple alignments, *hmmcompress* to compress the HMM profiles into binary data, and *hmmsearch* to contrast these profiles with *B. germanica* gene databases. Sequence searches were performed on the *B. germanica* genome CDS set, which was obtained from GCA_003018175.1 (PYGN01 project [3]), and on several transcriptomes: SRR5458588.1, SRR5458589.1, SRR5458590.1, SRR5458591.1 (PRJNA382128 project) and SRR6784710 (PRJNA389591 project). All samples corresponded to the Orlando Normal from American Cyanamid strain and to whole-body sequences of females at different developmental stages (nymphs or adults), to cover various gene expression patterns and to identify as many genes as possible. This process was automated with a Python 3.9.5 pipeline (<https://docs.python.org/3/reference/>, accessed on 13 June

2022) to recursively perform the processing for each gene and for each database used. The final results were confirmed with blastx searches and modified when required.

4.2. Transcriptome Assembly

Transcriptomes used as databases in which to search for homologous sequences were assembled from Sequence Read Archive (SRA) projects. The assembling was performed de novo rather than with mapping due to the fact that the genome of *B. germanica* contains many assembly gaps. Data extraction from SRA projects was performed with fasterq-dump (NCBI SRA Toolkit v2.11.0 package). The quality of reads was checked with FastQc v0.2.3 (<https://github.com/labsquare/fastQc>, accessed on 13 June 2022). Correction of poor quality reads was performed with Rcorrector v1.0.4 [53]. Trimmomatic v0.39 [54] was used to remove residual adapters that may not have been removed in the standard sequencing protocol. Finally, Trinity v2.12.0 [55] was used to perform de novo assemblies with paired-end reads. Data processing was performed on a CentOS v7.9.2009 server.

4.3. Review of Results and CDS Gathering

While HMMER was able to detect which sequences had a higher degree of homology in *B. germanica*, it was not able to identify the proteins unequivocally. To avoid false positives and hits with incomplete sequences, the results were thoroughly checked. The CDS harbored by the individual transcripts were obtained with getorf, from the EMBOSS v6.5.7 package [56], with a minimum sequence size of 300 nucleotides. Incomplete genome sequences were filled in with data from SRA projects. The resulting sequences were aligned to the genome with Splign v2.0.1 [57] and Spidey in Unipro UGENE v37 [58]. The resulting sequences were also aligned to reviewed 4IN sequences using MAFFT. The online version of Pfam v34.0 [59] was used to contrast the Pfam domains of the resulting sequences with those of the reviewed 4IN sequences.

4.4. Insect Rearing

A population of *B. germanica* originating from a laboratory population housed by Dr. X. Bellés's group at the Institute of Evolutionary Biology (CSIC-UPF, Barcelona, Spain) was raised in plastic containers at a temperature of 25 °C and a relative humidity of 60%, in 12/12 h light/dark cycles in climatic chambers at the Institute for Integrative Systems Biology (University of Valencia-CSIC). The diet was based on dog food (Teklad Global 21% protein dog diet, 2021C, Envigo) and water provided ad libitum. Adult individuals were collected from different containers to avoid population bias, between 0 and 48 h after the transition from nymph to adult, which is considered day 0. Only females were selected to avoid bias due to sexual dimorphism, and all the sampling was carried out at day 5 of adult life.

4.5. Insect Dissection and Collection of Samples

The cockroaches were subjected to CO₂ sedation and cleaned with 10% bleach, 70% ethanol and double washed with Type II water. They were fixed with entomological needles in supine position on a silicone plate, and the following tissues were extracted: salivary glands (pool of ten individuals, four replicates), foregut (pool of five individuals, four replicates), midgut (pool of five individuals, four replicates), hindgut (pool of five individuals, four replicates), Malpighian tubules (pool of six individuals, four replicates), and fat body (pool of three individuals, four replicates). Samples were collected in previously cooled Eppendorf tubes, treated overnight with RNAlater, frozen the next morning with liquid nitrogen and stored at −80 °C until RNA extraction.

For the collection of hemolymph samples (pool of forty individuals, three replicates), individuals were sedated with CO₂ and fixed with entomological needles in supine position on a silicone plate. Subsequently, one of the legs of the third pair was cut at the beginning of the second segment. Hemolymph was collected by capillarity by massaging the abdominal

area. Hemolymph was collected directly in TRI Reagent (hemolymph to reagent ratio 1:3), and samples were frozen in liquid nitrogen and stored at -80°C until RNA extraction.

4.6. RNA Extraction and Sequencing

RNA was extracted from the tissues with the NucleoSpin RNA kit (Macherey-Nagel), while hemolymph RNA was extracted with Direct-zol RNA miniprep kit (Zymo Research), both according to the manufacturer's recommended instructions. Samples were quantified with Nanodrop and Qubit. Sequencing was carried out at the *Servei Central de Suport a la Investigació Experimental* (SCSIE, University of Valencia). RNA integrity and quality, as well as mRNA enrichment, were analyzed before proceeding with library preparation. Sequencing was performed with an Illumina NextSeq 550, resulting in single-end reads of 150 nucleotides each.

4.7. Generation of Normalized Expression Levels

To make a side-by-side comparison of the expression levels of the different AMP genes (genes with identical CDS sequences were grouped) in each of the tissues, the reads were quantified and normalized. Residual adapters were cleaned with Trimmomatic, and poor-quality bases were corrected with Rcorrector, both using standard parameters. Quantification was performed by pseudoalignment with Kallisto v0.46.2 [60], with a mean sequence size of 136 nucleotides and a standard deviation of 10 nucleotides. The index used by Kallisto was generated by adding the mRNA sequences of *B. germanica* AMPs to the database.

To be able to make both an intersample comparison and an intrasample comparison, gene length corrected trimmed mean of M-values (GeTMM) normalization was used [61]. This normalization accounts for gene length, sequencing depth and RNA composition. An estimation of GeTMM values in all CDS of the *B. germanica* genome (those annotated in the genome [3] plus the sets of AMPs and IMD pathway genes from this study) showed that, on average, only 0.44% of the genes displayed an expression equal to or higher than 10^3 , 3.42% higher than 10^2 and 19.15% higher than 10^1 . We have used these threshold values to define very high, high and moderate expression. Data processing was performed with R v4.1 and the edgeR v3.34.0 package [62]. Figures were created with ggplot2 v3.3.4 [63] and lattice v0.20-44 [64] in R. In order to compare expression levels in salivary glands and hemolymph, a differential expression analysis was performed using DESeq2 1.36.0 [65] with standard parameters, using the quantification values from Kallisto. AMP genes were considered differentially expressed if $|\log_2\text{FoldChange}| > 2$, and adjusted *p*-value cutoff was set to 0.001.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23158444/s1>.

Author Contributions: Conceptualization, C.G.-F. and F.J.S.; methodology, R.D.-S. and L.Z.; software, L.Z.; formal analysis, L.Z., C.G.-F. and F.J.S.; investigation, R.D.-S. and L.Z.; resources, R.D.-S.; data curation, L.Z.; writing—original draft preparation, C.G.-F. and F.J.S.; writing—review and editing, L.Z., R.D.-S., C.G.-F. and F.J.S.; visualization, L.Z.; supervision, C.G.-F. and F.J.S.; funding acquisition, C.G.-F. and F.J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the European Regional Development Fund (ERDF) and the Ministerio de Ciencia, Innovación y Universidades (Spain) (PGC2018-099344-B-I00) and Generalitat Valenciana (Prometeo/2018/A/133 and CIPROM/2021/042).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB52531.

Acknowledgments: We would like to thank Mariano Collantes-Alegre for providing the drawing of the gut for the graphical abstract.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bell, W.J.; Roth, L.M.; Nalepa, C.A. *Cockroaches: Ecology, Behavior, and Natural History*; Johns Hopkins University Press: Baltimore, MD, USA, 2007.
2. Li, S.; Zhu, S.; Jia, Q.; Yuan, D.; Ren, C.; Li, K.; Liu, S.; Cui, Y.; Zhao, H.; Cao, Y.; et al. The genomic and functional landscapes of developmental plasticity in the American cockroach. *Nat. Commun.* **2018**, *9*, 1008. [[CrossRef](#)] [[PubMed](#)]
3. Harrison, M.C.; Jongepier, E.; Robertson, H.M.; Arning, N.; Bitard-Feidel, T.; Chao, H.; Childers, C.P.; Dinh, H.; Doddapaneni, H.; Dugan, S.; et al. Hemimetabolous genomes reveal molecular basis of termite eusociality. *Nat. Ecol. Evol.* **2018**, *2*, 557–566. [[CrossRef](#)] [[PubMed](#)]
4. Silva, F.J.; Muñoz-Benavent, M.; García-Ferris, C.; Latorre, A. *Blattella germanica* displays a large arsenal of antimicrobial peptide genes. *Sci. Rep.* **2020**, *10*, 21058. [[CrossRef](#)] [[PubMed](#)]
5. Latorre, A.; Domínguez-Santos, R.; García-Ferris, C.; Gil, R. Of cockroaches and symbionts: Recent advances in the characterization of the relationship between *Blattella germanica* and its dual symbiotic system. *Life* **2022**, *12*, 290. [[CrossRef](#)] [[PubMed](#)]
6. Bandi, C.; Sironi, M.; Damiani, G.; Magrassi, L.; Nalepa, C.; Laudani, U.; Sacchi, L. The establishment of intracellular symbiosis in an ancestor of cockroaches and termites. *Proc. R. Soc. Lond. B* **1995**, *259*, 293–299. [[CrossRef](#)]
7. Evangelista, D.A.; Wipfler, B.; Béthoux, O.; Donath, A.; Fujita, M.; Kohli, M.K.; Legendre, F.; Liu, S.; Machida, R.; Misof, B.; et al. An integrative phylogenomic approach illuminates the evolutionary history of cockroaches and termites (Blattodea). *Proc. R. Soc. B* **2019**, *286*, 20182076. [[CrossRef](#)]
8. López-Sánchez, M.J.; Neef, A.; Peretó, J.; Patiño-Navarrete, R.; Pignatelli, M.; Latorre, A.; Moya, A. Evolutionary convergence and nitrogen metabolism in *Blattabacterium* strain Bge, primary endosymbiont of the cockroach *Blattella germanica*. *PLoS Genet.* **2009**, *5*, e1000721. [[CrossRef](#)]
9. Sabree, Z.L.; Kambhampati, S.; Moran, N.A. Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 19521–19526. [[CrossRef](#)]
10. Neef, A.; Latorre, A.; Peretó, J.; Silva, F.J.; Pignatelli, M.; Moya, A. Genome economization in the endosymbiont of the wood roach *Cryptocercus punctulatus* due to drastic loss of amino acid synthesis capabilities. *Genome Biol. Evol.* **2011**, *3*, 1437–1448. [[CrossRef](#)]
11. Sabree, Z.L.; Huang, C.Y.; Arakawa, G.; Tokuda, G.; Lo, N.; Watanabe, H.; Moran, N.A. Genome shrinkage and loss of nutrient-providing potential in the obligate symbiont of the primitive termite *Mastotermes darwiniensis*. *Appl. Environ. Microbiol.* **2012**, *78*, 204–210. [[CrossRef](#)]
12. Patiño-Navarrete, R.; Moya, A.; Latorre, A.; Peretó, J. Comparative genomics of *Blattabacterium cuenoti*: The frozen legacy of an ancient endosymbiont genome. *Genome Biol. Evol.* **2013**, *5*, 351–361. [[CrossRef](#)] [[PubMed](#)]
13. González-Domenech, C.M.; Belda, E.; Patiño-Navarrete, R.; Moya, A.; Peretó, J.; Latorre, A. Metabolic stasis in an ancient symbiosis: Genome-scale metabolic networks from two *Blattabacterium cuenoti* strains, primary endosymbionts of cockroaches. *BMC Microbiol.* **2012**, *12*, S5. [[CrossRef](#)] [[PubMed](#)]
14. Patiño-Navarrete, R.; Piulachs, M.D.; Bellés, X.; Moya, A.; Latorre, A.; Peretó, J. The cockroach *Blattella germanica* obtains nitrogen from uric acid through a metabolic pathway shared with its bacterial endosymbiont. *Biol. Lett.* **2014**, *10*, 20140407. [[CrossRef](#)] [[PubMed](#)]
15. Moran, N.A.; Ochman, H.; Hammer, T.J. Evolutionary and ecological consequences of gut microbial communities. *Annu. Rev. Ecol. Evol. Syst.* **2019**, *50*, 451–475. [[CrossRef](#)] [[PubMed](#)]
16. Nishide, Y.; Kageyama, D.; Yokoi, K.; Jouraku, A.; Tanaka, H.; Futahashi, R.; Fukatsu, T. Functional crosstalk across IMD and Toll pathways: Insight into the evolution of incomplete immune cascades. *Proc. R. Soc. B* **2019**, *286*, 20182207. [[CrossRef](#)] [[PubMed](#)]
17. Hanson, M.A.; Lemaitre, B. New insights on *Drosophila* antimicrobial peptide function in host defense and beyond. *Curr. Opin. Immunol.* **2020**, *62*, 22–30. [[CrossRef](#)]
18. Mergaert, P.; Kikuchi, Y.; Shigenobu, S.; Nowack, E.C.M. Metabolic integration of bacterial endosymbionts through antimicrobial peptides. *Trends Microbiol.* **2017**, *25*, 703–712. [[CrossRef](#)]
19. Mergaert, P. Role of antimicrobial peptides in controlling symbiotic bacterial populations. *Nat. Prod. Rep.* **2018**, *35*, 336–356. [[CrossRef](#)] [[PubMed](#)]
20. Marra, A.; Hanson, M.A.; Kondo, S.; Erkosar, B.; Lemaitre, B. *Drosophila* antimicrobial peptides and lysozymes regulate gut microbiota composition and abundance. *mBio* **2021**, *12*, e0082421. [[CrossRef](#)]
21. Login, F.H.; Balmand, S.; Vallier, A.; Vincent-Monégat, C.; Vigneron, A.; Weiss-Gayet, M.; Rochat, D.; Heddi, A. Antimicrobial peptides keep insect endosymbionts under control. *Science* **2011**, *334*, 362–365. [[CrossRef](#)] [[PubMed](#)]
22. Oakeson, K.F.; Gil, R.; Clayton, A.L.; Dunn, D.M.; von Niederhausern, A.C.; Hamil, C.; Aoyagi, A.; Duval, B.; Baca, A.; Silva, F.J.; et al. Genome degeneration and adaptation in a nascent stage of symbiosis. *Genome Biol. Evol.* **2014**, *6*, 76–93. [[CrossRef](#)] [[PubMed](#)]
23. Masson, F.; Zaidman-Rémy, A.; Heddi, A. Antimicrobial peptides and cell processes tracking endosymbiont dynamics. *Phil. Trans. R. Soc. B* **2016**, *371*, 20150298. [[CrossRef](#)] [[PubMed](#)]

24. Kondorosi, E.; Mergaert, P.; Kereszt, A. A paradigm for endosymbiotic life: Cell differentiation of *Rhizobium* bacteria provoked by host plant factors. *Ann. Rev. Microbiol.* **2013**, *67*, 611–628. [[CrossRef](#)] [[PubMed](#)]
25. Carro, L.; Pujic, P.; Alloisio, N.; Fournier, P.; Boubakri, H.; Hay, A.E.; Poly, F.; François, P.; Hocher, V.; Mergaert, P.; et al. *Alnus* peptides modify membrane porosity and induce the release of nitrogen-rich metabolites from nitrogen-fixing *Frankia*. *ISME J.* **2015**, *9*, 1723–1733. [[CrossRef](#)] [[PubMed](#)]
26. Kleino, A.; Silverman, N. The *Drosophila* IMD pathway in the activation of the humoral immune response. *Dev. Comp. Immunol.* **2014**, *42*, 25–35. [[CrossRef](#)]
27. Ryu, J.-H.; Kim, S.-H.; Lee, H.-Y.; Bai, J.Y.; Nam, Y.-D.; Bae, J.-W.; Lee, D.G.; Shin, S.C.; Ha, E.-M.; Lee, W.-J. Innate immune homeostasis by the homeobox gene *caudal* and commensal-gut mutualism in *Drosophila*. *Science* **2008**, *319*, 777–782. [[CrossRef](#)]
28. Park, M.S.; Park, P.; Takeda, M. Starvation induces apoptosis in the midgut nidi of *Periplaneta americana*: A histochemical and ultrastructural study. *Cell Tissue Res.* **2009**, *335*, 631–638. [[CrossRef](#)]
29. Huang, X.; Warren, J.T.; Buchanan, J.; Gilbert, L.I.; Scott, M.P. *Drosophila* Niemann-Pick Type C-2 genes control sterol homeostasis and steroid biosynthesis: A model of human neurodegenerative disease. *Development* **2007**, *134*, 3733–3742. [[CrossRef](#)]
30. Erkosar, B.; Defaye, A.; Bozonnet, N.; Puthier, D.; Royet, J.; Leulier, F. *Drosophila* microbiota modulates host metabolic gene expression via IMD/NF- κ b signaling. *PLoS ONE* **2014**, *9*, e94729. [[CrossRef](#)]
31. Buonocore, F.; Fausto, A.M.; Della Pelle, G.; Roncevic, T.; Gerdol, M.; Picchiatti, S. Attacins: A promising class of insect antimicrobial peptides. *Antibiotics* **2021**, *10*, 212. [[CrossRef](#)]
32. Tanji, T.; Hu, X.; Weber, A.N.; Ip, Y.T. Toll and IMD pathways synergistically activate an innate immune response in *Drosophila melanogaster*. *Mol. Cell. Biol.* **2007**, *27*, 4578–4588. [[CrossRef](#)] [[PubMed](#)]
33. Lamberty, M.; Zachary, D.; Lanot, R.; Bordereau, C.; Robert, A.; Hoffmann, J.A.; Bulet, P. Insect immunity. Constitutive expression of a cysteine-rich antifungal and a linear antibacterial peptide in a termite insect. *J. Biol. Chem.* **2001**, *276*, 4085–4092. [[CrossRef](#)] [[PubMed](#)]
34. Bulet, P.; Stöcklin, R. Insect antimicrobial peptides: Structures, properties and gene regulation. *Protein Pept. Lett.* **2005**, *12*, 3–11. [[CrossRef](#)]
35. Harrison, M.C.; Arning, N.; Kremer, L.P.M.; Ylla, G.; Belles, X.; Bornberg-Bauer, E.; Huylmans, A.K.; Jongepier, E.; Piulachs, M.D.; Richards, S.; et al. Expansions of key protein families in the German cockroach highlight the molecular basis of its remarkable success as a global indoor pest. *J. Exp. Zool. B Mol. Dev. Evol.* **2018**, *330*, 254–264. [[CrossRef](#)] [[PubMed](#)]
36. Sackton, T.B.; Werren, J.H.; Clark, A.G. Characterizing the infection-induced transcriptome of *Nasonia vitripennis* reveals a preponderance of taxonomically-restricted immune genes. *PLoS ONE* **2013**, *8*, e83984. [[CrossRef](#)]
37. Sackton, T.B.; Lazzaro, B.P.; Clark, A.G. Rapid expansion of immune-related gene families in the house fly, *Musca domestica*. *Mol. Biol. Evol.* **2017**, *34*, 857–872. [[CrossRef](#)] [[PubMed](#)]
38. Sackton, T. Comparative genomics and transcriptomics of host–pathogen interactions in insects: Evolutionary insights and future directions. *Curr. Opin. Insect Sci.* **2019**, *31*, 106–113. [[CrossRef](#)]
39. Engel, P.; Moran, N.A. The gut microbiota of insects—Diversity in structure and function. *FEMS Microbiol. Rev.* **2013**, *37*, 699–735. [[CrossRef](#)]
40. Lee, Y.K.; Mazmanian, S.K. Has the microbiota played a critical role in the evolution of the adaptive immune system? *Science* **2010**, *330*, 1768–1773. [[CrossRef](#)]
41. Fraune, S.; Augustin, R.; Anton-Erxleben, F.; Wittlieb, J.; Gelhaus, C.; Klimovich, V.B.; Samoilovich, M.P.; Bosch, T.C. In an early branching metazoan, bacterial colonization of the embryo is controlled by maternal antimicrobial peptides. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 18067–18072. [[CrossRef](#)]
42. Bosch, T.C.G.; Zasloff, M. Antimicrobial peptides—Or how our ancestors learned to control the microbiome. *mBio* **2021**, *12*, e01847-21. [[CrossRef](#)] [[PubMed](#)]
43. Wang, Z.-Z.; Shi, M.; Ye, X.-Q.; Chen, M.-Y.; Chen, X.-X. Identification, characterization and expression of a defensin-like antifungal peptide from the whitefly *Bemisia tabaci* (Gennadius) (Hemiptera: Aleyrodidae). *Insect Mol. Biol.* **2013**, *22*, 297–305. [[CrossRef](#)] [[PubMed](#)]
44. Armstrong, B.A.; Kneubehl, A.R.; Mitchell, R.D.; Krishnavajhala, A.; Teel, P.D.; Pérez de León, A.A.; Lopez, J.E. Differential expression of putative *Ornithodoros turicata* defensins mediated by tick feeding. *Front. Cell. Infect. Microbiol.* **2020**, *10*, 152. [[CrossRef](#)] [[PubMed](#)]
45. Hoffmann, J.A. Innate immunity of insects. *Curr. Opin. Immunol.* **1995**, *7*, 4–10. [[CrossRef](#)]
46. Buchon, N.; Silverman, N.; Cherry, S. Immunity in *Drosophila melanogaster*—From microbial recognition to whole-organism physiology. *Nat. Rev. Immunol.* **2014**, *14*, 796–810. [[CrossRef](#)] [[PubMed](#)]
47. Shigenobu, S.; Stern, D.L. Aphids evolved novel secreted proteins for symbiosis with bacterial endosymbiont. *Proc. R. Soc. B* **2013**, *280*, 20121952. [[CrossRef](#)] [[PubMed](#)]
48. Uchi, N.; Fukudome, M.; Nozaki, N.; Suzuki, M.; Osuki, K.I.; Shigenobu, S.; Uchiumi, T. Antimicrobial activities of cysteine-rich peptides specific to bacteriocytes of the pea aphid *Acyrtosiphon pisum*. *Microbes Environ.* **2019**, *34*, 155–160. [[CrossRef](#)] [[PubMed](#)]
49. Yan, Y.; Hillyer, J. The immune and circulatory systems are functionally integrated across insect evolution. *Sci. Adv.* **2020**, *6*, eabb3164. [[CrossRef](#)]
50. Turner, M.; Pietri, J.E. Antimicrobial peptide expression in the cockroach gut during enterobacterial infection is specific and influenced by type III secretion. *Biol. Open.* **2022**, *11*, bio059414. [[CrossRef](#)] [[PubMed](#)]

51. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
52. Eddy, S.R. A new generation of homology search tools based on probabilistic inference. *Genome Inform.* **2009**, *23*, 205–211. [[CrossRef](#)] [[PubMed](#)]
53. Song, L.; Florea, L. Rcorrector: Efficient and accurate error correction for Illumina RNA-seq reads. *GigaScience* **2015**, *4*, 48. [[CrossRef](#)] [[PubMed](#)]
54. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
55. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q.; et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **2011**, *29*, 644–652. [[CrossRef](#)]
56. Rice, P.; Longden, I.; Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **2000**, *16*, 276–277. [[CrossRef](#)] [[PubMed](#)]
57. Kapustin, Y.; Souvorov, A.; Tatusova, T.; Lipman, D. Splign: Algorithms for computing spliced alignments with identification of paralogs. *Biol. Direct* **2008**, *3*, 20. [[CrossRef](#)]
58. Okonechnikov, K.; Golosova, O.; Fursov, M.; UGENE team. Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **2012**, *28*, 1166–1167. [[CrossRef](#)]
59. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.A.; Sonnhammer, E.L.L.; Tosatto, S.C.E.; Paladin, L.; Raj, S.; Richardson, L.J.; et al. Pfam: The protein families database in 2021. *Nucleic Acids Res.* **2021**, *49*, D412–D419. [[CrossRef](#)] [[PubMed](#)]
60. Bray, N.L.; Pimentel, H.; Melsted, P.; Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **2016**, *34*, 525–527. [[CrossRef](#)] [[PubMed](#)]
61. Smid, M.; Coebergh van den Braak, R.R.J.; van de Werken, H.J.G.; van Riet, J.; van Galen, A.; de Weerd, V.; van der Vlugt-Daane, M.; Bril, S.I.; Lalmahomed, Z.S.; Kloosterman, W.P.; et al. Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. *BMC Bioinform.* **2018**, *19*, 236. [[CrossRef](#)]
62. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [[CrossRef](#)] [[PubMed](#)]
63. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016.
64. Sarkar, D. *Lattice: Multivariate Data Visualization with R*; Springer: New York, NY, USA, 2008.
65. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [[CrossRef](#)] [[PubMed](#)]