BAYERISCHE JULIUS-MAXIMILIANS-UNIVERSITÄT WÜRZBURG

FAKULTÄT FÜR BIOLOGIE

LEHRSTUHL FÜR MIKROBIOLOGIE

Immune responses of the ant *Camponotus floridanus* towards pathogens and its obligate mutualistic endosymbiont *Blochmannia floridanus*

DISSERTATION

zur Erlangung des naturwissenschaftlichen Doktorgrades

der bayerischen Julius-Maximilians-Universität Würzburg

vorgelegt von

Carolin Ratzka

aus Regensburg

Würzburg, Oktober 2012

Erklärungen (gem. § 4 Abs. 3 S. 3, 5 und 8 der Promotionsordnung)

Hiermit erkläre ich, dass die vorliegende Arbeit von mir selbständig und nur unter der Verwendung der angegebenen Quellen und Hilfsmittel angefertigt wurde.

Diese Dissertation hat weder in gleicher noch in ähnlicher Form in einem anderen Prüfungsverfahren vorgelegen.

Ich habe früher, außer den mit dem Zulassungsgesuch urkundlich vorgelegten Graden, keine weiteren akademischen Grade erworben oder zu erwerben versucht.

Carolin Ratzka

Würzburg im Oktober 2012

Eingereicht am:

Mitglieder der Promotionskommission:

Vorsitzender:

Erstgutachter: Prof. Dr. Roy Gross

Zweitgutachter: Prof. Dr. Heike Feldhaar

Tag des Promotionskolloquiums:

Doktorurkunde ausgehändigt am:

Danksagung

An dieser Stelle möchte ich mich bei all jenen Menschen bedanken, die durch ihre Unterstützung zum Gelingen dieser Arbeit beigetragen haben.

Zu allererst gilt mein Dank Herrn Prof. Dr. Roy Gross für die interessante Aufgabenstellung und die nötige Freiheit bei der Ausgestaltung dieser Arbeit sowie für die vielfältige Unterstützung durch anregende Diskussionen und beratende Gespräche. Ebenso danke ich Frau Prof. Dr. Heike Feldhaar für die freundliche Übernahme des Zweitgutachtens dieser Arbeit sowie für die vielen wertvollen Diskussionen und für die gute Zusammenarbeit.

Herrn Prof. Dr. Thomas Rudel danke ich für die Bereitstellung des Arbeitsplatzes am Lehrstuhl für Mikrobiologie der Universität Würzburg. Dem SFB 567 der DFG sowie dem „Programm Chancengleichheit für Frauen in Forschung und Lehre“ des bayerischen Staates danke ich für die finanzielle Unterstützung dieser Arbeit.

Mein Dank gilt auch Herrn Prof. Dr. Thomas Dandekar für die gute Kooperation mit dem Lehrstuhl für Bioinformatik der Universität Würzburg. Herrn Dr. Frank Förster danke ich für die bioinformatische Auswertung der Transkriptomdaten sowie für die phylogenetische Analyse der antimikrobiellen Peptide. Ferner danke ich Herrn Dr. Chunguang Liang für die Hilfe bei der bioinformatischen Analyse der Hymenoptaecin-Sequenzen sowie der differentiell exprimierten Sequenzen aus der SSH.

Herrn Prof. Dr. Andreas Vilcinskas danke ich für die gute Kooperation sowie für die Möglichkeit, die Anwendung der RNA-Interferenz- (RNAi-) Methode am Institut für Phytopathologie und Angewandte Zoologie der Universität Gießen zu erlernen. Frau Dr. Eileen Knorr danke ich für die Hilfe bei der RNAi-Methode sowie für die schöne Zeit in Gießen.

Vielen Dank an den Lehrstuhl für Verhaltensphysiologie und Soziobiologie der Universität Würzburg, insbesondere an Frau Annette Laudahn, für die Pflege der Ameisenkolonien sowie für ihre Hilfsbereitschaft.

Des Weiteren möchte ich mich bei allen Mitgliedern des Lehrstuhls für Mikrobiologie für die angenehme Arbeitsatmosphäre sowie für die hilfreichen fachlichen Diskussionen und Ratschläge im Laufe meiner Doktorarbeit bedanken. Meiner Laborkollegin Manuela Geier danke ich für die unterhaltsame Laboratmosphäre, ihre Hilfsbereitschaft sowie für die schöne Zeit auch außerhalb des Labors. Bei Maria Kupper, die die Immunitätsforschung bei Rossameisen weiterführen wird, bedanke ich mich für das Korrekturlesen dieser Arbeit sowie für die vielen freundschaftlichen Gespräche und Unternehmungen. Meinem Vorgänger Sascha Stoll danke ich für die hilfreiche Einführung ins molekularbiologische Arbeiten zu Beginn meiner Doktorarbeit. Viele meiner Kollegen sind meine Freunde geworden, und ich möchte mich bei Birgit Bergmann, Kristina Keidel, Frederike Fritsch, Norman Mauder, Angela Baljuls, Jo-Ana Herweg und Claudia Sibilski nicht nur für die oben erwähnte Hilfe,

sondern auch für die netten Abende und gemeinsamen Unternehmungen bedanken. Außerdem danke ich Dagmar Beier für die schöne Zeit auf der Reitanlage Prosselsheim.

Allen meinen Freunden, insbesondere Natalie Dreeßen und Christine Zellmann, möchte ich an dieser Stelle für ihr Verständnis, ihre aufmunternden Worte und die erholsamen gemeinsamen Freizeitaktivitäten danken. Meinem Freund Christian Scherer danke ich für seine liebevolle Unterstützung sowie für die vielen lustigen Unternehmungen, die mir geholfen haben abzuschalten und neue Kraft zu schöpfen.

Zu guter Letzt möchte ich meiner gesamten Familie danken. Meinem Bruder Michael Ratzka danke ich für das Korrekturlesen dieser Arbeit und die Bereitstellung zahlreicher Serien und Filme. Meiner Schwester Christiane Diederich und ihrem Mann Michael Diederich danke ich für die erholsamen Kurzurlaube in Landau. Mein besonderer Dank gilt meinen Eltern Kurt und Eveline Ratzka, die mich stets uneingeschränkt unterstützt und an mich geglaubt haben. Danke für alles, was ihr für mich getan habt - ohne euch wäre ich wohl kaum so weit gekommen, wie ich heute bin!

„A curious ant lays down a trail for new findings."

(Christian Scherer)

Table of contents

1 Summary

The ant species *Camponotus floridanus* belongs to one of the most successful ant genera in terms of species richness as well as distribution (Bolton 1995). These animals live in huge colonies composed of genetically identical or closely related animals, which should predispose them to an increased vulnerability towards infection by pathogens (Cremer et al. 2007). Therefore the question is how ants (or social insects in general) can nevertheless efficiently combat infections.

In order to investigate the immune response of the ant *C. floridanus*, the present study initially focused on the identification of possible immune factors, encoded by the ant´s genome. By using the method "suppression subtractive hybridization" as well as by Illumnia sequencing technology, several genes could be identified, which were transcriptionally induced after bacterial challenge and thus should be immune-related. Among these were genes encoding proteins involved in pathogen recognition (e.g. a Gram-negative binding protein), signal transduction (e.g. a mitogen-activated protein-kinase), antimicrobial activity (e.g. the antimicrobial peptides defensin and hymenoptaecin), or general stress response (e.g. a heat shock protein). The exact function of single immune factors should be investigated in prospective studies by using RNA interference technology. For this purpose, a protocol for knockdown of gene expression by feeding of double stranded RNA (dsRNA) to worker ants was established in the context of this thesis.

In accordance with the ant´s genome sequence (Bonasio et al. 2010), only three antimicrobial peptide (AMP) genes could be identified in *C. floridanus*, in contrast to about 20 AMP genes in *Drosophila melanogaster* (Imler and Bulet 2005). The gene and cDNA sequences of these AMPs were established and their expression was shown to be induced by microbial challenge. Two different *defensin* genes were characterized. The *defensin-2* gene has a single intron, whereas the *defensin-1* gene has two introns. The deduced amino acid sequence of the *C. floridanus* defensins is very similar to other known ant defensins with the exception of a short C-terminal extension of defensin-1. A detailed characterization of the mRNA and gene sequence of the other AMP, a hymenoptaecin, revealed a special repeat structure which resembles the multipeptide precursor structure from the *Apis mellifera* apidaecin (Casteels-Josson et al. 1993). The *C. floridanus* hymenoptaecin has a signal and a pro-sequence followed by a hymenoptaecin-like domain and six directly repeated hymenoptaecin domains. Each of the hymenoptaecin domains is flanked by an EAEP-spacer sequence and a RR-site known to be a proteolytic processing site. Thus, proteolytic processing of the precursor protein may generate several mature AMPs leading to an amplification of the immune response and thereby to a massive increase in immune competence of the animals despite having an extremely reduced repertoire of AMP genes. Bioinformatical analyses revealed the presence of *hymenoptaecin* genes with similar multipeptide precursor structure in genomes of other ant species suggesting an evolutionary conserved important role of this gene in ant immunity.

C. floridanus ants harbor the obligate intracellular bacterium, *Blochmannia floridanus*, in specialized cells (so-called bacteriocytes), which are intercalated between midgut cells as well as in ovaries of females (Blochmann 1882; Sauer et al. 2002; Schröder et al. 1996). Previous studies showed that these symbiotic bacteria can supplement the nutrition of their hosts with essential amino acids and also seem to be involved in nitrogen recycling via an urease (Feldhaar et al. 2007). Thus, the ant species *C. floridanus* probably gains a significant fitness advantage through the symbiosis with *B. floridanus* in comparison to other ant species lacking such an endosymbiont. However, for the ant hosts this also raises the problem that on the one hand they have to maintain the beneficial symbiotic bacteria and on the other hand they need to raise an immune response against harmful pathogenic bacteria during an infection. This leads to the question, if the host immune system can discriminate between symbionts and potential pathogens. Therefore, it was investigated, if *B. floridanus* is still recognized as „non-self" by the ant´s immune system. A quantitative analysis of immune gene expression revealed different expression kinetics of individual factors and characteristic expression profiles after injection of different bacterial species (Gram-negative and Gram-positive bacteria). Injection of the own endosymbiont *B. floridanus* also induced an immune response of its host *C. floridanus*, which was comparable to the one towards pathogens. This means that, despite the evolutionary established cooperation of the endosymbionts and their hosts, these bacteria are still recognized as „non-self" by the host immune system.

This finding led to the question, if the ant immune system might be involved in regulation of the endosymbiont number in the midgut tissue in order to avoid their uncontrolled replication. During the holometabolous life cycle of the ant hosts the distribution of bacteriocytes and of *Blochmannia* endosymbionts is remarkably dynamic and peaks in late pupal stages, in which the entire midgut is transformed into a symbiotic organ (Stoll et al. 2010). It was hypothesized that hosts could regulate the number of endosymbionts present in their tissues via the innate immune system. A quantitative gene expression analysis of assumed symbiosis-relevant candidate genes revealed distinct expression patterns of some genes according to developmental stage and tissue. Interestingly, two peptidoglycan recognition protein (PGRP) genes, *PGRP-LB* and *PGRP-SC2*, reached high expression levels in the midgut tissue at the late pupal stages, particularly at the time, when the highest number of *Blochmannia* is present in this tissue. The encoded proteins are supposed to exhibit an amidase activity and studies on other insects suggest that they downregulate the immune response by cleavage of bacterial peptidoglycan into non-immuno-stimulatory fragments (Anselme et al. 2006; Bischoff et al. 2006; Zaidman-Remy et al. 2006). In accordance with this result the transcription of immune factors, like *hymenoptaecin*, was significantly less induced in the midgut tissue than in the residual body parts after an immune-challenge in the pupal stage, whereas *PGRP-LB* expression was simultaneously induced in the midgut tissue. These data suggested that amidase PGRPs, like PGRP-LB and -SC, might facilitate the toleration of the high endosymbiont number present at the pupal stage in the midgut tissue. Moreover, high expression of host genes encoding an invertebrate-type lysozyme and a serine

carboxypeptidase vitellogenic-like (CPVL) in the pupal midgut tissue, suggests lysosomal degradation of *Blochmannia* endosymbionts possibly to supply the ant host with essential nutrients, as already suggested for aphids (Nishikori et al. 2009a; Nishikori et al. 2009b).

In sum, this thesis provides a first description of the immune response of the ant *C. floridanus*. A comprehensive set of immune-relevant genes was determined. Especially, the identification and molecular characterization of the *hymenoptaecin* gene delivered new insights into the immune competence of ants in general. Moreover, first indications could be gathered for the involvement of the immune system in controlling the endosymbiont *B. floridanus*. Thus, a broad basis was established for prospective detailed analyses of the connection between immunity and symbiosis in *C. floridanus*.

2 Zusammenfassung

Die Ameisenart *Camponotus floridanus* gehört zu einer der erfolgreichsten Ameisengattungen hinsichtlich Artenreichtum und Verbreitung (Bolton 1995). Diese Tiere leben in großen Kolonien, welche sich aus genetisch identischen oder nahe verwandten Individuen zusammensetzen. Demnach sollten diese Tiere eine erhöhte Anfälligkeit gegenüber Infektionen durch Pathogene haben (Cremer et al. 2007). Somit stellt sich die Frage, wie Ameisen (oder allgemein soziale Insekten) Infektionen dennoch effizient bekämpfen können.

Um die Immunantwort der Ameise *C. floridanus* zu untersuchen, befasste sich die vorliegende Arbeit zunächst mit der Identifizierung von möglichen Immunfaktoren, welche im Genom der Ameise kodiert sind. Unter Verwendung der Methode "Suppression Subtractive Hybridization" sowie durch Illumnia Sequenzierungstechnologie konnten mehrere Gene ermittelt werden, welche nach bakterieller Anregung transkriptionell induziert wurden und demnach mit dem Immunsystem in Beziehung stehen sollten. Darunter befanden sich Gene, welche Proteine kodieren, die eine Rolle bei der Erkennung von Pathogenen (z.B. ein Protein zur Erkennung von Gram-negativen Bakterien), der Signalübertragung (z.B. eine mitogen-aktivierte Proteinkinase-Kinase), der antimikrobiellen Aktivität (z.B. die antimikrobiellen Peptide Defensin und Hymenoptaecin) oder der allgemeinen Stressantwort (z.B. ein Hitzeschockprotein) spielen. Die genaue Funktion von einzelnen Faktoren könnte in zukünftigen Studien unter Verwendung der RNA-Interferenz-Technologie untersucht werden. Zu diesem Zweck wurde im Rahmen dieser Doktorarbeit bereits ein Protokoll für den Knockdown der Genexpression durch Verfütterung von doppelsträngiger RNA an Arbeiterinnen etabliert.

In Übereinstimmung mit der Genomsequenz der Ameise (Bonasio et al. 2010), konnten nur drei antimikrobielle Peptid (AMP) Gene in *C. floridanus* identifiziert werden, im Gegensatz zu ungefähr 20 AMP Genen in *Drosophila melanogaster* (Imler and Bulet 2005). Die Gen- und cDNA-Sequenzen dieser AMPs wurden charakterisiert und es wurde gezeigt, dass ihre Expression durch mikrobielle Attacken induziert wird. Zwei verschiedene *Defensin* Gene wurden charakterisiert. Das *Defensin-2* Gen hat ein einziges Intron, wohingegen das *Defensin-1* Gen zwei Introns besitzt. Die abgeleitete Aminosäuresequenz der Defensine von *C. floridanus* ist sehr ähnlich zu anderen bekannten Defensinen von Ameisen mit Ausnahme einer kurzen C-terminalen Verlängerung des Defensin-1. Eine detaillierte Charakterisierung der mRNA- und Gen-Sequenz des anderen AMPs, eines Hymenoptaecins, ergab eine besondere Wiederholungsstruktur, welche der Multipeptid-Vorläufer-Struktur des Apidaecins von *Apis mellifera* ähnelt (Casteels-Josson et al. 1993). Das *C. floridanus* Hymenoptaecin hat eine Signal- und eine Pro-Sequenz gefolgt von einer Hymenoptaecin-ähnlichen Domäne und sechs direkt wiederholten Hymenoptaecin Domänen. Jede der Hymenoptaecin-Domänen wird flankiert von einer EAEP-Abstandssequenz und einer RR-Sequenz, welche eine bekannte proteolytische Schnittstelle ist. Demnach könnte die proteolytische Bearbeitung des Vorläufer Proteins mehrere reife AMPs hervorbringen. Dies würde zu einer Verstärkung der

Immunantwort führen und dadurch zu einem massiven Anstieg der Immun-Kompetenz dieser Tiere, obwohl sie ein stark reduziertes Repertoire an AMP Genen besitzen. Bioinformatische Analysen enthüllten die Anwesenheit von *Hymenoptaecin* Genen mit ähnlicher Multipeptid-Vorläufer-Struktur in den Genomen von anderen Ameisenarten, was eine evolutionär konservierte wichtige Aufgabe dieses Gens bei der Immunität von Ameisen andeutet.

C. floridanus Ameisen beherbergen das obligat intrazelluläre Bakterium, *Blochmannia floridanus*, in speziellen Zellen (den sogenannten Bakteriozyten), welche sich zwischen Zellen des Mitteldarms befinden sowie in den Ovarien von Weibchen (Blochmann 1882; Sauer et al. 2002; Schröder et al. 1996). Frühere Studien zeigten, dass diese symbiontischen Bakterien die Nahrung ihrer Wirte mit essentiellen Aminosäuren ergänzen können und auch eine Rolle beim Stickstoff-Recycling mittels einer Urease spielen könnten (Feldhaar et al. 2007). Demnach erlangt die Ameisenart *C. floridanus* durch die Symbiose zu *B. floridanus* vermutlich einen Fitnessvorteil gegenüber anderen Ameisenarten ohne diesen Endosymbionten. Für die Wirtstiere ergibt sich damit aber auch das Problem, dass sie einerseits die für sie nützlichen symbiontischen Bakterien erhalten müssen und andererseits bei einer Infektion eine Immunantwort gegenüber schädlichen pathogenen Bakterien aufbringen müssen. Dies führt zu der Frage, ob das Wirtsimmunsystem zwischen symbiontischen und potentiell pathogenen Bakterien unterscheiden kann. Daher wurde untersucht, ob *B. floridanus* noch als „nicht-selbst“ vom Immunsystem der Ameise erkannt wird. Eine quantitative Analyse der Immungenexpression ergab eine unterschiedliche Expressionskinetik von einzelnen Faktoren und charakteristische Expressionsprofile nach der Injektion von unterschiedlichen Arten von Bakterien (Gram-negative und Gram-positive Bakterien). Injektion des eigenen Endosymbionten *B. floridanus* induzierte ebenfalls eine Immunantwort seines Wirtes *C. floridanus*, welche vergleichbar war mit der gegenüber Pathogenen. Dies bedeutet, dass trotz der Koevolution zwischen den Endosymbionten und ihren Wirte, diese Bakterien immer noch als „nicht-selbst“ vom Wirtsimmunsystem erkannt werden.

Dieses Ergebnis warf die Frage auf, ob das Ameisen-Immunsystem an der Regulation der Anzahl von Endosymbionten im Mitteldarmgewebe beteiligt sein könnte, um deren unkontrollierte Replikation zu vermeiden. Während des holometabolen Lebenszyklus der Ameisen-Wirte ist die Verteilung der Bakteriozyten und der *Blochmannia* Endosymbionten bemerkenswert dynamisch und erreicht den Höhepunkt in den späten Puppenstadien, in welchen der gesamte Mitteldarm zu einem symbiotischen Organ umgewandelt wird (Stoll et al. 2010). Es wurde vermutet, dass die Wirte die vorhandene Anzahl von Endosymbionten in ihren Geweben durch das angeborene Immunsystem regulieren könnten. Eine quantitative Genexpressionsanalyse von vermeintlich Symbiose-relevanten Kandidatengenen ergab verschiedene Expressionsmuster von einigen Genen in Abhängigkeit vom Entwicklungsstadium und Gewebe. Interessanterweise erreichten zwei Peptidoglykan-Erkennungsprotein (PGRP) Gene, *PGRP-LB* und *PGRP-SC2*, hohe Expressionslevel im

Mitteldarmgewebe in den späten Puppenstadien, genau dann, wenn die höchste Anzahl an *Blochmannia* in diesem Gewebe vorhanden ist. Die kodierten Proteine besitzen vermutlich eine Amidaseaktivität und Studien an anderen Insekten lassen vermuten, dass diese Proteine die Immunantwort durch das Zerschneiden von Peptidoglykan in nicht-immunstimulierende Fragmente herunterregulieren (Anselme et al. 2006; Bischoff et al. 2006; Zaidman-Remy et al. 2006). In Übereinstimmung mit diesem Ergebnis, wurde nach einer Immunanregung im Puppenstadium die Transkription von Immunfaktoren, wie *Hymenoptaecin*, signifikant weniger im Mitteldarmgewebe als in den übrigen Körperteilen induziert, wohingegen gleichzeitig die Expression von *PGRP-LB* im Mitteldarmgewebe angeregt wurde. Diese Daten deuten an, dass Amidase PGRPs, wie PGRP-LB und -SC2, die Tolerierung der hohen Anzahl an Endosymbionten, welche im Puppenstadium in Mitteldarmgewebe vorhanden ist, ermöglichen könnten. Außerdem könnte die hohe Expression von Wirtsgenen, die ein für Invertebraten-typisches Lysozym und eine Vitellogenin-ähnliche Serin-Carboxypeptidase (CPVL) kodieren, im Mitteldarmgewebe von Puppen den lysosomalen Abbau von *Blochmannia* Endosymbionten anzeigen. Dies könnte dazu dienen, den Ameisenwirt mit essentiellen Nährstoffen zu versorgen, wie es bereits für Aphiden vorgeschlagen wurde (Nishikori et al. 2009a; Nishikori et al. 2009b).

Diese Doktorarbeit ist eine erste Charakterisierung der Immunantwort der Ameise *C. floridanus*. Eine Vielzahl an immun-relevanten Genen wurde bestimmt. Insbesondere die Identifizierung und molekulare Charakterisierung des *Hymenoptaecin* Gens lieferte neue Einblicke in die Immun-Kompetenz von Ameisen im Allgemeinen. Außerdem konnten erste Hinweise für die Beteiligung des Immunsystems an der Kontrolle des Endosymbionten *B. floridanus* gesammelt werden. Demnach wurde eine breite Grundlage geschaffen, welche zukünftige detaillierte Analysen der Verbindung zwischen Immunität und Symbiose bei *C. floridanus* ermöglicht.

3 Introduction

This thesis was aimed to investigate, if the immune system of the ant species *Camponotus floridanus* is involved in controlling the obligate endosymbiont *Blochmannia floridanus*, which is located in specialized midgut cells (Blochmann 1882; Sauer et al. 2002; Schröder et al. 1996). For this purpose the immune system of the ant host had to be characterized.

3.1 The insect immune system

Insects have multiple innate defense mechanisms to act against microbial intruders (Ganesan et al. 2011; Lemaitre and Hoffmann 2007; Strand 2008). The insect innate immune response is composed of cellular and humoral components, which are triggered by the recognition of microbe-associated molecular patterns (MAMPs). Upon MAMP detection microbial invaders are immediately fought by "constitutive" defense mechanisms including production of reactive oxygen species (ROS) and cellular defenses such as phagocytosis. At later time points of infection, a humoral immune response is mounted which mainly leads to the production of antimicrobial peptides (AMPs) (Bulet et al. 2004; Haine et al. 2008). It seems that this inducible response is required to kill those bacteria that have survived the immediate host´s constitutive defenses (Haine et al. 2008).

The cellular immune response of insects involves phagocytosis and encapsulation and is mediated through the activity of insect haemolymph cells (so-called haemocytes). Haemocytes of *Drosophila melanogaster* larvae can be divided into plasmatocytes, crystal cells and lamellocytes based on their structure and function (Strand 2008). 90-95 % of all larval haemocytes are plasmatocytes, which remove pathogens and dead cells via phagocytosis. Crystal cells represent 5 % of the haemocyte population and are involved in melanization reactions through expression and storage of components of the phenoloxidase cascade (Rizki et al. 1980). Lamellocytes encapsulate parasitoids and other objects, which are too large to be phagocytized. They are hardly found in healthy larvae, but their differentiation from haemocyte precursors is strongly induced upon infection of larvae with parasitoid eggs (Lanot et al. 2001).

The humoral immune response includes the production and release of antimicrobial peptides (AMPs), as well as melanization and clotting processes (Feldhaar and Gross 2008; Lemaitre and Hoffmann 2007). In *Drosophila* MAMP recognition mainly leads to signal production via the pathways Toll, IMD (immune deficiency), JAK/STAT (janus kinase/signal transduction and activator of transcription) and/or JNK (c-jun N-terminal kinase) and results in the production of antimicrobial effectors (Boutros et al. 2002; Delaney et al. 2006; Ganesan et al. 2011; Lemaitre and Hoffmann 2007). The expression of these defense molecules is primarily induced in the insect fat body and mainly regulated on the transcriptional level through NF-κB-like transcription factors like Relish (Rel) or Dorsal (Figure 1).

The distinction between the cellular and the humoral immune response is not quite strict, since on the one hand haemocytes can produce humoral effector proteins and on the other hand humoral factors can influence haemocyte functions (Feldhaar and Gross 2008).

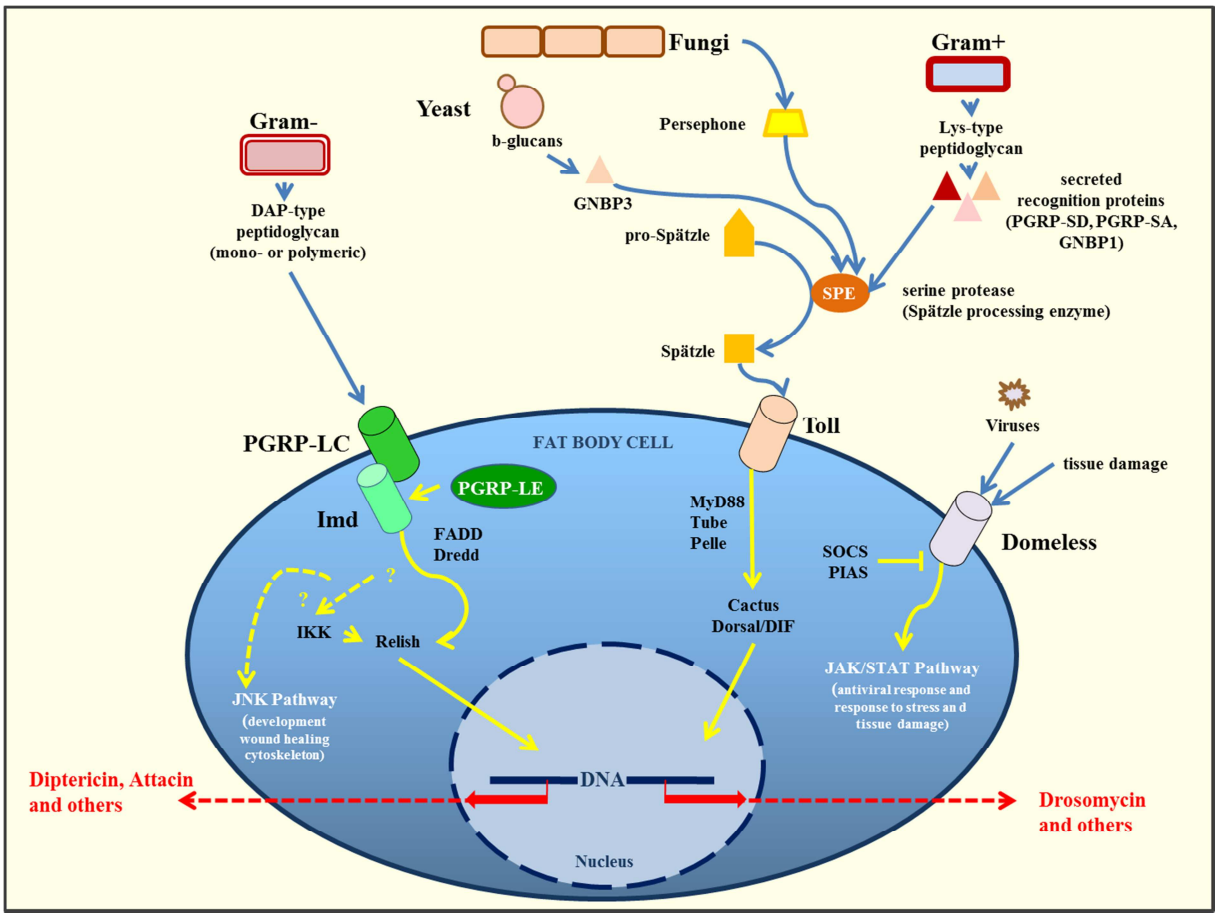

Figure 1. Activation of signaling pathways upon MAMP recognition in *D. melanogaster* (Figure modified from Feldhaar and Gross 2008). The Toll pathway is mainly activated through recognition of fungi and Gram-positive bacteria. These microorganisms are sensed by distinct extracellular pattern recognition receptors (PRRs) (Royet et al. 2011), which activate a proteolytic cascade leading to cleavage of the cytokine Spätzle (Weber et al. 2003). Mature Spätzle binds to Toll causing recruitment of the intracellular Death domain-containing proteins MyD88, Tube, and Pelle (Tauszig-Delamasure et al. 2002). This leads to phosphorylation of the inhibitor Cactus, which thus releases Dorsal and DIF (Nicolas et al. 1998). These Rel-homology domain-containing transcription factors then translocate from the cytoplasm to the nucleus, where they initiate transcription of certain antimicrobials (like Drosomycin) (Lemaitre and Hoffmann 2007). The IMD pathway is primarily activated upon recognition of Gram-negative bacteria by membrane-located peptidoglycan-recognition protein-LC (PGRP-LC) or by intracellular PGRP-LE (Kurata 2010). The adaptor protein IMD then interacts with FADD, which binds to the caspase Dredd (Hu and Yang 2000). This caspase causes activation of Relish, which induces transcription of immune relevant genes (like Diptericin) within the nucleus (Stöven et al. 2000; Stöven et al. 2003). The IMD pathway is furthermore linked to the JNK pathway, which is required for proper wound healing and also plays a role in immune gene expression (Delaney et al. 2006; Rämet et al. 2002). Viral infection as well as cell damage trigger the JAK/STAT pathway, which consists of three main cellular components: the receptor Domeless, the Janus Kinase (JAK) Hopscotch, and the STAT transcription factor (Agaisse and Perrimon 2004; Dostert et al. 2005; Souza-Neto et al. 2009). This results in transcriptional induction of several immune- and stress-responsive genes as well as antiviral genes. SOCS (suppressor of cytokine signaling) and PIAS (protein inhibitor of activated STAT) are known negative regulators of the JAK/STAT pathway in *Drosophila* (Arbouzova and Zeidler 2006).

3.1.1 Recognition of microorganisms

Defense reactions are triggered by the recognition of MAMPs. This recognition of non-self is facilitated by pattern recognition receptors (PRRs), which specifically bind different MAMPs. Upon binding, these PRRs mediate microbial killing through encapsulation and phagocytosis and/or trigger serine protease (SP) cascades and intracellular signaling pathways leading to transcriptional induction of effector genes (Ganesan et al. 2011; Lemaitre and Hoffmann 2007). Insects have various PRRs for the recognition of different types of microorganisms.

Gram-negative binding proteins (GNBPs) are PRRs with similarities to bacterial glucanases. They have lost their enzymatic activity, but still bind to bacterial ligands such as lipopolysaccharides, lipoteichoic acids, or fungal β-1,3-glucans (Gottar et al. 2006; Kim et al. 2000; Lee et al. 1996). The *Drosophila* genome encodes three full length *GNBP* genes (*GNBP1*, *GNBP2*, *GNBP3*) and two short genes with high similarity to the N-terminal part of the GNBPs (De Gregorio et al. 2001; Irving et al. 2001). The other family of PRRs are peptidoglycan-recognition proteins (PGRPs), which specifically bind lysine-type (Lys-type) peptidoglycan (found in almost all Gram-positive bacteria) or diaminopimelic acid-type (DAP-type) peptidoglycan (typical for Gram-negative bacteria) (Dziarski 2004; Royet and Dziarski 2007; Royet et al. 2011). PGRPs can be grouped into two subfamilies. Short *PGRP* genes (*PGRP-S*) have short transcripts and short 5´-untranslated regions (UTRs) and encode extracellular proteins, whereas long *PGRP* genes (*PGRP-L*) exhibit long transcripts and long 5´-UTRs and code for intracellular or membrane-spanning proteins. 13 *PGRP* genes are encoded in the genome of *Drosophila*, which are spliced into 19 proteins (Werner et al. 2003; Werner et al. 2000). In *Drosophila*, some PRRs activate the Toll pathway upon recognition of Lys-type peptidoglycan (PGRP-SA, -SD and GNBP1) or glucan (GNBP3), whereas others trigger the IMD pathway upon recognition of DAP-type peptidoglycan (PGRP-LC and -LE) (Lemaitre and Hoffmann 2007; Royet et al. 2011).

Seven of the Drosophila PGRPs (PGRP-LB-A, PGRP-LB-B, PGRP-LB-C, PGRP-SB1, PGRP-SB2, PGRP-SC1 and PGRP-SC2) have zinc-dependent amidase activity cleaving DAP-type peptidoglycan into non-immunostimulatory muropeptides (Bischoff et al. 2006; Mellroth et al. 2003; Mellroth and Steiner 2006; Zaidman-Remy et al. 2006; Zaidman-Remy et al. 2011). These amidase PGRPs seem to act as scavengers preventing excessive activation of the immune response via downregulation of the IMD pathway. Especially PGRP-LB and -SC1/2, which are mainly expressed in the gut, are supposed to modulate the immune response towards commensal gut bacteria (Bischoff et al. 2006; Zaidman-Remy et al. 2006).

3.1.2 Antimicrobial effectors

Antimicrobial peptides (AMPs) are the best characterized immune effector molecules from insects that are produced in response to infection. The first insect AMPs were described in 1981 from the moth *Hyalophora cecropia* (Steiner et al. 1981). Since then a multitude of

AMPs have been identified, which are produced by many different organisms ranging from animals to plants. Most AMPs have a quite low molecular weight (< 10 kDa), are membrane-active and display hydrophobic and cationic properties. Insect AMPs can be divided into several groups based on their structural characteristics. The main groups are α-helical peptides (e.g. cecropin), cysteine-rich peptides (e.g. defensin), proline-rich peptides (e.g. drosocin), and glycine-rich peptides (e.g. hymenoptaecin) (Reddy et al. 2004).

Several AMPs from Hymenopteran species have been reported so far (Casteels-Josson et al. 1994; Choi et al. 2008; Evans et al. 2006; Tian et al. 2010; Viljakainen and Pamilo 2008). The *Apis mellifera* genome encodes six AMPs. These are two proline-rich (abaecin and apidaecin) and two cysteine-rich AMPs (defensin-1 and defensin-2) as well as one serine-valine-rich (apisimin) and one glycine-rich (hymenoptaecin) AMP (Evans et al. 2006). A bioinformatical study also identified members of these AMPs in the genomes of the seven ants species sequenced so far (Zhang and Zhu 2012). Moreover, the existence of tachystatin-like and crustin-like AMPs was predicted from ant genomes by computational approaches (Zhang and Zhu 2012). Generally, these AMPs are synthesized as an inactive precursor peptide with a signal- and a pro-sequence. Processing of these precursors leads to the active mature peptides (Casteels-Josson et al. 1994).

Defensins known from Hymenopterans are short cationic peptides that contain six cysteine residues forming three stabilizing disulfide bridges. These peptides seem to act primarily against Gram-positive bacteria by interference with acidic phospholipids of the cytoplasmic membrane and the formation of voltage-dependent channels (Cornet et al. 1995; Maget-Dana and Ptak 1997). *A. mellifera* has two structurally different *defensin* genes (*def-1* and *def-2*) (Klaudiny et al. 2005).The *def-1* gene is characterized by the presence of two introns and three exons. The last exon encodes a short amidated C-terminal extension only found in bee defensins. In contrast, the *def-2* gene has only one intron (Klaudiny et al. 2005). Several *def-2* genes of a variety of ant species have been described previously (Viljakainen and Pamilo 2008). These ant *def-2* genes including those of *Formica*, *Lasius* and *Myrmica* species possess a single intron only. The comparison and determination of codon substitution frequencies revealed positive selection in the mature region of the ant defensins, while the signal- and pro-regions of the AMPs seem to evolve neutrally (Viljakainen and Pamilo 2008).

Hymenoptaecins have so far only been found in Hymenopteran species and are glycine-rich AMPs with activity against Gram-negative and Gram-positive bacteria (Casteels et al. 1993). *A. mellifera* hymenoptaecin seems to be lethal for *E. coli* due to sequential permeabilization of the outer and inner membrane (Casteels et al. 1993). Bee *hymenoptaecin* genes encode precursor peptides with a signal- and a pro-sequence, which are processed to the mature and active AMP (Casteels-Josson et al. 1994; Choi et al. 2008; Xu et al. 2009). In contrast, the wasp *Nasonia vitripennis* has two orthologous *hymenoptaecin* genes, which code for multipeptide precursors with an AMP-like region at the position corresponding to the propeptide of bee hymenoptaecins (Gao and Zhu 2010; Tian et al. 2010).

Compared to other AMPs from *A. mellifera* apidaecins are characterized by a special multipeptide precursor structure (Casteels-Josson et al. 1993). Each *apidaecin* gene consists of a conserved N-terminus encoding a signal- and a pro-sequence followed by several exons each encoding a mature apidaecin peptide. These exons appear to exhibit a large intraspecific variation in terms of both number and encoded amino acid sequence (Evans et al. 2006). Furthermore, alternative splicing of *apidaecin* gene produces several mRNAs of different length encoding different numbers and variants of apidaecin peptides. The corresponding multipeptide precursors seem to be posttranslationally processed by specific endoproteases leading to several mature apidaecins. Thus, the expression of a single gene leads to several mature AMPs allowing amplification of the immune response towards bacterial infection (Casteels-Josson et al. 1993).

In addition to AMPs, insects can produce several lysozymes. These are antibacterial enzymes, which are generally characterized by their ability to hydrolyze the β-(1,4)-glycosidic-bond between N-acetylmuramic acid and N-acetylglucosamine in the peptidoglycan layer of the bacterial cell wall resulting in cell lysis (Callewaert and Michiels 2010; Herreweghe and Michiels 2012; Jollès and Jollès 1984). Lysozymes from animals can be grouped into three major distinct types, which are the c-type (chicken- or conventional-type), the g-type (goose-type) and the i-type (invertebrate-type) lysozymes. C-type and i-type lysozymes have been identified in several insect species and likely play an important role in the insect´s antibacterial defense (Callewaert and Michiels 2010). Besides, some lysozymes seem to be adapted for a digestive function in several insect species (Daffre et al. 1994; Lemos et al. 1993; Li et al. 2005; Regel et al. 1998).

3.1.3 ROS production and the antioxidant system

The production of reactive oxygen species (ROS) is one of the immediately acting defense mechanisms of barrier epithelia such as the insect gut, which are in continuous contact with microorganisms. In *Drosophila*, the NADPH oxidase enzyme, dual oxidase (Duox), is essential for ROS production in the gut in response to ingested bacteria (Ha et al. 2005a). The enzyme has an N-terminal extracellular peroxidase domain (PHD), which was shown to convert H_2O_2 into highly microbicidal HOCl. Besides, local AMP production, mediated by the IMD pathway, serves as a backup system against ROS-resistant pathogens in the gut (Ryu et al. 2006).

Excessive ROS production is harmful to host cells and is therefore counteracted by the production of antioxidant enzymes such as superoxide dismutases (SODs) and immune responsive catalase (IRC). SODs reduce superoxide radicals to oxygen and hydrogen peroxide (Fridovich 1995). The latter one is then further reduced to water by IRC (Ha et al. 2005b). Moreover, ROS-caused cell damage has been shown to induce epithelial renewal in the gut through stem cell proliferation, which is mediated by the JAK/STAT and JNK signaling

pathways (Buchon et al. 2009a; Buchon et al. 2009b; Lee 2009). Taken together, gut homeostasis seems to be achieved by balancing the cell damage caused as a side effect of bacterial killing and the epithelial renewal by stem cell division.

Free iron can also catalyze the production of toxic ROS through the Fenton reaction. Iron is usually bound to specific host proteins such as transferrins and ferritins, which are also upregulated during infection (Pham and Winzerling 2010; Yoshiga et al. 1999; Yoshiga et al. 1997). Besides, both hosts and their microbial pathogens require iron as an essential cofactor in basic metabolic pathways and thus compete for available iron (Collins 2003; Schaible and Kaufmann 2004). Interestingly, the endosymbiont *Wolbachia* was shown to interfere with ferritin expression and iron metabolism of its insect hosts (Kremer et al. 2009).

3.1.4 Melanization

Melanization is an innate defense mechanism against invading parasites and is involved in encapsulation, pigmentation and wound healing (Lemaitre and Hoffmann 2007; Söderhäll and Cerenius 1998). The first step in the production of black melanin pigments is the hydroxylation of tyrosine (a monophenol) to produce 3,4-dihydroxyphenylalanine (DOPA, an ortho-phenol), which is catalyzed by tyrosine hydroxylase (TyrOH) or phenoloxidase (PO) (De Gregorio et al. 2001; Gorman et al. 2007). PO furthermore oxidizes ortho-phenols such as DOPA and Dopamine to their corresponding ortho-quinones, which then polymerize to melanin. During melanization several cytotoxic melanin precursors and also ROS are produced that presumably harm invading parasites and pathogens (Nappi and Vass 1993). Furthermore, synthesized melanin sequesters parasites into melanotic capsules. The enzyme PO is produced in crystal cells as an enzymatically inactive precursor called prophenoloxidase (proPO). Activation of PO occurs through proteolytic cleavage of proPO by a serine protease (SP) called prophenoloxidase activating enzyme (PPAE) (Cerenius et al. 2008; Cerenius and Söderhäll 2004). PPAE itself is produced as an inactive zymogen and is activated via cleavage at the end of a proteolytic cascade involving other SPs. This proPO cascade is triggered by wounding or by PRRs upon MAMP recognition. Immoderate activation of the melanization cascade is prevented by SP inhibitors called serpins (Cerenius et al. 2008; Cerenius and Söderhäll 2004; Kanost et al. 2004).

3.2 Antiviral immunity and RNA interference

Just like other organisms, insects can be infected by several different viruses. The insect antiviral response is mainly based on degradation of viral RNA through the RNA interference (RNAi) machinery as well as on the induction of inflammatory signaling pathways, especially JAK/STAT and potentially also Toll and IMD (Sabin et al. 2010; Vodovar and Saleh 2012). The JAK/STAT pathway likely responds to cytokine signaling upon cell damage and induces

expression of some genes counteracting viral infection (Dostert et al. 2005; Souza-Neto et al. 2009).

The most potent antiviral strategy in insects is RNAi, which is a small RNA-dependent gene silencing pathway. RNAi is induced upon recognition of viral double-stranded RNA (dsRNA) in the cytosol of infected cells. A specific endonuclease, the RNaseIII-enzyme Dicer-2, cleaves the dsRNA into small interfering RNA (siRNA) duplexes of 21-23 nucleotides in length (Bernstein et al. 2001). The protein R2D2 separates the two siRNA strands and loads the guide strand onto the Argonaute protein AGO-2 within the RNA-induced silencing complex (RISC). This RISC complex then targets RNAs with complementary sequence to the loaded siRNA and specifically degrades these viral RNAs by catalytic activity of AGO-2 (Martinez et al. 2002).

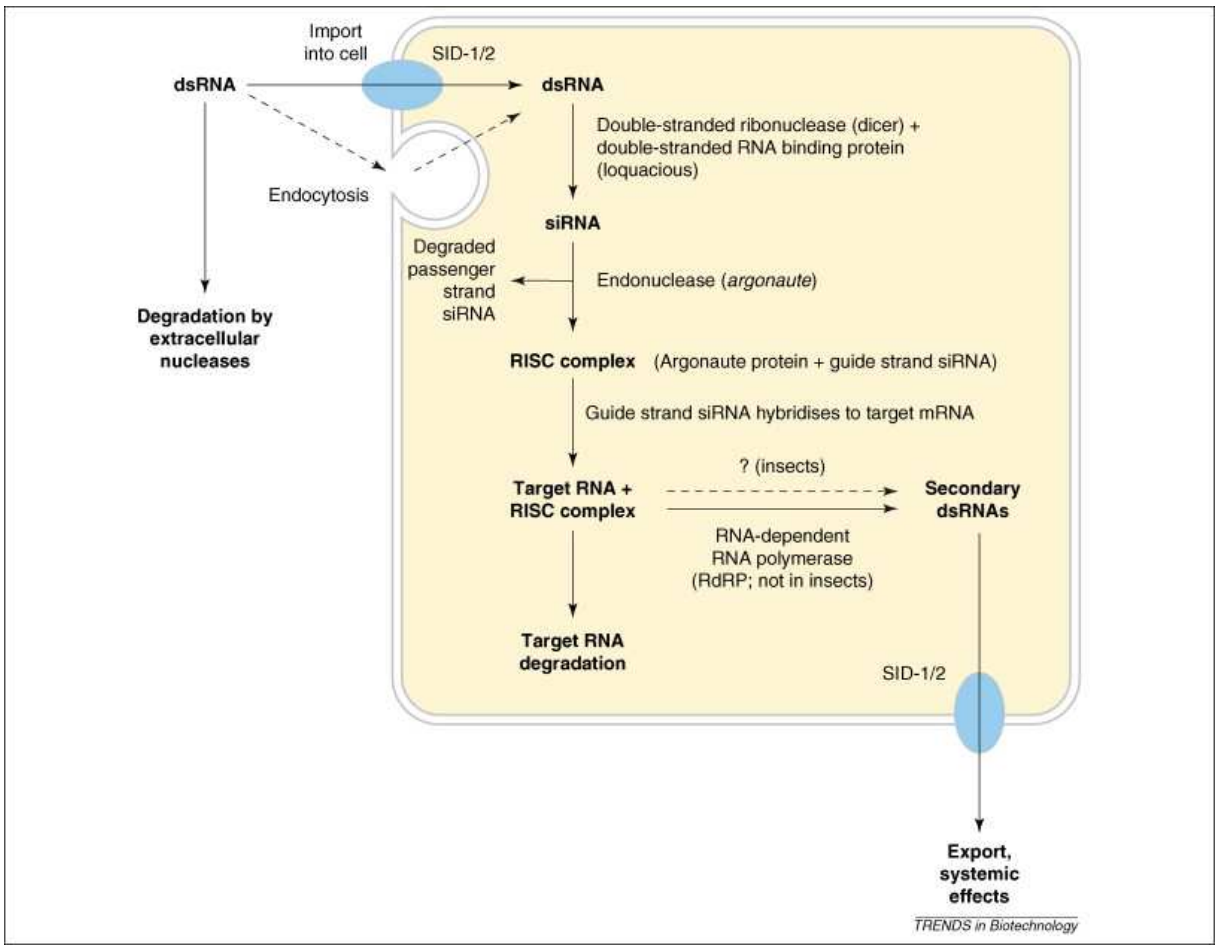

Figure 2. Schematic overview of dsRNA-mediated gene silencing in cells of lower animals (Price and Gatehouse 2008).

Despite its involvement in antiviral immunity, RNAi is furthermore a conserved gene silencing mechanism in many eukaryotes (Figure 2) (Fire et al. 1998; Mello and Conte 2004). In this case the downregulation of gene expression is also triggered by specific dsRNA and

target mRNA is degraded by the RISC-complex, as described above. Thus, RNAi is also a post-transcriptional control mechanism involving degradation of a target mRNA. RNAi-mediated gene silencing can even have systemic effects meaning that gene knockout effect disperses throughout the organism, persists over development and/or is transmitted to the progeny (Fire et al. 1998; Mello and Conte 2004). In plants and nematodes, a cellular RNA-dependent RNA polymerase (RdRP) mediates these systemic RNAi effects (Dalmay et al. 2000; Sijen et al. 2001). The enzyme RdRP is primed by siRNA strands and uses the target RNA as a template for the production of new dsRNAs. Processing of these synthesized dsRNAs by Dicer then generates new siRNAs (secondary siRNAs) and thus amplifies the silencing signal. Furthermore, the secondary dsRNAs can be exported from the cell to neighbouring cells and thus spread the gene knockout effect through the organism. In *Caenorhabditis elegans* SID (systemic RNA interference deficient) proteins mediate the transport of dsRNA between cells (Feinberg and Hunter 2003; Winston et al. 2002; Winston et al. 2007). Orthologs of the transmembrane protein SID-1 have been identified in several insect species, but do not seem to be essential for dsRNA uptake in insects (Huvenne and Smagghe 2010). The presence of a *C. elegans*-like RdRP enzyme could not be confirmed in the genomes of insects sequenced so far. However, systemic RNAi effects have been reported in several insect orders, including Diptera, Coleoptera and Hymenoptera suggesting that the observed signal amplification is based on another mechanism than in *C. elegans* (Huvenne and Smagghe 2010; Price and Gatehouse 2008). The practicability of RNAi in insects and the increasing availability of insect genomes have made this technology a powerful tool to study gene functions by loss-of-function analyses and thus allows to perform detailed genetic analysis in animals, which were so far not amenable to genetic investigations (Mito et al. 2011).

In the last two years genome sequencing projects have made genes of seven ant species accessible for knockdown studies using RNAi (Bonasio et al. 2010; Nygaard et al. 2011; Smith et al. 2011a; Smith et al. 2011b; Suen et al. 2011; Wurm et al. 2011). However, it is unclear, how easily the application of RNAi can be adopted for ants, as several studies on other insects indicate that different species vary strongly in their response to dsRNA as well as in their ability to distribute the RNAi signal systemically through the whole body (Huvenne and Smagghe 2010; Mito et al. 2011; Terenius et al. 2011). In several studies it was shown that in the honey bee *A. mellifera* dsRNA injection of adult workers results in the induction of RNAi (Amdam et al. 2003; Ihle et al. 2010; Nelson et al. 2007; Wang et al. 2010). In honey bee larvae effective gene knockdown can also be achieved via direct feeding of a dsRNA-containing diet (Nunes and Simoes 2009; Patel et al. 2007; Wolschin et al. 2011), whereas RNAi through feeding of dsRNA seems to be hardly attainable in adult worker bees (Kate E. Ihle, pers. comm.). So far, only two studies described a successful knockdown of gene expression in ants however after injection of interfering RNAs (Bonasio et al. 2010; Lu et al. 2009).

Injection as delivery method for interfering RNAs might implicate problems for the interpretation of the results, since injection itself entails an alteration of the animal´s physiology. For example, injection of a sterile Ringer solution already caused upregulation of immune genes in the ant *Camponotus floridanus* and in other insects (Boutros et al. 2002; Choi et al. 2008; Ratzka et al. 2011). Accordingly, feeding of interfering RNAs seems to be a much more suitable approach, especially for the study of immune genes. Direct feeding of bacteria expressing dsRNA is the most easily performable delivery method, but may require large quantities of dsRNA. Feeding of dsRNA has been standardized for *C. elegans* (Timmons et al. 2001; Timmons and Fire 1998) and has recently also been adopted for several insect species (Tian et al. 2009; Zhu et al. 2011). For *in vivo* production of dsRNA, *E. coli* HT115(DE3) strain is used, which lacks the dsRNA-degrading enzyme RNase III and exhibits IPTG-inducible T7 RNA polymerase activity for transcription of target genes (Timmons et al. 2001; Timmons and Fire 1998). The critical point of this method is that the bacteria and therewith the dsRNA need to be taken up in a high concentration. This may be a problem in insect species that usually do not feed on bacteria. In this case, the dsRNA needs to be extracted from the bacteria before feeding or synthesized *in vitro*, which is however more expensive. Moreover, feeding of dsRNA does not always achieve as strong RNAi effects as injection, as insects seem to differ in their ability to respond to ingested dsRNA (Terenius et al. 2011).

It is anticipated that RNAi should work in the ant species *C. floridanus*, as genes coding for core components of the RNAi pathway, such as Dicer or RISC proteins, are well conserved in the ant´s genome (Bonasio et al. 2010). The presence of a *SID-1* (systemic RNA interference deficient) orthologue furthermore suggests the possibility of systemic RNAi in *C. floridanus*. Accordingly, injection of siRNAs provoked downregulation of a cGMP-dependent protein kinase in *C. floridanus* (Bonasio et al. 2010). Nonetheless, it is doubtful if feeding of dsRNA would result in RNAi in *C. floridanus*, since the crop content of adult workers has a quite low pH (Feldhaar, unpublished results) and likely contains several digestive enzymes (Hamilton et al. 2011) possibly degrading ingested dsRNAs. Moreover, studies targeting gene expression in larvae face the additional problem, that the larval brood of ants is helpless and has to be fed by workers (Hölldobler and Wilson 1990). Thus, dsRNA is first taken up by worker ants, whereas it is unclear, whether it becomes too diluted or degraded before transmission to larvae by regurgitation.

3.3 Social immunity

Eusocial insects such as some bees and wasps, ants and termites are characterized by a reproductive division of labor meaning the formation of (more or less) sterile and reproductive castes. Moreover, these insects stand out by the living together of overlapping generations usually with division of labor regarding brood care and food supply (Hölldobler and Wilson 1990; Wilson 1971). Colonies of (eu)social insects may contain up to several

thousand genetically closely related individuals and thus are faced with specific problems concerning hygiene issues and pathogen defense. To counteract the increased risk of infection and to complement their individual defenses, social insects have evolved collective defense mechanisms, often referred to as social immunity (Cremer et al. 2007). These social defenses include mechanisms that decrease disease transmission and susceptibility such as the intermixture of antimicrobial substances into the nest material, the spread of antimicrobial glandular secretions on their cuticles as well as the detection and removal of infected brood or dead animals from the nest (Cremer et al. 2007; Wilson-Rich et al. 2009). Furthermore, social insects may accumulate biofilms with antibiotic producing microbes on their cuticle as it is known from attine ants (Currie et al. 2006; Currie et al. 1999). These fungus-farming ants seem to selectively accept certain microbes like *Pseudonocardia* and *Streptomyces* into the biofilms of their integument as well as of their fungus gardens, which protect their colonies against ant and garden diseases, respectively (Mueller 2012).

In addition to this antiseptic or hygienic behavior, a phenomenon called social immunization occurs in social insect colonies meaning lower susceptibility of uninfected nestmates to a certain pathogen after contact to nestmates infected with the same pathogen (Hamilton et al. 2011; Konrad et al. 2012; Rosengaus et al. 1999; Traniello et al. 2002; Ugelvig and Cremer 2007). Social immunization may occur actively via triggering of own immune responses in uninfected nestmates after contact with infected individuals and/or passively via social transfer of immune effectors (Konrad et al. 2012; Rosengaus et al. 1999; Traniello et al. 2002). Passive social immunization was recently suggested as protective mechanism against bacterial infection in ant colonies of *Camponotus pennsylvanicus* (Hamilton et al. 2011). *C. pennsylvanicus* workers that had received an abdominal injection of bacteria or lipopolysaccharides showed increased trophallactic behavior and the exchanged regurgitate droplets had elevated levels of antimicrobial activity. Moreover, the authors identified a lysosomal aspartic protease that was more abundant in regurgitate droplets of immunized workers (Hamilton et al. 2011). Apart from that, there is also evidence for active social immunization in ant colonies after fungus infection of single individuals (Konrad et al. 2012). *Lasius* ant colonies exhibited intensive allogrooming behavior between fungus-exposed and healthy nestmates, which seemed to trigger low-level infections in healthy nestmates through pathogen transfer. These low-level infections are not lethal, but likely activate antifungal immune responses in healthy animals resulting in active immunization of the colony. It was suggested that the observed differences in social immunization mechanisms after fungal and bacterial infection are due to the distinct infection modes of these pathogens. Entomopathogenic fungi are usually externally transmitted through spores on the cuticle allowing active immunization via social transfer of immune elicitors, whereas bacterial infections generally happen via oral uptake (Konrad et al. 2012).

3.4 Bacterial endosymbioses in insects

3.4.1 Primary and secondary endosymbioses

Insects are the most diverse and successful animal group on earth, which is reflected in the variety of habitats they live in, their abundance and species richness. The successful occupation of a large variety of ecological niches is often facilitated by symbiosis with several microorganisms (Feldhaar 2011; Ratzka et al. 2012b). Symbiosis is defined as the close and constant living together of organisms belonging to different species, whereas the quality of the association can range from mutualistic to parasitic (De Bary 1879). When symbionts reside within the animal´s body, they are referred to as endosymbionts, in contrast to ectosymbionts, which are located on the outer surface of the animals. Bacterial endosymbionts of insects have been shown to confer various fitness advantages to their hosts such as nutritional upgrading (Akman et al. 2002; Douglas 1998; Feldhaar et al. 2007; Zientz et al. 2004), thermal tolerance (Dunbar et al. 2007) or enhanced pathogen/parasitoid resistance (Brownlie and Johnson 2009; Currie et al. 2003; Kaltenpoth 2009; Kaltenpoth et al. 2005; Oliver et al. 2010; Oliver et al. 2009; Oliver et al. 2005; Teixeira et al. 2008). Traditionally, bacterial endosymbionts of insects have been classified into primary and secondary symbionts.

Associations between primary endosymbionts and their insect hosts are evolutionary long established (for up to estimated 250 million years) (Baumann 2005). Frequently, these obligate symbionts reside within specialized host cells (so-called bacteriocytes) that may form an organ called the bacteriome (Braendle et al. 2003). Primary endosymbionts are mainly vertically transmitted to the host progeny via the germ line, which prevents endosymbiont recombination with free-living bacteria. As a consequence, genomes of primary endosymbionts are subjected to a relatively high rate of DNA evolution resulting in accumulation of deleterious mutations, a drastic reduction of the genome size and an extremely high adenine-thymine content of the genome (Itoh et al. 2002; McCutcheon and Moran 2012; Moran 1996; Wernegreen 2002; Wernegreen 2012). During host-symbiont coevolution most of the bacterial genes, which were redundant with host genes or not required in the stable intracellular environment, were deleted, whereas bacterial genes encoding factors that complemented metabolic pathways of the host were retained (Gil et al. 2004; Moya et al. 2008; Zientz et al. 2004). Thus, primary endosymbionts often provide their hosts with nutrients lacking from its diet and are necessary for host survival and reproduction. Because of their extreme adaption to the intracellular life style, most of these primary endosymbionts are not culturable *in vitro* (Kikuchi 2009).

In contrast, secondary endosymbionts are usually facultative symbionts, that are not restricted to bacteriocytes, but may also be found extra- or intracellularly in other host tissues or in the haemocoel (Dale and Moran 2006). These secondary endosymbionts are often culturable *in vitro* (Kikuchi 2009). Generally, their transmission occurs maternally, but horizontal transmission is also possible (Dale and Moran 2006; Oliver et al. 2010; Russell et al. 2003).

Since secondary endosymbionts are not required for host survival and are in some cases even detrimental for their host, they are in contrast to primary endosymbionts usually not carried by all individuals of a host species (Harris et al. 2010). The genomes of secondary endosymbionts seem to be not (yet) as reduced as those of primary endosymbionts (Burke and Moran 2011a; Degnan et al. 2009; Toh et al. 2006; Wernegreen 2012). It was even shown that secondary endosymbionts may complement beneficial functions of primary endosymbionts that were otherwise lost due to uncontrolled genome reduction of primary endosymbionts (Koga et al. 2003; Perez-Brocal et al. 2006). Accordingly, it was suggested that secondary endosymbionts might even replace primary endosymbionts, when the destructive degeneration of the latter genomes has proceeded too far and they have lost essential functions for their hosts (Perez-Brocal et al. 2006).

So far, the most widely-studied model organisms in terms of host-symbiont interactions are the tsetse fly *Glossina morsitans* (Diptera), the aphid species *Acyrthosiphon pisum* (Hemiptera) as well as the weevil *Sitophilus zeamais* (Coleoptera) and its sibling species *Sitophilus oryzae*.

G. morsitans flies harbour several different bacterial symbionts. These include the secondary symbionts *Sodalis glossinidius* and *Wolbachia* as well as the obligate mutualist *Wigglesworthia glossinidia*, which are maternally transmitted to the host offspring throughout milk gland secretions during the intrauterine development (Aksoy 2000; Attardo et al. 2008). In adult flies the primary endosymbiont *Wigglesworthia*, a Gram-negative γ-Proteobacterium, resides free in the cytosol of bacteriocytes forming a bacteriome at the anterior gut and it provides its host with certain vitamins lacking from its vertebrate blood diet (Akman et al. 2002; Nogge 1981). Based on the concordant evolution of *Wigglesworthia* and its tsetse host the origin of this symbiosis was estimated to date back about 80 million years (Chen et al. 1999).

The primary endosymbiont of *A. pisum* is *Buchnera aphidicola*, an obligate intracellular Gram-negative γ-Proteobacterium that supplies its aphid hosts with essential amino acids lacking from their strict plant phloem diet (Wilkinson 1998). It is located in vacuole-like organelles within bacteriocytes forming a bilobed bacteriome in the aphid haemocoel (Baumann 2005; Douglas 1998; Griffith and Beck 1973). The symbiotic association between aphids and *Buchnera* was estimated to be about 150-250 million years old (Munson et al. 1991). Besides this primary endosymbiont, *A. pisum* often harbors one or more secondary facultative endosymbionts extracellularly in the haemolymph, especially the Gram-negative bacteria *Hamiltonella defensa*, *Serratia symbiotica* and *Regiella insecticola* (Moran et al. 2005b; Oliver et al. 2010).

Cereal weevils of the genus *Sitophilus*, including the three species *S. zeamais*, *S. oryzae* and *S. granarius*, harbor an obligate Gram-negative γ-Proteobacterium called *Sitophilus* primary endosymbiont (or SPE) within bacteriocytes (Heddi et al. 1998; Lefevre et al. 2004). In comparison to other primary endosymbionts the genome of SPE is rather unreduced and the

association of *Sitophilus* with SPE has a relatively recent origin (less than 25 million years ago), probably by symbiont replacement (Conord et al. 2008; Gil et al. 2008; Lefevre et al. 2004). In larval insects the SPE-bearing bacteriocytes form a bacteriome, which derives from the insect gut (Heddi et al. 1999b). During nymphal metamorphosis, this larval bacteriome dissociates and the symbionts are released into the haemolymph. The symbionts must then reach and infect the new adult bacteriome, which is associated with the hindgut (Anselme et al. 2006; Buchner 1965). In addition to SPE, weevils are often infected with *Wolbachia*, which are distributed throughout the whole body with particularly high density in the germ cells, where they cause nucleocytoplasmic incompatibility (Heddi et al. 1999b).

3.4.2 The symbiosis between *Camponotus floridanus* and *Blochmannia floridanus*

The symbiotic association between carpenter ants (genus *Camponotus*) and bacteria was already described at the end of the 19th century by Friedrich Blochmann as the first known obligate endosymbiosis in insects (Blochmann 1882). The ant genus *Camponotus* includes more than 1000 species with worldwide distribution, and thus belongs to the most diverse and successful ant genera (Bolton 1995; Hölldobler and Wilson 1990). Most *Camponotus* species are omnivorous and live in monogynous colonies (with only one reproducing queen per colony), while one colony may contain thousands of workers. All *Camponotus* species investigated to date as well as some related genera of the tribe Camponotini (including *Polyrhachis*, *Echinopla* and *Calomyrmex*) harbor obligate intracellular Gram-negative γ-Proteobacteria of the genus *Blochmannia*. The occurrence of *Blochmannia* within the genera *Camponotus*, *Polyrhachis* and *Colobopsis* suggests, that the *Blochmannia*-ant association originated before the divergence of these three taxa about 40 million years ago (Degnan et al. 2004; Sameshima et al. 1999; Sauer et al. 2000; Wernegreen et al. 2009). *Blochmannia* endosymbionts live free in the cytosol of specialized bacteriocytes of the ant host, which are intercalated between midgut cells as well as in ovaries of females (Sauer et al. 2002; Schröder et al. 1996). Since *Blochmannia* is only vertically transmitted to the next generation, host species and symbiont strictly coevolve allowing for classification of *Blochmannia* species according to the corresponding host species (Degnan et al. 2004).

So far, the best characterized *Blochmannia*-ant association is that between *C. floridanus* and *B. floridanus*. The genomes of the endosymbiont and recently also of the ant host have been sequenced facilitating the investigation of this symbiosis on the molecular level (Bonasio et al. 2010; Gil et al. 2003). *B. floridanus* has a strongly reduced genome, which consists of a circular chromosome of about 705 kb (636 genes, guanine-cytosine content of about 27.4 %), corresponding to one seventh of the *E. coli* genome (Gil et al. 2003). The endosymbiont has lost biosynthetic pathways for several non-essential amino acids, while those for all essential amino acids, with exception of arginine, were retained. However, based on *Blochmannia*-synthesized citrulline arginine may be produced by the host itself (Gil et al. 2003; Zientz et al. 2004). In contrast to the genome of its insect host, *B. floridanus* genome encodes all enzymes

necessary for reduction of sulfate to sulfide, which can be incorporated into biomolecules and is for example needed for cysteine production. Moreover, *B. floridanus* contains a complete urease gene cluster. The encoded urease enzyme hydrolyzes urea to CO_2 and ammonia. The latter one may be incorporated into glutamate via *Blochmannia*-synthesized glutamine synthetase resulting in glutamine (Figure 3) (Gil et al. 2003; Zientz et al. 2004).

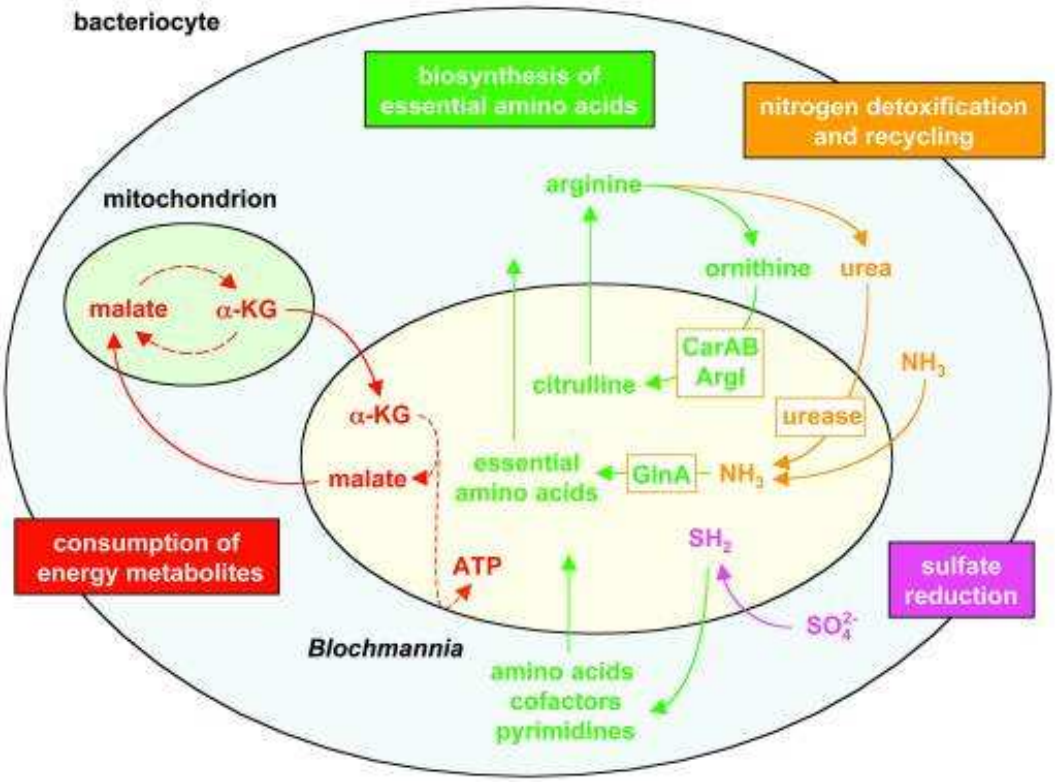

Figure 3. Hypothetic metabolic interactions of *B. floridanus* with the host bacteriocyte (Gil et al. 2003, Copyright (2003) National Academy of Sciences, U.S.A.). In this simplified model the endosymbiont receives energy metabolites (possibly in form of α-ketoglutarate (α-KG)) as well as non-essential amino acids, cofactors and pyrimidines from its host cell. *Blochmannia* is able to catabolize urea to ammonia (NH_3) via its urease and furthermore it can reduce sulfate (SO_4^{2-}) to sulfide (SH_2). With these components the endosymbiont may produce essential amino acids and citrulline (for production of arginine by the host cell) and provide them to its host cell.

The supposed nutritional role of *B. floridanus* for its host *C. floridanus* was proven via *in vivo* analyses (Feldhaar et al. 2007; Zientz et al. 2006). Groups of worker ants were cleared from their endosymbionts by feeding the antibiotic Rifampicin. When such groups of aposymbiotic workers were exclusively fed with a diet lacking essential amino acids, they raised significantly less brood than groups of endosymbiont-bearing workers fed on diets with and without essential amino acids. This effect could be reversed, when aposymbiotic workers were also provided with essential amino acids in their diet suggesting that endosymbiotic *Blochmannia* can complement a deficient diet of its host with essential amino acids. Moreover, a role of *B. floridanus* in nitrogen recycling via urease function was confirmed through proof of ^{15}N-enrichment in free essential amino acids of workers, which were

previously fed with ^{15}N-labelled urea. It was suggested that nutritional upgrading of the diet by *Blochmannia* may confer a significant fitness advantage to its host, as it allows endosymbiont-bearing ant species to mainly fed on nitrogen-poor plant derived resources (Feldhaar et al. 2007).

The factors that mediate the ant-*Blochmannia* symbiosis on the molecular level are so far unknown. Strikingly, *Blochmannia*, as well as *Wigglesworthia*, seem to lack the *dnaA* gene, which encodes an essential protein for DNA replication initiation in bacteria. Moreover, *Blochmannia* also lacks *priA* and *recA* gene encoding the known alternative proteins for DNA replication initiation (Gil et al. 2003). It was suggested that the cytosolic localization of *Blochmannia* and *Wigglesworthia* may require the direct control of DNA replication by the host via so far unknown host factors. Accordingly, *Buchnera*, which still retains its *dnaA* gene, is enclosed with host membranes within bacteriocytes (Gil et al. 2003). Furthermore, the symbiont densities seem to be regulated somehow in *C. floridanus*, as the number of *B. floridanus* endosymbionts per host and also the distribution of bacteriocytes varies strongly over different developmental stages (Figure 4) (Stoll et al. 2010; Stoll et al. 2008). The percentage of endosymbiont-bearing midgut cells increases strongly during host ontogeny and peaks in late pupal stages, where the entire midgut is transformed into a symbiotic organ. Thus, symbionts seem to have an important function (possibly for host nutrition) in this developmental phase during which ant hosts are separated from the environment by the puparium and no external food is ingested. Accordingly, transcriptional profiling of *B. floridanus* genes during the various developmental stages of its host revealed higher expression of putative symbiotic genes (e.g. amino acid biosynthesis genes, urease associated genes) in comparison to housekeeping genes in particular during the pupal stage (Stoll et al. 2008). After eclosion of workers the bacterial number decreases with increasing age of workers (Stoll et al. 2010; Stoll et al. 2008). This variation of symbiont densities during host development might be regulated via host immune factors.

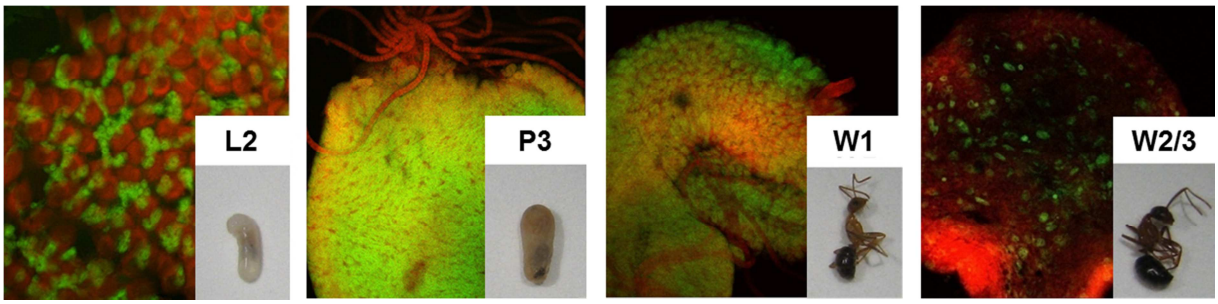

Figure 4. Confocal images of the *C. floridanus* midgut from different developmental stages: larvae (L2), late pupae (P3), young (W1) and old (W2,3) workers (modified from (Stoll et al. 2010)). *Blochmannia* cells within bacteriocytes were stained with a 16S rRNA specific green-fluorescent oligonucleotide and host cells with red-fluorescent SYTO Orange 83. The highest number of *Blochmannia*-cells per host can be found during late pupal stages, when the whole midgut is transformed into a symbiotic bacteriome.

3.5 Immunity and symbiosis

At first glance, it seems surprising that bacterial endosymbionts are not eradicated by the host immune system, as they are often very closely related to pathogenic bacteria (Moran et al. 2008), which are recognized and fought by the insect immune system. This raises the question, which mechanisms allow insect hosts to maintain a chronic infection with bacterial symbionts (Ratzka et al. 2012b). Are bacterial endosymbionts not recognized as non-self by the insect immune system or just hidden from it via compartmentalization? Do they actively manipulate and/or overcome the immune response? Recently, several studies have addressed these questions in different model organisms (Anselme et al. 2008; Anselme et al. 2006; Gerardo et al. 2010; Kremer et al. 2012; Laughton et al. 2011; Login et al. 2011; Login and Heddi 2012; Ratzka et al. 2012b; Ratzka et al. 2011; Vigneron et al. 2012; Wang and Aksoy 2012; Wang et al. 2009; Weiss et al. 2011).

3.5.1 Bacterial adaptations

Special adaptions may allow bacterial endosymbionts to avoid or resist the host immune response (Ratzka et al. 2012b). Altered surface structures usually used for pattern recognition can enable evasion of the insect´s immune system. For example, *Sodalis glossinidius*, a secondary facultative endosymbiont of the tsetse fly *Glossina morsitans morsitans*, has amino acid substitutions and insertions in the exposed loop domains of its major outer membrane protein (OmpA), which are absent from pathogenic bacteria (Weiss et al. 2008). It was shown that such polymorphisms contribute to host tolerance of symbiotic bacteria. In case of *Sodalis* OmpA seems to trigger a strong host immune response, which the bacterium probably can survive, because it is resistant to several host AMPs (Hao et al. 2001; Hu and Aksoy 2005). Furthermore, it was argued that *Sodalis* may also escape elimination by encapsulation or melanization processes by invading host haemocytes via a type three secretion system (Dale et al. 2001; Weiss et al. 2008).

In contrast to *Sodalis*, which seems to tolerate the host immune response, other bacterial endosymbionts, such as *Wolbachia* (Bourtzis et al. 2000) and *Spiroplasma* (Hurst et al. 2003) may either not elicit an immune response or may actively interfere with the host immune system, thus preventing an immune response (Ratzka et al. 2012b).

Wolbachia are a group of obligate intracellular and maternally transmitted bacteria, which are extremely widespread in arthropods and are presumably found in up to 66 % of all insects (Hilgenboecker et al. 2008). The secret of the success of this group of bacteria is the ability to manipulate host reproduction by inducing cytoplasmic incompatibility, parthenogenesis, feminization and male-killing (Stouthamer et al. 1999; Werren 1997; Werren et al. 2008). Several studies indicate that the interactions between *Wolbachia* and its host are highly specific for combinations of *Wolbachia*-strain and host. In some combinations, *Wolbachia* seems to neither induce nor suppresse AMP expression (Bourtzis et al. 2000), whereas it

strongly induces immune gene expression in other combinations (Kambris et al. 2010; Moreira et al. 2009). A recent transcriptomic approach in the parasitoid wasp *Asobara tabida* furthermore indicated that immune genes induced by *Wolbachia* in symbiotic wasps are mainly located upstream in the immune pathways, whereas downstream effectors, such as AMPs, are rather downregulated (Kremer et al. 2012). This finding highlighted the fact that *Wolbachia* is indeed detected by the host organism, but subsequently extensively modifies the host immune response to escape elimination (Ratzka et al. 2012b). The mechanisms of how *Wolbachia* interferes with the host immune system are yet not well characterized, but may be mediated through ankyrin repeat (ANK) proteins possibly secreted by a type four secretion system, both encoded by the *Wolbachia* genome (McGraw and O'Neill 2004; Wu et al. 2004).

Spiroplasma endosymbionts of insects can be found both extra- and intracellularly and can form a broad range from mutualistic to pathogenic associations with their hosts (Regassa and Gasparich 2006). It was shown that infection of *D. melanogaster* with male-killing *Spiroplasma* did not induce immune gene expression, whereas infected hosts were still able to induce immune gene expression after septic injury (Herren and Lemaitre 2011; Hurst et al. 2003; Hutchence et al. 2011). The lack of immune recognition may be due to the fact that *Spiroplasma* are wall-less bacteria (Whitcomb and Williamson 1975) and therefore lack MAMPs like peptidoglycan that usually trigger insect immune responses (Aggrawal and Silverman 2007). A recent study on a male-killing *Spiroplasma* strain, however, demonstrated that the bacterium actively interferes with the host immune system as it was capable to resist an immune response and to suppress constitutive expression of some immune genes in immune-unchallenged insects (Anbutsu and Fukatsu 2010). In contrast, a closely related non male-killing strain did not exhibit immune suppression (Anbutsu and Fukatsu 2010).

3.5.2 Modulation of the host immune response towards symbionts

Fine-tuning of the host immune response towards endosymbiotic bacteria can facilitate their maintenance (Login and Heddi 2012; Ratzka et al. 2012b). In *Drosophila* the local immune response of the gut seems to be accurately adjusted for the maintenance of a commensal gut microbiota. The direct contact between gut epithelia and ingested or resident gut bacteria constantly triggers the IMD signaling pathway in the gut. This does not result in AMP production in the gut though, since the intestinal homeobox gene *caudal* represses nuclear factor Kappa B- (NF-κB-) dependent AMP genes (Ryu et al. 2008). Furthermore, PRRs with an amidase activity, such as the PGRP-LB and PGRP-SC2, play an important role in bacterial tolerance at epithelial surfaces of *Drosophila* (Bischoff et al. 2006; Zaidman-Remy et al. 2006). These catalytic proteins are amidases that specifically cleave peptidoglycan of Gram-negative bacteria and thereby remove microbial immune elicitors. This negative immune regulation presumably allows modulation of the immune response towards commensal gut bacteria.

Another possibility to avoid immune actions against symbionts is their compartmentalization into specialized host cells. These so-called bacteriocytes may have an altered gene expression allowing symbionts to persist without triggering an immune response. In the maize weevil *S. zeamais* the bacteriocytes harboring SPE form an organ, the bacteriome (Heddi et al. 1999b). Interestingly, two putative negative regulators of the immune response, *tollip* and *wPGRP-1*, were found to be constitutively overexpressed within the bacteriome (Anselme et al. 2008; Heddi et al. 2005). The *wPGRP-1* gene, which encodes the weevil homolog of PGRP-LB, apparently plays a key role in preventing activation of the weevil immune system against SPE, as it was not only overexpressed in the bacteriome tissue, but also strongly induced in the symbiotic nymphal phase of *S. zeamais*, where symbionts were released from bacteriocytes (Anselme et al. 2006). Thus, upregulation of the amidase *wPGRP-1* at this stage is likely to prevent immune defenses against the extracellular symbionts through removal of bacterial elicitors (Anselme et al. 2006). A recent transcriptome study on *S. oryzae*, the sibling species of *S. zeamais*, also demonstrated that the gene expression profile of bacteriocytes is precisely adjusted to allow tolerance of SPE (Vigneron et al. 2012). In the bacteriome tissue genes involved in cell growth and survival were highly expressed, whereas the expression of immune-relevant genes possibly harming SPE was significantly lower in the bacteriome in comparison to that in larvae without symbionts (aposymbiotic larvae) or to that in larvae challenged with *E. coli*.

A study on the tsetse fly *G. morsitans morsitans* demonstrated the importance of PGRP-LB for symbiotic tolerance (Wang et al. 2009). Tsetse flies host the primary obligate endosymbiont *Wigglesworthia glossinidia* within a bacteriome consisting of differentiated midgut cells (Aksoy 1995). The *PGRP-LB* gene was strongly expressed in this organ and its expression level correlated with symbiont numbers (Wang et al. 2009). As aforementioned, tsetse progeny receive *Wigglesworthia* symbionts within the mother´s milk during intrauterine development (Aksoy 2000; Attardo et al. 2008). PGRP-LB was shown to be a major component of this milk (Wang and Aksoy 2012). RNAi-mediated knock down of *PGRP-LB* expression resulted in AMP production and finally in decreased *Wigglesworthia* number within the bacteriome as well as in the mother´s milk (Wang and Aksoy 2012; Wang et al. 2009). Accordingly, PGRP-LB likely cleaves peptidoglycan released by *Wigglesworthia* under normal conditions and thus prevents elimination of the symbiont through the host immune system. Furthermore, it may also contribute to symbiont density regulation, since recombinantly expressed PGRP-LB was shown to exhibit antimicrobial activities (Wang and Aksoy 2012).

An extreme example for symbiont tolerance is the pea aphid *Acyrthosiphon pisum*. Comparison of the *A. pisum* genome sequence with the known genomes of other insects revealed, that aphids lack most of the genes of the IMD pathway involved in recognition of bacteria (e.g. PGRPs), signal transduction (e.g. IMD or Relish) and antimicrobial response (e.g. AMPs like defensin) (Consortium 2010; Gerardo et al. 2010). This reduced immune gene

repertoire possibly allows the presence of several different symbiont species (Gerardo et al. 2010). Several other studies confirmed that in comparison to other insects aphids do not raise a strong humoral immune response against potential pathogens (Altincicek et al. 2008a; Burke and Moran 2011b; Gerardo et al. 2010; Laughton et al. 2011). It was discussed that the stably established primary symbiosis with *Buchnera* might have selected for the loss of immune genes involved in combating Gram-negative bacteria and, as a consequence, also allows the presence of a suite of secondary symbionts (Douglas et al. 2011; Gerardo et al. 2010). However, other insects also harbour several symbionts and still have a complex immune gene repertoire. In fact, a recent study suggested that more sophisticated mechanisms than a mere gene loss are necessary for the maintenance of symbionts within their aphid hosts (Laughton et al. 2011). It was shown that the aphid cellular immune system is able to eradicate at least some microbial intruders, while leaving symbionts unharmed (Laughton et al. 2011). This finding implied that aphid symbionts are adapted in a way that they are not recognized or harmed by aphid haemocytes and/or that the aphid immune system is able to distinguish between symbionts and potential pathogens.

3.5.3 Control of endosymbionts

Although endosymbionts are beneficial for their insect hosts and need to be maintained, hosts have to control symbiont number within their tissues in order to prevent their uncontrolled proliferation (Ratzka et al. 2012b). Compartmentalization not only allows maintenance of symbionts unharmed from immune attacks, but also helps to restrict their occurrence to a specialized host tissue. For weevils it was shown that intracellular SPE bacteria (Anselme et al. 2008) are still recognized and possibly attacked by the host immune system, when injected into the haemocoel. Thus, the immune system is probably also involved in controlling the symbiont population and prevents symbionts from escaping bacteriocytes. In the past, it has often been argued that intracellular symbionts are shielded from the immune system, because recognition and effector proteins of known innate immune signaling pathways, like Toll or IMD, are secreted into the haemolymph and thus only act on extracellular bacteria (Ratzka et al. 2012b). However, in *Drosophila* it was also shown that autophagy can act as an innate defense mechanism against intracellular pathogens, such as *Listeria monocytogenes*, and is induced independently from Toll and IMD signaling pathways (Yano et al. 2008). Autophagy is a cellular pathway for protein and organelle turnover involving degradation of cell components through the lysosomal system (Amano et al. 2006; Kirkegaard et al. 2004; Mizushima et al. 2008). The PRR PGRP-LE recognizes DAP-type peptidoglycan of intracellular bacteria, which induces autophagy and independently from that also activates the IMD pathway resulting in AMP production (Yano et al. 2008). Thus, autophagy and degradation via the lysosomal system might also be involved in controlling intracellular symbionts (Ratzka et al. 2012b). Accordingly, weevil bacteriocytes highly express several genes involved in vesicular formation and trafficking possibly supporting autophagocytic

processes (Heddi et al. 2005; Vigneron et al. 2012). A recent study showed that *Wolbachia* populations are in fact regulated by autophagy in several distinct symbiotic relationships (Voronin et al. 2012). In aphids several studies furthermore indicated a connection between *Buchnera* degradation and age- and morph-dependent activation of the lysosomal machinery (Nakabachi et al. 2005; Nishikori et al. 2009b). In winged aphids the *Buchnera* density decreased drastically around the final ecdysis (Nishikori et al. 2009b). Simultaneously, the number of lysosome-like acidic organelles increased strongly in the bacteriocytes and also the expression of lysosome-related genes (encoding lysozyme and cathepsin L) was strongly induced, suggesting *Buchnera* degradation proceeds through the lysosomal system (Nishikori et al. 2009b). In addition, a host serine carboxypeptidase was identified that seems to be involved in proteolytic activation of *Buchnera*-degrading enzymes and/or directly in digesting *Buchnera* proteins in the lysosomes of bacteriocytes in alatae (Nishikori et al. 2009a). Degradation of *Buchnera* to satisfy host demands during this important developmental stage most likely enhances host fitness and thereby indirectly also the fitness of *Buchnera* symbionts (Nishikori et al. 2009a).

Intracellular endosymbionts may also be controlled via specialized host AMPs. Several studies revealed overexpression of the AMP gene *coleoptericin-A* (*colA*) in endosymbiont-bearing tissues of *Sitophilus* weevils (Anselme et al. 2008; Heddi et al. 1998; Login et al. 2011; Vigneron et al. 2012). The ColA peptide was even shown to be located inside the SPE cytoplasm and sometimes also found to be attached to the bacterial membrane surface (Login et al. 2011). In antimicrobial activity assays ColA showed bactericidal activity against Gram-positive *Micrococcus luteus* and Gram-negative *E. coli*. Interestingly, at bacteriostatic conditions ColA inhibited cell division and caused giant cell phenotype in *E. coli* resembling the elongated morphology of SPE. ColA was found to interact with bacterial Omps and with GroEL (Login et al. 2011). Presumably, bacterial Omps facilitate entering of ColA to the bacterial cytoplasm. There the peptide causes cell elongation through interaction with GroEL, because *groEL* mutations in *E. coli* are also known to elicit cell gigantism (Chapman et al. 2006). Knock down of *colA* transcripts in larval weevils by RNAi, resulted in resumption of cytokinesis and loss of the elongated form of SPE (Login et al. 2011). As a consequence SPE was able to escape from bacteriocytes and spread through larval tissues. Taken together, the weevil ColA peptide has a special symbiotic function in weevil endosymbiosis by controlling endosymbiont number and location (Login et al. 2011; Login and Heddi 2012).

3.5.4 Protective symbiosis and immune priming by endosymbionts

Downregulation of the host immune response in order to maintain endosymbionts might have negative effects on host immune competence against pathogens. However, endosymbionts have been reported to confer resistance towards pathogens and parasites in several insects (reviewed in (Feldhaar 2011; Haine 2008; Ratzka et al. 2012b)). Prominent examples for such symbiont-mediated resistance have been described in aphids. Secondary endosymbionts,

especially *H. defensa* and partially also *S. symbiotica* and *R. insecticola*, have been shown to protect several aphid host species against attacks from the parasitoids *Aphidius ervi* and *A. eadyi* (Ferrari et al. 2004; Oliver et al. 2010; Oliver et al. 2005; Oliver et al. 2003; von Burg et al. 2008; Vorburger et al. 2010). Recent studies indicated that the protective effect of *H. defensa* depends on the presence and on the variant of a bacteriophage called APSE (*A. pisum* secondary endosymbiont bacteriophage), which infects the symbiont (Degnan and Moran 2008; Moran et al. 2005a; Oliver et al. 2010; Oliver et al. 2009). APSE variants encode different putative toxins with homology to known toxins from vertebrate pathogens. Amongst them are a cytolethal distending toxin, a Shiga-like toxin and a YD-repeat toxin, which possibly target eukaryotic tissue, in this case presumably the developing parasitoid wasp (Degnan and Moran 2008; Moran et al. 2005a). Besides the protection against wasps, *R. insecticola*, another secondary endosymbiont of aphids, has also been shown to enhance host resistance to the aphid-specific fungal entomopathogen *Pandora neoaphidis* (Ferrari et al. 2004; Scarborough et al. 2005). The diverse protective symbiotic interactions of aphids may compensate for the limited humoral immune response resulting from the reduction of the immune gene repertoire (Gerardo et al. 2010).

Symbiont-mediated protection has also been described in several other insect-symbiont systems (Feldhaar 2011; Ratzka et al. 2012b). In *Drosophila hydei* for example, resistance against parasitic wasps is mediated via *Spiroplasma* (Xie et al. 2010). These symbionts furthermore protect *Drosophila neotestacea* against sterilization through a parasitic nematode (Jaenike et al. 2010) and may thus rapidly spread through a host population (Jaenike and Brekke 2011). Endosymbionts could furthermore aid their hosts against fungal and bacterial pathogens through the production of antibiotics, which has already been shown for digger wasps (Kaltenpoth et al. 2005; Kaltenpoth et al. 2010) as well as for fungus-growing attine ants (Currie et al. 1999; Mueller 2012) and pine beetles (Scott et al. 2008). Aside from the protection against pathogens and parasites endosymbionts may also protect their hosts from predators. For example endosymbionts of *Paederus* beetles produce the polyketide pederin, a toxin that protects *Paederus* larvae from wolf spiders by reducing their palatability as prey (Kellner 2003; Kellner and Dettner 1996; Piel et al. 2004).

As aforementioned, *Wolbachia* bacteria are remarkably widespread reproductive parasites of arthropods (Hilgenboecker et al. 2008; Werren et al. 2008). Recently, it has been shown that *Wolbachia* infection renders *Drosophila* more resistant against diverse RNA viruses such as *Drosophila* C virus, Flock House Virus and Nora Virus (Hedges et al. 2008; Teixeira et al. 2008). This antiviral effect of *Wolbachia* infection is also effective in other dipterans like *Culex quinquefasciatus* or *Aedes aegypti* and might therefore be applied to reduce their competence as vector for viral diseases like West Nile, Dengue or Chikungunya (Glaser and Meola 2010; Moreira et al. 2009). Aside from virus protection the *Wolbachia* strain *w*MelPop furthermore inhibits *Plasmodium* development in important *Anopheles* vector species and also has an inhibitory effect on filarial nematodes in *Aedes* mosquitos (Kambris et al. 2010;

Kambris et al. 2009). Thus *Wolbachia*-mediated protection might become a new tool to develop novel control strategies for vector-borne diseases like malaria or filariasis. The inhibitory effects of *Wolbachia* infection against various pathogens seem to result from priming of the host immune system by these bacteria. Host gene expression analyses revealed upregulation of several immune genes in response to *Wolbachia* infection including genes that encode thioester-containing proteins (TEPs), C-type lectins fibrinogen-related proteins (FREPs), cecropins, defensins and a leucine-rich repeat immune protein (LRIM1) (Kambris et al. 2010; Kambris et al. 2009; Moreira et al. 2009). A recent study showed that infection with *Wolbachia* led to increased ROS levels and oxidative stress in *Aedes* mosquitoes (Pan et al. 2012). As a consequence the Toll pathway was activated, which mediates the expression of antioxidant proteins to counteract oxidative stress. Besides, Toll activation also led to expression of AMPs like defensins and cecropins, which were shown to be involved in resistance of *Wolbachia*-infected mosquitoes to Dengue virus infection (Pan et al. 2012). Thus, in certain host species *Wolbachia* infection seems to trigger expression of some host immune genes that prevent colonization of host tissues by other pathogens.

There is emerging evidence that symbionts play an important role in shaping diverse immune functions. Recent studies on *Glossina morsitans* showed that tsetse flies require the presence of endosymbionts (especially *Wigglesworthia*) during larval development for maturation and subsequent proper function of the immune system in adult tsetse flies (Weiss et al. 2012; Weiss et al. 2011). As aforesaid, *G. morsitans* hosts harbour several different bacterial symbionts including *Wigglesworthia, Sodalis* and *Wolbachia* (Aksoy 2000; Attardo et al. 2008). Tetracycline treatment resulted in aposymbiotic tsetse lacking all endosymbionts (Pais et al. 2008; Weiss et al. 2012). Aposymbiotic flies derived from symbiont-cured mothers had a highly compromised cellular immune system and were highly susceptible to *E. coli* infection in comparison to wild-type flies (Weiss et al. 2012). However, immune system development could be partially restored in aposymbiotic offspring, when their symbiont-free mothers were fed with a diet supplemented with *Wigglesworthia* cell extracts (Weiss et al. 2012). Feeding of *Sodalis* cell extracts did not yield the same effect suggesting that specific molecular components of the obligate endosymbiont *Wigglesworthia* exhibit immunostimulatory activity within tsetse hosts (Weiss et al. 2012). This astonishing function of *Wigglesworthia* may have evolved based on the relatively aseptic life style of tsetse flies due to their sterile vertebrate blood diet and viviparous reproduction (Weiss et al. 2011).

3.6 Objectives of this work

In the *Blochmannia*-ant symbiosis, the endosymbiont´s genome as well as its gene expression during host development have already been intensively studied during the last decade (Feldhaar et al. 2007; Gil et al. 2003; Stoll et al. 2008; Stoll et al. 2009; Zientz et al. 2004), whereas gene expression of symbiosis-relevant factors of the host *C. floridanus* has not been investigated, yet. It was suggested that the immune system of the insect host might be involved in controlling the endosymbiosis (Anselme et al. 2008). The ant species *C. floridanus* was chosen as a model organism to possibly reveal new aspects of insect immunity, as on the one hand it harbors an obligate intracellular endosymbiont and on the other hand it exhibits a social lifestyle. Thus, for example endosymbiont-mediated traits, like protection against pathogens, may not only affect the fitness of the individual ant, but rather be important for the fitness of the whole colony. The main aim of this study was to first identify key components of the *C. floridanus* immune system in order to subsequently enable detailed characterization of the ant´s immune response towards pathogens as well as towards its endosymbiont *B. floridanus*.

For this purpose a preferably broad spectrum of immune-relevant genes was identified using suppression subtractive hybridization (SSH) method. Furthermore, a comprehensive examination of *C. floridanus* transcriptome in response to infection was performed via application of Illumina high throughput sequencing technology in order to establish a basis for ongoing detailed analyses of single immune factors and their interaction. Kinetics of several immune genes of *C. floridanus* were determined using quantitative real time RT-PCR. Moreover, genes encoding AMPs and their full length mRNAs were characterized on the molecular level.

An essential question of this study was, if *B. floridanus* is detected via the host immune system, because this is the prerequisite for involvement of the immune system in endosymbiont control. Comparative genomics predicted a reduction in murein and LPS biosynthetic pathways of *B. floridanus* (Zientz et al. 2004) putting into question, if MAMPs of *B. floridanus* could still be recognized as non-self via *C. floridanus*. Immune gene expression was investigated after the injection of *B. floridanus* in order to reveal, whether this endosymbiont, which is usually intracellularly located, can elicit an immune response of its host, when it gets into the haemocoel.

Factors of the host immune system, which might be involved in endosymbiont tolerance and control, were examined via expression analysis of candidate genes known to mediate symbiosis in other insects. The symbiotic function of such genes should be investigated via knockdown analyses. To this end, RNAi technology was adapted for application in *C. floridanus*.

4 Material

4.1 *Camponotus floridanus* colonies

Independent biological replicates of experiments were performed with different *C. floridanus* colonies, which are listed below (Table 1).

Table 1. *C. floridanus* colonies used in this work. A.E. = Annett Endler, C.S. = Christian Strehl

Colony	Original location	Lab. rearing since	Collected by
C03	Florida, Saddlebunch Keys	19.08.2001	A.E., C.S.
C79	Florida, Orchid Island	31.08.2001	A.E., C.S.
C90	Florida, Orchid Island	31.08.2001	A.E., C.S.
C96	Florida, Orchid Island	31.08.2001	A.E., C.S.
C132	Florida, Sugarloaf Shores	05.07.2002	A.E.
C152	Florida, Orchid Island	09.07.2002	A.E.
C264	Florida, Tarpon Springs	23.06.2003	A.E.
C331	Florida, Sugarloaf Shores	01.07.2003	A.E.

4.2 Bacterial strains

The bacterial strains used in the context of this work are listed in the Table 2.

Table 2. Bacterial strains used in this work and their characteristic traits.

Bacterial strain	Genotype / Characteristics	Reference
Escherichia coli DH5α	F-, *endA1*, *glnV44*, *thi-1*, *recA1*, *relA1*, *gyrA96*, *deoR*, *nupG*, Φ80*dlac*ZΔM15, Δ(lacZYA-argF)U169, *hsdR17*(r_{K-} m_{K+}), λ–	Bethesda Research Laboratories (BRL)
E. coli D31	F-, *proA23*, *lac-28*, *tsx-81*, *trp-30*, *his-51*, *rpsL174*(strR), *rfe-229*, *ampCp-1*, Str^r, Amp^r	(Monner et al. 1971)
E. coli HT115(DE3)	F-, *mcrA*, *mcrB*, IN(*rrn*D-*rrn*E)1, *rnc14*::Tn10(DE3 lysogen: lavUV5 promoter -T7 polymerase) (IPTG-inducible T7 polymerase) (RNAse III minus), TET^r	(Timmons et al. 2001)
E. coli HT115(DE3)/L4440-*GFP*	*E. coli* HT115(DE3) carrying plasmid L4440-*GFP*, TET^r, Amp^r	this work
E. coli HT115(DE3)/L4440-*LB*	*E. coli* HT115(DE3) carrying plasmid L4440-*LB*, TET^r, Amp^r	this work
E. coli HT115(DE3)/L4440-*SC2*	*E. coli* HT115(DE3) carrying plasmid L4440-*SC2*, TET^r, Amp^r	this work
E. coli HT115(DE3)/L4440-*LS*	*E. coli* HT115(DE3) carrying plasmid L4440-*LS*, TET^r, Amp^r	this work
Bacillus subtilis DSM 10	test strain	Deutsche Sammlung von Mikroorganismen

Micrococcus flavus	test strain	Deutsche Sammlung von Mikroorganismen
Micrococcus luteus	test strain	gift from Prof. Vilcinskas (University of Gießen)

4.3 Oligonucleotides

4.3.1 For suppression subtractive hybridization (SSH) and rapid amplification of cDNA ends (RACE)

Oligonucleotides used for SSH and RACE were designed according to the instructions of the SMART™ PCR cDNA Synthesis Kit, the PCR-Select™ cDNA Subtraction Kit and the SMART™ RACE cDNA Amplification Kit (Clontech) (Table 3). Each long and short Adaptor 1 and 2R for SSH hybridization were annealed before use. Therefore the corresponding long and short Adaptor were mixed 1:1, heated for 2 min at 70°C and slowly cooled down to room temperature. The Universal Primer Mix A for RACE contained 0.4 µM UniversalPA_Long and 2 µM UniversalPA_Short.

Table 3. Oligonucleotides used for SSH and RACE

Primer name	**Sequence 5´-3´**	**Application**
3´Pho Smart Oligo	AAGCAGTGGTATCAACGCAGAGTACGCGGG-3´phosphorylated	5´Adaptor for 5´RACE cDNA synthesis
Smart Oligo II A	AAGCAGTGGTATCAACGCAGAGTACGCGGG	5´ and 3´Adaptor for SMART cDNA synthesis
3´Smart CDS Primer IIA	AAGCAGTGGTATCAACGCAGAGTAC$T_{(30)}$VN (N=A,C,G or T; V= A,G or C)	
5´PCR Primer II A	AAGCAGTGGTATCAACGCAGAGT	SMART cDNA amplification
Adapter 1 long Adapter 1 short	CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT ACCTGCCCGG	Adaptor 1 and 2R for SSH hybridization
Adapter 2R long Adapter 2R short	CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGT ACCTCGGCCG	
SSH PCR Primer 1	CTAATACGACTCACTATAGGGC	First PCR from SSH
SSH nested PCR 1 SSH nested PCR 2R	TCGAGCGGCCGCCCGGGCAGGT AGCGTGGTCGCGGCCGAGGT	Second PCR from SSH
UniversalPA_Long UniversalPA_Short	CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT CTAATACGACTCACTATAGGGC	Universal Primer Mix A for RACE
M13 F1 M13 R1	GTTTTCCCAGTCACGAC CAGGAAACAGCTATGAC	Screening pGEM-T vector

4.3.2 For *hymenoptaecin* (*hym*) sequences

Full length *C. floridanus hym* gene and mRNA sequence (GenBank Acc. No. HQ315784) were characterized by using the following primers:

Table 4. Primers used for *hym* characterization

Gene	Primer name	Sequence 5´-3´	Tm (°C)	Application
hym	Cfl_HymRA_R5 (GSP1)	GGCTGTGTGCCCTGGAAACTTGTCGAACCC	76	5´RACE
	Cfl_HymRA_F3 (GSP2)	GACCCGATGCGCAAAGGAGGAGCTT	71	3´RACE
	Cfl_Hshortc_F2	GAAGTAGCTTCACAGTAGAAACGAAAA	64	Full length cDNA
	Cfl_Hshortc_R3	TAATATTTTGTGAAACAGCCTCAAA	58	
	Cfl_Hym_seqF1	AGCTTCCAGAGCCCGATGC	62	Sequencing (Seq.) primer
	Cfl_Hym_seqR1	GCCGAAGTTCCTTTCAGC	56	
	Cfl_Hym_introF1	CGCTTTCGCAAATGACTTCT	56	
	Cfl_Hym_5´F1	TGGCCTTATTGTGCGCTATC	58	5´-probe
	Cfl_Hym_5´R1	CGGTTGTGCGACCATTGTTA	58	
	Cfl_Hym_repF1	GGCATCCAGGCTGAAAGAAG	60	Repeat-probe
	Cfl_Hym_repR1	TCCTCCGTAGACATCCGCTG	63	

4.3.3 For *defensins*

C. floridanus defensin-1 (*def-1*, GenBank Acc. No.: JN989495) and *defensin-2* (*def-2*, GenBank Acc. No.: JQ693412) sequences were determined using the following primers:

Table 5. Primers used for *def-1* and *def-2* characterization

Gene	Primer name	Sequence 5´-3´	Tm (°C)	Application
def-1	Cfl_DefRA_R2 (GSP1)	TGCAGTGAGCTGCGCAAGCGCTATGAT	71	5´RACE
	Cfl_DefRA_F2 (GSP2)	TCATAGCGCTTGCGCAGCTCACTGCAT	71	3´RACE
	Cfl_Def_flsF1	ATTCTACTGCAAGATTTGAAAGACG	61	Full length cDNA
	Cfl_Def_flsR1	AATTTCTGGTACATATACATTTGTTGA	59	
def-2	Cfl_Def22RA_R1 (GSP1)	GACGACGCTGCACCAGGCATTTAGC	71	5´RACE
	Cfl_Def22RA_F1 (GSP2)	CACCATCGACGAGCCGCAATACGAC	71	3´RACE
	Cfl_Def22_flsF1	TCATTTCAGGGGTATTCGAGTG	60	Full length cDNA
	Cfl_Def22_flsR1	GCGATTGAGAAATTAAACTAACTGG	61	
	Cfl_Def22_F788	TTTCACTGTGGCAAACACACA	57	Seq. primer

4.3.4 For *tyrosine hydroxylase (TyrOH)* sequences

Full length *C. floridanus TyrOH* mRNA sequence (GenBank Acc. No. EFN71001) was characterized by using the following primers:

Table 6. Primers used for *TyrOH* mRNA characterization

Gene	Primer name	Sequence 5´-3´	Tm (°C)	Application
TyrOH	Cfl_TyHRA_F1	GCCGCCCAGAAGAACCGTGAAATGT	69	5´RACE
	Cfl_TyHRA_R1	TGACGACCAATGTCTCAAAGCGAGCA	68	3´RACE
	Cfl_TyH_flsF1	GCTCGTACGCTCCGTCTAAA	60	Full length cDNA
	Cfl_TyH_flsR3	ATGTCATATGTATCAACCGCCTTTA	61	
	Cfl_TyH_flsF2	GAGACATTGGTCGTCAAGCA	58	
	Cfl_TyHRA_F3	GCCGCCGGTCTCTTGACGGCGCGTGACTT	79	Seq. primer
	Cfl_TyHRA_F4	GATGACCCACCTCACAAATGCCGTCAATA	70	

4.3.5 For quantitative real time RT-PCR (qRT-PCR)

Oligonucleotide pairs for qPCR were designed on the corresponding mRNA sequences with Primer3 v. 0.4.0 (Rozen and Skaletsky 2000) to yield products of 120-140 bp with melting temperatures around 56°C (Table 7).

Table 7. Primers used for qRT-PCR.

Gene name	GenBank Acc. No.	Primer name	Sequence 5´-3´
Housekeeping genes:			
Elongation factor-1 alpha (*EF1a*)	EFN72500	Cfl_EF1_rtF2	ACCCTTGGCGTTAAGCAGTT
		Cfl_EF1_rtR2	CCGGGTTGTAGCCAATCTTT
60S ribosomal protein L32	EFN68969	Cfl_rpL32_rtF1	GCCAACAGGCTTCCGTAAAG
		Cfl_rpL32_rtR1	TGAGCACGTTCGACGATAGC
glyceraldehyde-3-P dehydrogenase (*GAPDH*)	EFN69158	Cfl_GAPDH_rtF1	TTCACGACCATCGAGAAAGC
		Cfl_GAPDH_rtR1	TTGAAGCTCGGATCGTAAGC
60S ribosomal protein L18	EFN68908	Cfl_rpL18_rtF1	AAACCTGGACGGGAGAACTG
		Cfl_rpL18_rtR1	TGCTCGAGCTCTTTCGGTTA
Recognition:			
Gram-negative binding protein (*GNBP*)	HS410970	Cfl_GNBP_rtF1	ATTTCCGGATAACACGACCA
		Cfl_GNBP_rtR1	CGCCGTTCCTATGTTTCCTA
Peptidoglycan-recognition protein-LB (*PGRP-LB*)	EFN73971	Cfl_PGRP-LB_rtF2	ACAGCATGAATTTCGTGAATGC
		Cfl_PGRP-LB_rtR2	GCCATGATGGATCACCACATA
PGRP-SC2	EFN73970	Cfl_PGRP-SC2_rtF1	TTTCATGGTCGGAGAAGACG
		Cfl_PGRP-SC2_rtR1	GTGGTGTTCGAGAGCGGTAG
		Cfl_PGRP-SC2_rtF2	AACTGATCGCTTATGGAGTGG
		Cfl_PGRP-SC2_rtR2	CCAATGAGGCCATGATTGTAT
PGRP-2	EFN70060	Cfl_PGRP-2_rtF2	TGAGGGATGTGGATGGACAC
		Cfl_PGRP-2_rtR2	TTAATTGATGCGCAGCGTTC
PGRP-LE	EFN63542	Cfl_PGRP-LE_rtF1	GAATTCTGCGAGACGCAATC
		Cfl_PGRP-LE_rtR1	CTGCCAATGTATGCGAGACC
Signaling and immune defense:			
Relish (*Rel*)	EFN61437	Cfl_Rel_rtF1	CACCTTTGCAATTAGCTGCTG
		Cfl_Rel_rtR1	ACCTCCTTCGACTGCGATATG
Tollip	EFN63773	Cfl_Tollip_rtF1	GGATCAACAAGCTGCTCTGG
		Cfl_Tollip_rtR1	GGGTCCATCCTTGTCATTCC

Lipoma preferred partner	HS410960	Cfl_lpp_rtF2 Cfl_lpp_rtR2	GGACGGCATACCATTCACTG CTGACCAGGCTCAGGCATAA
MAPKK	HS410965	Cfl_MAPKAP2_rtF1 Cfl_MAPKAP2_rtR1	ACCGAGGAAACCAAGACGAG TTTCCCATTGATGCCAAGAC
Sequestosome-1	HS410961	Cfl_oxidSP_rtF2 Cfl_oxidSP_rtR2	CAAGAGGCGACTAGGCACAG GGCTTATTCTTGGCGGAATC
Serpin (*spn*)	HS410959	Cfl_sp4B_rtF1 cfl_sp4B_rtR1	TGTCTGCTGGTCAGATCGTC ATTGCTGACTGTTTGCAACG
Peroxidase	HS410958	Cfl_Peroxi_rtF1 Cfl_Peroxi_rtR1	TCCCAGGATCGATGGAACTA TTGTTTCTCAGCCGGACTCT
Hymenoptaecin (*hym*)	HS410972	Cfl_HymRT_F4 Cfl_HymRT_R4	TGGCCTTATTGTGCGCTATC AACTTGTCGAACCCTTTTCG
		Cfl_HymRT_F5 Cfl_HymRT_R5	CACAGTAGAAACGAAAACATTCC ATGAAGTTTCCTGGGCACTCG
Defensin-1 (*def-1*)	JN989495	Cfl_Def1_rtF1 Cfl_Def1_rtR1	CGGTAGAGTCTCCGGACTTTT CGCTATGATTAACACCGAAGC
Defensin-2 (*def-2*)	JQ693412	Cfl_Def2_rtF3 Cfl_Def2_rtR3	TGCTCGTTATCTTCGCCACT CAGATTCGATGCCATCTCGT
Punch	HS410964	Cfl_Punch_rtF2 Cfl_Punch_rtR2	GAAGCAAATCGCGATAGCAG AGCATCGTCGAGGTCACTGT
Tyrosine Hydroxylase (*TyrOH*)	HS410957	Cfl_TyHRT_F2 Cfl_TyHRT_R2	GCCCAGAAGAACCGTGAAAT TGTTTGCTTGACGACCAATGTC
Malvolio (*Mvl*)	HS410963	Cfl_Nramp_rtF1 Cfl_Nramp_rtR1	GACAACCCGCAAAAATCAGA GTCAAAAAGGCCATGAGCAA
Serine proteinases:			
Stubble (*stu*)	HS410979	Cfl_PPOaF_rtF2 Cfl_PPOaF_rtR2	CGAATGCGTGCCCTATTATC TGGGTGGTGGCTGTAGATTC
Serine Protease 1	HS410980	Cfl_SerPro_rtF1 Cfl_SerPro_rtR1	TATATCGTTCGTGCCGGTTC GACGCAATAAACCAGCGACT
Snake 1 (*sna*)	HS410982	Cfl_SerPro2_rtF2 Cfl_SerPro2_rtR2	GTCACAGAGCATGCTGATCG ACCTTGCTCGCGTGAATTATC
Snake 2	HS410986	Cfl_SerPro4_rtF2 Cfl_SerPro4_rtR2	CTCACTGCGTGAACAACGTG GTAACGCGGATGCGTTATCA
Snake 3	HS410991	Cfl_SerPro7_rtF1 Cfl_SerPro7_rtR1	AGCTCAGACTCGATGGAACG GTCCCGATCTCGTTCTCGTA
Peroxisomal membrane protein PEX16	HS410981	Cfl_SerPro3_rtF2 Cfl_SerPro3_rtR2	ATGTGCGCTGGAGGAGAAG CCTGCAGCTTCACAGCCTAT
Mite allergen Der f 3	HS410990	Cfl_SerPro6_rtF1 Cfl_SerPro6_rtR1	CACCGATTTGGGATACAACG ATTCGCCCAGGATACCAATC
Stress-response:			
Apolipoprotein D	HS410994	Cfl_ApoD_rtF1 Cfl_ApoD_rtR1	AACACCATCGCCAATCTTCT ATTTCGCAATCTGGTTGTCC
Heat shock protein (*hsp*)	HS410992	Cfl_HSP40_rtF1 Cfl_HSP40_rtR1	ATGCTACCGGATACGACTGG TCCATGACCCGAACGATAAT
Cytochrome P450 6a2	HS410993	Cfl_450moxi_rtF1 Cfl_450moxi_rtR1	TCACTTCGCGACAAGTTGAT AGGTCCCATTGTTCTGCTGT
BI-1-like protein	HS410998	Cfl_BaxI_rtF1 Cfl_BaxI_rtR1	CTTCTTGAGGTGGCGGATAC GATCCGCGTAAAGTGACGA
Metabolism:			
KAT AadAT	HS411004	Cfl_KATAadAT_rtF1 Cfl_KATAadAT_rtR1	TCTTCTATCCTGGCGTCGAA CACCAGAAATCTGGCGATTC

OSBP	HS411006	Cfl_OSBP_rtF2 Cfl_OSBP_rtR2	ATGGTTCGTCCTGTCAAACG TCGCGCAAGTTGAAGAATTT
SCaMC-2	HS411001	Cfl_smc_rtF1 Cfl_smc_rtR1	CCGAGATTTCTCCAGCATGA TTGTGAATTCTTCGGGAACG
Neurochondrin-like protein	HS411005	Cfl_HAD_rtF1 Cfl_HAD_rtR1	GAGACCCAAGGAATGCATCA TAGGATCGTCTCCCGAAAGG
Cytochrome b5	HS411002	Cfl_cb5_rtF1 Cfl_cb5_rtR1	GGTCACTCATCAGACGCAAA TTCGTGTATCCGTTCGTTCA
SERAC1	HS411003	Cfl_cb561_rtF1 Cfl_cb561_rtR1	GGGCATTGCTTCTCTGGTAA TGCCCAGAAAATTGTGAAGG
Translation:			
IF4E	HS411008	Cfl_TIF_rtF1 Cfl_TIF_rtR1	AGCGCGGATAGTGTTATGGA TGTTTGACTGCCAGCTTTGA
IF2C 2	HS411009	Cfl_Argo_rtF2 Cfl_Argo_rtR2	GTCGCGAACGGGAGATTAAC TTTGGGCGGAGGTAATACAC
Lysosomal system, Autophagy:			
Lysozyme c-type 1 (*c-type lyso 1*)	EFN74565	Cfl_Lyso_rtF1 Cfl_Lyso_rtR1	GCCGGTAGAGGGAAAGAAAC GTCATTAGCGCCGTATCCAT
Lysozyme i-type (*i-type lyso*)	EFN71839	Cfl_2Lyso_rtF3 Cfl_2Lyso_rtR3	GCTCAATCACCGTCCAATGA GCCCAGTAGCCCCATGTTAT
Cathepsin L (*cath L*)	EFN65237	Cfl_CathL_rtF1 Cfl_CathL_rtR1	ATAGGTCCGATCTCGGTTGC CCGAATCTGTGCCATAACCA
Lysosomal aspartic protease (*lap*)	EFN61279	Cfl_lap1_rtF1 Cfl_lap1_rtR1	TGAGCCAGGTTTAGCATTCG TGAAAACGGCTTTTGGAACC
Carboxypeptidase vitellogenic-like (*CPVL*)	EFN61703	Cfl_CPVL_rtF1 Cfl_CPVL_rtR1	GAAGGCTACCGTCCAACACA ATTGGCGCAGTCTTTGGATT
Autophagy protein 5 (*apg-5*)	EFN68118	Cfl_Apg5_rtF1 Cfl_Apg5_rtR1	GCGATTTACAAGGGCCAGAT TTCCAACCACATCTCATGCTC
ROS regulation:			
Superoxide dismutase (*SOD*)	EFN700851	Cfl_SOD1_rtF1 Cfl_SOD1_rtR1	TGCGCCACATTGGAGATTTA GGATCAGCGTGAACCACAAC
Dual oxidase (*Duox*)	EFN74201	Cfl_Dual1_rtF1 Cfl_Dual1_rtR1	AATTCCACAAGATCGCAGCA GGAAAGCTGATCTCGCTGGT
Ferritin light chain	EFN61070	Cfl_ferri_rtF1 Cfl_ferri_rtR1	GGTGGAAAGATGGACCTCAA TAACGCTTCTTTGGCGAGTT
Hypothetical proteins:			
hypothetical protein EAG_12212	HS411013	Cfl_hypo1_rtF2 Cfl_hypo1_rtR2	GCCACCATCACCATAACTCC AAGCGGATGTTCAAGCTCAA
hypothetical protein EAG_14061	HS411014	Cfl_hypo2_rtF1 Cfl_hypo2_rtR1	CGATGCATTCCCTGTGAAGT GGGTAACATCGCCTCGTGTA
hypothetical protein 3	HS411017	Cfl_hypo3_rtF1 Cfl_hypo3_rtR1	CTTCAACGAGCGAAGCGATA CTTTGCATCCTGGTTTGCAG
hypothetical protein 4	HS411015	Cfl_hypo4_rtF1 Cfl_hypo4_rtR1	GGATCTTTCGGCCGCTATTA AAGCATCGACGTAGCTTGGA
hypothetical protein 5	HS411016	Cfl_hypo5_rtF1 Cfl_hypo5_rtR1	GAAACCGGTGATTCTCTCGAC TCGCAAGAGAAAGAGGTGTGA
hypothetical protein 6	HS411012	Cfl_hypo6_rtF1 Cfl_hypo6_rtR1	GAACGAAGGACGCGAAAAC GCTCGCCCTCGCTAGTAAT
unknown UTR	HS411011	Cfl_alcam_rtF1 Cfl_alcam_rtR1	TAAACGCTCATCCGTGTCG GCGAAATCGATTCTTCGTCA

4.3.6 For validation of Illumina sequencing results by qRT-PCR

Oligonucleotide pairs for validation of Illumina sequencing results by qRT-PCR were designed on the corresponding mRNA sequences with Primer3 v. 0.4.0 (Rozen and Skaletsky 2000) to yield products of 120-140 bp with melting temperatures around 56°C (Table 8).

Table 8. Primers used for validation of Illumina sequencing results by qRT-PCR.

Gene name	GenBank Acc. No.	Primer name	Sequence 5´-3´
Downregulated genes:			
Scavenger receptor class B member 1 (*scav*)	EFN70524	Cfl_scav_rtF1 Cfl_scav_rtR1	ATGGCTTGAATCCTCGACAG CAAGAAACGGTACGCCAATC
Probable phenoloxidase subunit CG8193 (*POsub*)	EFN74080	Cfl_POsub_rtF2 Cfl_POsub_rtR2	CATAATCTCGGTCACGTTGC GACGAAAGCGTGGAATCTGT
Zinc carboxypeptidase A 1 (*zcp*)	EFN74038	Cfl_zcp_rtF1 Cfl_zcp_rtR1	ACGCGAGATCAAAGGTGTCA TGCTGGTCAGCACTTGATGT
Lipase member H-A (*lip*)	EFN63276	Cfl_lip_rtF1 Cfl_lip_rtR1	ACGCCTTTATTCAGGGCAAG AAAGTATTCCGCTGCCCTGT
Sushi, von Willebrand factor type A(*sushi*)	EFN63579	Cfl_sushi_rtF1 Cfl_sushi_rtR1	TATCTGGAGCGGACACCAAC TGAATCCATTTTGGCAGGAG
Upregulated genes:			
Metalloproteinase inhibitor 3(*MPI*)	EFN65977	Cfl_MPI_rtF1 Cfl_MPI_rtR1	ATAAGGTGGGCCGATACCAC TTTCCGCCACTGGACTCTAG
Suppressor of cytokine signaling 2 (*SOCS2*)	EFN67156	Cfl_SOCS2_rtF1 Cfl_SOCS2_rtR1	TTTGTCAGTGTCACGGCAAG TCGGCAGTCACTCATGTTCA
Transferrin (*transf*)	EFN62546	Cfl_transf_rtF1 Cfl_transf_rtR1	ATTCCAAGCACCCATCAATG TATAAGGTGCTCGCGGCTAA
Parathyroid hormone-related peptide receptor (*PHR*)	EFN73318	Cfl_PHR_rtF1 Cfl_PHR_rtR1	GGTTCACGACCATCAACACA TGATCTTGGAACATGCAAGG
Esterase FE4 (*ester*)	EFN65474	Cfl_ester_rtF1 Cfl_ester_rtR1	CTTCGATGGGTCAAGAGGAA TGGAACAATCCCTTGGACAT
hypothetical protein (*hp67112*)	EFN67112	Cfl_hp67112_rtF1 Cfl_hp67112_rtR1	GGCAGTTTGCAGGAAATACG AAATTACACCGTTGCCTTCG
Peptidoglycan-recognition protein-LB (*PGRP-LB*)	EFN73971	Cfl_PGRP-LB_rtF2 Cfl_PGRP-LB_rtR2	ACAGCATGAATTTCGTGAATGC GCCATGATGGATCACCACATA
Relish (*Rel*)	EFN61437	Cfl_Rel_rtF1 Cfl_Rel_rtR1	CACCTTTGCAATTAGCTGCTG ACCTCCTTCGACTGCGATATG
Defensin-1 (*def-1*)	HS410966	Cfl_Defl_rtF1 Cfl_Defl_rtR1	CGGTAGAGTCTCCGGACTTTT CGCTATGATTAACACCGAAGC
Tyrosine Hydroxylase (*TyrOH*)	HS410957	Cfl_TyHRT_F2 Cfl_TyHRT_R2	GCCCAGAAGAACCGTGAAAT TGTTTGCTTGACGACCAATGTC

4.3.7 For RNA interference (RNAi)

Primers used for RNAi experiments are listed in the Table 9. For *in vitro* dsRNA production T7 promoter sequence (5´-TAATACGACTCACTATAGGGAGA-3´) was added to the 5´end of each primer.

Table 9. Primers used for RNAi experiments.

Gene name	Primer name	Sequence 5´-3´
TyrOH	Cfl_TyrOH_rnaiF2	TAATACGACTCACTATAGGGAGAGTAGACGATGCTCGCTTTGAG
	Cfl_TyrOH_rnaiR2	TAATACGACTCACTATAGGGAGAGGAGATTCCGCTATGGTCTTC
pEGFP	dsGFP_T7_rnaiF1	TAATACGACTCACTATAGGGAGAGTAAACGGCCACAAGTTCAG
	dsGFP_T7_rnaiR1	TAATACGACTCACTATAGGGAGACGGCCATGATATAGACGTTG
	dsGFP_rnaiF1	GTAAACGGCCACAAGTTCAG
	dsGFP_rnaiR1	CGGCCATGATATAGACGTTG
PGRP-LB	Cfl_PGRP-LB_rnaiF1	GATCTTACCAGGACCTTCACA
	Cfl_PGRP-LB_rnaiR1	GCTGTTCTTCTCGATGTCGTTT
	Cfl_PGRP-LB_HincIIF	GAGTCGACGATCTTACCAGGACCTTCACA
	Cfl_PGRP-LB_ApaIR	TTGGGCCCGCTGTTCTTCTCGATGTCGTTT
PGRP-SC2	Cfl_PGRP-SC2_rnaiF1	CGTTGATTTTCACGATCTACTT
	Cfl_PGRP-SC2_rnaiR1	GCTACGTGTGGTGTTCGAGAG
pL4440	L4440_2720_seqF	GCCTTTGAGTGAGCTGATAC
	L4440_312_seqR	GGCGATTAAGTTGGGTAACG

4.4 Media, buffers and other solutions

All chemicals used for production of media, buffers and other solutions were ordered from the companies Applichem, Bio-Rad, Eppendorf, GE Healthcare, Greiner, Invitrogen, Merck, Roche, Roth, Serva and Sigma.

LB medium (agar):
10 g tryptone
5 g yeast extract
10 g NaCl
(15 g agar)
Ad 1l dH_2O
Autoclave sterile.

2x TY medium (agar):
16 g tryptone
10 g yeast extract
5 g NaCl
(15 g agar)
Ad 1l dH_2O
Autoclave sterile.

Ant ringer solution (after M. Obermayer):
Solution A in 900 ml dest. H_2O:
7.4 g NaCl
0.5 g KCl
0.22 g $CaCl_2$
Mix solution A and B and then add:
1.1 g TES
1.2 g trehalose
Adjust to pH 7.0

Solution B in 100 ml dest. H_2O:
0.11 g Na_2HPO_4
0.05 g KH_2PO_4

Ant homogenization buffer:
10 mM Tris/HCl, pH 7.5
60 mM NaCl
10 mM EDTA
0.15 mM spermine
0.15 mM spermidine

Ant solubilization buffer:
0.2 M Tris/HCl, pH 9
30 mM EDTA
2 % SDS

10 x PBS:
80 g NaCl
2 g KCl
11.5 g Na_2HPO_4
2 g KH_2PO_4
ad 1 l dH_2O
Adjust to pH 7.4

10x TBE buffer:
109.3 g Tris
55.62 g boric acid
9.31 g EDTA
Ad 1 l dH_2O

20 x SSPE:
175.3 g NaCl
27.6 g NaH_2PO_4
7.4 g EDTA
ad 1 l dH_2O
Adjust to pH 7.4

RNA Loading Dye:
0.72 ml formamide
0.16 ml 10x MOPS buffer
0.26 ml formaldehyde
0.26 ml RNase-free H_2O
0.1 ml 80% glycerol
Point of a spatula bromophenol blue
Store at -20°C

4x Laemmli buffer:
20 ml glycerine
15 ml 10% SDS
10 ml 2 M Tris HCl, pH 6.8
3.75 ml β-mercaptoethanol
50 mg bromphenol blue
ad 50 ml dH_2O

***Blochmannia* isolation buffer:**
35 mM TrisCl pH 7.6
25 mM KCl
250 mM sucrose
Filtrate for sterilization.

20x SSC buffer:
175.32 g NaCl
88.23 g Trisodium citrate (dihydrate)
Ad 1 l dH_2O (pH 7.0)

100 x Denhardt's solution:
2 % BSA
2 % Ficoll 400
2 % PVP 360

10x MOPS buffer:
41.85 g 3-(N-morpholino)propanesulfonic acid
8.2 g sodium acetate
2.92 g EDTA
Ad 1 l RNase-free H_2O
Adjust to pH 7.0
Protect from light.

10x SDS-PAGE Buffer:
144 g glycine
30.2 g Tris
10 g SDS
Ad 1 l dH_2O

5 Methods

5.1 Ant culture

C. floridanus colonies were kept in artificial plaster nests in climate chambers of the University of Würzburg as described before (Feldhaar et al. 2007).

For the experiments in the context of this work the different developmental stages of *C. floridanus* were defined as follows (Stoll et al. 2008) (Figure 5): very young larvae (several individuals still clumped together, L1), older larvae (about 2-4 mm long, L2), young pupae (before metamorphosis, P1), older pupae (shortly after metamorphosis, still uncolored, P2), late pupae (slightly melanized, shortly before eclosion, P3), young callow workers (not fully melanized, no aggressive behavior, W1), adult workers from the nest surface (fully melanized, age not distinguishable, W2). Furthermore, *C. floridanus* adult workers exhibit a distinct size polymorphism. In the following the two worker castes are referred to as minor and major workers.

Figure 5. Developmental stages of *C. floridanus* (Stoll et al. 2008). The lower row shows the pupal stages after removal of the pupal case.

5.2 Immune-challenge of different *C. floridanus* stages

5.2.1 Production of bacterial solutions with defined concentrations for injection

A single colony from the particular bacterial strain was transferred from a fresh agar plate into 6 ml LB medium (containing 100 µg/ml ampicillin (Amp100) and 100 µg/µl streptomycin (Str100) in case of *E. coli* D31) and incubated at 37°C and 190 rpm overnight (to an OD_{600} of about 2). On the next day an aliquot of the bacterial solution was taken off for determination of the cell concentration and the residual culture was immediately autoclaved. A serial dilution of the aliquot was made up to the dilution factor 10^{-8}. Subsequently, 100 µl of the dilution factors 10^{-6} to 10^{-8} were each plated twice on LB agar plates and incubated at 37°C (30°C in case of *Micrococcus*) overnight. On the next day the colonies per plate were counted

and out of it the initial concentration of the bacterial culture was calculated. The dead-autoclaved bacteria from the initial culture were pelleted through centrifugation for 3 min at 6000 rcf and washed twice with sterile PBS. Bacterial pellets were directly used for picking of *C. floridanus* or diluted with ant Ringer solution to the required concentration for injection experiments.

In order to obtain a defined *B. floridanus*-solution, 100 midguts from pupae were dissected, emptied and homogenized in 2 ml of isolation buffer as described before (Stoll 2009). Bacteria were pelleted by centrifugation for 3 min at 6000 rcf, washed twice with isolation buffer and resuspended in 1 ml PBS. Number of *B. floridanus* cells per ml was estimated by Neubauer haemocytometer counting. Cell concentration was adjusted to 6-8 x 10^9 cells/ml within ant Ringer solution. Contamination with other bacteria is very unlikely, as previous studies did not reveal any secondary microbiota in the midgut of *C. floridanus* (Feldhaar et al. 2007).

5.2.2 Immune-challenge via injection or picking

Injection of *C. floridanus* workers was performed as described before (Geier 2009). Briefly, the solution for injection (0.3-0.5 µl) was heaved up using a micropipette and the micropipette tip was carefully removed. The liquid was then pulled from the micropipette tip into a glass capillary, which was produced in a DMZ-Universal Puller (Zeitz). The glass capillary was put in a plastic hollow cylinder placed in a micromanipulator (Science Products GmbH) for injection. Adult workers were anesthetized on ice and then fixed in a micropipette tip glued on a plastic small bowl. Using the micromanipulator, workers were carefully injected between the first and the second segment of the gaster, whereat it was taken care that no liquid leaked from the puncture after injection. Pupae after metamorphosis (P2, P3) were injected in the same manner through the pupal case. Larvae were injected ventrally on the level of the midgut. For the gene expression studies (see sections 6.6 and 6.7) *C. floridanus* major workers were injected with 0.3 µl ant Ringer solution, either sterile or with heat killed bacteria (*B. floridanus*, *M. flavus* or *S. marcescens*) at 6-8 x 10^9 cells/ml.

For immune-challenge via picking, a minutiae needle (Minutiennadel Sphinx V2A 0,1x12mm, bioform) was glued in the opening of a cannula as a fixture for better handling and directly dipped into a bacterial pellet. Such as by injections adult worker ants and pupae after metamorphosis were picked between the first and second segment of the gaster and larvae were picked ventrally on the level of the midgut.

5.3 RNA isolation

For total RNA extraction from insect tissue anesthetized animals were dissected in PBS and the corresponding tissue (Figure 6) was immediately put into RNAlater® (Ambion), a reagent

that inactivates RNases and stabilizes RNA within tissues. Tissues from same treatment groups were pooled in 500-1000 µl RNAlater and stored at 4°C until RNA isolation (for maximal 3 days). For total RNA extraction from whole workers anesthetized individuals were dissected in PBS, whereat the cuticles were carefully opened at the head, thorax and abdomen in order to allow complete permeation of tissues with RNAlater. Furthermore the poison gland was removed and emptied into PBS before putting the animals into RNAlater. Depending on the experimental design, whole larvae or pupae were directly homogenized (see below) without previous dissection and storage in RNAlater.

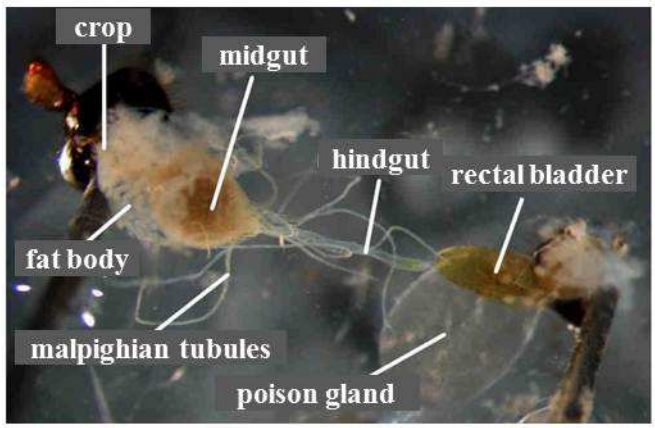

Figure 6. Anatomy of the intestinal tract of a *C. floridanus* worker (modified from (Stoll 2009))

RNAlater-stabilized tissues (up to 50 mg per sample) were transferred into 600 µl buffer RLT (included in the RNeasy mini kit from Qiagen) plus 6 µl β-mercaptoethanol within a 2 ml tube containing lysing matrix D (MP Biomedicals). Samples were homogenized within a FastPrep®-24 Instrument (MP Biomedicals) at 6.0 m/s for 45 s. In order to pellet insoluble debris such as exoskeleton samples were centrifuged at 14 000 rcf for 3 min at 4°C. The supernatant was transferred into a new 2 ml microcentrifuge tube and 1 ml TRIzol® (Invitrogen) or 1 ml TRI Reagent® LS (Sigma) was added to each sample. Then RNA isolation was performed as described in the manufacturer's procedures, whereas the RNA pellet was finally resuspended in 100 µl RNase-free water. The RNA concentration and purity of each sample were measured with a NanoDrop™ ND-100 (Peqlab) spectrophotometer. A 260/280 ratio below 2 indicated a high DNA contamination of the sample. In this case as well as for samples experientially known to be highly DNA-contaminated (e.g. samples from whole larvae) an optional DNase digestion was performed using Turbo DNA-free kit (Ambion) according to the manufacturer's procedures.

After TRIzol®- / TRI Reagent®- RNA isolation samples were often still contaminated with protein or phenol carryovers, which was indicated through a 260/230 ratio below 2. Therefore samples were additionally purified through RNeasy mini kit columns (Qiagen) with on-column DNase digestion (RNase-Free DNase Set, Qiagen) as described in the manufacturer's

procedures. The RNA was eluted in 20-40 µl RNase-free water. RNA concentration and purity of each sample were measured with a NanoDrop™ spectrophotometer. RNA used in real-time qRT-PCR analysis was additionally checked for gDNA contamination by PCR with specific oligonucleotides (about 200-500 ng RNA per PCR reaction). RNA samples were stored at -80°C until further use.

5.4 DNA isolation

Genomic DNA for large scale applications (like Southern blotting) was extracted via phenol/chloroform isolation (Heinze et al. 1994). Whole larvae or workers (up to six individuals per sample) were crushed in liquid nitrogen and then resuspended in 100 µl of ant homogenization buffer. Then 1 µl RNase A (10 mg/ml) and subsequently another 150 µl buffer A were added. After addition of 250 µl of ant solubilization buffer and 5 µl of proteinase K (12.5 mg/ml), the samples were incubated for 2-3 h at 37°C. In order to remove proteins from the DNA samples 250 µl phenol were added under a hood and samples were shaken gently for about 5 min. For phase separation samples were centrifuged in a microcentrifuge at 8000 rpm for 10 min. The aqueous, DNA-containing supernatant was transferred to a new tube without carrying over any of the milky substance at the interface of the phenol and aqueous solution. For high purity of the DNA the phenol extraction step was repeated once again. As a next step 250 µl of chloroform:isoamyl alcohol (1:1) were added and samples were shaken until the solution was milky. After an incubation of 5 min in the -20°C freezer samples were shaken again for 5 s and then centrifuged at 8000 rpm for 10 min. The aqueous supernatant was carefully transferred into a new tube and then adjusted to 1 M NaCl and 1% SDS. An equal volume of chloroform was added and samples were shaken until the solution was milky. After incubation for 5 min at RT and for 5 min at -20°C with occasionally shaking, samples were centrifuged at 8000 rcf for 10 min at 4°C. Finally the aqueous supernatant was transferred into a new tube and the DNA was precipitated by adding two volumes of 100% ethanol and incubation at -20°C overnight. DNA was pelleted by centrifugation at 14 000 rcf for 10 min. The pellet was washed with 700µl 70% ethanol and centrifuged at 14 000 rcf for 5 min. The supernatant was removed as much as possible and the pellets were dried under a hood for 5-10 min. Finally the DNA pellet was dissolved in 30 µl of low TE buffer (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA) and the DNA concentration and purity were checked using a NanoDrop™ spectrophotometer.

Genomic DNA from one individual per sample or from a small amount of tissue was purified through the DNeasy blood and tissue kit (Qiagen) as described in the manufacturer's procedures. DNA samples were eluted in 30-100 µl elution buffer AE and checked photometrically.

5.5 Polymerase chain reaction and agarose gel electrophoresis

DNA target sequences were specifically amplified through polymerase chain reaction (PCR) using MolTaq polymerase (Molzym) for most applications. For a 25 µl reaction 2.5 µl 10x PCR buffer, 0.125 µl MolTaq and 0.5 µl 10mM dNTPs were used with 5-200 ng template DNA. A typical PCR program started with an initial denaturation at 95°C for 3 min followed by 32 cycles with defined temperature changes, which were programmed in a Thermocycler T3 (Biometra). Each cycle consisted of a denaturation step (30 s at 95°C) followed by a primer annealing step (20°s at a primer specific temperature around the lowest melting temperature of the used primer pair) and a DNA synthesis step (1 min per kb at 72°C). In the end a final extension step of 7 min at 72°C was performed.

For applications, in which high sequence accuracy was necessary, and for amplification of long DNA templates (> 3 kb) (e.g. for RACE, SSH or sequencing) PCR was performed using Advantage® 2 Polymerase Mix (Clontech) according to the manufacturer's procedures.

The correct length of the PCR products was checked by TBE agarose gel electrophoresis. 5 µl of each PCR reaction were combined with 1 µl 6x loading dye and electrophoresed on a 1.2 % agarose gel in 0.5x TBE buffer alongside quantitative DNA markers (1kb DNA ladder, Fermentas). DNA fragments were separated according to their size at 140-170 V for 25-40 min. Subsequently, DNA was stained in an ethidium bromide bath (0.5 µg/ml), rinsed with H_2O and photographed under UV light.

5.6 Purification of PCR products

PCR products were cleaned up using the QIAquick PCR purification kit (Qiagen) or using the GeneJET PCR purification kit (Fermentas) as described in the manufacturer's procedures. PCR products were eluted in 20-50 µl of elution buffer. DNA concentration and purity were checked using a NanoDrop™ spectrophotometer and the DNA was stored at -20°C.

5.7 Preparation of competent *E. coli* cells

Chemical competent *E. coli* DH5α (Invitrogen) and HT115(DE3) cells were produced for cloning of different vector plasmids. A single colony from a fresh plate was transferred into 3 ml LB medium (containing 12.5 µg/ml tetracycline (TET12.5) in case of HT115(DE3) cells) and incubated at 37°C and 190 rpm overnight. On the next day 2 x 50 ml LB medium were inoculated each with 500 µl of the overnight culture and incubated at 37°C and 190 rpm until an OD_{600} of 0.5. After 15 min cooling on ice the cells were centrifuged for 10 min at 5000 rpm and 4°C. Each pellet was carefully resuspended in 10 ml ice-cold 0.1 M $CaCl_2$. After incubation for 30 min on ice cells were pelleted at 4000 rpm and 4°C for 10 min.

Finally each pellet was carefully resuspended in 2.5 ml 0.1 M $CaCl_2$ with 20% glycerin and aliquots of 250 µl were stored at -80°C.

5.8 Ligation of PCR fragments into pGEM®-T vector and transformation into competent *E. coli* cells

The pGEM®-T vector is a linearized vector with 3´-T overhangs at both ends. The Taq polymerases used in this work characteristically add 3´-A overhangs to the amplified DNA fragments in a template-independent fashion. Therefore the 3´-A overhangs of PCR templates can be directly ligated to the 3´-T overhangs of the pGEM®-T vector without previous restriction. Fragments are inserted into the *lacZ* gene of the plasmid allowing blue/white screening of the produced recombinant clones on indicator plates.

Ligation was performed as described in the manufacturer´s procedures with a 3-8 molar ratio of PCR product: vector. After incubation for 1-2 h at room temperature the ligation reaction was carefully mixed with 100 µl chemical competent *E. coli* DH5α cells and incubated for 20 min on ice. Cells were heat-shocked for 50 s at 42°C and then immediately cooled on ice for 2 min. After addition of 900 µl LB medium the culture was incubated for 1.5 h at 37°C and 190 rpm. LB Amp100 agar plates were applied with 10 µl 1 M IPTG and 50 µl 2% X-Gal (solved in DMSO). Subsequently, 50-100 µl of the transformation culture was spread on each plate and incubated at 37°C overnight. On the next day white colonies were smeared on fresh LB / Amp100 agar plates and simultaneously screened for insert with correct size via PCR with vector specific primers M13 F1 (5´-GTTTTCCCAGTCACGAC-3´) und M13 R1 (5´-CAGGAAACAGCTATGAC-3´).

5.9 Isolation of plasmid DNA from *E. coli* cells

For isolation of a small amount of plasmid DNA (about 6-10 µg) mini-plasmid-preparations were carried using the UltraPrep Plasmid Mini (Molzym) or the AxyPrep™ Plasmid Miniprep (Axygen) kit as described in the manufacturers´ protocols. Plasmid DNA was isolated from 4 to 6 ml overnight culture (grown with the corresponding antibiotics) and finally eluted in 30-50 µl elution buffer.

If a larger amount of plasmid DNA was required (about 20-100 µg), midi-plasmid-preparations from 200 ml overnight culture were performed with the Nucleobond kit (Machery-Nagel) according to the manufacturer´s procedures. Plasmid DNA concentration and purity were measured photometrically and plasmid DNA was stored at -20°C.

5.10 DNA sequencing of purified PCR products and plasmids

Sequencing of PCR products and plasmids was done by the company Seqlab in Göttingen. For sequencing of PCR products about 250 ng per kb and for plasmids about 700 ng DNA were mixed with 20 pmol of a specific primer in a total volume of 7 µl and sent per Post to Seqlab. The obtained sequences were analyzed using Bioedit, FastPCR or NCBI Blast.

5.11 Construction of subtracted cDNA libraries using suppression subtractive hybridization (SSH)

Suppression subtractive hybridization (SSH) is a suitable method to identify differentially expressed genes between two cDNA pools (Diatchenko et al. 1996). The basic principle of SSH is to hybridize cDNA that contains differentially expressed transcripts (tester) with a reference cDNA (driver) and then suppress amplification of hybrid sequences. The remaining unhybridized cDNAs represent genes that are specifically expressed in the tester but not in the driver mRNA. In the context of this work cDNA samples from immune-challenged animals (tester) were compared with cDNA pools from untreated animals (driver) in order to obtain immune-relevant genes.

5.11.1 RNA sample preparation

For the generation of the subtractive libraries *C. floridanus* major workers were injected with ant Ringer solution containing either a mix of heat killed *Escherichia coli* and *Micrococcus flavus* or heat-killed *Serratia marcescens* at 6-8 x 10^9 cells/ml. Ants (for tester cDNAs) were anesthetized on ice and 0.3µl of bacteria solution was injected into the gaster. Control ants (for driver cDNAs) were cooled on ice, but not inoculated. After injection workers were kept in artificial nests until dissection. Non-inoculated workers were treated in parallel as controls. In case of *E. coli*- and *M. flavus*-challenge, midgut and fat body of eight individuals per sample were dissected in sterile PBS 6 hours (h) after injection and kept in RNAlater (Ambion) until RNA isolation. After *S. marcescens*-challenge 2 individuals each were dissected at 3, 6, 9 and 12 h after challenge and their midguts and fat bodies were pooled before RNA-isolation.

For each sample total RNA from midgut and fat body of eight major workers was extracted using TRIzol® Reagent (Invitrogen) and purified through RNeasy mini kit columns (Qiagen) with on-column DNase digestion (RNase-Free DNase Set, Qiagen) as described in section 5.3. A subtractive library from midgut / fat body tissue in response to *E. coli*- and *M. flavus*-challenge 6 h after injection and a midgut / fat body library in response to *S. marcescens*-challenge at 3, 6, 9 and 12 h after injection was generated. Midgut tissue was chosen for

library construction as the bacteriocytes are intercalated between midgut cells and fat body as most of the immune gene transcription takes place in this tissue.

5.11.2 SMART™ cDNA synthesis and cDNA amplification

SMART™ (Switching Mechanism at 5' End of RNA Template) is a unique technology developed by Clontech, which allows the efficient incorporation of known adaptor sequences at both ends of cDNA during first strand synthesis. A modified oligo(dT) primer (containing the sequence of the SMART oligonucleotide as an adapter overhang) primes the first-strand synthesis reaction. When the reverse transcriptase (RT) reaches the 5´end of the mRNA, it adds a few additional deoxycytidines to the 3´-end of the cDNA allowing base-pairing with the oligo(dG) sequence of the SMART™ II A oligonucleotide. RT then switches templates and replicates to the end of the oligonucleotide (Chenchik et al. 1998). Accordingly, the resulting full-length cDNA contains sequences that are complementary to the SMART oligonucleotide at both ends. These SMART adaptor sequences serve as universal priming sites for end-to-end cDNA amplification. The terminal transferase activity and subsequent template-switching of RT occur preferentially at the 5'-cap structure of eukaryotic mRNAs. Thus, cDNAs from truncated mRNAs generally do not incorporate the SMART oligonucleotide and are not amplified during PCR. Accordingly, the amplified SMART cDNA is enriched for full-length cDNA.

In this work driver and tester cDNAs for both libraries were synthesized in each case from 1 µg total RNA using the SMART™ PCR cDNA Synthesis Kit (Clontech) according to the protocols of the manufacturer. Produced cDNA was amplified in 19 PCR cycles using Advantage® 2 Polymerase Mix (Clontech) and purified via CHROMA SPIN™ 1000 Columns as described in the manual of the SMART™ PCR cDNA Synthesis Kit (Clontech).

5.11.3 Rsa I digestion and adapter ligation

Purified cDNA was digested with 20 u RsaI (Fermentas) in 1x Tango restriction buffer (Fermentas) for 3-5 h at 37°C in order to generate shorter, blunt-ended cDNA fragments, which are required for adapter ligation as well as for subtraction. After successful digestion cDNA was purified using the QIAquick PCR purification kit (Qiagen) and subsequently precipitated by addition of ½ volume of 4 M ammonium acetate and 2.5 volumes of 100 % ethanol. Samples were thoroughly mixed by vortexing and then centrifuged at 14 000 rpm for 20 min at room temperature. Supernatant was carefully removed and pellets were washed with 500 µl of 80 % ethanol. After centrifugation at 14 000 rpm for 10 min supernatant was removed and pellets were dried in a SpeedVac Centrifuge (Concentrator 5301, Eppendorf) for 5-10 min. Pellets were dissolved in 6.7 µl of 1x TNE (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.1 mM EDTA) buffer and the concentration of a 1:10 dilution was measured

photometrically. Samples were then diluted to a final DNA concentration of 300 ng/µl for adaptor ligation.

Each cDNA was aliquotted into two tubes, whereas one aliquot was ligated with Adaptor 1 and the other with Adaptor 2R. Adaptor ligation for forward subtraction was performed at 16°C overnight as described in the manual of the PCR-Select™ cDNA Subtraction Kit (Clontech) using T4 DNA Ligase (Fermentas). After inactivation of the ligase at 72°C for 5 min, ligation efficiency was analyzed according to the PCR-Select™ cDNA Subtraction protocols.

5.11.4 First and second hybridization

Next in the SSH procedure two hybridizations are performed (Diatchenko et al. 1996). In the first hybridization reaction an excess of driver cDNA is added to each of the two tester cDNA pools (ligated to Adaptor 1 or 2R). Samples are then heat denatured and allowed to anneal leading to enrichment and equalization of differentially expressed single stranded cDNA molecules with Adaptor 1 or 2R sequence at one end. The two primary hybridization samples are then mixed together without denaturing. Thus, during the second hybridization the differentially expressed single stranded cDNAs can reassociate and form new hybrid molecules with different Adaptor sequences (1 and 2R) at both ends.

When adaptor ligation was successful, hybridization reactions were performed as described in the protocols of the PCR-Select™ cDNA Subtraction Kit (Clontech). Preliminary screening of the subtracted libraries showed the redundant representation of cDNA fragments with sequences from *hymenoptaecin* (*hym*) and *tyrosine hydroxylase* (*TyrOH*) gene. Therefore new SSH experiments were performed, in which these cDNA fragments were artificially added to the driver cDNA pool prior to hybridization in order to suppress their amplification during the following PCR steps and to enrich the library for other differentially expressed genes. For that purpose full length cDNAs from *hym* and *TyrOH* gene were amplified from cDNA using specific primer pairs Cfl_Hym_flsF1 & Cfl_Hym_flsR1 as well as Cfl_TyrOH_flsF1 & Cfl_TyrOH_flsR1. PCR products were digested with RsaI (Fermentas) and 10 ng of digested and purified fragments from each gene were added to driver cDNAs before hybridizations.

5.11.5 First and second PCR amplification

After second hybridization cDNA samples are used as template for PCR with Adaptor specific primers. During first PCR only the desired, differentially expressed cDNAs with two different adaptors amplify exponentially, while amplification of background molecules is suppressed through formation of pan-like secondary structures that prevent primer annealing (Siebert et al. 1995). Finally, a second PCR amplification with nested primers 1 and 2R is

performed to further reduce background PCR products and enrich for differentially expressed sequences.

For both subtracted libraries first PCR amplification was performed with 27 cycles and second PCR with 10 cycles according to the PCR-Select™ cDNA Subtraction protocols. PCR products from the second PCR were purified with the PCR Purification Kit (Qiagen), inserted into the plasmid vector pGEM® (Promega) and transformed into chemically competent *E. coli* DH5α cells (Invitrogen).

5.12 Differential screening of subtracted cDNA libraries via colony blotting

A total of 320 clones from each subtracted library were differentially screened using dot blots. For each blot 80 clones were selected randomly and screened for differentially ESTs, principally as described in the PCR-select cDNA subtraction screening kit (Clontech). Clones were grown in 100 µl of LB Amp100 for 6 h at 37°C with moderate shaking in 96-well plates. Two microliters of bacterial culture were spotted in duplicate on Amersham Hybond-N^+ membranes (GE Healthcare, UK) and allowed to grow overnight at 37°C on a LB Amp100 agar plate. Membranes were then denatured in 0.5 M NaOH; 1.5 M NaCl for 4 min, neutralized in 1.5 M NaCl; 0.5 M Tris/HCl pH 7.5 for another 4 min and allowed to dry for 30 min at room temperature. Nucleic acids were fixed to the membrane by baking for 2 h at 80°C.

As probes for hybridization 100 ng of purified secondary PCR products of subtracted tester or unsubtracted driver were radioactively labelled with 60 mCi of [α-^{33}P]-dATP with the DecaLabel Kit (Fermentas). Labelled cDNA probes were purified with Illustra Microspin S-200 HR Columns (GE Healthcare, UK). For each new blot the already obtained differentially ESTs were artificially added to the unsubtracted driver probes before radioactive labeling in order to avoid redundant sequencing.

Spotted membranes were equilibrated for 5 min in 2x SSPE and prehybridized in roller bottles with 10 ml of hybridization buffer (5x SSPE, 2 % SDS, 1x Denhardt's solution, 100 mg/ml salmon sperm DNA) for 2 h at 62°C. In each case two membranes, spotted with DNA from identical clones, were hybridized in individual tubes with subtracted tester and unsubtracted driver probes. After addition of the heat-denatured probes (5 min at 95°C, cooled down on ice) hybridization was continued for 20 h at 62°C in a rotatory oven. Membranes were rinsed and then washed three times for 20 min at 62°C in 25 ml of wash buffer (0.5x SSPE, 0.2 % SDS), sealed in saran wrap and exposed to a storage phosphor screen (GE Healthcare, UK) for 2 days. Screens were scanned on a Typhoon 9200 Variable Mode Imager (GE Healthcare, UK) with a resolution of 50 microns.

Colonies that showed a strong signal with the subtracted tester cDNA probe and no or only a low signal with unsubtracted driver cDNA probe were chosen for colony PCR with vector specific primers M13 F1 and M13 R1 using PCR cycler and the MolTaq PCR system (Molzym). PCR conditions were as follows: denaturation at 95°C for 3 min followed by 32 cycles of denaturation at 95°C for 15 s, annealing at 50°C for 15 s, and extension at 72°C for 2 min 30 s. PCR products were purified with the PCR Purification Kit (Qiagen) and sequenced. The sequences were generated by Seqlab Sequence Laboratories Göttingen with the vector primers M13 forward and reverse. After removal of flanking vector sequences and adapters, Mega BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) was used to identify corresponding gene sequences in the *Camponotus* genome and homologous sequences from other insects. For rapid function annotation and a general overview on the proteins encoded by the ESTs an analysis of the encoded COGs (clusters of orthologous groups) was applied (performed by Chunguang Liang) (Tatusov et al. 2003), using the routine COGmaster (Huson and Bryant 2006) which also includes annotation of eukaryotic COGs (KOGs).

The differential expression of the genes identified via SSH was additionally proven by quantitative real time RT-PCR (qRT-PCR) on three independent biological samples using cDNA derived from pooled total RNA from midgut and fat body of eight major workers per sample.

5.13 Southern blotting

Digested DNA from immune-challenged and healthy workers (about 30 µg per lane) was separated on a 1.0 % agarose gel and transferred onto a nylon blotting membrane (Amersham Hybond N^+, GE Healthcare, UK). Membranes were prehybridized in roller bottles with 10 ml of hybridization buffer (0.5 M $NaHPO_4$ (pH 7.4), 1 mM EDTA, 0.7 % SDS) for 30 min at 65°C. A hymenoptaecin repeat-fragment of 270 bp was amplified with the primers Cfl_Hym_repF1 and Cfl_Hym_repR1 (Table 4) and used as hybridization probe. This hymenoptaecin repeat-probe was labeled with Rediprime II DNA Labeling System (Amersham) according to the manufacturer's specifications. After addition of the heat-denatured probe (5 min at 95°C, cooled down on ice) hybridization was continued for 20 h at 62°C in a rotatory oven. Membranes were rinsed and then washed two times for 30 min at 65°C in 20 ml of wash buffer (0.04 M $NaHPO_4$ (pH 7.4), 1 mM EDTA, 0.5 % SDS), sealed in saran wrap and exposed to a storage phosphor screen (GE Healthcare, UK) for 2 days. Screens were scanned on a Typhoon 9200 Variable Mode Imager (GE Healthcare, UK) with a resolution of 50 microns.

5.14 Northern blotting

Each 25 µg total RNA from immune-challenged and untreated workers were concentrated to a volume of 5 µl within a SpeedVac centrifuge and then mixed with 15 µl RNA Loading Dye. After denaturation for 10 min at 65°C, samples were separated in 1x MOPS buffer on a 1.0 % formaldehyde agarose gel alongside to a size marker (RiboRuler™ High Range RNA Ladder, Fermentas). RNA was transferred onto a nylon blotting membrane (Amersham Hybond N^+, GE Healthcare, UK). Membranes were prehybridized in roller bottles with 10 ml of Amersham Rapid-hyp Buffer (GE Healthcare, UK) for 30 min at 65°C. A hymenoptaecin 5´-fragment of 270 bp was amplified with the primers Cfl_Hym_5´F1 and Cfl_Hym_5´R1 (Table 4) and used as hybridization probe. This hymenoptaecin 5´-fragment was radioactively labeled with 60 mCi of [α-^{32}P]-dATP with the DecaLabel Kit (Fermentas, Germany) according to the manufacturer's specifications. The labeled cDNA probe was purified with Illustra Microspin S-200 HR Columns (GE Healthcare, UK). Membranes were rinsed and then washed two times for 15 min at 65°C in 50 ml of wash buffer (2 x SSC, 0.1 % SDS), sealed in saran wrap and exposed to a storage phosphor screen (GE Healthcare, UK) for 2 days. Screens were scanned on a Typhoon 9200 Variable Mode Imager (GE Healthcare, UK) with a resolution of 50 microns.

5.15 Rapid amplification of cDNA ends (RACE)

SMART™ technology (Clontech) (see section 5.11.2) can also be used to perform both 5´- and 3´-rapid amplification of cDNA ends (RACE). For that purpose two separate cDNA populations are created from one RNA sample. The 5´RACE cDNA is synthesized using Smart II A TS-Oligo-3´P, which contains a oligo(dG) sequence for template switching and is phosphorylated at the 3´-end preventing unwanted priming of this oligonucleotide (Dai et al. 2007). The 3´RACE cDNA is synthesized using the 3´SMART CDS Primer A, which is a modified oligo(dT) Primer with the SMART sequence at its 5´end. Accordingly, resulting 5´RACE cDNA has the SMART adaptor sequence incorporated at its 5´end and 3´RACE cDNA at its 3´end. Thus, the 5´- and 3´-end of a specific transcript can be amplified from the corresponding cDNA with a gene specific primer and the adaptor sequence primer (Universal Primer Mix A). The 5'- product and the 3'- product are then sequenced to obtain the sequences of the extreme ends of the transcript. With this information 5'- and 3'-gene-specific primers can be designed to amplify the full-length cDNA.

5.15.1 cDNA synthesis for RACE

The procedure for RACE cDNA synthesis was adapted from the SMART™ RACE cDNA Amplification Kit (Clontech). The two first-strand cDNA populations used for 5' and 3'-RACE were synthesized each from 1 µg of total RNA from *Serratia*-injected workers

prepared for the SSH method (tester). For 5´RACE cDNA synthesis 1 µg template RNA was mixed with 1 µl oligo(dT)$_{20}$ primer (Invitrogen) and 1 µl Smart II A TS-Oligo-3´P (adapted from Clontech and (Dai et al. 2007)) in a total volume of 4 µl. For 3´RACE cDNA synthesis 1 µg template RNA was mixed with 1 µl 3´RACE CDS Primer A (Clontech) in a total volume of 4 µl. For denaturation of RNA secondary structures both samples were incubated for 2 min at 70°C within a Thermocycler T3 (Biometra) and then immediately cooled on ice for 2 min. For reverse transcription 2 µl 5x First Strand Buffer (Promega M531A: 250mM TrisHCl, 375mM KCl, 15mM $MgCl_2$, 50mM DTT), 0.25 µl RNaseOut (10 u) (Invitrogen), 1 µl 10 mM dNTPs, 1 µl BSA (2 mg/ml), 0.2 µl 100 mM $MnCl_2$ and 0.75 µl Superscript II RT (Invitrogen) were added to each tube. $MnCl_2$ and BSA ensured 5´cap-dependent addition of (dC)-nucleotides by RT, which is essential for annealing of the TS-oligo in 5´RACE cDNA synthesis (Schmidt and Mueller 1999). After incubation for 1.5 h at 42°C in a Thermocycler T3 (Biometra) samples were diluted with 100 µl Tricine-EDTA (10 mM Tricine-KOH pH 8.5, 1 mM EDTA) and subsequently heated for 7 min at 72°C. Produced cDNA samples were stored at -20°C.

5.15.2 RACE-PCR

The complete sequences of the transcripts of interest were obtained by 3'- and 5'- RACE, performed with the SMART RACE cDNA Amplification Kit including the Advantage II PCR kit (Clontech). For *TyrOH*, *hym* and *def-1* the nucleotide sequences of the 3'- and 5'- primers (GSP1 and GSP2) were designed on the corresponding EST (GenBank Acc. No. for EST from *TyrOH*: HS410957, for EST from *hym*: HS410972 and for EST from *def-1*: HS410966) obtained from the SSH experiments. For *def-2* RACE primers were designed according to the *C. floridanus* genome sequence. RACE PCR was carried out using touchdown PCR program and Universal Primer Mix A according to the instructions of the SMART™ RACE cDNA Amplification Kit (Clontech).

5.15.3 Amplification of full length cDNAs

5'- and 3'-gene-specific primers were designed from the obtained RACE cDNA sequences and used for PCR amplification of the full length cDNAs und genes. In each case resulting PCR-products were purified with the PCR Purification Kit (Qiagen), inserted into the plasmid vector pGEM®-T (Promega) and transformed into *E. coli* DH5α cells (Invitrogen). The plasmids from several (at least three) different clones for each gene were then extracted for sequencing with the UltraPrep Kit (Molzym). The sequences were generated by Seqlap with the vector-specific (M13 F1 and M13 R1) as well as with gene-specific primers.

5.16 Quantitative real time RT-PCR (qRT-PCR)

5.16.1 cDNA synthesis for qRT-PCR

For the gene expression studies RNA extractions of four to six independent biological samples were carried out for each condition. In each case 500 ng - 2 µg of total RNA were reverse-transcribed with oligo $(dT)_{18}$ primers using the RevertAid First Strand cDNA Synthesis Kit (Fermentas) according to the manufacturer´s protocols. Resulting cDNA was purified with the QIAquick PCR purification kit (Qiagen) and eluted in 50-100 µl of buffer EB (10 mM TrisHCl, pH 8.5) or directly diluted to a final cDNA concentration of 20-10 ng/µl. Produced cDNA was stored at -20°C.

5.16.2 qRT-PCR using the ddCt method

Oligonucleotide pairs for qRT-PCR were designed on the corresponding mRNA sequences with Primer3 v. 0.4.0 (Rozen and Skaletsky 2000) to yield products of 120-140 bp with T_m values around 56°C (Table 7 and Table 8).

For gene expression studies qRT-PCR experiments were performed using the Absolute™ QPCR SYBR® Green Mix (Thermo Fisher Scientific Inc., Epsom, Surrey, UK) or the PerfeCta™ SYBR® Green FastMix™, Rox (Quanta Biosciences) on a StepOnePlus™ Real-Time PCR System (Applied Biosystems). Both PCR master mixes contain the fluorescent dye SYBR Green I, which incorporates into the DNA produced during PCR. The fluorescence of each reaction batch is proportional to the generated amount of DNA and is measured after each cycle. For comparison of different samples the so-called Ct- (cycle threshold) value of each sample is determined, which is the number of cycles at which the fluorescence significantly exceeds the background fluorescence. With these Ct-values the relative transcription level of one sample (treatment) in comparison to another sample (control) can be calculated using the ddCt method (Livak and Schmittgen 2001). Briefly, the Ct-value of a constitutively expressed reference gene (housekeeping gene) is subtracted from the Ct of the target gene (= dCt-value) for standardization between different samples. Next, the dCt-values from the control samples are subtracted from the treatment samples to gain ddCt-values. The ratio of the relative expression between control and treatment sample can then be calculated by the arithmetic formula 2^{-ddCt}. If no difference in gene expression occurs, the ratio is equal to one.

When using the PCR master mix from Thermo Fisher Scientific, qRT-PCR samples contained 1x ABsolute™ QPCR SYBR® Green Mix, gene-specific oligonucleotides in a concentration of 100 nM each, 1 µl of the cDNA and water to a final volume of 25 µl. After 15 min of enzyme activation at 95°C, 40 cycles of 15 s denaturation at 95°C, 60 s of annealing at 56°C and 30 s of extension at 60°C were run.

Utilization of the FastMix™ from Quanta Biosciences allowed shortening the different steps of the PCR cycles. In this case, the enzyme was initially activated for 5 min at 95°C and then 45 cycles of 5 s denaturation at 95°C, 10 s of annealing at 56°C and 20 s of extension at 60°C were run. qRT-PCR samples contained 1x ABsolute™ PerfeCta™ SYBR® Green FastMix™ (Rox), gene-specific oligonucleotides in a concentration of 250 nM each, 1 µl of the cDNA and water to a final volume of 20 µl.

Fragment specificity was checked in a melting curve. Each biological sample was run in duplicate in the qRT-PCR and results were averaged. The relative transcription levels were calculated using the ddCt method (Livak and Schmittgen 2001). The results were standardized using the expression level of a constitutively expressed reference gene. No differences in reference gene expression levels were found between treatments, time points or developmental stages.

5.16.3 Reference gene selection

The prerequisite for internal standardization of target gene expression levels is the stable expression of the reference gene. *BestKeeper* is an excel-based tool using pair-wise correlations, that determines the best suited housekeeping gene from up to ten candidate genes and combines them into an index (Pfaffl et al. 2004).

The suitability of the four reference genes *ribosomal protein L32* (*rpL32*), *ribosomal protein L18* (*rpL18*), *elongation factor 1-alpha* (*EF1α*) and *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) was determined using *BestKeeper* software tool (Pfaffl et al. 2004) on the basis of their expression levels within different tissues (midgut and residual body parts) and developmental stages (L2, P3, W2) of *C. floridanus* (see section 6.8.1). All four candidate genes were considered as suitable genes for normalization in *C. floridanus* and were therefore used as internal standards in other experiments.

5.16.4 Statistical analyses of qRT-PCR data

Statistical analyses were performed with Statistica 9.0.

For the studies in section 6.6 and 6.7 average gene expression levels relative to the constitutively expressed *EF1α* gene (dCt = Ct(target) - Ct(*EF1α*)) of four independent samples were determined from bacteria injected and untreated ants at each time point and tested for significant differences using two sided t-tests. In addition normalized changes in gene expression in injected animals (-ddCt = -(dCt(injected) - dCt(untreated)) were calculated from four independent samples for each treatment and time point. A two factorial ANOVA followed by Tukey´s HSD post hoc test was performed on these (-ddCt)-values in order to investigate the influence of the factors time and injected bacterial species.

For the studies in section 6.8 and 6.9 average gene expression levels relative to the *BestKeeper* index or to *rpL32* (dCt = Ct(target) - Ct(internal standard)) of six independent samples were determined from the midgut tissue and the residual body of different stages. A two factorial ANOVA followed by Tukey´s HSD post hoc test was performed on these (dCt)-values in order to investigate the influence of the factors tissue and stage or tissue and immune-challenge, respectively.

5.17 Transcriptome analysis using Illumina sequencing technology

5.17.1 Preparation of RNA samples for Illumina sequencing

Dead-autoclaved *E. coli* D31 and *M. luteus* were mixed in a 1:1 proportion and cells were pelleted by centrifuging for 3 min at 6000 rcf. The pellet was washed twice with sterile 1x PBS and hold on ice during the experiment. Adult minor workers (W2) and larvae (L2) were picked with a minutiae needle (bioform) dipped into the bacterial pellet. After the immune-challenge animals were kept in artificial nests for 12 h. Non-inoculated animals were treated in parallel as controls. 12 h after immune-challenge RNA from 5 picked and 5 untreated workers as well as from 5 picked and 5 untreated larvae was extracted in separate samples using TRIzol® Reagent (Invitrogen) and purified through RNeasy mini kit columns (Qiagen) with on-column DNase digestion (RNase-Free DNase Set, Qiagen) as described above (see section 5.3). RNA of each sample was eluted in 30 µl RNase free H_2O, a 2 µl aliquot was taken off for quality control and the samples were immediately frozen at -80°C. Aliquots were checked for contamination of gDNA via PCR. Total RNA was quantified and integrity assessed on an Agilent 2100 Bioanalyzer using the Agilent RNA 6000 Nano Chip kit (Agilent Technologies) according to the manufacturer´s instructions. Each 7.5 µg total RNA from immune-challenged workers and larvae as well as each 5 µg from untreated workers and larvae were mixed and the two resulting RNA samples were sent to Eurofins MWG GmbH for sequencing. The sequencing company prepared an amplified short insert cDNA library of 150-250 bp size for each sample and then sequenced these 2 samples in 2 channels of Illumina HiSeq 2000 using 2x 50 bp paired-end sequencing.

5.17.2 Analysis of sequencing data

The sequencing data was obtained from Eurofins MWG GmbH and subsequently analyzed by Frank Förster (Department of Bioinformatics, University of Würzburg). In short, the 2x 50 bp paired-end reads from both conditions (untreated and immune-challenged) were aligned to the *C. floridanus* genome sequence using TopHat (http://tophat.cbcb.umd.edu/) (Trapnell et al. 2009; Trapnell et al. 2012). Then Cufflinks (http://cufflinks.cbcb.umd.edu/) was used to assemble the reads into transcripts (Trapnell et al. 2012; Trapnell et al. 2010). Differentially expressed genes between the two conditions were identified using Cuffdiff, a part of the

Cufflinks package. “Cufflinks and Cuffdiff implement a linear statistical model to estimate an assignment of abundance to each transcript that explains the observed reads with maximum likelihood” (Trapnell et al. 2012). Additionally, *DESeq* package was used to identify genes, which are differentially expressed after an immune-challenge (Anders and Huber 2010). The *DESeq* software provides a method to test for differential gene expression based on the negative binonial distribution and a shrinkage estimator for the distribution's variance (Anders and Huber 2010).

5.17.3 Validation of Illumina sequencing results by qRT-PCR

Six RNA samples from untreated and immune-challenged animals were prepared in each case from L2 and W2, as described above (see section 5.17.1). In each case cDNA was reverse-transcribed from 1 µg of total RNA and gene transcripts were analyzed by qRT-PCR (see sections 5.16.1 and 5.16.2). Five downregulated and ten upregulated candidate genes from the Illumina sequencing results were chosen for validation by qRT-PCR (for primers see section 4.3.6, Table 8). Average gene expression levels relative to the constitutive *EF1α* gene (dCt = Ct(target) - Ct(*EF1α*)) were determined separately in samples from untreated and immune-challenged L2 and W2, respectively. Comparisons of the dCt-values between the six samples from untreated and immune-challenged animals were made using two-sided t-tests for samples from L2 and W2, respectively. The ratio of the gene relative expression between samples from untreated and from immune-challenged animals was calculated by the arithmetic formula 2^{-ddCt}.

5.18 Phylogenetic analyses

Phylogenetic analyses were performed by Frank Förster (Department of Bioinformatics, University of Würzburg).

5.18.1 Bioinformatical prediction of proteins in ant genomes

Gene prediction was performed for the published genomes of the ant species *C. floridanus, Atta cephalotes, Harpegnathos saltator, Pogonomyrmex barbatus, Solenopsis invicta, Linepithema humile, and Acromyrmex echinator* using the gene prediction pipeline maker (version 2.11-beta) (Cantarel et al. 2008). Therefore, the genomic contigs were prefiltered by BLASTX (version 2.2.24). All contigs having hits against published hymenoptaecin or defensin proteins were used for the gene prediction pipeline. The identified *C. floridanus* cDNA sequences and all available protein sequences from *A. mellifera* or other ants were used as EST and protein evidence for the gene predictions. Augustus with its *Nasonia* model was

used for *de novo* gene predictor (Stanke et al. 2008). The obtained gene predictions were manually curated.

5.18.2 Phylogenetic tree reconstruction for mature hymenoptaecin peptides

The sequences for the cDNAs and proteins resulting from the gene prediction and the real sequences obtained from *C. floridanus* were analysed by the ProP server (version 1.0) (Duckert et al. 2004). All cDNAs were fragmented according to the cleavage sites predicted by ProP. All obtained single domain cDNA fragments were aligned by translator (Abascal et al. 2010) with default settings and the resulting alignment was cleaned by Gblocks (Talavera and Castresana 2007) with the default settings from the translatorX website. The phylogenetic tree was reconstructed by PhyML (version 3.0.1) (Guindon et al. 2010) under the GTR+I+G+F model with 100 bootstrap replicates as implemented in seaview (version 4.3.0) (Gouy et al. 2010). Branches with a bootstrap support below 40 were combined using iTOL (Letunic and Bork 2007) and the tree was drawn with the software FigTree (version 1.3.1).

5.18.3 Phylogenetic tree reconstruction and tree reconciliation for defensin peptides

For two ant species, two defensin peptides were predicted. Therefore, *Ixodes scapularis* was added to the sequence set as outgroup (GenBank Acc. No.: XP_002436104.1) and the two *A. mellifera* defensins defensin-1 (GenBank Acc. No.: NM_001011616.2) and defensin-2 (GenBank Acc. No.: NM_001011638.1). The whole sequence set was aligned by muscle (version 3.8.31) (Edgar 2004). The tree was reconstructed by BioNJ (Gascuel 1997). Therefore, the observed amino acid frequencies were used. Branches with a bootstrap support below 40 were combined using iTOL (Letunic and Bork 2007). The species tree for tree reconciliation was derived from Brady et al. 2006 and Gadau et al. 2012 (Brady et al. 2006; Gadau et al. 2012). For tree reconciliation the software Notung (version 2.6) (Durand et al. 2006; Vernot et al. 2008) was used. The gene tree and the reconciled tree were drawn using FigTree (version 1.3.1).

5.19 RNA interference (RNAi)

RNAi is a cellular process, which leads to the specific downregulation of gene expression by double stranded RNA (dsRNA) (Fire et al. 1998). Within the cell the dsRNA is cleaved into small interfering RNAs (siRNAs) of 19-29 nucleotides (Martinez et al. 2002). These siRNAs serve as template for degradation of complementary mRNAs leading to knockdown of gene expression. In order to avoid off-target effects, target sequences for dsRNA production must not contain more than 18-nucleotides perfect matches with all other expressed sequences of the target organism (Kulkarni et al. 2006).

5.19.1 *In vitro* synthesis of dsRNA

DsRNA can be synthesized in a single in vitro transcription reaction with a T7 RNA polymerase from a PCR product that contains opposing T7 RNA polymerase promoter sites at the 5' ends of each strand.

Specific primers were designed for *C. floridanus TyrOH* gene as well as for *GFP* gene (Table 9) and corresponding product sequences (300-400 bp) were checked for off-target effects using NCBI Nucleotide Blast. All primer sequences were fused with the T7 promoter sequence at the 5´-end. *TyrOH* gene fragment was amplified from *C. floridanus* cDNA and *GPF* fragment from vector pEGFP-N1 (Clontech) with specific primers through PCR in 35 cycles using MolTaq polymerase. PCR products were purified, cloned into pGEM®-T vector and sequenced. DNA template for dsRNA transcription was amplified from plasmids with correct sequence via PCR with the T7 promoter-fused primers. DsRNA was prepared in 20 µl reactions using the MEGAscript kit (Ambion) according to the manufacturer's procedures. Sense and antisense strands were transcribed from 1 µg DNA template in the same reaction. For purification dsRNA was phenol-chloroform extracted and isopropanol precipitated as described in section 5.19.4.

5.19.2 Creation of dsRNA-producing *E. coli* strains

In order to obtain large amounts of dsRNA for feeding experiments, the dsRNA was produced in *E. coli* HT115(DE3) strain, which was provided by Lisa Timmons (University of Kansas, USA) & Andrew Fire and obtained from CGC (Caenorhabditis Genetics Center, Minneapolis, USA) (Kamath et al. 2001; Timmons et al. 2001). L4440 plasmid, a gift of Andrew Fire (Addgene plasmid 1654), is suitable to produce dsRNA within this bacterial strain. It has two T7 promoters in inverted orientation flanking the multiple cloning site and confers ampicillin resistance (Timmons and Fire 1998). For A/T-cloning T-tailed L4440 vector was constructed as described before (Kamath and Ahringer 2003). Briefly, 10 µg L4440 vector (obtained from a midi-plasmid-preparation) were digested in 1x Buffer R (Fermentas) with 5 µl EcoRV enzyme (50 u) (Fermentas) in a 100 µl reaction at 37°C overnight to create blunt ends for 3´-T addition. DNA was precipitated through addition of 10 µl 3 M sodium acetate and 300 µl ethanol. After incubation at -80°C for 20 min samples were centrifuged for 10 min at 14 000 rcf and 4°C. DNA pellets were washed with 70 % ethanol and centrifuged for 10 min at 7500 rcf and 4°C. Ethanol was removed, pellets were air-dried under a hood for 5-10 min and finally resuspended in 79 µl H_2O. T-tailing of the digested vector was performed by addition of 10 µl 10x PCR buffer (Molzym), 8 µl 25 mM $MgCl_2$, 2 µl 100 mM dTTP, 1 µl MolTaq Polymerase (Molzym) and incubation for 2 h at 72°C. For dephosphorylation 1 µl FastAP™ Thermosensitive Alkaline Phosphatase (Fermentas) was added and the reaction mix

was incubated for 30 min at 37°C. The dephosphorylated, T-tailed vector was purified using the QIAquick PCR purification kit (Qiagen) and eluted in 30 µl elution buffer.

Fragments approximately 400 bp in length of *C. floridanus PGRP-LB* and *PGRP-SC2* mRNA and of *GFP* mRNA were chosen as inserts for dsRNA production (for primers see Table 9) and checked for off-target effects using NCBI Nucleotide Blast. *PGRP* gene fragments were amplified from *C. floridanus* cDNA and *GPF* fragment was amplified from vector pEGFP-N1 (Clontech) with gene specific primers through PCR in 35 cycles using MolTaq polymerase (Molzym). PCR products were analyzed on a 1.2 % agarose gel and purified with PCR Purification Kit (Qiagen). 200 ng of purified insert and 100 ng of T-tailed L4440 were ligated overnight at 16°C in 1x T4 DNA Ligase Buffer (Fermentas) with 5 u T4 DNA Ligase (Fermentas) in a 20 µl reaction. 10 µl of ligation reaction were transformed into competent *E. coli* DH5α cells via heat-shock (see section 5.8). Transformed bacteria were seeded on LB Amp100 agar plates. Single colonies were screened for inserts with correct size via PCR with vector specific T7 promoter primer. Plasmids were isolated and inserts were sequenced at Seqlab. L4440 plasmids containing the designated insert were transformed into competent *E. coli* HT115(DE3) cells via heat-shock (see section 5.8). Positive clones were selected on LB Amp100 / TET12.5 agar plates. To confirm insert identity to *C. floridanus PGRP* genes, L4440 plasmids from single colonies were isolated and sequenced using vector specific-primers (L4440_2720_seqF and L4440_312_seqR). Aliquots of corresponding clones were stored in LB medium with 50 % glycerin at -80°C.

For *PGRP-LB* / *-SC2* double knockdown *PGRP-LB* gene fragment was amplified with specific primers Cfl_PGRP-LB_HincIIF and Cfl_PGRP-LB_HincIIR (Table 9). About 4 µg of the PCR product as well as 4 µg of the plasmid L4440-SC2 were each digested in 1x Buffer B (Fermentas) with 15 u ApaI and 30 u HincII in a 40 µl reaction at 37°C overnight. Digested *PGRP-LB* fragment and vector L4440-SC2 were ligated and subsequently transformed into competent *E. coli* HT115(DE3) cells as described in section 5.8. Correct insert sequence was confirmed via sequencing at Seqlab.

5.19.3 Phenol extraction of dsRNA from *E. coli* HT115(DE3) cells

For dsRNA production single colonies of HT115(DE3) bacteria containing L4440 vector plus appropriate insert were grown overnight at 37°C and 190 rpm in LB medium with 100 µg/ml Amp plus 12.5 µg/ml TET. The culture was diluted 50-fold in 100 ml of 2x YT medium and allowed to grow to an OD_{600} of 0.4. Synthesis of T7 polymerase was induced by addition of IPTG to 1 mM and the bacteria were incubated with shaking for an additional 5 h at 37°C.

Total nucleic acids were extracted on the basis of the protocols from Timmons (Timmons et al. 2001). 50 ml of each bacteria culture were centrifuged at 5000 rcf for 10 min. Cell pellets were resuspended in 5 ml 1 M ammonium acetate (NH_4Ac) / 10 mM EDTA and split in 5x 1 ml aliquots. 1 ml of phenol:chloroform:isoamyl alcohol (P:C:I, 25:24:1) was added to each

aliquot and the samples were incubated at 65°C for 30 min with occasionally shaking. After centrifugation at 12 000 rcf and 4°C for 15 min the upper aqueous phase was extracted once again with 1 ml of P:C:I (25:24:1). The upper aqueous phase was mixed with 1 volume of isopropanol (1 ml) and incubated at -20°C overnight. Nucleic acids were pelleted through centrifugation at 12 000 rcf and 4°C for 30 min. The pellet was washed with 70 % EtOH and centrifuged for 5 min at 7500 rcf and 4°C. The supernatant was removed and the pellet was air-dried for 5-10 min under a hood. Each pellet was resuspended in 70 µl nuclease-free water and the corresponding aliquots were combined (5 x 70 µl = 350 µl). A 1:10 dilution of each sample was checked photometrically and about 1 µg was analyzed through agarose gel electrophoresis.

In order to obtain pure dsRNA, the nucleic acids were treated with RNase free DNase (Ambion) for digestion of gDNA and with RNase A solution (Ambion) for digestion of single stranded (ss) RNA. Each sample was digested in 1x Turbo DNase Buffer (Ambion) with 6 µl Turbo DNase (Ambion) and 4 µl RNase A (Ambion) in a 400 µl reaction for 1-2 h at 37°C. Afterwards, 140 µl H_2O and 60 µl 5 M NH_4Ac were added to each reaction and the dsRNA was extracted with an equal volume (600 µl) of phenol:chloroform (P:C, 1:1) as described in section 5.19.4.

5.19.4 Purification of dsRNA via phenol/chloroform extraction and isopropanol precipitation

For purification of dsRNA 35 µl H_2O and 15 µl 5 M NH_4Ac were added per 100 µl sample containing dsRNA (from *in vitro* transcription or after extraction from *E. coli* HT115(DE3) cells) and extracted with an equal volume (150 µl) of phenol:chloroform (P:C, 1:1). Samples were vortexed for 1 min and spun at 16 000 rcf in a microcentrifuge for 2 min. The upper, aqueous phase was transferred to a fresh tube and mixed with 1 volume chloroform. Again, samples were vortexed for 1 min and centrifuged at 16 000 rcf for 2 min. RNA was precipitated from the upper aqueous phase by adding 1 volume of isopropanol and mixing well. The mixture was chilled for at least 20 min at -20°C and then centrifuged for 15 min at 14 000 rpm and 4°C to pellet the RNA. The supernatant was carefully removed and the pellet was washed with 500 µl 70 % EtOH. After centrifugation for 5 min at 7500 rcf and 4°C the supernatant solution was carefully removed. Afterwards, samples were briefly centrifuged again and residual EtOH was removed using a micropipette. DsRNA pellets were dried in a SpeedVac Centrifuge for at least 10-20 min.

For dsRNA injection experiments dsRNA pellets were resuspended in nuclease-free H_2O to a final concentration of 10 µg/µl. For feeding experiments dsRNA was dissolved in 300 mM saccharose solution to a final concentration of 2 µg/µl. In each case concentration and purity of a 1:10 dilution of the dsRNA were measured by a NanoDrop™ spectrophotometer and dsRNA quality was checked through agarose gel electrophoresis.

5.19.5 Delivery of dsRNA to *C. floridanus*

DsRNA from in vitro transcription was injected using sterile glass capillaries as described in section 5.2.2. For *TyrOH* gene knockdown (see section 6.5.3) 5 µg dsRNA in a volume of 0.3-0.5 µl nuclease free H_2O were injected per individual.

In order to check the stability of the dsRNA within the crop of *C. floridanus*, adult minor workers were starved for 3 days and then fed with 300 mM saccharose solution containing specific dsRNA at a concentration of 2 µg/µl ad libitum. At 6 h and 24 h after feeding, individuals were shortly anesthetized on ice and then light pressure was applied to the gaster, until they regurgitated their crop content. 1 µl of the resulting regurgitate per ant was analyzed by TBE agarose gel electrophoresis.

For *PGRP* knockdown experiments in untreated animals, dsRNA was extracted from the four different *E. coli* HT115(DE3) strains, which either expressed dsRNA of the *GFP* gene, the *C. floridanus PGRP-LB* or *PGRP-SC2* gene or of both genes (in this case the *PGRP* genes sequences were cloned in tandem into the L4440 expression vector, see section 5.19.2).

In the preliminary study (section 6.10.1.3) the purified dsRNA was fed to subcolonies of *C. floridanus* consisting of 50 adult minor workers and 15 larvae (presumed early L2, about 1-2 mm long). Over a period of 10 days (d), ants were supplied once a day with about 2 µl 300 mM saccharose solution per individual, either pure (H) or containing 2 µg/µl dsRNA of *GFP* (G), *PGRP-LB* (L), *PGRP-SC2* (S) or of *PGRP-LB* and -*SC2* (LS). Afterwards, all subcolonies were fed with 300 mM saccharose solution for five additional days. After 5 d, 10 d and 15 d feeding period, total RNA was extracted in each case from 10 midguts and 5 residual bodies from animals of the five different groups and transcription of the target genes was determined using qRT-PCR. Furthermore, target gene expression was determined in cDNA probes from larvae (each derived from 3 whole larvae of each group) at 10 d and 15 d, in order to determine, if a RNAi effect could also be achieved in larvae.

For *PGRP-LB* knockdown during an immune response (section 6.10.2), subcolonies consisting of 25 adult minor workers and 7 larvae (L2, about 2 mm long) were fed with dsRNA from *GFP* or *PGRP-LB* gene for 6 d, as described above. At day 4, workers and larvae were picked with a needle dipped into dead bacteria (1:1 mix of *E. coli* D31 and *M. luteus*, see section 5.2.2) in order to elicit an immune response. As controls, two subcolonies were supplied with 300 mM pure saccharose solution for 6 d. On day 4, the individuals from one control subcolony (I) were likewise picked with dead bacteria just like the dsGFP- and dsPGRP-LB-receiving groups, whereas the individuals from the other control subcolony (H) were only anesthetized on ice, but not picked. At day 6, total RNA was extracted in each case from 10 midguts and 5 residual bodies as well as from 3 whole larvae of the four different groups and immune gene expression was investigated using qRT-PCR.

5.19.6 Statistics for *PGRP-LB* knockdown experiments

Statistical analyses were performed with Statistica 9.0. For the study in section 6.10.2 average gene expression levels relative to the constitutively expressed *rpL32* gene (dCt = Ct(target) - Ct(*rpL32*)) of six independent samples were determined from immune-challenged and untreated ants at each time point. A factorial ANOVA followed by Tukey´s HSD post hoc test was performed on these (dCt)-values in order to investigate the influence of the dsRNA-feeding and the immune-challenge. Then, normalized changes in gene expression relative to untreated animals (H) (-ddCt = -(dCt(immune-challenged) - dCt(untreated)) were calculated from six independent samples for each treatment and time point.

5.20 Analysis of *C. floridanus* haemolymph

5.20.1 Haemolymph sampling

Haemolymph samples from untreated major workers were compared with those from animals 24 h after picking with a needle dipped into a 1:1 mix of *E. coli* D31 / *M. luteus*. For haemolymph collection major worker ants were cooled on ice and then head and gaster were cut off with micro-scissors. Head and thorax were slightly pressed using forceps and the leaking out haemolymph was soaked up with a calibrated glass capillary. The collected haemolymph (0.5-3 µl per individual) was transferred to reaction tubes containing 1 µl of a mixture of 1-Phenyl-2-thiourea (PTU, P-7629) and aprotinin (A-4529) (Sigma-Aldrich) each at a concentration of 0.1 mg/ml to prevent melanization. Haemolymph from 8-15 individuals was pooled in order to obtain 15 µl aliquots. Samples were centrifuged for 10 min at 14 000 rcf and 4°C to pellet cell debris. The supernatant was kept on ice or frozen at -20°C until further analyses.

5.20.2 Inhibition-zone assays

Inhibition-zone assays were performed in order to determine antibacterial activity of the *C. floridanus* haemolymph. Microorganisms used in the inhibition-zone assays were *E. coli* D31 (Monner et al. 1971) as a gram-negative bacterium as well as *Micrococcus luteus* and *Bacillus subtilis* as gram-positive bacteria. In each case 200 µl of a fresh overnight culture (with an OD_{600} of about 2) were plated on LB agar plates. As soon as the liquid was permeated, 3 µl haemolymph were pipetted as a droplet on the plates. Furthermore 5 µg kanamycin were applied as a positive control. The plates were incubated at 30°C overnight. On the next day the clear zone of inhibition was documented by photography.

5.20.3 SDS-polyacrylamide analysis of haemolymph proteins

SDS-polyacrylamide-gel-electrophoresis (SDS-PAGE) is a method to separate proteins on the basis of their molecular weight. Proteins are linearized through heat denaturation and negatively charged in a SDS containing buffer, so that they migrate towards the anode within an electric field. The higher the acrylamide concentration of the gel is, the smaller are the gel pores and the better is the separation of small proteins. A typical gel (10 % polyacrylamide) consists of a stacking and a resolving gel with the components listed in Table 10.

Table 10. Composition of SDS-polyacrylamide stacking and resolving gel.

Stacking gel (5 %, 5 ml)	Resolving gel (10 %, 20 ml)
3.4 ml H_2O	8 ml H_2O
0.83 ml 30 % acrylamide mix	6.6 ml 30 % acrylamide mix
0.63 ml 1.0 M Tris (pH 6.8)	5 ml 1.5 M Tris (pH 8.8)
0.05 ml 10 % SDS	0.2 ml 10 % SDS
0.05 ml 10 % APS	0.2 ml 10 % APS
0.005 ml TEMED	0.008 ml TEMED

In this work aliquots of the same haemolymph samples were electrophoresed on both 10 % and 15 % polyacrylamide gels for the separation of proteins in the range of 30-200 kDa and 3-30 kDa, respectively. For 15 % gels the portion of acrylamide was increased to 10 ml and the portion of water was accordingly diminished to 4.6 ml. After casting the resolving gel was overlaid with isopropanol in order to get a straight polymerization front. After polymerization isopropanol was decanted, the stacking gel was poured on top of the resolving gel and a comb was put into the gel for formation of wells for the samples.

Each 5 µl of the haemolymph samples were mixed with 2.5 µl 4x Laemmli buffer and denatured for 5 min at 95°C. Samples were loaded onto the gel and electrophoresed together with a size marker (PageRuler™ Prestained Protein Ladder from Fermentas). For separation of large proteins samples were electrophoresed on 10 % acrylamide gels in 1x Tris-glycine buffer at 140 V until the dye front had reached the lower end of the gel. For better disbanding of small proteins samples were run on 15 % acrylamide gels at 25 mA using Tricine-SDS-PAGE (Schägger 2006; Schägger and von Jagow 1987). In this case the cathode buffer (in the upper buffer tank) was composed of 0.1 M Tris, 0.1 M Tricine (pH 8.3) and 0.1 % SDS, whereas the anode buffer (in the lower tank) consisted of 0.2 M Tris-HCl (pH 8.9) (Randolt et al. 2008).

After electrophoresis proteins were visualized via colloidal Coomassie staining using Roti®-Blue (Roth) according to the manufacturer's instructions. Initially, gels were fixed for 30 min in 0.85 % o-phosphoric acid / 20 % methanol and then stained overnight in 1x Roti®-Blue solution with 20 % methanol. Gels were destained in 25 % methanol.

6 Results

6.1 Identification of immune-relevant genes from *C. floridanus* by suppression subtractive hybridization (SSH)

Using SSH, 65 differentially expressed ESTs were identified (GenBank Acc. No. HS410953 - HS411017) corresponding to 35 infection-inducible genes from *C. floridanus* presumably linked to immune defense and immunity-related processes (Table 11) (Ratzka et al. 2011). With this approach a first experimental complement to the bioinformatic identification of immune defense genes from the recently published *C. floridanus* genome sequence was provided (Bonasio et al. 2010).

In *Drosophila melanogaster* the recognition of invasive microorganisms by pattern recognition receptors (PRRs) leads to immune responses via four signaling pathways (Toll, IMD, JAK/STAT and JNK) (Lemaitre and Hoffmann 2007). Here several ESTs were identified, which encode proteins possibly involved in signaling, such as a protein annotated as a MAP-kinase activated protein kinase. Furthermore ESTs were found coding for a putative Gram-negative binding protein (GNBP). This may trigger Toll activation via recognition of Lys-type peptidoglycan of Gram-positive bacteria. Toll is activated through a proteolytic cascade involving several serine proteases and resulting in the cleavage of the cytokine Spaetzle (Lemaitre and Hoffmann 2007; Ryu et al. 2010). The subtracted library contained many ESTs encoding putative serine proteases e.g. snake-like proteins, which could be involved in this proteolytic Toll activation, as suggested for *Drosophila* (Irving et al. 2001). Six ESTs were also very similar to a prophenoloxidase activating factor from *Tribolium* (GenBank Acc. No.: XP_968658) and the serine protease homolog 42 from *Apis* (GenBank Acc. No.: XP_623150), which was shown to be induced after septic injury and could be involved in prophenoloxidase cleavage and activation (Zou et al. 2006). The corresponding *C. floridanus* gene, annotated as serine protease *stubble* (*stu*), could play a role in triggering melanization. In addition, two transcripts encode a putative serine protease inhibitor (*serpin*), which may regulate serine protease cascades.

Activation of immune signaling triggers immune defense, e.g. antimicrobial peptides (AMPs) (Lemaitre and Hoffmann 2007). Several ESTs were found coding for the AMPs defensin-1 (def-1) and hymenoptaecin (hym). The latter is a glycine-rich AMP known from other Hymenopteran insects, but escaped detection in the *C. floridanus* genome so far as the sequence is not included in the draft genome sequence (Bonasio et al. 2010). However, it seems to play a key role in the ant´s immune response as transcription was strongly induced after an immune-challenge (see section 6.3).

MAMP recognition and wounding lead to melanization reactions in *Drosophila* (Lemaitre and Hoffmann 2007). Concordantly, ESTs were found for the enzymes tyrosine hydroxylase (TyrOH) and GTP cyclohydrolase, shown to be induced after infection in *Drosophila* and

Manduca and probably involved in immune-associated melanization (De Gregorio et al. 2001; Gorman et al. 2007). TyrOH catalyzes the conversion of tyrosine into DOPA, which via dopamine is converted into melanin. GTP cyclohydrolase (*Punch*) partcipates in tetrahydrobiopterin formation, which is an essential cofactor for TyrOH. JNK signaling is required for activation of *Punch* (Silverman et al. 2003). Interestingly, one EST was found belonging to a gene with homology to *malvolio* (*mvl*), which is also JNK controlled (Silverman et al. 2003). *Malvolio* is the *Drosophila* homologue of NRAMP-1 (natural resistance-associated macrophage protein). This ion metal transporter is probably involved in macrophage defense against microbial invasion in vertebrates (Nelson 1999). This gene is upregulated during an immune response in *Drosophila* and implicated in the insect cellular immune response (Silverman et al. 2003). JNK signaling is required for the activation of immune-inducible genes like *punch* and *malvolio*, suggesting an important role in insect cellular immunity and stress response (Silverman et al. 2003).

Stress and immune responses in insects seem to be partly linked (Brun et al. 2006; Ekengren and Hultmark 2001). Several ESTs were found encoding stress-related proteins, induced upon septic injury in *Tribolium castaneum* (Altincicek et al. 2008b). These included a heat shock protein, a cytochrome P450 and a lipocalin (Ratzka et al. 2011).

Table 11. ESTs from the subtracted *C. floridanus* library (see next 3 pages) (Ratzka et al. 2011). ESTs were analyzed as described in section 5.12 and classified according to function. Transcripts were quantitated by qRT-PCR in midgut and fat body of untreated major worker ants and in animals six hours after injection of 0.3 µl ant ringer solution containing a 1:1 mix of heat killed *E. coli* and *M. flavus* at 6-8 x 10^9 cells/ml. Midgut and fat body tissue from eight workers was pooled for each sample. Comparisons of transcript levels between the three samples were made using two-sided t-tests. A significant increase in transcript level six hours after bacteria injection is indicated by an asterisk ($p < 0.05$). Acc. No.: accession number; n.s.m.: no significant match

Annotated as	GenBank Acc. No. of EST	GenBank Acc. No. of protein	E-value	COG	Description	median (-ddCt)	homologous sequence	E-value
Signaling and immune defense								
Beta-1,3-glucan-binding protein	HS410953	EFN66519	2E-29	COG2273	Gram-negative binding protein (GNBP), Beta-glucanase/Beta-glucan synthetase	4.31 (*)	ACT66879	1E-17
	HS410954	EFN66519	2E-82				XP_001121634	2E-48
	HS410969	EFN66519	6E-78				XP_001121634	2E-44
	HS410970	EFN66519	1E-155				XP_001121634	7E-52
Lipoma-preferred partner-like protein	HS410962	EFN64790	5E-15	KOG4577	LIM-domain containing, lipoma preferred partner (LPP)	3.04 (*)	XP_396072	5E-09
	HS410967	EFN64790	5E-126				XP_396072	1E-105
	HS410960	EFN64790	2E-77				XP_396072	7E-76
MAP kinase-activated protein kinase 2	HS410965	EFN68856	1E-41	KOG0604	MAP kinase-activated protein kinase 2, Serine/Threonine protein kinase	1.14 (*)	XP_394995	2E-18
Sequestosome-1	HS410961	EFN70930	0.0	KOG0457, KOG1280	Oxidative stress protein, sequestosome 1, containing ZZ-type Zinc finger	1.60 (*)	XP_001608187	3E-45
Leukocyte elastase inhibitor / Serpin	HS410959	EFN69733	4E-141	COG4826, KOG2392	Serine protease inhibitor (Serpin)	3.89 (*)	XP_001603946	3E-85
	HS410968	EFN69733	5E-103				XP_001603946	4E-24
Peroxidase	HS410958	EFN68225	3E-99	-	Peroxidase, Peroxinectin	0.84	XP_392481	5E-53
Hymenoptaecin	HS410972	-	-	-	Hymenoptaecin precursor	7.29 (*)	ACA04899	7E-06
Hymenoptaecin	HS410973	-	-	-	Hymenoptaecin precursor		ACA04899	4E-05
Hymenoptaecin	HS410976	-	-	-	Hymenoptaecin precursor		ACA04899	4E-19
Hymenoptaecin	HS410974	-	-	-	Hymenoptaecin precursor		ACA04899	4E-20
Hymenoptaecin	HS410975	-	-	-	Hymenoptaecin precursor		NP_001165830	4E-09
Hymenoptaecin	HS410978	-	-	-	Hymenoptaecin precursor		NP_001165830	2E-09
Hymenoptaecin	HS410971	-	-	-	Hymenoptaecin precursor		ADB29130	4E-08
Hymenoptaecin	HS410977	-	-	-	Hymenoptaecin precursor		ADB29130	4E-08

Annotated as	GenBank Acc. No. of EST	GenBank Acc. No. of protein	E-value	COG	Description	median (-ddCt)	homologous sequence	E-value
Defensin-1	HS410966	EFN62355	9E-19	-	Defensin	5.66 (*)	ACB46512	2E-15
GTP cyclohydrolase 1	HS410964	EFN69560	5E-70	COG0302, KOG2698	Punch, GTP cyclohydrolase I	4.95 (*)	XP_393086	3E-69
Tyrosine 3-monooxygenase	HS410957	EFN71001	2E-36	KOG3820, COG3186	Tyrosine hydroxylase	4.18 (*)	NP_001011633	8E-20
	HS410956	EFN71001	7E-62				XP_392459	5E-56
	HS410955	EFN71001	1E-148				XP_001600616	7E-132
Malvolio	HS410963	EFN64791	2E-61	COG1914, KOG1291	Natural resistance-associated macrophage protein (Nramp)	2.70 (*)	XP_001600967	6E-41
Serine proteinases								
Serine proteinase stubble	HS410979	EFN72618	1E-85	COG5640, KOG3627	Secreted trypsin-like serine protease, similar to serine protease homolog 21 (*Nasonia*), serine protease homolog 42 (*Apis*) and prophenoloxidase activating factor (*Tribolium*)	2.49 (*)	NP_001155060 / XP_001600151 (XP_623150 and XP_968658)	5E-103
	HS410988	EFN72618	1E-85					2E-104
	HS410989	EFN72618	1E-85					6E-104
	HS410987	EFN72618	1E-85					2E-101
	HS410984	EFN72618	9E-52					8E-40
	HS410985	EFN72618	1E-79					2E-96
Serine protease	HS410980	-	-	-	Serine protease with chitin binding Peritrophin-A domain	2.40 (*)	XP_001604250	8E-05
Serine protease snake	HS410982	EFN71579	4E-30	-	Regulatory CLIP domain of proteinases	5.77 (*)	XP_001120043	8E-09
	HS410983	EFN71579	4E-28	-				2E-07
Serine protease snake	HS410986	EFN63907	2E-28	COG5640, KOG3627	Secreted trypsin-like serine protease	3.86 (*)	XP_001121032	1E-19
Serine protease snake	HS410991	EFN63907	1E-75				XP_001121032	1E-30
Peroxisomal membrane protein PEX16	HS410981	EFN71262	1E-35	COG5640, KOG3627	Secreted trypsin-like serine protease	5.30 (*)	XP_002069996	7E-21
Mite allergen Der f 3	HS410990	EFN71150	2E-131	COG5640, KOG3627	Secreted trypsin-like serine protease	1.84 (*)	XP_624813	4E-82
Stress response-associated proteins								
Apolipoprotein D	HS410994	EFN66334	3E-94	-	Lipocalin, apolipoprotein D	2.19 (*)	ADA82598	2E-65
	HS411000	EFN66334	2E-93	-				1E-64
J domain-containing protein (HSP)	HS410992	EFN67201	1E-100	COG2214, KOG0550, KOG0691	DnaJ-class molecular chaperone, heat shock protein	2.98 (*)	XP_392393	2E-62
	HS410997	EFN67201	9E-101					2E-72

Annotated as	GenBank Acc. No. of EST	GenBank Acc. No. of protein	E-value	COG	Description	median (-ddCt)	homologous sequence	E-value
Cytochrome P450 6a2	HS410993 HS410995 HS410999	EFN65184 EFN65184 EFN65184	1E-77 4E-70 1E-69	KOG0158	Cytochrome P450 CYP3/CYP5/CYP6/CYP9 subfamilies	1.74 (*)	NP_001035324	5E-57 3E-51 3E-49
-	HS410998 HS410996	- -	- -	-	BAX inhibitor (BI)-1 like protein	2.09 (*)	HP544217, XM_391854	4E-22 5E-06
Metabolism								
KAT AadAT	HS411007 HS411004	EFN62920 EFN62920	7E-64 1E-92	COG0436	Kynurenine/alpha-aminoadipate aminotransferase mitochondrial	5.12 (*)	XP_001603888 XP_001120032	1E-51 3E-69
Oxysterol-binding protein 1	HS411006	EFN68395	8E-20	KOG1737, KOG2210	Oxysterol-binding protein	0.79 (*)	XP_392480	5E-17
SCaMC-2	HS411001	EFN71751	2E-47	KOG0036	Small calcium-binding mitochondrial carrier	3.04 (*)	XP_001603181	2E-12
Neurochondrin-like protein	HS411005	EFN61035		COG1011, KOG3085	Predicted hydrolase (HAD superfamily)	0.58 (*)	XP_562137	1E-59
Cytochrome b5	HS411002	EFN73453	3E-18	KOG0537, KOG4576	Cytochrome b5 /Heme steroid binding	6.19 (*)	XP_001120546	2E-22
SERAC1	HS411003	EFN73214	6E-54	-	Eukaryotic cytochrome b(561)	3.82 (*)	XP_974652	1E-08
Translation								
IF4E	HS411008	EFN73765	6E-79	KOG1669, KOG1670	Translation initation factor	0.64 (*)	XP_624290	2E-65
Eukaryotic translation initiation factor 2C 2	HS411009	EFN62163	1E-54	KOG1041	Argonaute 1 (translation initiation factor 2c)	0.50 (*)	XP_624444	1E-53
Hypothetical proteins / unknown UTR								
hypothetical protein EAG_12212	HS411013	EFN75036	1E-61	-	Hypothetical protein	3.32 (*)	XP_625261	4.4
hypothetical protein EAG_14061	HS411014	EFN74149	2E-65	-	Hypothetical protein	2.43 (*)	XP_393274	5E-48
-	HS411017	-	-	-	Hypothetical protein	1.62 (*)	XP_001120071	0.06
-	HS411015	-	-	-	Hypothetical protein	4.19 (*)	EFN76145	1E-09
-	HS411016	-	-	-	Hypothetical protein	3.92 (*)	n.s.m.	-
-	HS411012	-	-	-	Hypothetical protein	3.88	n.s.m.	-
- -	HS411011 HS411010	- -	- -	- -	Unknown	4.53 (*)	BX248506	3E-06

6.2 Characterization of the transcriptional response to bacterial challenge in *C. floridanus* larvae and workers by Illumina sequencing

Illumina sequencing technology was used in order to attain a more comprehensive view of the transcriptional changes after an immune-challenge than obtained by SSH.

6.2.1 Quality of RNA samples

The quality of the total RNA for this sequencing project was checked using an Agilent 2100 Bioanalyzer. The virtual gels and electrophoretic profiles of both samples were quite similar (Figure 7).

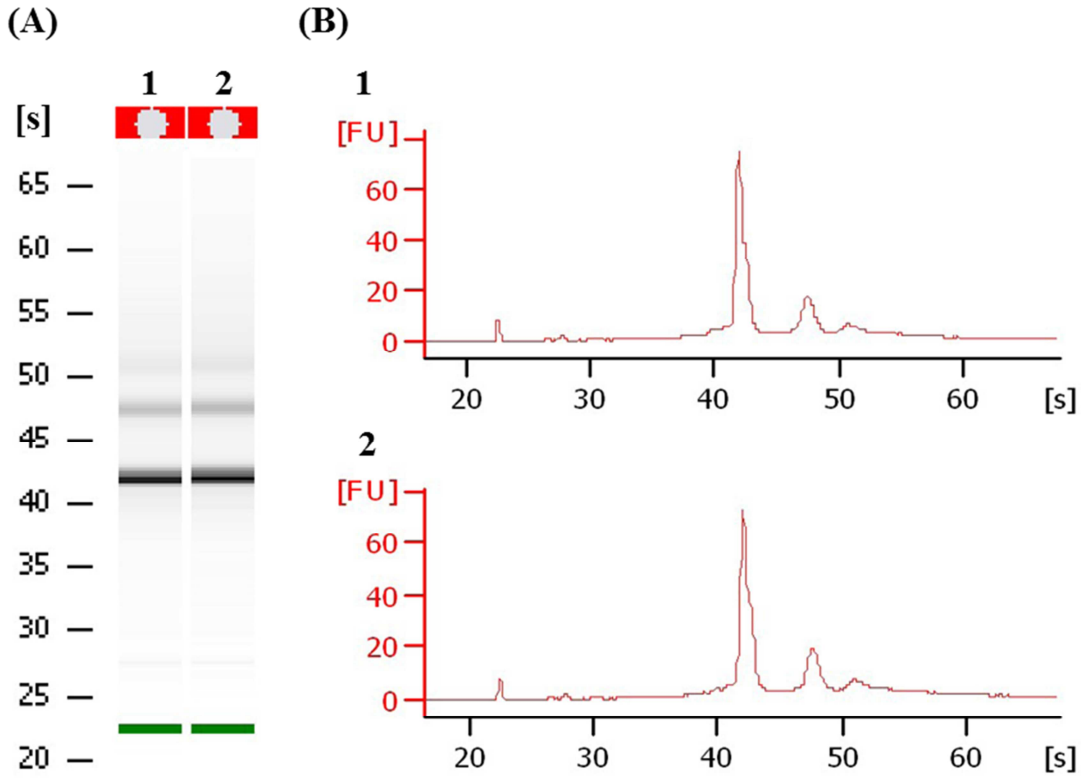

Figure 7. (A) The image shows a total RNA gel like-image produced by the Agilent 2100 Bioanalyzer. Lane 1: RNA from *C. floridanus* workers and larvae (1:1 mix) at 12 h after bacterial challenge, Lane 2: RNA from untreated workers (W2) and larvae (L2) (1:1 mix). (B) Electrophoretic profiles of RNA from immune-challenged (1) and untreated (2) W2 and L2 (1:1 mix). [FU]: Fluorescence units, [s]: seconds

The band / peak at about 47 [s] run time likely represents the 28S rRNA. In most insects the 28S rRNA consists of two fragments, which are linked by a hydrogen-bond. Depending on the pretreatment (e.g. heat denaturation) and electrophoresis conditions (native or denaturing), disruption of this hydrogen-bond can occur, which was traditionally referred to as “hidden break” of the RNA (Gould 1967; Ishikawa and Newburgh 1972). The resulting two fragments

have about the same size as the 18S rRNA and thus co-migrate on the gel (Gillespie et al. 2006; Winnebeck et al. 2010). Thus, the band / peak at about 43 [s] run time is likely composed of the 18S rRNA and separated 28S rRNA fragments (Figure 7). Since both samples had two sharp bands / peaks, the RNA quality was considered to be good and the samples were sent to Eurofins for library preparation and sequencing.

6.2.2 Transcript expression analysis of the Illumina sequencing data using TopHat and Cufflinks

Illumnia sequencing technology was used to sequence the cDNAs of transcripts from immune-challenged and untreated animals (whole L2 and W2). Sequencing was performed by Eurofins MWG GmbH and resulted in 251,747,794 reads for immune-challenged and 236,285,674 reads for untreated animals. A transcriptome assembly of the reads from both samples was performed by Frank Förster using the software tools TopHat and Cufflinks (Trapnell et al. 2012). The merged assembly from both conditions (immune-challenged and untreated) resulted in the prediction of 15,730 genes (provided with new gene identifiers = XLOCs) encoded in the *C. floridanus* genome (Table 12), in contrast to the 17,064 genes predicted in the current Official Gene Set (OGS) v3.3 (original gene identifiers = EAGs) (Bonasio et al. 2010). Based on the Illumina sequencing data, Cufflinks could identify 26,924 sites of alternative splicing, whereas only 7,583 had been predicted so far (Bonasio et al. 2010). Sequence information was furthermore used to define the transcription start sites (TSSs), protein coding DNA sequences (CDSs) and untranslated regions (UTRs) of the predicted genes resulting in the identification of 31,624 gene isoforms. Detailed examination of these isoforms will be accomplished in prospective studies.

Table 12. Results of the transcript expression analysis by Cufflinks. The table shows the predicted genes, isoforms, transcription start sites (TSS), coding sequences (CDS), promoters and alternative splicing (splicing) events of the *C. floridanus* genome based on the Cufflinks analysis of the Illumina sequencing data.

CuffSet instance with:
2 samples
15730 genes
31624 isoforms
26924 TSS
15384 CDS
15730 promoters
26924 splicing
13723 relCDS

6.2.3 Differentially expressed genes after bacterial challenge

Cuffdiff, a program included in the Cufflinks package (Trapnell et al. 2012), detected 136 differentially expressed genes (XLOCs) between the sample from untreated and immune-

challenged animals. In addition, *DESeq* program was used, which is suitable to identify significantly differentially expressed genes between two samples without replicates (Anders and Huber 2010). *DESeq* determined 263 differentially expressed XLOCs ($p < 0.001$), which for the most part included those identified by Cuffdiff. Taken together, 292 XLOCs were identified, which were differentially expressed in response to bacterial challenge (see section 8.1, Table 16).

The differentially expressed genes identified by Cufflinks and/or *DESeq* included most of the genes previously identified by SSH (e.g. *GNBP*, *def-1*, *malvolio* and *TyrOH*) and many others known to be immune-regulated in other insects. The upregulated factors comprised genes encoding pattern recognition receptors (e.g. Hemolymph lipopolysaccharide-binding protein, VIP36-like protein (L-type lectin), GNBP, PGRP), serine proteases (e.g. snake, stubble-like, Limulus clotting factor C), proteins involved in signaling and transcription (e.g. Nuclear factor NF-kappa-B p110 subunit (Relish), NF-kappa-B inhibitor cactus, Leucine-rich repeat-containing protein 20, HIV Tat-specific factor 1-like protein, Parathyroid hormone-related peptide receptor) as well as stress-related proteins (e.g. Cytochromes P450, Cytochrome b5, J-domain-containing protein, Transferrin, Apolipoprotein D). Interestingly, a gene encoding a 23 kDa integral membrane protein belonging to the tetraspanin family was significantly induced upon infection. Tetraspanin proteins typically contain four transmembrane domains and have been implicated in a multitude of biological processes including cellular immunity (Levy and Shoham 2005a; Levy and Shoham 2005b). Moreover, genes encoding two WD repeat-containing proteins and an ankyrin repeat-containing protein (Transient receptor potential channel pyrexia) were significantly induced in immune-challenged animals and possibly represent adaptor proteins involved in protein-protein-interactions (Sedgwick and Smerdon 1999; Smith 2008). Among the upregulated genes were also some known negative regulators of the immune response e.g. *PGRP-LB*, *cactus* and *serine protease inhibitor* genes, which are known to be upregulated after an immune-challenge in *Drosophila* (Nicolas et al. 1998; Reichhart 2005; Zaidman-Remy et al. 2006).

In contrast to SSH, analysis of the Illumina sequencing data revealed not only upregulated, but also downregulated genes. Several of the significantly downregulated genes encode proteins involved in digestion (e.g. Chymotrypsin, Trypsin, Maltase, Lipase) and storage (e.g. Hexamerin), which were already described to be downregulated upon immune-challenge in other insects (Aguilar et al. 2005; Lourenco et al. 2009; Meng et al. 2008).

Among the significantly regulated genes were several members of the cytochrome P450 superfamily, which encode detoxifying enzymes with diverse functions in insecticide resistance and infection (Chung et al. 2009; Félix and Silveira 2012; Feyereisen 1999). In accordance with studies on other insects, some cytochrome P450 genes were significantly up- and some were downregulated after infection indicating complex adjustments in cytochrome P450 gene expression are needed for detoxifying processes (De Gregorio et al. 2001; De Gregorio et al. 2002; Johansson et al. 2005; Yang and Liu 2011). Similarly, the significantly

upregulated *esterase* genes might play a role in abolishing toxic molecules generated during microbial infection, as suggested for *Bombyx mori* (Shiotsuki and Kato 1999), *Galleria mellonella* (Serebrov et al. 2001) and *Manduca sexta* (Zhu et al. 2003).

The set of significantly differentially expressed genes contained several hypothetical proteins, of which some were up- and some were downregulated after the immune-challenge.

6.2.4 Validation of Illumina sequencing results by qRT-PCR

For validation of the results obtained from Illumina sequencing 15 differentially expressed genes were chosen and their expression was determined separately in six biological replicates from untreated and immune-challenged larvae (L2) and workers (W2) by qRT-PCR.

The five supposedly downregulated genes were chosen from a list of significantly differentially expressed genes based on preliminary analysis of the Illumina sequencing with rather mild excluding criteria ($p < 0.1$). In the final strict analyses ($p < 0.001$) by Cufflinks (Trapnell et al. 2012) and *DESeq* (Anders and Huber 2010) these genes (except *phenoloxidase subunit CG8193* (*POsub*) gene) were not considered as significantly downregulated. This fits to the qRT-PCR analysis, in which these genes were not significantly downregulated (*sushi*) or only in L2 (*POsub*, *zinc carboxypeptidase* (*zcp*), *lipase* (*lip*)) or in W2 (*scavenger receptor class B member 1* (*scav*)) (Table 13).

Prominent immune-relevant genes like *PGRP-LB*, *Rel*, *def-1* and *TyrOH*, which were considered as significantly induced by Cufflinks and/or *DESeq*, were confirmed to be significantly upregulated in samples from immune-challenged L2 and W2 by qRT-PCR results (Table 13). When median fold changes in gene expression from L2 and W2 were averaged (median L2,W2), the genes exhibited similar induction levels in qRT-PCR as well as in Cufflinks and *DESeq* analysis (Table 13). Interestingly, qRT-PCR analysis revealed that several genes are more strongly induced in L2 than in W2, e.g. the genes *MPI* (*metalloproteinase inhibitor*), *SOCS2* (*suppressor of cytokine signaling 2*), *transf* (*transferrin*), *ester* (*esterase FE4*) and *hp67112* (*hypothetical protein, EFN67112.1*) were only significantly induced in samples from immune-challenged L2, but not from W2.

In sum, the qRT-PCR results confirmed the data obtained from the Illumina sequencing analysis showing similar trends in up- or downregulation of ant genes after an immune-challenge.

Table 13. Five downregulated and ten upregulated genes were chosen for validation by qRT-PCR from a list of differentially expressed genes determined by a preliminary analysis of Illumina sequencing results. Six RNA samples from each untreated and immune-challenged larvae (L2) and workers (W2) were prepared under same conditions as for Illumina sequencing. Using qRT-PCR average gene expression levels relative to the constitutive *EF1a* gene (dCt = Ct(target) - Ct(*EF1a*)) were determined separately in six samples from L2 and W2. Comparisons of transcript levels between the six samples from untreated and immune-challenged animals were made using two-sided t-tests (asterisks indicate significant differences with $p < 0.05$; ns: not significant). The table shows the median gene expression levels normalized to the expression in untreated animals using the 2^{-ddCt} method. Moreover, fold changes in gene expression after immune-challenge are shown based on the analysis of Illumina sequencing data by Cufflinks (Trapnell et al. 2012) and by *DESeq* (Anders and Huber 2010). Statistical analyses were performed by the software packages Cufflinks and *DESeq* (asterisks indicate significant differences with $p < 0.001$; ns: not significant). The merged transcriptome assembly by Cufflinks resulted in new gene identifiers (XLOCs) for the reannotated genes, which were assigned to the originally annotated genes (EAGs).

Gene name	EAG	XLOC	median L2	median W2	median (L2,W2)	Cufflinks	DESeq
downreg.							
scav	EAG_01232	XLOC_004574	0.54 (ns)	0.24 (*)	0.25 (ns/*)	0.15 (ns)	0.46 (ns)
POsub	EAG_04466	XLOC_001021	0.35 (*)	0.45 (ns)	0.42 (*/ns)	0.17 (*)	0.17 (*)
zcp	EAG_14622	XLOC_001088	0.40 (*)	1.27 (ns)	0.87 (*/ns)	0.29 (ns)	0.28 (ns)
lip	EAG_09693	XLOC_011955	0.38 (*)	1.00 (ns)	0.88 (*/ns)	0.43 (ns)	0.44 (ns)
sushi	EAG_02378	XLOC_011646	0.88 (ns)	0.84 (ns)	0.85 (ns/ns)	0.48 (ns)	0.46 (ns)
upreg.							
MPI	EAG_12870	XLOC_009154	4.75 (*)	0.96 (ns)	1.28 (*/ns)	2.14 (ns)	1.90 (ns)
SOCS2	EAG_09027	XLOC_007856	3.91 (*)	1.48 (ns)	1.92 (*/ns)	2.45 (ns)	2.13 (ns)
transf	EAG_13861	XLOC_012710	12.50 (*)	1.21 (ns)	3.29 (*/ns)	5.58 (ns)	5.28 (*)
PHR	EAG_13089	XLOC_001812	42.15 (*)	5.19 (*)	14.55 (*/*)	13.64 (*)	12.78 (*)
ester	EAG_15281	XLOC_009670	34.48 (*)	1.35 (ns)	8.04 (*/ns)	27.21 (*)	26.37 (*)
hp67112	EAG_12674	XLOC_008001	742.06 (*)	3.74 (ns)	32.16 (*/ns)	276.28 (*)	213.29 (*)
PGRP-LB	EAG_05215	XLOC_001151	12.19 (*)	3.24 (*)	6.04 (*/*)	6.33 (*)	6.48 (*)
Rel	EAG_10813	XLOC_013816	2.09 (*)	3.98 (*)	3.54 (*/*)	2.90 (ns)	2.53 (*)
def-1	EAG_01545	XLOC_012789	9.47 (*)	6.68 (*)	9.47 (*/*)	27.12 (*)	24.15 (*)
TyrOH	EAG_09890	XLOC_004129	12.18 (*)	7.04 (*)	8.76 (*/*)	9.18 (ns)	8.68 (*)

6.3 Characterization of the *C. floridanus hymenoptaecin* (*hym*) gene and mRNA structure

6.3.1 Cloning and sequence analysis of hymenoptaecin encoding cDNA

The search for immune inducible genes in *C. floridanus* by the SSH approach revealed the presence of a cDNA encoding a homologue of hymenoptaecin, an AMP known from other hymenopteran species (Ratzka et al. 2011). As this gene was not annotated in the genome sequence (Bonasio et al. 2010), the full length *hymenoptaecin* (*hym*) cDNA sequence was determined using RACE. The attempts to define the 5`- and 3`-ends of the cDNA resulted in complex patterns. As expected, the 5´RACE of the *hym* cDNA revealed one product. However, the 3´RACE resulted in several products of different length. Further investigation of these products revealed that all were *hym* derived 3´RACE products with an identical 3´UTR and a poly-A tail. The various molecules with different length resulted from a 309-nucleotide sequence, which was repeated several times but in different copy number in the different amplification products. To confirm the existence of these deduced cDNAs the full length *hym* cDNAs were amplified with primers binding near the 5´and 3´-end. Several clones with insert size varying from 681 bp to 2536 bp were identified and analyzed (Figure 8). All of the examined cDNA sequences contained a constant 5´- and 3´-end, embracing a region of variable length containing one to six copies of the 309-nucleotide repeat sequence (Ratzka et al. 2012a).

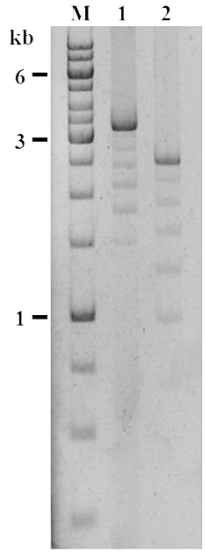

Figure 8. PCR amplification of full length *hym* gene and cDNA (Ratzka et al. 2012a). The PCR-products from gDNA (lane 1) and cDNA (lane 2) were separated on a 1.2 % agarose gel alongside molecular size markers (lane M, GeneRuler 1 kb DNA Ladder, Fermentas) and analyzed with EtBr staining. The major bands correspond to the full length *hym* gene- (3356 bp, lane 1) and cDNA-product (2536 bp, lane 2). The minor bands are technical artefacts with variable repeat numbers caused by the tandem repeats.

All putative precursor peptides deduced from the different amplification products are composed of a signal peptide of 19 amino acids (aa), a propeptide of 26 aa and a mature peptide region of differing length in dependence of the repeat number. The latter seems to be further processed into multiple mature AMPs in accordance with the presence of proprotein convertase cleavage sites Arg (R) / Lys (K), as predicted by ProP 1.0 (Duckert et al. 2004). Several different hymenoptaecin peptide variants could be deduced from the analyzed ESTs. Considering all obtained cDNA sequences and the size of the major product, a major *hym* mRNA of 2536 bp was deduced (Acc. No.: HQ315784) containing an ORF of 2373 bp (781 aa) corresponding to six repeats of the putative mature peptide sequence (performed by Chunguang Liang, Ratzka et al. 2012a). Each of the repeated units consists of a coding sequence for the mature hymenoptaecin peptide, preceded by coding regions for a spacer sequence (EAEP) and a putative proprotein cleavage site (RR or KR) (Figure 10). Further analysis of the deduced mature peptides showed that all putative hymenoptaecin peptides are 97 aa long and start with a glutamine (Q) at their N-terminus. Only the first peptide of each precursor, the so called hymenoptaecin-like peptide, displayed an exception consisting of 108 aa due to an N-terminal insertion (Figure 9) and starting with a glycine (G). In sum, the domain composition of the *Camponotus* hymenoptaecin resembled the multipeptide precursor structure of bee apidaecins (Casteels-Josson et al. 1993).

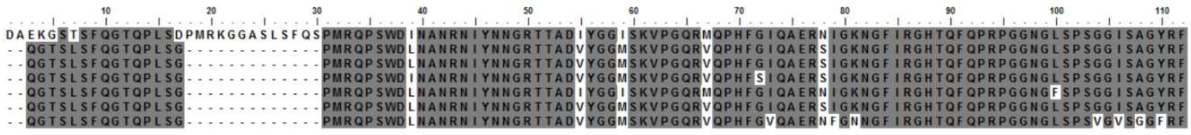

Figure 9. Alignment of HLD (hymenoptaecin-like domain) and all HDs (hymenoptaecin domains) from the same *C. floridanus* hymenoptaecin multipeptide precursor protein (Ratzka et al. 2012a). Grey boxes indicate conserved residues. The insertion in the HLD (top) is clearly visible.

6.3.2 Genomic organization of *hym*

As described above, by comparison of all possible repeat versions from the sequenced hymenoptaecin cDNAs, several different deduced hymenoptaecin peptide variants were found. Therefore the question was raised whether this diversity was caused by the existence of multiple *hym* genes in the *C. floridanus* genome or by alternative splicing of large transcripts from a single gene. To solve this question, the *hym* gene(s) was amplified by PCR using conditions for the amplification of large fragments. Similar to the amplification of *hym* cDNAs the PCR with gDNA yielded a ladder of amplified products ranging from 1500 bp to 3356 bp in length (Figure 8). Comparing gDNA to mRNAs, one phase 0 intron of 820 bp in size was identified, which is located after the codon coding for histidine number 39 (Figure 10) (Ratzka et al. 2012a).

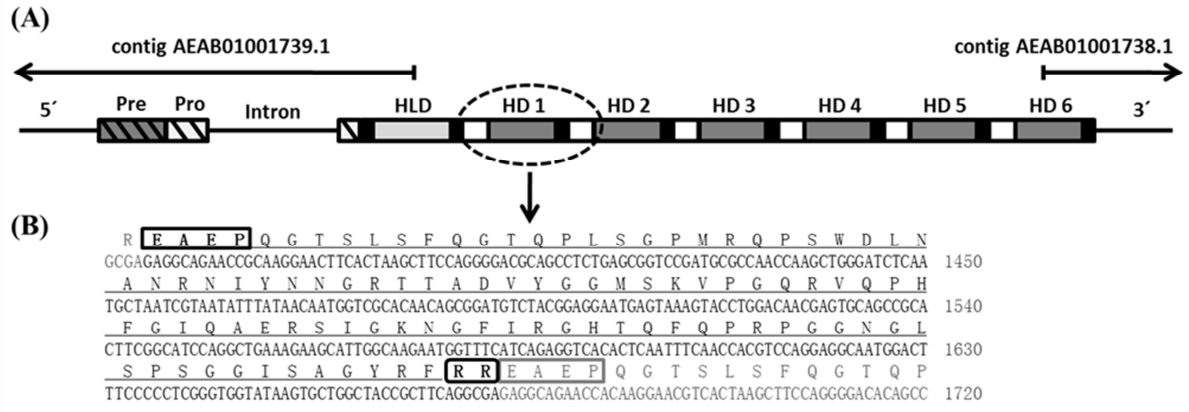

Figure 10. (A) Schematic structure of the *hym* gene containing a single intron within the region coding for the hymenoptaecin propeptide (Ratzka et al. 2012a). The deduced multipeptide precursor peptide consists of a signal-sequence (Pre, grey hatched box) and a pro-sequence (Pro, white hatched box), followed by a hymenoptaecin-like domain (HLD, light grey box) and six repeated hymenoptaecin domains (HD 1-6, dark grey boxes). The hymenoptaecin domains are flanked by the two putative processing sites EAEP (white boxes) and RR (black boxes). (B) The nucleotide and deduced amino acid sequence of a hymenoptaecin repeat unit are shown and the putative processing sites are boxed.

As no other introns were found, the observed variation in repeat numbers could not be explained by alternative splicing of exons coding for the repeats. Therefore, it was investigated, whether hymenoptaecin variants may be encoded by a multigene family. In Southern Blot analysis of *Bsu15*I-digested DNA from multiple ants a fragment of about 2.9 kb in size hybridized with a cDNA probe derived from the repeat sequences (Figure 11A). The size of this fragment was consistent to that expected from digestion of the main PCR-product containing six repeats (Ratzka et al. 2012a).

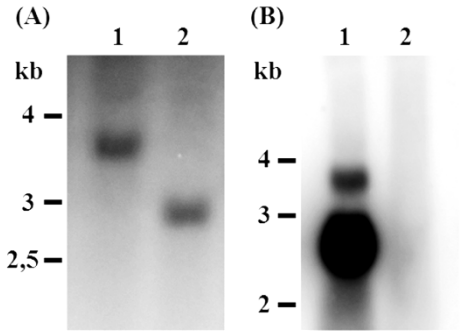

Figure 11. (A) Southern blot with *C. floridanus* genomic DNA using a ^{32}P-labelled *hym* fragment corresponding to one of the repeats as a probe (see section 5.13) (Ratzka et al. 2012a). Genomic DNA (35 µg per lane) was digested with EcoRI (lane 1) and with Bsu15I (lane 2), separated by gel electrophoresis and hybridized with the above mentioned DNA fragment. (B) Northern blot with total RNA of *C. floridanus* using a ^{32}P-labelled cDNA fragment corresponding to the 5`end of the *hym* gene as a probe (see section 5.14). Total RNA (25 µg per lane) was isolated from midgut and fat body of major workers injected with heat-killed *S. marcescens* (2 x 10^5 bacteria/ant) in the haemocoel (lane 1) or untreated animals (lane 2). The major band corresponds to the spliced mature transcript, while the minor band very likely is the unspliced precursor. The position of molecular size markers is indicated on the left side of each figure. All hybridizing bands have the expected molecular size.

Since no other signal was found, this result suggested that hymenoptaecin is encoded by a single gene which comprises six repeated sequence motifs (Figure 10). This was also confirmed by Northern Blot analysis which resulted in a major band of the expected size of the mature transcript (2624 bp) and a minor band of larger molecular size, which probably represents the unspliced primary transcript, since its size perfectly matches the predicted size of 3444 bp (Figure 11B). In sum, the data suggested the existence of a single *C. floridanus hym* gene (Acc. No.: HQ315784) and it was supposed that the above described variable repeat numbers after PCR amplification were a technical artefact caused by the tandem repeats (Ratzka et al. 2012a).

6.3.3 Phylogenetic analysis of hymenoptaecin peptides and comparison of the different hymenoptaecin multipeptide precursors

The recently established genome sequences of six different ant species revealed the presence of at least one antimicrobial peptide belonging to the hymenoptaecin family in each ant species (Bonasio et al. 2010; Nygaard et al. 2011; Smith et al. 2011a; Smith et al. 2011b; Suen et al. 2011). For some of the ant species hymenoptaecin proteins were annotated, but the sequences were incomplete, e.g. the predicted proteins showed a lack of crucial elements like recognition sequences for signal- and propeptides. Therefore, the predicted proteins were investigated on genome level using the published genome drafts. On genome level the problems, which lead to the wrong prediction results could be identified. The sequence region, which apparently codes for the hymenoptaecin peptide(s) in the species *Atta cephalotes*, *Linepithema humile*, *Pogonomyrmex barbatus* and *Solenopsis invicta* seems to span contig boundaries, which were filled with N's during the scaffolding (Ratzka et al. 2012a). This is based on the general problem to assemble short read sequences from next generation sequencing methods through regions with repetitive sequence elements. Three different hymenoptaecin precursor proteins were predicted for the genome of the ant species *Acromyrmex echinatior* (GenBank Acc. No.: EGI65977, EGI65978 and EGI65979) (Nygaard et al. 2011). However, from the analysis of the *A. echinatior* genome one putative hymenoptaecin multipeptide precursor could be deduced, which combines the three predicted ones due to missing stop codons between the predicted proteins (Ratzka et al. 2012a). The obtained hymenoptaecin peptide region of this precursor is extremely long and codes for 23 putative mature hymenoptaecin peptides (Figure 12F). The two annotated hymenoptaecin precursor proteins from *Harpegnathos saltator* seemed to be plausible and contain either four (GenBank Acc. No.: EFN79831) or six (GenBank Acc. No.: EFN79832) mature AMPs, respectively.

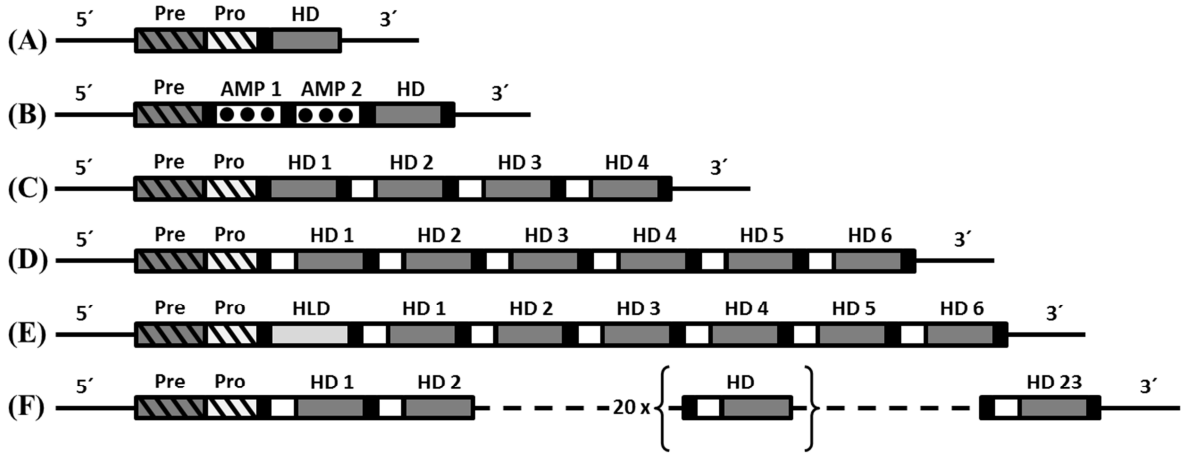

Figure 12. Schematic structure of the hymenoptaecin precursors from different hymenopteran species (Ratzka et al. 2012a): (A) *A. mellifera* (GenBank Acc. No.: NP_001011615) or *Bombus ignitus* (GenBank Acc. No.: ACA04900); (B) *N. vitripennis* (GenBank Acc. No.: NP_001165829 XP_001607881); (C) *Harpegnathos saltator* 1 (GenBank Acc. No.: EFN79831); (D) *H. saltator* 2 (GenBank Acc. No.: EFN79832); (E) *C. floridanus* (GenBank Acc. No.: HQ315784); (F) *Acromyrmex echinatior* (hymenoptaecin multipeptide precursor deduced from genome draft). The various domains are marked as follows: signal-sequence (grey hatched box), pro sequence (white hatched box), hymenoptaecin-like domain (HLD, light grey box), hymenoptaecin domains (HD 1-6, dark grey boxes), proline-rich AMP-like peptide (AMP 1-2, white dotted boxes). The hymenoptaecin domains are flanked by the putative processing sites EAEP (EANP for *Harpegnathos*) (white box) and RR (or RxxR) (black box).

Interestingly, the deduced hymenoptaecin precursor proteins from the examined ant species are all multidomain proteins with remarkable similarities to the multipeptide precursor of *C. floridanus* hymenoptaecin (Figure 12). They possess varying numbers of hymenoptaecin domains (HDs), which are all flanked by the putative spacer region EAEP (EANP for *Harpegnathos*) and processing sites RR (or RxxR, as predicted by ProP 1.0). The phylogenetic analysis of the hymenoptaecins based on the existing set of proteins suggests an intra-species accumulation of the single domains within the proteins (Figure 13; see also section 8.2, Figure 42).

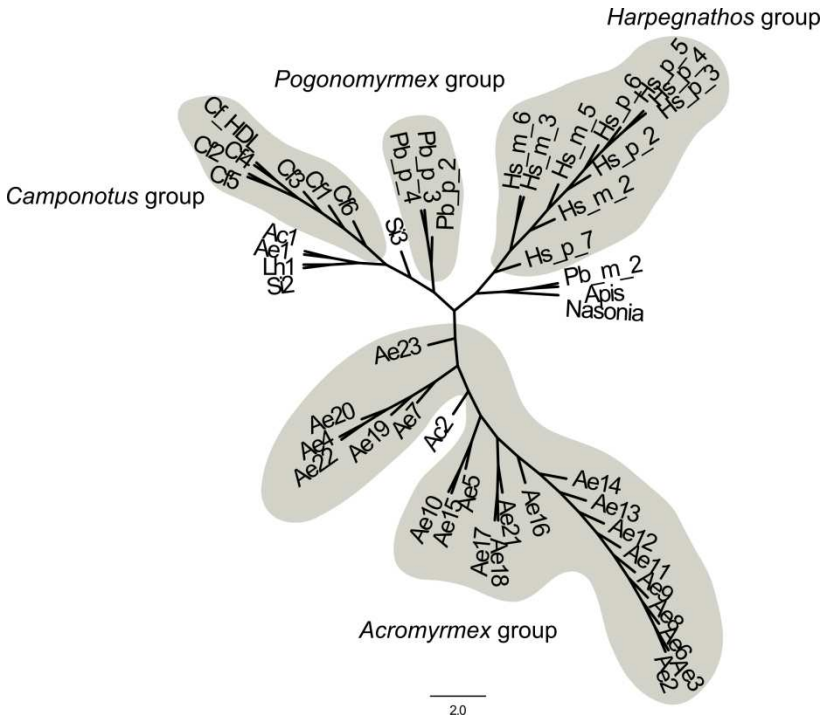

Figure 13. Phylogenetic analysis of hymenoptaecin domains from different ant species (performed by Frank Förster) (Ratzka et al. 2012a). Shown is the unrooted tree of the single domains of the hymenoptaecins of the ant species, *N. vitripennis*, and *A. mellifera*. The proteins were cleaved at the sites predicted by ProP, followed by the alignment by translatorX. The tree was reconstructed by PhyML with a GTR+I+G+F model with 100 bootstrap replicates. The domains of the species with a complete hymenoptaecin protein form clades and are named as groups according their genus name and are indicated by their grey background. The domains which are outside these groups result from missing data from the predicted genes. The gene models are incomplete due to long N-stretches in the genomic sequences based on the scaffolding process. Nevertheless, the distinct groups formed by the complete proteins suggest an intra-species mechanism for the accumulation of the single domains.

6.4 Characterization of *defensin* (*def*) genes from *C. floridanus*

6.4.1 Cloning and sequence analysis of *C. floridanus defensins*

The SSH approach to identify immune inducible genes as well as the screening of the genome sequence resulted in the identification of two different sequences coding for defensin-like AMPs in *C. floridanus*. Phylogenetic analysis allocated these sequences as homologues to defensin-1 and defensin-2 from *A. mellifera*.

5´RACE and 3´RACE of the *defensin-1* (*def-1*) cDNA suggested a full length mRNA sequence of 535 bp (Acc. No.: JN989495), which was verified by amplification with primers near the 5´- and 3´-ends. The deduced *C. floridanus* defensin-1 prepropeptide is 102 amino acids (aa) long, including a signal peptide of 17 aa and a propeptide of 40 aa, followed by a mature peptide of 44 aa, as predicted by ProP 1.0 (Duckert et al. 2004). The subsequent cloning of the corresponding defensin gene revealed the presence of three exons (64, 229 and 13 bp) and two introns (399 and 360 bp). The first intron is a phase 1 intron, which is located within the codon of glutamic acid number 22. The second intron is a phase 2 intron at the alanine residue number 98 following a so-called CXC motif characteristic for defensins (Figure 14A).

In contrast the *defensin-2* (*def-2*) gene contains only two exons (97 and 194 bp), which are separated through one phase 1 intron (1043 bp) located within the codon of threonine number 33 (Figure 14B). The full length mRNA sequence of *def-2* (Acc. No.: JQ693412) is 1238 bp long and encodes a prepropeptide of 97 aa consisting of a signal peptide of 18 aa, a propeptide of 36 aa and a mature defensin-2 peptide of 43 aa, as predicted by ProP 1.0 (Duckert et al. 2004).

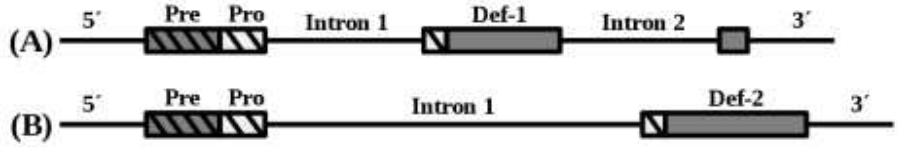

Figure 14. Schematic structure of the *def* genes from *C. floridanus* (Ratzka et al. 2012a). (A) The gene encoding *def-1* (GenBank Acc. No.: JN989495) is composed of three exons and two introns. The first intron is located within the region coding for the propeptide and the second is located within the region coding for the mature defensin-1 peptide. (B) The gene encoding *def-2* (GenBank Acc. No.: JQ693412) contains only one intron, which is also located within the propeptide coding region. Both deduced precursor peptides consist of a signal-sequence (Pre, grey hatched box) and a pro-sequence (Pro, white hatched box), followed by the mature defensin peptide (Def, grey box).

6.4.2 Phylogenetic analysis of defensin peptides from different ant species

The comparison of the deduced amino acid sequences of the *C. floridanus* defensins with other ant defensins suggests a duplication event in the last common ancestor (LCA) of the bee *A. mellifera* and the ant species (Figure 16). Nevertheless, almost all ant species have lost one

of their defensin proteins. Only *C. floridanus* and *S. invicta* possess two defensins. Moreover, most of the proteins forming a clade with the *A. mellifera* defensin-1 were formerly described as defensin-2 homologues. Therefore, these proteins should be renamed to defensin-1 (Figure 15; see also section 8.3, Figure 43). Furthermore, an additional duplication event of defensin-1 peptide seems to have taken place in the LCA of *Lasius niger*, *C. floridanus* and *Formica aquilonia*, due to the grouping of the defensin-1 proteins in the phylogenetic tree (Figure 15), which does not represent the species tree (Figure 16). Nevertheless, the species lost either their defensin-1a (*C. floridanus*) or their defensin-1b (*L. niger* and *F. aquilonia*; Figure 16) (Ratzka et al. 2012a).

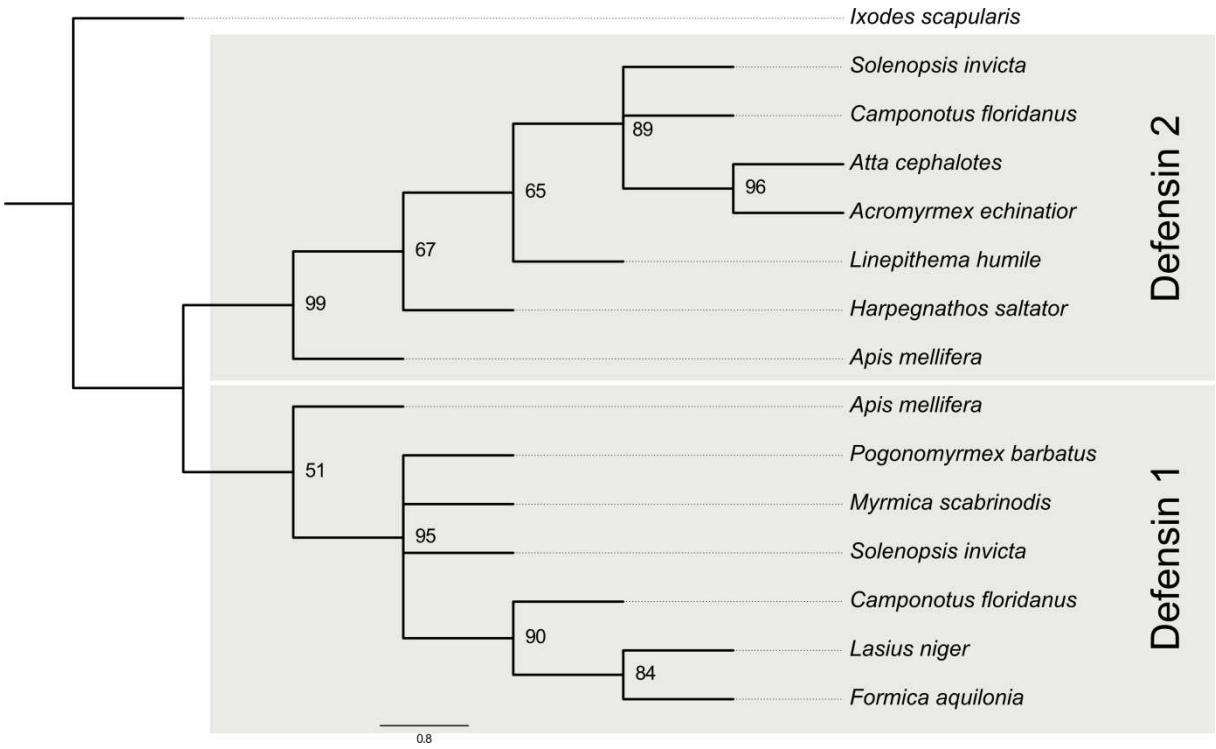

Figure 15. Phylogenetic analysis of defensins from different ant species (performed by Frank Förster) (Ratzka et al. 2012a). All defensin sequences were aligned by muscle and a BioNJ-tree with 100 bootstrap replicates was calculated. Branches with a bootstrap support below 40 were removed. Other bootstrap values are indicated. The genes for the gene tree were generated using the gene prediction pipeline maker and hand curated. The gene tree was rooted at the defensin from *Ixodes scapularis* (GenBank Acc. No.: XP_002436104.1). The *A. mellifera* defensin-1 forms a clade with proteins formerly described as defensin-2, which gives a first indication that they could be renamed accordingly. However, some of the proteins form a clade with the *A. mellifera* defensin-2. Moreover, two species *S. invicta* and *C. floridanus*, own both defensins. Therefore, the data suggests a duplication event at the LCA of *A. mellifera* and the ant species.

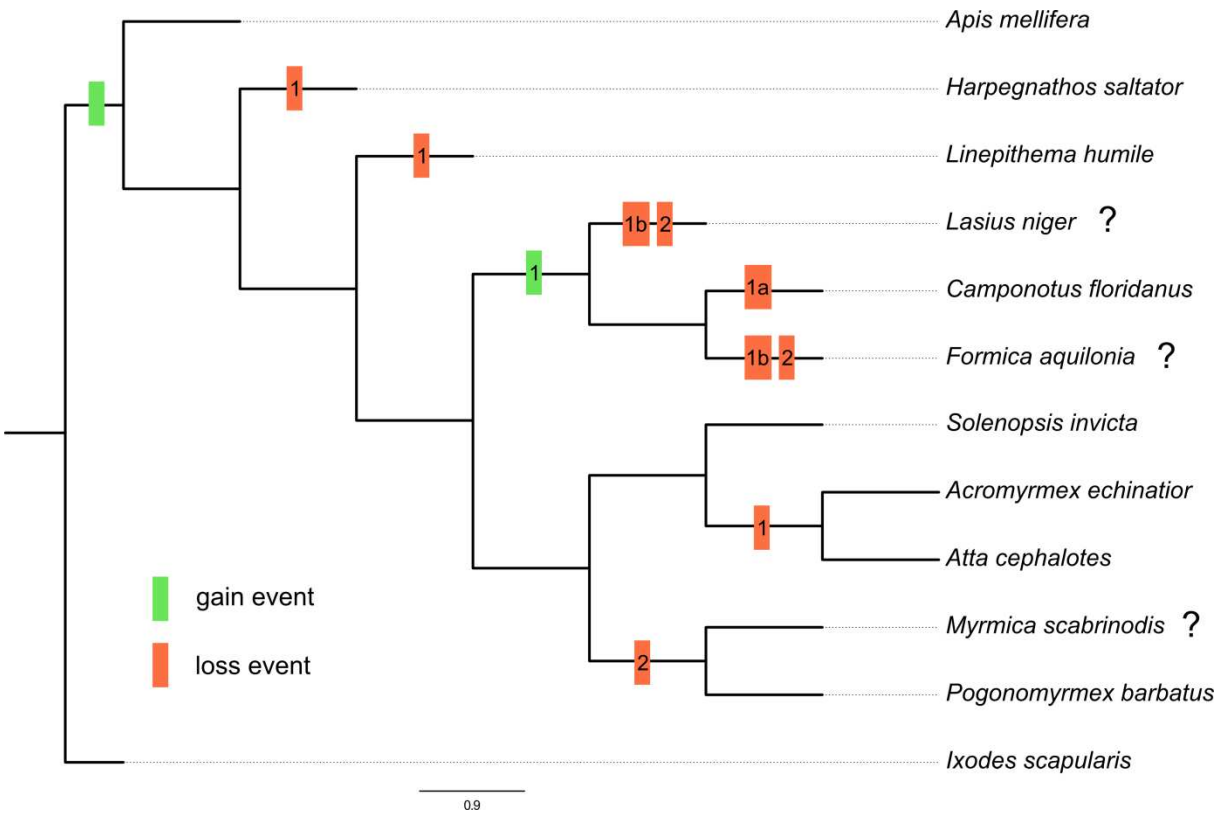

Figure 16. Reconciled species tree of ant defensins (performed by Frank Förster) (Ratzka et al. 2012a). This tree was generated by Notung. The ant species *L. niger* (GenBank Acc. No.: ACB46517.1), *Myrmica scabrinodis* (GenBank Acc. No.: ACB46524.1), and *F. aquilonia* (GenBank Acc. No.: Q5BU36.1) are examples for the sequences generated by (Viljakainen and Pamilo 2005; Viljakainen and Pamilo 2008). For these species no whole genome is available, which is indicated by the question mark behind the species. The gain and loss events are indicated by green and red boxes. The gain event at the LCA of *A. mellifera* and the ant species generated the defensin-1/2 peptide. The LCA of *L. niger*, *C. floridanus* and *F. aquilonia* had an additional duplication event of its defensin-1 peptide, but the species lost either their defensin-1a or their defensin-1b.

6.5 Characterization of *C. floridanus tyrosine hydroxylase* (*TyrOH*) full length mRNA

6.5.1 Cloning of *TyrOH* encoding full length cDNA

5´RACE and 3´RACE of *TyrOH* cDNA suggested a full length mRNA sequence of 3531 bp including 387 bp of 5´untranslated region (UTR), an open reading frame of 1716 bp and 1428 bp of 3´UTR (with polyA-tail). Full length mRNA sequence was verified by amplification with primers near the 5´- and 3´-ends, cloning into pGEM and sequencing of three independent clones. Furthermore, full length mRNA was also partially covered by three ESTs (GenBank Acc. No. HS410955, HS410956 and HS410957) obtained from SSH.

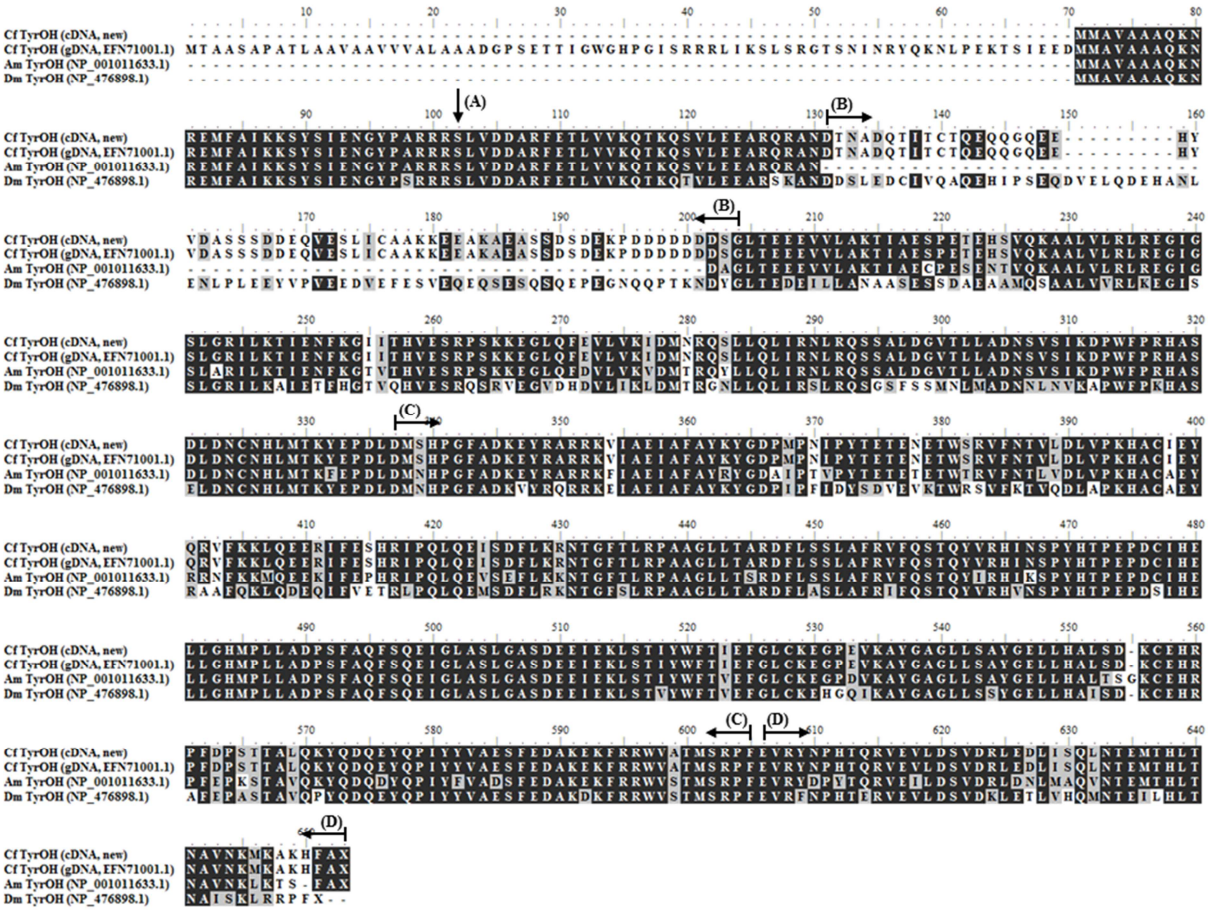

Figure 17. Amino acid alignment of TyrOH sequences from *C. floridanus* (*Cf*) (top: TyrOH protein sequence deduced from cDNA obtained in this work, below: annotated TyrOH protein sequence from genome (GenBank Acc. No.: EFN71001.1)), from *A. mellifera* (*Am*: GenBank Acc. No.: NP_001011633.1) and from *D. melanogaster* isoform B (*Dm*: GenBank Acc. No.: NP_476898.1). Identical residues are underlaid with a black background and similar ones with a grey background. (A) The vertical arrow indicates a conserved serine residue, which is known to be phosphorylated by cAMP-dependent protein kinase in *D. melanogaster* (Vie et al. 1999). The arrows (B) mark an acidic region with low sequence conservation, which is encoded by two alternatively spliced exons in *D. melanogaster* and missing in *A. mellifera*. The arrows (C) comprise the putative catalytic domain. The carboxyl-terminal region defined by arrows (D) is a predicted tetramerization domain. (The putative domain boundaries were selected based on the experimentally determined domain boundaries of rat TyrOH (Goodwill et al. 1997), as already suggested by (Gorman et al. 2007).)

The deduced *C. floridanus* TyrOH protein is 572 aa long. Comparison of the determined coding sequence with the annotated *C. floridanus* TyrOH (GenBank Acc. No. EFN71001.1**)** revealed that the 5´UTR of *TyrOH* gene is spliced differently than predicted. The resulting ORF encodes a TyrOH protein, which is 70 aa shorter than the previously assumed one (Figure 17). Alignment of different TyrOH proteins also suggests that *C. floridanus* TyrOH is 572 aa long (instead of 642 aa) and starts with MM. In *D. melanogaster* TyrOH is regulated via phosphorylation of serine residue 34 by a cAMP-dependent protein kinase (Vie et al. 1999). This serine residue is conserved in *C. floridanus* as well as in *A. mellifera* TyrOH (GenBank Acc. No. NP_001011633.1) (Figure 17A). Like the hypodermal form (isoform B) of *D. melanogaster* TyrOH (GenBank Acc. No. NP_476898.1), *C. floridanus* TyrOH contains an acidic region (isoelectric point = 3.67, calculated by http://web.expasy.org/compute_pi/), which is supposed to stabilize the enzyme in an activated conformation (Vie et al. 1999) (Figure 17B). The putative catalytic domain as well as the carboxyl-terminal region, which is probably needed for tetramer formation (Goodwill et al. 1997), are both well conserved between different insects (Figure 17C and D).

6.5.2 Expression level of *TyrOH* gene in different developmental stages

Expression of *TyrOH* gene was determined in whole untreated *C. floridanus* larvae (L2), pupae (P3) and adult workers (W2) (Figure 18). Two-sided t-test detected that relative *TyrOH* expression levels (dCt-values) were significantly increased in P3 in comparison to expression levels in L2 ($p = 0.0104$) and W2 ($p = 0.0007$).

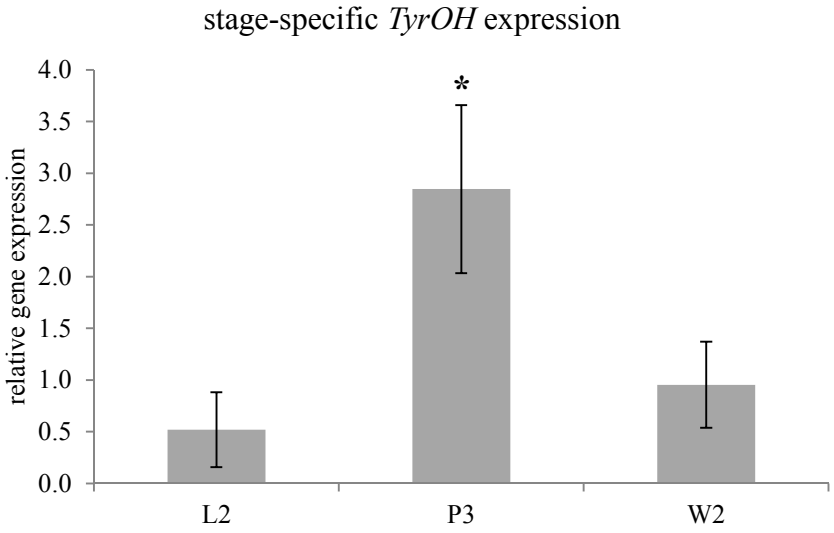

Figure 18. Relative *TyrOH* gene expression in *C. floridanus* whole larvae (L2), pupae (P3) and adult workers (W2). Average *TyrOH* expression levels relative to the constitutive *rpL32* gene (dCt = Ct(target) - Ct(*rpL32*)) of six independent samples were determined and tested for significant differences (indicated as asterisks, two-sided t-test, $p < 0.05$). Columns show *TyrOH* expression levels normalized to the expression of one sample from larvae using the 2^{-ddCt} method.

6.5.3 Knockdown of *TyrOH* gene expression through RNAi

The significantly increased expression levels of *TyrOH* in the late pupal stage suggested a role for cuticle sclerotization and pigmentation as already shown in *T. castaneum* (Gorman and Arakane 2010). In order to test, whether reduction of TyrOH results in decreased cuticle pigmentation in *C. floridanus*, early pupae (P2: shortly after metamorphosis, but still uncolored) were injected with 5 µg dsRNA of *TyrOH* gene. Negative control insects were injected with similar amounts of dsRNA from *GFP* gene or with pure H_2O. Cuticle pigmentation and *TyrOH* expression levels were determined 4 and 7 days post injection. However, no different phenotype could be observed in dsTyrOH-injected individuals in comparison to controls (data not shown) and TyrOH expression levels were also not decreased indicating that RNAi effect could not be achieved.

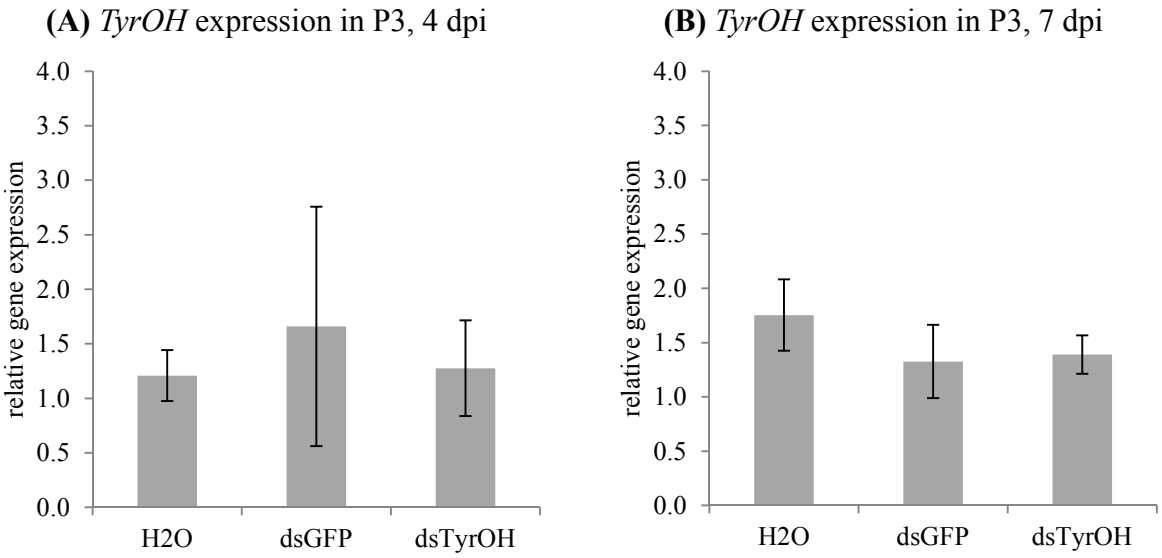

Figure 19. Relative *TyrOH* gene expression in *C. floridanus* pupae 4 days post injection (dpi) (A) and 7 dpi (B) of 0.5 µl nuclease-free H_2O either pure or containing dsRNA of *GFP* or *TyrOH* gene at a concentration of 10 µg/µl. Average *TyrOH* expression levels relative to the constitutive *rpL32* gene (dCt = Ct(target) - Ct(*rpL32*)) of four independent samples were determined. No significant differences of *TyrOH* gene expression levels after injection of different solutions could be detected using two-sided t-test ($p < 0.05$). Columns show *TyrOH* expression levels normalized to the sample with lowest expression using the 2^{-ddCt} method.

6.6 Expression kinetics of some selected genes

For time courses of immune gene expression, the transcriptional upregulation of selected genes was determined in response to a Gram-positive (*Micrococcus flavus, Mf*) and a Gram-negative bacterium (*Serratia marcescens, Sm*) at different time points after injection by qRT-PCR. To achieve a broad overview of the expression kinetics, several genes were chosen that may contribute to different aspects of the immune response: recognition (*Gram-negative binding protein* (*GNBP*), GenBank Acc. No.: HS410970), signaling (serine proteases *snake* (*sna*), GenBank Acc. No.: HS410982 and *stubble* (*stu*), GenBank Acc. No.: HS410979), regulation (*serpin* (*spn*), GenBank Acc. No.: HS410959), immune defense (*defensin* (*def-1*), GenBank Acc. No.: HS410966 and *hymenoptaecin* (*hym*), GenBank Acc. No.: HS410972), melanization (*tyrosine hydroxylase*, GenBank Acc. No.: HS410957), cellular immune response (*malvolio* (*mvl*), GenBank Acc. No.: HS410963) and stress response (*heat shock protein* (*hsp*), GenBank Acc. No.: HS410992).

Two factorial ANOVA detected significant differences in normalized gene expression changes ((-ddCt)-values) for the genes *GNBP* and *hsp* over time (*GNBP*: $F = 89.99$, $p < 0.0001$; *hsp*: $F = 29.86$, $p < 0.0001$) and between bacteria injected (*GNBP*: $F = 12.34$, $p = 0.0018$; *hsp*: $F = 5.92$, $p = 0.0227$), as well as in the interaction time X bacteria (*GNBP*: $F = 10.79$, $p = 0.0001$; *hsp*: $F = 11.95$, $p = 0.0001$). Both genes showed a significant increase of relative expression levels from 2 to 8 hpi (Figure 20A and B, Tukey´s HSD: *GNBP*(2h*Mf*/8h*Mf*): $p = 0.0001$; *GNBP*(2h*Sm*/8h*Sm*): $p = 0.0001$; *hsp*(2h*Mf*/8h*Mf*): $p = 0.0002$; *hsp*(2h*Sm*/8h*Sm*): $p = 0.0003$). From 8 to 24 hpi the expression level of *GNBP* already declined significantly for *Serratia*-injected ants (Tukey´s HSD: (8h*Sm*/24h*Sm*): $p = 0.0006$), whereas its expression remained on a significantly higher level in *Micrococcus*-injected animals (Tukey´s HSD: (24h*Mf*/24h*Sm*): $p = 0.0026$). This influence of the injected bacterial species on gene expression was even more pronounced at 48 hpi. At this time point the expression of the both genes had decreased to levels similar to 2 hpi in *Serratia*-injected ants. In contrast, these genes were still expressed on a significantly higher level in *Micrococcus*-injected individuals (*GNBP*(48h*Mf*/48h*Sm*): $p = 0.0036$; *hsp*(48h*Mf*/48h*Sm*): $p = 0.0002$). The prolonged upregulation of *GNBP* after injection of the Gram-positive *Micrococcus* fits the assumed function of *GNBP* in sensing Gram-positive bacteria via activation of the Toll pathway (see section 6.1). Furthermore it is known that Toll target genes exhibit a late and sustained expression pattern (Lemaitre and Hoffmann 2007). The prolonged activation of *hsp* after *Micrococcus*-injection, which indicates a sustained stress response, is perhaps also linked to this retained immune activation via *GNBP*.

For the serine protease *snake* (*sna*) (Figure 20C) the influence of the factor time on gene expression (ANOVA: $F = 23.36$, $p < 0.0001$), as well as the interaction time X bacteria (ANOVA: $F = 3.85$, $p = 0.022$) was significant. Independent from the injected bacterial species, the normalized expression levels ((-ddCt)-values) rose significantly from 2 to 8 hpi (Tukey´s HSD: (2h*Mf*/8h*Mf*): $p = 0.0007$, (2h*Sm*/8h*Sm*): $p = 0.0028$) and remained on this

higher levels leastwise till 24 hpi (2h*Mf*/24h*Mf*): p = 0.0003; (2h*Sm*/24h*Sm*): p = 0.0033). At 48 hpi Tukey´s HSD post hoc test revealed a significantly higher expression level in *Micrococcus*- than in *Serratia*-injected individuals at 48 hpi (*sna*(48h*Mf*/48h*Sm*): p = 0.036). This sustained induction of *sna* after injection of the Gram-positive *Micrococcus* fits to the assumed function of this gene in the proteolytic activation of the Toll pathway after sensing of Gram-positive bacteria via *GNBP*.

Two factorial ANOVA detected significant differences in normalized gene expression changes ((-ddCt)-values) for the gene *spn* (Figure 20D) over time (F = 18.74, p < 0.0001) and between bacteria injected (F = 15.13, p = 0.0007), as well as in the interaction time X bacteria (F = 6.76, p = 0.0018). The normalized expression levels increased significantly from 2 to 8 hpi (Tukey´s HSD: *spn*(2h*Mf*/8*Mf*): p = 0.0023; *spn*(2h*Sm*/8h*Sm*): p = 0.004), whereas the expression changes between *Micrococcus*- and *Serratia*-injection were similar. However, an influence of the injected bacterial species on gene expression of *spn* could be detected in the later time course of infection, as this gene was expressed on significantly higher levels in *Micrococcus*- than in *Serratia*-injected individuals at 24 hpi (*spn*(24h*Mf*/24h*Sm*): p = 0.0487) and at 48 hpi (*spn*(48h*Mf*/48h*Sm*): p = 0.0012).

The AMP genes *hym* and *def-1* (Figure 20E, F) showed a similar upregulation in response to both *Serratia*- and *Micrococcus*-injection. Factorial ANOVA analysis revealed an overall influence of time on expression changes (*hym*: F = 3.68, p = 0.0261; *def-1*: F = 12.18, p < 0.0001). Normalized expression changes of *def-1* increased during the time course after injection, which was significant at 24 hpi (Tukey´s HSD: *def-1*(2h*Sm*/24h*Sm*): p = 0.0118; *def-1*(2h*Mf*/24h*Mf*): p = 0.0157). Therefore the *def-1* gene seems to be more strongly induced later in the time course of infection, which has also been shown for *def* in *Apis mellifera* (Casteels-Josson et al. 1994). For *hym* gene post hoc testing did not discover distinct groups. Comparison of relative expression levels in bacteria injected and untreated ants revealed that *hym* was significantly induced at all measured time points (two-sided t-tests: p < 0.01). Furthermore, this AMP displayed the highest expression differences of all investigated genes. The median (-ddCt) reached a maximum of 11.09, corresponding to about 2100 fold gene induction (2^{-ddCt}). In comparison the gene *def-1* was also strongly induced, however the median (-ddCt) only reached a maximum of 8.65, corresponding to about 400 fold gene induction (2^{-ddCt}).

For the genes *stu*, *TyrOH* and *mvl* (Figure 20G, H, I) the two-factorial ANOVA analysis detected no significant impact of the injected bacterial species on the normalized expression changes. This was expected, as these genes are likely to encode proteins involved in the melanization process and cellular immune response (see section 6.1), which are generally induced upon injury and not specifically triggered by a certain class of microorganisms. However, all three genes exhibited distinct expression kinetics over time (*stu*: F = 34.83, p < 0.0001; *TyrOH*: F = 5.04, p = 0.0075; *mvl*: F = 26.04, p < 0.0001). The normalized expression levels ((-ddCt)-values) of *stu* gene rose significantly from 2 to 8 hpi (Tukey´s HSD:

(2h*Mf*/8h*Mf*): p = 0.0003, (2h*Sm*/8h*Sm*): p = 0.0027). Afterwards induction levels decreased again, which was significant at 48 hpi (Tukey´s HSD: (24h*Mf*/48h*Mf*): p = 0.0289, (24h*Sm*/48h*Sm*): p = 0.0004). At this time point the relative expression levels (dCt-values) of bacteria injected ants did not differ from untreated groups. The normalized expression levels ((-ddCt)-values) of *TyrOH* gene were already quite high at 2 hpi (2h*Mf*: median(-ddCt) = 2.93; 2h*Sm*: median(-ddCt) = 3.78). For *Micrococcus*-injected ants the expression of *TyrOH* slightly increased from 2 to 48 hpi (Tukey´s HSD: *TyrOH*(2h*Mf*/48h*Mf*): p = 0.0409), whereas for *Serratia*-injected groups *TyrOH* expression remained similar over all measured time points. In contrast, the normalized gene expression changes of *mvl* increased significantly during the time course and reached maximal levels at 24 hrs after *Micrococcus*-(*mvl*(2h*Mf*/24h*Mf*): p = 0.0002) and *Serratia*-injection (*mvl*(2h*Sm*/24h*Sm*): p = 0.0003).

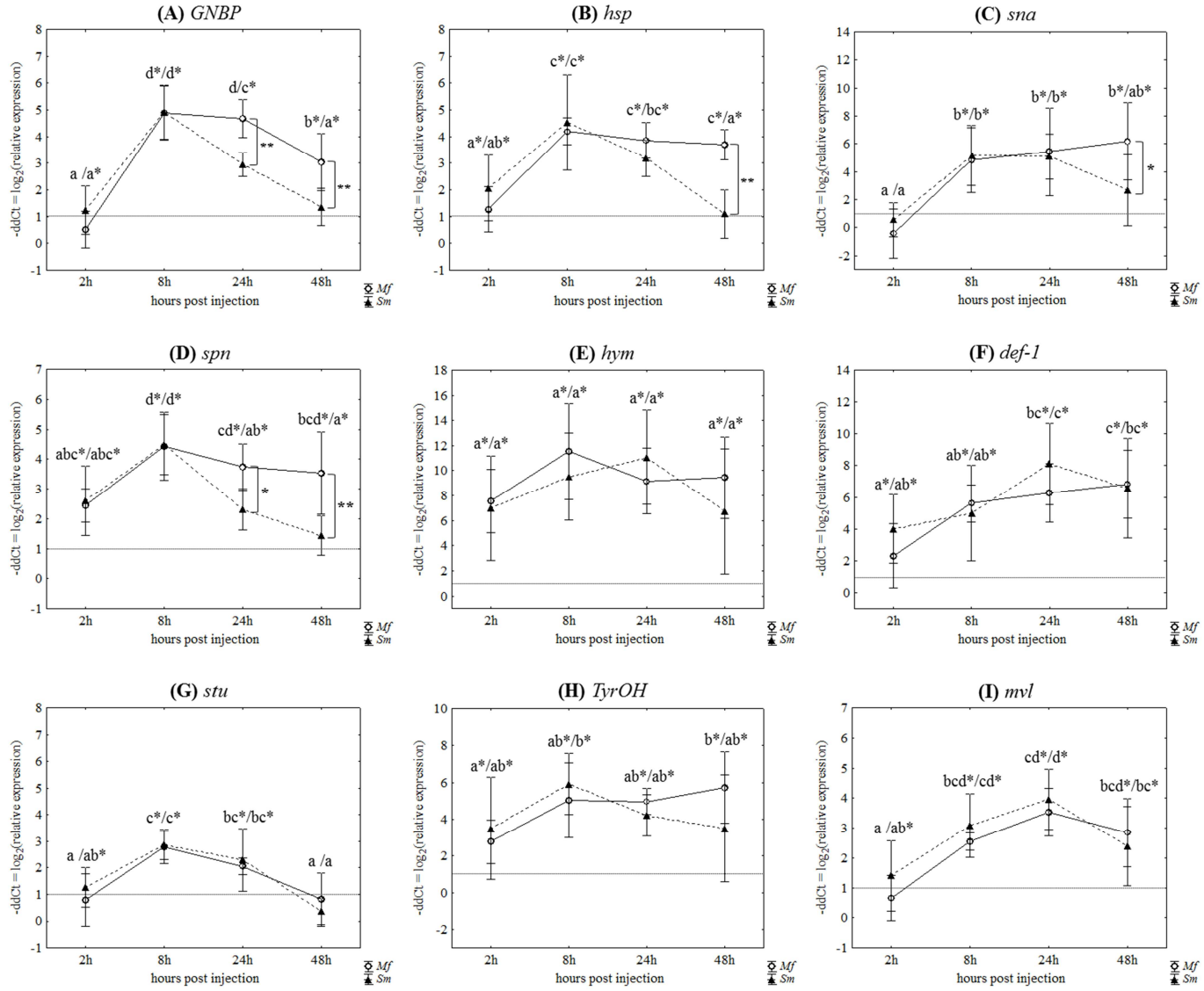

Figure 20. Expression kinetics of infection-inducible genes in fat body and midgut of *C. floridanus* (Ratzka et al. 2011). Major worker ants were injected with a solution of dead *M. flavus* (*Mf*) or *S. marcescens* (*Sm*). Average gene expression levels relative to the constitutive *EF1a* gene (dCt = Ct(target) - Ct(*EF1a*)) of four independent samples were determined from bacteria injected and untreated ants at each time point and tested for significant differences (indicated as asterisks at each letter, two-sided t-test, $p < 0.05$) of the genes (A) *Gram-negative binding protein* (*GNBP*), (B) *heat shock protein* (*hsp*), (C) *snake* (*sna*), (D) *serpin* (*spn*), (E) *hymenoptaecin* (*hym*), (F) *defensin-1* (*def-1*), (G) *stubble* (*stu*), (H) *tyrosine hydroxylase* (*TyrOH*), (I) *malvolio* (*mvl*). In addition normalized changes in gene expression in animals injected with different bacteria (-ddCt = -(dCt(injected) - dCt(untreated)) were tested for influence of time and bacterial species using factorial ANOVA. Groups differing significantly ($p < 0.05$) from each other in Tukey´s HSD post hoc test are marked with different letters (*Mf*-injected / *Sm*-injected ants) (for statistics see section 8.4, Table 17 and Table 18). The dotted line indicates two fold gene induction in bacteria injected groups (2^{-ddCt} = ratio = 2).

6.7 Immune gene expression in response to injection of *B. floridanus*

As the kinetics study (section 6.6) revealed that all genes were significantly upregulated at 8 hpi and 24 hpi, these time points were chosen to further investigate the gene expression in response to injection of the endosymbiont *B. floridanus* (*Bf*). Furthermore the gene induction after injection of a sterile ant Ringer (aR) solution was determined as an additional control.

Relative expression levels of the target genes (dCt-values) from injected and untreated ants revealed that all target genes were significantly induced in bacteria injected ants at both measured time points after injection (two-sided t-test, $p < 0.05$) (Figure 21). In contrast, after sterile Ringer-injection, only the gene *def-1* was significantly upregulated at both time points after injection, whereas the genes *GNBP*, *hsp*, *stu*, *hym* and *mvl* were only significantly induced at 8 hpi (two-sided t-test, $p < 0.05$) and had expression levels similar to untreated ants at 24 hpi. For the genes *sna*, *spn* and *TyrOH* the expression levels after sterile injection were not significantly different from those of untreated ants at 8 and 24 hpi.

A two factorial ANOVA revealed a significant influence of the injected solution on normalized gene expression changes ((-ddCt)-values) for all genes (*GNBP*: $F = 64.27$, $p < 0.0001$; *hsp*: $F = 46.53$, $p < 0.0001$; *sna*: $F = 12.78$, $p < 0.0001$; *spn*: $F = 17.88$, $p < 0.0001$; *hym*: $F = 5.44$, $p = 0.0053$; *def-1*: $F = 3.30$, $p = 0.0376$; *stu*: $F = 32.22$, $p < 0.0001$; *TyrOH*: $F = 12.55$, $p < 0.0001$; *mvl*: $F = 28.54$, $p < 0.0001$). The normalized gene expression changes ((-ddCt)-values) of the genes *GNBP*, *hsp*, *spn*, *hym* and *stu* were also significantly influenced by the factor time (*GNBP*: $F = 105.97$, $p < 0.0001$; *hsp*: $F = 28.69$, $p < 0.0001$; *spn*: $F = 26.5$, $p < 0.0001$; *hym*: $F = 6.77$, $p = 0.0156$; *stu*: $F = 10.53$, $p = 0.0037$). The interaction time X bacteria was significant only for the genes *GNBP* ($F = 10.69$, $p = 0.0001$), *hsp* ($F = 3.29$, $p = 0.0380$) and *spn* ($F = 4.85$, $p = 0.0089$).

Tukey´s HSD post hoc test revealed that the differences were highest between sterile Ringer and bacteria injection (Figure 21). At 8 hpi the genes *GNBP* and *stu* were induced on a lower level after sterile than after injection of bacteria (Tukey´s HSD: *GNBP*: p(8haR/8h*Mf,Sm,Bf*) < 0.004; *stu*: p(8haR/8h*Mf,Sm,Bf*) < 0.02). For the genes *hsp* and *spn* this difference was only significant for injection of the potential pathogens *Micrococcus* (*hsp*: p(8haR/8h*Mf*) = 0.0007; *spn*: p(8haR/8h*Mf*) = 0.0018) and *Serratia* (*hsp*: p(8haR/8h*Sm*) = 0.0002; *spn*: p(8haR/8h*Sm*) = 0.0015), not for the endosymbiont *Blochmannia*. At 24 hpi the expression levels of the genes *GNBP*, *hsp*, *sna, TyrOH* and *mvl* were significantly lower after sterile Ringer injection than after injection of *Micrococcus* (*GNBP*: p(24haR/24h*Mf*) = 0.0001; *hsp*: p(24haR/24h*Mf*) = 0.0001; *sna*: p(24haR/24h*Mf*) = 0.0036; *TyrOH*: p(24haR/24h*Mf*) = 0.0031; *mvl*: p(24haR/24h*Mf*) = 0.0001), *Serratia* (*GNBP*: p(24haR/24h*Sm*) = 0.0001; *hsp*: p(24haR/24h*Sm*) = 0.0001; *sna*: p(24haR/24h*Sm*) = 0.0082; *TyrOH*: p(24haR/24h*Sm*) = 0.021; *mvl*: p(24haR/24h*Sm*) = 0.0001) or *Blochmannia* (*GNBP*: p(24haR/24h*Bf*) = 0.0079; *hsp*: p(24haR/24h*Bf*) = 0.0008; *sna*: p(24haR/24h*Bf*) = 0.0253; *TyrOH*: p(24haR/24h*Bf*) = 0.0063; *mvl*: p(24haR/24h*Bf*) = 0.0002). Thus, sterile injection itself induced an immune response, albeit one that was not as enduring as after septic injury. Furthermore,

Blochmannia-injection caused changes in gene expression, which were significantly higher than sterile injection alone.

Only the AMP genes *hym* and *def-1* exhibited almost the same induction levels after sterile and bacteria injection. Such unspecific AMP induction has also been described for other insects e.g. for *Apis mellifera* (Evans et al. 2006) or for *Drosophila melanogaster* (Boutros et al. 2002).

The response of the genes *GNBP* and *hsp* differed in dependence of the bacteria injected. At 24 hpi the expression of *GNBP* was significantly stronger for the Gram-positive bacterium *Micrococcus* than for the Gram-negative bacteria *Serratia* (24h*Mf*/24h*Sm*: $p = 0.0037$) and *Blochmannia* (24h*Mf*/24h*Bf*: $p = 0.0001$), probably because it plays a role in the recognition of Gram-positive bacteria. Interestingly, the expression levels of *GNBP* 24 h after *Blochmannia*-injection were significantly lower than after *Serratia*-injection (24h*Bf*/24h*Sm*: $p = 0.0084$). A similar pattern was observed for *hsp* expression, which is likely an indicator of the ant´s stress response (see section 6.1). In this case a significantly weaker induction in response to *Blochmannia* was observed at 8 hpi in comparison to *Serratia* (8h*Bf*/8h*Sm* $p = 0.0085$) and at 24 hpi in comparison to *Micrococcus* (24h*Bf*/24h*Sm*: $p = 0.0059$).

In sum, the data indicated that *B. floridanus* was still recognized as non-self by the ant immune system. This was an important result, as the recognition of *Blochmannia* was a prerequisite for the involvement of the host immune system in controlling the endosymbiont.

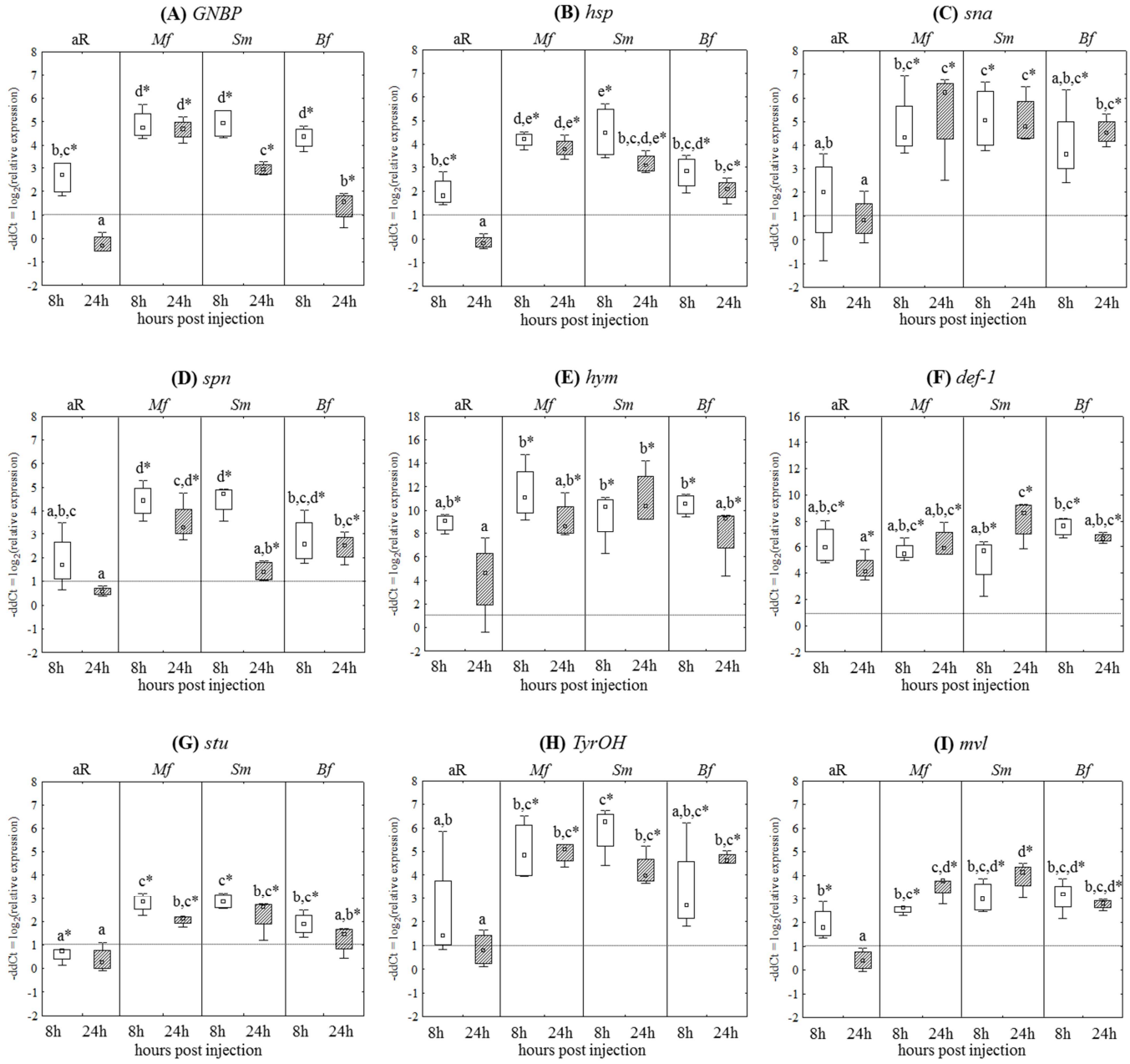

Figure 21. Immune gene expression in fat body and midgut of *C. floridanus* major workers of (A) *Gram-negative binding protein* (*GNBP*), (B) *heat shock protein* (*hsp*), (C) *snake* (*sna*), (D) *serpin* (*spn*), (E) *hymenoptaecin* (*hym*), (F) *defensin-1* (*def-1*), (G) *stubble* (*stu*), (H) *tyrosine hydroxylase* (*TyrOH*), (I) *malvolio* (*mvl*) in response to injection of 0.3 µl ant Ringer solution either sterile (aR) or containing 6-8 x 10^8 dead bacteria/ml (*M. flavus* (*Mf*), *S. marcescens* (*Sm*), or *B. floridanus* (*Bf*)). Average gene expression levels relative to the constitutive *EF1a* gene (dCt = Ct(target) - Ct(*EF1a*)) were determined from bacteria injected and untreated ants at each time point and tested for significant differences (* two-sided t-test, $p < 0.05$). In addition normalized changes in gene expression in animals injected with different bacteria (-ddCt = -(dCt(injected) - dCt(untreated)) were tested for influence of time and bacterial species using factorial ANOVA with Tukey´s HSD post hoc test. Boxes marked with different letters are significantly different ($p < 0.05$) (for statistics see section 8.5, Table 19 and Table 20). The dotted line indicates two fold gene induction in bacteria injected groups (2^{-ddCt} = ratio = 2) (Ratzka et al. 2011).

6.8 Gene expression analysis of endosymbiont-bearing midgut tissue during ontogeny of *C. floridanus*

The number of *Blochmannia* cells per individual ant is quite variable during host ontogeny (Stoll et al. 2010; Stoll et al. 2008). As the previous study revealed that *B. floridanus* is still recognized as non-self (section 6.7), the ant immune system might be involved in controlling the endosymbiont number. In order to detect differences in gene expression, which might be associated with changes in endosymbiont number during host ontogeny, the expression of several genes was investigated in dependence of developmental stage and tissue by qRT-PCR. For this purpose several candidate genes were chosen, which might be involved in host-symbiont-interactions. These included the *C. floridanus* pattern recognition receptor (PRR) genes *PGRP-LB* (GenBank Acc. No.: EFN73971), *PGRP-SC2* (GenBank Acc. No.: EFN73970), *PGRP-LE* (GenBank Acc. No.: EFN63542), *PGRP-2* (GenBank Acc. No.: EFN70060) and *GNBP* (GenBank Acc. No.: HS410970) as well as the NF-κB transcription factor *Relish* (*Rel*, GenBank Acc. No.: EFN61437), the antimicrobial peptide genes *hym* (GenBank Acc. No.: HQ315784) and *def-1* (GenBank Acc. No.: JN989495) and the supposed negative regulator of the immune response *tollip* (GenBank Acc. No.: EFN63773). Furthermore, the expression of candidate genes, possibly involved in the lysosomal system, was analyzed, which included *C. floridanus chicken-type lysozyme 1* (*c-type lyso 1*, GenBank Acc. No.: EFN74565), *invertebrate-type lysozyme* (*i-type lyso*, GenBank Acc. No.: EFN71839), *cathepsin L* (*cath L*, GenBank Acc. No.: EFN65237), a *lysosomal aspartic protease* (*lap*, GenBank Acc. No.: EFN61279) and a *serine carboxypeptidase vitellogenic-like* (*CPVL*, GenBank Acc. No.: EFN61703). To check for induction of autophagy the expression of the *autophagy protein 5* gene (*apg-5*, GenBank Acc. No.: EFN68118) was examined and moreover the expression of genes involved in production and detoxification of reactive oxygen species (ROS) like *ferritin* (GenBank Acc. No.: EFN61070), *Dual oxidase* (*Duox*, GenBank Acc. No.: EFN74201) and *superoxide dismutase* (*SOD*, GenBank Acc. No.: EFN700851) was investigated. For each gene the expression levels in the body tissues (without midgut) were compared with those of the midgut tissue of late larvae (L2), late pupae (P3) and adult workers (W2), as maximal differences in endosymbiont number are present in the midguts of these stages.

6.8.1 Reference gene selection

In order to obtain reliable gene expression levels of target genes, it was necessary to first identify appropriate reference genes of *C. floridanus* for normalization of the qRT-PCR data. Such reference genes should be expressed on similar levels in the investigated stages and tissues. In order to determine the stability of the expression of different reference genes for this study, Ct-values of the genes *ribosomal protein L32* (*rpL32*, GenBank Acc. No.: EFN68969), *ribosomal protein L18* (*rpL18*, GenBank Acc. No.: EFN68908), *elongation*

factor 1-alpha (*EF1α*, GenBank Acc. No.: EFN72500) and *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*, GenBank Acc. No.: EFN69158) were compared for all cDNA samples derived from different stages (L2, P3, W2) and tissues (midgut and residual body) using the *BestKeeper* (*BK*) software tool (Pfaffl et al. 2004). The SD- (standard deviation-) value is the most important criterion for evaluating the stability of reference genes by the *BK* program. The expressions of all reference genes tested in this study showed low Ct variations (0.52 < SD [± Ct] < 0.64) as well as up-/ downregulation (1.43 < SD [± x-fold] < 1.56), indicating that all genes were well suited as reference genes (Table 14). Moreover, the program combined all candidate standards to an index (*BK* index), which was most appropriate (SD [± Ct] = 0.40; SD [± x-fold] = 1.32) and therefore used for internal standardization of target gene expression data in this study.

Table 14. Statistical analyses of four candidate reference genes based on their Ct-values. In the last column the *BestKeeper* (*BK*) index for these reference genes is computed together with the same descriptive parameters.

	rpL32	*EF1a*	*GAPDH*	*rpL18*	*BK*
N	36	36	36	36	36
Geo Mean [Ct]	20.21	19.67	22.36	20.10	20.56
Ar Mean [Ct]	20.22	19.69	22.37	20.11	20.57
Min [Ct]	18.42	18.33	21.07	19.08	19.65
Max [Ct]	21.52	21.64	23.81	21.08	21.55
SD [± Ct]	**0.55**	**0.63**	**0.54**	**0.53**	**0.40**
CV [% Ct]	2.71	3.21	2.40	2.64	1.96
Min [x-fold]	-3.45	-2.54	-2.44	-2.03	1.89
Max [x-fold]	2.48	3.92	2.73	1.97	1.98
SD [± x-fold]	**1.46**	**1.55**	**1.45**	**1.44**	**1.32**

Abbreviations: N: sample size; Geo Mean [Ct]: geometric means of the threshold cycle (CT); Ar Mean [Ct]: the arithmetic mean of Ct; Min [Ct] and Max [Ct]: the extreme values of Ct; SD [±Ct] : the standard deviation of the Ct; CV [%Ct]: the coefficient of variance expressed as a percentage at the Ct level; Min [x-fold] and Max [x-fold]: the extreme values of expression levels expressed as an absolute x-fold over/under-regulation coefficient; SD [±x-fold]: standard deviation of the absolute regulation coefficients.

6.8.2 Expression of immune-related genes in midgut and residual body of different developmental stages

For the PRR genes *PGRP-LB* and -*SC2* two factorial ANOVA detected significant differences in the relative gene expression levels (dCt-values) between different tissues (midgut (m) and residual body parts (b)) (*PGRP-LB*: $F = 254.99$, $p < 0.0001$; *PGRP-SC2*: $F = 4.32$, $p = 0.0462$) and between different developmental stages (*PGRP-LB*: $F = 37.18$, $p < 0.0001$; *PGRP-SC2*: $F = 41.11$, $p < 0.0001$), as well as in the interaction tissue X stage (*PGRP-LB*: $F = 99.39$, $p < 0.0001$; *PGRP-SC2*: $F = 133.47$, $p < 0.0001$) (Figure 22A, B). Both genes were significantly stronger expressed in the midgut of P3 than of L2 (Tukey´s HSD: *PGRP-LB*(mP3/mL2): $p = 0.0005$; *PGRP-SC2*(mP3/mL2): $p = 0.0001$), whereas *PGRP-SC2* was also expressed on a significantly higher level in the midgut of P3 than of W2 (Tukey´s HSD: *PGRP-SC2*(mP3/mW2): $p = 0.0001$). In the residual body, both genes were lowest expressed in P3 (Tukey´s HSD: *PGRP-LB*: p (bL2/bP3) = 0.0001, p (bP3/bW2) = 0.0001; *PGRP-SC2*: p (bL2/bP3) = 0.0003, p (bP3/bW2) = 0.0001). In contrast to *PGRP-LB*, which was significantly stronger expressed in the body of L2 than of W2 (Tukey´s HSD: *PGRP-LB*(bL2/bW2): $p = 0.0261$), *PGRP-SC2* was higher expressed in the body of W2 than of L2 (Tukey´s HSD: *PGRP-SC2*(bL2/bW2): $p = 0.0001$). In sum, both PRR genes were highly expressed in the midgut tissue, especially in P3, where highest *Blochmannia* number is present in this tissue. Most strikingly, in this stage expression of *PGRP-LB* was about 970 fold higher in the midgut tissue than in the residual body parts. Similarly, also *PGRP-SC2* followed this expression pattern, albeit its expression difference was somewhat milder. This gene was expressed about 20 fold more strongly in the midgut tissue than in the residual body during the P3 stage.

For the PRR gene *PGRP-LE* the influence of the factor tissue on gene expression (ANOVA: $F = 12.63$, $p < 0.0013$), as well as the interaction tissue X stage (ANOVA: $F = 5.58$, $p = 0.0087$) was significant. Expression of *PGRP-LE* was lowest in body of L2 (Tukey´s HSD: *PGRP-LE*: p(bL2/mL2) = 0.0007, p(bL2/mP3) = 0.0167, p(bL2/bP3) = 0.0431, p(bL2/mW2 = 0.005). Apart from that, *PGRP-LE* was quite equally expressed in the investigated stages and tissues (Figure 22C).

In case of the other two PRR genes *PGRP-2* and *GNBP* two factorial ANOVA detected significant differences in the relative gene expression levels (dCt-values) between tissues (*PGRP-2*: $F = 502.38$, $p < 0.0001$; *GNBP*: $F = 101.06$, $p < 0.0001$) and between developmental stages (*PGRP-2*: $F = 38.71$, $p < 0.0001$; *GNBP*: $F = 6.55$, $p = 0.0044$), as well as in the interaction tissue X stage (*PGRP-2*: $F = 16.20$, $p < 0.0001$; *GNBP*: $F = 7.06$, $p = 0.0031$) (Figure 22D, E). Both genes exhibited a similar expression pattern. In each stage, expression of both genes was significantly lower in the midgut than in the residual body (Tukey´s HSD: *PGRP-2*: p(mL2/bL2) = 0.0001, p(mP3/bP3) = 0.0001, p(mW2/bW2) = 0.0001; *GNBP*: p(mL2/bL2) = 0.0074, p(mP3/bP3) = 0.0007, p(mW2/bW2) = 0.0001), whereas both genes had maximal expression levels in the body of W2 (Tukey´s HSD:

PGRP-2: p(bL2/bW2) = 0.0001, p(bP3/bW2) = 0.0001; *GNBP*: p(bL2/bW2) = 0.0231, p(bP3/bW2) = 0.0012). This is probably due to the fact, that the body samples included the immune-responsive fat body tissue, where several PRR genes are constitutively expressed on a high level.

Expression of the NF-κB-like transcription factor *Rel* was significantly influenced through the factors tissue (ANOVA: F = 16.44, p = 0.0003) and stage (ANOVA: F = 25.66, p < 0.0001), as well as through the interaction tissue X stage (ANOVA: F = 13.46, p < 0.0001) (Figure 22F). *Rel* was significantly stronger expressed in the midgut of L2 than in the residual body (Tukey´s HSD: p(mL2/bL2) = 0.0001). In P3 and W2 *Rel* expression was equal in midgut and body, whereas on a significantly higher level in W2 than in P3 (Tukey´s HSD: p(mP3/mW2) = 0.0007, p(mP3/bW2) = 0.0013, p(bP3/mW2) = 0.0005, p(bP3/bW2) = 0.0009).

For the AMP genes *hym* and *def-1* two factorial ANOVA detected significant differences in the relative gene expression levels (dCt-values) between different developmental stages (*hym*: F = 30.30, p < 0.0001; *def-1*: F = 22.86, p < 0.0001), as well as in the interaction tissue X stage (*hym*: F = 12.95, p < 0.0001; *def-1*: F = 4.32, p = 0.0216), and in case of *hym* also between different tissues (F = 31.50, p < 0.0001) (Figure 22G, H). *Hym* gene was strongest expressed in body of W2 (Tukey´s HSD: p = 0.0001 for all combinations), however with high variations between the six biological replicates of bW2. Similar variations were also obtained for expression of *def-1* in body of W2 and might indicate different health status of individual ants. Apart from these variations, expression levels of *def-1* were similar in the midgut of L2 and W2, as well as in the body of L2 and W2. *Def-1* expression was lowest in body of P3 (Tukey´s HSD: p(bL2/bP3) = 0.0011, p(bP3/bW2) = 0.0001) and besides *def-1* was significantly less expressed in midgut of P3 than of L2 (Tukey´s HSD: p(mL2/mP3) = 0.0040).

For the *tollip* gene the two-factorial ANOVA analysis detected no significant impact of the factors tissue and stage and of their interaction on the relative expression levels (Figure 22I). Thus, the *tollip* gene was equally expressed in the midgut and body of different stages and does not seem to be involved in mediating host-symbiont-interactions in *C. floridanus*. In contrast, the *tollip* homologue of weevils was strongly expressed in the bacteriome and it was suggested to be a negative regulator of the immune response towards endosymbionts (Anselme et al. 2008).

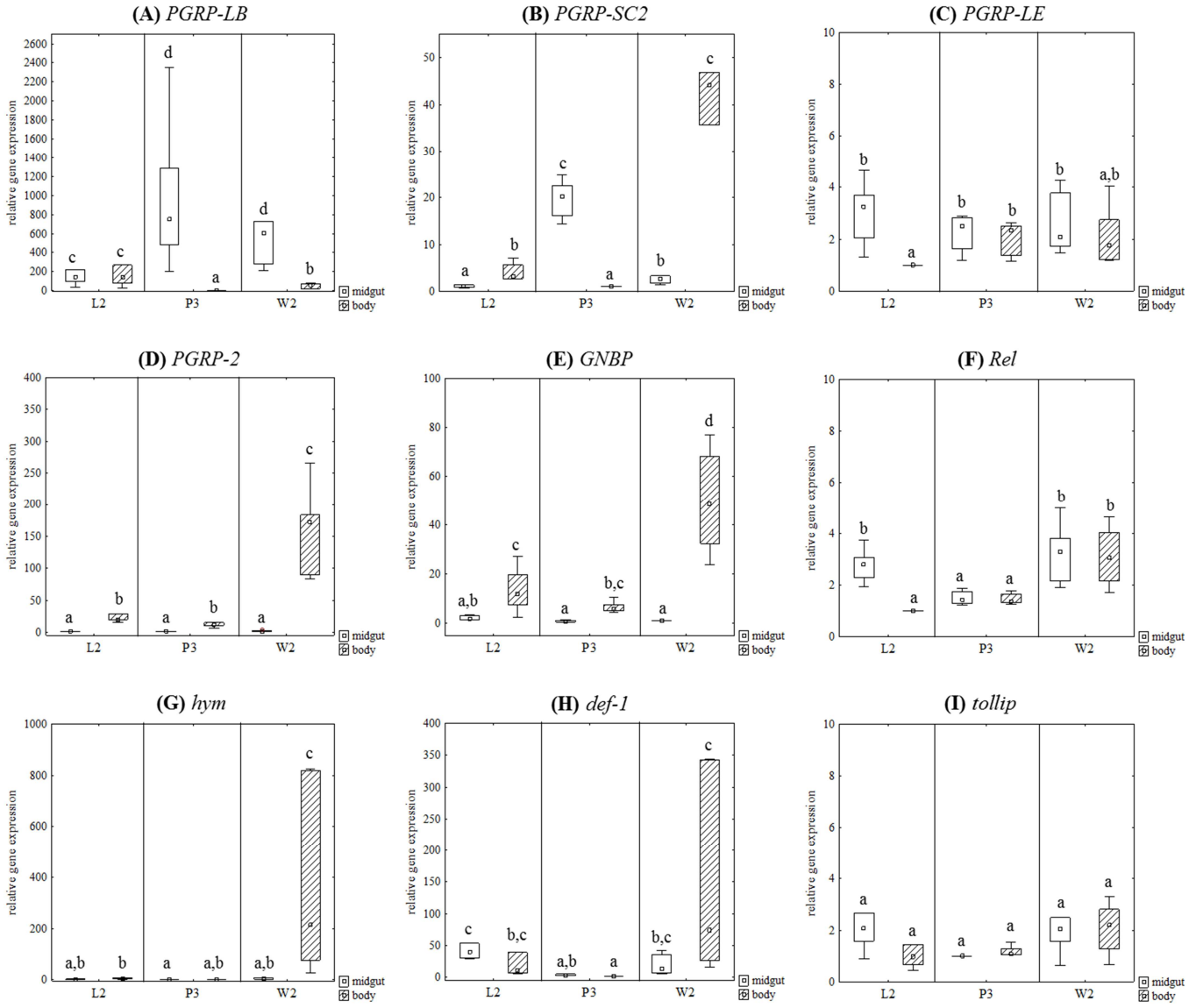

Figure 22. Expression of immune-related genes in midgut (white boxes) and residual body (hatched boxes) of different *C. floridanus* stages (L2 = larvae, P3 = late pupae, W2 = adult workers). Average gene expression levels relative to the *BestKeeper* (*BK*) index (dCt = Ct(target) – Ct(BK)) of six independent samples were determined from midgut and residual body of different stages. These dCt-values of the genes (A) *peptidoglycan recognition protein-LB* (*PGRP-LB*), (B) *PGRP-SC2*, (C) *PGRP-LE,* (D) *PGRP-2*, (E) *Gram-negative binding protein* (*GNBP*), (F) *Relish* (*Rel*), (G) *hymenoptaecin* (*hym*), (H) *defensin-1* (*def-1*), (I) *tollip* were tested for influence of tissue and stage using factorial ANOVA. Groups differing significantly ($p < 0.05$) from each other in Tukey´s HSD post hoc test are marked with different letters (for statistics see section 8.6, Table 21 and Table 22). Box-plots show normalized changes in gene expression relative to the tissue with lowest expression level (2^{-ddCt}-values).

6.8.3 Alignment of conserved PGRP domains

The distinct expression patterns of *C. floridanus PGRP* genes in the midgut of different stages suggested that especially *PGRP-LB* and *PGRP-SC2* might be linked to regulation of endosymbiont number. For further characterization, the *C. floridanus* PGRPs were aligned with other insect PGRPs as well as with T7 lysozyme (Figure 23). The latter one is a zinc-dependent amidase, which cleaves the amide bond between the lactyl group of N-acetylmuramic acid and the α-amino group of the L-alanine residue in the crosslinking peptide (Cheng et al. 1994; Mellroth and Steiner 2006).

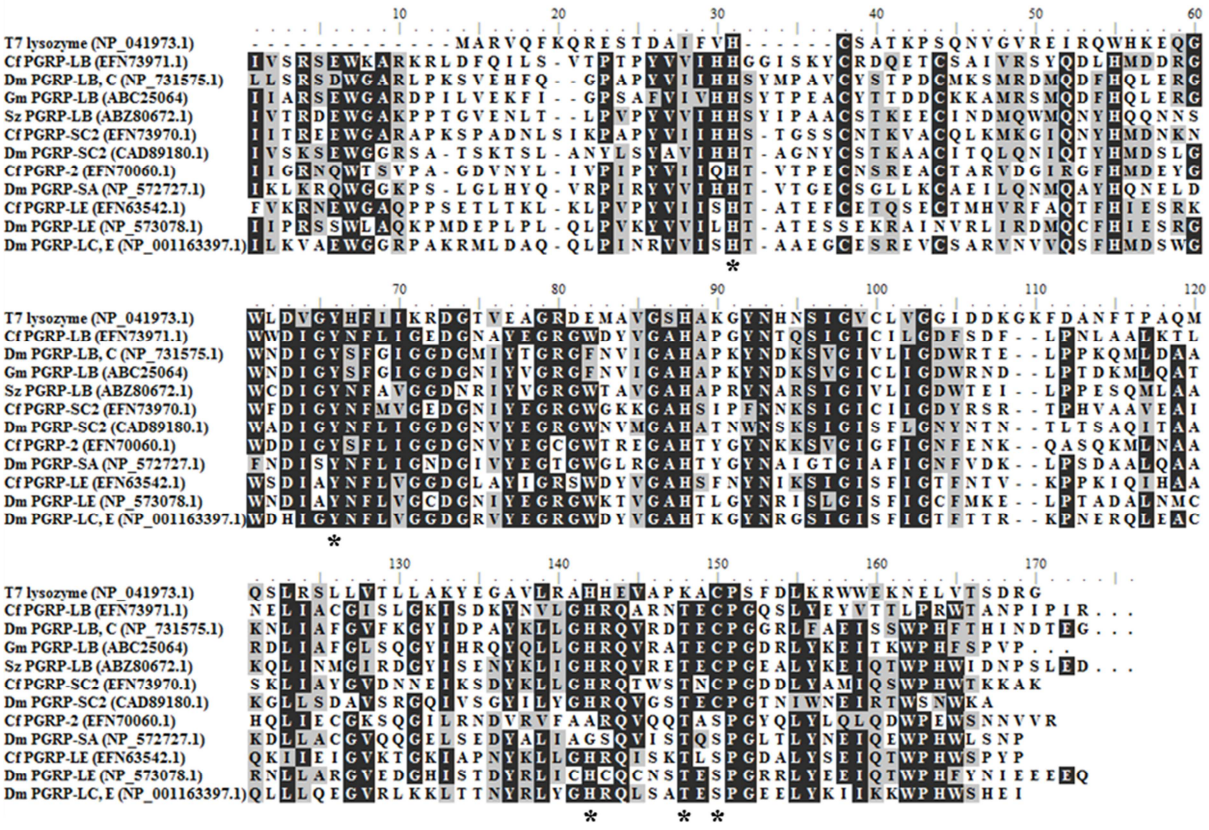

Figure 23. Alignment of lysozyme from Enterobacteria phage T7 (GenBank Acc. No.: NP_041973.1) with conserved PGRP domains from *C. floridanus* (*Cf*: PGRP-LB: EFN73971.1, PGRP-SC2: EFN73970.1, PGRP-2: EFN70060.1, PGRP-LE: EFN63542.1), *D. melanogaster* (*Dm*: PGRP-LB (isoform C): NP_731575.1, PGRP-SC2: CAD89180.1, PGRP-SA: NP_572727.1, PGRP-LE: NP_573078.1, PGRP-LC (isoform E): NP_001163397.1), *G. morsitans morsitans* (*Gm*: PGRP-LB: ABC25064) and *S. zeamais* (*Sz*: PGRP-LB: ABZ80672.1). Identical residues are underlaid with a black background and similar ones with a grey background. Residues shown to be necessary for T7 lysozyme activity are indicated by an asterisk below the corresponding position. PGRP members with proven or supposed amidase activity all possess a cysteine (C) residue at the position corresponding to C-130 in T7 lysozyme, whereas in contrast the other PGRPs have a serine (S) residue at the same position.

For the amidase activity of T7 lysozyme the residues histidine number 17 and 122 (H-17, H-122) as well as cysteine-130 (C-130), tyrosine-46 (Y-46) and lysine-128 (K-128) were shown

to be necessary, whereas replacement of K-128 by a threonine (T) residue results in retained but reduced activity (Cheng et al. 1994). The residues required for amidase activity are highly conserved in insect PGRPs with proven amidase activity like PGRP-LB and -SC2 from *D. melanogaster* (Bischoff et al. 2006; Mellroth et al. 2003; Zaidman-Remy et al. 2006) or PGRP-LB from *G. morsitans morsitans* (Wang and Aksoy 2012) as well as in PGRP-LB from *S. zeamais*, which is supposed to mediate host-symbiont-interactions (Anselme et al. 2006). Most notably, the known amidase PGRPs have a conserved cysteine at the position corresponding to C-130 in T7 lysozyme. This cysteine residue, together with the two highly conserved histidine residues (H-17 and H-122), functions as a zinc ion (Zn^{2+}) ligand, which is essential for amidase activity.

Strikingly, this cysteine residue is also highly conserved in PGRP-LB and PGRP-SC2 from *C. floridanus*, whereas PGRP-2 and PGRP-LE have a serine residue at the same position (Figure 23). This finding suggested that PGRP-LB and -SC2 are functional amidases and might have a scavenger function for peptidoglycan in *C. floridanus* as already shown for the corresponding PGRPs from *Drosophila* (Bischoff et al. 2006; Mellroth et al. 2003; Zaidman-Remy et al. 2006).

The *C. floridanus* PRR designated as PGRP-LE showed highest sequence homology to *Drosophila* PGRP-LC (GenBank Acc. No. NP_001163397.1, E-value: 1E-61), but was also significantly similar to *Drosophila* PGRP-LE (GenBank Acc. No. NP_573078.1, E-value: 1E-49). In *Drosophila* both PGRP-LC and -LE activate the IMD pathway after recognition of DAP-type peptidoglycan (Kurata 2010; Royet et al. 2011). PGRP-LC is a transmembrane protein, whereas full length PGRP-LE acts as an intracellular receptor, which induces autophagy of intracellular pathogens (Kurata 2010; Yano et al. 2008). TMHMM Server v. 2.0 (Krogh et al. 2001; Sonnhammer et al. 1998) (http://www.cbs.dtu.dk/services/TMHMM-2.0/) predicts a transmembrane helix for *C. floridanus* PGRP-LE from amino acid position 31 to 53. Accordingly, the protein seems to be a transmembrane receptor similar to PGRP-LC from *Drosophila*.

PGRP-2 from *C. floridanus* was significantly similar to *PGRP-SA* from *Drosophila* (GenBank Acc. No. NP_572727.1, E-value: 4E-52), which activates the Toll pathway upon recognition of Gram-positive bacteria (Michel et al. 2001).

6.8.4 Stage- and tissue-specific expression of genes involved in the lysosomal system, in autophagy and in production / detoxification of ROS

The relative expression levels of the genes *c-type lyso 1*, *cath L* and *lap* were significantly influenced by the factor stage (ANOVA: *c-type lyso 1*: $F = 81.27$, $p < 0.0001$; *cath L*: $F = 48.22$, $p < 0.0001$; *lap*: $F = 177.82$, $p < 0.0001$) and by the interaction tissue X stage (ANOVA: *c-type lyso 1*: $F = 51.61$, $p < 0.0001$; *cath L*: $F = 17.55$, $p < 0.0001$; *lap*: $F = 73.96$, $p < 0.0001$). In case of *c-type lyso 1* and *lap* two factorial ANOVA also detected a significant impact of the factor tissue on the relative expression levels (*c-type lyso 1*: $F = 5.57$, $p = 0.0250$; *lap*: $F = 10.68$, $p = 0.0027$) (Figure 24A, C, D). Regarding the midgut tissue, all three genes were expressed on the lowest level in P3 (Tukey´s HSD: *c-type lyso 1*: p(mL2/mP3) = 0.0001, p(mP3/mW2) = 0.0001; *cath L*: p(mL2/mP3) = 0.0063, p(mP3/mW2) = 0.0001; *lap*: p(mL2/mP3) = 0.0001, p(mP3/mW2) = 0.0001), whereat *cath L* was particularly high expressed in the midgut of W2 (Tukey´s HSD: p(mL2/mW2) = 0.0004) and *lap* in the midgut of L2 (Tukey´s HSD: p(mL2/mW2) = 0.0001). In the residual body, these genes were all highly expressed in W2 (Tukey´s HSD: *c-type lyso 1*: p(bL2/bW2) = 0.0001, p(bP3/bW2) = 0.0001; *cath L*: p(bL2/bW2) = 0.0001; *lap*: p(bL2/bW2) = 0.0001, p(bP3/bW2) = 0.0001).

For the genes *i-type lyso* and *CPVL* two factorial ANOVA detected significant differences in the relative gene expression levels between different tissues (*i-type lyso*: $F = 229.36$, $p < 0.0001$; *CPVL*: $F = 80.25$, $p < 0.0001$), as well as in the interaction tissue X stage (*i-type lyso*: $F = 130.27$, $p < 0.0001$; *CPVL*: $F = 14.48$, $p < 0.0001$), and in case of *i-type lyso* also between different developmental stages ($F = 52.76$, $p < 0.0001$) (Figure 24B, E). *I-type lyso* was significantly stronger expressed in the body of L2 and W2 than of P3 (Tukey´s HSD: p(bL2/bP3) = 0.0161, p(bP3/bW2) = 0.0035), whereas the expression of *CPVL* was only significantly elevated in the body of W2 (Tukey´s HSD: p(bL2/bW2) = 0.0037, p(bP3/bW2) = 0.0014). Interestingly, both genes were strongly expressed in the midgut tissue of P3, which was significant against mL2 and mW2 for *i-type lyso* (Tukey´s HSD: $p < 0.0001$ for both combinations) and against mW2 for *CPVL* (Tukey´s HSD: $p = 0.0483$). This high expression of *i-type lyso* and *CPVL* in midgut of P3 might be linked to the presence of high *Blochmannia* number, as in aphids homologues of both genes have been shown to be involved in control of *Buchnera* endosymbionts through the lysosomal system (Nishikori et al. 2009a; Nishikori et al. 2009b).

Expression of the *apg-5* gene was significantly influenced through the factors tissue (ANOVA: $F = 9.49$, $p = 0.0044$) and stage (ANOVA: $F = 9.41$, $p = 0.0007$), as well as through the interaction tissue X stage (ANOVA: $F = 13.13$, $p = 0.0001$) (Figure 24F). *Apg-5* was significantly expressed on the lowest level in the midgut of P3 (Tukey´s HSD: p(mL2/mP3) = 0.0392, p(mP3/mW2) = 0.0237, p(mP3/bP3) = 0.0008, p(mP3/bW2) = 0.0001), whereas the overall scale of changes in gene expression levels was quite small (median of expression in bW2 only 2.7 times higher than in mP3).

For the genes *ferritin*, *Duox* and *SOD* the influence of the factors tissue (ANOVA: *ferritin*: $F = 47.95$, $p < 0.0001$; *Duox*: $F = 915.68$, $p < 0.0001$; *SOD*: $F = 4.73$, $p = 0.0377$) and stage (ANOVA: *ferritin*: $F = 18.58$, $p < 0.0001$; *Duox*: $F = 104.27$, $p < 0.0001$; *SOD*: $F = 32.03$, $p < 0.0001$) on gene expression was significant, and in case of *ferritin* and *Duox* also the interaction tissue X stage (ANOVA: *ferritin*: $F = 37.43$, $p < 0.0001$; *Duox*: $F = 11.96$, $p = 0.0002$) (Figure 24G, H, I). *Ferritin* reached maximal expression in the midgut of L2 (Tukey´s HSD: $p < 0.002$ for all combinations). In contrast, *Duox* expression in the midgut of P3 was about ten times stronger than in the midgut of L2 and about two times stronger than in the midgut of W2 (Tukey´s HSD: p(mL2/mP3) = 0.0001, p(mP3/mW2) = 0.0015). In the residual body *Duox* was definitely strongest expressed in P3 (Tukey´s HSD: p(bL2/bP3) = 0.0001, p(bP3/bW2) = 0.0001), whereas *ferritin* expression was significantly elevated in the body of W2 in comparison to that in the body of L2 and of P3 (Tukey´s HSD: p(bL2/bW2) = 0.0001, p(bP3/bW2) = 0.0003). *SOD* gene was equally expressed in the midgut and the body of each stage, whereas on a significantly higher level in W2 (Tukey´s HSD: p(mL2/mW2) = 0.0042, p(mP3/mW2) = 0.0002, p(bL2/mW2) = 0.0002, p(bP3/mW2) = 0.0001, p(mP3/bW2) = 0.0026, p(bL2/bW2) = 0.0021, p(bP3/bW2) = 0.0002).

In sum, the high expression of *i-type lyso* and *CPVL* in the midgut of P3 suggested that these host genes might be involved in regulation of endosymbiont number, as already shown for aphids (Nishikori et al. 2009a; Nishikori et al. 2009b). Apart from that, the gene expression data of this study did not indicate an involvement of autophagy or production of ROS in host-symbiont-interactions in *C. floridanus*.

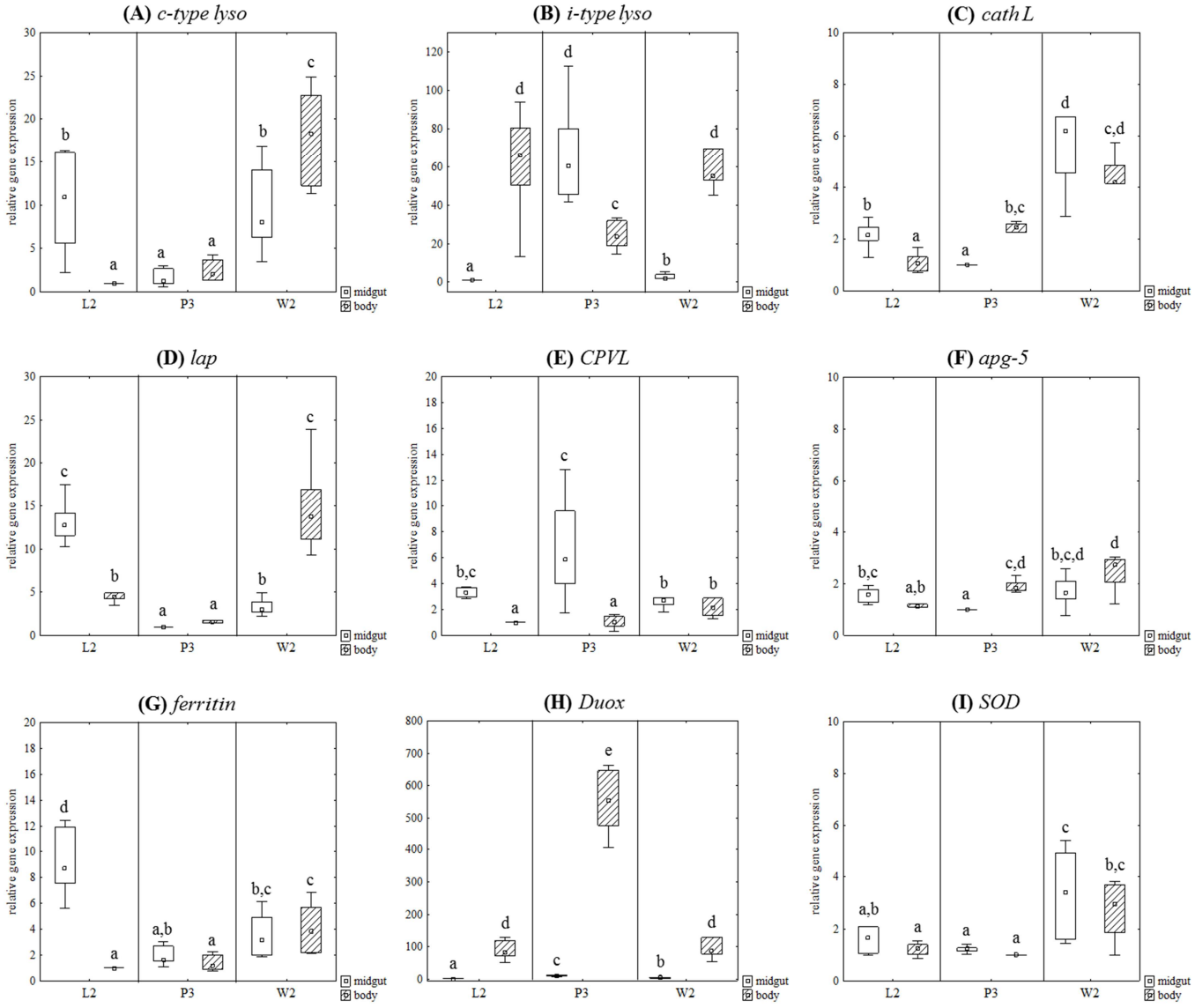

Figure 24. Gene expression in midgut (white boxes) and residual body (hatched boxes) of different *C. floridanus* stages (L2 = larvae, P3 = late pupae, W2 = adult workers). Average gene expression levels relative to the *BestKeeper* (*BK*) index (dCt = Ct(target) – Ct(BK)) of six independent samples were determined from midgut and residual body of different stages. These dCt-values of the genes (A) *c-type lysozyme 1* (*c-type lyso 1*), (B) *i-type lyso*, (C) *cathepsin L* (*cath L*), (D) *lysosomal aspartic protease* (*lap*), (E) *carboxypeptidase vitellogenic-like* (*CPVL*), (F) *autophagy protein 5* (*apg-5*), (G) *ferritin*, (H) *Dual oxidase* (*Duox*), (I) *superoxide dismutase* (*SOD*) were tested for influence of tissue and stage using factorial ANOVA. Groups differing significantly ($p < 0.05$) from each other in Tukey´s HSD post hoc test are marked with different letters (for statistics see section 8.7, Table 23 and Table 24). Box-plots show normalized changes in gene expression relative to the tissue with lowest expression level (2^{-ddCt}-values).

Figure 25 shows an overview of the gene expression approach to identify symbiosis-relevant candidate genes in the form of a heatmap.

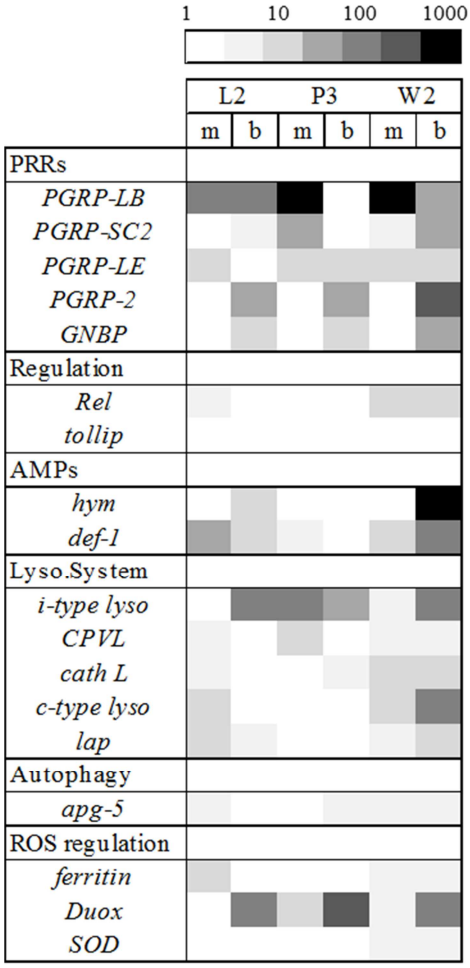

Figure 25. Overview of the candidate gene expression data, visualized in a heatmap. The median gene expression values in the midgut (m) and body (b) tissue of larvae (L2), pupae (P3) and workers (W2) range from of 1 (white) to 1000 (black).

6.8.5 Protein sequence analysis of lysozymes from *C. floridanus*

C. floridanus encodes one i-type and two c-type lysozymes in its genome (Table 15). The latter ones are quite similar to each other (only 9 aa different, see Figure 26). Both proteins possess with a signal sequence, which is followed by a mature peptide of 138 aa and a molecular weight (MW) of about 15 kDa. Both mature c-type lysozymes are basic proteins (predicted isoelectric points (pI) of 9.05 and 8.63). Comparison with c-type lysozymes from other insects revealed high conservation of 8 cysteines, which presumably form four intramolecular disulphide bonds (Figure 26). In hen egg white lysozyme (HEWL) the residue glutamic acid number 35 (E35) and partially also the residue aspartic acid number 52 (D52) were shown to be essential for muramidase function (Malcolm et al. 1989). The glutamic acid residue is conserved in both c-type lysozymes from *C. floridanus*, whereas both proteins have a serine residue at the position corresponding to D52 in HEWL. The same serine residue was also found to be present in c-type lysozymes from other ant species (e.g. *A. echinatior* and *H. saltator*) as well as in homologues from bees (e.g. *A. mellifera*, *Bombus terrestris*) and wasps (e.g. *N. vitripennis*). This is very interesting since this D52 is conserved in most of the other c-type lysozymes known to date (Bachali et al. 2002; Callewaert and Michiels 2010; Daffre et al. 1994; Li et al. 2005). Accordingly, the exchange of aspartic acid to serine seems to be characteristic for c-type lysozymes from hymenopteran species.

Table 15. Characteristics of the three lysozymes encoded in the *C. floridanus* genome. Existence of a signal sequence was predicted by SignalP 4.0 Server (http://www.cbs.dtu.dk/services/SignalP/). Accordingly, the number of amino acids (aa) in the mature protein, as well as in the predicted signal sequence are shown. The molecular weight (MW) and the isoelectric point (pI) were predicted for the putative mature peptides (by http://web.expasy.org/compute_pi/).

Lysozyme	GenBank Acc. No.	Signal / processed sequence (aa)	length aa (mature)	MW (kDa) (mature)	pI (mature)	Number of Cys (mature)
i-type	EFN71839.1	no	149	16.17	5.76	12
c-type 1	EFN74565.1	yes (20+2)	138	15.44	9.05	8
c-type 2	EFN64268.1	yes (20+2)	138	15.45	8.63	8

In contrast to the c-type lysozymes, the *C. floridanus* i-type lysozyme has no signal sequence and thus the mature protein is 149 aa long and has a MW of about 16 kDa (Table 15). The predicted pI of this lysozyme is rather acidic, which might be an adaption to its intracellular location (Callewaert and Michiels 2010). Comparison with i-type lysozymes from other insects revealed high conservation of 12 cysteine residues presumably forming six disulphide bonds (Figure 27). The residues E18 and D30 were identified as the catalytic residues in the i-type lysozyme of the bivalve *Tapes japonica* by site-directed mutagenesis as well as crystal structure analysis (Goto et al. 2007). The corresponding E and D residues were also found to be highly conserved in several invertebrate lysozymes from non-arthropod species. However, these residues are not conserved in i-type lysozyme from *C. floridanus* (the corresponding positions are marked with an asterisk in Figure 27), as well as in those from other arthropods.

Thus, it is unclear whether i-type lysozymes from arthropods have lost their catalytic activity and / or are bactericidal through a muramidase-independent mechanism (Masschalck and Michiels 2003) and / or might achieve another, so far unknown, function (Bachali et al. 2002; Callewaert and Michiels 2010; Herreweghe and Michiels 2012).

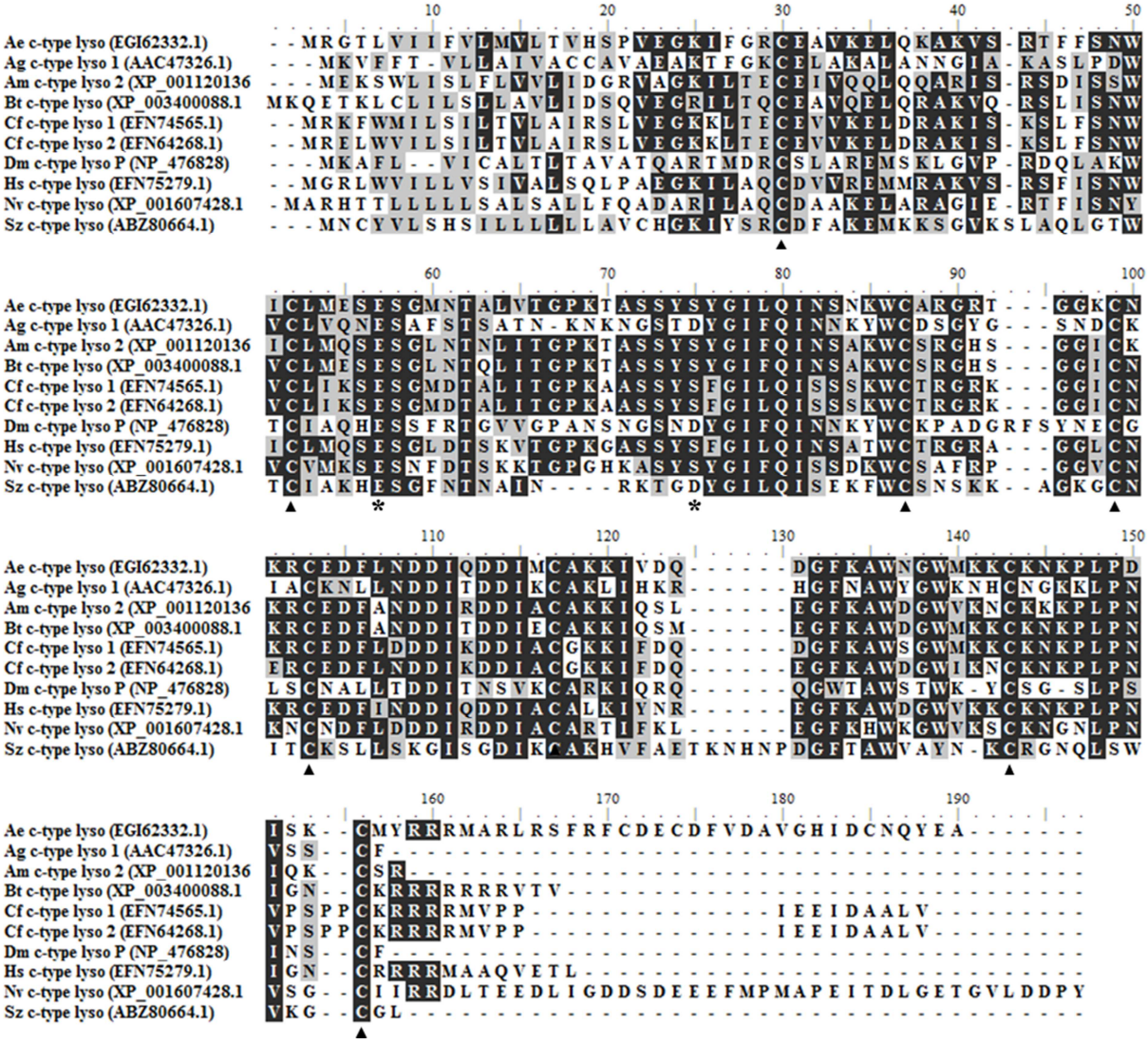

Figure 26. Alignment of c-type lysozymes from several insect species including *A. echinatior* (*Ae*), *Anopheles gambiae* (*Ag*), *A. mellifera* (*Am*), *Bombus terrestris* (*Bt*) *C. floridanus* (*Cf*), *D. melanogaster* (*Dm*), *H. saltator* (*Hs*), *N. vitripennis* (*Nv*) and *S. zeamais* (*Sz*). GenBank accession numbers are provided for each sequence in parentheses after the species name. Identical residues are underlaid with a black background and similar ones with a grey background. The eight conserved cysteine (C) residues are indicated by a triangle. Residues supposed to be important for muramidase function are indicated by an asterisk below the corresponding position. Lysozymes from bees and ants all possess a serine (S) residue at the position corresponding to aspartate 52 (D52) in chicken egg white lysozyme.

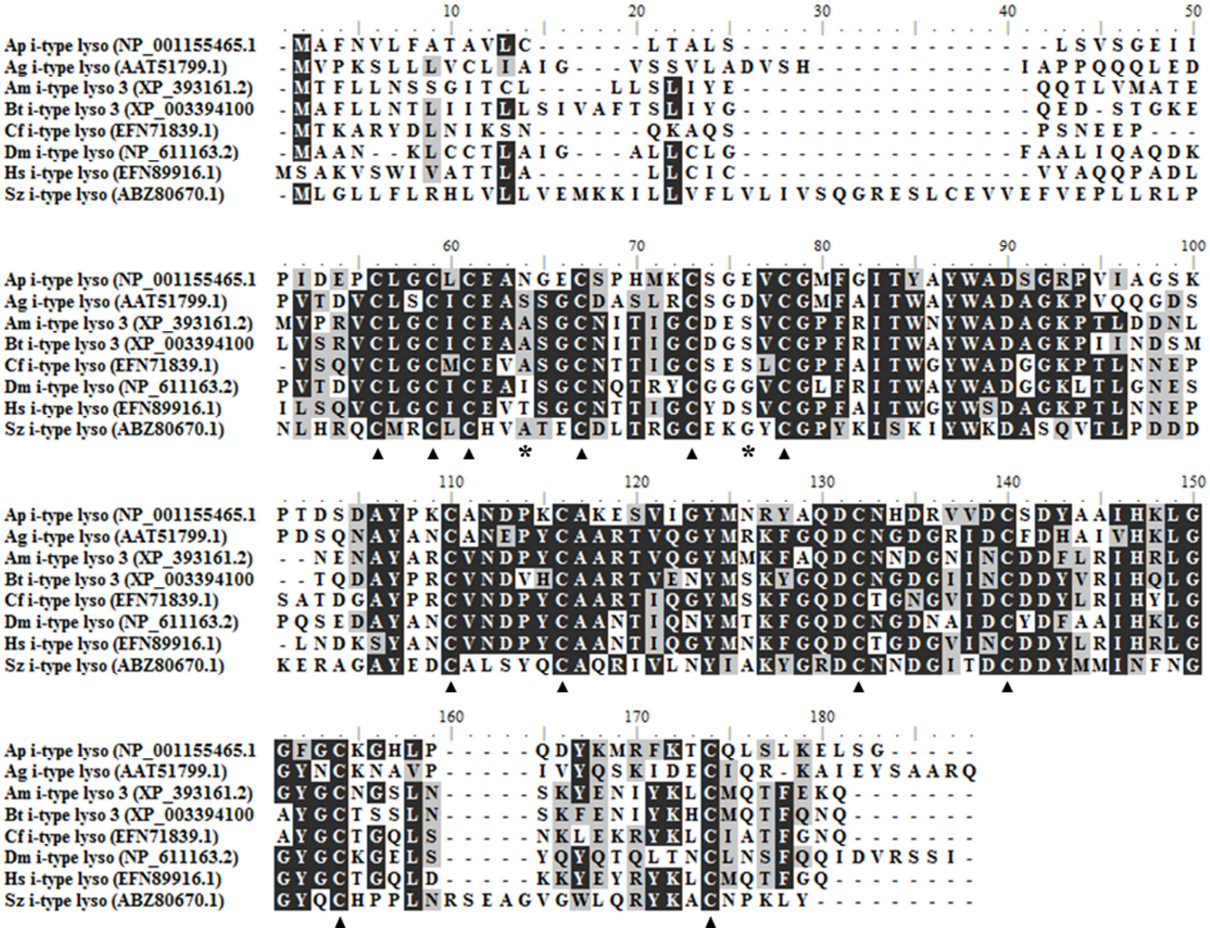

Figure 27. Alignment of i-type lysozymes from several insect species including *Acyrthosiphon pisum* (*Ap*), *Anopheles gambiae* (*Ag*), *A. mellifera* (*Am*), *Bombus terrestris* (*Bt*) *C. floridanus* (*Cf*), *D. melanogaster* (*Dm*), *H. saltator* (*Hs*) and *S. zeamais* (*Sz*). GenBank accession numbers are provided for each sequence in parentheses after the species name. Identical residues are underlaid with a black background and similar ones with a grey background. The twelve conserved cysteine (C) residues are indicated by a triangle. Position of residues known to be important for muramidase function in non-arthropod i-type lysozymes are indicated by an asterisk below the corresponding position.

6.9 Immune responses the in late pupal stage

The previous study revealed that the PRR genes with amidase domains (*PGRP-LB* and -*SC2*, see section 6.8.3) as well as the *i-type lyso* gene are highly expressed in the midgut of P3. Accordingly, these genes might be involved in the control of high endosymbiont number, which is present in the midgut at this stage. During the pupal stage the individuals are separated from the environment through the puparium, and thus they do not ingest food and usually do not have to fight an infection. Therefore, it was investigated how the immune system would react to an artificial immune-challenge. In order to get an insight into the immune response of pupae, expression of the PRR genes (*PGRP-LB*, *-SC2*, *-2*, *-LE* and *GNBP*), the NF-κB-like transcription factor *Rel* as well as of the AMP genes *hym* and *def-1* and of the *i-type lyso* gene in the midgut and the residual body of untreated controls was compared with that of individuals 6 h after picking with a 1:1 mix of heat-killed *E. coli* D31 and *M. luteus*.

For the PRR genes *PGRP-LB* and -*SC2* two factorial ANOVA detected significant differences in the relative gene expression levels (dCt-values) between different tissues (*PGRP-LB*: $F = 486.62$, $p < 0.0001$; *PGRP-SC2*: $F = 262.41$, $p < 0.0001$) (Figure 28A, B). Furthermore, expression of *PGRP-LB* was also affected by an immune-challenge ($F = 55.60$, $p < 0.0001$), as well as by the interaction tissue X treatment (immune-challenged (i) vs. untreated control (c)) ($F = 17.49$, $p = 0.0005$). In accordance with the previous results, both genes were significantly stronger expressed in the midgut of untreated P3 than in the residual body (Tukey´s HSD: *PGRP-LB*: p(mc/bc) = 0.0002; *PGRP-SC2*: p(mc/bc) = 0.0002). While expression of *PGRP-SC2* did not change at 6 h post immune-challenge (6 hpi) in comparison to controls, *PGRP-LB* gene expression was induced in response to bacteria injection, which was significant only for the body tissues (Tukey´s HSD: p(bc/bi) = 0.0002). Nonetheless, the median of *PGRP-LB* gene expression in the midgut increased from about 530 to 1040.

The relative expression levels of the other PRR genes *PGRP-LE*, *PGRP-2* and *GNBP* were significantly influenced by the factors tissue (ANOVA: *PGRP-LE*: $F = 11.73$, $p = 0.0027$; *PGRP-2*: $F = 86.96$, $p < 0.0001$; *GNBP*: $F = 27.73$, $p < 0.0001$) and immune-challenge (ANOVA: *PGRP-LE*: $F = 5.53$, $p < 0.0291$; *PGRP-2*: $F = 99.25$, $p < 0.0001$; *GNBP*: $F = 77.96$, $p < 0.0001$) (Figure 28C, D, E). For *PGRP-LE* gene the changes in gene expression in response to an immune-challenge were very small and only significantly elevated in the midgut of bacteria-injected pupae in comparison to the body of healthy individuals (Tukey´s HSD: p(mi/bc) = 0.0031). In contrast, *PGRP-2* and *GNBP* were both strongly induced at 6 hpi in the midgut as well as in the body (Tukey´s HSD: *PGRP-2*: p(mc/mi) = 0.0002, p(bc/bi) = 0.0002; *GNBP*: p(mc/mi) = 0.0002, p(bc/bi) = 0.0002). Most strikingly, both genes were significantly stronger induced in the body than in the midgut (Tukey´s HSD: *PGRP-2*: p(mi/bi) = 0.0002; *GNBP*: p(mi/bi) = 0.0002).

For *Rel* two factorial ANOVA detected only a significant impact of the factor immune-challenge ($F = 9.09$, $p = 0.0069$) on relative expression levels. Accordingly *Rel* expression was increased at 6 hpi, which was however only significant in the body tissues (Tukey´s HSD: p(bc/bi) = 0.0244) (Figure 28F).

For the AMP genes *hym* and *def-1* the influence of the factors tissue (ANOVA: *hym*: $F = 20.56$, $p = 0.0002$; *def-1*: $F = 9.86$, $p = 0.0052$) and immune-challenge (ANOVA: *hym*: $F = 392.69$, $p < 0.0001$; *def-1*: $F = 65.58$, $p < 0.0001$) on gene expression was significant, as well as the interaction tissue X treatment (ANOVA: *hym*: $F = 6.78$, $p = 0.0170$; *def-1*: $F = 15.42$, $p = 0.0008$) (Figure 28G, H). At 6 hpi the expression of *hym* and *def-1* was significantly increased in both midgut and body in comparison to expression in untreated animals (Tukey´s HSD: *hym*: p(mc/mi) = 0.0002, p(bc/bi) = 0.0002; *def-1*: p(mc/mi) = 0.0367, p(bc/bi) = 0.0002). In contrast to *def-1* gene expression, which was elevated on a similar level in midgut and body at 6 hpi, *hym* gene expression was significantly less induced in the midgut than in the residual body (Tukey´s HSD: p(mi/bi) = 0.0005).

Expression of the *i-type lyso* gene was only significantly influenced by the factor tissue (ANOVA: $F = 23.71$, $p < 0.0001$) (Figure 28I). Accordingly, *i-type lyso* was not induced after immune-challenge, but it was significantly stronger expressed in the pupal midgut than in the residual body (Tukey´s HSD: p(mc/bc) = 0.0045; p(mi/bi) = 0.0348, p(mi/bc) = 0.0186, p(mc/bi) = 0.0087), as already shown in the previous study (Figure 24B).

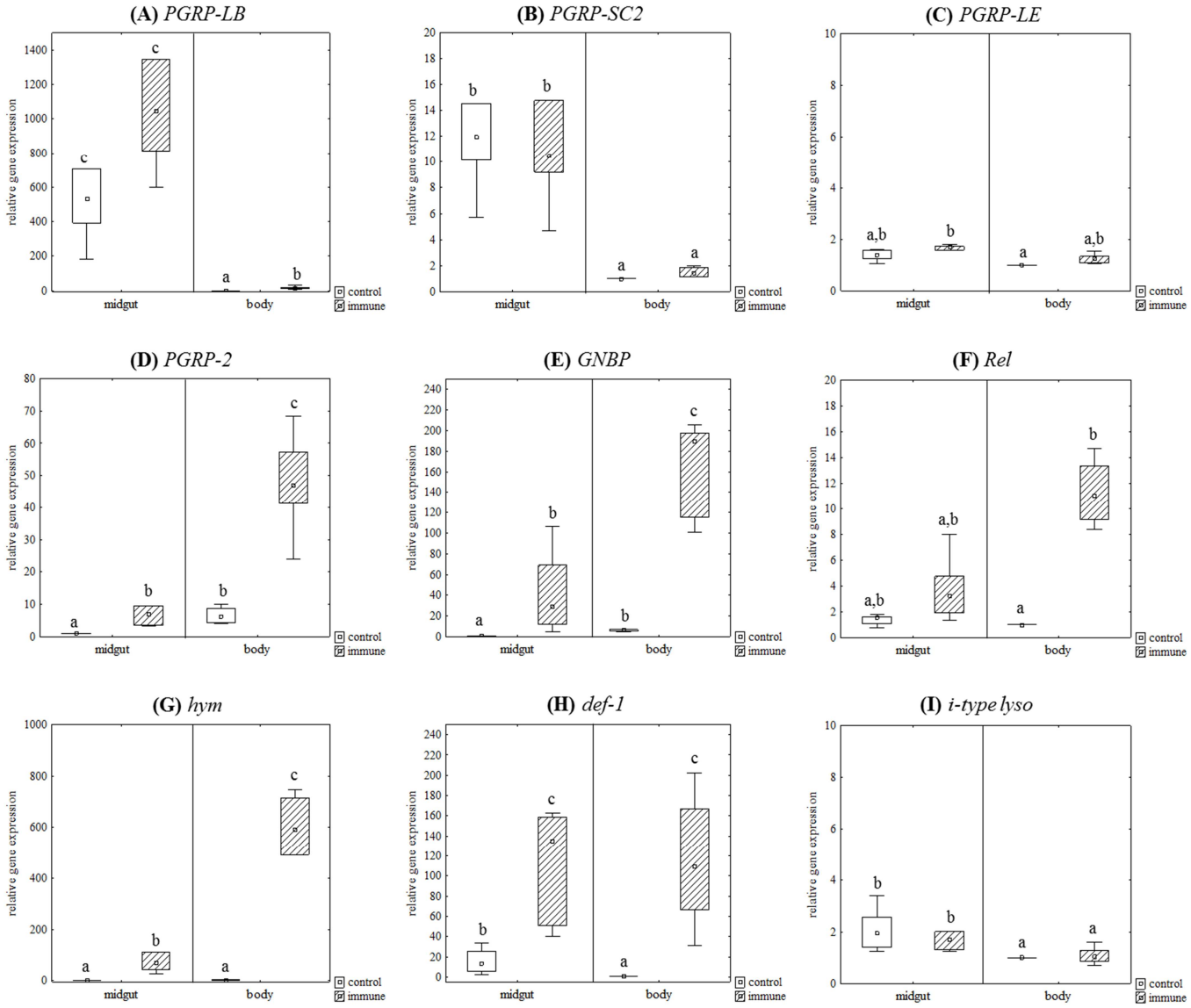

Figure 28. Gene expression in midgut and residual body of untreated (white boxes) and immune-challenged (hatched boxes) *C. floridanus* pupae (P3). For immune-challenge P3 were picked with a needle dipped into a pellet of dead *M. luteus* and *E. coli* D31 (1:1). Average gene expression levels relative to the constitutive *rpL32* gene (dCt = Ct(target) - Ct(*rpL32*)) of six independent samples were determined from midgut and residual body of untreated P3 and of individuals at 6 h after immune-challenge. These dCt-values of the genes (A) *peptidoglycan recognition protein-LB* (*PGRP-LB*), (B) *PGRP-SC2*, (C) *PGRP-LE,* (D) *PGRP-2*, (E) *Gram-negative binding protein* (*GNBP*), (F) *Relish* (*Rel*), (G) *hymenoptaecin* (*hym*), (H) *defensin-1* (*def-1*), (I) *i-type lysozyme* (*lyso*) were tested for influence of tissue (midgut vs. residual body) and treatment (immune-challenged vs. untreated) using factorial ANOVA. Groups differing significantly ($p < 0.05$) from each other in Tukey´s HSD post hoc test are marked with different letters (for statistics see section 8.8, Table 25 and Table 26). Box-plots show normalized changes in gene expression relative to the tissue with lowest expression level (2^{-ddCt}-values).

6.10 Knockdown of amidase *PGRP* gene expression in *C. floridanus*

6.10.1 Preliminary experiments for knockdown of amidase *PGRP* gene expression through feeding of the corresponding dsRNA

The supposed role of *C. floridanus PGRP-LB* and *-SC2* as negative regulators of the immune response as well as their involvement in host-symbiont-interactions should be further characterized by knockdown of the corresponding transcripts using RNAi. In order to achieve a significant knockdown effect, it is essential to get the specific dsRNA directly into the target tissue. As both *PGRP* genes were highly expressed in the midgut tissue, where the *B. floridanus* harboring bacteriocytes are located, feeding the dsRNA to the ants seemed to be the most promising delivery method to reach the midgut tissue. A cheap and easily performable method of delivery is the feeding of bacteria expressing dsRNA, which has been standardized for *C. elegans* (Timmons et al. 2001; Timmons and Fire 1998) and has recently also been adopted for some insect species (Tian et al. 2009; Zhu et al. 2011). The critical point of this method is that the bacteria and therewith the dsRNA must be taken up in a high concentration. Thus, it was initially tested, if the *E. coli* strain HT115(DE3), which is used for dsRNA expression, is sufficiently ingested by *C. floridanus*. However, the ants did not eat food containing these bacteria and, if at all, only after a long starvation period. Thus, the dsRNA had to be purified from the *E. coli* HT115(DE3) for feeding experiments.

6.10.1.1 Extraction of dsRNA from E. coli *HT115(DE3)*

A protocol for purification of the dsRNA from the *E. coli* HT115(DE3) was adapted for direct feeding of dsRNA. Four different bacterial strains were created, which either expressed dsRNA of the *green fluorescent protein* (*GFP*) gene, *C. floridanus PGRP-LB* or *PGRP-SC2* genes or of both genes (in this case the *PGRP* genes sequences were cloned in tandem into the L4440 expression vector, see section 5.19.5). The production of the corresponding dsRNAs was induced by adding IPTG to the bacterial growth medium. Five hours after induction, the dsRNA produced in the bacteria was analyzed by preparation of total RNA followed by treatment with DNase to remove residual genomic DNA, as well as with RNase A to remove single-stranded RNA. As a negative control, total RNA from bacteria, which were previously not treated with IPTG, was also extracted in the same way. As shown in Figure 29, the different dsRNAs were successfully synthesized in large quantities in the different bacterial strains after IPTG induction. In each case a sharp band of dsRNA with expected size was clearly visible indicating good quality and no degradation of the dsRNA. Thus, with this protocol dsRNA could be produced in very large amounts, which were needed for feeding experiments.

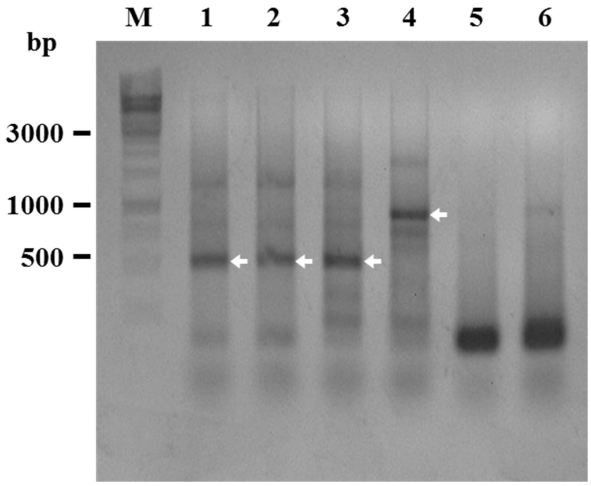

Figure 29. Quality of dsRNA produced in bacteria. In each case 1 µg of purified dsRNA was separated on a 1.2 % agarose gel alongside molecular size markers (lane M, GeneRuler 1 kb DNA Ladder, Fermentas) and analyzed with EtBr staining. White arrows indicate the major RNase/DNase resistant bands (presumed dsRNA). These major bands had the expected size of the corresponding dsRNA: Lane 1 dsGFP (473 bp), lane 2: dsPGRP-LB (483 bp), lane 3: dsPGRP-SC2 (482 bp), lane 4 (dsPGRP-LB/-SC2: 921 bp). Lanes 5 and 6 show negative control samples, which were extracted from the same bacterial clones as lane 1 (dsPGRP-LB) and lane 4 (dsPGRP-LB/-SC2), but in this case the bacteria were previously not treated with IPTG for induction of dsRNA production.

6.10.1.2 Stability of dsRNA within the crop of C. floridanus

PGRP-LB and/or *-SC2* knockdown should have the strongest impact on endosymbiont number during the pupal stage, when these genes normally reach maximal expression in the midgut (Figure 22) and the highest *Blochmannia* number is present in this tissue (Stoll et al. 2010). At the pupal stage no food is ingested, so that the dsRNA had to be fed to the late larvae shortly before pupation. However, larvae don´t take up their food by themselves, but they have to be fed by worker ants. Therefore, it was also important to determine, if sufficient amounts of the dsRNA are passed from workers to larvae.

The crop content of adult *Camponotus* workers has a quite low pH (Feldhaar, unpublished results) and likely contains several digestive enzymes (Hamilton et al. 2011). Thus, it was tested, whether ingested dsRNA was stable in the crop of *C. floridanus* workers, as this was the prerequisite for the proper transfer of non-degraded dsRNA to larvae. At 6 h and 24 h after ingestion of dsRNA, 1 µl regurgitate of three different workers per time point was examined by electrophoresis and checked for integrity of the ingested dsRNA. As a control dsRNA was simultaneously incubated at room temperature and 1 µg was also analyzed after 6 h and 24 h. As shown in Figure 30, the dsRNA remained stable in the crop of *C. floridanus*. Even 24 h after ingestion sharp bands of dsRNA could be detected on agarose gels (Figure 30B). Therefore it was assumed that the non-degraded dsRNA might also be passed to larvae by feeding via the workers.

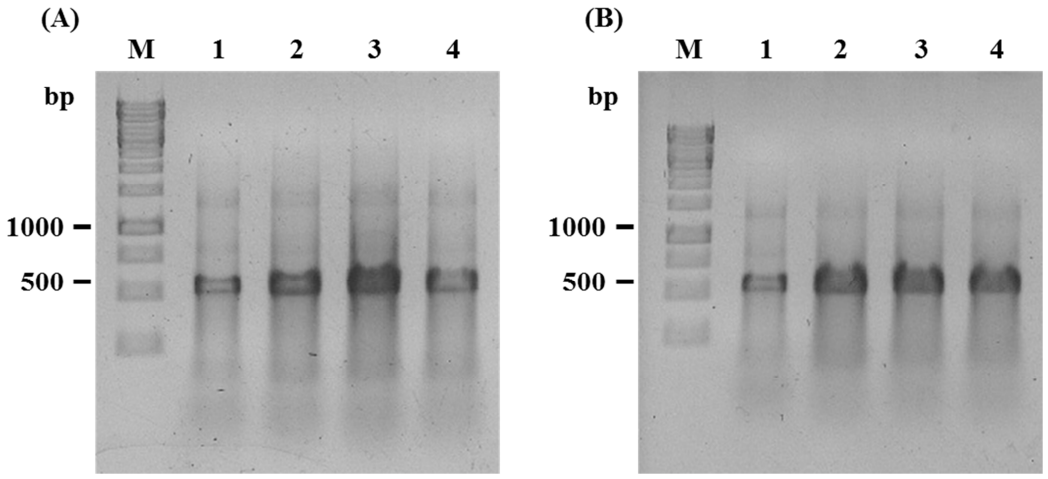

Figure 30. Stability of dsRNA within the crop of *C. floridanus*. Adult workers were fed with dsRNA of *GFP* gene. At 6 h (A) and 24 h (B) after feeding the regurgitate was collected from 3 individuals per time point ((A) and (B): lanes 2 - 4). 1 µl regurgitate was separated on a 1.2 % agarose gel alongside molecular size markers (lane M, GeneRuler 1 kb DNA Ladder, Fermentas) and analyzed with EtBr staining. As a control, 1 µg of dsGFP was applied, which was previously incubated at room temperature for the same time ((A) and (B): lane 1).

6.10.1.3 Knockdown of amidase PGRP *gene expression by feeding of dsRNA*

The previous experiments (see sections 6.8.2 and 6.9) suggested that in *C. floridanus* the amidase PGRPs, PGRP-LB and / or PGRP-SC, might be involved in the toleration of the high endosymbiont number present at the pupal stage in the midgut tissue. In order to test this hypothesis, the aim of the following experiments was to knock down the expression of one or both of the *PGRP* genes in the midgut tissue during the late larval stage using RNAi. Subsequently, the endosymbiont number present in *PGRP*-knockdown and control ants should be compared in the late pupal stage and the transcription of the target genes should be determined by qRT-PCR. It was assumed that knockdown of the amidase PGRPs would lead to an induction of immune genes like *hym* in the midgut possibly resulting in a decrease of the endosymbiont number per host.

To this end, the successful application of RNAi technology had to be initially established for *C. floridanus*. Preliminary feeding experiments were performed with subcolonies of minor worker ants, because they can be fed directly. With these experiments it should be ensured that feeding dsRNA can produce a strong and long-lasting knockdown effect in *C. floridanus*. Furthermore, larvae were added to the subcolonies in order to check, if knockdown of gene expression could also be achieved in this stage by feeding dsRNA mediated by worker ants. After 5 d, 10 d and 15 d feeding period (see section 5.19.5), transcription of the genes *PGRP-LB*, *-SC2*, *-2*, *hym* and *def-1* was determined in the midgut and the residual body of adult workers using qRT-PCR. Moreover, target gene expression was also analyzed in whole larvae at 10 d and 15 d, in order to investigate, if an RNAi effect could also be achieved in larvae. For this preliminary study, only two experimental series were performed and thus no statistical analysis of the expression data was conducted.

Analysis of target gene expression revealed that a RNAi effect could only be achieved for *PGRP-LB*, and that the knockdown was especially pronounced at 5 d in the residual body parts, but not in the midgut tissue (Figure 31A, D). At this time point the mean normalized *PGRP-LB* expression level in the body of ants fed with dsPGRP-LB (L) was 0.04 (± 0.01), in contrast to 1.00 in ants fed with pure saccharose solution (H) and 1.12 (± 0.37) in those fed with dsGFP (G). At the same time point *PGRP-LB* expression was also noticeably reduced to a mean value of 0.35 (± 0.18) in ants, which received dsRNA from *PGRP-LB* and -*SC2* gene (LS). The knockdown was also evident at 10 d, as normalized *PGRP-LB* expression was reduced to 0.26 (± 0.09) in L samples (Figure 31E). On day 15, after the 5 d dsRNA-feeding break, *PGRP-LB* gene expression in L samples was back on control levels (Figure 31F). At the measured time points, no RNAi effect could be observed for *PGRP-LB* gene expression in larvae (Figure 31G, H).

Knockdown of *PGRP-SC2* gene expression could not be achieved by feeding of the corresponding specific dsRNA (S samples) at all measured time points, neither in workers nor in larvae (Figure 32).

To check for specificity of the RNAi effect, the expression of another *PGRP* gene, *PGRP-2*, was also determined. As expected, *PGRP-2* gene expression was not reduced by feeding of dsRNA specific for the *PGRP-LB* and/or -*SC2* gene (Figure 33). However, the data indicated, that feeding of dsRNA might non-specifically induce expression of this gene in the midgut, as it was slightly stronger expressed in the midgut of all ant groups, which were fed with dsRNA (samples G, L, S, LS) in comparison to pure saccharose solution (H), especially at day 10 (Figure 33B).

It was assumed that knockdown of the amidase PGRPs would lead to an induction of immune effectors like the AMP genes *hym* or *def-1* as a response to the endosymbiont within the midgut tissue and this would probably result in a decrease of the endosymbiont number per host. However, although in L probes from body tissue, in which *PGRP-LB* transcripts were strongly reduced (Figure 31D, E), gene expression of AMP genes was not evidently affected (Figure 34 and Figure 35). It was remarkable, that at 5 d and 10 d expression of *hym* and *def-1* was quite low in body samples from L, S and LS in comparison to H (Figure 34D, E and Figure 35D, E). Yet it is doubtful, whether this effect was specific for feeding of dsRNA from amidase *PGRP* genes, as expression of AMP genes varied strongly in G control samples. Moreover, in several cases the variation in AMP gene expression levels of the biological replicates was quite strong. As the expression levels of *hym* and *def-1* were already shown to be highly variable in untreated animals (see section 6.8.2, Figure 22), a possible effect of dsRNA feeding on these genes could hardly be determined.

To determine the presumed effect of amidase *PGRP*-knockdown on endosymbiont survival, expression of *B. floridanus GroEL* gene relative to *C. floridanus rpL32* gene was determined in all samples from midgut tissue and from larvae. However, since amidase *PGRP* gene

expression could not be knocked down in midgut tissue, also no effect on endosymbiont number could be measured (Figure 36).

In sum, a specific RNAi effect could only be determined for *PGRP-LB* gene in the body (except the midgut) of adult workers. Moreover, the data suggested that it was not possible to achieve a strong and long-lasting RNAi effect in larvae through feeding of dsRNA (mediated by worker ants). Thus, in this way knockdown of amidase *PGRP* transcripts could also not be accomplished in the late pupal stage.

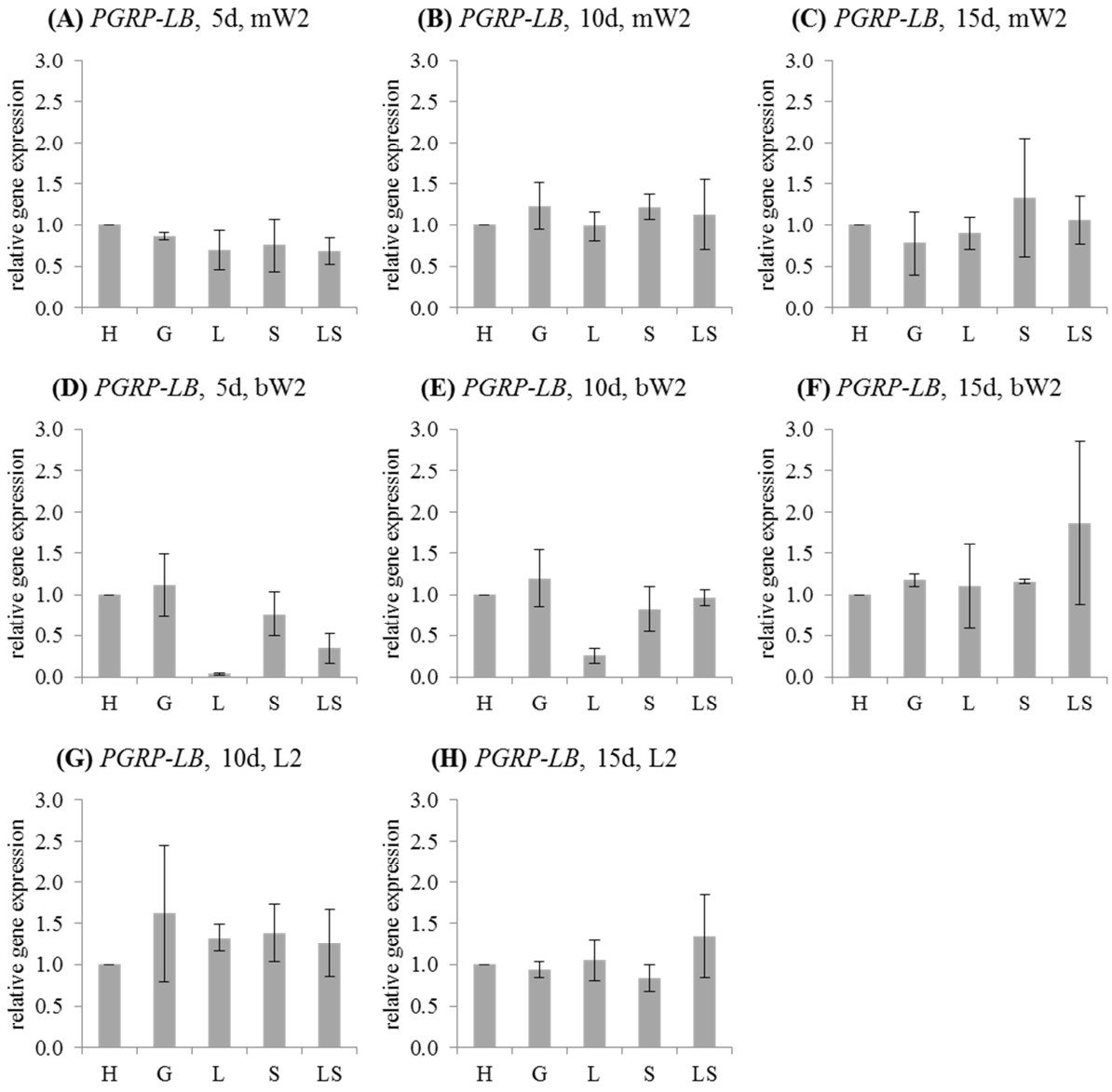

Figure 31. Relative expression of *PGRP-LB* gene in midgut (m, (A) - (C)) and residual body (b, (D) - (F)) of W2 as well as in larvae ((G) and (H)) at 5 d, 10 d and 15 d of the dsRNA feeding experiment. Subcolonies were fed for 10 d with saccharose dissolved in either pure H_2O (H) or containing 2 µg/µl dsRNA of *GFP* (G), *PGRP-LB* (L), *PGRP-SC2* (S) or of *PGRP-LB* and -*SC2* (LS). Afterwards, all groups were fed with saccharose solution for five additional days. Average *PGRP-LB* expression levels relative to the constitutive *rpL32* gene (dCt = Ct(*PGRP-LB*) - Ct(*rpL32*)) of two independent samples were determined and then normalized to the corresponding control samples (H) using the 2^{-ddCt} method.

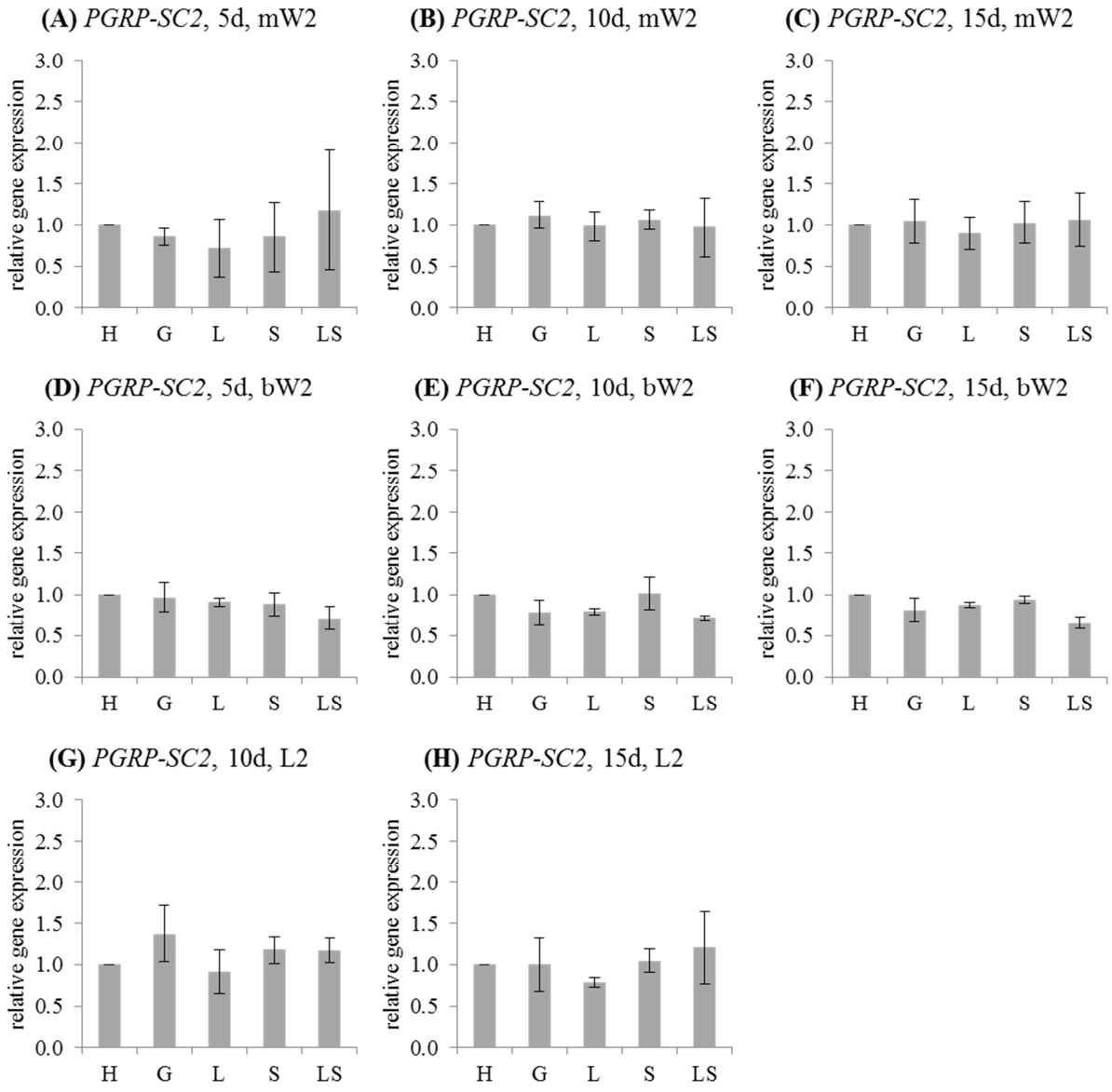

Figure 32. Relative expression of *PGRP-SC2* gene in midgut (m, (A) - (C)) and residual body (b, (D) - (F)) of W2 as well as in larvae ((G) and (H)) at 5 d, 10 d and 15 d of the dsRNA feeding experiment. Subcolonies were fed for 10 d with saccharose dissolved in either pure H_2O (H) or containing 2 µg/µl dsRNA of *GFP* (G), *PGRP-LB* (L), *PGRP-SC2* (S) or of *PGRP-LB* and -*SC2* (LS). Afterwards, all groups were fed with saccharose solution for five additional days. Average *PGRP-SC2* expression levels relative to the constitutive *rpL32* gene (dCt = Ct(*PGRP-SC2*) - Ct(*rpL32*)) of two independent samples were determined and then normalized to the corresponding control samples (H) using the 2^{-ddCt} method.

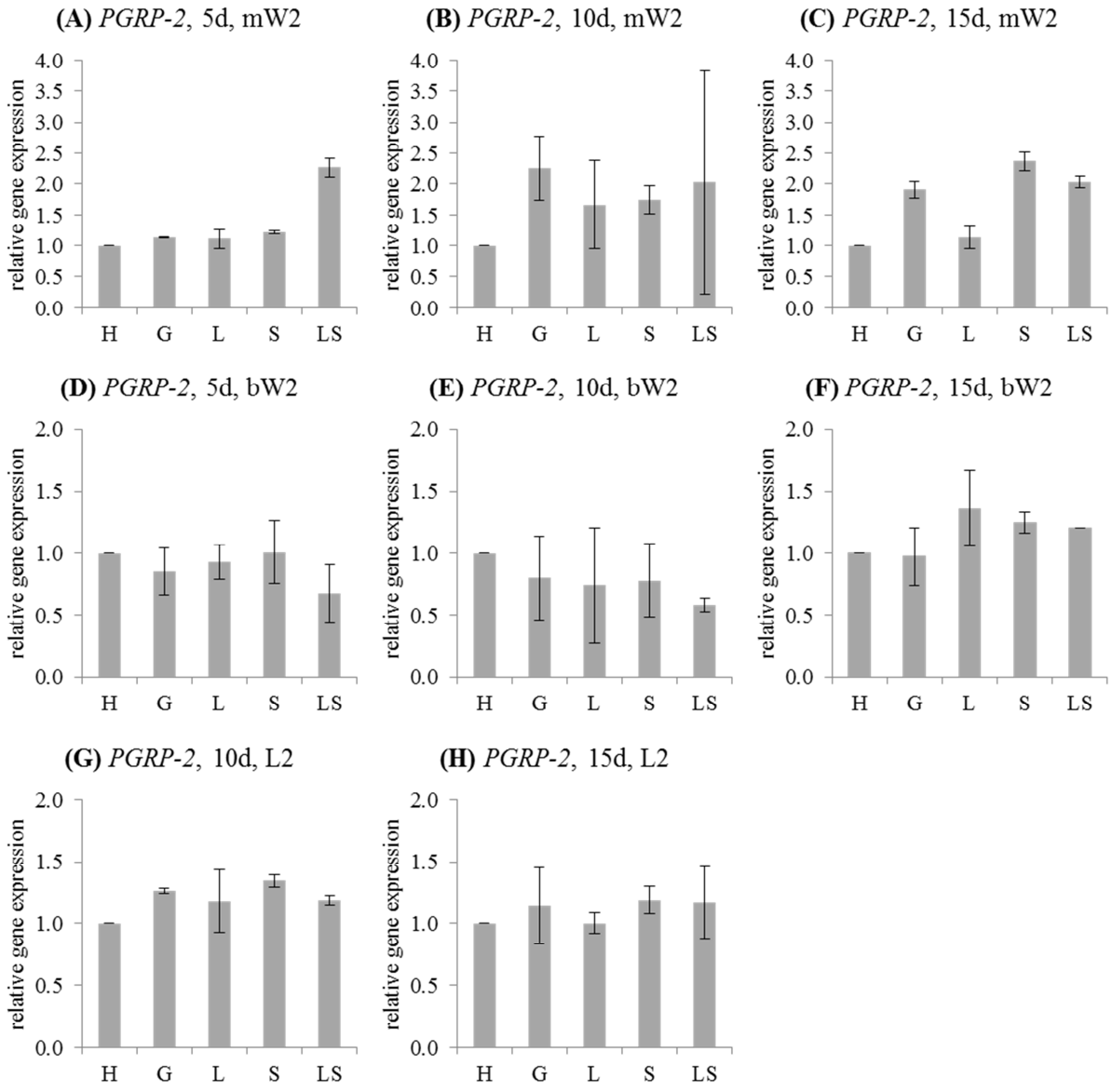

Figure 33. Relative expression of *PGRP-2* gene in midgut (m, (A) - (C)) and residual body (b, (D) - (F)) of W2 as well as in larvae ((G) and (H)) at 5 d, 10 d and 15 d of the dsRNA feeding experiment. Subcolonies were fed for 10 d with saccharose dissolved in either pure H_2O (H) or containing 2 µg/µl dsRNA of *GFP* (G), *PGRP-LB* (L), *PGRP-SC2* (S) or of *PGRP-LB* and -*SC2* (LS). Afterwards, all groups were fed with saccharose solution for five additional days. Average *PGRP-2* expression levels relative to the constitutive *rpL32* gene (dCt = Ct(*PGRP-2*) - Ct(*rpL32*)) of two independent samples were determined and then normalized to the corresponding control samples (H) using the 2^{-ddCt} method.

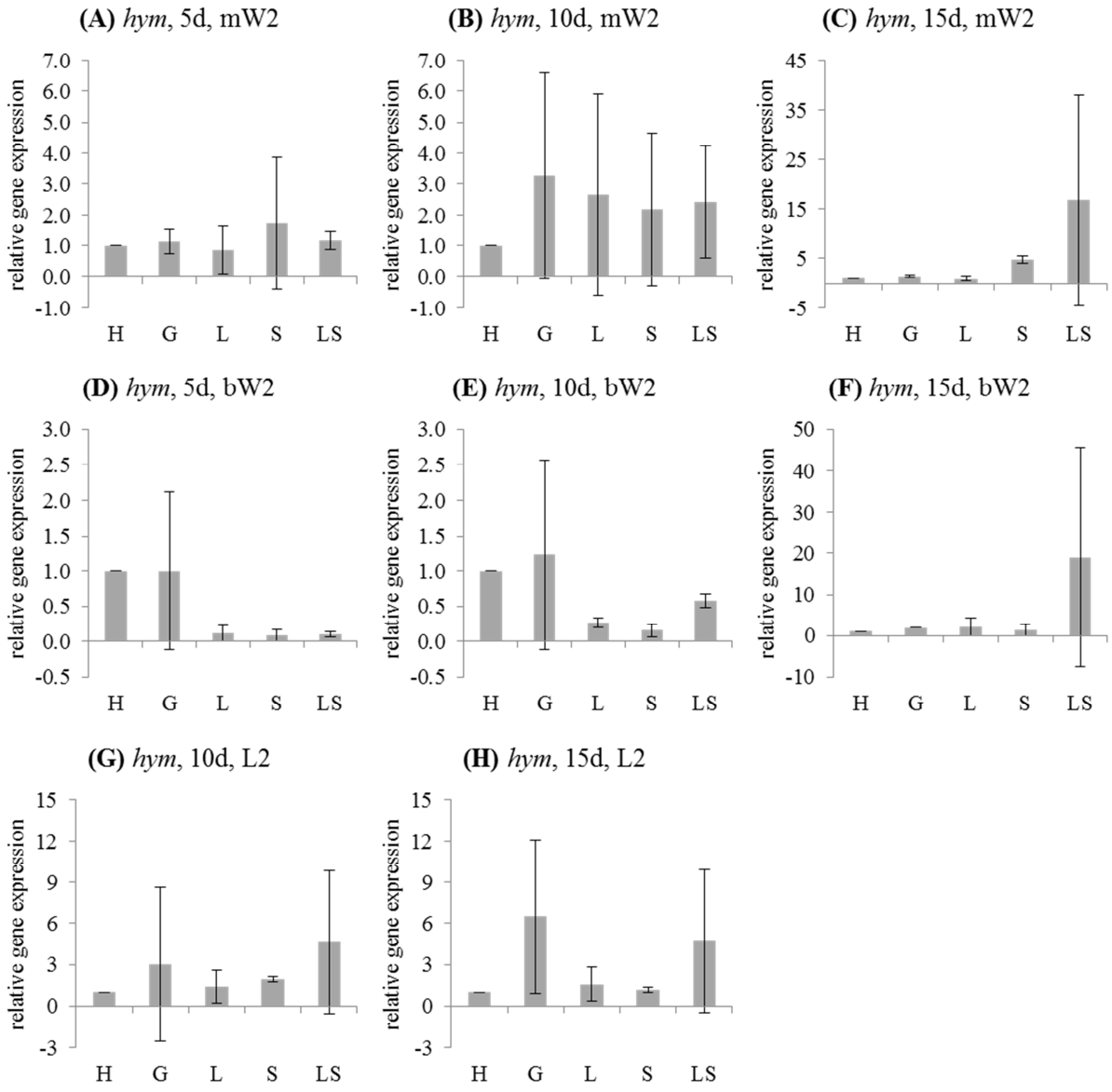

Figure 34. Relative expression of *hym* gene in midgut (m, (A) - (C)) and residual body (b, (D) - (F)) of W2 as well as in larvae ((G) and (H)) at 5 d, 10 d and 15 d of the dsRNA feeding experiment. Subcolonies were fed for 10 d with saccharose dissolved in either pure H_2O (H) or containing 2 µg/µl dsRNA of *GFP* (G), *PGRP-LB* (L), *PGRP-SC2* (S) or of *PGRP-LB* and -*SC2* (LS). Afterwards, all groups were fed with saccharose solution for five additional days. Average *hym* expression levels relative to the constitutive *rpL32* gene (dCt = Ct(*hym*) - Ct(*rpL32*)) of two independent samples were determined and then normalized to the corresponding control samples (H) using the 2^{-ddCt} method.

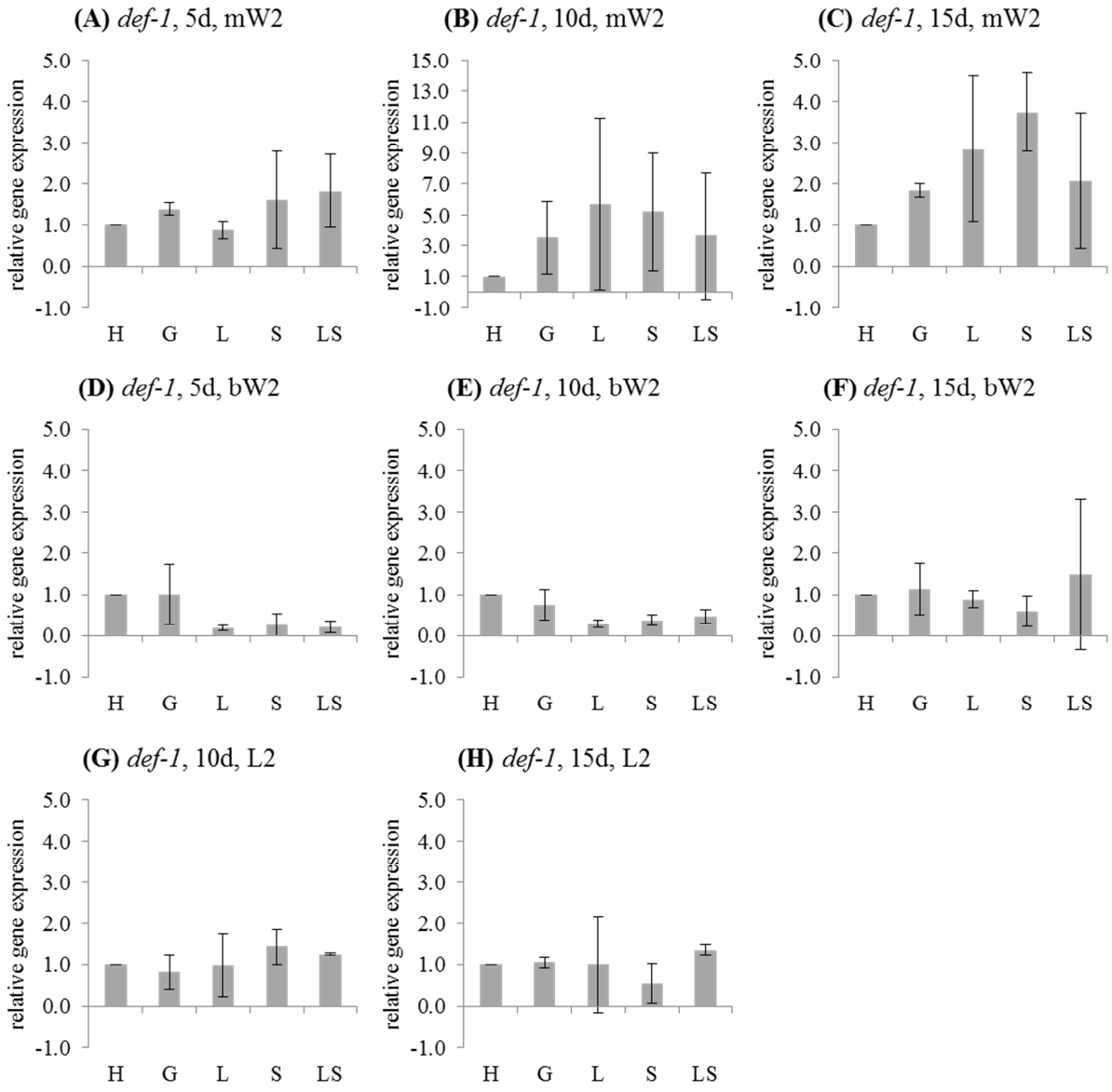

Figure 35. Relative expression of *def-1* gene in midgut (m, (A) - (C)) and residual body (b, (D) - (F)) of W2 as well as in larvae ((G) and (H)) at 5 d, 10 d and 15 d of the dsRNA feeding experiment. Subcolonies were fed for 10 d with saccharose dissolved in either pure H_2O (H) or containing 2 µg/µl dsRNA of *GFP* (G), *PGRP-LB* (L), *PGRP-SC2* (S) or of *PGRP-LB* and -*SC2* (LS). Afterwards, all groups were fed with saccharose solution for five additional days. Average *def-1* expression levels relative to the constitutive *rpL32* gene (dCt = Ct(*def-1*) - Ct(*rpL32*)) of two independent samples were determined and then normalized to the corresponding control samples (H) using the 2^{-ddCt} method.

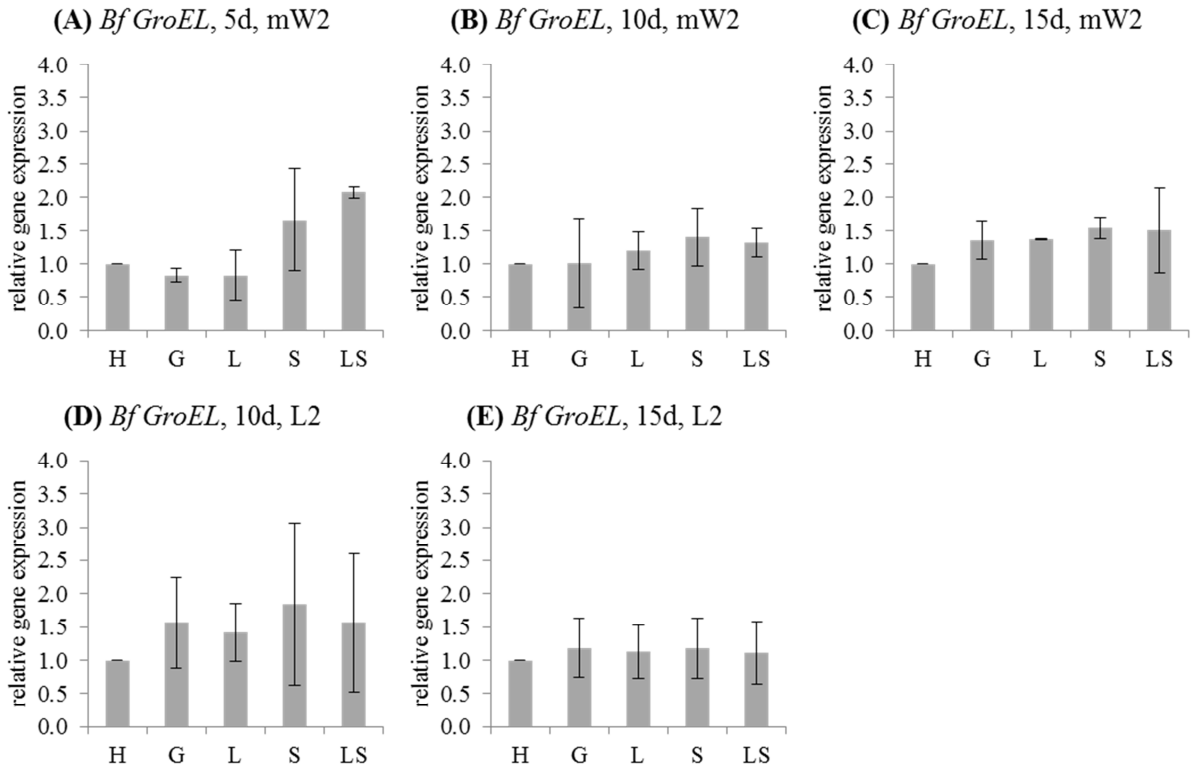

Figure 36. Relative expression of *B. floridanus* (*Bf*) *GroEL* gene in midgut (m, (A) - (C)) as well as in larvae ((D) and (E)) at 5 d, 10 d and 15 d of the dsRNA feeding experiment. Subcolonies were fed for 10 d with saccharose dissolved in either pure H_2O (H) or containing 2 µg/µl dsRNA of *GFP* (G), *PGRP-LB* (L), *PGRP-SC2* (S) or of *PGRP-LB* and -*SC2* (LS). Afterwards, all groups were fed with saccharose solution for five additional days. Average expression levels of the constitutive *Bf GroEL* gene relative to the constitutive host gene *rpL32* (dCt = Ct(*Bf GroEL*) - Ct(*rpL32*)) of two independent samples were determined and then normalized to the corresponding control samples (H) using the 2^{-ddCt} method.

6.10.2 Knockdown of *PGRP-LB* during an immune response

Preliminary experiments suggested that knockdown of amidase *PGRP* gene expression in the midgut of the late pupal stage was hardly attainable and thus the involvement of these genes in regulation of the endosymbiont number could not be investigated. Alternatively, it was decided to prove the hypothesis that PGRP-LB is in fact a negative regulator of the ant´s immune response. In the previous experiment, knockdown of *PGRP-LB* in the body of adult worker ants was accomplished, but did not affect immune gene expression. However, the downregulation of the immune response by PGRP-LB was possibly not evident from this experiment, as the ants did not suffer from an infection. Therefore, a new experiment was performed, in which *C. floridanus* subcolonies were fed with dsRNA from *GFP* or *PGRP-LB* gene for 6 d, as described above, but at day 4 workers were inoculated with bacteria in order to elicit an immune response. Immune gene expression in the midgut and the residual body parts was investigated at day 6 by qRT-PCR, as it was previously shown that 48 h after bacterial challenge immune gene expression has already declined again and is only slightly higher than in control animals (see section 6.6). Furthermore, the preliminary RNAi experiments suggested a strong knockdown of *PGRP-LB* expression around day 5 (Figure 31D). Accordingly, it was assumed that *PGRP-LB* knockdown during an ongoing immune response would result in prolonged induction of immune gene expression, since immune elicitors (peptidoglycan molecules) should not be removed anymore as efficiently as in the control animals.

Expression of *PGRP-LB* gene was significantly affected by the treatment in the body of adult workers (ANOVA: $F = 25.97$, $p < 0.0001$), but not in the midgut tissue (Figure 37A, B). *PGRP-LB* transcripts were significantly reduced in the body of immune-challenged animals, which were fed with saccharose solution containing dsPGRP-LB (L), in comparison to untreated animals, which did not take up dsRNA (H) (Tukey´s HSD: $p(H/L) = 0.0010$). In contrast, *PGRP-LB* expression was significantly induced in response to the immune-challenge in the two other control groups, which were fed with saccharose solution either pure (I) (Tukey´s HSD: $p(H/I) = 0.0346$) or containing dsGFP (G) (Tukey´s HSD: $p(H/L) = 0.0314$). Thus, *PGRP-LB* could efficiently be knocked down in the body of adult workers by feeding of dsPGRP-LB. Yet in larvae, no RNAi effect for *PGRP-LB* could be achieved, as the gene was significantly induced in all immune-challenged individuals (ANOVA: $F = 12.05$, $p = 0.0001$; Tukey´s HSD: $p(H/I) = 0.0008$, $p(H/G) = 0.0004$), including L groups (Tukey´s HSD: $p(H/L) = 0.0007$) (Figure 37C).

In the midgut of W2, the expression of the other amidase *PGRP* gene (*PGRP-SC2*) was neither affected by feeding of dsRNA nor by the immune-challenge (Figure 37D), whereas in the body of workers (bW2) *PGRP-SC2* expression was significantly induced in all immune-challenged groups (Figure 37E, ANOVA: $F = 5.72$, $p = 0.0054$; Tukey´s HSD: $p(H/I) = 0.0434$, $p(H/G) = 0.0041$, $p(H/L) = 0.0491$). Thus, knockdown of *PGRP-LB* expression in bW2 did not affect *PGRP-SC2* expression. As in bW2, *PGRP-SC2* expression in larvae was

significantly induced in response to the immune-challenge (ANOVA: $F = 8.33$, $p = 0.0009$, Tukey´s HSD: p(H/I) = 0.0038, p(H/G) = 0.0102, p(H/L) = 0.0013), but was not affected by feeding of dsRNA (Figure 37F).

Also expression of the *PGRP-2* gene was not affected by feeding of dsPGRP-LB, as its expression level in L groups was not significantly different from that in I and G control groups (Figure 37G-I). In W2 *PGRP-2* expression was slightly induced in immune-challenged groups (I, G, L) in comparison to untreated animals (H) (ANOVA: mW2: $F = 5.24$, $p = 0.0078$; bW2: $F = 4.36$, $p = 0.0162$), which was significant for G and L in mW2 (Tukey´s HSD: p(H/G) = 0.0403, p(H/L) = 0.0059) as well as for L in bW2 (Tukey´s HSD: p(H/L) = 0.0113). In larvae, *PGRP-2* expression was significantly induced in all immune-challenged groups (Figure 37I, ANOVA: $F = 8.23$, $p = 0.0009$, Tukey´s HSD: p(H/I) = 0.0071, p(H/G) = 0.0021, p(H/L) = 0.0029), whereas also no impact of dsRNA feeding could be detected.

The AMP gene *hym* was significantly induced at 48 h after the immune-challenge in worker bodies (Figure 38B, ANOVA: $F = 13.50$, $p < 0.0001$, Tukey´s HSD: p(H/I) < 0.0001, p(H/G) = 0.0001, p(H/L) < 0.0001) as well as in L2 (Figure 38C, ANOVA: $F = 16.80$, $p < 0.0001$, Tukey´s HSD: p(H/I) = 0.0002, p(H/G) = 0.0002, p(H/L) = 0.0003). However other than expected, the *hym* gene was not significantly stronger expressed in L groups than in I and G control groups even in bW2, where *PGRP-LB* expression was efficiently reduced. In mW2 *hym* was indeed significantly induced in L groups in comparison to H (Figure 38A, ANOVA: $F = 3.67$, $p < 0.0296$, Tukey´s HSD: p(H/L) = 0.0239), but not in comparison to I and G. Thus in sum, no effect of *PGRP-LB* knockdown on *hym* expression could be detected.

Expression of *def-1* gene in bW2 was significantly induced in response to the immune-challenge (ANOVA: $F = 6.83$, $p = 0.0024$, Tukey´s HSD: p(H/I) = 0.0055, p(H/G) = 0.0103, p(H/L) = 0.0071), but not significantly stronger in L groups than in I and G control groups (Figure 38E). Thus, the data indicated no effect of *PGRP-LB* knockdown on *def-1* expression.

For completeness, the other *defensin* gene of *C. floridanus*, *def-2*, was also included in this study. However, the expression of *def-2* was not significantly affected by the treatments, neither by feeding of dsRNA nor by immune-challenge (Figure 38G-I). Accordingly, also no effect of *PGRP-LB* knockdown on *def-2* gene expression could be detected.

In sum, *PGPR-LB* gene expression was efficiently knocked down during an immune response through feeding of the corresponding dsRNA. However, no impact of this knockdown on the expression of other immune genes could be detected.

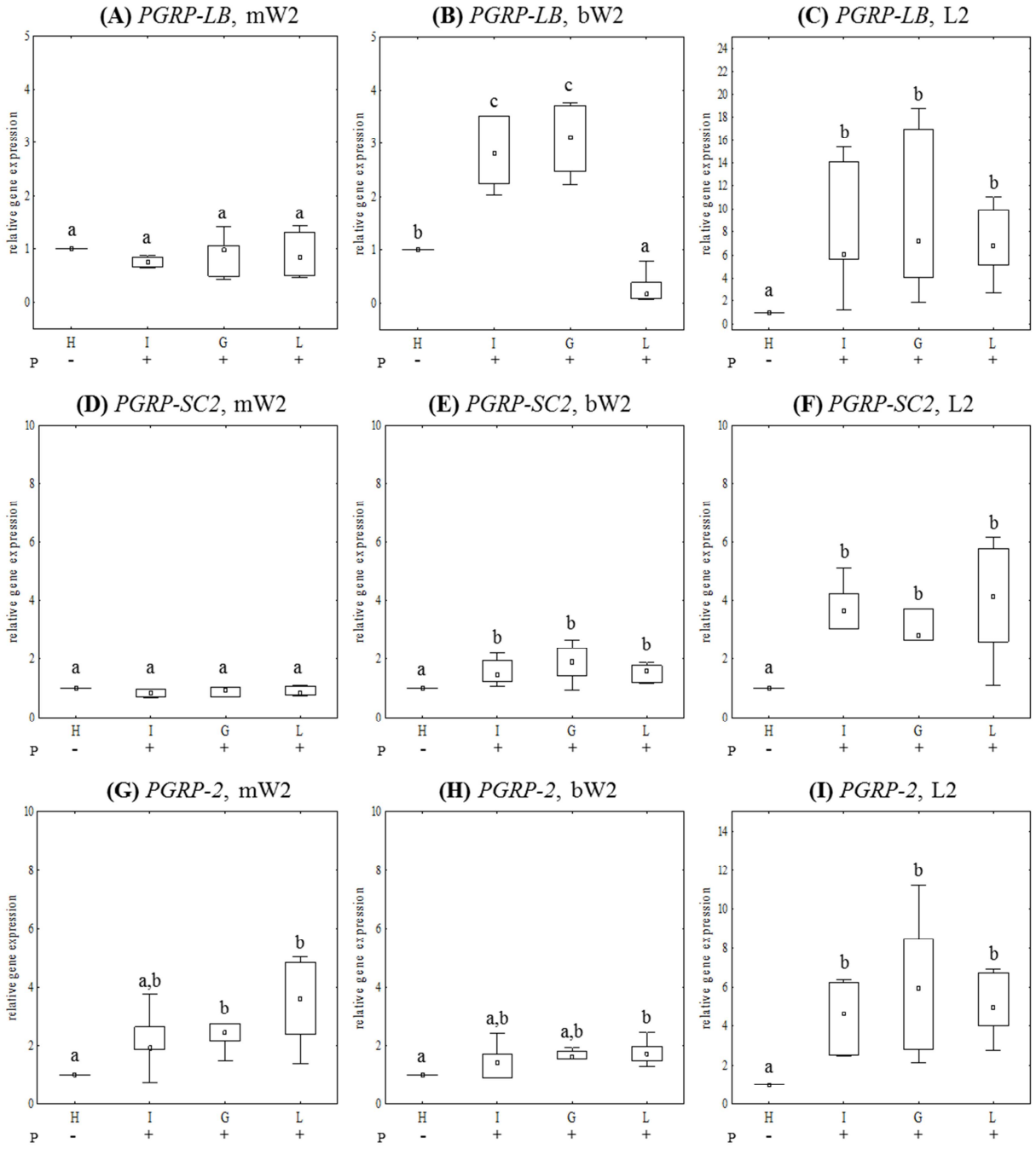

Figure 37. Relative expression of *PGRP* genes in midgut (m) and residual body (b) of minor workers (W2) as well as in whole larvae (L2) at day 6 of the dsRNA feeding experiment. Subcolonies were fed for 6 d with saccharose dissolved in either pure H_2O (H and I) or containing 2 µg/µl dsRNA of *GFP* (G) or *PGRP-LB* (L). At day 4 individuals of the groups I, G and L were picked (P +) with a needle dipped into a pellet of dead *M. luteus* and *E. coli* D31 (1:1). Average gene expression levels relative to the constitutive *rpL32* gene (dCt = Ct(target) - Ct(*rpL32*)) of six independent samples were determined at 48 h after the immune-challenge (on day 6 of the feeding experiment) from mW2 and bW2 as well as from L2 of all different groups. These dCt-values of the genes *PGRP-LB* (A-C), *PGRP-SC2* (D-F) and *PGRP-2* (G-I) were tested for influence of the treatment using factorial ANOVA. Groups differing significantly ($p < 0.05$) from each other in Tukey´s HSD post hoc test are marked with different letters (for statistics see section 8.9, Table 27 - Table 32). Box-plots show normalized changes in gene expression relative to the corresponding control samples (H) (2^{-ddCt}-values).

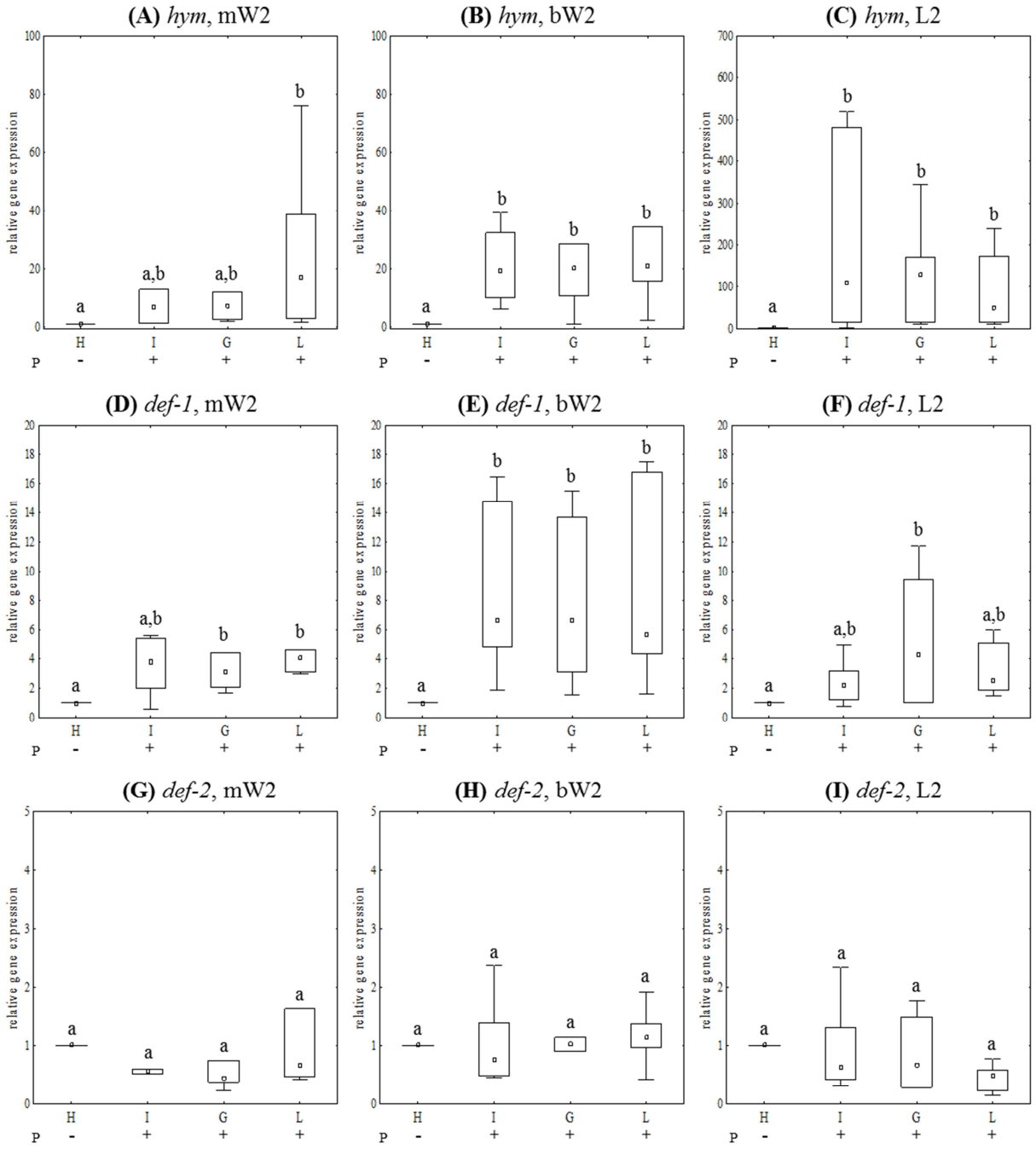

Figure 38. Relative expression of AMP genes in midgut (m) and residual body (b) of minor workers (W2) as well as in whole larvae (L2) at day 6 of the dsRNA feeding experiment. Subcolonies were fed for 6 d with saccharose dissolved in either pure H_2O (H and I) or containing 2 µg/µl dsRNA of *GFP* (G) or *PGRP-LB* (L). At day 4 individuals of the groups I, G and L were picked (P +) with a needle dipped into a pellet of dead *M. luteus* and *E. coli* D31 (1:1). Average gene expression levels relative to the constitutive *rpL32* gene (dCt = Ct(target) - Ct(*rpL32*)) of six independent samples were determined at 48 h after the immune-challenge (on day 6 of the feeding experiment) from mW2 and bW2 as well as from L2 of all different groups. These dCt-values of the genes *hym* (A-C), *def-1* (D-F) and *def-2* (G-I) were tested for influence of the treatment using factorial ANOVA. Groups differing significantly ($p < 0.05$) from each other in Tukey´s HSD post hoc test are marked with different letters (for statistics see section 8.9, Table 27 - Table 32). Box-plots show normalized changes in gene expression relative to the corresponding control samples (H) (2^{-ddCt}-values).

6.11 Analysis of *C. floridanus* haemolymph

After an immune-challenge insects produce several immune effector proteins/peptides, which are mainly secreted into the haemolymph (Boman et al. 1974; Boman et al. 1972; Hultmark et al. 1980; Lemaitre and Hoffmann 2007). The antimicrobial properties of the haemolymph from *C. floridanus* were investigated by inhibition-zone assays. Moreover, haemolymph proteins/peptides from immune-challenged animals were characterized by SDS-PAGE in order to identify immune factors on the protein level.

6.11.1 Antimicrobial activity of *C. floridanus* haemolymph

Antimicrobial activity in the haemolymph of *C. floridanus* major workers was tested from untreated individuals (C1-C4) as well as 24 h post injection of heat-killed bacteria (1:1 mix of *E. coli* D31 and *M. luteus*) (I1-I4) via inhibition-zone assays using different indicator organisms (Figure 39). No zone of inhibition could be detected for Gram-negative *E. coli*, whereas growth of Gram-positive *M. luteus* was already inhibited by haemolymph from untreated animals and even stronger by haemolymph of bacteria-injected individuals. Gram-positive *B. subtilis* was not inhibited by haemolymph of untreated animals and only slightly by two haemolymph samples (I3 and I4) from bacteria-injected animals. Brownish coloration of application sites, especially from untreated haemolymph samples, indicates melanization of the haemolymph.

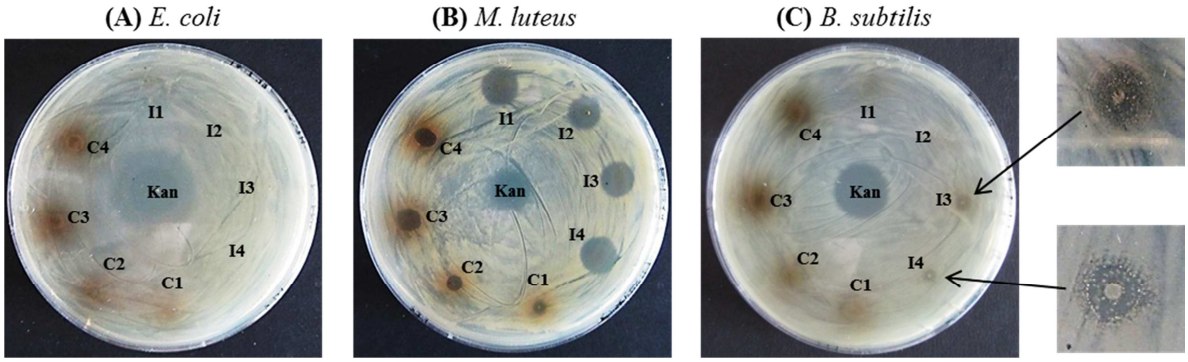

Figure 39. Inhibition-zone assay for the detection of antimicrobial activities in the haemolymph of major workers 24 h post bacteria injection. An aliquot of fresh overnight cultures of (A) *E. coli* D31, (B) *M. luteus* or (C) *B. subtilis* was spread on an agar plate. 3 µl aliquots of pooled haemolymph samples from untreated controls (C1-C4) and bacteria-injected individuals (I1-I4) were directly applied onto the agar plate with pipette tips and the plate was subsequently incubated overnight at 30°C. As a positive control, 5 µg kanamycin (Kan) was placed in the center of each agar plate.

6.11.2 Analysis of *C. floridanus* haemolymph proteins

Aliquots from haemolymph samples used for inhibition-zone assays (C1-C3 and I1-I3) were also analyzed via SDS-PAGE of proteins/peptides. Large proteins (30-200 kDa) and small

proteins/peptides were separated on 10 % (Figure 40) and 15 % (Figure 41) polyacrylamide/0.1 % SDS gels, respectively. Comparison of samples from untreated and bacteria-injected individuals revealed no different protein/peptide-bands.

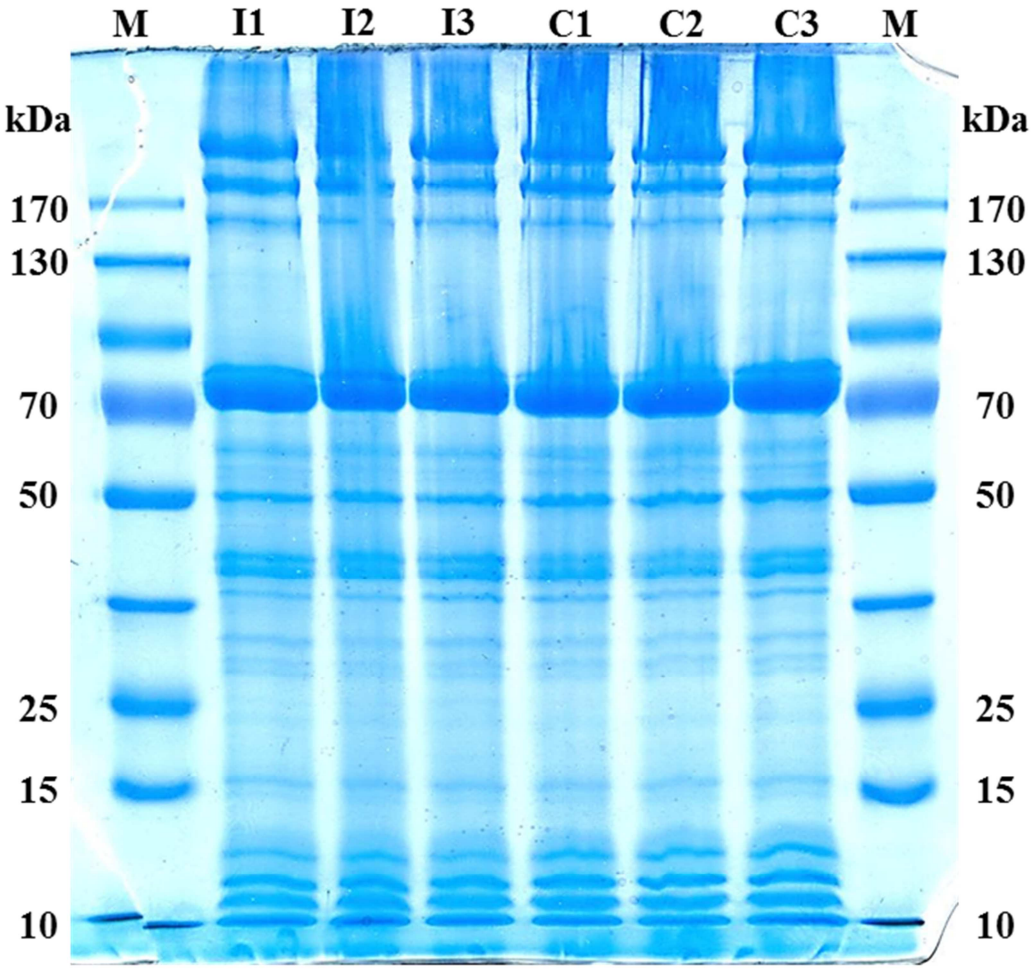

Figure 40. Gel electrophoretic separation of large haemolymph proteins. Haemolymph samples from major workers were collected 24 h post bacteria injection. Aliquots of 5 µl were mixed with Laemmli buffer and three biological replicates of pooled haemolymph samples from untreated control animals (C1-C3) and bacteria-injected individuals (I1-I3) were applied on a 10 % polyacrylamide/0.1 % SDS gel and electrophoresed in 1x Tris-glycine buffer alongside a size marker (PageRuler™ Prestained Protein Ladder from Fermentas). Gels were stained via colloidal Coomassie staining (Roti®-Blue, Roth). Differentially expressed immune proteins/peptides could not be detected.

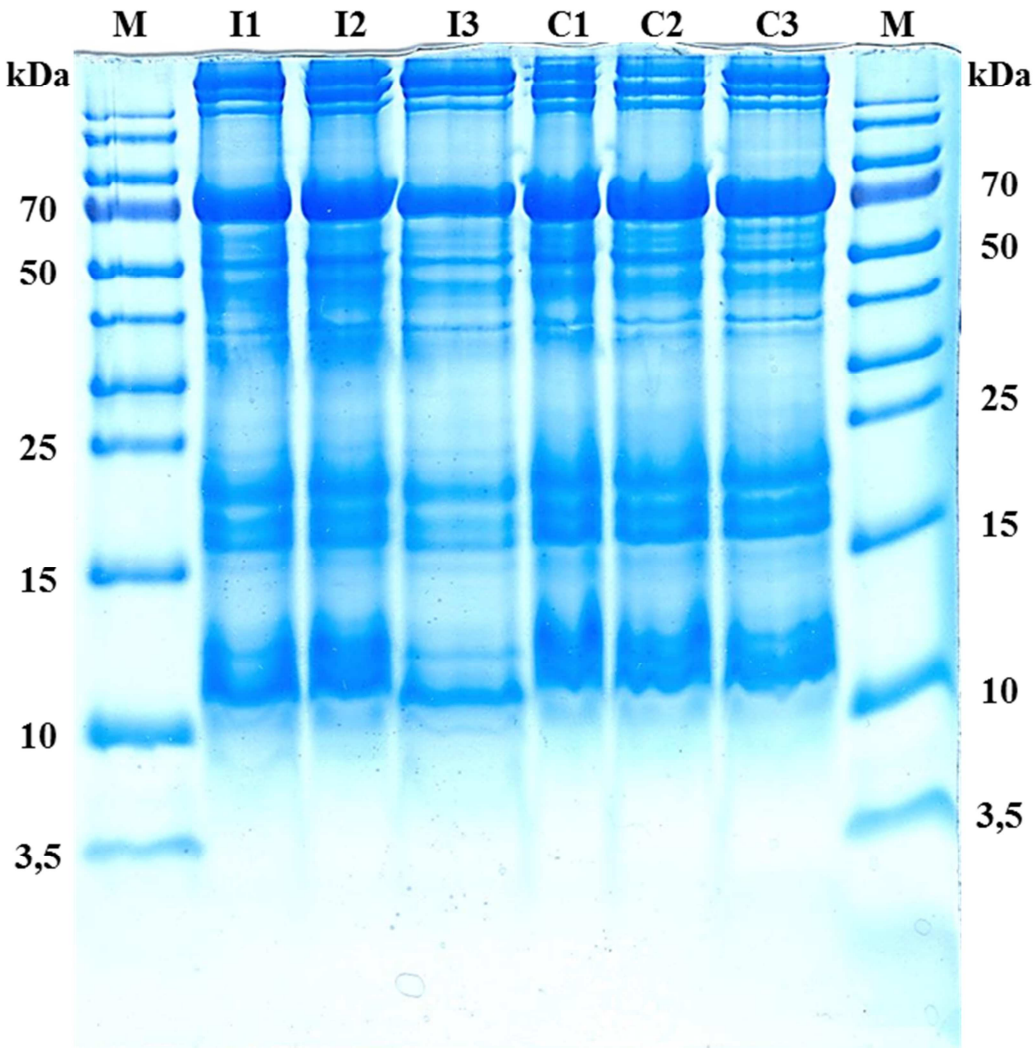

Figure 41. Gel electrophoretic separation of small haemolymph proteins/peptides. Haemolymph samples from major workers were collected 24 h post bacteria injection. Aliquots of 5 µl were mixed with Laemmli buffer and three biological replicates of pooled haemolymph samples from untreated control animals (C1-C3) and bacteria-injected individuals (I1-I3) were applied on a 15 % polyacrylamide/0.1 % SDS gel. Electrophoresis was performed alongside a size marker (PageRuler™ Prestained Protein Ladder from Fermentas) and according to (Schägger 2006; Schägger and von Jagow 1987). Gels were stained via colloidal Coomassie staining (Roti®-Blue, Roth). Differentially expressed immune proteins/peptides could not be detected.

7 Discussion

7.1 Identification of immune-relevant genes from *C. floridanus* by SSH

Using SSH, 35 infection-inducible genes from *C. floridanus* were identified, which encode proteins that share significant sequence similarities with proteins known to be involved in diverse aspects of the immune response in other insects (Table 11). Among these were genes encoding proteins involved in pathogen recognition (e.g. GNBP), signal transduction (e.g. MAPK-kinase), antimicrobial activity (e.g. defensin and hymenoptaecin), or general stress response (e.g. heat shock protein).

A big advantage of the SSH method is, that it can be used to identify differentially ESTs between two conditions from animals, for which no reference genome is available. Even if a reference genome already exists, SSH is still quite useful to identify prominent candidate genes for continuative expression studies as well as so far not annotated genes, as seen for *hym* gene in the context of this study. In addition to the subtraction of common cDNAs between two samples, SSH includes a normalization step to equalize the abundance of cDNAs within the target population and to increase the chance for detection of rare transcripts (Diatchenko et al. 1996). In the light of the overrepresentation of ESTs belonging to *hym* and *TyrOH* genes within the subtracted libraries, the normalization in this study seemed to be incomplete. Similar incomplete normalization has also been described in a study for identification of immune genes from *Manduca sexta* (Zhu et al. 2003). In the present study this problem was successfully dealt with by artificial addition of ESTs of *hym* and *TyrOH* to the cDNAs derived from untreated animals.

It should also be noted, that this initial SSH study did not provide a saturating genome-wide analysis of immune genes due to the limitations of this method. In this approach only prominent candidate genes were identified, which are upregulated after an immune-challenge, since cDNAs derived from untreated animals were subtracted from cDNAs derived from immune-challenged individuals. Genes, which are downregulated after an immune-challenge, could be obtained by reciprocal subtraction of the two cDNA populations.

As insect hosts need to maintain a chronic infection of their mutualistic endosymbionts, immune gene regulation of the bacteriocyte tissue may differ from other tissues though. In other insects it has been shown that bacteriocytes seem to constitutively express some molecules with antimicrobial activity, e.g. an i-type lysozyme in aphids (Nakabachi et al. 2005). Moreover a *tollip* (Toll-interacting protein) gene and a *PGRB-LB* gene were shown to be more strongly expressed in the bacteriocytes of weevils than in other body parts (Anselme et al. 2008). These genes might regulate the immune responses of bacteriocytes in such a way that the endosymbionts are not cleared but confined within the bacteriocytes. Homologues of these genes are also found in the *C. floridanus* genome. Since such factors would be constitutively expressed rather than being induced upon an immune-challenge, these genes

might have not been detected in this study, as SSH method mainly provides ESTs from differentially expressed genes. In order to identify such possible symbiosis-relevant genes a candidate gene profiling study was conducted (see section 6.8).

7.2 Characterization of the *C. floridanus* transcriptome in response to bacterial challenge using Illumina sequencing technology

Illumina sequencing technology was used in order to identify a broader spectrum of immune-regulated genes than by SSH. For this purpose, RNA samples from whole larvae and workers (1:1 mix) were used for sequencing, whereas RNA samples for SSH were extracted only from midgut and fatbody tissue of workers.

Using TopHat and Cufflinks (Trapnell et al. 2012) the Illumina sequencing reads were assembled to produce a transcriptome annotation of the *C. floridanus* genome. Based on this analysis 15,730 genes were predicted, which is 1334 genes less than in the current Official Gene Set (OGS) v3.3 (Bonasio et al. 2010). Additional analyses are needed to examine the function of these missing genes. Some of these genes may be exclusively expressed in pupae or in males, which were not included in the Illumina samples. In several cases, Cufflinks fused two (or more) originally separately annotated EAGs to one new XLOC (name of the new gene identifiers) resulting in a lower number of predicted genes in this study. Exemplary examination of such EAGs revealed that they were often just annotated as pseudogenes (e.g. EAG_12213). Thus, several of the originally annotated genes might also be annotation artefacts. In accordance with that, the RNA sequencing data published together with the *C. floridanus* genome also only supported about 80 % of the predicted genes (Bonasio et al. 2010).

In addition to the gene prediction, Cufflinks also compiled the transcription start site(s) as well as the splicing structure of each gene from the Illumina reads resulting in the identification of 31,624 isoforms, which need to be further characterized in prospective studies.

Using Cuffdiff and *DESeq* a set of 292 genes (XLOCs) was determined, which were significantly differentially expressed between the sample from immune-challenged and untreated animals (see section 6.2.3 and 8.1, Table 16). Results of the Illumina sequencing were validated by qRT-PCR for 15 genes (Table 13). Moreover, most of the genes identified by SSH (e.g. *GNBP*, *def-1*, *TyrOH*) were confirmed to be significantly upregulated by the Illumina results. The list of upregulated genes determined by SSH could be complemented with several additional factors from diverse aspects of the insect immune response. The *hym* gene is absent in the list of differentially expressed genes, since Illumina sequencing reads could not be properly mapped due to the missing repeat part in the genome (see sections 6.3 and 7.4). However, separate analysis of Illumina sequencing reads showed upregulation of

this gene in the sample from immune-challenged animals (data not shown). For some genes the mapping of the reads to the genome seems to be ambiguous, since different expression fold changes were determined by Cufflinks and *DESeq*. Prospective studies need to determine expression fold changes of such genes using qRT-PCR.

Among the upregulated genes were also some known negative regulators of the immune response e.g. *PGRP-LB*, *cactus* and *serine protease inhibitor* genes. The sample from immune-challenged animals was taken 12 h after the picking with bacteria. In *Drosophila*, it was shown that several immune genes are already strongly induced at 30 min after an immune-challenge, whereas 12 h later immune gene expression already starts to decline (Boutros et al. 2002; Irving et al. 2001; Lemaitre and Hoffmann 2007). Since prolonged induction of immune effector genes might also harm the insect host, upregulation of negative immune regulators likely prevents harmful over-activation of the immune system (Bischoff et al. 2006; Nicolas et al. 1998; Reichhart 2005; Zaidman-Remy et al. 2006). Moreover, a gene encoding a prophenoloxidase subunit (POsub) was found to be significantly downregulated in Illumina results as well as in qRT-PCR analysis (Table 13). This downregulation of POsub might counteract excessive melanization reactions producing harmful ROS (Sadd and Siva-Jothy 2006).

Several of the annotated genes encoding proteins from immune signaling pathways like Imd, MyD88, Toll or JAK (Bonasio et al. 2010) were only slightly, but not significantly upregulated in immune-challenged animals. These proteins belong to the middle parts of different immune signaling pathways and thus seem to be rather constitutively expressed than induced upon immune-challenge. Transcriptional upregulation of the corresponding genes is not necessary for an immune response, since the activation of several immune pathways is controlled on the posttranslational level and/or via transcriptional upregulation of genes encoding recognition proteins (e.g. GNBP) (Lemaitre and Hoffmann 2007; Ryu et al. 2010).

The downregulation of genes involved in digestion and storage upon immune-challenge fits to the studies on other insects and indicates that during an infection insects seem to temporarily shut down synthesis of non-essential proteins in order to use resources for costly defense reactions (Aguilar et al. 2005; Lourenco et al. 2009; Meng et al. 2008).

The number of differentially regulated genes upon an immune-challenge is presumably higher than the 292 genes identified in this study. However, so far only one sample per condition was sequenced und thus the usual variation between biological replicates was only estimated by Cufflinks / *DESeq* leading to an underestimation of significantly regulated genes. Sequencing of more replicates per condition would be necessary to identify a higher number of significantly regulated genes.

In addition to simple changes in gene expression, Cuffdiff also reports genes that are differentially spliced or differentially regulated via promoter switching (Trapnell et al. 2012). The 292 differentially expressed genes identified from the Illumina sequencing correspond to

722 isoforms. The analysis of these isoforms could not be completed in the context of this thesis. Thus, prospective studies are needed to examine these isoforms and to investigate whether some genes are possibly differentially regulated in response to bacterial challenge via alternative splicing or promoter switching. Moreover, it should be investigated, if the immune response of larvae differs from that of workers, since qRT-PCR analysis indicated that several genes seem to be more strongly induced after an immune-challenge in larvae than in workers.

7.3 Analysis of *C. floridanus* haemolymph

The antibacterial activity of the haemolymph from *C. floridanus* major workers was tested against *E. coli* D31, *M. luteus* and *B. subtilis* in inhibition-zone assays (Figure 39). Growth of Gram-positive *M. luteus* was clearly inhibited by the haemolymph from bacteria-injected as well as from untreated ants. Haemolymph from bacteria-injected animals caused larger zones of inhibition suggesting that additional defense molecules had been produced after the bacterial challenge. Accordingly, growth of Gram-positive *B. subtilis* was not inhibited by haemolymph from healthy animals, but by two haemolymph samples from bacteria-injected ants.

No zone of inhibition could be detected on plates with *E. coli* D31. This Gram-negative bacterium is commonly used as a test organism in inhibition-zone assays with insect haemolymph, because its mutated lipopolysaccharide makes it quite susceptible to killing by haemolymph (Boman et al. 1974). Therefore, it was surprising, that the growth of *E. coli* D31 was not inhibited by *C. floridanus* haemolymph samples. It was assumed that *C. floridanus* haemolymph contains several molecules with activity against Gram-negative bacteria. For example, the AMP hymenoptaecin is known to be active against several Gram-positive and Gram-negative bacteria (Casteels et al. 1993). Moreover, a recombinantly expressed mature hymenoptaecin peptide from *C. floridanus* peptide was already shown to be lethal for *E. coli* D31 (Ratzka et al. 2012a). Thus, the negative result might be caused by the haemolymph sampling procedure. During melanization reactions ROS are produced, which can damage proteins. To prevent melanization of haemolymph samples, phenylthiourea and aprotinin were added to the samples, as described previously (Randolt et al. 2008). Nonetheless, melanization of the haemolymph was observed. When the haemolymph of bacteria-injected animals was leaking out of the wound and collected with a glass capillary, it was already considerably more brownish than that of untreated animals. Further darkening of the samples was observed after application on top of the agar layer. This might have altered the antibacterial properties of *C. floridanus* haemolymph.

For identification of immune-relevant factors on the protein level haemolymph proteins were separated by SDS polyacrylamide electrophoresis on two different gels with conditions for separation of proteins in the range of 30-200 kDa and 3-30 kDa, respectively (Figure 40 and Figure 41). No enhanced protein bands could be detected in the haemolymph samples from

bacteria-injected ants, whereas using the same method several immune-induced proteins were identified in haemolymph samples from honey bees, e.g. hymenoptaecin, defensin-1 and PGRP-S2 (Randolt et al. 2008).

A number of different bands were detected on the 10 % polyacrylamide gel, which was run for separation of larger proteins (30-200 kDa). The majority of these haemolymph proteins are presumably not inducible defense proteins, but rather constitutively expressed storage proteins, e.g. hexamerins (Martinez et al. 2000; Wheeler and Martinez 1995). A differential regulation of some proteins in response to immune-challenge may not have been detectable, as the protein bands were quite densely packed and the gel lanes were partially blurry. Possibly 2D gel electrophoresis might result in a better separation of proteins and facilitate to identify immune-related proteins in the haemolymph of *C. floridanus*.

The transcriptional data suggested enhanced synthesis of the AMPs hymenoptaecin and defensin-1 in the fatbody tissue of *C. floridanus* after an immune-challenge. Since both precursor proteins have a predicted signal sequence, it was assumed that the mature AMPs would be secreted into the haemolymph, as already shown for honey bees (Randolt et al. 2008). The predicted mature AMPs have a rather low molecular weight of about 10.4 kDa (hymenoptaecin) and 4.4 kDa (defensin-1). Therefore, the haemolymph samples were also analyzed on a 15 % polyacrylamide gel with adapted conditions for separation of small proteins/peptides (Schägger 2006; Schägger and von Jagow 1987). However, no induced peptides could be detected in the haemolymph samples from bacteria-injected individuals. Conspicuously, no peptide band with a size smaller than 10 kDa could be detected at all. It is recommended to fix small proteins on polyacrylamide gels before staining, because otherwise they might get flushed out during washing steps. In this study, polyacrylamide gels were fixed with methanol as described previously. Fixing of small proteins with glutaraldehyde might achieve better results, as peptides get cross-linked through this reagent (Westermeier and Gronau 2005). Apart from that, AMPs might be present at too low concentration in haemolymph samples for detection with Coomassie staining. Accordingly, *C. floridanus* mature hymenoptaecin peptides could be detected with more sensitive Western Blots using specific antibodies (Maria Kupper, unpublished results). Application of larger haemolymph sample volumes would probably not improve AMP detection, since small peptides can get caught in high molecular weight proteins. AMPs might be visible on polyacrylamide gels after deprivation of high molecular weight proteins through acidic/methanolic extraction of the haemolymph, as already described for *Galleria mellonella* (Cytrynska et al. 2007).

7.4 Characterization of the *C. floridanus hym* gene and mRNA structure

In the recently published genome sequence of *C. floridanus* the *hym* gene escaped detection possibly due to sequencing problems of this gene carrying multiple direct repeats (Bonasio et al. 2010). However, two contigs (AEAB01001738.1 and AEAB01001739.1) were found

which harbour the 5`- and 3`- ends of the gene (Figure 10). Since no other contigs encoding DNA sequences resembling the *hym* gene were discovered, the genomic data also supports the existence of a single *hym* gene, thus confirming that the above described variable repeat numbers after PCR amplification were a technical artefact caused by the tandem repeats. The characterization of the hymenoptaecin revealed a very peculiar modular composition of the deduced peptide(s) as compared to hymenoptaecins of other Hymenoptera. The hymenoptaecins known from other hymenopterans such as *A. mellifera* (Casteels-Josson et al. 1994) and *Bombus ignitus* (Choi et al. 2008) show significant sequence homology to the hymenoptaecin domains repeated several times in the multipeptide precursor of the *C. floridanus* hymenoptaecin, suggesting structural and functional similarities. The *A. mellifera* hymenoptaecin is 93 aa long, including a 2-pyrrolidone-5-caboxylic acid at the N-terminus, which is derived from glutamine (Casteels et al. 1993). With the exception of the first so-called hymenoptaecin-like domain, the six deduced mature *C. floridanus* hymenoptaecins are 97 aa long and all start with a glutamine residue. Therefore, an amino-terminal blocking by forming 2-pyrrolidone-5-carboxylic acid is very likely also for the *C. floridanus* hymenoptaecin peptides. One of the putative mature hymenoptaecin peptides was overexpressed in *E. coli* and shown to exhibit moderate antibacterial activity (performed by Maria Kupper, Ratzka et al. 2012a). The possibility of an amino-terminal blocking might further increase the antibacterial potency of *Cfl*-hym peptides.

The exact mode of action of hymenoptaecin peptides is unknown so far, but is likely to involve membrane permeabilization processes (Casteels et al. 1993). The structure prediction of a similar hymenoptaecin peptide from *Nasonia* revealed a beta-sheet structure without disulfide bonds, in which the majority of glycines was exposed on the molecular surface (Gao and Zhu 2010). The investigated ant hymenoptaecin peptides contain many highly conserved GxxxG and GxxxG-like motifs (substitution of one or both glycine residues by other small residues, such as alanine or serine). These motifs are known to mediate transmembrane domain association (Russ and Engelman 2000; Senes et al. 2004) and might therefore interact with bacterial proteins or with each other resulting in pore formation as known for class II-b bacteriocins (Nissen-Meyer et al. 2010).

In contrast to the *A. mellifera* hymenoptaecin, the *C. floridanus* hymenoptaecin has a complex precursor organization. A comparable precursor organization is known for the *N. vitripennis* hymenoptaecin, which encodes three AMP-like peptides, including one with similarity to the hymenoptaecin domains of *C. floridanus* (Gao and Zhu 2010). Overall, the *C. floridanus* hymenoptaecin precursor structure is more similar to the multipeptide precursor structure of apidaecins, consisting of several repeated units (Casteels-Josson et al. 1993). As for apidaecin precursors the mature hymenoptaecin peptides might be released by a three step mechanism, which is similar to maturation procedure of the yeast alpha-mating factor, since the repeats are flanked by repeating -X-A- (or -X-P-) sequences (Julius et al. 1983). The initial processing is probably mediated by the *KEX2*-encoded endoprotease, which cuts at the C-terminus of the

basic dipeptides Arg/Lys (RK) or Arg/Arg (RR) (Fuller et al. 1989). The next step is the C-terminal maturation via the *KEX1*-encoded carboxypeptidase, which removes both basic residues (Dmochowska et al. 1987). The last step is the N-terminal maturation of the spacer-mature peptides by a dipeptidyl aminopeptidase that removes E/D-A/P dipeptides (Julius et al. 1983). Homologues of the respective enzymes are present in *C. floridanus* (GenBank Acc. No. EFN61704 to CAA96915 (E-value: 3E-35), EFN64345 to CAA96143 (E-value: 7E-94) and EFN67964 to NP_014862 (E-value: 5E-55)).

Despite the similarities in the multipeptide precursors, the *C. floridanus hym* differs from the *A. mellifera apidaecin* with regard to the gene structure. The latter one consists of several exons, each encoding a functional and distinct apidaecin peptide (Evans et al. 2006). In contrast, the hymenoptaecin mature peptide regions are encoded by a single exon only. This intronless gene structure prohibits the possibility of generating different transcripts by splice variation. Nevertheless, the multipeptide precursor structure of the *Camponotus hym* gene would allow the amplification of the antibacterial response despite the presence of only a single gene, as already suggested for apidaecins (Casteels-Josson et al. 1993). Furthermore, a surprisingly high number of mature hymenoptaecin peptide variants with slightly different amino acid sequence was found in the pooled samples from *C. floridanus* workers deriving from a single colony. This high level of individual sequence variation has also been described for apidaecin exons from different bees (Casteels-Josson et al. 1993; Evans et al. 2006).

The genome sequences of other ant species (Bonasio et al. 2010; Nygaard et al. 2011; Smith et al. 2011a; Smith et al. 2011b; Suen et al. 2011) revealed the presence of at least one gene locus encoding a hymenoptaecin precursor with similar domain structure as the *C. floridanus* hymenoptaecin. In a recent study the complete *hym* gene from *S. invicta* was amplified and sequenced revealing seven repeated hymenoptaecin domains within this gene (Zhang and Zhu 2012). The presence of such multidomain *hym* genes in all ant species indicates an ancient origin of this gene structure early in the evolution of ants. However, the genome sequences indicate problems during the assembly and the gene prediction step of the genome projects and only a few sequences seem to be complete. All investigated complete hymenoptaecin peptide regions from ants are encoded by a single exon only, which is evidence for exon duplication (Street et al. 2006). The preliminary phylogenetic analysis of mature hymenoptaecin peptides in the context of this work (performed by Frank Förster) suggests that the duplication event occurred independently in each species after separation (Figure 13), which is in accordance with another recent study (Zhang and Zhu 2012). Prospective studies will reveal the full length hymenoptaecin precursor sequences from other ant species by direct sequencing and will deliver insights into the evolutionary history of the hymenoptaecin protein family in ants.

7.5 Characterization of *defensin* genes from *C. floridanus*

Bioinformatical prediction of defensins revealed the presence of at least one *defensin* gene in all investigated ant genomes with homology to *def-1* or *def-2* from *A. mellifera* (Figure 15). In the context of this study it was shown that *C. floridanus* and *S. invicta* encode both *defensin* genes. Therefore, it seems that the last common ancestor of the ants and *A. mellifera* encoded both *defensins*. Based on the assumption that the genomes of *L. niger, F. aquilonia,* and *M. scabrinodis* contain only the published *defensin* genes, the reconciled gene tree (Figure 16) exhibits many gain and loss events. Multiple duplication events indicate a high adaptive potential and evolutionary plasticity of the antimicrobial peptides in ants. The *C. floridanus* mature defensin-1 and defensin-2 peptide sequences are well conserved with other ant defensins. However, the *def-1* gene comprises three exons and two introns, in contrast to other characterized ant *defensin* genes (Viljakainen and Pamilo 2005; Viljakainen and Pamilo 2008), which have two exons and one intron. Interestingly, a similar intron-exon composition is also known for other hymenopteran *def-1* genes, e.g. from *A. mellifera* (Klaudiny et al. 2005), *B. ignitus* (Choi et al. 2008) and *N. vitripennis* (Tian et al. 2010), while the *Drosophila defensin* gene does not carry any intron at all (Dimarcq et al. 1994). In contrast to other insect defensins, the bee defensin-1 has an extra stretch of 11 amino acids at its C-terminus, which encodes an additional C-terminal α-helical domain (Casteels-Josson et al. 1994). The *C. floridanus* defensin-1 has a short C-terminal extension of three amino acids in length. The precursors of the bee defensins have an extra amino acid, a glycine (G), at their C-termini, which seem to be amidated as suggested in the mature *A. mellifera* defensin (Casteels-Josson et al. 1994). As the deduced *C. floridanus* defensin-1 also ends with a G, it may as well be amidated. According to this, the mature *C. floridanus* defensin-1 is 3 amino acids longer than all other known ant defensins and it is so far the only known ant defensin which has an additional exon that is lacking from most other insects (Froy and Gurevitz 2003). Further investigations will reveal, if this C-terminal extension can also be found in defensins from other ant species or if it is a special feature of *C. floridanus*.

7.6 Characterization of *C. floridanus TyrOH* mRNA

The full length *TyrOH* mRNA sequence of 3531 bp was determined by RACE experiments. The deduced *C. floridanus* TyrOH protein is 572 aa long (Figure 17). A serine residue known to be essential for activation of TyrOH as well as the catalytic domain are highly conserved in *C. floridanus* TyrOH suggesting enzymatic activity of the protein. Tyrosine hydroxylases are highly conserved enzymes from vertebrates and invertebrates (E.C. 1.14.16.2), which catalyze the hydroxylation of tyrosine to 3,4-dihydroxyphenylalanine (DOPA). In insects DOPA is required as a precursor molecule for melanization and cuticle sclerotization (Andersen 2010) and furthermore acts as a neurotransmitter in the central nervous system (Osborne 1996). Two TyrOH isoforms are expressed in *D. melanogaster*, which are generated by tissue-specific

alternative splicing of the primary transcript from the gene *pale* (Birman et al. 1994). The shorter isoform is expressed in neural cells (Birman et al. 1994) and required for neural function (Budnik and White 1987; Friggi-Grelin et al. 2003). The longer isoform, which contains an additional acidic region of 71 aa in the regulatory domain, is expressed in non-nervous tissues, especially in the epidermis (Birman et al. 1994) and is required for cuticle tanning (sclerotization and melanization) (Neckameyer and White 1993; True et al. 1999). The *C. floridanus* TyrOH cDNA, which was cloned from midgut and fatbody mRNA, is more similar to the epidermal TyrOH isoform from *D. melanogaster* than to the neural form, as it also contains a similar acidic region of 65 aa. This acidic region seems to constitutively activate the epidermal form of *D. melanogaster* TyrOH and to make it less sensitive to feedback inhibition by dopamine, which likely improves TyrOH function during cuticle tanning (Vie et al. 1999). Thus, TyrOH might also be required for cuticle tanning in *C. floridanus*, as already suggested for other insects (Gorman et al. 2007; Gorman and Arakane 2010; Hwang et al. 2010). In accordance with that, the *TyrOH* gene expression was significantly elevated in *C. floridanus* late pupae, in which the newly formed cuticle begins to darken and to sclerotize, in comparison to larvae and workers.

The supposed role of *TyrOH* for cuticle tanning in *C. floridanus* was investigated by RNAi experiments. In *Tribolium castaneum* injection of 200 ng dsTyrOH in larvae resulted in death of the animals before adult eclosion, whereas injection of only 2 ng dsTyrOH in pupae was already sufficient to obtain a pale phenotype in adults (Gorman and Arakane 2010). Since *C. floridanus* pupae are considerably bigger than those of *T. castaneum*, a comparatively large amount of 4 µg dsTyrOH was injected per individual. Early pupae were chosen as test animals, in order to achieve a gene knockdown effect during the late pupal stage, at which *TyrOH* levels are normally elevated. However, at the tested time points no knockdown of *TyrOH* transcripts could be detected in dsTyrOH-injected individuals in comparison to control animals (Figure 19). Thus, the impact of *TyrOH* knockdown on cuticle tanning could not be investigated. So far, application of RNAi has not been extensively adapted for *C. floridanus*, whereas this method is already well established for *T. castaneum* (Posnien et al. 2009). Maybe knockdown of *TyrOH* expression could be achieved after modification of the experimental design. For example, injection of a TyrOH-specific siRNA might produce a stronger RNAi effect than injection of dsRNA, since the cleavage step of the dsRNA into siRNAs within the ant is omitted. Accordingly, the injection of specific siRNAs caused the downregulation of a cGMP-dependent protein kinase in *C. floridanus* (Bonasio et al. 2010). However, not achieving *TyrOH* knockdown might also be a species-specific problem. *TyrOH* was shown to be expressed in epithelial cells that secrete brown, sclerotized cuticle, but not in the fat body of healthy animals (Gorman et al. 2007; Gorman and Arakane 2010). Whereas systemic gene knockdown has been reported for several genes in *T. castaneum* (Bucher et al. 2002; Posnien et al. 2009; Tomoyasu and Denell 2004), a recent study in *A. mellifera* indicated that an RNAi effect could only be achieved in fatbody cells of honey bees (Jarosch and Moritz 2011).

Accordingly, knockdown of *TyrOH* expression in epidermal cells of *C. floridanus* might be hardly attainable.

The SSH approach as well as the Illumina sequencing data showed upregulation of *C. floridanus TyrOH* gene expression in response to bacterial challenge. Moreover, the kinetics study revealed strong induction of *TyrOH* in midgut and fatbody after injection of Gram-positive and Gram-negative bacteria suggesting an immune function of TyrOH. A role of TyrOH in insect immunity was also indicated in studies on other insects including *D. melanogaster* (De Gregorio et al. 2001), *Galleria mellonella* (Seitz et al. 2003), *Manduca sexta* (Gorman et al. 2007) and *Plutella xylostella* (Hwang et al. 2010). It was assumed that TyrOH is involved in immune-associated melanization processes through the production of the melanin precursor molecule DOPA (De Gregorio et al. 2001; Gorman et al. 2007).

7.7 Expression kinetics of some selected immune genes and of the immune response towards *B. floridanus*

Following injection of Gram-positive and Gram-negative bacteria all genes included in the study of kinetics of the immune response were significantly induced, showing peak levels 8 h to 24 h post injection. Induction levels differed strongly among genes in comparison to untreated controls (Figure 20). Expression levels of genes *GNBP*, *hsp*, *snake* and *serpin* remained on a higher expression level after injection of Gram-positive *M. flavus* in comparison to the Gram-negative *S. marcescens* (Figure 20A-D). Gram-positive bacteria predominantly trigger activation of the Toll pathway, which regulates many late-responsive genes (Lemaitre et al. 1997). Sterile Ringer-injection caused a significant induction of immune genes, which was however not as enduring as after injection of bacteria (Figure 21). This phenomenon has also been described in other insects and implies that wounding itself might cause a temporary innate immune response (Boutros et al. 2002; Choi et al. 2008).

The immune response towards the endosymbiont *Blochmannia floridanus* injected into the haemocoel was almost as strong as to other bacteria, implying that *Blochmannia* does not escape immune detection outside the bacteriocytes. In contrast to the Gram-negative *Serratia*, the endosymbiont had a slightly weaker (albeit not significant) effect on the induction of the genes *GNBP* and *hsp* as expression levels decreased earlier (*GNBP*) or did not reach such high levels (*hsp*) (Figure 21A, B). The observed differences may not be due to endosymbiotic bacteria inducing a less strong immune response though. An equal number of bacteria injected may not translate into an equal quantity of MAMPs, e.g. when bacteria differ in size, larger bacteria may possess more recognition patterns due to their increased surface area.

It is often argued that inherited endosymbionts like *Blochmannia* may not interact with the host´s immune system due to their intracellular location (Wurm et al. 2011). However, the insect immune system can detect intracellular bacteria via peptidoglycan recognition proteins

(PGRPs) (Kurata 2010). *C. floridanus* encodes several PGRPs in its genome (Bonasio et al. 2010), including a homologue of PGRP-LE known to detect intracellular bacteria (Kurata 2010). *B. floridanus* still has the genomic capacity to synthesize cell-wall related MAMPs (Zientz et al. 2004), and thus the endosymbiont might permanently trigger the expression of host immune factors.

In weevils and aphids the bacteriocytes are combined into a specific organ, the bacteriome, which is attached to the foregut or within the body cavity (Baumann 2005; Heddi et al. 1999a). In contrast, the bacteriocytes in *Camponotus* are intercalated between midgut cells in larvae and adults. In pupae the majority of the midgut cells carry *Blochmannia* and during this developmental stage the midgut itself can be regarded as a bacteriome-like organ (Sauer et al. 2002; Stoll et al. 2010). The midgut of insects is usually filled with a consortium of bacteria (Dillon and Dillon 2004), which are probably tolerated by repression of NFκB–dependent AMP genes (Ryu et al. 2008). Interestingly, no gut microbiota other than *B. floridanus* could be detected in midguts of *C. floridanus* (Feldhaar et al. 2007), which might indicate an elevated immune status of this endosymbiont-bearing tissue in comparison to midgut tissue of insects species lacking endosymbionts.

7.8 Gene expression according to developmental stage and tissue

The *C. floridanus* genome sequence (Bonasio et al. 2010) in combination with the Illumina sequencing data of the present study made it possible to identify putative candidate genes involved in host-symbiont interactions based on their significant homology to previously investigated symbiosis-relevant genes from other insects. The expression of these genes in the endosymbiont-bearing midgut tissue of different stages was compared with that in the residual body parts. The aim of this expression study was to reveal differences in the expression of candidate genes, which might causatively correlate with the changes of endosymbiont number during host ontogeny.

7.8.1 Reference gene selection

In order to obtain reliable gene expression results, it was very important to validate the stable expression of the reference genes used for normalization of the qRT-PCR data. Such reference genes are often just selected based on their homology to already validated reference genes from other organisms. For example, housekeeping genes, like *β-actin* or *GAPDH*, were commonly used for normalization, as they are required for maintenance of basic cellular functions and thus were considered to be constitutively expressed in all cells of an organism. However, there is emerging evidence that also expression of such well-known housekeeping genes may be affected by certain biological treatments or even vary in different tissues (Bustin 2000; Bustin 2002). Accordingly, it is recommended to first validate the stability of

intended reference genes in a certain organism in dependence of the particular treatment before using them for normalization, in order to prevent falsely biasing of the results. Since this study comprised several different stages and tissues of *C. floridanus*, the expression stability of several reference genes was initially validated using BestKeeper software (Pfaffl et al. 2004). The expression of the genes *rpL32*, *rpL18*, *EF1α* and *GAPDH* was shown to be quite stable across all different samples and thus these genes were appropriate for normalization of the expression data (Table 14). Thus, this study provides a validation of different candidate genes, which may also be used for normalization of qRT-PCR data in prospective studies on *C. floridanus*.

7.8.2 Immune-relevant genes

C. floridanus contains four *PGRP* genes in its genome. Sequence analysis revealed that two of them, *PGRP-LB* and *PGRP-SC2*, very likely encode proteins with functional amidase activity (Figure 23). The *Drosophila* homologues of these proteins were already shown to downregulate the IMD pathway via cleavage of the trigger molecule, DAP-type peptidoglycan. By doing so, these amidase PGRPs seem to particularly suppress the immune response towards commensal bacteria in the gut (Bischoff et al. 2006; Zaidman-Remy et al. 2006). Furthermore, studies on other insects indicated a key role for PGRP-LB in regulation of the immune response towards endosymbionts (Anselme et al. 2006; Wang and Aksoy 2012; Wang et al. 2009). Strikingly, both *PGRP-LB* and -*SC2* were highly expressed in the midgut tissue of late pupae (Figure 22A, B) correlating with the highest number of *Blochmannia* cells present in this tissue (Stoll et al. 2010). During the pupal stage, no food is ingested and thus there is no need to downregulate the immune response towards ingested bacteria. However, exactly during this developmental stage, the entire midgut is transformed into a huge symbiotic organ resembling true bacteriomes of aphids and tsetse flies, which consist exclusively of endosymbiont-bearing cells (Stoll et al. 2010). This strongly suggests the involvement of PGRP-LB and -SC2 in the control of *Blochmannia* number in *C. floridanus*. Especially in pupae around metamorphosis, *Blochmannia* was found to be present in midgut cells other than bacteriocytes, which indicates that *Blochmannia* may be able to actively invade other cell types within the midgut tissue (Stoll et al. 2010). PGRP-LB and -SC2 might prevent activation of immune responses against *Blochmannia* via cleavage of released peptidoglycan. A recent study showed that PGRP-LB from tsetse flies itself exhibits antimicrobial activities (Wang and Aksoy 2012). Thus, PGRP-LB might additionally directly regulate number of invasive *Blochmannia* in the midgut lumen and the outer surroundings of the midgut and prevent excessive proliferation of the endosymbiont. In contrast, expression of *PGRP-LB* and -*SC2* was extremely low in residual body parts of pupae. This might keep *Blochmannia* from spreading into other tissues, as it is recognized and presumably killed by the ant´s immune system (Ratzka et al. 2011). The remarkably high expression level of *PGRP-SC2* in the body of adult workers possibly indicates a higher tolerance level for

peptidoglycan before triggering an immune response and might be an indication of immunosenescence in foraging workers (Amdam et al. 2005).

The *C. floridanus* PRR designated as PGRP-LE (GenBank Acc. No. EFN63542) showed high sequence homology to both *Drosophila* PGRP-LC and PGRP-LE, which activate the IMD pathway after recognition of DAP-type peptidoglycan (Kurata 2010; Royet et al. 2011). Since a transmembrane domain is predicted for PGRP-LE (-LC) from *C. floridanus*, the protein seems to be a transmembrane receptor similar to PGRP-LC from *Drosophila* (Werner et al. 2000). However, further studies are needed to investigate the exact location and function of PGRP-LE (-LC) in *C. floridanus*, since the corresponding gene may also be differentially spliced leading to several protein isoforms with distinct function, as already shown in *Drosophila* (Kaneko et al. 2006; Werner et al. 2003; Werner et al. 2000). The *C. floridanus PGRP-LE* (*-LC*) gene was quite evenly expressed in the midgut and the residual body of different stages. Only in the larval body, it was expressed on a marginally lower level (Figure 22C). This lower *PGRP-LE* (*-LC*) expression in *C. floridanus* larvae might be linked to the finding that excessive activation of JNK pathway mediated through IMD activation by PGRP-LC/-LE impairs normal development in *Drosophila* larvae (Maillet et al. 2008). The expression pattern of *C. floridanus PGRP-LE* gene did not indicate an involvement in host-symbiont interactions.

Activation of the IMD pathway leads to transcription of effector genes via the NF-κb transcription factor Relish (Rel) (Hedengren et al. 1999). Expression of the corresponding *Rel* gene in *C. floridanus* did not exhibit dramatic changes during development or in the midgut tissue, which might hint to endosymbiont regulation (Figure 22F). However, expression level of *Rel* gene does not necessarily correlate with induction of AMP synthesis, since Rel protein needs to be activated by endoproteolytic cleavage before it promotes transcription of NF-κb-regulated genes (Stöven et al. 2000; Stöven et al. 2003).

The expression of the *C. floridanus PGRP-2* and *GNBP* gene was significantly higher in the body of all stages than in the midgut tissue. These genes encode PRRs with homology to PGRP-SA and GNBP1 from *Drosophila*, which jointly activate the Toll pathway upon recognition of Lys-type peptidoglycan from Gram-positive bacteria (Gobert et al. 2003; Michel et al. 2001; Park et al. 2007; Pili-Floury et al. 2004). Since Toll activation occurs though the cooperation of both PRRs, the corresponding genes likely show a similar expression pattern in *C. floridanus* (Figure 22D, E). Both genes are probably mainly produced in the fat body tissue (Royet and Dziarski 2007; Royet et al. 2011), which was included in the body samples. A role of these PRRs in recognition of Gram-negative *Blochmannia* is rather unlikely, since they are supposed to recognize Gram-positive bacteria and their expression is quite low in the midgut tissue. The significant higher expression of these genes in the body indicates preparedness of the immune system to detect and immediately act against invading bacteria.

Expression of the AMP genes *hym* and *def-1* varied strongly in the body of workers (Figure 22G, H). The worker ants included in the pooled samples were just taken from the nest surface of the different colonies. Thus, the variations in AMP expression may reflect different age or nutrition status of the investigated animals possibly affecting immunocompetence (Alaux et al. 2010; Amdam et al. 2005; Armitage and Boomsma 2010), or some individuals were just staging an immune response towards an infection. Accordingly, the PRR genes such as *PGRP-2* and *GNBP*, which are located upstream of the signaling pathway, exhibited a similar variability of the expression levels in the body of workers, whereas the effect was more pronounced for the downstream effector genes *hym* and *def-1*.

A recent study on weevils reported that the antimicrobial peptide ColA controls cell division of SPZ endosymbionts within bacteriocytes (Login et al. 2011). No homologue of *colA* gene could be identified in the *C. floridanus* genome using NCBI BLASTX. The expression pattern of *hym* and *def-1* during development did not reveal any indication for a role in the control of the intracellular endosymbionts. However, both genes were shown to be induced in response to extracellular *Blochmannia* and thus may prevent their uncontrolled spreading outside of the bacteriocytes (Ratzka et al. 2011).

The *tollip* gene was found to be highly expressed in the bacteriome of *Sitophilus* weevils (Anselme et al. 2008). As the gene exhibited significant homologies to a mammalian negative regulator of the Toll pathway (Zhang and Ghosh 2002), it was supposed to be involved in mediating host-symbiont-interactions in weevils (Anselme et al. 2008). However, the *tollip* homologue of *C. floridanus* was evenly expressed across tissues and stages, and thus there was no evidence for a role of this gene in endosymbiosis of *C. floridanus* (Figure 22I).

7.8.3 Genes involved in the lysosomal system, autophagy and ROS- production / detoxification

In the pea aphid *Acyrthosiphon pisum*, lysosome-related genes like *i-type lysozyme* (*i-type lyso*) and *cathepsin L* (*cath L*) were found to be highly expressed in the bacteriome tissue and were shown to be induced in bacteriocytes of post-reproductive aphids suggesting degradation of *Buchnera* cells via the lysosomal system (Nakabachi et al. 2005; Nishikori et al. 2009b). Moreover, expression of a host serine carboxypeptidase vitellogenic-like (CPVL) was linked to degradation of *Buchnera* endosymbionts (Nishikori et al. 2009a). Interestingly, homologues of *i-type lyso* and *CPVL* were highly expressed in the midgut of *C. floridanus* pupae, whereas *cath L* was strongly expressed in the midgut of adult workers (Figure 24B, C, E). In contrast to other analyzed endosymbionts, *B. floridanus* has completely lost the genes encoding the flagellar apparatus presumably involved in transport functions (Gil et al. 2003). Due to the reduced transport capability of *Blochmannia*, it is ambiguous, how endosymbiont-produced nutrients are properly exported to the host cell. Thus, degradation of the endosymbiont seems to be the easiest way to make nutrients available for the ant host. In older workers symbiosis

does not seem to be very important anymore, as it degenerates with the age of the workers (Stoll et al. 2010; Stoll et al. 2008) possibly also by lysosomal degradation of endosymbionts. Thus, the expression patterns of *i-type lyso*, *CPVL* and *cath L* might indicate lysosomal digestion of endosymbionts. Accordingly, in previous studies a high number of vesicles possibly derived from the lysosomal system was observed in pupal midgut cells (Stoll et al. 2010). In fact, vesicle-contained *Blochmannia* were occasionally observed by electron microscopy in previous work (Wolschin et al. 2004).

Apart from *i-type lyso*, the *C. floridanus* genome encodes two c-type lysozymes. The deduced amino acid sequence of these proteins is almost identical and thus only *c-type lyso 1* was included in the expression study. Its expression was comparatively high in the midguts of larvae and adults as well as in the body of adults and quite low in pupae (Figure 24A). Several c-type lysozymes from *Drosophila* have already been shown to be highly expressed in the midguts of larvae and adults, but not in pupae (Daffre et al. 1994). Moreover, c-type lysozymes were constitutively expressed in most tissues of adult silkworms (Lee and Brey 1995) and mosquitoes (Li et al. 2005) and it was assumed that they might be involved in several processes including digestion and immunity. In hen egg white lysozyme (HEWL) the residues E35 and D52 are required for muramidase activity (Malcolm et al. 1989). While the corresponding E is conserved, the D residue is exchanged by a S residue in *C. floridanus* c-type lyso 1 as well as in homologues from other hymenopterans (Figure 26). Exchange of D52 with a S in HEWL abolished its catalytic activity, whereas it was still as bactericidal as wildtype HEWL (Ibrahim et al. 2001). Muramidase-independent bactericidal activity has also been described for other lysozymes (Düring et al. 1999; Masschalck and Michiels 2003). Thus, c-type lysozymes from *C. floridanus* and other hymenopterans might still achieve a defensive function. Since c-type lyso 1 from *C. floridanus* possesses a signal sequence, it may be secreted e.g. into the gut lumen. In the larval and adult midgut it might be involved in the antibacterial defense against ingested bacteria. Since no food is ingested during pupal stage, the gene might be downregulated in the pupal midgut. This might additionally facilitate the spreading of *Blochmannia* endosymbionts throughout the whole midgut tissue during this stage. Accordingly, c-type lysozyme was also found to be downregulated in the bacteriome of weevils (Vigneron et al. 2012).

Expression of a *lysosomal aspartic protease* gene (*lap*) was investigated in this study, since it encodes a homologue of a protein from *C. pennsylvanicus*, which was identified in the crop content and shown to be upregulated in response to bacteria injection (Hamilton et al. 2011). Accordingly, the *lap* gene was highly expressed in the body samples from adults, which included the crop tissue. A connection of *lap* expression to endosymbiont regulation could not be detected (Figure 24D).

Lysosomal degradation of endosymbionts may be preceded by autophagocytic processes, as already described for regulation of *Wolbachia* populations in diverse symbioses (Voronin et al. 2012). However, the expression pattern of *C. floridanus apg-5*, encoding a protein needed

for autophagosome formation, did not indicate an involvement of autophagy in control of endosymbionts (Figure 24F).

In *C. floridanus* the endosymbiont-bearing bacteriocytes are intercalated in the midgut tissue (Sauer et al. 2002; Schröder et al. 1996; Stoll et al. 2010). Except intracellular *Blochmannia*, there is no significant microflora present in the midgut of *C. floridanus* (Bonasio et al. 2010; Feldhaar et al. 2007). In contrast, other insects including various ants (Anderson et al. 2012; Poulsen and Sapountzis 2012) generally harbour diverse gut microbiomes (Dillon and Dillon 2004). In *Drosophila* Dual oxidase (Duox) essentially contributes to antimicrobial activities in the gut through production of ROS after ingestion of bacteria (Ha et al. 2005a). Accordingly, Duox might also account for sterilization of the midgut lumen in *C. floridanus*. Expression of the corresponding gene was comparatively low in the midgut tissue of untreated animals, but significantly elevated in pupae (Figure 24H). Presence of maximal *Blochmannia* numbers in the midgut of pupae might trigger *Duox* expression, as shown for *Wolbachia*-infected mosquitoes (Pan et al. 2012). However, the elevated *Duox* expression levels in the midgut of pupae might also be linked to its suggested role in melanization and sclerotization processes (Anh et al. 2011), which would also explain its strong expression in the body of pupae.

Wolbachia-induced ROS production leads to upregulation of antioxidant genes like *superoxide dismutase* (*SOD*) and *ferritin* in mosquitoes (Brennan et al. 2008; Pan et al. 2012). The expression data of this study did not indicate an increased expression of *SOD* or *ferritin* in response to higher *Blochmannia* number in *C. floridanus* pupae in comparison to larvae or workers (Figure 24G, I).

7.9 Immune responses in the late pupal stages

The high expression levels of *PGRP-LB* and -*SC2* gene in the midguts of untreated pupae suggested that the immune response of this tissue was down-modulated to enable endosymbiont tolerance. In contrast, the comparatively low expression levels of the same genes in the residual body parts of pupae indicated the readiness of the immune system to possibly counteract undesirable spreading of endosymbionts. By an artificial immune-challenge of pupae it was confirmed that in fact the immune response of the midgut tissue differs from that of other body parts. The PRR genes *PGRP-2* and *GNBP*, which presumably trigger an immune response after microbial challenge (Gobert et al. 2003; Michel et al. 2001; Park et al. 2007; Pili-Floury et al. 2004), were significantly stronger induced in the body than in the midgut tissue (Figure 28D, E). In contrast, expression of the two PRR genes *PGRP-LB* and *PGRP-SC2*, which have a presumably active amidase domain and thus possibly counteract an activation of the immune response, was comparatively low in the pupal body, even after an immune-challenge (Figure 28A, B). Both genes were much stronger expressed in the midgut tissue than in other body parts. The median of *PGRP-LB* expression in the midgut even increased after an immune-challenge, albeit not significantly. Expression of the

PGRP-LE (*-LC*) gene was only slightly induced in immune-challenged animals (Figure 28C), but activation of the IMD pathway through this PRRs is likely regulated on the posttranscriptional level (Choe et al. 2005; Mellroth et al. 2005; Schmidt et al. 2008).

The gene encoding the NF-κb transcription factor Rel was upregulated in immune-challenged *C. floridanus* pupae, which was significant in the body but not in the midgut tissue (Figure 28F). After endoproteolytic activation Rel induces the transcription of AMP genes (Stöven et al. 2000; Stöven et al. 2003). In accordance with that, the transcription of the key antimicrobial effector gene *hym*, which is likely controlled by Relish (Schlüns and Crozier 2007), was significantly stronger induced in the pupal body than in the midgut (Figure 28G). In the context of this thesis it was shown that a mature hymenoptaecin peptide of *C. floridanus* exhibits an antimicrobial effect of against Gram-negative bacteria (performed by Maria Kupper, Ratzka et al. 2012a) and thus excessive synthesis of these AMPs in the midgut might harm the Gram-negative endosymbionts. In contrast, the gene encoding defensin-1, which presumably acts mainly against Gram-positive bacteria (Cociancich et al. 1993; Hetru et al. 1998), was similarly induced in the pupal midgut and residual body parts (Figure 28H).

Expression of *i-type lyso* was not affected by an immune-challenge, but it was expressed significantly stronger in the midgut than in the residual body (Figure 28I). This might indicate a digestive rather than a defensive function of this lysozyme. Accordingly, the encoded i-type lysozyme has an acidic isoelectric point of 5.67, which is typical for digestive lysozymes (Callewaert and Michiels 2010). Since the encoded i-type lysozyme has no predicted signal sequence, it does not seem to be secreted into the midgut lumen, but might act intracellularly within acidic lysosomes.

7.10 Knockdown of *PGRP* gene expression by RNAi

RNAi technology is a suitable method to study gene functions in ants (Formicidae, Hymenoptera), for which so far no transgenic individuals can be obtained. However, the application of this method first needs to be adopted for ants, since to my knowledge until today only two studies exist describing a successful knockdown of gene expression after injection of dsRNAs in the ant species *Solenopsis invicta* (Lu et al. 2009) and after injection of siRNAs in *C. floridanus* (Bonasio et al. 2010).

The present study provides the first demonstration of a successful gene knockdown in Formicidae (Hymenoptera) by feeding of dsRNA to subcolonies of *C. floridanus* workers. Knockdown of *PGRP-LB* gene in the body of minor worker ants was achieved for up to 10 days by continuous feeding with the corresponding dsRNA (Figure 31D, E). In this way, *PGRP-LB* expression could even be knocked down during an immune response, which normally induces expression of this gene (Figure 37B). Feeding as a delivery system for dsRNA has many advantages in comparison to injection. First, it saves time to just feed a

group of animals with dsRNA instead of injecting each test animal. Second, injection often causes higher mortality rates, so that extra individuals have to be included to guarantee survival of enough test animals per group, which again costs time and money. Furthermore, injection is an invasive method that considerably stresses the test insects. The damage of the cuticle and the wounding caused by injection of dsRNA is sufficient to induce immune responses (Boutros et al. 2002; Choi et al. 2008; Ratzka et al. 2011), which hampers the interpretation of knockdown studies on immune genes. In social insects such as ants, reintroduction of injected animals into the colony can also be problematic as they may be rejected or harmed by their nestmates (Carlin and Hölldobler 1986). Thus, RNAi by feeding of dsRNA is a time-saving alternative approach for loss-of-function studies in social insects. Apart from that, the dsRNA for the feeding experiments of this study was purified from dsRNA-producing bacteria (Timmons et al. 2001), which was cheaper than *in vitro* transcription using commercial kits.

Knockdown of *PGRP-LB* expression was not achieved in the midgut tissue (Figure 31A-C). This was surprising, since studies on the tsetse fly *Glossina morsitans* (Diptera) (Walshe et al. 2009) as well as on the moth *Spodoptera litura* (Lepidoptera) (Rajagopal et al. 2002) suggested that feeding dsRNA results in gene knockdown in the midgut tissue rather than in the residual body parts. However, a study on the honey bee *A. mellifera* (Hymenoptera) indicated that gene knockdown in this insect could not be achieved in any other tissue than the fat body (Jarosch and Moritz 2011). As the *C. floridanus* body samples of this study included the fat body and *PGRP-LB* is generally expressed in this tissue (Zaidman-Remy et al. 2006), the detected knockdown likely occurred mainly in the fat body. Nonetheless, it is unclear how the RNAi signal might have reached the fat body tissue without effecting the expression in the midgut. As the dsRNA was shown to be quite stable in the crop content of *C. floridanus* workers, it may perhaps been taken up directly from this tissue.

In larvae no knockdown of *PGRP-LB* expression could be detected (Figure 31G, H). In *A. mellifera* feeding of dsRNA to larvae has been shown to induce RNAi in several studies (Nunes and Simoes 2009; Patel et al. 2007; Wolschin et al. 2011). However, honey bee larvae take up food by themselves, whereas *C. floridanus* larvae have to be fed by brood-tending workers. It was difficult to determine, how much dsRNA was passed from workers to larvae, and thus the ingested amount of dsPGRP-LB might not have been sufficient to elicit an RNAi effect in larvae. Apart from that, the digestive properties of midguts from larvae and workers may be considerably different (Erthal et al. 2007; Terra et al. 1988), which may impair the efficacy of ingested dsRNA to trigger gene knockdown. For example, dsRNA was not stable and thus not active in the midgut juice of starved *Spodoptera frugiperda* larvae (Rodriguez-Cabrera et al. 2010). Moreover, a study on leaf cutting ants revealed that the midgut of larvae contains more different digestive enzymes and exhibits higher enzyme activity than that of adults (Erthal et al. 2007). The effectiveness of feeding dsRNA to trigger RNAi in larvae may also be dependent from the chosen gene and might work for other target genes. Since no long-

lasting RNAi effect could be achieved in larvae by worker-mediated feeding of dsRNA, the effect of *PGRP-LB* / *-SC2* knockdown on endosymbiont number in pupae could not be investigated. Injection of dsRNA in late larvae or early pupae would presumably also not improve knockdown of *PGRP* gene expression in midguts of late pupae, as it is doubtful, that dsRNA would properly reach the midgut tissue and *PGRP* gene expression is quite low in the residual body parts of pupae, anyway.

The impact of *PGRP-LB* knockdown during an immune response was investigated (Figure 37 and Figure 38). It was assumed that the amidase PGRP-LB downregulates the immune response through cleavage of peptidoglycan molecules, which thereby lose their immunostimulatory properties, as already shown in other insects (Wang et al. 2009; Zaidman-Remy et al. 2006). Contrary to expectations, no enhanced or prolonged induction of immune effector genes such as *hym* and *def-1* could be detected in *PGRP-LB* knockdown animals at 48 h post immune-challenge. This negative result might be due to the fact that only dead bacteria were injected in this experiment. This means that the amount of immune-eliciting peptidoglycan did not increase during the time course of infection, since bacteria did not replicate within the ant´s body cavity. Thus the role of PGRP-LB as a negative regulator of the ant´s immune response might not have been obvious from this experimental design. Apart from that, *PGRP-SC2* gene, which encodes the other amidase PGRP of *C. floridanus*, could not be knocked down. The encoded PGRP-SC2 might fulfill a similar function as PGRP-LB in *C. floridanus* and thus might compensate for the knockdown of *PGRP-LB*. For clearness it is recommended to repeat this experiment with injection of living bacteria.

The failure of *PGRP-SC2* knockdown after feeding of the corresponding dsRNA (Figure 32) furthermore demonstrates that the efficiency of RNAi varies between different genes and might be dependent from the particular cell type expressing the target gene (Jarosch and Moritz 2011).

Depending on the investigated ant species, knockdown of gene expression might also work in larvae after feeding of a more concentrated dsRNA solution to worker ants. Moreover, it needs to be tested, if effective knockdown of gene expression in larvae may be achieved for other genes. However, in this study it was shown that feeding of dsRNA is a practicable method for RNAi in ants, at least suitable for adult workers.

7.11 Conclusions and outlook

The experiments of the present study provide a first insight into the humoral immune response of the ant species *C. floridanus*. The SSH approach together with the Illumina sequencing project delivered a broad overview of the transcriptional changes in response to bacterial challenge. Several factors with homology to immune-relevant genes from other insects as well as so far uncharacterized genes were identified and shown to be differentially regulated after an immune-challenge. A protocol for gene knockdown in worker ants by feeding of dsRNA was elaborated in the context of this work and could be used to further study the exact function and/or interaction of single factors in prospective studies.

The data reported here in combination with the recently published ant genome sequences indicate that the hypothesis of a reduced immune gene repertoire in social insects cannot easily be adopted for ant species (Ratzka et al. 2012a). The genome drafts of *C. floridanus* and *H. saltator* (Bonasio et al. 2010) suggest indeed a comparable low number of genes encoding AMPs. However, this low number may to a certain extent be counteracted by the amplification of hymenoptaecin domains which are encoded as large precursor proteins with multiple bioactive domains. Sequence variations in the mature peptides may also lead to diversification of the immune response. Furthermore, *P. barbatus* even has more antimicrobial peptide genes than *A. mellifera* (Smith et al. 2011b). Detailed analyses of the complete antimicrobial repertoire from different ant species will deliver a better classification of the individual defense capabilities of these social insects (Ratzka et al. 2012a).

Identification of the ant´s immune gene repertoire facilitated first investigations of the immune responses towards the endosymbiont *B. floridanus*. It was shown that *B. floridanus* is still recognized as non-self by the ant´s immune system. Thus, the immune system is likely to prevent colonization of tissues other than the bacteriocytes by *Blochmannia* and may also play a role in the regulation of the number of bacteria present within bacteriocytes. The few other insect-symbiont systems investigated so far indicate highly specific mechanisms leading to symbiont tolerance in dependence of symbiont tissue location and of the proportion of intra- to extracellular phases (Ratzka et al. 2012b). Recently published transcriptomic studies revealed that chronic infection with endosymbionts affects several different cellular functions, such as oxidative stress regulation, immune pathways, apoptosis and autophagy (Kremer et al. 2012; Vigneron et al. 2012). As all these functions are also known to be affected in host-pathogen interactions, it was suggested that a common language exists between bacteria and their hosts. The cellular pathways may be affected differently as a function of the mode of the bacteria-host-interaction as well as according to the bacterial location within the host (Kremer et al. 2012). The transcriptional profiling of symbiosis-relevant candidate genes in the context of this work indicated that amidase PGRPs, especially PGRP-LB, likely play a key role in regulation of the immune response towards endosymbionts in *C. floridanus*, as already shown for other insects (Anselme et al. 2006; Wang et al. 2009; Zaidman-Remy et al. 2006). PGRP-LB might achieve a dual function by on the one hand directly controlling endosymbionts

through its antimicrobial activities (Wang and Aksoy 2012) and on the other hand indirectly through downregulation of the IMD pathway (Zaidman-Remy et al. 2006). Besides, lysosomal degradation of *Blochmannia* might be used as a control mechanism for these endosymbionts and might additionally supply the ant host with essential nutrients, especially during metamorphosis. Further adaptations of the RNAi protocols are needed to possibly achieve gene knockdown in larvae and pupae as well as in the midgut tissue in order to further characterize the role of single factors in endosymbiont-host interactions. Besides, the investigation of gene expression in aposymbiotic ants, which have been cleared from *Blochmannia* endosymbionts with antibiotics, might deliver new insights into the regulation of this symbiosis.

Comparing the data from different model systems reveals that the presence of endosymbionts definitely affects the host immune system and that endosymbionts likely played an important role in the evolution of the insect immune system (Ratzka et al. 2012b). However, so far we still do not have a comprehensive view of the mechanisms leading to the establishment and the maintenance of bacterial endosymbiosis in insects. More detailed and comparative analyses between host-pathogen- and host-symbiont-interactions as well as between ancient and more recent symbioses are needed to elucidate the molecular mechanisms enabling the establishment of symbiosis. Next generation sequencing in combination with RNAi technology provide powerful tools to identify and characterize new symbiosis-relevant genes, even in species, where no reference genome is available. Comparison of such gene sets with data from many other insects with different evolutionary background may help to identify conserved genes with relevance for symbiosis (Ratzka et al. 2012b). A better understanding of the role of endosymbionts in shaping host immune functions will contribute to the development of novel symbiont-based strategies for the control of insect-borne diseases and can lead to new insights how microorganisms interact with the innate immune system.

8 Appendix

8.1 Identification of differentially expressed genes based on the analysis of the Illumina sequencing data by Cufflinks and *DESeq*

Table 16. List of significantly differentially expressed genes based on the analysis of the Illumina sequencing data using Cufflinks and *DESeq*. Table shows significantly down- and upregulated genes 12 h after picking of larvae and workers with a 1:1 mix of Gram-negative and Gram-positive bacteria. Cufflinks analysis resulted in new gene identifiers (XLOCs), which correspond to one or more of the originally annotated genes (EAGs).

XLOC	EAG	Annotation	Cufflinks		DESeq	
			fold change	sig.	fold change	sig.
XLOC_002318	EAG_03738	hypothetical protein	0.0039	yes	0.0040	yes
XLOC_012977	EAG_12392	hypothetical protein	0.0080	yes	0.0059	yes
XLOC_001608	EAG_09527	Lipase member H	0.0116	yes	0.0141	yes
XLOC_006837	EAG_13248	Chymotrypsin-1	0.0138	yes	0.0128	yes
XLOC_004463	EAG_03822	Chymotrypsin-2	0.0144	yes	0.0137	yes
XLOC_004641	EAG_13046	hypothetical protein	0.0268	yes	0.0135	yes
XLOC_011732	EAG_01890	Lipase member H-A	0.0208	no	0.0210	yes
XLOC_007464	EAG_11680	hypothetical protein	0.0264	yes	0.0271	yes
XLOC_004338	EAG_10571	Probable guanine deaminase	0.0374	yes	0.0424	yes
XLOC_010119	EAG_12959	Flexible cuticle protein 12	0.0420	yes	0.0381	yes
XLOC_010628	EAG_14481	hypothetical protein	0.0680	yes	0.0413	yes
XLOC_011629	EAG_04860	hypothetical protein	0.0613	yes	0.0581	yes
XLOC_001153	EAG_11874	Arylphorin subunit alpha	0.0791	yes	0.0744	yes
XLOC_006924	EAG_13262	A disintegrin and metalloproteinase with thrombospondin motifs 1	0.0792	yes	0.0770	yes
XLOC_003974	EAG_05387	Mite allergen Der f 3	0.0915	yes	0.0826	yes
XLOC_005632	EAG_11556	hypothetical protein	0.1412	no	0.0423	yes
XLOC_005577	EAG_04522	hypothetical protein	0.0970	yes	0.0961	yes
XLOC_001444	EAG_08392	Trypsin-1	0.0900	yes	0.1089	yes
XLOC_001444	EAG_08393	Trypsin-3	0.0900	yes	0.1089	yes
XLOC_006201	EAG_02267	Hexamerin-1.1	0.0479	yes	0.1511	yes
XLOC_004681	EAG_08739	Transcription factor MafA	0.0970	yes	0.1025	yes
XLOC_007623	EAG_02852	Carbonic anhydrase 6	0.1028	yes	0.0983	yes
XLOC_006632	EAG_09966	Sugar transporter ERD6-like 6	0.1036	yes	0.0978	yes
XLOC_002942	EAG_06050	Guanine deaminase	0.1051	yes	0.1004	yes
XLOC_000608	EAG_05879	hypothetical protein	0.1064	yes	0.1067	yes
XLOC_014452	EAG_04122	Endocuticle structural glycoprotein SgAbd-1	0.1106	yes	0.1035	yes
XLOC_001271	EAG_06493	hypothetical protein	0.1193	yes	0.1048	yes
XLOC_013466	EAG_06943	hypothetical protein	0.1172	yes	0.1172	yes
XLOC_013195	EAG_04570	-	0.0000	no	0.2359	yes
XLOC_003793	EAG_05602	Acidic mammalian chitinase	0.1319	no	0.1304	yes
XLOC_015102	EAG_03190	-	0.1349	yes	0.1278	yes
XLOC_015297	BX248583.1	5S rRNA	0.2493	no	0.0319	yes
XLOC_006735	EAG_10341	hypothetical protein	0.1513	yes	0.1312	yes
XLOC_007461	EAG_11675	Glycerate kinase	0.1656	no	0.1203	yes
XLOC_005924	EAG_00549	Maltase 1	0.1624	yes	0.1387	yes
XLOC_000609	EAG_05880	hypothetical protein	0.1490	no	0.1534	yes
XLOC_005804	EAG_01127	Cytochrome P450 6k1	0.1585	yes	0.1470	yes
XLOC_014813	EAG_12494	Lysosomal acid phosphatase	0.1716	yes	0.1576	yes
XLOC_012643	EAG_06404	hypothetical protein	0.1734	no	0.1570	yes

XLOC_001828	EAG_16344	Putative trypsin-6	0.1682	yes	0.1649	yes
XLOC_003092	EAG_14000	hypothetical protein	0.1794	yes	0.1543	yes
XLOC_001021	EAG_04466	Probable phenoloxidase subunit CG8193 (POsub)	0.1732	yes	0.1707	yes
XLOC_013084	EAG_07578	Lipase 3	0.1648	yes	0.1800	yes
XLOC_011269	EAG_03375	Guanine deaminase	0.1764	yes	0.1694	yes
XLOC_005567	EAG_04522	hypothetical protein	0.1835	no	0.1648	yes
XLOC_010625	EAG_14478	Maltase 1	0.1762	yes	0.1734	yes
XLOC_002533	EAG_16114	hypothetical protein	0.3132	no	0.0415	yes
XLOC_010627	EAG_14480	Maltase 1	0.1703	yes	0.1882	yes
XLOC_002321	EAG_05013	Lipase 3	0.1808	yes	0.1840	yes
XLOC_007597	EAG_02852	Carbonic anhydrase 6	0.2484	no	0.1185	yes
XLOC_007767	EAG_04819	hypothetical protein	0.2045	yes	0.1693	yes
XLOC_004937	EAG_01471	Alpha-glucosidase	0.1973	yes	0.1875	yes
XLOC_010118	EAG_12957	Sodium-dependent phosphate transport protein 1, chloroplastic	0.1137	no	0.2750	yes
XLOC_010118	EAG_12958	Flexible cuticle protein 12	0.1137	no	0.2750	yes
XLOC_001006	EAG_04456	Fatty acid synthase	0.1940	no	0.2158	yes
XLOC_003976	EAG_05389	Trypsin-7	0.2358	no	0.1881	yes
XLOC_000756	EAG_13292	Chymotrypsin-2	0.1601	no	0.2706	yes
XLOC_012894	EAG_08005	Probable G-protein coupled receptor Mth-like 1	0.1943	yes	0.2418	yes
XLOC_006056	EAG_05756	Sorbitol dehydrogenase	0.2208	no	0.2257	yes
XLOC_009071	EAG_14558	Allantoinase, mitochondrial	0.2256	yes	0.2223	yes
XLOC_001805	EAG_13079	Putative oxidoreductase yrbE	0.2316	no	0.2200	yes
XLOC_004794	EAG_03740	-	0.2686	yes	0.1832	yes
XLOC_003958	EAG_14168	Spermatogenesis-associated protein 17	0.2354	yes	0.2185	yes
XLOC_006289	EAG_08152	Fatty acyl-CoA reductase 1	0.2049	yes	0.2604	no
XLOC_009525	EAG_10596	Organic cation transporter protein	0.2540	yes	0.2130	yes
XLOC_009349	EAG_09173	Cytochrome P450 4C1	0.2455	yes	0.2271	yes
XLOC_004855	EAG_11919	Larval cuticle protein LCP-17	0.2776	no	0.2029	yes
XLOC_012382	EAG_11791	Solute carrier family 2, facilitated glucose transporter member 8	0.2403	no	0.2498	yes
XLOC_005801	EAG_01122	Alkaline phosphatase 4	0.2589	no	0.2432	yes
XLOC_008190	EAG_02446	L-asparaginase	0.2692	yes	0.2331	yes
XLOC_010558	EAG_09662	Guanine deaminase	0.2632	yes	0.2461	yes
XLOC_006030	EAG_15547	Lysosomal alpha-mannosidase	0.2619	no	0.2505	yes
XLOC_000323	EAG_12154	Fatty acyl-CoA reductase 1	0.2595	yes	0.2571	yes
XLOC_013513	EAG_05868	Putative vitellogenin receptor	0.2816	no	0.2396	yes
XLOC_010923	EAG_08920	Lysozyme c-1	0.3197	no	0.2116	yes
XLOC_006881	EAG_13177	Cytochrome P450 6k1	0.2512	no	0.2806	yes
XLOC_007980	EAG_10833	hypothetical protein	0.2599	yes	0.2720	yes
XLOC_003114	EAG_10142	-	0.3494	no	0.1851	yes
XLOC_002679	EAG_12542	hypothetical protein	0.2818	no	0.2605	yes
XLOC_001443	EAG_08391	Trypsin-4	0.2795	yes	0.2633	yes
XLOC_004697	EAG_08730	5'-nucleotidase	0.2751	no	0.2689	yes
XLOC_003975	EAG_05388	Mite allergen Der p 3	0.2156	yes	0.3308	no
XLOC_001089	EAG_14623	Zinc carboxypeptidase A 1	0.2812	no	0.2690	yes
XLOC_006197	EAG_05772	Probable G-protein coupled receptor Mth-like 10	0.2877	no	0.2645	yes
XLOC_013194	EAG_04570	-	0.2797	yes	0.2782	yes
XLOC_006619	EAG_09947	Peritrophin-1	0.2877	no	0.2712	yes
XLOC_009592	EAG_08011	Puromycin-sensitive aminopeptidase	0.2889	yes	0.2741	yes
XLOC_000521	EAG_11108	Cytochrome P450 18a1	0.2837	yes	0.2801	yes

XLOC_001862	EAG_11415	Beta-galactosidase	0.2928	yes	0.2725	yes
XLOC_010498	EAG_09663	hypothetical protein	0.2969	no	0.2757	yes
XLOC_008529	EAG_15588	Probable cytochrome P450 4aa1	0.2592	no	0.3176	yes
XLOC_010206	EAG_04701	Hexokinase-2	0.2969	no	0.2874	yes
XLOC_014241	EAG_14357	hypothetical protein	0.2994	no	0.2925	yes
XLOC_014241	EAG_14358	hypothetical protein	0.2994	no	0.2925	yes
XLOC_006239	EAG_03488	Uricase	0.3242	no	0.2704	yes
XLOC_005161	EAG_14320	-	0.2939	no	0.3183	yes
XLOC_000755	EAG_13290	Chymotrypsin-1	0.0000	no	0.3066	yes
XLOC_000755	EAG_13291	Chymotrypsin-1	0.0000	no	0.3066	yes
XLOC_011919	EAG_03143	hypothetical protein	0.3079	no	0.3061	yes
XLOC_011666	EAG_04562	Synaptic vesicle glycoprotein 2A	0.3337	no	0.2984	yes
XLOC_005473	EAG_11986	hypothetical protein	0.3230	no	0.3105	yes
XLOC_014773	EAG_12252	Alpha-tocopherol transfer protein	0.3467	no	0.2927	yes
XLOC_005230	EAG_13595	Sphingomyelin phosphodiesterase	0.3232	no	0.3211	yes
XLOC_012211	EAG_05732	Cytochrome P450 6k1	0.3264	no	0.3196	yes
XLOC_011681	EAG_16430	hypothetical protein	0.3246	yes	0.3249	yes
XLOC_007049	EAG_01286	hypothetical protein	0.4761	no	0.1805	yes
XLOC_013136	EAG_15708	Probable alpha-ketoglutarate-dependent hypophosphite dioxygenase	0.3298	no	0.3273	yes
XLOC_013136	EAG_15709	MAD2L1-binding protein	0.3298	no	0.3273	yes
XLOC_007113	EAG_08120	hypothetical protein	0.3791	no	0.2789	yes
XLOC_006603	EAG_07155	Sugar transporter ERD6-like 4	0.3535	no	0.3307	yes
XLOC_001841	EAG_16348	hypothetical protein	0.3518	no	0.3338	yes
XLOC_001841	EAG_16350	hypothetical protein	0.3518	no	0.3338	yes
XLOC_000543	EAG_11109	Cytochrome P450 306a1	0.3618	no	0.3271	yes
XLOC_013560	EAG_12717	Trypsin-1	0.3999	no	0.2924	yes
XLOC_013560	EAG_12718	Carboxypeptidase B	0.3999	no	0.2924	yes
XLOC_001442	EAG_08388	Dipeptidase 1	0.3573	yes	0.3406	yes
XLOC_014735	EAG_03763	Prostatic acid phosphatase	0.3613	yes	0.3378	yes
XLOC_009564	EAG_14815	hypothetical protein	0.4622	no	0.2372	yes
XLOC_006727	EAG_10448	Esterase E4	0.3587	no	0.3470	yes
XLOC_000245	EAG_16201	Estradiol 17-beta-dehydrogenase 8	0.3534	no	0.3538	yes
XLOC_000245	EAG_16202	Estradiol 17-beta-dehydrogenase 8	0.3534	no	0.3538	yes
XLOC_004740	EAG_11161	Putative gustatory receptor 43a	0.3655	no	0.3455	yes
XLOC_006882	EAG_13178	Cytochrome P450 6k1	0.3600	no	0.3521	yes
XLOC_001449	EAG_08401	Ecdysteroid UDP-glucosyltransferase	0.3618	no	0.3538	yes
XLOC_009627	EAG_10329	Transient receptor potential channel pyrexia	0.3622	no	0.3633	yes
XLOC_007834	EAG_06225	hypothetical protein	0.3765	no	0.3560	yes
XLOC_011858	EAG_08688	Solute carrier family 2, facilitated glucose transporter member 8	0.3722	no	0.3670	yes
XLOC_010223	EAG_07433	hypothetical protein	0.3984	no	0.3462	yes
XLOC_008889	EAG_05025	Membrane metallo-endopeptidase-like	0.3895	no	0.3591	yes
XLOC_003325	EAG_03997	Serine/threonine-protein kinase RIO2	0.3632	no	0.3854	yes
XLOC_003325	EAG_03998	Sorbitol dehydrogenase	0.3632	no	0.3854	yes
XLOC_014138	EAG_05609	RCC1 and BTB domain-containing protein 1	0.3770	no	0.3721	yes
XLOC_007073	EAG_05260	Adenine phosphoribosyltransferase	0.3822	no	0.3676	yes
XLOC_009489	EAG_10908	Plasma glutamate carboxypeptidase	0.3468	no	0.4062	yes
XLOC_002911	EAG_16321	Membrane alanyl aminopeptidase	0.4317	no	0.3341	yes
XLOC_001833	EAG_16354	Tenascin-X	0.3766	no	0.3898	yes
XLOC_005805	EAG_01128	Probable cytochrome P450 6g2	0.3818	no	0.3848	yes
XLOC_000612	EAG_05884	Vitamin K-dependent γ-carboxylase	0.4086	no	0.3611	yes

XLOC_003764	EAG_06072	hypothetical protein	0.3902	no	0.3803	yes
XLOC_006963	EAG_06163	Probable cytochrome P450 6a13	0.3964	no	0.3787	yes
XLOC_009840	EAG_07237	hypothetical protein	0.4150	no	0.3815	yes
XLOC_009840	EAG_07239	hypothetical protein	0.4150	no	0.3815	yes
XLOC_001848	EAG_11390	Septin-4	0.3969	no	0.4016	yes
XLOC_009057	EAG_14527	hypothetical protein	0.4627	no	0.3408	yes
XLOC_006864	EAG_13139	Calexcitin-2	0.4066	no	0.4005	yes
XLOC_002477	EAG_16147	4-coumarate--CoA ligase 4	0.3947	no	0.4142	yes
XLOC_008926	EAG_13613	Retinol dehydrogenase 11	0.4328	no	0.3786	yes
XLOC_008926	EAG_13614	Retinol dehydrogenase 11	0.4328	no	0.3786	yes
XLOC_008926	EAG_13617	Retinol dehydrogenase 11	0.4328	no	0.3786	yes
XLOC_007662	EAG_01784	Beta-glucuronidase	0.4191	no	0.3929	yes
XLOC_014089	EAG_06520	Kynurenine/alpha-aminoadipate aminotransferase mitochondrial	0.4334	no	0.3820	yes
XLOC_011218	EAG_10546	L-xylulose reductase	0.4247	no	0.3920	yes
XLOC_008545	EAG_15583	Epididymal secretory protein E1	0.4116	no	0.4181	yes
XLOC_006731	EAG_10454	Esterase FE4	0.4398	no	0.3936	yes
XLOC_003728	EAG_06683	hypothetical protein	0.4307	no	0.4031	yes
XLOC_014706	EAG_15098	Solute carrier family 2, facilitated glucose transporter member 8	0.4293	no	0.4128	yes
XLOC_014008	EAG_05288	hypothetical protein	0.4375	no	0.4051	yes
XLOC_004472	EAG_03831	hypothetical protein	0.4404	no	0.4092	yes
XLOC_007501	EAG_06930	Liver carboxylesterase 31	0.4857	no	0.3984	yes
XLOC_001592	EAG_09491	hypothetical protein	0.5119	no	0.3816	yes
XLOC_001592	EAG_09492	hypothetical protein	0.5119	no	0.3816	yes
XLOC_011522	EAG_09647	Suppressor of tumorigenicity protein 14	0.5994	no	0.3898	yes
XLOC_015011	EAG_14498	hypothetical protein	0.6273	no	0.3743	yes
XLOC_015011	EAG_14499	Metallophosphoesterase domain-containing protein 1	0.6273	no	0.3743	yes
XLOC_005568	EAG_04523	hypothetical protein	1.0000	no	0.0335	yes
XLOC_009694	EAG_13572	-	0.6285	no	0.4246	yes
XLOC_004137	EAG_09904	Protein G12	1.0000	no	0.1798	yes
XLOC_001152	EAG_11872	Hexamerin-1.1	1.0000	no	0.1862	yes
XLOC_001152	EAG_11873	Hexamerin	1.0000	no	0.1862	yes
XLOC_006202	EAG_02268	Hexamerin	1.0000	no	0.2010	yes
XLOC_010142	EAG_01178	hypothetical protein	1.0000	no	0.2863	yes
XLOC_014700	EAG_15115	Calsenilin	0.9996	no	0.3848	yes
XLOC_014700	EAG_15116	hypothetical protein	0.9996	no	0.3848	yes
XLOC_014700	EAG_15117	hypothetical protein	0.9996	no	0.3848	yes
XLOC_010140	EAG_01177	hypothetical protein	1.0000	no	0.4262	yes
XLOC_010140	EAG_01179	-	1.0000	no	0.4262	yes
XLOC_013876	EAG_06702	hypothetical protein	1.1610	no	0.3161	yes
XLOC_013876	EAG_06703	Uncharacterized family 31 glucosidase KIAA1161	1.1610	no	0.3161	yes
XLOC_006061	EAG_05757	Lachesin	1.1668	no	0.3114	yes
XLOC_007035	EAG_00269	hypothetical protein	3.2378	yes	0.5521	no
XLOC_014816	EAG_00117	hypothetical protein	3.3441	yes	0.7247	no
XLOC_003360	EAG_15770	-	1.6314	no	2.8368	yes
XLOC_003360	EAG_15771	Hemolymph lipopolysaccharide-binding protein	1.6314	no	2.8368	yes
XLOC_003360	EAG_15772	hypothetical protein	1.6314	no	2.8368	yes
XLOC_009422	EAG_05452	Serine protease inhibitor 3/4	2.0123	no	2.6450	yes
XLOC_015058	EAG_04362	23 kDa integral membrane protein (tetraspanin)	2.4483	no	2.3088	yes

XLOC_013060	EAG_12307	L-lactate dehydrogenase	2.3852	no	2.4132	yes
XLOC_006926	EAG_13264	Ornithine decarboxylase	2.5294	no	2.3008	yes
XLOC_006926	EAG_13266	Vacuolar protein sorting-associated protein 37B	2.5294	no	2.3008	yes
XLOC_012591	EAG_16376	Transient receptor potential channel pyrexia	2.4353	no	2.4586	yes
XLOC_012964	EAG_03980	ATP-binding cassette sub-family A member 13	2.5026	no	2.4128	yes
XLOC_002016	EAG_04385	hypothetical protein	2.5500	no	2.3670	yes
XLOC_008349	EAG_08024	NF-kappa-B inhibitor cactus	2.4859	no	2.4353	yes
XLOC_008349	EAG_08025	NF-kappa-B inhibitor cactus	2.4859	no	2.4353	yes
XLOC_002015	EAG_04384	hypothetical protein	2.5964	no	2.3571	yes
XLOC_002197	EAG_11319	Leucine-rich repeat-containing protein 20	2.4658	no	2.5123	yes
XLOC_008017	EAG_03461	WD repeat-containing protein 68	1.3490	no	3.7091	yes
XLOC_010413	EAG_03909	Lysosomal acid phosphatase	1.3872	no	3.7362	yes
XLOC_012308	EAG_15813	Kynurenine/alpha-aminoadipate aminotransferase mitochondrial	2.6565	no	2.4875	yes
XLOC_014658	EAG_12169	Cytochrome b-c1 complex subunit 8	4.4789	yes	0.7459	no
XLOC_014466	EAG_00563	Probable maltase H	2.6972	yes	2.6274	yes
XLOC_008639	EAG_11037	Tyrocidine synthetase 3	2.7440	no	2.6157	yes
XLOC_003804	EAG_11642	Lysosomal aspartic protease	3.2996	yes	2.0752	no
XLOC_010184	EAG_08524	hypothetical protein	2.7539	no	2.6662	yes
XLOC_013816	EAG_10813	Nuclear factor NF-kappa-B p110 subunit (Relish)	2.9037	no	2.5325	yes
XLOC_009258	EAG_06989	5-aminolevulinate synthase, erythroid-specific, mitochondrial	2.7832	no	2.6542	yes
XLOC_003294	EAG_07546	Apoptosis-inducing factor 3	2.8546	no	2.6364	yes
XLOC_000006	EAG_12212	hypothetical protein	2.8659	no	2.6313	yes
XLOC_000006	EAG_12213	-	2.8659	no	2.6313	yes
XLOC_011353	EAG_11776	hypothetical protein	2.7901	no	2.7172	yes
XLOC_006897	EAG_13206	MIT domain-containing protein 1	2.8423	no	2.6799	yes
XLOC_011327	EAG_14994	Serine protease snake	2.5435	no	3.0053	yes
XLOC_011327	EAG_14995	UPF0195 protein CG30152	2.5435	no	3.0053	yes
XLOC_000964	EAG_07937	Innexin shaking-B	2.9043	no	2.7702	yes
XLOC_000964	EAG_07938	hypothetical protein	2.9043	no	2.7702	yes
XLOC_014972	EAG_00151	hypothetical protein	4.4983	yes	1.1957	no
XLOC_001446	EAG_08397	hypothetical protein	2.9219	no	2.7838	yes
XLOC_002280	EAG_00908	hypothetical protein	2.9255	yes	2.8068	yes
XLOC_012690	EAG_13851	Cell division cycle protein 16-like protein	2.6803	no	3.0629	yes
XLOC_012690	EAG_13852	Cytoplasmic polyadenylation element-binding protein 1-B	2.6803	no	3.0629	yes
XLOC_013413	EAG_10755	hypothetical protein	3.0756	yes	2.7048	yes
XLOC_009395	EAG_16049	Lysosomal-trafficking regulator	2.9860	no	2.8021	yes
XLOC_013972	EAG_08759	Lysosomal aspartic protease	3.0086	no	2.9121	yes
XLOC_001084	EAG_14608	hypothetical protein	3.2215	no	2.7239	yes
XLOC_002500	EAG_03531	Acyl-CoA desaturase	3.0290	no	2.9420	yes
XLOC_001188	EAG_09983	Protein phosphatase 1H	2.4913	no	3.4874	yes
XLOC_000458	EAG_15950	hypothetical protein	5.0441	no	1.0000	yes
XLOC_014342	EAG_06041	hypothetical protein	3.1767	no	2.9903	yes
XLOC_006077	EAG_02845	Lysosomal acid phosphatase	3.5501	no	2.8085	yes
XLOC_007544	EAG_16219	hypothetical protein	5.1907	yes	1.1757	no
XLOC_014860	EAG_02682	Hormone-sensitive lipase	3.2367	yes	3.1649	yes
XLOC_001082	EAG_14606	hypothetical protein	3.4191	yes	3.0011	yes

XLOC_004363	EAG_04718	15-hydroxyprostaglandin dehydrogenase [NAD+]	3.9976	no	2.4692	yes
XLOC_001155	EAG_11877	Succinate-semialdehyde dehydrogenase, mitochondrial	-	-	3.2526	yes
XLOC_002466	EAG_08609	Serine proteinase stubble	3.5269	no	3.3653	yes
XLOC_007871	EAG_09072	J domain-containing protein	3.5486	yes	3.4339	yes
XLOC_002895	EAG_12848	VIP36-like protein (L-type lectin)	5.8512	yes	1.1393	no
XLOC_000863	EAG_00349	Septin-4	5.8533	yes	1.4620	no
XLOC_004142	EAG_09927	Krueppel-like factor 6	6.4460	yes	0.8737	no
XLOC_012699	EAG_13836	hypothetical protein	3.6813	no	3.6689	yes
XLOC_010279	EAG_14638	Vitellogenin	3.8335	no	3.6387	yes
XLOC_000688	EAG_08906	Brain-specific angiogenesis inhibitor 1	4.9461	no	2.7031	yes
XLOC_000688	EAG_08907	hypothetical protein	4.9461	no	2.7031	yes
XLOC_010462	EAG_15426	hypothetical protein	6.6531	yes	1.0000	no
XLOC_006614	EAG_07160	Cytochrome P450 307a1	3.7038	no	3.9881	yes
XLOC_010530	EAG_08319	Calcium-independent phospholipase A2-gamma	3.8592	yes	3.8693	yes
XLOC_013555	EAG_12710	Vitellogenic carboxypeptidase	3.9463	yes	3.8481	yes
XLOC_009245	EAG_09209	ETS-like proteinous factor	3.9320	yes	3.9141	yes
XLOC_000031	EAG_03877	Three prime repair exonuclease 2	3.8265	yes	4.1654	yes
XLOC_003882	EAG_13369	Peroxisomal membrane protein PEX16	4.7039	no	3.3703	yes
XLOC_013559	EAG_12716	Cationic trypsin-3	4.2462	no	3.9202	yes
XLOC_011055	EAG_08243	Bone morphogenetic protein 6	4.3086	no	3.9039	yes
XLOC_009044	EAG_03720	-	3.8996	yes	4.4217	yes
XLOC_000491	EAG_15564	Sodium-independent sulfate anion transporter	4.6516	no	3.8846	yes
XLOC_013431	EAG_07823	Major royal jelly protein 1	4.5848	yes	4.2491	yes
XLOC_010412	EAG_03907	Lipoma-preferred partner-like protein	5.2900	no	3.6175	yes
XLOC_001430	EAG_08466	hypothetical protein	4.6674	no	4.2448	yes
XLOC_003931	EAG_13504	Limulus clotting factor C	4.5306	yes	4.4041	yes
XLOC_010974	EAG_15165	Protein Dom3Z	4.4286	yes	4.5288	yes
XLOC_008774	EAG_07078	Apolipoprotein D	4.9565	no	4.0781	yes
XLOC_008774	EAG_07079	LIM domain kinase 1	4.9565	no	4.0781	yes
XLOC_009073	EAG_14524	Exonuclease GOR	5.1973	yes	3.8416	yes
XLOC_005403	EAG_10054	Leukocyte elastase inhibitor	4.5752	yes	4.4671	yes
XLOC_003239	EAG_00061	-	7.8370	yes	1.2343	no
XLOC_000029	EAG_03877	Three prime repair exonuclease 2	5.1964	no	4.1017	yes
XLOC_001717	EAG_07099	hypothetical protein	8.1632	yes	1.1625	no
XLOC_003540	EAG_14241	Pupal cuticle protein Edg-78E	5.2861	yes	4.0627	no
XLOC_001904	EAG_06589	hypothetical protein	5.0609	yes	4.3814	no
XLOC_000648	EAG_08871	Vanin-like protein 1	4.8329	yes	4.6896	yes
XLOC_000648	EAG_08872	Cytochrome b-c1 complex subunit Rieske, mitochondrial	4.8329	yes	4.6896	yes
XLOC_000881	EAG_14061	hypothetical protein	5.1534	yes	4.4519	yes
XLOC_000881	EAG_14062	hypothetical protein	5.1534	yes	4.4519	yes
XLOC_002586	EAG_13429	hypothetical protein	5.1005	no	4.7362	yes
XLOC_006233	EAG_03480	L-ascorbate oxidase	5.5612	yes	4.7375	yes
XLOC_001210	EAG_09982	Uncharacterized protein C9orf82-like protein	5.2241	yes	5.1940	yes
XLOC_012710	EAG_13861	Transferrin	5.5831	no	5.2797	yes
XLOC_005966	EAG_01672	hypothetical protein	5.7993	yes	5.3220	yes
XLOC_006885	EAG_13181	Putative methyltransferase METT10D	5.9757	yes	5.5092	yes
XLOC_003509	EAG_14847	Serine protease snake	6.0647	yes	5.6451	yes

XLOC_010470	EAG_08838	Transient receptor potential channel pyrexia	5.2099	yes	6.5622	yes
XLOC_007827	EAG_06223	Circadian clock-controlled protein	5.9418	yes	6.6125	yes
XLOC_010400	EAG_03908	Protein Malvolio (mvl)	6.7739	no	5.8646	yes
XLOC_001151	EAG_05215	Peptidoglycan-recognition protein-LB	6.3318	yes	6.4798	yes
XLOC_008013	EAG_03460	Transmembrane protein 205	7.1772	yes	5.6771	yes
XLOC_008592	EAG_11013	Beta-1,3-glucan-binding protein	6.6809	yes	6.3711	yes
XLOC_001909	EAG_06591	Protein SERAC1	7.6040	yes	5.6207	yes
XLOC_011001	EAG_09018	hypothetical protein	7.4239	yes	5.9082	yes
XLOC_004107	EAG_09917	Protein G12	7.5138	no	6.8266	yes
XLOC_012704	EAG_13843	hypothetical protein	8.1714	no	7.1241	yes
XLOC_005978	EAG_14394	hypothetical protein	8.7544	yes	7.6319	yes
XLOC_001189	EAG_09984	hypothetical protein	8.2591	no	8.2803	yes
XLOC_009059	EAG_14530	G-protein coupled receptor Mth2	7.7639	yes	9.0026	yes
XLOC_013834	EAG_10804	hypothetical protein	9.2254	yes	7.8741	yes
XLOC_004129	EAG_09890	Tyrosine 3-monooxygenase (TyrOH)	9.1825	no	8.6824	yes
XLOC_000419	EAG_00033	hypothetical protein	16.1587	yes	1.8339	no
XLOC_006861	EAG_13135	Synaptic vesicle glycoprotein 2B	9.3415	yes	9.0432	yes
XLOC_006960	EAG_06157	Serine protease inhibitor dipetalogastin	11.5522	yes	7.4906	yes
XLOC_011891	EAG_15482	Synaptic vesicle glycoprotein 2B	10.4066	yes	9.7505	yes
XLOC_007169	EAG_02183	hypothetical protein	17.3506	yes	4.2320	no
XLOC_000664	EAG_08866	Fatty acyl-CoA reductase 1	17.7601	yes	6.6503	no
XLOC_010268	EAG_03314	Aromatic-L-amino-acid decarboxylase	12.7527	yes	12.8090	yes
XLOC_001668	EAG_11729	Cytochrome b5	13.2463	yes	12.7245	yes
XLOC_001812	EAG_13089	Parathyroid hormone/parathyroid hormone-related peptide receptor	13.6416	yes	12.7760	yes
XLOC_003910	EAG_00025	hypothetical protein	25.4454	yes	1.8137	no
XLOC_000776	EAG_08564	HIV Tat-specific factor 1-like protein	20.1951	yes	15.9196	yes
XLOC_000776	EAG_08566	HIV Tat-specific factor 1-like protein	20.1951	yes	15.9196	yes
XLOC_007210	EAG_02762	hypothetical protein	34.7724	yes	3.1035	no
XLOC_000592	EAG_08659	hypothetical protein	20.4227	yes	-	-
XLOC_002300	EAG_02637	hypothetical protein	33.7244	yes	8.0408	no
XLOC_003683	EAG_16000	-	27.3025	yes	15.2352	yes
XLOC_010580	EAG_16160	hypothetical protein	21.5901	yes	-	-
XLOC_012789	EAG_01545	Defensin-2	27.1229	yes	24.1543	yes
XLOC_009670	EAG_15280	Esterase FE4	27.2141	yes	26.3745	yes
XLOC_009670	EAG_15281	Esterase FE4	27.2141	yes	26.3745	yes
XLOC_007631	EAG_02878	Aminopeptidase N	16.4583	yes	37.2415	no
XLOC_012711	EAG_13862	WD repeat-containing protein C10orf79	27.7557	yes	32.6977	yes
XLOC_013569	EAG_12981	hypothetical protein	34.7724	yes	-	-
XLOC_004128	EAG_09889	hypothetical protein	35.7537	yes	37.7567	yes
XLOC_013904	EAG_02018	hypothetical protein	32.8657	yes	47.4277	yes
XLOC_000526	EAG_11073	hypothetical protein	43.0667	yes	-	-
XLOC_013895	EAG_02018	hypothetical protein	48.1903	yes	39.3679	yes
XLOC_000188	EAG_04105	Serine proteinase stubble	56.7233	yes	36.3951	yes
XLOC_001784	EAG_05562	hypothetical protein	92.6237	yes	2.1361	no
XLOC_008001	EAG_12674	hypothetical protein	276.2824	yes	213.2921	yes

8.2 Alignment of hymenoptaecin domains

Alignment of hymenoptaecin domains from different ant species was prepared by Frank Förster. The cDNAs of the ants were fragmented according to the cleavage sites predicted by ProP. Afterwards, all cDNA fragments were aligned by translatorX with default settings by muscle and the resulting alignment was cleaned by Gblocks with the default settings from the translatorX website.

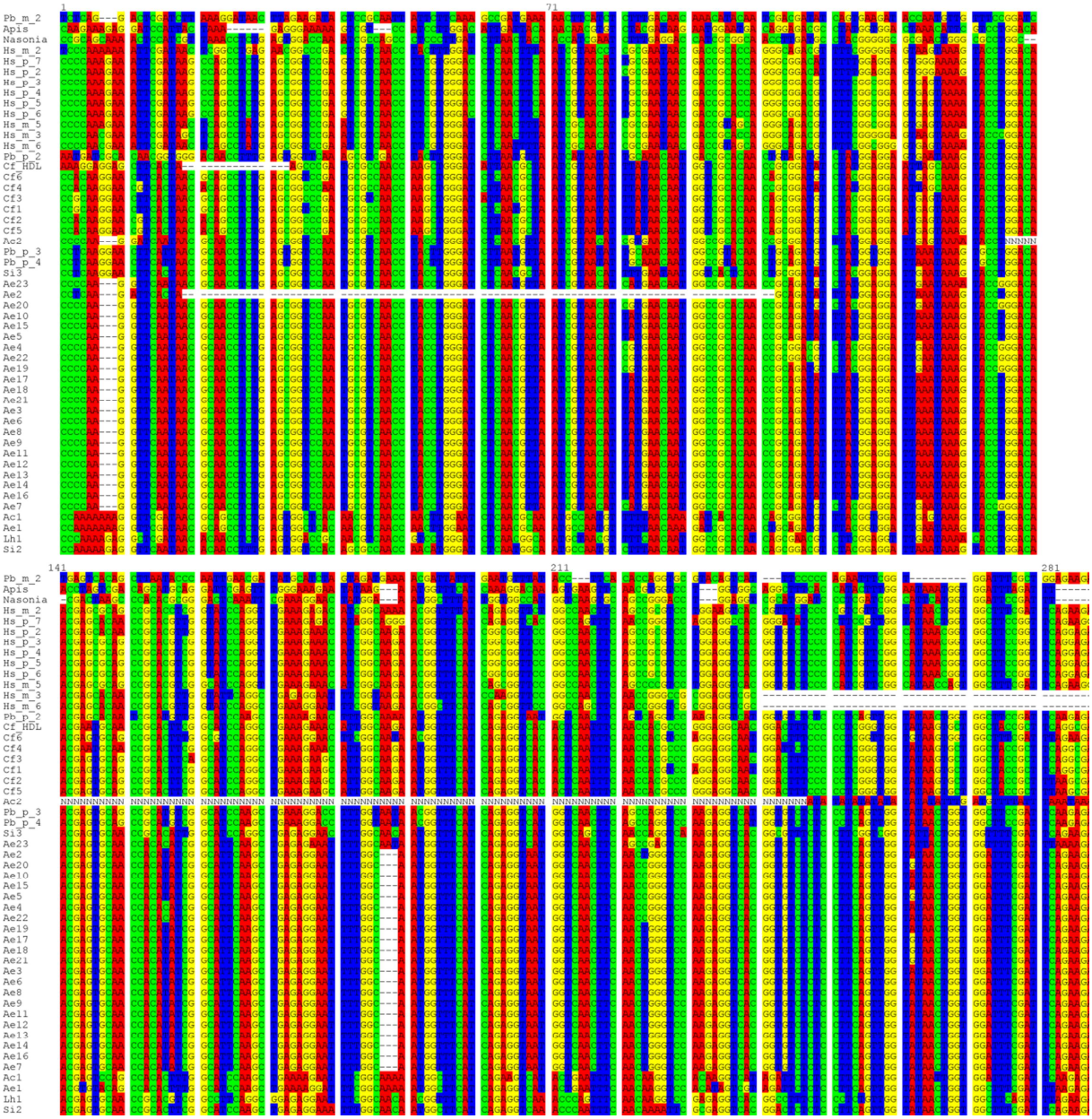

Figure 42. Alignment of hymenoptaecin domains from different ant species (by Frank Förster).

8.3 Alignment of defensin peptides

Alignment of defensin peptides from different ant species was prepared by Frank Förster. The predicted peptide sequences for the seven ant species, the defensin-1 and defensin-2 of *Apis mellifera* (NM_001011616.2, NM_001011638.1), and the defensin of *Ixodes scapularis* (XP_002436104.1) were aligned by muscle with default settings.

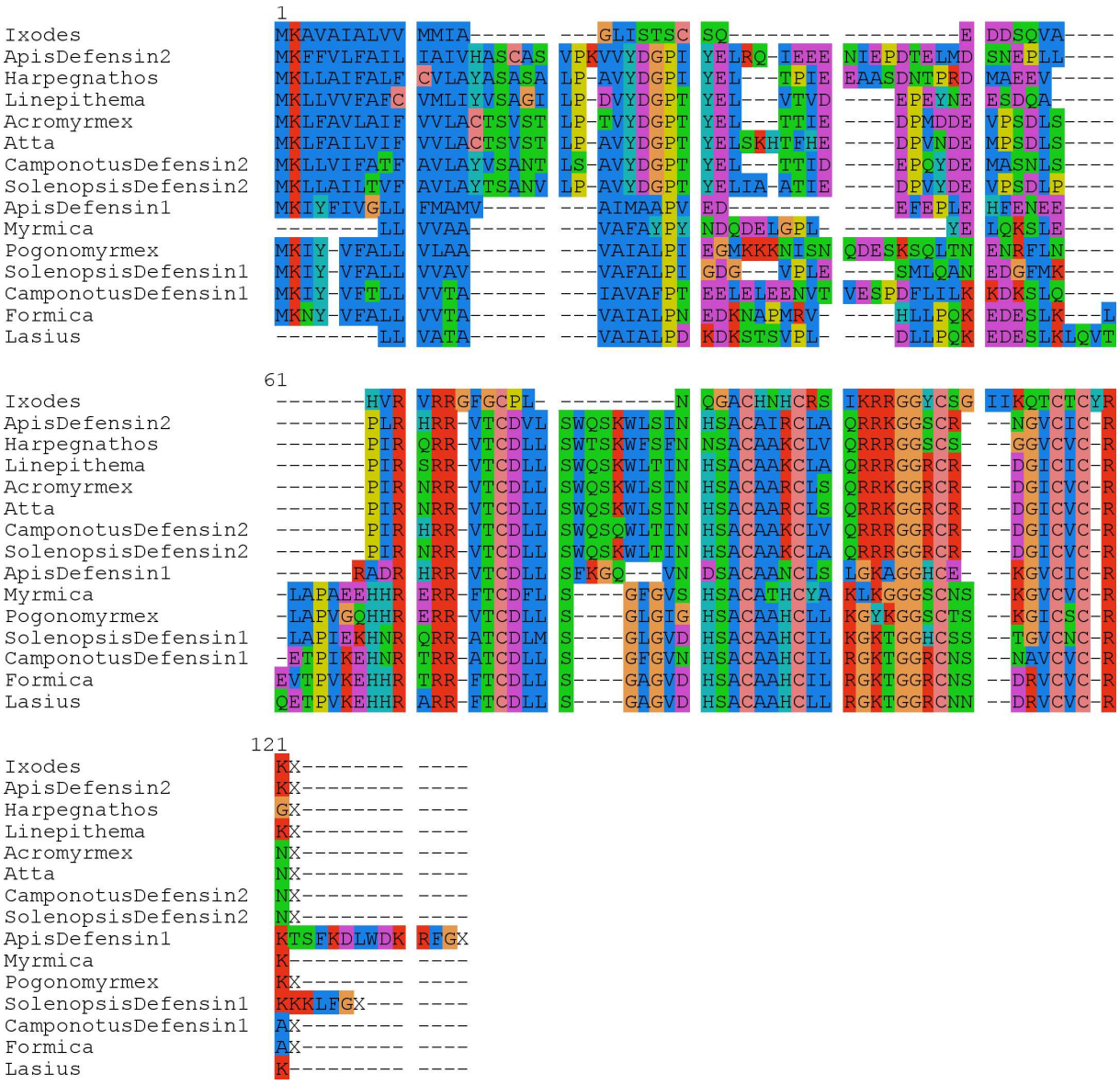

Figure 43. Alignment of defensin peptides from different ant species (by Frank Förster).

8.4 Statistical analysis of gene expression data from the kinetics study

Table 17. Factorial ANOVAs of normalized gene expression levels (ddCt-values) for nine immune genes with time (2, 8, 24 or 48 hours post injection) and bacteria (*M. flavus*- or *S. marcescens*-injection) as categorical predictors (see section 6.6). df: degree of freedom; SS: sum of squares; MS: mean square.

Gene		SS	df	MS	F	p
GNBP	Intercept	278.28	1	278.28	996.39	**0.0000**
	time (t)	75.40	3	25.13	89.99	**0.0000**
	bacteria (b)	3.45	1	3.45	12.34	**0.0018**
	t x b	9.04	3	3.01	10.79	**0.0001**
	error	6.70	24	0.28		
Hsp	Intercept	285.70	1	285.70	751.06	**0.0000**
	time (t)	34.08	3	11.36	29.86	**0.0000**
	bacteria (b)	2.25	1	2.25	5.92	**0.0227**
	t x b	13.64	3	4.55	11.95	**0.0001**
	error	9.13	24	0.38		
Sna	Intercept	434.78	1	434.78	213.18	**0.0000**
	time (t)	142.91	3	47.64	23.36	**0.0000**
	bacteria (b)	3.25	1	3.25	1.59	0.2190
	t x b	23.59	3	7.86	3.85	**0.0220**
	error	48.95	24	2.04		
Spn	Intercept	311.77	1	311.77	853.31	**0.0000**
	time (t)	20.54	3	6.85	18.74	**0.0000**
	bacteria (b)	5.53	1	5.53	15.13	**0.0007**
	t x b	7.41	3	2.47	6.76	**0.0018**
	error	8.77	24	0.37		
Hym	Intercept	2595.44	1	2595.44	490.05	**0.0000**
	time (t)	58.43	3	19.48	3.68	**0.0261**
	bacteria (b)	5.62	1	5.62	1.06	0.3132
	t x b	24.48	3	8.16	1.54	0.2297
	error	127.11	24	5.30		
Def-1	Intercept	1000.98	1	1000.98	466.44	**0.0000**
	time (t)	78.41	3	26.14	12.18	**0.0000**
	bacteria (b)	3.54	1	3.54	1.65	0.2116
	t x b	9.78	3	3.26	1.52	0.2350
	error	51.50	24	2.15		
Stu	Intercept	87.60	1	87.60	359.55	**0.0000**
	time (t)	25.46	3	8.49	34.83	**0.0000**
	bacteria (b)	0.05	1	0.05	0.22	0.6457
	t x b	0.92	3	0.31	1.26	0.3118
	error	5.85	24	0.24		
TyrOH	Intercept	634.72	1	634.72	428.96	**0.0000**
	time (t)	22.38	3	7.46	5.04	**0.0075**
	bacteria (b)	0.95	1	0.95	0.64	0.4301
	t x b	12.51	3	4.17	2.82	0.0605
	error	35.51	24	1.48		
Mvl	Intercept	208.76	1	208.76	535.70	**0.0000**
	time (t)	30.44	3	10.15	26.04	**0.0000**
	bacteria (b)	0.80	1	0.80	2.06	0.1639
	t x b	1.69	3	0.56	1.44	0.2546
	error	9.35	24	0.39		

Table 18. Tukey´s HSD post hoc test of normalized gene expression levels (ddCt-values) for nine immune genes with the fixed factors time (2, 8, 24 or 48 hours post injection (hpi)) and bacteria (*M. flavus* (*Mf*)- or *S. marcescens* (*Sm*)-injection) (see section 6.6).

Gene	factors	2h *Mf*	2h *Sm*	8h *Mf*	8h *Sm*	24h *Mf*	24h *Sm*	48h *Mf*	48h *Sm*
GNBP	2h *Mf*		0.5510	**0.0001**	**0.0001**	**0.0001**	**0.0002**	**0.0002**	0.3482
	2h *Sm*	0.5510		**0.0001**	**0.0001**	**0.0001**	**0.0026**	**0.0016**	1.0000
	8h *Mf*	**0.0001**	**0.0001**		1.0000	0.9991	**0.0007**	**0.0011**	**0.0001**
	8h *Sm*	**0.0001**	**0.0001**	1.0000		0.9971	**0.0006**	**0.0009**	**0.0001**
	24h *Mf*	**0.0001**	**0.0001**	0.9991	0.9971		**0.0026**	**0.0042**	**0.0001**
	24h *Sm*	**0.0002**	**0.0026**	**0.0007**	**0.0006**	**0.0026**		1.0000	**0.0059**
	48h *Mf*	**0.0002**	**0.0016**	**0.0011**	**0.0009**	**0.0042**	1.0000		**0.0036**
	48h *Sm*	0.3482	1.0000	**0.0001**	**0.0001**	**0.0001**	**0.0059**	**0.0036**	
Gene	factors	2h *Mf*	2h *Sm*	8h *Mf*	8h *Sm*	24h *Mf*	24h *Sm*	48h *Mf*	48h *Sm*
Hsp	2h *Mf*		0.6054	**0.0002**	**0.0001**	**0.0002**	**0.0043**	**0.0004**	0.9999
	2h *Sm*	0.6054		**0.0014**	**0.0003**	**0.0088**	0.2235	**0.0207**	0.3625
	8h *Mf*	**0.0002**	**0.0014**		0.9925	0.9925	0.3263	0.9376	**0.0002**
	8h *Sm*	**0.0001**	**0.0003**	0.9925		0.7640	0.0784	0.5407	**0.0001**
	24h *Mf*	**0.0002**	**0.0088**	0.9925	0.7640		0.7835	0.9999	**0.0002**
	24h *Sm*	**0.0043**	0.2235	0.3263	0.0784	0.7835		0.9351	**0.0017**
	48h *Mf*	**0.0004**	**0.0207**	0.9376	0.5407	0.9999	0.9351		**0.0002**
	48h *Sm*	0.9999	0.3625	**0.0002**	**0.0001**	**0.0002**	**0.0017**	**0.0002**	
Gene	factors	2h *Mf*	2h *Sm*	8h *Mf*	8h *Sm*	24h *Mf*	24h *Sm*	48h *Mf*	48h *Sm*
Sna	2h *Mf*		0.9754	**0.0007**	**0.0004**	**0.0003**	**0.0004**	**0.0002**	0.0815
	2h *Sm*	0.9754		**0.0061**	**0.0028**	**0.0015**	**0.0033**	**0.0003**	0.4334
	8h *Mf*	**0.0007**	**0.0061**		1.0000	0.9985	1.0000	0.8716	0.4328
	8h *Sm*	**0.0004**	**0.0028**	1.0000		1.0000	1.0000	0.9667	0.2663
	24h *Mf*	**0.0003**	**0.0015**	0.9985	1.0000		1.0000	0.9947	0.1635
	24h *Sm*	**0.0004**	**0.0033**	1.0000	1.0000	1.0000		0.9526	0.2985
	48h *Mf*	**0.0002**	**0.0003**	0.8716	0.9667	0.9947	0.9526		**0.0360**
	48h *Sm*	0.0815	0.4334	0.4328	0.2663	0.1635	0.2985	**0.0360**	
Gene	factors	2h *Mf*	2h *Sm*	8h *Mf*	8h *Sm*	24h *Mf*	24h *Sm*	48h *Mf*	48h *Sm*
Spn	2h *Mf*		1.0000	**0.0023**	**0.0018**	0.0931	1.0000	0.2236	0.2978
	2h *Sm*	1.0000		**0.0052**	**0.0040**	0.1835	0.9973	0.3899	0.1620
	8h *Mf*	**0.0023**	**0.0052**		1.0000	0.7230	**0.0011**	0.4387	**0.0002**
	8h *Sm*	**0.0018**	**0.0040**	1.0000		0.6547	**0.0009**	0.3749	**0.0002**
	24h *Mf*	0.0931	0.1835	0.7230	0.6547		**0.0487**	0.9997	**0.0005**
	24h *Sm*	1.0000	0.9973	**0.0011**	**0.0009**	**0.0487**		0.1269	0.4640
	48h *Mf*	0.2236	0.3899	0.4387	0.3749	0.9997	0.1269		**0.0012**
	48h *Sm*	0.2978	0.1620	**0.0002**	**0.0002**	**0.0005**	0.4640	**0.0012**	
Gene	factors	2h *Mf*	2h *Sm*	8h *Mf*	8h *Sm*	24h *Mf*	24h *Sm*	48h *Mf*	48h *Sm*
Hym	2h *Mf*		1.0000	0.2712	0.9255	0.9729	0.4177	0.9405	0.9996
	2h *Sm*	1.0000		0.1476	0.7780	0.8791	0.2467	0.8062	1.0000
	8h *Mf*	0.2712	0.1476		0.9125	0.8249	1.0000	0.8936	0.1105
	8h *Sm*	0.9255	0.7780	0.9125		1.0000	0.9776	1.0000	0.6919
	24h *Mf*	0.9729	0.8791	0.8249	1.0000		0.9353	1.0000	0.8110
	24h *Sm*	0.4177	0.2467	1.0000	0.9776	0.9353		0.9698	0.1903
	48h *Mf*	0.9405	0.8062	0.8936	1.0000	1.0000	0.9698		0.7237
	48h *Sm*	0.9996	1.0000	0.1105	0.6919	0.8110	0.1903	0.7237	

Gene	factors	2h *Mf*	2h *Sm*	8h *Mf*	8h *Sm*	24h *Mf*	24h *Sm*	48h *Mf*	48h *Sm*
Def-1	2h *Mf*		0.7231	0.0633	0.2019	**0.0157**	**0.0003**	**0.0044**	**0.0078**
	2h *Sm*	0.7231		0.7667	0.9750	0.3910	**0.0118**	0.1621	0.2469
	8h *Mf*	0.0633	0.7667		0.9986	0.9981	0.2900	0.9341	0.9803
	8h *Sm*	0.2019	0.9750	0.9986		0.9167	0.0987	0.6499	0.7885
	24h *Mf*	**0.0157**	0.3910	0.9981	0.9167		0.6494	0.9993	1.0000
	24h *Sm*	**0.0003**	**0.0118**	0.2900	0.0987	0.6494		0.9164	0.8183
	48h *Mf*	**0.0044**	0.1621	0.9341	0.6499	0.9993	0.9164		1.0000
	48h *Sm*	**0.0078**	0.2469	0.9803	0.7885	1.0000	0.8183	1.0000	
Gene	factors	2h *Mf*	2h *Sm*	8h *Mf*	8h *Sm*	24h *Mf*	24h *Sm*	48h *Mf*	48h *Sm*
Stu	2h *Mf*		0.8695	**0.0003**	**0.0002**	**0.0252**	**0.0050**	1.0000	0.9171
	2h *Sm*	0.8695		**0.0041**	**0.0027**	0.3480	0.1033	0.8938	0.2136
	8h *Mf*	**0.0003**	**0.0041**		1.0000	0.4256	0.8307	**0.0003**	**0.0002**
	8h *Sm*	**0.0002**	**0.0027**	1.0000		0.3325	0.7396	**0.0002**	**0.0002**
	24h *Mf*	**0.0252**	0.3480	0.4256	0.3325		0.9965	**0.0289**	**0.0015**
	24h *Sm*	**0.0050**	0.1033	0.8307	0.7396	0.9965		**0.0058**	**0.0004**
	48h *Mf*	1.0000	0.8938	**0.0003**	**0.0002**	**0.0289**	**0.0058**		0.8959
	48h *Sm*	0.9171	0.2136	**0.0002**	**0.0002**	**0.0015**	**0.0004**	0.8959	
Gene	factors	2h *Mf*	2h *Sm*	8h *Mf*	8h *Sm*	24h *Mf*	24h *Sm*	48h *Mf*	48h *Sm*
TyrOH	2h *Mf*		0.9894	0.1978	**0.0249**	0.2294	0.7114	**0.0409**	0.9894
	2h *Sm*	0.9894		0.6342	0.1425	0.6864	0.9896	0.2137	1.0000
	8h *Mf*	0.1978	0.6342		0.9683	1.0000	0.9762	0.9923	0.6345
	8h *Sm*	**0.0249**	0.1425	0.9683		0.9516	0.5210	1.0000	0.1426
	24h *Mf*	0.2294	0.6864	1.0000	0.9516		0.9859	0.9859	0.6867
	24h *Sm*	0.7114	0.9896	0.9762	0.5210	0.9859		0.6593	0.9897
	48h *Mf*	**0.0409**	0.2137	0.9923	1.0000	0.9859	0.6593		0.2138
	48h *Sm*	0.9894	1.0000	0.6345	0.1426	0.6867	0.9897	0.2138	
Gene	factors	2h *Mf*	2h *Sm*	8h *Mf*	8h *Sm*	24h *Mf*	24h *Sm*	48h *Mf*	48h *Sm*
Mvl	2h *Mf*		0.6829	**0.0054**	**0.0004**	**0.0002**	**0.0001**	**0.0011**	**0.0118**
	2h *Sm*	0.6829		0.2117	**0.0172**	**0.0016**	**0.0003**	0.0547	0.3622
	8h *Mf*	**0.0054**	0.2117		0.9207	0.3753	0.0636	0.9965	1.0000
	8h *Sm*	**0.0004**	**0.0172**	0.9207		0.9687	0.5094	0.9994	0.7755
	24h *Mf*	**0.0002**	**0.0016**	0.3753	0.9687		0.9732	0.7851	0.2209
	24h *Sm*	**0.0001**	**0.0003**	0.0636	0.5094	0.9732		0.2391	**0.0307**
	48h *Mf*	**0.0011**	0.0547	0.9965	0.9994	0.7851	0.2391		0.9656
	48h *Sm*	**0.0118**	0.3622	1.0000	0.7755	0.2209	**0.0307**	0.9656	

8.5 Statistical analysis of the gene expression data in response to injection of *B. floridanus*

Table 19. Factorial ANOVAs of normalized gene expression levels for nine immune genes with time (8 or 24 hours post injection) and bacteria (ant Ringer-, *M. flavus*- , *S. marcescens*- or *B. floridanus*-injection) as categorical predictors (see section 6.7). df: degree of freedom; SS: sum of squares; MS: mean square.

Gene		SS	df	MS	F	p
GNBP	Intercept	325.42	1	325.42	1089.22	**0.0000**
	time (t)	31.66	1	31.66	105.97	**0.0000**
	bacteria (b)	57.60	3	19.20	64.27	**0.0000**
	t x b	9.58	3	3.19	10.69	**0.0001**
	error	7.17	24	0.30		
Hsp	Intercept	252.03	1	252.03	697.49	**0.0000**
	time (t)	10.37	1	10.37	28.69	**0.0000**
	bacteria (b)	50.44	3	16.81	46.53	**0.0000**
	t x b	3.56	3	1.19	3.29	**0.0380**
	error	8.67	24	0.36		
Sna	Intercept	501.81	1	501.81	242.81	**0.0000**
	time (t)	0.05	1	0.05	0.03	0.8719
	bacteria (b)	79.23	3	26.41	12.78	**0.0000**
	t x b	2.62	3	0.87	0.42	0.7382
	error	49.60	24	2.07		
Spn	Intercept	231.89	1	231.89	400.00	**0.0000**
	time (t)	15.36	1	15.36	26.50	**0.0000**
	bacteria (b)	31.09	3	10.36	17.88	**0.0000**
	t x b	8.43	3	2.81	4.85	**0.0089**
	error	13.91	24	0.58		
Hym	Intercept	2662.05	1	2662.05	564.30	**0.0000**
	time (t)	31.95	1	31.95	6.77	**0.0156**
	bacteria (b)	77.05	3	25.68	5.44	**0.0053**
	t x b	41.76	3	13.92	2.95	0.0530
	error	113.22	24	4.72		
Def-1	Intercept	1237.96	1	1237.96	830.91	**0.0000**
	time (t)	0.56	1	0.56	0.38	0.5454
	bacteria (b)	14.74	3	4.91	3.30	**0.0376**
	t x b	27.42	3	9.14	6.14	**0.0030**
	error	35.76	24	1.49		
Stu	Intercept	100.28	1	100.28	442.79	**0.0000**
	time (t)	2.34	1	2.34	10.35	**0.0037**
	bacteria (b)	21.89	3	7.30	32.22	**0.0000**
	t x b	0.31	3	0.10	0.45	0.7172
	error	5.44	24	0.23		
TyrOH	Intercept	492.48	1	492.48	296.27	**0.0000**
	time (t)	1.96	1	1.96	1.18	0.2881
	bacteria (b)	62.58	3	20.86	12.55	**0.0000**
	t x b	12.07	3	4.02	2.42	0.0908
	error	39.89	24	1.66		
Mvl	Intercept	227.97	1	227.97	777.14	**0.0000**
	time (t)	0.00	1	0.00	0.00	0.9749
	bacteria (b)	25.12	3	8.37	28.54	**0.0000**
	t x b	8.45	3	2.82	9.60	**0.0002**
	error	7.04	24	0.29		

Table 20. Tukey´s HSD post hoc test of normalized gene expression levels (ddCt-values) for nine immune genes with the fixed factors time (8 or 24 hours post injection (hpi)) and bacteria (ant Ringer (aR)-, *M. flavus* (*Mf*)- , *S. marcescens* (*Sm*)- or *B. floridanus* (*Bf*)-injection) (see section 6.7).

Gene	factors	8h aR	8h *Mf*	8h *Sm*	8h *Bf*	24h aR	24h *Mf*	24h *Sm*	24h *Bf*
GNBP	8h aR		**0.0002**	**0.0002**	**0.0038**	**0.0001**	**0.0005**	0.9827	0.0647
	8h *Mf*	**0.0002**		1.0000	0.8212	**0.0001**	0.9993	**0.0011**	**0.0001**
	8h *Sm*	**0.0002**	1.0000		0.7666	**0.0001**	0.9977	**0.0008**	**0.0001**
	8h *Bf*	**0.0038**	0.8212	0.7666		**0.0001**	0.9817	**0.0308**	**0.0001**
	24h aR	**0.0001**	**0.0001**	**0.0001**	**0.0001**		**0.0001**	**0.0001**	**0.0079**
	24h *Mf*	**0.0005**	0.9993	0.9977	0.9817	**0.0001**		**0.0037**	**0.0001**
	24h *Sm*	0.9827	**0.0011**	**0.0008**	**0.0308**	**0.0001**	**0.0037**		**0.0084**
	24h *Bf*	0.0647	**0.0001**	**0.0001**	**0.0001**	**0.0079**	**0.0001**	**0.0084**	
Gene	factors	8h aR	8h *Mf*	8h *Sm*	8h *Bf*	24h aR	24h *Mf*	24h *Sm*	24h *Bf*
Hsp	8h aR		**0.0007**	**0.0002**	0.5437	**0.0011**	**0.0039**	0.1297	1.0000
	8h *Mf*	**0.0007**		0.9912	0.0526	**0.0001**	0.9913	0.2972	**0.0009**
	8h *Sm*	**0.0002**	0.9912		**0.0085**	**0.0001**	0.7411	0.0663	**0.0002**
	8h *Bf*	0.5437	0.0526	**0.0085**		**0.0002**	0.2482	0.9816	0.6511
	24h aR	**0.0011**	**0.0001**	**0.0001**	**0.0002**		**0.0001**	**0.0001**	**0.0008**
	24h *Mf*	**0.0039**	0.9913	0.7411	0.2482	**0.0001**		0.7618	**0.0059**
	24h *Sm*	0.1297	0.2972	0.0663	0.9816	**0.0001**	0.7618		0.1790
	24h *Bf*	1.0000	**0.0009**	**0.0002**	0.6511	**0.0008**	**0.0059**	0.1790	
Gene	factors	8h aR	8h *Mf*	8h *Sm*	8h *Bf*	24h aR	24h *Mf*	24h *Sm*	24h *Bf*
Sna	8h aR		0.0816	**0.0410**	0.3469	0.9929	**0.0220**	**0.0478**	0.1304
	8h *Mf*	0.0816		1.0000	0.9912	**0.0148**	0.9985	1.0000	1.0000
	8h *Sm*	**0.0410**	1.0000		0.9434	**0.0069**	1.0000	1.0000	0.9991
	8h *Bf*	0.3469	0.9912	0.9434		0.0867	0.8446	0.9593	0.9990
	24h aR	0.9929	**0.0148**	**0.0069**	0.0867		**0.0036**	**0.0082**	**0.0253**
	24h *Mf*	**0.0220**	0.9985	1.0000	0.8446	**0.0036**		1.0000	0.9888
	24h *Sm*	**0.0478**	1.0000	1.0000	0.9593	**0.0082**	1.0000		0.9996
	24h *Bf*	0.1304	1.0000	0.9991	0.9990	**0.0253**	0.9888	0.9996	
Gene	factors	8h aR	8h *Mf*	8h *Sm*	8h *Bf*	24h aR	24h *Mf*	24h *Sm*	24h *Bf*
Spn	8h aR		**0.0018**	**0.0015**	0.7415	0.2825	0.0822	0.9894	0.9590
	8h *Mf*	**0.0018**		1.0000	0.0719	**0.0001**	0.7037	**0.0003**	**0.0218**
	8h *Sm*	**0.0015**	1.0000		0.0595	**0.0001**	0.6490	**0.0003**	**0.0177**
	8h *Bf*	0.7415	0.0719	0.0595		**0.0102**	0.8160	0.2684	0.9992
	24h aR	0.2825	**0.0001**	**0.0001**	**0.0102**		**0.0004**	0.7594	**0.0356**
	24h *Mf*	0.0822	0.7037	0.6490	0.8160	**0.0004**		**0.0131**	0.4932
	24h *Sm*	0.9894	**0.0003**	**0.0003**	0.2684	0.7594	**0.0131**		0.5685
	24h *Bf*	0.9590	**0.0218**	**0.0177**	0.9992	**0.0356**	0.4932	0.5685	
Gene	factors	8h aR	8h *Mf*	8h *Sm*	8h *Bf*	24h aR	24h *Mf*	24h *Sm*	24h *Bf*
Hym	8h aR		0.6997	0.9999	0.9704	0.0700	1.0000	0.8605	0.9994
	8h *Mf*	0.6997		0.8861	0.9968	**0.0015**	0.7807	1.0000	0.3862
	8h *Sm*	0.9999	0.8861		0.9981	**0.0315**	1.0000	0.9694	0.9843
	8h *Bf*	0.9704	0.9968	0.9981		**0.0074**	0.9875	0.9999	0.7907
	24h aR	0.0700	**0.0015**	**0.0315**	**0.0074**		0.0518	**0.0031**	0.1924
	24h *Mf*	1.0000	0.7807	1.0000	0.9875	0.0518		0.9145	0.9973
	24h *Sm*	0.8605	1.0000	0.9694	0.9999	**0.0031**	0.9145		0.5678
	24h *Bf*	0.9994	0.3862	0.9843	0.7907	0.1924	0.9973	0.5678	

Gene	factors	8h aR	8h *Mf*	8h *Sm*	8h *Bf*	24h aR	24h *Mf*	24h *Sm*	24h *Bf*
Def-1	8h aR		0.9980	0.8716	0.7494	0.4600	1.0000	0.3660	0.9992
	8h *Mf*	0.9980		0.9957	0.3723	0.8306	0.9940	0.1242	0.9287
	8h *Sm*	0.8716	0.9957		0.1081	0.9954	0.8158	**0.0275**	0.5642
	8h *Bf*	0.7494	0.3723	0.1081		**0.0229**	0.8146	0.9978	0.9629
	24h aR	0.4600	0.8306	0.9954	**0.0229**		0.3900	**0.0051**	0.1960
	24h *Mf*	1.0000	0.9940	0.8158	0.8146	0.3900		0.4340	0.9998
	24h *Sm*	0.3660	0.1242	**0.0275**	0.9978	**0.0051**	0.4340		0.6966
	24h *Bf*	0.9992	0.9287	0.5642	0.9629	0.1960	0.9998	0.6966	
Gene	factors	8h aR	8h *Mf*	8h *Sm*	8h *Bf*	24h aR	24h *Mf*	24h *Sm*	24h *Bf*
Stu	8h aR		**0.0002**	**0.0002**	**0.0156**	0.9975	**0.0052**	**0.0010**	0.5442
	8h *Mf*	**0.0002**		1.0000	0.1788	**0.0001**	0.3813	0.8043	**0.0027**
	8h *Sm*	**0.0002**	1.0000		0.1280	**0.0001**	0.2913	0.7047	**0.0018**
	8h *Bf*	**0.0156**	0.1788	0.1280		**0.0034**	0.9997	0.9293	0.5631
	24h aR	0.9975	**0.0001**	**0.0001**	**0.0034**		**0.0012**	**0.0003**	0.2094
	24h *Mf*	**0.0052**	0.3813	0.2913	0.9997	**0.0012**		0.9956	0.3010
	24h *Sm*	**0.0010**	0.8043	0.7047	0.9293	**0.0003**	0.9956		0.0809
	24h *Bf*	0.5442	**0.0027**	**0.0018**	0.5631	0.2094	0.3010	0.0809	
Gene	factors	8h aR	8h *Mf*	8h *Sm*	8h *Bf*	24h aR	24h *Mf*	24h *Sm*	24h *Bf*
TyrOH	8h aR		0.1134	**0.0143**	0.9568	0.6924	0.1326	0.4953	0.2309
	8h *Mf*	0.1134		0.9769	0.5985	**0.0025**	1.0000	0.9827	0.9999
	8h *Sm*	**0.0143**	0.9769		0.1424	**0.0004**	0.9642	0.5900	0.8741
	8h *Bf*	0.9568	0.5985	0.1424		0.1520	0.6484	0.9786	0.8216
	24h aR	0.6924	**0.0025**	**0.0004**	0.1520		**0.0031**	**0.0210**	**0.0063**
	24h *Mf*	0.1326	1.0000	0.9642	0.6484	**0.0031**		0.9899	1.0000
	24h *Sm*	0.4953	0.9827	0.5900	0.9786	**0.0210**	0.9899		0.9995
	24h *Bf*	0.2309	0.9999	0.8741	0.8216	**0.0063**	1.0000	0.9995	
Gene	factors	8h aR	8h *Mf*	8h *Sm*	8h *Bf*	24h aR	24h *Mf*	24h *Sm*	24h *Bf*
Mvl	8h aR		0.7748	0.1073	0.0958	**0.0099**	**0.0082**	**0.0006**	0.4336
	8h *Mf*	0.7748		0.8502	0.8236	**0.0003**	0.2195	**0.0217**	0.9990
	8h *Sm*	0.1073	0.8502		1.0000	**0.0002**	0.9350	0.3379	0.9896
	8h *Bf*	0.0958	0.8236	1.0000		**0.0002**	0.9494	0.3673	0.9847
	24h aR	**0.0099**	**0.0003**	**0.0002**	**0.0002**		**0.0001**	**0.0001**	**0.0002**
	24h *Mf*	**0.0082**	0.2195	0.9350	0.9494	**0.0001**		0.9436	0.5079
	24h *Sm*	**0.0006**	**0.0217**	0.3379	0.3673	**0.0001**	0.9436		0.0751
	24h *Bf*	0.4336	0.9990	0.9896	0.9847	**0.0002**	0.5079	0.0751	

8.6 Statistical analysis of the immune gene expression data according to developmental stage and tissue

Table 21. Factorial ANOVAs of relative gene expression levels (dCt-values) for nine immune genes with stage (L2, P3 or W2) and tissue (midgut or residual body) as categorical predictors (see section 6.8.2). df: degree of freedom; SS: sum of squares; MS: mean square.

Gene		SS	df	MS	F	p
PGRP-LB	Intercept	2234.86	1	2234.86	3240.56	**0.0000**
	tissue	175.85	1	175.85	254.99	**0.0000**
	stage	51.29	2	25.64	37.18	**0.0000**
	tissue x stage	137.09	2	68.54	99.39	**0.0000**
	error	20.69	30	0.69		
PGRP-SC2	Intercept	834.43	1	834.43	2086.44	**0.0000**
	tissue	1.73	1	1.73	4.32	**0.0462**
	stage	32.88	2	16.44	41.11	**0.0000**
	tissue x stage	106.76	2	53.38	133.47	**0.0000**
	error	12.00	30	0.40		
PGRP-LE	Intercept	1896.43	1	1896.43	6436.59	**0.0000**
	tissue	3.72	1	3.72	12.63	**0.0013**
	stage	0.83	2	0.41	1.41	0.2610
	tissue x stage	3.29	2	1.64	5.58	**0.0087**
	error	8.84	30	0.29		
PGRP-2	Intercept	5241.68	1	5241.68	13091.24	**0.0000**
	tissue	201.15	1	201.15	502.38	**0.0000**
	stage	31.00	2	15.50	38.71	**0.0000**
	tissue x stage	12.97	2	6.49	16.20	**0.0000**
	error	12.01	30	0.40		
GNBP	Intercept	5367.84	1	5367.84	4598.60	**0.0000**
	tissue	117.97	1	117.97	101.06	**0.0000**
	stage	15.28	2	7.64	6.55	**0.0044**
	tissue x stage	16.48	2	8.24	7.06	**0.0031**
	error	35.02	30	1.17		
Rel	Intercept	1047.43	1	1047.43	7293.30	**0.0000**
	tissue	2.36	1	2.36	16.44	**0.0003**
	stage	7.37	2	3.69	25.66	**0.0000**
	tissue x stage	3.86	2	1.93	13.46	**0.0001**
	error	4.31	30	0.14		
Hym	Intercept	3772.73	1	3772.73	2045.65	**0.0000**
	tissue	58.10	1	58.10	31.50	**0.0000**
	stage	111.74	2	55.87	30.30	**0.0000**
	tissue x stage	47.75	2	23.88	12.95	**0.0001**
	error	60.86	33	1.84		
Def-1	Intercept	491.71	1	491.71	209.92	**0.0000**
	tissue	6.51	1	6.51	2.78	0.1050
	stage	107.12	2	53.56	22.86	**0.0000**
	tissue x stage	20.22	2	10.11	4.32	**0.0216**
	error	77.30	33	2.34		
Tollip	Intercept	3370.19	1	3370.19	1532.90	**0.0000**
	tissue	0.85	1	0.85	0.38	0.5397
	stage	3.87	2	1.93	0.88	0.4254
	tissue x stage	2.85	2	1.42	0.65	0.5306
	error	65.96	30	2.20		

Table 22. Tukey´s HSD post hoc test of relative gene expression levels (dCt-values) for nine immune genes with the fixed factors stage (L2, P3 or W2) and tissue (midgut (m) or residual body (b)) (see section 6.8.2).

Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
PGRP-LB	m L2		**0.0005**	**0.0017**	1.0000	**0.0001**	**0.0233**
	m P3	**0.0005**		0.9961	**0.0005**	**0.0001**	**0.0001**
	m W2	**0.0017**	0.9961		**0.0015**	**0.0001**	**0.0001**
	b L2	1.0000	**0.0005**	**0.0015**		**0.0001**	**0.0261**
	b P3	**0.0001**	**0.0001**	**0.0001**	**0.0001**		**0.0001**
	b W2	**0.0233**	**0.0001**	**0.0001**	**0.0261**	**0.0001**	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
PGRP-SC2	m L2		**0.0001**	**0.0176**	**0.0008**	0.9956	**0.0001**
	m P3	**0.0001**		**0.0001**	**0.0001**	**0.0001**	0.0746
	m W2	**0.0176**	**0.0001**		0.8104	**0.0049**	**0.0001**
	b L2	**0.0008**	**0.0001**	0.8104		**0.0003**	**0.0001**
	b P3	0.9956	**0.0001**	**0.0049**	**0.0003**		**0.0001**
	b W2	**0.0001**	0.0746	**0.0001**	**0.0001**	**0.0001**	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
PGRP-LE	m L2		0.8077	0.9698	**0.0007**	0.5733	0.4957
	m P3	0.8077		0.9967	**0.0167**	0.9986	0.9949
	m W2	0.9698	0.9967		**0.0050**	0.9494	0.9140
	b L2	**0.0007**	**0.0167**	**0.0050**		**0.0432**	0.0573
	b P3	0.5733	0.9986	0.9494	**0.0432**		1.0000
	b W2	0.4957	0.9949	0.9140	0.0573	1.0000	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
PGRP-2	m L2		0.9490	0.1197	**0.0001**	**0.0001**	**0.0001**
	m P3	0.9490		0.5066	**0.0001**	**0.0001**	**0.0001**
	m W2	0.1197	0.5066		**0.0001**	**0.0001**	**0.0001**
	b L2	**0.0001**	**0.0001**	**0.0001**		0.1220	**0.0001**
	b P3	**0.0001**	**0.0001**	**0.0001**	0.1220		**0.0001**
	b W2	**0.0001**	**0.0001**	**0.0001**	**0.0001**	**0.0001**	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
GNBP	m L2		0.3104	0.5907	**0.0074**	0.1166	**0.0001**
	m P3	0.3104		0.9964	**0.0002**	**0.0007**	**0.0001**
	m W2	0.5907	0.9964		**0.0002**	**0.0024**	**0.0001**
	b L2	**0.0074**	**0.0002**	**0.0002**		0.8419	**0.0231**
	b P3	0.1166	**0.0007**	**0.0024**	0.8419		**0.0012**
	b W2	**0.0001**	**0.0001**	**0.0001**	**0.0231**	**0.0012**	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
Rel	m L2		**0.0050**	0.9669	**0.0001**	**0.0032**	0.9949
	m P3	**0.0050**		**0.0007**	0.1296	1.0000	**0.0013**
	m W2	0.9669	**0.0007**		**0.0001**	**0.0005**	0.9998
	b L2	**0.0001**	0.1296	**0.0001**		0.1818	**0.0001**
	b P3	**0.0032**	1.0000	**0.0005**	0.1818		**0.0009**
	b W2	0.9949	**0.0013**	0.9998	**0.0001**	**0.0009**	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
Hym	m L2		0.0942	1.0000	0.8782	0.7201	**0.0001**
	m P3	0.0942		0.1603	**0.0156**	0.8278	**0.0001**
	m W2	1.0000	0.1603		0.9057	0.8012	**0.0001**
	b L2	0.8782	**0.0156**	0.9057		0.2261	**0.0001**
	b P3	0.7201	0.8278	0.8012	0.2261		**0.0001**
	b W2	**0.0001**	**0.0001**	**0.0001**	**0.0001**	**0.0001**	

Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
Def-1	m L2		**0.0040**	0.3927	0.4774	**0.0001**	0.9808
	m P3	**0.0040**		0.3249	0.2567	0.2183	**0.0093**
	m W2	0.3927	0.3249		1.0000	**0.0015**	0.7063
	b L2	0.4774	0.2567	1.0000		**0.0011**	0.7954
	b P3	**0.0001**	0.2183	**0.0015**	**0.0011**		**0.0001**
	b W2	0.9808	**0.0093**	0.7063	0.7954	**0.0001**	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
Tollip	m L2		0.7214	0.9993	0.7982	0.8520	0.9986
	m P3	0.7214		0.8924	1.0000	0.9999	0.9114
	m W2	0.9993	0.8924		0.9373	0.9630	1.0000
	b L2	0.7982	1.0000	0.9373		1.0000	0.9507
	b P3	0.8520	0.9999	0.9630	1.0000		0.9722
	b W2	0.9986	0.9114	1.0000	0.9507	0.9722	

8.7 Statistical analysis of the expression data from genes involved in the lysosomal system, in autophagy and in ROS regulation

Table 23. Factorial ANOVAs of relative gene expression levels (dCt-values) for selected for genes involved in the lysosomal system, autophagy and ROS regulation with stage (L2, P3 or W2) and tissue (midgut or residual body) as categorical predictors (see section 6.8.4). df: degree of freedom; SS: sum of squares; MS: mean square.

Gene		SS	df	MS	F	p
C-type lyso	Intercept	961.93	1	961.93	3128.60	**0.0000**
	tissue	1.71	1	1.71	5.57	**0.0250**
	stage	49.98	2	24.99	81.27	**0.0000**
	tissue x stage	31.74	2	15.87	51.61	**0.0000**
	error	9.22	30	0.31		
I-type lyso	Intercept	808.96	1	808.96	2392.34	**0.0000**
	tissue	77.56	1	77.56	229.36	**0.0000**
	stage	35.68	2	17.84	52.76	**0.0000**
	tissue x stage	88.10	2	44.05	130.27	**0.0000**
	error	10.14	30	0.34		
Cath L	Intercept	12.14	1	12.14	55.06	**0.0000**
	tissue	0.02	1	0.02	0.11	0.7463
	stage	21.25	2	10.63	48.22	**0.0000**
	tissue x stage	7.74	2	3.87	17.55	**0.0000**
	error	6.61	30	0.22		
Lap	Intercept	33.82	1	33.82	239.26	**0.0000**
	tissue	1.51	1	1.51	10.68	**0.0027**
	stage	50.27	2	25.13	177.82	**0.0000**
	tissue x stage	20.91	2	10.45	73.96	**0.0000**
	error	4.24	30	0.14		
CPVL	Intercept	410.23	1	410.23	1537.30	**0.0000**
	tissue	21.41	1	21.41	80.25	**0.0000**
	stage	1.67	2	0.84	3.13	0.0581
	tissue x stage	7.73	2	3.87	14.49	**0.0000**
	error	8.01	30	0.27		
Apg-5	Intercept	1345.93	1	1345.93	11602.80	**0.0000**
	tissue	1.10	1	1.10	9.49	**0.0044**
	stage	2.18	2	1.09	9.41	**0.0007**
	tissue x stage	3.05	2	1.52	13.13	**0.0001**
	error	3.48	30	0.12		
Ferritin	Intercept	0.08	1	0.08	0.34	0.5642
	tissue	11.91	1	11.91	47.95	**0.0000**
	stage	9.23	2	4.62	18.58	**0.0000**
	tissue x stage	18.60	2	9.30	37.43	**0.0000**
	error	7.45	30	0.25		
Duox	Intercept	2909.17	1	2909.17	10133.45	**0.0000**
	tissue	262.88	1	262.88	915.68	**0.0000**
	stage	59.87	2	29.93	104.27	**0.0000**
	tissue x stage	6.86	2	3.43	11.96	**0.0002**
	error	8.61	30	0.29		
SOD	Intercept	99.41	1	99.41	571.72	**0.0000**
	tissue	0.82	1	0.82	4.73	**0.0377**
	stage	11.14	2	5.57	32.03	**0.0000**
	tissue x stage	0.01	2	0.00	0.03	0.9734
	error	5.22	30	0.17		

Table 24. Tukey´s HSD post hoc test of relative gene expression levels (dCt-values) for genes involved in the lysosomal system, autophagy and ROS regulation with the fixed factors stage (L2, P3 or W2) and tissue (midgut (m) or residual body (b)) (see section 6.8.4).

Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
C-type lyso	m L2		**0.0001**	1.0000	**0.0001**	**0.0001**	**0.0358**
	m P3	**0.0001**		**0.0001**	0.7627	0.2399	**0.0001**
	m W2	1.0000	**0.0001**		**0.0001**	**0.0002**	**0.0284**
	b L2	**0.0001**	0.7627	**0.0001**		**0.0131**	**0.0001**
	b P3	**0.0001**	0.2399	**0.0002**	**0.0131**		**0.0001**
	b W2	**0.0358**	**0.0001**	**0.0284**	**0.0001**	**0.0001**	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
I-type Lyso	m L2		**0.0001**	**0.0030**	**0.0001**	**0.0001**	**0.0001**
	m P3	**0.0001**		**0.0001**	0.9734	**0.0024**	1.0000
	m W2	**0.0030**	**0.0001**		**0.0001**	**0.0001**	**0.0001**
	b L2	**0.0001**	0.9734	**0.0001**		**0.0161**	0.9902
	b P3	**0.0001**	**0.0024**	**0.0001**	**0.0161**		**0.0035**
	b W2	**0.0001**	1.0000	**0.0001**	0.9902	**0.0035**	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
Cath L	m L2		**0.0063**	**0.0004**	**0.0130**	0.9907	**0.0145**
	m P3	**0.0063**		**0.0001**	0.9997	**0.0014**	**0.0001**
	m W2	**0.0004**	**0.0001**		**0.0001**	**0.0016**	0.6993
	b L2	**0.0130**	0.9997	**0.0001**		**0.0029**	**0.0001**
	b P3	0.9907	**0.0014**	**0.0016**	**0.0029**		0.0589
	b W2	**0.0145**	**0.0001**	0.6993	**0.0001**	0.0589	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
Lap	m L2		**0.0001**	**0.0001**	**0.0001**	**0.0001**	0.9943
	m P3	**0.0001**		**0.0001**	**0.0001**	0.0523	**0.0001**
	m W2	**0.0001**	**0.0001**		0.3139	**0.0008**	**0.0001**
	b L2	**0.0001**	**0.0001**	0.3139		**0.0001**	**0.0001**
	b P3	**0.0001**	0.0523	**0.0008**	**0.0001**		**0.0001**
	b W2	0.9943	**0.0001**	**0.0001**	**0.0001**	**0.0001**	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
CPVL	m L2		0.1625	0.9918	**0.0002**	**0.0001**	0.5561
	m P3	0.1625		**0.0483**	**0.0001**	**0.0001**	**0.0032**
	m W2	0.9918	**0.0483**		**0.0003**	**0.0002**	0.8775
	b L2	**0.0002**	**0.0001**	**0.0003**		0.9988	**0.0037**
	b P3	**0.0001**	**0.0001**	**0.0002**	0.9988		**0.0014**
	b W2	0.5561	**0.0032**	0.8775	**0.0037**	**0.0014**	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
Apg-5	m L2		**0.0392**	0.9999	0.2345	0.6368	**0.0409**
	m P3	**0.0392**		**0.0237**	0.9507	**0.0008**	**0.0001**
	m W2	0.9999	**0.0237**		0.1592	0.7628	0.0661
	b L2	0.2345	0.9507	0.1592		**0.0074**	**0.0002**
	b P3	0.6368	**0.0008**	0.7628	**0.0074**		0.6169
	b W2	**0.0409**	**0.0001**	0.0661	**0.0002**	0.6169	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
Ferritin	m L2		**0.0001**	**0.0003**	**0.0001**	**0.0001**	**0.0016**
	m P3	**0.0001**		0.0822	0.0552	0.5113	**0.0171**
	m W2	**0.0003**	0.0822		**0.0002**	**0.0011**	0.9831
	b L2	**0.0001**	0.0552	**0.0002**		0.8098	**0.0001**
	b P3	**0.0001**	0.5113	**0.0011**	0.8098		**0.0003**
	b W2	**0.0016**	**0.0171**	0.9831	**0.0001**	**0.0003**	

Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
Duox	m L2		**0.0001**	**0.0001**	**0.0001**	**0.0001**	**0.0001**
	m P3	**0.0001**		**0.0015**	**0.0001**	**0.0001**	**0.0001**
	m W2	**0.0001**	**0.0015**		**0.0001**	**0.0001**	**0.0001**
	b L2	**0.0001**	**0.0001**	**0.0001**		**0.0001**	0.9999
	b P3	**0.0001**	**0.0001**	**0.0001**	**0.0001**		**0.0001**
	b W2	**0.0001**	**0.0001**	**0.0001**	0.9999	**0.0001**	
Gene	factors	mL2	mP3	mW2	bL2	bP3	bW2
SOD	m L2		0.7526	**0.0042**	0.7008	0.1558	0.0722
	m P3	0.7526		**0.0002**	1.0000	0.8578	**0.0026**
	m W2	**0.0042**	**0.0002**		**0.0002**	**0.0001**	0.8452
	b L2	0.7008	1.0000	**0.0002**		0.8941	**0.0021**
	b P3	0.1558	0.8578	**0.0001**	0.8941		**0.0002**
	b W2	0.0722	**0.0026**	0.8452	**0.0021**	**0.0002**	

8.8 Statistical analysis of the gene expression data in pupae in response to an immune-challenge

Table 25. Factorial ANOVAs of relative gene expression levels (dCt-values) in pupae for selected genes with tissue (midgut or residual body) and immune status (immune-challenged and untreated) as categorical predictors (see section 6.9). df: degree of freedom; SS: sum of squares; MS: mean square.

Gene		SS	df	MS	F	p
PGRP-LB	Intercept	1161.82	1	1161.82	1648.97	**0.0000**
	tissue	342.86	1	342.86	486.62	**0.0000**
	immune	39.18	1	39.18	55.60	**0.0000**
	tissue x immune	12.33	1	12.33	17.49	**0.0005**
	error	14.09	20	0.70		
PGRP-SC2	Intercept	448.01	1	448.01	1752.00	**0.0000**
	tissue	67.10	1	67.10	262.41	**0.0000**
	immune	0.06	1	0.06	0.22	0.6478
	tissue x immune	0.34	1	0.34	1.32	0.2635
	error	5.11	20	0.26		
PGRP-LE	Intercept	589.00	1	589.00	6863.26	**0.0000**
	tissue	1.01	1	1.01	11.73	**0.0027**
	immune	0.47	1	0.47	5.53	**0.0291**
	tissue x immune	0.01	1	0.01	0.16	0.6947
	error	1.72	20	0.09		
PGRP-2	Intercept	953.34	1	953.34	1957.44	**0.0000**
	tissue	42.35	1	42.35	86.96	**0.0000**
	immune	48.34	1	48.34	99.26	**0.0000**
	tissue x immune	0.00	1	0.00	0.00	0.9628
	error	9.74	20	0.49		
GNBP	Intercept	2301.76	1	2301.76	1420.95	**0.0000**
	tissue	44.93	1	44.93	27.73	**0.0000**
	immune	126.29	1	126.29	77.96	**0.0000**
	tissue x immune	0.02	1	0.02	0.01	0.9203
	error	32.40	20	1.62		
Rel	Intercept	250.18	1	250.18	68.71	**0.0000**
	tissue	2.90	1	2.90	0.80	0.3824
	immune	33.08	1	33.08	9.09	**0.0069**
	tissue x immune	7.46	1	7.46	2.05	0.1679
	error	72.82	20	3.64		
Hym	Intercept	867.11	1	867.11	1072.01	**0.0000**
	tissue	16.63	1	16.63	20.56	**0.0002**
	immune	317.63	1	317.63	392.69	**0.0000**
	tissue x immune	5.48	1	5.48	6.78	**0.0170**
	error	16.18	20	0.81		
Def-1	Intercept	269.84	1	269.84	149.96	**0.0000**
	tissue	17.74	1	17.74	9.86	**0.0052**
	immune	118.00	1	118.00	65.58	**0.0000**
	tissue x immune	27.74	1	27.74	15.42	**0.0008**
	error	35.99	20	1.80		
I-type lyso	Intercept	343.70	1	343.70	1832.25	**0.0000**
	tissue	4.45	1	4.45	23.71	**0.0001**
	immune	0.01	1	0.01	0.06	0.8067
	tissue x immune	0.08	1	0.08	0.44	0.5151
	error	3.75	20	0.19		

Table 26. Tukey´s HSD post hoc test of relative gene expression levels (dCt-values) in pupae for selected genes with the fixed factors tissue (midgut (m) or residual body (b)) and immune status (immune-challenged (i) and untreated control (c)) (see section 6.9).

Gene	factors	mc	mi	bc	bi
PGRP-LB	m c		0.1281	**0.0002**	**0.0002**
	m i	0.1281		**0.0002**	**0.0002**
	b c	**0.0002**	**0.0002**		**0.0002**
	b i	**0.0002**	**0.0002**	**0.0002**	
Gene	factors	mc	mi	bc	bi
PGRP-SC2	m c		0.9615	**0.0002**	**0.0002**
	m i	0.9615		**0.0002**	**0.0002**
	b c	**0.0002**	**0.0002**		0.6691
	b i	**0.0002**	**0.0002**	0.6691	
Gene	factors	mc	mi	bc	bi
PGRP-LE	m c		0.5254	0.0608	0.8717
	m i	0.5254		**0.0031**	0.1749
	b c	0.0608	**0.0031**		0.2423
	b i	0.8717	0.1749	0.2423	
Gene	factors	mc	mi	bc	bi
PGRP-2	m c		**0.0002**	**0.0002**	**0.0002**
	m i	**0.0002**		0.9688	**0.0002**
	b c	**0.0002**	0.9688		**0.0002**
	b i	**0.0002**	**0.0002**	**0.0002**	
Gene	factors	mc	mi	bc	bi
GNBP	m c		**0.0002**	**0.0059**	**0.0002**
	m i	**0.0002**		0.0873	**0.0080**
	b c	**0.0059**	0.0873		**0.0002**
	b i	**0.0002**	**0.0080**	**0.0002**	
Gene	factors	mc	mi	bc	bi
Rel	m c		0.6821	0.9808	0.0539
	m i	0.6821		0.4562	0.3784
	b c	0.9808	0.4562		**0.0244**
	b i	0.0539	0.3784	**0.0244**	
Gene	factors	mc	mi	bc	bi
Hym	m c		**0.0002**	0.5342	**0.0002**
	m i	**0.0002**		**0.0002**	**0.0005**
	b c	0.5342	**0.0002**		**0.0002**
	b i	**0.0002**	**0.0005**	**0.0002**	
Gene	factors	mc	mi	bc	bi
Def-1	m c		**0.0367**	**0.0005**	**0.0111**
	m i	**0.0367**		**0.0002**	0.9437
	b c	**0.0005**	**0.0002**		**0.0002**
	b i	**0.0111**	0.9437	**0.0002**	
Gene	factors	mc	mi	bc	bi
I-type lyso	m c		0.9165	**0.0045**	**0.0087**
	m i	0.9165		**0.0186**	**0.0348**
	b c	**0.0045**	**0.0186**		0.9910
	b i	**0.0087**	**0.0348**	0.9910	

8.9 Statistical analysis of the gene expression data after feeding of dsRNA during an immune response

Table 27. Factorial ANOVAs of relative gene expression levels (dCt-values) in worker midgut tissue for selected genes with treatment as categorical predictor (see section 6.10.2). df: degree of freedom; SS: sum of squares; MS: mean square.

Gene		SS	df	MS	F	p
PGRP-LB	Intercept	302.39	1	302.39	1061.95	**0.0000**
	treatment	0.55	3	0.18	0.64	0.5952
	error	5.69	20	0.28		
PGRP-SC2	Intercept	660.38	1	660.38	3476.21	**0.0000**
	treatment	0.16	3	0.05	0.29	0.8343
	error	3.80	20	0.19		
PGRP-2	Intercept	1131.62	1	1131.62	1960.39	**0.0000**
	treatment	9.08	3	3.03	5.24	**0.0078**
	error	11.54	20	0.58		
Hym	Intercept	1619.03	1	1619.03	398.44	**0.0000**
	treatment	44.71	3	14.90	3.67	**0.0296**
	error	81.27	20	4.06		
Def-1	Intercept	27.16	1	27.16	25.48	**0.0001**
	treatment	15.95	3	5.32	4.99	**0.0096**
	error	21.32	20	1.07		
Def-2	Intercept	3400.32	1	3400.32	2251.02	**0.0000**
	treatment	3.07	3	1.02	0.68	0.5761
	error	30.21	20	1.51		

Table 28. Tukey´s HSD post hoc test of relative gene expression levels (dCt-values) in worker midgut tissue for selected genes with the following fixed factors: (H) untreated animals fed with pure saccharose solution, (I) immune-challenged animals fed with pure saccharose solution, (G) immune-challenged animals fed with saccharose solution containing 2 µg/µl dsGFP, (L) immune-challenged animals fed with saccharose solution containing 2 µg/µl dsPGRP-LB (see section 6.10.2).

Gene	factor	H	I	G	L
PGRP-LB	H		0.5491	0.7884	0.7882
	I	0.5491		0.9763	0.9764
	G	0.7884	0.9763		1.0000
	L	0.7882	0.9764	1.0000	
Gene	factor	H	I	G	L
PGRP-SC2	H		0.8515	0.8873	0.8817
	I	0.8515		0.9998	0.9999
	G	0.8873	0.9998		1.0000
	L	0.8817	0.9999	1.0000	
Gene	factor	H	I	G	L
PGRP-2	H		0.1694	**0.0403**	**0.0059**
	I	0.1694		0.8769	0.3853
	G	**0.0403**	0.8769		0.8136
	L	**0.0059**	0.3853	0.8136	
Gene	factor	H	I	G	L
Hym	H		0.1654	0.1087	**0.0239**
	I	0.1654		0.9955	0.7622
	G	0.1087	0.9955		0.8764
	L	**0.0239**	0.7622	0.8764	
Gene	factor	H	I	G	L
Def-1	H		0.1024	**0.0426**	**0.0079**
	I	0.1024		0.9704	0.6192
	G	**0.0426**	0.9704		0.8617
	L	**0.0079**	0.6192	0.8617	
Gene	factor	H	I	G	L
Def-2	H		0.7425	0.6504	0.9969
	I	0.7425		0.9986	0.8492
	G	0.6504	0.9986		0.7694
	L	0.9969	0.8492	0.7694	

Table 29. Factorial ANOVAs of relative gene expression levels (dCt-values) in worker body tissue for selected genes with treatment as categorical predictor (see section 6.10.2). df: degree of freedom; SS: sum of squares; MS: mean square.

Gene		SS	df	MS	F	p
PGRP-LB	Intercept	1138.66	1	1138.66	1374.02	**0.0000**
	treatment	64.57	3	21.52	25.97	**0.0000**
	error	16.57	20	0.83		
PGRP-SC2	Intercept	52.97	1	52.97	420.45	**0.0000**
	treatment	2.16	3	0.72	5.72	**0.0054**
	error	2.52	20	0.13		
PGRP-2	Intercept	110.53	1	110.53	720.72	**0.0000**
	treatment	2.01	3	0.67	4.36	**0.0162**
	error	3.07	20	0.15		
Hym	Intercept	55.57	1	55.57	30.14	**0.0000**
	treatment	74.67	3	24.89	13.50	**0.0000**
	error	36.87	20	1.84		
Def-1	Intercept	360.81	1	360.81	233.22	**0.0000**
	treatment	31.71	3	10.57	6.83	**0.0024**
	error	30.94	20	1.55		
Def-2	Intercept	581.57	1	581.57	1105.30	**0.0000**
	treatment	0.29	3	0.10	0.19	0.9049
	error	10.52	20	0.53		

Table 30. Tukey´s HSD post hoc test of relative gene expression levels (dCt-values) in worker body tissue for selected genes with the following fixed factors: (H) untreated animals fed with pure saccharose solution, (I) immune-challenged animals fed with pure saccharose solution, (G) immune-challenged animals fed with saccharose solution containing 2 µg/µl dsGFP, (L) immune-challenged animals fed with saccharose solution containing 2 µg/µl dsPGRP-LB (see section 6.10.2).

Gene	factor	H	I	G	L
PGRP-LB	H		**0.0346**	**0.0314**	**0.0010**
	I	**0.0346**		1.0000	**0.0002**
	G	**0.0314**	1.0000		**0.0002**
	L	**0.0010**	**0.0002**	**0.0002**	
Gene	factor	H	I	G	L
PGRP-SC2	H		**0.0434**	**0.0041**	**0.0491**
	I	**0.0434**		0.6994	0.9999
	G	**0.0041**	0.6994		0.6642
	L	**0.0491**	0.9999	0.6642	
Gene	factor	H	I	G	L
PGRP-2	H		0.2187	0.0829	**0.0113**
	I	0.2187		0.9485	0.4614
	G	0.0829	0.9485		0.7778
	L	**0.0113**	0.4614	0.7778	
Gene	factor	H	I	G	L
Hym	H		**0.0000**	**0.0001**	**0.0000**
	I	**0.0000**		0.6616	0.8387
	G	**0.0001**	0.6616		0.5227
	L	**0.0000**	0.8387	0.5227	
Gene	factor	H	I	G	L
Def-1	H		**0.0055**	**0.0103**	**0.0071**
	I	**0.0055**		0.9919	0.9995
	G	**0.0103**	0.9919		0.9984
	L	**0.0071**	0.9995	0.9984	
Gene	factor	H	I	G	L
Def-2	H		0.9543	0.9935	0.9979
	I	0.9543		0.9942	0.8972
	G	0.9935	0.9942		0.9702
	L	0.9979	0.8972	0.9702	

Table 31. Factorial ANOVAs of relative gene expression levels (dCt-values) in whole larvae for selected genes with treatment as categorical predictor (see section 6.10.2). df: degree of freedom; SS: sum of squares; MS: mean square.

Gene		SS	df	MS	F	p
PGRP-LB	Intercept	538.20	1	538.20	588.26	**0.0000**
	treatment	33.06	3	11.02	12.05	**0.0001**
	error	18.30	20	0.91		
PGRP-SC2	Intercept	304.11	1	304.11	635.98	**0.0000**
	treatment	11.95	3	3.98	8.33	**0.0009**
	error	9.56	20	0.48		
PGRP-2	Intercept	521.29	1	521.29	574.18	**0.0000**
	treatment	22.42	3	7.47	8.23	**0.0009**
	error	18.16	20	0.91		
Hym	Intercept	398.83	1	398.83	126.70	**0.0000**
	treatment	158.70	3	52.90	16.80	**0.0000**
	error	62.96	20	3.15		
Def-1	Intercept	0.18	1	0.18	0.17	0.6829
	treatment	11.69	3	3.90	3.76	**0.0273**
	error	20.73	20	1.04		
Def-2	Intercept	1765.49	1	1765.49	692.53	**0.0000**
	treatment	5.68	3	1.89	0.74	0.5392
	error	50.99	20	2.55		

Table 32. Tukey´s HSD post hoc test of relative gene expression levels (dCt-values) in whole larvae for selected genes with the following fixed factors: (H) untreated animals fed with pure saccharose solution, (I) immune-challenged animals fed with pure saccharose solution, (G) immune-challenged animals fed with saccharose solution containing 2 µg/µl dsGFP, (L) immune-challenged animals fed with saccharose solution containing 2 µg/µl dsPGRP-LB (see section 6.10.2).

Gene	factor	H	I	G	L
PGRP-LB	H		**0.0008**	**0.0004**	**0.0007**
	I	**0.0008**		0.9851	0.9997
	G	**0.0004**	0.9851		0.9946
	L	**0.0007**	0.9997	0.9946	
Gene	factor	H	I	G	L
PGRP-SC2	H		**0.0038**	**0.0102**	**0.0013**
	I	**0.0038**		0.9678	0.9608
	G	**0.0102**	0.9678		0.7819
	L	**0.0013**	0.9608	0.7819	
Gene	factor	H	I	G	L
PGRP-2	H		**0.0071**	**0.0021**	**0.0029**
	I	**0.0071**		0.9433	0.9760
	G	**0.0021**	0.9433		0.9989
	L	**0.0029**	0.9760	0.9989	
Gene	factor	H	I	G	L
Hym	H		**0.0002**	**0.0002**	**0.0003**
	I	**0.0002**		0.9956	0.9928
	G	**0.0002**	0.9956		0.9575
	L	**0.0003**	0.9928	0.9575	
Gene	factor	H	I	G	L
Def-1	H		0.3432	**0.0246**	0.0772
	I	0.3432		0.4986	0.8205
	G	**0.0246**	0.4986		0.9437
	L	0.0772	0.8205	0.9437	
Gene	factor	H	I	G	L
Def-2	H		0.9577	0.9059	0.4743
	I	0.9577		0.9982	0.7699
	G	0.9059	0.9982		0.8562
	L	0.4743	0.7699	0.8562	

9 References

Abascal, F, Zardoya, R and Telford, MJ (2010) TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res* **38**: W7-13.

Agaisse, H and Perrimon, N (2004) The roles of JAK/STAT signaling in *Drosophila* immune responses. *Immunol Rev* **198**: 72-82.

Aggrawal, K and Silverman, N (2007) Peptidoglycan recognition in *Drosophila*. *Biochem Soc T* **35**: 1496-1500.

Aguilar, R, Jedlicka, AE, Mintz, M, Mahairaki, V, Scott, AL and Dimopoulos, G (2005) Global gene expression analysis of *Anopheles gambiae* responses to microbial challenge. *Insect Biochem Mol Biol* **35**: 709-719.

Akman, L, Yamashita, A, Watanabe, H, Oshima, K, Shiba, T, Hattori, M and Aksoy, S (2002) Genome sequence of the endocellular obligate symbiont of tsetse flies, *Wigglesworthia glossinidia*. *Nat Gen* **32**: 402-407.

Aksoy, S (1995) *Wigglesworthia* gen. nov. and *Wigglesworthia glossinidia* sp. nov., taxa consisting of the mycetocyte-associated, primary endosymbionts of tsetse flies. *Int J Syst Bacteriol* **45**: 848-851.

Aksoy, S (2000) Tsetse-A haven for microorganisms. *Parasitol Today* **16**: 114-118.

Alaux, C, Ducloz, F, Crauser, D and Le Conte, Y (2010) Diet effects on honeybee immunocompetence. *Biol Lett* **6**: 562-565.

Altincicek, B, Gross, J and Vilcinskas, A (2008a) Wounding-mediated gene expression and accelerated viviparous reproduction of the pea aphid *Acyrthosiphon pisum*. *Insect Mol Biol* **17**: 711-716.

Altincicek, B, Knorr, E and Vilcinskas, A (2008b) Beetle immunity: Identification of immune-inducible genes from the model insect *Tribolium castaneum*. *Dev Comp Immunol* **32**: 585-595.

Amano, A, Nakagawa, I and Yoshimori, T (2006) Autophagy in innate immunity against intracellular bacteria. *J Biochem* **140**: 161-166.

Amdam, GV, Aase, AL, Seehuus, SC, Kim Fondrk, M, Norberg, K and Hartfelder, K (2005) Social reversal of immunosenescence in honey bee workers. *Exp Gerontol* **40**: 939-947.

Amdam, GV, Simoes, ZL, Guidugli, KR, Norberg, K and Omholt, SW (2003) Disruption of vitellogenin gene function in adult honeybees by intra-abdominal injection of double-stranded RNA. *BMC Biotechnol* **3**: 1.

Anbutsu, H and Fukatsu, T (2010) Evasion, suppression and tolerance of *Drosophila* innate immunity by a male-killing *Spiroplasma* endosymbiont. *Insect Mol Biol* **19**: 481-488.

Anders, S and Huber, W (2010) Differential expression analysis for sequence count data. *Genome Biol* **11**: R106.

Andersen, SO (2010) Insect cuticular sclerotization: a review. *Insect Biochem Mol Biol* **40**: 166-178.

Anderson, KE, Russell, JA, Moreau, CS, Kautz, S, Sullam, KE, Hu, Y, Basinger, U, Mott, BM, Buck, N and Wheeler, DE (2012) Highly similar microbial communities are shared among related and trophically similar ant species. *Mol Ecol* **21**: 2282-2296.

Anh, NT, Nishitani, M, Harada, S, Yamaguchi, M and Kamei, K (2011) Essential role of Duox in stabilization of *Drosophila* wing. *J Biol Chem* **286**: 33244-33251.

Anselme, C, Perez-Brocal, V, Vallier, A, Vincent-Monegat, C, Charif, D, Latorre, A, Moya, A and Heddi, A (2008) Identification of the weevil immune genes and their expression in the bacteriome tissue. *BMC Biol* **6**: 43.

Anselme, C, Vallier, A, Balmand, S, Fauvarque, MO and Heddi, A (2006) Host *PGRP* gene expression and bacterial release in endosymbiosis of the weevil *Sitophilus zeamais*. *Appl Environ Microbiol* **72**: 6766-6772.

Arbouzova, NI and Zeidler, MP (2006) JAK/STAT signalling in *Drosophila*: insights into conserved regulatory and cellular functions. *Development* **133**: 2605-2616.

Armitage, SA and Boomsma, JJ (2010) The effects of age and social interactions on innate immunity in a leaf-cutting ant. *J Insect Physiol* **56**: 780-787.

Attardo, GM, Lohs, C, Heddi, A, Alam, UH, Yildirim, S and Aksoy, S (2008) Analysis of milk gland structure and function in *Glossina morsitans*: milk protein production, symbiont populations and fecundity. *J Insect Physiol* **54**: 1236-1242.

Bachali, S, Jager, M, Hassanin, A, Schoentgen, F, Jolles, P, Fiala-Medioni, A and Deutsch, JS (2002) Phylogenetic analysis of invertebrate lysozymes and the evolution of lysozyme function. *J Mol Evol* **54**: 652-664.

Baumann, P (2005) Biology bacteriocyte-associated endosymbionts of plant sap-sucking insects. *Annu Rev Microbiol* **59**: 155-189.

Bernstein, E, Caudy, AA, Hammond, SM and Hannon, GJ (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**: 363-366.

Birman, S, Morgan, B, Anzivino, M and Hirsh, J (1994) A novel and major isoform of tyrosine hydroxylase in *Drosophila* is generated by alternative RNA processing. *J Biol Chem* **269**: 26559-26567.

Bischoff, V, Vignal, C, Duvic, B, Boneca, IG, Hoffmann, JA and Royet, J (2006) Downregulation of the *Drosophila* immune response by peptidoglycan-recognition proteins SC1 and SC2. *PLoS Pathog* **2**: e14.

Blochmann, F (1882) Über das Vorkommen bakterienähnlicher Gebilde in den Geweben und Eiern verschiedener Insekten. *Zbl Bakt* **11**: 234-240.

Bolton, B (1995) *A New and General Catalogue of the Ants of the World*. Harvard University Press, Cambridge, MA, USA.

Boman, HG, Nilsson-Faye, I, Paul, K and Rasmuson, T, Jr. (1974) Insect immunity. I. Characteristics of an inducible cell-free antibacterial reaction in hemolymph of *Samia cynthia* pupae. *Infect Immun* **10**: 136-145.

Boman, HG, Nilsson, I and Rasmuson, B (1972) Inducible antibacterial defence system in *Drosophila*. *Nature* **237**: 232-235.

Bonasio, R, Zhang, G, Ye, C, Mutti, NS, Fang, X, et al. (2010) Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Science* **329**: 1068-1071.

Bourtzis, K, Pettigrew, MM and O'Neill, SL (2000) *Wolbachia* neither induces nor suppresses transcripts encoding antimicrobial peptides. *Insect Mol Biol* **9**: 635-639.

Boutros, M, Agaisse, H and Perrimon, N (2002) Sequential activation of signaling pathways during innate immune responses in *Drosophila*. *Dev Cell* **3**: 711-722.

Brady, SG, Schultz, TR, Fisher, BL and Ward, PS (2006) Evaluating alternative hypotheses for the early evolution and diversification of ants. *Proc Natl Acad Sci U S A* **103**: 18172-18177.

Braendle, C, Miura, T, Bickel, R, Shingleton, AW, Kambhampati, S and Stern, DL (2003) Developmental origin and evolution of bacteriocytes in the aphid-*Buchnera* symbiosis. *PLoS Biol* **1**: E21.

Brennan, LJ, Keddie, BA, Braig, HR and Harris, HL (2008) The endosymbiont *Wolbachia pipientis* induces the expression of host antioxidant proteins in an *Aedes albopictus* cell line. *PLoS ONE* **3**: e2083.

Brownlie, JC and Johnson, KN (2009) Symbiont-mediated protection in insect hosts. *Trends Microbiol* **17**: 348-354.

Brun, S, Vidal, S, Spellman, P, Takahashi, K, Tricoire, H and Lemaitre, B (2006) The MAPKKK Mekk1 regulates the expression of Turandot stress genes in response to septic injury in *Drosophila*. *Genes Cells* **11**: 397-407.

Bucher, G, Scholten, J and Klingler, M (2002) Parental RNAi in *Tribolium* (Coleoptera). *Curr Biol* **12**: R85-86.

Buchner, P (1965) *Endosymbiosis of animals with plant microorganisms*. Interscience Publishers, New York.

Buchon, N, Broderick, NA, Chakrabarti, S and Lemaitre, B (2009a) Invasive and indigenous microbiota impact intestinal stem cell activity through multiple pathways in *Drosophila*. *Genes Dev* **23**: 2333-2344.

Buchon, N, Broderick, NA, Poidevin, M, Pradervand, S and Lemaitre, B (2009b) *Drosophila* intestinal response to bacterial infection: activation of host defense and stem cell proliferation. *Cell Host Microbe* **5**: 200-211.

Budnik, V and White, K (1987) Genetic dissection of dopamine and serotonin synthesis in the nervous system of *Drosophila melanogaster*. *J Neurogenet* **4**: 309-314.

Bulet, P, Stocklin, R and Menin, L (2004) Anti-microbial peptides: from invertebrates to vertebrates. *Immunol Rev* **198**: 169-184.

Burke, GR and Moran, NA (2011a) Massive genomic decay in *Serratia symbiotica*, a recently evolved symbiont of aphids. *Genome Biol Evol* **3**: 195-208.

Burke, GR and Moran, NA (2011b) Responses of the pea aphid transcriptome to infection by facultative symbionts. *Insect Mol Biol* **20**: 357-365.

Bustin, SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* **25**: 169-193.

Bustin, SA (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *J Mol Endocrinol* **29**: 23-39.

Callewaert, L and Michiels, CW (2010) Lysozymes in the animal kingdom. *J Biosci* **35**: 127-160.

Cantarel, BL, Korf, I, Robb, SM, Parra, G, Ross, E, Moore, B, Holt, C, Sanchez Alvarado, A and Yandell, M (2008) MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res* **18**: 188-196.

Carlin, NF and Hölldobler, B (1986) The kin recognition system of carpenter ants (Camponotus Spp) .1. Hierarchical Cues in Small Colonies. *Behav Ecol Sociobiol* **19**: 123-134.

Casteels-Josson, K, Capaci, T, Casteels, P and Tempst, P (1993) Apidaecin multipeptide precursor structure: a putative mechanism for amplification of the insect antibacterial response. *EMBO J* **12**: 1569-1578.

Casteels-Josson, K, Zhang, W, Capaci, T, Casteels, P and Tempst, P (1994) Acute transcriptional response of the honeybee peptide-antibiotics gene repertoire and required post-translational conversion of the precursor structures. *J Biol Chem* **269**: 28569-28575.

Casteels, P, Ampe, C, Jacobs, F and Tempst, P (1993) Functional and chemical characterization of Hymenoptaecin, an antibacterial polypeptide that is infection-inducible in the honeybee (*Apis mellifera*). *J Biol Chem* **268**: 7044-7054.

Cerenius, L, Lee, BL and Söderhäll, K (2008) The proPO-system: pros and cons for its role in invertebrate immunity. *Trends Immunol* **29**: 263-271.

Cerenius, L and Söderhäll, K (2004) The prophenoloxidase-activating system in invertebrates. *Immunol Rev* **198**: 116-126.

Chapman, E, Farr, GW, Usaite, R, Furtak, K, Fenton, WA, Chaudhuri, TK, Hondorp, ER, Matthews, RG, Wolf, SG, Yates, JR, Pypaert, M and Horwich, AL (2006) Global aggregation of newly translated proteins in an *Escherichia coli* strain deficient of the chaperonin GroEL. *Proc Natl Acad Sci USA* **103**: 15800-15805.

Chen, X, Li, S and Aksoy, S (1999) Concordant evolution of a symbiont with its host insect species: molecular phylogeny of genus *Glossina* and its bacteriome-associated endosymbiont, *Wigglesworthia glossinidia*. *J Mol Evol* **48**: 49-58.

Chenchik, A, Zhu, YY, Diatchenko, L, Li, R, Hill, J and Siebert, PD (1998) Generation and use of high-quality cDNA from small amounts of total RNA by SMART PCR. In: *Gene cloning and analysis by RT-PCR (Bio Techniques Books, MA)*. pp. 305-319.

Cheng, X, Zhang, X, Pflugrath, JW and Studier, FW (1994) The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. *Proc Natl Acad Sci USA* **91**: 4034-4038.

Choe, K-M, Lee, H and Anderson, KV (2005) Drosophila peptidoglycan recognition protein LC (PGRP-LC) acts as a signal-transducing innate immune receptor. *Proc Natl Acad Sci USA* **102**: 1122-1126.

Choi, YS, Choo, YM, Lee, KS, Yoon, HJ, Kim, I, Je, YH, Sohn, HD and Jin, BR (2008) Cloning and expression profiling of four antibacterial peptide genes from the bumblebee *Bombus ignitus*. *Comp Biochem Physiol B Biochem Mol Biol* **150**: 141-146.

Chung, H, Sztal, T, Pasricha, S, Sridhar, M, Batterham, P and Daborn, PJ (2009) Characterization of *Drosophila melanogaster* cytochrome P450 genes. *Proc Natl Acad Sci USA*.

Cociancich, S, Ghazi, A, Hetru, C, Hoffmann, JA and Letellier, L (1993) Insect defensin, an inducible antibacterial peptide, forms voltage-dependent channels in *Micrococcus luteus*. *J Biol Chem* **268**: 19239-19245.

Collins, HL (2003) The role of iron in infections with intracellular bacteria. *Immunol Lett* **85**: 193-195.

Conord, C, Despres, L, Vallier, A, Balmand, S, Miquel, C, Zundel, S, Lemperiere, G and Heddi, A (2008) Long-term evolutionary stability of bacterial endosymbiosis in Curculionoidea: Additional evidence of symbiont replacement in the Dryophthoridae family. *Mol Biol Evol* **25**: 859-868.

Consortium, IAG (2010) Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol* **8**: e1000313.

Cornet, B, Bonmatin, JM, Hetru, C, Hoffmann, JA, Ptak, M and Vovelle, F (1995) Refined three-dimensional solution structure of insect defensin A. *Structure* **3**: 435-448.

Cremer, S, Armitage, SA and Schmid-Hempel, P (2007) Social immunity. *Curr Biol* **17**: R693-702.

Currie, CR, Poulsen, M, Mendenhall, J, Boomsma, JJ and Billen, J (2006) Coevolved crypts and exocrine glands support mutualistic bacteria in fungus-growing ants. *Science* **311**: 81-83.

Currie, CR, Scott, JA, Summerbell, RC and Malloch, D (1999) Fungus-growing ants use antibiotic-producing bacteria to control garden parasites. *Nature* **398**: 701-704.

Currie, CR, Wong, B, Stuart, AE, Schultz, TR, Rehner, SA, Mueller, UG, Sung, GH, Spatafora, JW and Straus, NA (2003) Ancient tripartite coevolution in the attine ant-microbe symbiosis. *Science* **299**: 386-388.

Cytrynska, M, Mak, P, Zdybicka-Barabas, A, Suder, P and Jakubowicz, T (2007) Purification and characterization of eight peptides from *Galleria mellonella* immune hemolymph. *Peptides* **28**: 533-546.

Daffre, S, Kylsten, P, Samakovlis, C and Hultmark, D (1994) The lysozyme locus in *Drosophila melanogaster*: an expanded gene family adapted for expression in the digestive tract. *Mol Gen Genet* **242**: 152-162.

Dai, ZM, Zhu, XJ, Chen, Q and Yang, WJ (2007) PCR-suppression effect: kinetic analysis and application to representative or long-molecule biased PCR-based amplification of complex samples. *J Biotechnol* **128**: 435-443.

Dale, C and Moran, NA (2006) Molecular interactions between bacterial symbionts and their hosts. *Cell* **126**: 453-465.

Dale, C, Young, SA, Haydon, DT and Welburn, SC (2001) The insect endosymbiont *Sodalis glossinidius* utilizes a type III secretion system for cell invasion. *Proc Natl Acad Sci USA* **98**: 1883-1888.

Dalmay, T, Hamilton, A, Rudd, S, Angell, S and Baulcombe, DC (2000) An RNA-dependent RNA polymerase gene in Arabidopsis is required for posttranscriptional gene silencing mediated by a transgene but not by a virus. *Cell* **101**: 543-553.

De Bary, A (1879) *Die Erscheinung der Symbiose*. Verlag v. K. J. Trübner, Straßburg.

De Gregorio, E, Spellman, PT, Rubin, GM and Lemaitre, B (2001) Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. *Proc Natl Acad Sci USA* **98**: 12590-12595.

De Gregorio, E, Spellman, PT, Tzou, P, Rubin, GM and Lemaitre, B (2002) The Toll and Imd pathways are the major regulators of the immune response in *Drosophila*. *Embo J* **21**: 2568-2579.

Degnan, PH, Lazarus, AB, Brock, CD and Wernegreen, JJ (2004) Host-symbiont stability and fast evolutionary rates in an ant-bacterium association: cospeciation of *Camponotus* species and their endosymbionts, Candidatus *Blochmannia*. *Syst Biol* **53**: 95-110.

Degnan, PH and Moran, NA (2008) Diverse phage-encoded toxins in a protective insect endosymbiont. *Appl Environ Microbiol* **74**: 6782-6791.

Degnan, PH, Yu, Y, Sisneros, N, Wing, RA and Moran, NA (2009) *Hamiltonella defensa*, genome evolution of protective bacterial endosymbiont from pathogenic ancestors. *Proc Natl Acad Sci USA* **106**: 9063-9068.

Delaney, JR, Stoven, S, Uvell, H, Anderson, KV, Engstrom, Y and Mlodzik, M (2006) Cooperative control of *Drosophila* immune responses by the JNK and NF-kappaB signaling pathways. *EMBO J* **25**: 3068-3077.

Diatchenko, L, Lau, YF, Campbell, AP, Chenchik, A, Moqadam, F, Huang, B, Lukyanov, S, Lukyanov, K, Gurskaya, N, Sverdlov, ED and Siebert, PD (1996) Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci USA* **93**: 6025-6030.

Dillon, RJ and Dillon, VM (2004) The gut bacteria of insects: nonpathogenic interactions. *Annu Rev Entomol* **49**: 71-92.

Dimarcq, JL, Hoffmann, D, Meister, M, Bulet, P, Lanot, R, Reichhart, JM and Hoffmann, JA (1994) Characterization and transcriptional profiles of a *Drosophila* gene encoding an insect defensin. A study in insect immunity. *Eur J Biochem* **221**: 201-209.

Dmochowska, A, Dignard, D, Henning, D, Thomas, DY and Bussey, H (1987) Yeast KEX1 gene encodes a putative protease with a carboxypeptidase B-like function involved in killer toxin and alpha-factor precursor processing. *Cell* **50**: 573-584.

Dostert, C, Jouanguy, E, Irving, P, Troxler, L, Galiana-Arnoux, D, Hetru, C, Hoffmann, JA and Imler, JL (2005) The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of *Drosophila*. *Nat Immunol* **6**: 946-953.

Douglas, AE (1998) Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria *Buchnera*. *Annu Rev Entomol* **43**: 17-37.

Douglas, AE, Bouvaine, S and Russell, RR (2011) How the insect immune system interacts with an obligate symbiotic bacterium. *Proc R Soc Lond B Biol Sci* **278**: 333-338.

Duckert, P, Brunak, S and Blom, N (2004) Prediction of proprotein convertase cleavage sites. *Protein Eng Des Sel* **17**: 107-112.

Dunbar, HE, Wilson, AC, Ferguson, NR and Moran, NA (2007) Aphid thermal tolerance is governed by a point mutation in bacterial symbionts. *PLoS Biol* **5**: e96.

Durand, D, Halldorsson, BV and Vernot, B (2006) A hybrid micro-macroevolutionary approach to gene tree reconstruction. *J Comput Biol* **13**: 320-335.

Düring, K, Porsch, P, Mahn, A, Brinkmann, O and Gieffers, W (1999) The non-enzymatic microbicidal activity of lysozymes. *FEBS Lett* **449**: 93-100.

Dziarski, R (2004) Peptidoglycan recognition proteins (PGRPs). *Mol Immunol* **40**: 877-886.

Edgar, RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792-1797.

Ekengren, S and Hultmark, D (2001) A family of Turandot-related genes in the humoral stress response of *Drosophila*. *Biochem Biophys Res Commun* **284**: 998-1003.

Erthal, M, Jr., Peres Silva, C and Ian Samuels, R (2007) Digestive enzymes in larvae of the leaf cutting ant, *Acromyrmex subterraneus* (Hymenoptera: Formicidae: Attini). *J Insect Physiol* **53**: 1101-1111.

Evans, JD, Aronstein, K, Chen, YP, Hetru, C, Imler, JL, Jiang, H, Kanost, M, Thompson, GJ, Zou, Z and Hultmark, D (2006) Immune pathways and defence mechanisms in honey bees *Apis mellifera*. *Insect Mol Biol* **15**: 645-656.

Feinberg, EH and Hunter, CP (2003) Transport of dsRNA into cells by the transmembrane protein SID-1. *Science* **301**: 1545-1547.

Feldhaar, H (2011) Bacterial symbionts as mediators of ecologically important traits of insect hosts. *Ecol Entomol* **36**: 533-543.

Feldhaar, H and Gross, R (2008) Immune reactions of insects on bacterial pathogens and mutualists. *Microbes Infect* **10**: 1082-1088.

Feldhaar, H, Straka, J, Krischke, M, Berthold, K, Stoll, S, Mueller, MJ and Gross, R (2007) Nutritional upgrading for omnivorous carpenter ants by the endosymbiont *Blochmannia*. *BMC Biol* **5**: 48.

Félix, R and Silveira, H (2012) The Role of *Anopheles gambiae* P450 Cytochrome in insecticide resistance and infection. In: *Insecticides - Pest Engineering* Perveen, F, ed., pp. 503-518. InTech.

Ferrari, J, Darby, AC, Daniell, TJ, Godfray, HCJ and Douglas, AE (2004) Linking the bacterial community in pea aphids with host-plant use and natural enemy resistance. *Ecol Entomol* **29**: 60-65.

Feyereisen, R (1999) Insect P450 enzymes. *Annu Rev Entomol* **44**: 507-533.

Fire, A, Xu, S, Montgomery, MK, Kostas, SA, Driver, SE and Mello, CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**: 806-811.

Fridovich, I (1995) Superoxide radical and superoxide dismutases. *Annu Rev Biochem* **64**: 97-112.

Friggi-Grelin, F, Coulom, H, Meller, M, Gomez, D, Hirsh, J and Birman, S (2003) Targeted gene expression in *Drosophila* dopaminergic cells using regulatory sequences from tyrosine hydroxylase. *J Neurobiol* **54**: 618-627.

Froy, O and Gurevitz, M (2003) Arthropod and mollusk defensins--evolution by exon-shuffling. *Trends Genet* **19**: 684-687.

Fuller, RS, Brake, A and Thorner, J (1989) Yeast prohormone processing enzyme (KEX2 gene product) is a Ca^{2+}-dependent serine protease. *Proc Natl Acad Sci USA* **86**: 1434-1438.

Gadau, J, Helmkampf, M, Nygaard, S, Roux, J, Simola, DF, Smith, CR, Suen, G, Wurm, Y and Smith, CD (2012) The genomic impact of 100 million years of social evolution in seven ant species. *Trends Genet* **28**: 14-21.

Ganesan, S, Aggarwal, K, Paquette, N and Silverman, N (2011) NF-kappaB/Rel proteins and the humoral immune responses of *Drosophila melanogaster*. *Curr Top Microbiol* **349**: 25-60.

Gao, B and Zhu, S (2010) Characterization of a hymenoptaecin-like antimicrobial peptide in the parasitic wasp *Nasonia vitripennis*. *Process Biochem* **45**: 139-146.

Gascuel, O (1997) BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol Biol Evol* **14**: 685-695.

Geier, M. 2009. Experimentelle Ansiedlung von Pathogenen im Darm von *Camponotus floridanus*. Diploma Thesis, Universität Würzburg.

Gerardo, NM, Altincicek, B, Anselme, C, Atamian, H, Barribeau, SM, et al. (2010) Immunity and other defenses in pea aphids, *Acyrthosiphon pisum*. *Genome Biol* **11**: R21.

Gil, R, Belda, E, Gosalbes, MJ, Delaye, L, Vallier, A, Vincent-Monegat, C, Heddi, A, Silva, FJ, Moya, A and Latorre, A (2008) Massive presence of insertion sequences in the genome of SOPE, the primary endosymbiont of the rice weevil *Sitophilus oryzae*. *Int Microbiol* **11**: 41-48.

Gil, R, Latorre, A and Moya, A (2004) Bacterial endosymbionts of insects: insights from comparative genomics. *Environ Microbiol* **6**: 1109-1122.

Gil, R, Silva, FJ, Zientz, E, Delmotte, F, Gonzalez-Candelas, F, Latorre, A, Rausell, C, Kamerbeek, J, Gadau, J, Hölldobler, B, van Ham, RC, Gross, R and Moya, A (2003)

The genome sequence of *Blochmannia floridanus*: comparative analysis of reduced genomes. *Proc Natl Acad Sci USA* **100**: 9388-9393.

Gillespie, JJ, Johnston, JS, Cannone, JJ and Gutell, RR (2006) Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization, and retrotransposable elements. *Insect Mol Biol* **15**: 657-686.

Glaser, RL and Meola, MA (2010) The native *Wolbachia* endosymbionts of *Drosophila melanogaster* and *Culex quinquefasciatus* increase host resistance to West Nile virus infection. *PLoS ONE* **5**: e11977.

Gobert, V, Gottar, M, Matskevich, AA, Rutschmann, S, Royet, J, Belvin, M, Hoffmann, JA and Ferrandon, D (2003) Dual activation of the *Drosophila* Toll pathway by two pattern recognition receptors. *Science* **302**: 2126-2130.

Goodwill, KE, Sabatier, C, Marks, C, Raag, R, Fitzpatrick, PF and Stevens, RC (1997) Crystal structure of tyrosine hydroxylase at 2.3 A and its implications for inherited neurodegenerative diseases. *Nat Struct Biol* **4**: 578-585.

Gorman, MJ, An, C and Kanost, MR (2007) Characterization of tyrosine hydroxylase from *Manduca sexta*. *Insect Biochem Mol Biol* **37**: 1327-1337.

Gorman, MJ and Arakane, Y (2010) Tyrosine hydroxylase is required for cuticle sclerotization and pigmentation in *Tribolium castaneum*. *Insect Biochem Mol Biol* **40**: 267-273.

Goto, T, Abe, Y, Kakuta, Y, Takeshita, K, Imoto, T and Ueda, T (2007) Crystal structure of *Tapes japonica* Lysozyme with substrate analogue: structural basis of the catalytic mechanism and manifestation of its chitinase activity accompanied by quaternary structural change. *J Biol Chem* **282**: 27459-27467.

Gottar, M, Gobert, V, Matskevich, AA, Reichhart, JM, Wang, C, Butt, TM, Belvin, M, Hoffmann, JA and Ferrandon, D (2006) Dual detection of fungal infections in *Drosophila* via recognition of glucans and sensing of virulence factors. *Cell* **127**: 1425-1437.

Gould, HJ (1967) The nature of high molecular weight fragments of ribosomal RNA. *J Mol Biol* **29**: 307-313.

Gouy, M, Guindon, S and Gascuel, O (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**: 221-224.

Griffith, GW and Beck, SD (1973) Intracellular symbiotes of pea aphid, *Acyrthosiphon pisum*. *J Insect Physiol* **19**: 75-84.

Guindon, S, Dufayard, JF, Lefort, V, Anisimova, M, Hordijk, W and Gascuel, O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**: 307-321.

Ha, EM, Oh, CT, Bae, YS and Lee, WJ (2005a) A direct role for dual oxidase in *Drosophila* gut immunity. *Science* **310**: 847-850.

Ha, EM, Oh, CT, Ryu, JH, Bae, YS, Kang, SW, Jang, IH, Brey, PT and Lee, WJ (2005b) An antioxidant system required for host protection against gut infection in *Drosophila*. *Dev Cell* **8**: 125-132.

Haine, ER (2008) Symbiont-mediated protection. *Proc R Soc Lond B Biol Sci* **275**: 353-361.

Haine, ER, Moret, Y, Siva-Jothy, MT and Rolff, J (2008) Antimicrobial defense and persistent infection in insects. *Science* **322**: 1257-1259.

Hamilton, C, Lejeune, BT and Rosengaus, RB (2011) Trophallaxis and prophylaxis: social immunity in the carpenter ant *Camponotus pennsylvanicus*. *Biol Lett* **7**: 89-92.

Hao, Z, Kasumba, I, Lehane, MJ, Gibson, WC, Kwon, J and Aksoy, S (2001) Tsetse immune responses and trypanosome transmission: implications for the development of tsetse-based strategies to reduce trypanosomiasis. *Proc Natl Acad Sci USA* **98**: 12648-12653.

Harris, HL, Brennan, LJ, Keddie, BA and Braig, HR (2010) Bacterial symbionts in insects: balancing life and death. *Symbiosis* **51**: 37-53.

Heddi, A, Charles, H, Khatchadourian, C, Bonnot, G and Nardon, P (1998) Molecular characterization of the principal symbiotic bacteria of the weevil *Sitophilus oryzae*: a peculiar G + C content of an endocytobiotic DNA. *J Mol Evol* **47**: 52-61.

Heddi, A, Grenier, AM, Khatchadourian, C, Charles, H and Nardon, P (1999a) Four intracellular genomes direct weevil biology: nuclear, mitochondrial, principal endosymbiont, and Wolbachia. *Proc Natl Acad Sci U S A* **96**: 6814-6819.

Heddi, A, Grenier, AM, Khatchadourian, C, Charles, H and Nardon, P (1999b) Four intracellular genomes direct weevil biology: nuclear, mitochondrial, principal endosymbiont, and *Wolbachia*. *Proc Natl Acad Sci USA* **96**: 6814-6819.

Heddi, A, Vallier, A, Anselme, C, Xin, H, Rahbe, Y and Wackers, F (2005) Molecular and cellular profiles of insect bacteriocytes: mutualism and harm at the initial evolutionary step of symbiogenesis. *Cell Microbiol* **7**: 293-305.

Hedengren, M, Asling, B, Dushay, MS, Ando, I, Ekengren, S, Wihlborg, M and Hultmark, D (1999) *Relish*, a central factor in the control of humoral but not cellular immunity in *Drosophila*. *Mol Cell* **4**: 827-837.

Hedges, LM, Brownlie, JC, O'Neill, SL and Johnson, KN (2008) *Wolbachia* and virus protection in insects. *Science* **322**: 702.

Heinze, J, Gadau, J, Hölldobler, B, Nanda, I, Schmid, M and Scheller, K (1994) Genetic-variability in the ant *Camponotus floridanus* detected by multilocus DNA-fingerprinting. *Naturwissenschaften* **81**: 34-36.

Herren, JK and Lemaitre, B (2011) *Spiroplasma* and host immunity: activation of humoral immune responses increases endosymbiont load and susceptibility to certain Gram-negative bacterial pathogens in *Drosophila melanogaster*. *Cell Microbiol* **13**: 1385-1396.

Herreweghe, JM and Michiels, CW (2012) Invertebrate lysozymes: diversity and distribution, molecular mechanism and *in vivo* function. *J Biosci* **37**: 327-348.

Hetru, C, Hoffmann, D and Bulet, P (1998) Antimicrobial peptides from insects. In: *Molecular mechanisms of immune responses in insects* Brey, PT and Hultmark, D, eds., pp. 40-66. Chapman & Hall, London.

Hilgenboecker, K, Hammerstein, P, Schlattmann, P, Telschow, A and Werren, JH (2008) How many species are infected with *Wolbachia*?-A statistical analysis of current data. *FEMS Microbiol Lett* **281**: 215-220.

Hölldobler, B and Wilson, EO (1990) *The ants*. Belknap Press of Harvard University Press, Cambridge, Mass. xii, 732 pp.

Hu, S and Yang, X (2000) dFADD, a novel death domain-containing adapter protein for the *Drosophila* caspase DREDD. *J Biol Chem* **275**: 30761-30764.

Hu, Y and Aksoy, S (2005) An antimicrobial peptide with trypanocidal activity characterized from *Glossina morsitans morsitans*. *Insect Biochem Mol Biol* **35**: 105-115.

Hultmark, D, Steiner, H, Rasmuson, T and Boman, HG (1980) Insect immunity. Purification and properties of three inducible bactericidal proteins from hemolymph of immunized pupae of *Hyalophora cecropia*. *Eur J Biochem* **106**: 7-16.

Hurst, GD, Anbutsu, H, Kutsukake, M and Fukatsu, T (2003) Hidden from the host: *Spiroplasma* bacteria infecting *Drosophila* do not cause an immune response, but are suppressed by ectopic immune activation. *Insect Mol Biol* **12**: 93-97.

Huson, DH and Bryant, D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**: 254-267.

Hutchence, KJ, Fischer, B, Paterson, S and Hurst, GD (2011) How do insects react to novel inherited symbionts? A microarray analysis of *Drosophila melanogaster* response to the presence of natural and introduced *Spiroplasma*. *Mol Ecol* **20**: 950-958.

Huvenne, H and Smagghe, G (2010) Mechanisms of dsRNA uptake in insects and potential of RNAi for pest control: a review. *J Insect Physiol* **56**: 227-235.

Hwang, SH, Cho, S and Park, YC (2010) cDNA cloning and induction of tyrosine hydroxylase gene from the diamondback moth, *Plutella xylostella*. *Arch Insect Biochem Physiol* **75**: 107-120.

Ibrahim, HR, Matsuzaki, T and Aoki, T (2001) Genetic evidence that antibacterial activity of lysozyme is independent of its catalytic function. *FEBS Lett* **506**: 27-32.

Ihle, KE, Page, RE, Frederick, K, Fondrk, MK and Amdam, GV (2010) Genotype effect on regulation of behaviour by vitellogenin supports reproductive origin of honeybee foraging bias. *Anim Behav* **79**: 1001-1006.

Imler, JL and Bulet, P (2005) Antimicrobial peptides in *Drosophila*: structures, activities and gene regulation. *Chem Immunol Allergy* **86**: 1-21.

Irving, P, Troxler, L, Heuer, TS, Belvin, M, Kopczynski, C, Reichhart, JM, Hoffmann, JA and Hetru, C (2001) A genome-wide analysis of immune responses in *Drosophila*. *Proc Natl Acad Sci USA* **98**: 15119-15124.

Ishikawa, H and Newburgh, RW (1972) Studies of the thermal conversion of 28 s RNA of *Galleria mellonella* (L.) to an 18 s product. *J Mol Biol* **64**: 135-144.

Itoh, T, Martin, W and Nei, M (2002) Acceleration of genomic evolution caused by enhanced mutation rate in endocellular symbionts. *Proc Natl Acad Sci USA* **99**: 12944-12948.

Jaenike, J and Brekke, TD (2011) Defensive endosymbionts: a cryptic trophic level in community ecology. *Ecol Lett* **14**: 150-155.

Jaenike, J, Unckless, R, Cockburn, SN, Boelio, LM and Perlman, SJ (2010) Adaptation via symbiosis: recent spread of a *Drosophila* defensive symbiont. *Science* **329**: 212-215.

Jarosch, A and Moritz, RF (2011) Systemic RNA-interference in the honeybee *Apis mellifera*: tissue dependent uptake of fluorescent siRNA after intra-abdominal application observed by laser-scanning microscopy. *J Insect Physiol* **57**: 851-857.

Johansson, KC, Metzendorf, C and Soderhall, K (2005) Microarray analysis of immune-challenged *Drosophila* hemocytes. *Exp Cell Res* **305**: 145-155.

Jollès, P and Jollès, J (1984) What's new in lysozyme research? Always a model system, today as yesterday. *Mol Cell Biochem* **63**: 165-189.

Julius, D, Blair, L, Brake, A, Sprague, G and Thorner, J (1983) Yeast alpha factor is processed from a larger precursor polypeptide: the essential role of a membrane-bound dipeptidyl aminopeptidase. *Cell* **32**: 839-852.

Kaltenpoth, M (2009) Actinobacteria as mutualists: general healthcare for insects? *Trends Microbiol* **17**: 529-535.

Kaltenpoth, M, Gottler, W, Herzner, G and Strohm, E (2005) Symbiotic bacteria protect wasp larvae from fungal infestation. *Curr Biol* **15**: 475-479.

Kaltenpoth, M, Schmitt, T, Polidori, C, Koedam, D and Strohm, E (2010) Symbiotic streptomycetes in antennal glands of the South American digger wasp genus *Trachypus* (Hymenoptera, Crabronidae). *Physiol Entomol* **35**: 196-200.

Kamath, RS and Ahringer, J (2003) Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods* **30**: 313-321.

Kamath, RS, Martinez-Campos, M, Zipperlen, P, Fraser, AG and Ahringer, J (2001) Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biol* **2**: RESEARCH0002.

Kambris, Z, Blagborough, AM, Pinto, SB, Blagrove, MS, Godfray, HC, Sinden, RE and Sinkins, SP (2010) *Wolbachia* stimulates immune gene expression and inhibits *Plasmodium* development in *Anopheles gambiae*. *PLoS Pathog* **6**: e1001143.

Kambris, Z, Cook, PE, Phuc, HK and Sinkins, SP (2009) Immune activation by life-shortening *Wolbachia* and reduced filarial competence in mosquitoes. *Science* **326**: 134-136.

Kaneko, T, Yano, T, Aggarwal, K, Lim, JH, Ueda, K, Oshima, Y, Peach, C, Erturk-Hasdemir, D, Goldman, WE, Oh, BH, Kurata, S and Silverman, N (2006) PGRP-LC and PGRP-LE have essential yet distinct functions in the *Drosophila* immune response to monomeric DAP-type peptidoglycan. *Nat Immunol* **7**: 715-723.

Kanost, MR, Jiang, H and Yu, XQ (2004) Innate immune responses of a lepidopteran insect, *Manduca sexta*. *Immunol Rev* **198**: 97-105.

Kellner, RLL (2003) Stadium-specific transmission of endosymbionts needed for pederin biosynthesis in three species of *Paederus* rove beetles. *Entomol Exp Appl* **107**: 115-124.

Kellner, RLL and Dettner, K (1996) Differential efficacy of toxic pederin in deterring potential arthropod predators of *Paederus* (Coleoptera: Staphylinidae) offspring. *Oecologia* **107**: 293-300.

Kikuchi, Y (2009) Endosymbiotic Bacteria in Insects: Their Diversity and Culturability. *Microbes Environ.* **24**: 195-204.

Kim, YS, Ryu, JH, Han, SJ, Choi, KH, Nam, KB, Jang, IH, Lemaitre, B, Brey, PT and Lee, WJ (2000) Gram-negative bacteria-binding protein, a pattern recognition receptor for lipopolysaccharide and beta-1,3-glucan that mediates the signaling for the induction of innate immune genes in *Drosophila melanogaster* cells. *J Biol Chem* **275**: 32721-32727.

Kirkegaard, K, Taylor, MP and Jackson, WT (2004) Cellular autophagy: surrender, avoidance and subversion by microorganisms. *Nat Rev Microbiol* **2**: 301-314.

Klaudiny, J, Albert, S, Bachanova, K, Kopernicky, J and Simuth, J (2005) Two structurally different *defensin* genes, one of them encoding a novel defensin isoform, are expressed in honeybee *Apis mellifera*. *Insect Biochem Mol Biol* **35**: 11-22.

Koga, R, Tsuchida, T and Fukatsu, T (2003) Changing partners in an obligate symbiosis: a facultative endosymbiont can compensate for loss of the essential endosymbiont *Buchnera* in an aphid. *Proc R Soc Lond B Biol Sci* **270**: 2543-2550.

Konrad, M, Vyleta, ML, Theis, FJ, Stock, M, Tragust, S, Klatt, M, Drescher, V, Marr, C, Ugelvig, LV and Cremer, S (2012) Social transfer of pathogenic fungus promotes active immunisation in ant colonies. *PLoS Biol* **10**: e1001300.

Kremer, N, Charif, D, Henri, H, Gavory, F, Wincker, P, Mavingui, P and Vavre, F (2012) Influence of *Wolbachia* on host gene expression in an obligatory symbiosis. *BMC Microbiol* **12 Suppl 1**: S7.

Kremer, N, Voronin, D, Charif, D, Mavingui, P, Mollereau, B and Vavre, F (2009) *Wolbachia* interferes with ferritin expression and iron metabolism in insects. *PLoS Pathog* **5**: e1000630.

Krogh, A, Larsson, B, von Heijne, G and Sonnhammer, EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**: 567-580.

Kulkarni, MM, Booker, M, Silver, SJ, Friedman, A, Hong, P, Perrimon, N and Mathey-Prevot, B (2006) Evidence of off-target effects associated with long dsRNAs in *Drosophila melanogaster* cell-based assays. *Nat Methods* **3**: 833-838.

Kurata, S (2010) Extracellular and intracellular pathogen recognition by *Drosophila* PGRP-LE and PGRP-LC. *Int Immunol*.

Lanot, R, Zachary, D, Holder, F and Meister, M (2001) Postembryonic hematopoiesis in *Drosophila*. *Dev Biol* **230**: 243-257.

Laughton, AM, Garcia, JR, Altincicek, B, Strand, MR and Gerardo, NM (2011) Characterisation of immune responses in the pea aphid, *Acyrthosiphon pisum*. *J Insect Physiol* **57**: 830-839.

Lee, WJ (2009) Bacterial-modulated host immunity and stem cell activation for gut homeostasis. *Genes Dev* **23**: 2260-2265.

Lee, WJ and Brey, PT (1995) Isolation and characterization of the lysozyme-encoding gene from the silkworm *Bombyx mori*. *Gene* **161**: 199-203.

Lee, WJ, Lee, JD, Kravchenko, VV, Ulevitch, RJ and Brey, PT (1996) Purification and molecular cloning of an inducible gram-negative bacteria-binding protein from the silkworm, *Bombyx mori*. *Proc Natl Acad Sci USA* **93**: 7888-7893.

Lefevre, 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.

Lemaitre, B and Hoffmann, J (2007) The host defense of *Drosophila melanogaster*. *Annu Rev Immunol* **25**: 697-743.

Lemaitre, B, Reichhart, JM and Hoffmann, JA (1997) *Drosophila* host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. *Proc Natl Acad Sci U S A* **94**: 14614-14619.

Lemos, FJA, Ribeiro, AF and Terra, WR (1993) A bacteria-digesting midgut-lysozyme from *Musca domestica* (Diptera) larvae. Purification, properties and secretory mechanism. *Insect Biochem Mol Biol* **23**: 533-541.

Letunic, I and Bork, P (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**: 127-128.

Levy, S and Shoham, T (2005a) Protein-protein interactions in the tetraspanin web. *Physiology* **20**: 218-224.

Levy, S and Shoham, T (2005b) The tetraspanin web modulates immune-signalling complexes. *Nat Rev Immunol* **5**: 136-148.

Li, B, Calvo, E, Marinotti, O, James, AA and Paskewitz, SM (2005) Characterization of the c-type lysozyme gene family in *Anopheles gambiae*. *Gene* **360**: 131-139.

Livak, KJ and Schmittgen, TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**: 402-408.

Login, FH, Balmand, S, Vallier, A, Vincent-Monegat, C, Vigneron, A, Weiss-Gayet, M, Rochat, D and Heddi, A (2011) Antimicrobial peptides keep insect endosymbionts under control. *Science* **334**: 362-365.

Login, FH and Heddi, A (2012) Insect immune system maintains long-term resident bacteria through a local response. *J Insect Physiol*.

Lourenco, AP, Martins, JR, Bitondi, MM and Simoes, ZL (2009) Trade-off between immune stimulation and expression of storage protein genes. *Arch Insect Biochem Physiol*.

Lu, HL, Vinson, SB and Pietrantonio, PV (2009) Oocyte membrane localization of vitellogenin receptor coincides with queen flying age, and receptor silencing by RNAi disrupts egg formation in fire ant virgin queens. *FEBS J* **276**: 3110-3123.

Maget-Dana, R and Ptak, M (1997) Penetration of the insect defensin A into phospholipid monolayers and formation of defensin A-lipid complexes. *Biophys J* **73**: 2527-2533.

Maillet, F, Bischoff, V, Vignal, C, Hoffmann, J and Royet, J (2008) The *Drosophila* peptidoglycan recognition protein PGRP-LF blocks PGRP-LC and IMD/JNK pathway activation. *Cell Host Microbe* **3**: 293-303.

Malcolm, BA, Rosenberg, S, Corey, MJ, Allen, JS, de Baetselier, A and Kirsch, JF (1989) Site-directed mutagenesis of the catalytic residues Asp-52 and Glu-35 of chicken egg white lysozyme. *Proc Natl Acad Sci USA* **86**: 133-137.

Martinez, J, Patkaniowska, A, Urlaub, H, Luhrmann, R and Tuschl, T (2002) Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**: 563-574.

Martinez, T, Burmester, T, Veenstra, JA and Wheeler, D (2000) Sequence and evolution of a hexamerin from the ant *Camponotus festinatus*. *Insect Mol Biol* **9**: 427-431.

Masschalck, B and Michiels, CW (2003) Antimicrobial properties of lysozyme in relation to foodborne vegetative bacteria. *Crit Rev Microbiol* **29**: 191-214.

McCutcheon, JP and Moran, NA (2012) Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol* **10**: 13-26.

McGraw, EA and O'Neill, SL (2004) *Wolbachia pipientis*: intracellular infection and pathogenesis in *Drosophila*. *Curr Opin Microbiol* **7**: 67-70.

Mello, CC and Conte, D, Jr. (2004) Revealing the world of RNA interference. *Nature* **431**: 338-342.

Mellroth, P, Karlsson, J, Hakansson, J, Schultz, N, Goldman, WE and Steiner, H (2005) Ligand-induced dimerization of *Drosophila* peptidoglycan recognition proteins *in vitro*. *Proc Natl Acad Sci USA* **102**: 6455-6460.

Mellroth, P, Karlsson, J and Steiner, H (2003) A scavenger function for a *Drosophila* peptidoglycan recognition protein. *J Biol Chem* **278**: 7059-7064.

Mellroth, P and Steiner, H (2006) PGRP-SB1: an N-acetylmuramoyl L-alanine amidase with antibacterial activity. *Biochem Biophys Res Commun* **350**: 994-999.

Meng, Y, Omuro, N, Funaguma, S, Daimon, T, Kawaoka, S, Katsuma, S and Shimada, T (2008) Prominent down-regulation of storage protein genes after bacterial challenge in eri-silkworm, *Samia cynthia ricini*. *Arch Insect Biochem Physiol* **67**: 9-19.

Michel, T, Reichhart, JM, Hoffmann, JA and Royet, J (2001) *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. *Nature* **414**: 756-759.

Mito, T, Nakamura, T, Bando, T, Ohuchi, H and Noji, S (2011) The advent of RNA interference in Entomology. *Entomol Sci* **14**: 1-8.

Mizushima, N, Levine, B, Cuervo, AM and Klionsky, DJ (2008) Autophagy fights disease through cellular self-digestion. *Nature* **451**: 1069-1075.

Monner, DA, Jonsson, S and Boman, HG (1971) Ampicillin-resistant mutants of *Escherichia coli* K-12 with lipopolysaccharide alterations affecting mating ability and susceptibility to sex-specific bacteriophages. *J Bacteriol* **107**: 420-432.

Moran, NA (1996) Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proc Natl Acad Sci USA* **93**: 2873-2878.

Moran, NA, Degnan, PH, Santos, SR, Dunbar, HE and Ochman, H (2005a) The players in a mutualistic symbiosis: insects, bacteria, viruses, and virulence genes. *Proc Natl Acad Sci USA* **102**: 16919-16926.

Moran, NA, McCutcheon, JP and Nakabachi, A (2008) Genomics and evolution of heritable bacterial symbionts. *Annu Rev Genet* **42**: 165-190.

Moran, NA, Russell, JA, Koga, R and Fukatsu, T (2005b) Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects. *Appl Environ Microbiol* **71**: 3302-3310.

Moreira, LA, Iturbe-Ormaetxe, I, Jeffery, JA, Lu, G, Pyke, AT, et al. (2009) A *Wolbachia* symbiont in *Aedes aegypti* limits infection with dengue, *Chikungunya*, and *Plasmodium*. *Cell* **139**: 1268-1278.

Moya, A, Pereto, J, Gil, R and Latorre, A (2008) Learning how to live together: genomic insights into prokaryote-animal symbioses. *Nat Rev Genet* **9**: 218-229.

Mueller, UG (2012) Symbiont recruitment versus ant-symbiont co-evolution in the attine ant-microbe symbiosis. *Curr Opin Microbiol* **15**: 1-9.

Munson, MA, Baumann, P, Clark, MA, Baumann, L, Moran, NA, Voegtlin, DJ and Campbell, BC (1991) Evidence for the establishment of aphid-eubacterium endosymbiosis in an ancestor of four aphid families. *J Bacteriol* **173**: 6321-6324.

Nakabachi, A, Shigenobu, S, Sakazume, N, Shiraki, T, Hayashizaki, Y, Carninci, P, Ishikawa, H, Kudo, T and Fukatsu, T (2005) Transcriptome analysis of the aphid bacteriocyte, the symbiotic host cell that harbors an endocellular mutualistic bacterium, *Buchnera*. *Proc Natl Acad Sci USA* **102**: 5477-5482.

Nappi, AJ and Vass, E (1993) Melanogenesis and the generation of cytotoxic molecules during insect cellular immune reactions. *Pigm Cell Res* **6**: 117-126.

Neckameyer, WS and White, K (1993) *Drosophila* tyrosine hydroxylase is encoded by the *pale* locus. *J Neurogenet* **8**: 189-199.

Nelson, CM, Ihle, KE, Fondrk, MK, Page, RE and Amdam, GV (2007) The gene vitellogenin has multiple coordinating effects on social organization. *PLoS Biol* **5**: e62.

Nelson, N (1999) Metal ion transporters and homeostasis. *EMBO J* **18**: 4361-4371.

Nicolas, E, Reichhart, JM, Hoffmann, JA and Lemaitre, B (1998) *In vivo* regulation of the IkappaB homologue *cactus* during the immune response of *Drosophila*. *J Biol Chem* **273**: 10463-10469.

Nishikori, K, Kubo, T and Morioka, M (2009a) Morph-dependent expression and subcellular localization of host serine carboxypeptidase in bacteriocytes of the pea aphid associated with degradation of the endosymbiotic bacterium *Buchnera*. *Zoolog Sci* **26**: 415-420.

Nishikori, K, Morioka, K, Kubo, T and Morioka, M (2009b) Age- and morph-dependent activation of the lysosomal system and *Buchnera* degradation in aphid endosymbiosis. *J Insect Physiol* **55**: 351-357.

Nissen-Meyer, J, Oppegard, C, Rogne, P, Haugen, HS and Kristiansen, PE (2010) Structure and Mode-of-Action of the Two-Peptide (Class-IIb) Bacteriocins. *Probiotics Antimicrob Proteins* **2**: 52-60.

Nogge, G (1981) Significance of symbionts for the maintenance of an optimal nutritional state for successful reproduction in hematophagous arthropods. *Parasitology* **82**: 101-104.

Nunes, FM and Simoes, ZL (2009) A non-invasive method for silencing gene transcription in honeybees maintained under natural conditions. *Insect Biochem Mol Biol* **39**: 157-160.

Nygaard, S, Zhang, G, Schiott, M, Li, C, Wurm, Y, et al. (2011) The genome of the leaf-cutting ant *Acromyrmex echinatior* suggests key adaptations to advanced social life and fungus farming. *Genome Res* **21**: 1339-1348.

Oliver, KM, Degnan, PH, Burke, GR and Moran, NA (2010) Facultative symbionts in aphids and the horizontal transfer of ecologically important traits. *Annu Rev Entomol* **55**: 247-266.

Oliver, KM, Degnan, PH, Hunter, MS and Moran, NA (2009) Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science* **325**: 992-994.

Oliver, KM, Moran, NA and Hunter, MS (2005) Variation in resistance to parasitism in aphids is due to symbionts not host genotype. *Proc Natl Acad Sci USA* **102**: 12795-12800.

Oliver, KM, Russell, JA, Moran, NA and Hunter, MS (2003) Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proc Natl Acad Sci USA* **100**: 1803-1807.

Osborne, RH (1996) Insect neurotransmission: neurotransmitters and their receptors. *Pharmacol Ther* **69**: 117-142.

Pais, R, Lohs, C, Wu, Y, Wang, J and Aksoy, S (2008) The obligate mutualist *Wigglesworthia glossinidia* influences reproduction, digestion, and immunity processes of its host, the tsetse fly. *Appl Environ Microbiol* **74**: 5965-5974.

Pan, X, Zhou, G, Wu, J, Bian, G, Lu, P, Raikhel, AS and Xi, Z (2012) *Wolbachia* induces reactive oxygen species (ROS)-dependent activation of the Toll pathway to control dengue virus in the mosquito *Aedes aegypti*. *Proc Natl Acad Sci USA* **109**: E23-31.

Park, JW, Kim, CH, Kim, JH, Je, BR, Roh, KB, Kim, SJ, Lee, HH, Ryu, JH, Lim, JH, Oh, BH, Lee, WJ, Ha, NC and Lee, BL (2007) Clustering of peptidoglycan recognition protein-SA is required for sensing lysine-type peptidoglycan in insects. *Proc Natl Acad Sci USA* **104**: 6602-6607.

Patel, A, Fondrk, MK, Kaftanoglu, O, Emore, C, Hunt, G, Frederick, K and Amdam, GV (2007) The making of a queen: TOR pathway is a key player in diphenic caste development. *PLoS ONE* **2**: e509.

Perez-Brocal, V, Gil, R, Ramos, S, Lamelas, A, Postigo, M, Michelena, JM, Silva, FJ, Moya, A and Latorre, A (2006) A small microbial genome: the end of a long symbiotic relationship? *Science* **314**: 312-313.

Pfaffl, MW, Tichopad, A, Prgomet, C and Neuvians, TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnol Lett* **26**: 509-515.

Pham, DQ and Winzerling, JJ (2010) Insect ferritins: Typical or atypical? *Biochim Biophys Acta* **1800**: 824-833.

Piel, J, Hofer, I and Hui, DQ (2004) Evidence for a symbiosis island involved in horizontal acquisition of pederin biosynthetic capabilities by the bacterial symbiont of *Paederus fuscipes* beetles. *J Bacteriol* **186**: 1280-1286.

Pili-Floury, S, Leulier, F, Takahashi, K, Saigo, K, Samain, E, Ueda, R and Lemaitre, B (2004) In vivo RNA interference analysis reveals an unexpected role for GNBP1 in the defense against Gram-positive bacterial infection in *Drosophila* adults. *J Biol Chem* **279**: 12848-12853.

Posnien, N, Schinko, J, Grossmann, D, Shippy, TD, Konopova, B and Bucher, G (2009) RNAi in the red flour beetle (*Tribolium*). *Cold Spring Harb Protoc* **2009**: pdb prot5256.

Poulsen, M and Sapountzis, P (2012) Behind every great ant, there is a great gut. *Mol Ecol* **21**: 2054-2057.

Price, DR and Gatehouse, JA (2008) RNAi-mediated crop protection against insects. *Trends Biotechnol* **26**: 393-400.

Rajagopal, R, Sivakumar, S, Agrawal, N, Malhotra, P and Bhatnagar, RK (2002) Silencing of midgut aminopeptidase N of *Spodoptera litura* by double-stranded RNA establishes its role as *Bacillus thuringiensis* toxin receptor. *J Biol Chem* **277**: 46849-46851.

Rämet, M, Lanot, R, Zachary, D and Manfruelli, P (2002) JNK signaling pathway is required for efficient wound healing in *Drosophila*. *Dev Biol* **241**: 145-156.

Randolt, K, Gimple, O, Geissendorfer, J, Reinders, J, Prusko, C, Mueller, MJ, Albert, S, Tautz, J and Beier, H (2008) Immune-related proteins induced in the hemolymph after aseptic and septic injury differ in honey bee worker larvae and adults. *Arch Insect Biochem Physiol* **69**: 155-167.

Ratzka, C, Förster, F, Liang, C, Kupper, M, Dandekar, T, Feldhaar, H and Gross, R (2012a) Molecular characterization of antimicrobial peptide genes of the carpenter ant *Camponotus floridanus*. *PLoS ONE* **7**: e43036.

Ratzka, C, Gross, R and Feldhaar, H (2012b) Endosymbiont tolerance and control within insect hosts. *Insects* **3**: 553-572.

Ratzka, C, Liang, C, Dandekar, T, Gross, R and Feldhaar, H (2011) Immune response of the ant *Camponotus floridanus* against pathogens and its obligate mutualistic endosymbiont. *Insect Biochem Mol Biol* **41**: 529-536.

Reddy, KV, Yedery, RD and Aranha, C (2004) Antimicrobial peptides: premises and promises. *Int J Antimicrob Agents* **24**: 536-547.

Regassa, LB and Gasparich, GE (2006) *Spiroplasmas*: evolutionary relationships and biodiversity. *Front Biosci* **11**: 2983-3002.

Regel, R, Matioli, SR and Terra, WR (1998) Molecular adaptation of *Drosophila melanogaster* lysozymes to a digestive function. *Insect Biochem Mol Biol* **28**: 309-319.

Reichhart, JM (2005) Tip of another iceberg: *Drosophila* serpins. *Trends Cell Biol* **15**: 659-665.

Rizki, TM, Rizki, RM and Grell, EH (1980) A mutant affecting the crystal cells in *Drosophila melanogaster*. *Roux Arch Dev Biol* **188**: 91-99.

Rodriguez-Cabrera, L, Trujillo-Bacallao, D, Borras-Hidalgo, O, Wright, DJ and Ayra-Pardo, C (2010) RNAi-mediated knockdown of a *Spodoptera frugiperda* trypsin-like serine-protease gene reduces susceptibility to a *Bacillus thuringiensis* Cry1Ca1 protoxin. *Environ Microbiol* **12**: 2894-2903.

Rosengaus, RB, Traniello, JFA, Chen, T, Brown, JJ and Karp, RD (1999) Immunity in a social insect. *Naturwissenschaften* **86**: 588-591.

Royet, J and Dziarski, R (2007) Peptidoglycan recognition proteins: pleiotropic sensors and effectors of antimicrobial defences. *Nat Rev Microbiol* **5**: 264-277.

Royet, J, Gupta, D and Dziarski, R (2011) Peptidoglycan recognition proteins: modulators of the microbiome and inflammation. *Nat Rev Immunol* **11**: 837-851.

Rozen, S and Skaletsky, H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**: 365-386.

Russ, WP and Engelman, DM (2000) The GxxxG motif: a framework for transmembrane helix-helix association. *J Mol Biol* **296**: 911-919.

Russell, JA, Latorre, A, Sabater-Munoz, B, Moya, A and Moran, NA (2003) Side-stepping secondary symbionts: widespread horizontal transfer across and beyond the Aphidoidea. *Mol Ecol* **12**: 1061-1075.

Ryu, JH, Ha, EM, Oh, CT, Seol, JH, Brey, PT, Jin, I, Lee, DG, Kim, J, Lee, D and Lee, WJ (2006) An essential complementary role of NF-kappaB pathway to microbicidal oxidants in *Drosophila* gut immunity. *Embo J* **25**: 3693-3701.

Ryu, JH, Kim, SH, Lee, HY, Bai, JY, Nam, YD, Bae, JW, Lee, DG, Shin, SC, Ha, EM and Lee, WJ (2008) Innate immune homeostasis by the homeobox gene *caudal* and commensal-gut mutualism in *Drosophila*. *Science* **319**: 777-782.

Ryu, KH, Park, JW, Kurokawa, K, M., M and Lee, BL (2010) The molecular activation and regulation mechanisms of proteolytic Toll signaling cascade in insect innate immunity. *ISJ* **7**: 181-191.

Sabin, LR, Hanna, SL and Cherry, S (2010) Innate antiviral immunity in *Drosophila*. *Curr Opin Immunol* **22**: 4-9.

Sadd, BM and Siva-Jothy, MT (2006) Self-harm caused by an insect's innate immunity. *Proc Biol Sci* **273**: 2571-2574.

Sameshima, S, Hasegawa, E, Kitade, O, Minaka, N and Matsumoto, T (1999) Phylogenetic comparison of endosymbionts with their host ants based on molecular evidence. *Zool Sci* **16**: 993-1000.

Sauer, C, Dudaczek, D, Hölldobler, B and Gross, R (2002) Tissue localization of the endosymbiotic bacterium "*Candidatus* Blochmannia floridanus" in adults and larvae of the carpenter ant *Camponotus floridanus*. *Appl Environ Microbiol* **68**: 4187-4193.

Sauer, C, Stackebrandt, E, Gadau, J, Hölldobler, B and Gross, R (2000) Systematic relationships and cospeciation of bacterial endosymbionts and their carpenter ant host species: proposal of the new taxon *Candidatus* Blochmannia gen. nov. *Int J Syst Evol Microbiol* **50 Pt 5**: 1877-1886.

Scarborough, CL, Ferrari, J and Godfray, HC (2005) Aphid protected from pathogen by endosymbiont. *Science* **310**: 1781.

Schägger, H (2006) Tricine-SDS-PAGE. *Nat Protoc* **1**: 16-22.

Schägger, H and von Jagow, G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**: 368-379.

Schaible, UE and Kaufmann, SH (2004) Iron and microbial infection. *Nat Rev Microbiol* **2**: 946-953.

Schlüns, H and Crozier, RH (2007) Relish regulates expression of antimicrobial peptide genes in the honeybee, *Apis mellifera*, shown by RNA interference. *Insect Mol Biol* **16**: 753-759.

Schmidt, RL, Trejo, TR, Plummer, TB, Platt, JL and Tang, AH (2008) Infection-induced proteolysis of PGRP-LC controls the IMD activation and melanization cascades in *Drosophila*. *Faseb J* **22**: 918-929.

Schmidt, WM and Mueller, MW (1999) CapSelect: a highly sensitive method for 5' CAP-dependent enrichment of full-length cDNA in PCR-mediated analysis of mRNAs. *Nucleic Acids Res* **27**: e31.

Schröder, D, Deppisch, H, Obermayer, M, Krohne, G, Stackebrandt, E, Holldobler, B, Goebel, W and Gross, R (1996) Intracellular endosymbiotic bacteria of *Camponotus* species (carpenter ants): systematics, evolution and ultrastructural characterization. *Mol Microbiol* **21**: 479-489.

Scott, JJ, Oh, DC, Yuceer, MC, Klepzig, KD, Clardy, J and Currie, CR (2008) Bacterial protection of beetle-fungus mutualism. *Science* **322**: 63-63.

Sedgwick, SG and Smerdon, SJ (1999) The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends Biochem Sci* **24**: 311-316.

Seitz, V, Clermont, A, Wedde, M, Hummel, M, Vilcinskas, A, Schlatterer, K and Podsiadlowski, L (2003) Identification of immunorelevant genes from greater wax moth (*Galleria mellonella*) by a subtractive hybridization approach. *Dev Comp Immunol* **27**: 207-215.

Senes, A, Engel, DE and DeGrado, WF (2004) Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Curr Opin Struct Biol* **14**: 465-479.

Serebrov, VV, Alekseev, AA and Glupov, VV (2001) Changes in the activity and pattern of hemolymph esterases in the larvae of wax moth *Galleria mellonella* L. (Lepidoptera, Pyralidae) during mycosis. *Izv Akad Nauk Ser Biol*: 588-592.

Shiotsuki, T and Kato, Y (1999) Induction of carboxylesterase isozymes in *Bombyx mori* by *E. coli* infection. *Insect Biochem Mol Biol* **29**: 731-736.

Siebert, PD, Chenchik, A, Kellogg, DE, Lukyanov, KA and Lukyanov, SA (1995) An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Res* **23**: 1087-1088.

Sijen, T, Fleenor, J, Simmer, F, Thijssen, KL, Parrish, S, Timmons, L, Plasterk, RH and Fire, A (2001) On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **107**: 465-476.

Silverman, N, Zhou, R, Erlich, RL, Hunter, M, Bernstein, E, Schneider, D and Maniatis, T (2003) Immune activation of NF-kappaB and JNK requires *Drosophila* TAK1. *J Biol Chem* **278**: 48928-48934.

Smith, CD, Zimin, A, Holt, C, Abouheif, E, Benton, R, et al. (2011a) Draft genome of the globally widespread and invasive Argentine ant (*Linepithema humile*). *Proc Natl Acad Sci USA* **108**: 5673-5678.

Smith, CR, Smith, CD, Robertson, HM, Helmkampf, M, Zimin, A, et al. (2011b) Draft genome of the red harvester ant *Pogonomyrmex barbatus*. *Proc Natl Acad Sci USA* **108**: 5667-5672.

Smith, TF (2008) Diversity of WD-repeat proteins. *Subcell Biochem* **48**: 20-30.

Söderhäll, K and Cerenius, L (1998) Role of the prophenoloxidase-activating system in invertebrate immunity. *Curr Opin Immunol* **10**: 23-28.

Sonnhammer, EL, von Heijne, G and Krogh, A (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **6**: 175-182.

Souza-Neto, JA, Sim, S and Dimopoulos, G (2009) An evolutionary conserved function of the JAK-STAT pathway in anti-dengue defense. *Proc Natl Acad Sci USA* **106**: 17841-17846.

Stanke, M, Diekhans, M, Baertsch, R and Haussler, D (2008) Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24**: 637-644.

Steiner, H, Hultmark, D, Engstrom, A, Bennich, H and Boman, HG (1981) Sequence and specificity of two antibacterial proteins involved in insect immunity. *Nature* **292**: 246-248.

Stoll, S. 2009. Funktionelle Analyse von *Blochmannia floridanus*, dem primären Endosymbionten der Rossameise *Camponotus floridanus*. PhD Thesis, Universität Würzburg.

Stoll, S, Feldhaar, H, Fraunholz, MJ and Gross, R (2010) Bacteriocyte dynamics during development of a holometabolous insect, the carpenter ant *Camponotus floridanus*. *BMC Microbiol* **10**: 308.

Stoll, S, Feldhaar, H and Gross, R (2008) Transcriptional profiling of the endosymbiont *Blochmannia floridanus* during different developmental stages of its holometabolous ant host. *Environ Microbiol* **11**: 877-888.

Stoll, S, Feldhaar, H and Gross, R (2009) Promoter characterization in the AT-rich genome of the obligate endosymbiont "*Candidatus* Blochmannia floridanus". *J Bacteriol* **191**: 3747-3751.

Stouthamer, R, Breeuwer, JA and Hurst, GD (1999) *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. *Annu Rev Microbiol* **53**: 71-102.

Stöven, S, Ando, I, Kadalayil, L, Engstrom, Y and Hultmark, D (2000) Activation of the *Drosophila* NF-kappaB factor Relish by rapid endoproteolytic cleavage. *EMBO Rep* **1**: 347-352.

Stöven, S, Silverman, N, Junell, A, Hedengren-Olcott, M, Erturk, D, Engstrom, Y, Maniatis, T and Hultmark, D (2003) Caspase-mediated processing of the *Drosophila* NF-kappaB factor Relish. *Proc Natl Acad Sci USA* **100**: 5991-5996.

Strand, MR (2008) The insect cellular immune response. *Insect Sci* **15**: 1-14.

Street, TO, Rose, GD and Barrick, D (2006) The role of introns in repeat protein gene formation. *J Mol Biol* **360**: 258-266.

Suen, G, Teiling, C, Li, L, Holt, C, Abouheif, E, et al. (2011) The genome sequence of the leaf-cutter ant *Atta cephalotes* reveals insights into its obligate symbiotic lifestyle. *PLoS Genet* **7**: e1002007.

Talavera, G and Castresana, J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* **56**: 564-577.

Tatusov, RL, Fedorova, ND, Jackson, JD, Jacobs, AR, Kiryutin, B, et al. (2003) The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**: 41.

Tauszig-Delamasure, S, Bilak, H, Capovilla, M, Hoffmann, JA and Imler, JL (2002) Drosophila MyD88 is required for the response to fungal and Gram-positive bacterial infections. *Nat Immunol* **3**: 91-97.

Teixeira, L, Ferreira, A and Ashburner, M (2008) The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. *PLoS Biol* **6**: e2.

Terenius, O, Papanicolaou, A, Garbutt, JS, Eleftherianos, I, Huvenne, H, et al. (2011) RNA interference in Lepidoptera: an overview of successful and unsuccessful studies and implications for experimental design. *J Insect Physiol* **57**: 231-245.

Terra, WR, Espinozafuentes, FP and Ferreira, C (1988) Midgut amylase, lysozyme, aminopeptidase, and trehalase from larvae and adults of *Musca domestica*. *Arch Insect Biochem Physiol* **9**: 283-297.

Tian, C, Gao, B, Fang, Q, Ye, G and Zhu, S (2010) Antimicrobial peptide-like genes in *Nasonia vitripennis*: a genomic perspective. *BMC Genomics* **11**: 187.

Tian, H, Peng, H, Yao, Q, Chen, H, Xie, Q, Tang, B and Zhang, W (2009) Developmental control of a lepidopteran pest *Spodoptera exigua* by ingestion of bacteria expressing dsRNA of a non-midgut gene. *PLoS ONE* **4**: e6225.

Timmons, L, Court, DL and Fire, A (2001) Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene* **263**: 103-112.

Timmons, L and Fire, A (1998) Specific interference by ingested dsRNA. *Nature* **395**: 854.

Toh, H, Weiss, BL, Perkin, SA, Yamashita, A, Oshima, K, Hattori, M and Aksoy, S (2006) Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. *Genome Res* **16**: 149-156.

Tomoyasu, Y and Denell, RE (2004) Larval RNAi in *Tribolium* (Coleoptera) for analyzing adult development. *Dev Genes Evol* **214**: 575-578.

Traniello, JF, Rosengaus, RB and Savoie, K (2002) The development of immunity in a social insect: evidence for the group facilitation of disease resistance. *Proc Natl Acad Sci USA* **99**: 6838-6842.

Trapnell, C, Pachter, L and Salzberg, SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**: 1105-1111.

Trapnell, C, Roberts, A, Goff, L, Pertea, G, Kim, D, Kelley, DR, Pimentel, H, Salzberg, SL, Rinn, JL and Pachter, L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7**: 562-578.

Trapnell, C, Williams, BA, Pertea, G, Mortazavi, A, Kwan, G, van Baren, MJ, Salzberg, SL, Wold, BJ and Pachter, L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**: 511-515.

True, JR, Edwards, KA, Yamamoto, D and Carroll, SB (1999) *Drosophila* wing melanin patterns form by vein-dependent elaboration of enzymatic prepatterns. *Curr Biol* **9**: 1382-1391.

Ugelvig, LV and Cremer, S (2007) Social prophylaxis: group interaction promotes collective immunity in ant colonies. *Curr Biol* **17**: 1967-1971.

Vernot, B, Stolzer, M, Goldman, A and Durand, D (2008) Reconciliation with non-binary species trees. *J Comput Biol* **15**: 981-1006.

Vie, A, Cigna, M, Toci, R and Birman, S (1999) Differential regulation of *Drosophila* tyrosine hydroxylase isoforms by dopamine binding and cAMP-dependent phosphorylation. *J Biol Chem* **274**: 16788-16795.

Vigneron, A, Charif, D, Vincent-Monegat, C, Vallier, A, Gavory, F, Wincker, P and Heddi, A (2012) Host gene response to endosymbiont and pathogen in the cereal weevil *Sitophilus oryzae*. *BMC Microbiol* **12 Suppl 1**: S14.

Viljakainen, L and Pamilo, P (2005) Identification and molecular characterization of *defensin* gene from the ant *Formica aquilonia*. *Insect Mol Biol* **14**: 335-338.

Viljakainen, L and Pamilo, P (2008) Selection on an antimicrobial peptide defensin in ants. *J Mol Evol* **67**: 643-652.

Vodovar, N and Saleh, M (2012) Of Insects and Viruses: The Role of Small RNAs in Insect Defence. In: *Small RNAs their diversity, roles and practical uses* Jockusch, EL, ed., pp. 1-201.

von Burg, S, Ferrari, J, Muller, CB and Vorburger, C (2008) Genetic variation and covariation of susceptibility to parasitoids in the aphid *Myzus persicae*: no evidence for trade-offs. *Proc R Soc Lond B Biol Sci* **275**: 1089-1094.

Vorburger, C, Gehrer, L and Rodriguez, P (2010) A strain of the bacterial symbiont *Regiella insecticola* protects aphids against parasitoids. *Biol Lett* **6**: 109-111.

Voronin, D, Cook, DA, Steven, A and Taylor, MJ (2012) Autophagy regulates *Wolbachia* populations across diverse symbiotic associations. *Proc Natl Acad Sci USA* **109**: E1638-1646.

Walshe, DP, Lehane, SM, Lehane, MJ and Haines, LR (2009) Prolonged gene knockdown in the tsetse fly *Glossina* by feeding double stranded RNA. *Insect Mol Biol* **18**: 11-19.

Wang, J and Aksoy, S (2012) PGRP-LB is a maternally transmitted immune milk protein that influences symbiosis and parasitism in tsetse's offspring. *Proc Natl Acad Sci USA* **109**: 10552-10557.

Wang, J, Wu, Y, Yang, G and Aksoy, S (2009) Interactions between mutualist *Wigglesworthia* and tsetse peptidoglycan recognition protein (PGRP-LB) influence trypanosome transmission. *Proc Natl Acad Sci USA* **106**: 12133-12138.

Wang, Y, Mutti, NS, Ihle, KE, Siegel, A, Dolezal, AG, Kaftanoglu, O and Amdam, GV (2010) Down-regulation of honey bee *IRS* gene biases behavior toward food rich in protein. *PLoS Genet* **6**: e1000896.

Weber, AN, Tauszig-Delamasure, S, Hoffmann, JA, Lelievre, E, Gascan, H, Ray, KP, Morse, MA, Imler, JL and Gay, NJ (2003) Binding of the *Drosophila* cytokine Spätzle to Toll is direct and establishes signaling. *Nat Immunol* **4**: 794-800.

Weiss, BL, Maltz, M and Aksoy, S (2012) Obligate symbionts activate immune system development in the tsetse fly. *J Immunol* **188**: 3395-3403.

Weiss, BL, Wang, J and Aksoy, S (2011) Tsetse immune system maturation requires the presence of obligate symbionts in larvae. *PLoS Biol* **9**: e1000619.

Weiss, BL, Wu, Y, Schwank, JJ, Tolwinski, NS and Aksoy, S (2008) An insect symbiosis is influenced by bacterium-specific polymorphisms in outer-membrane protein A. *Proc Natl Acad Sci USA* **105**: 15088-15093.

Wernegreen, JJ (2002) Genome evolution in bacterial endosymbionts of insects. *Nat Rev* **3**: 850-861.

Wernegreen, JJ (2012) Strategies of genomic integration within insect-bacterial mutualisms. *Biol Bull* **223**: 112-122.

Wernegreen, JJ, Kauppinen, SN, Brady, SG and Ward, PS (2009) One nutritional symbiosis begat another: phylogenetic evidence that the ant tribe Camponotini acquired *Blochmannia* by tending sap-feeding insects. *BMC Evol Biol* **9**: 292.

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.

Werner, T, Liu, G, Kang, D, Ekengren, S, 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.

Werren, JH (1997) Biology of *Wolbachia*. *Annu Rev Entomol* **42**: 587-609.

Werren, JH, Baldo, L and Clark, ME (2008) *Wolbachia*: master manipulators of invertebrate biology. *Nat Rev Microbiol* **6**: 741-751.

Westermeier, R and Gronau, S (2005) *Electrophoresis in practice : a guide to methods and applications of DNA and protein separations*, 4th rev. and enl. edn. Wiley-VCH, Weinheim. xx, 406 p. pp.

Wheeler, DE and Martinez, T (1995) Storage proteins in ants (Hymenoptera:Formicidae). *Comp Biochem Physiol B Biochem Mol Biol* **112**: 15-19.

Whitcomb, RF and Williamson, DL (1975) Helical wall-free prokaryotes in insects: multiplication and pathogenicity. *Ann NY Acad Sci* **266**: 260-275.

Wilkinson, TL (1998) The elimination of intracellular microorganisms from insects: an analysis of antibiotic-treatment in the pea aphid (*Acyrthosiphon pisum*). *Comp Biochem Phys A* **119**: 871-881.

Wilson-Rich, N, Spivak, M, Fefferman, NH and Starks, PT (2009) Genetic, individual, and group facilitation of disease resistance in insect societies. *Annu Rev Entomol* **54**: 405-423.

Wilson, EO (1971) *The insect societies*. Belknap Press of Harvard University Press, Cambridge, Mass.,. 548 pp.

Winnebeck, EC, Millar, CD and Warman, GR (2010) Why does insect RNA look degraded? *J Insect Sci* **10**: 159.

Winston, WM, Molodowitch, C and Hunter, CP (2002) Systemic RNAi in *C. elegans* requires the putative transmembrane protein SID-1. *Science* **295**: 2456-2459.

Winston, WM, Sutherlin, M, Wright, AJ, Feinberg, EH and Hunter, CP (2007) *Caenorhabditis elegans* SID-2 is required for environmental RNA interference. *Proc Natl Acad Sci USA* **104**: 10565-10570.

Wolschin, F, Hölldobler, B, Gross, R and Zientz, E (2004) Replication of the endosymbiotic bacterium *Blochmannia floridanus* is correlated with the developmental and reproductive stages of its ant host. *Appl Environ Microbiol* **70**: 4096-4102.

Wolschin, F, Mutti, NS and Amdam, GV (2011) Insulin receptor substrate influences female caste development in honeybees. *Biol Lett* **7**: 112-115.

Wu, M, Sun, LV, Vamathevan, J, Riegler, M, Deboy, R, et al. (2004) Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. *PLoS Biol* **2**: E69.

Wurm, Y, Wang, J, Riba-Grognuz, O, Corona, M, Nygaard, S, et al. (2011) The genome of the fire ant *Solenopsis invicta*. *Proc Natl Acad Sci USA*.

Xie, J, Vilchez, I and Mateos, M (2010) Spiroplasma bacteria enhance survival of *Drosophila hydei* attacked by the parasitic wasp *Leptopilina heterotoma*. *PLoS ONE* **5**: e12149.

Xu, P, Shi, M and Chen, XX (2009) Antimicrobial peptide evolution in the Asiatic honey bee *Apis cerana*. *PLoS ONE* **4**: e4239.

Yang, T and Liu, N (2011) Genome analysis of cytochrome P450s and their expression profiles in insecticide resistant mosquitoes, *Culex quinquefasciatus*. *PLoS ONE* **6**: e29418.

Yano, T, Mita, S, Ohmori, H, Oshima, Y, Fujimoto, Y, Ueda, R, Takada, H, Goldman, WE, Fukase, K, Silverman, N, Yoshimori, T and Kurata, S (2008) Autophagic control of *Listeria* through intracellular innate immune recognition in *Drosophila*. *Nat Immunol* **9**: 908-916.

Yoshiga, T, Georgieva, T, Dunkov, BC, Harizanova, N, Ralchev, K and Law, JH (1999) *Drosophila melanogaster* transferrin. Cloning, deduced protein sequence, expression during the life cycle, gene localization and up-regulation on bacterial infection. *Eur J Biochem* **260**: 414-420.

Yoshiga, T, Hernandez, VP, Fallon, AM and Law, JH (1997) Mosquito transferrin, an acute-phase protein that is up-regulated upon infection. *Proc Natl Acad Sci USA* **94**: 12337-12342.

Zaidman-Remy, A, Herve, M, Poidevin, M, Pili-Floury, S, Kim, MS, Blanot, D, Oh, BH, Ueda, R, Mengin-Lecreulx, D and Lemaitre, B (2006) The *Drosophila* amidase PGRP-LB modulates the immune response to bacterial infection. *Immunity* **24**: 463-473.

Zaidman-Remy, A, Poidevin, M, Herve, M, Welchman, DP, Paredes, JC, Fahlander, C, Steiner, H, Mengin-Lecreulx, D and Lemaitre, B (2011) *Drosophila* Immunity: Analysis of PGRP-SB1 Expression, Enzymatic Activity and Function. *PLoS ONE* **6**: e17231.

Zhang, G and Ghosh, S (2002) Negative regulation of toll-like receptor-mediated signaling by Tollip. *J Biol Chem* **277**: 7059-7065.

Zhang, Z and Zhu, S (2012) Comparative genomics analysis of five families of antimicrobial peptide-like genes in seven ant species. *Dev Comp Immunol*.

Zhu, F, Xu, J, Palli, R, Ferguson, J and Palli, SR (2011) Ingested RNA interference for managing the populations of the Colorado potato beetle, *Leptinotarsa decemlineata*. *Pest management science* **67**: 175-182.

Zhu, Y, Johnson, TJ, Myers, AA and Kanost, MR (2003) Identification by subtractive suppression hybridization of bacteria-induced genes expressed in *Manduca sexta* fat body. *Insect Biochem Mol Biol* **33**: 541-559.

Zientz, E, Beyaert, I, Gross, R and Feldhaar, H (2006) Relevance of the endosymbiosis of *Blochmannia floridanus* and carpenter ants at different stages of the life cycle of the host. *Appl Environ Microbiol* **72**: 6027-6033.

Zientz, E, Dandekar, T and Gross, R (2004) Metabolic interdependence of obligate intracellular bacteria and their insect hosts. *Microbiol Mol Biol Rev* **68**: 745-770.

Zou, Z, Lopez, DL, Kanost, MR, Evans, JD and Jiang, H (2006) Comparative analysis of serine protease-related genes in the honey bee genome: possible involvement in embryonic development and innate immunity. *Insect Mol Biol* **15**: 603-614.