ANALYSIS OF ANTENNAL TRANSCRIPTOMES OF HERMIT CRABS COMPARING AQUATIC (*PAGURUS BERNHARDUS*) AND TERRESTRIAL (*COENOBITA CLYPEATUS*) SPECIES

Dissertation

To fulfill the
requirements for the degree of

„doctor rerum naturalium“ (Dr. rer. nat.)

Submitted to the council of the
Faculty of Biology and Pharmacy
of the Friedrich Schiller University Jena

by Dipl. biol. Katrin Christine Groh-Lunow

born 21st of March 1985 in Greiz

Reviewers:

1. Prof. Dr. Bill S. Hansson (Jena)

2. Prof. Dr. Steffen Harzsch (Greifswald)

3. Prof. Dr. Wolfgang Rössler (Würzburg)

Day of public defense: 04. 07. 2014

_____

Table of contents

1. Introduction

Understanding the evolution of animals is an overarching goal of science. From today's perspective it can be retraced by fossils acting as mute witnesses to the process and recent species that are the result of millennia of adaptation. Among the evolutionary milestones is the successful transition from water to land. Countless animals and plants became terrestrial in their ancestral lineages and met the various requirements of this step independently. One of the most important abilities they all had to keep was the ability to sense chemicals in their surroundings. It is well known how terrestrial animals such as vertebrates and insects sense and compute environmental odors from a molecular and an information integration perspective. However, to conclude what the ancestral state of olfaction was, it is necessary to compare those findings with species that never left water. This comparison proves difficult in insects, as all recent species are either terrestrial or they have gone back to water secondarily. Crustaceans, however, provide the opportunity to study the current state of a primary aquatic nose in direct comparison to the results of 20 million years of terrestrial olfaction in hermit crab species.

This thesis aims at understanding the molecular genetics of crustacean chemosensation and its adaptation during the terrestrialization of hermit crabs. It further compares those findings with insects, which are the arthropod sister phylum to crustaceans with a history of 500 million years of terrestrial olfaction.

Arthropod terrestrialization and chemosensation

All organisms approach their environment by their senses. Probably the most ancient of which is the chemical sense, present in all organisms from bacteria to mammals. Chemicals are released by almost all organic and inorganic matters to be found in nature and to the receiving organisms they serve as transmitters of habitat features, as well as the presence of nutrients, conspecifics and dangers like predators. Chemical sensing alone or in combination with visual, tactile and acoustic senses allows the organism to detect, discriminate, localize, evaluate and associate their origin. For example, this enables the organism to make a decision between attraction and aversion. The characteristics of chemical stimuli heavily depend on the environment, as the chemical composition and properties are influenced by the habitat background and the properties of the transporting medium (e.g. air or water). If

the conveying medium abruptly changes from water to air, then it has to be met by several adaptations. The altered water availability coming from water to land requires the prevention from desiccation, as well as changes in ion balance, respiration and metabolism. These requirements were independently met by numerous arthropods which succeeded in the transition from water to land. Figure 1 depicts the habitats of recent arthropod clades, which are indicated by blue triangles if it is aquatic, orange squares if it is terrestrial or a combination of both colors and symbols if some members are terrestrial and some aquatic.

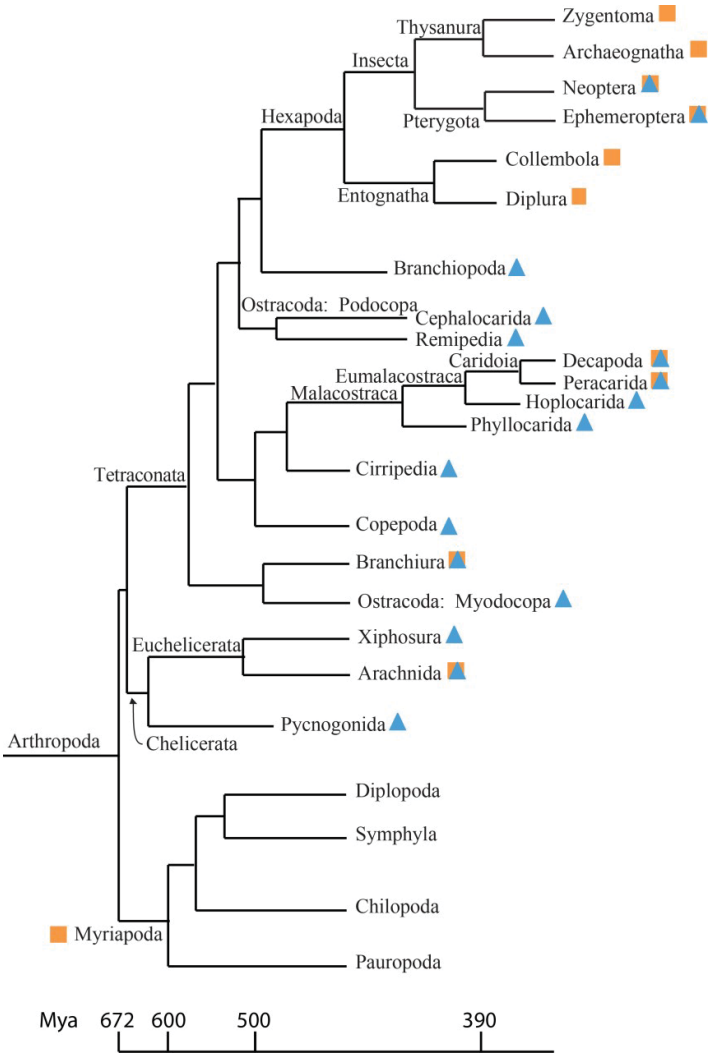


Figure 1: Phylogenetic relationship of Arthropods. Terrestrial clades indicated with orange squares, aquatic clades indicated with blue triangles. Clades with members of both habitats with blue triangle on orange square; modified after (Regier, Shultz, and Kambic 2005)

The most prominent terrestrial members of the arthropods are Insecta with an estimated species count of 5 million (Chapman 2009). Extant members of this class are primary land-living while some have aquatic larval stages and terrestrial adults (like mosquitoes, mayflies and dragonflies) or include members that returned to water secondarily (like predatory water beetles). Representatives of the main insect clades are well studied and this has led to a profound understanding of their sensory abilities. Besides dragonflies, where hunting adults mainly rely on vision, the most important sense for insects is the sense of smell. Nocturnal moths for example can use olfaction to discriminate and locate their hosts, and to judge its quality by the plant specific odor blend from an environmental bouquet of 1700 known volatiles emitted by plants (Knudsen and Gershenzon 2006; Allmann et al. 2013). In another example, ant species lay trails of chemical pheromones to lead their nest mates to food sources (Attygalle and Morgan 1985). Ants are also able to locate small food items in almost featureless environments, such as salt deserts, based on their sense of smell (Buehlmann, Hansson and Knaden, in preparation).

The crustacean sense of smell has been extensively studied mainly in decapods. Besides its involvement in orientation and foraging (Bell 1906; Guenther et al. 1996) olfaction plays an important role in intraspecific communication (Breithaupt and Atema 1993; Breithaupt and Thiel 2011). In hermit crabs, the sense of smell is employed to localize distant food items (Rittschof and Sutherland 1986; Thacker 1998; Stensmyr et al. 2005), empty snail shells (Small et al. 1994) and water. Moreover, crustaceans are able to discriminate between fresh- and salt water (Vannini and Ferretti 1997).

Phylogenetic context of hermit crabs

Crustaceans are the phylogenetic sistergroup of insects (Richter 2002). According to fossil records, insects split from the last common ancestor and left water in the very late Silurian (Grimaldi 2010). Members of at least five crustacean lineages independently succeeded in the transition from water to land as well, among them isopods (Lindqvist 1972), amphipods (Hurley 1968), astacids (Richardson 2007), brachyurans (Hartnoll 1988) and anomurans (Hartnoll 1988; Greenaway 2003). Within the anomurans, members of the taxon Paguroidea are called “hermit crabs” for they developed the ability to occupy empty snail shells to protect their pleon. This thesis is focused on two hermit crab species, each representing one

habitat. One species, belonging to the subtaxon Paguridae, is the common hermit crab *Pagurus bernhardus* (L. 1785). It is strictly aquatic, and inhabits the Atlantic Ocean around the European coast as well as the Mediterranean sea, the British Channel and Western Baltic Sea. The second species investigated in this thesis is the Caribbean hermit crab *Coenobita clypeatus* (Herbst 1791), subtaxon Coenobitidae. This subtaxon includes two genera which developed a fully terrestrial lifestyle (McLaughlin 2007): the coconut crab *Birgus latro* and 15 species of shell-carrying land hermit crabs (genus *Coenobita*). Larval stages of all land hermit crabs are aquatic while the juveniles and adults are fully terrestrial. Fossils date the transition from water to land to the lower Miocene, *ca* 20 mya resulting in various adaptations to the new habitat (Glaessner 1969; Bliss and Mantel 1968). The habitat of *Coenobita clypeatus* are the shores of the Caribbean mainland and islands, where populations of few to several dozen animals live either at the rim of vegetation on the beach or in lucid forests some distance inland. The last common ancestor of *Pagurus* and *Coenobita* dates approximately 173 Mya (Bracken-Grissom et al. 2013).

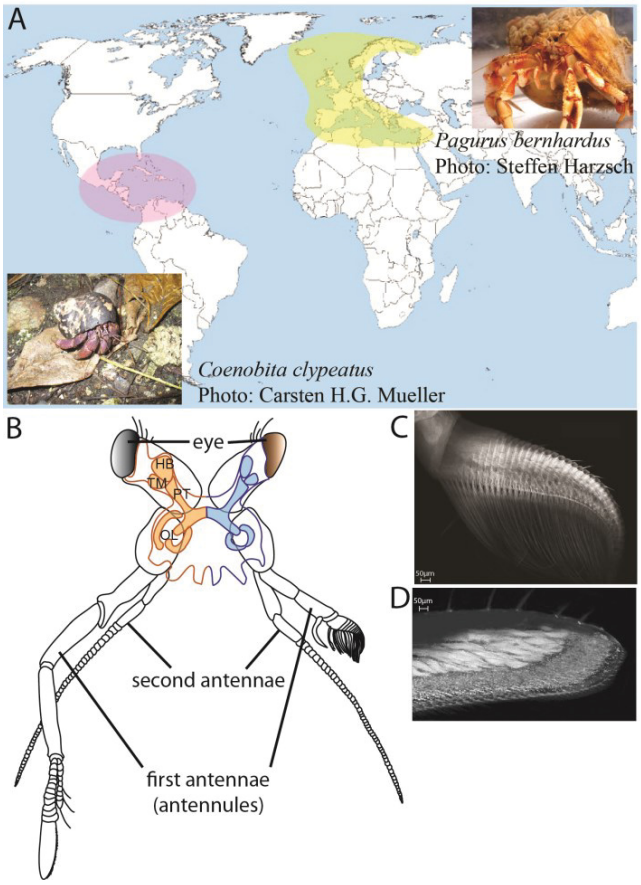


Figure 2: Species investigated; A: areas of origin, yellow *P. bernhardus* (upper right), red: *C. clypeatus* (lower left); B: combined scheme of both crabs heads and head appendices, left: *C. clypeatus*, right: *P. bernhardus*, olfactory relevant brain centers colored OL=Olfactory lobe, PT=Protocerebral Tract, TM=Terminal medulla, HB: Hemiellipsoid body; C/D last antennular segment; staining: Glutaraldehyd C: *P. bernhardus*, D: *C. clypeatus* (C/D from (Koczan 2012))

Reception and processing of chemical information

Reception of odors requires the contact between the odor bearing medium and a receptor embedded in the membrane of the receiving dendrite. This contact between the receptor and its respective ligand initiates a signal transduction cascade that can be modulated and ultimately leads to a signal that is forwarded to higher processing centers of the brain.

Insects

In insects the first contact between odor and receptor occurs after the odorant enters the sensillum lymph via pores in the cuticle (Steinbrecht 1997). Odorant binding proteins (OBPs) are present in this lymph space in high amounts and are believed to ferry otherwise water insoluble odorants to the dendrite surface where they bind to their respective receptors (Steinbrecht 1998). Various OBPs have been identified in insects but not all of them are restricted to chemosensory tissues and thus may have other functions.

Insect sensilla functionally differ by their individual OSNs as each expresses a distinct odor specific receptor that determines the spectrum of their ligands (Stensmyr 2003). Two distinct receptor classes are known to be involved in insect chemosensing. One is the superfamily of Olfactory Receptors (ORs) and Gustatory Receptors (GRs). While the latter ones are mainly responsible for sensing chemicals upon contact, except one GR sensing Carbon dioxide, the ORs represent one “nose” of insects. The second receptor family are Ionotropic Receptors (IRs). They are ligand gated ion channels derived from the ancient family of ionotropic Glutamate Receptors (iGluRs) and can be retraced to early Protostomian origin (Croset et al. 2010). The highest resemblance to their ancestral iGluRs and conservation across species can be found in the two IR co-receptors IR25a and IR8a (Benton et al. 2009; Croset et al. 2010). Both keep the aminoterminal domain (ATD) (see figure 3) that is lacking in all other IRs and have a broad and partially overlapping expression pattern in coeloconic sensilla of the *Drosophila* antennae (Benton et al. 2009). In contrast to IR8a, which seems to have first emerged in the insecta, IR25a has homologues in distantly related species such as snails (Croset et al. 2010). The IR family is further subdivided into the “antennal” IRs, a family of receptors evidently expressed in insect antennae, and

"divergent" IRs. The latter were identified in arthropod genomes but are not expressed in insect antennae. The tuning width of odorant specific antennal IRs differs between individual receptors. In *Drosophila*, neurons expressing IR31a are activated by 2-oxopentanoic acid while IR41a expressing neurons broadly respond to all 18 amines tested (Silbering et al. 2011).

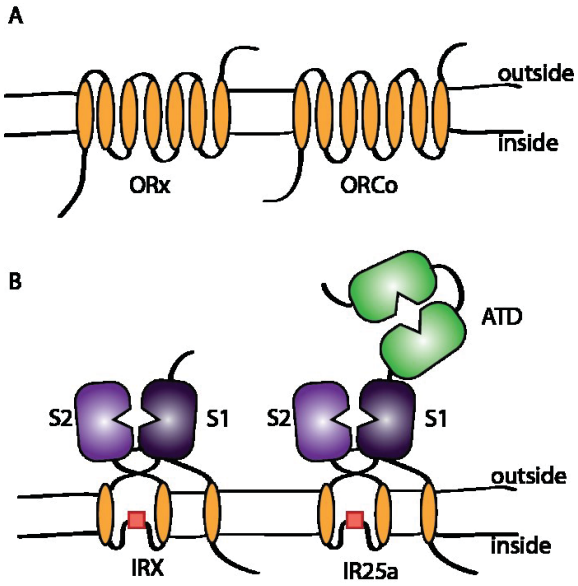


Figure 3: Scheme drawings of arthropod chemosensory receptors, A: odor specific OR and co-receptor ORCo, B: odor specific IR and coreceptor IR25a; ATD: aminoterminal domain; S1/S2 subunits of the ligand binding domain

Studies of the odor spectrum covered by each of the Drosophila "noses" discovered a partially overlapping but also partially distinct set of compounds eliciting responses in the respective OSNs. Best ligands to IR expressing OSNs are acids, alcohols, amines, esters and ketones, with most agonists chemically belonging to amines and carboxylic acids (Silbering et al. 2011).

Olfactory information gathered by the OSNs is forwarded by the antennal nerve to the first olfactory processing center of the insect brain, the paired antennal lobes (AL). The AL is subdivided into distinct spherical compartments, the glomeruli, each receiving information from an individual population of OSNs expressing one odor specific OR or IR (Gao, Yuan, and Chess 2000; Vosshall, Wong and Axel 2000; Scott et al. 2001; Benton et al. 2009; Silbering et al. 2011).

Crustaceans

Morphologically the crustacean sense of smell is situated on their first antennae, also called antennules. In *Pagurus* the antennules are much shorter than the second pair of antennae and bear ca. 700 uniform aesthetascs of long and hair-like shape on the last antennular segment (Koczan 2012). Each is innervated by ca. 300 bipolar OSNs with cell bodies arranged in approx. 270 clusters in the antennular lumen (Koczan 2012; Ghiradella, Case and Cronshaw 1968a). Crustacean aesthetascs are always poreless but in Coenobitids a thinner cuticle is exposed to the outside than to the protected side (Ghiradella, Case and Cronshaw 1968b; Stensmyr et al. 2005). It is assumed that odor molecules diffuse through the thin cuticle to reach the receptor lymph and the dendrites as the cuticle allows the diffusion of staining agents (Ghiradella, Case and Cronshaw 1968a). Presumably as a result of terrestrialization, the basal joints of Coenobita antennules are elongated to allow the flagella to be extended to the ground in normal walking postion (Ghiradella, Case and Cronshaw 1968b). The last antennular segment bears an erinaceous array of ca. 350 aesthetascs which are short and blunt in shape (Koczan 2012; Krång et al. 2012). The OSNs send about the same number of dendrites to one aesthetasc as in *Pagurus,* with their cell bodies arranged in approximately 140 spindle-like complexes with an average of 150 cell bodies each. It is not yet known how many aesthetascs are innervated by one OSN cell body complex and how the pattern of OSN types is distributed within the complexes and comparing different complexes. In contrast to *Pagurus* and other aquatic decapods, where OSN dendrites branch in different places including within the aesthetasc, the basal bodies of *Coenobita* are situated well under the cuticle of the aesthetasc pad in a distinct vacuole (Ghiradella, Case and Cronshaw 1968b). Studies in the lobsters *Homarus americanus* and *Panulirus argus* revealed chemoreceptive cells of the antennules to be narrowly tuned to single substances and either purinergic, taurinergic or glutamatergic, with ligand binding leading to a net inward current (Johnson and Atema 1983; Ache and Derby 1985; Carr, Ache and Gleeson 1987). Single odors can thereby excite one and inhibit another type of OSNs by second messenger pathways either including cAMP or inositol 1,4,5-trisphosphate (IP3) and activate two different channel types, a non-selective cation channel and a Ca^{2+} channel with additional signal modulation by a third inositol 1,3,4,5-tetrakisphosphate gated channel (Dionne and Dubin 1994; Boekhoff et al. 1994). Olfactory transduction and signaling in

lobsters seems to include also G-protein mediated signaling, as a single $G\alpha_q$ mRNA type is expressed in aesthetasc dendrites (McClintock et al. 1997). A histamine gated chloride channel ubiquitously expressed in somata of lobster OSNs indicates further the presence of a yet undescribed regulatory or feedback process (McClintock and Ache 1989; Bayer et al. 1989). Afferent axons from the OSN somata rise into the antennal nerve and converge into the first olfactory processing center, the olfactory lobe (OL) consisting of numerous conical shaped glomeruli (Schmidt and Ache 1996). In the aquatic *P. bernhardus*, reconstructions estimate 560 glomeruli organized in one lobe whereas in the terrestrial *C. clypeatus* the OL is enlarged and composed of two sublobes with a total of approximately 1000 glomeruli (Koczan 2012; Harzsch and Hansson 2008). As in hymenopterans, the sheer number of glomeruli makes a one-to-one correlation in receptor count appear unlikely. The glomeruli number or OL size on its own is independent of anatomy, taxon and habitat, whereas the convergence ratio between aesthetasc count and number of glomeruli might have functional meaning in terms of habitat and sensitivity (Beltz et al. 2003). The OL glomeruli are interconnected by thousands of local interneurons, with subpopulations innervating different regions of the glomeruli, and send projection neurons to the accessory lobe, a higher order integration center (Harzsch and Hansson 2008). Olfactory information terminates in both main neuropils of the lateral protocerebrum: the medulla terminalis and the hemiellipsoid body (Sullivan and Beltz 2001; Harzsch and Hansson 2008).

Compared to insects, molecular information about chemoreceptors in crustaceans is sparse. The only crustacean genome sequenced so far is *Daphnia pulex*, the water flea (Colbourne et al. 2011). This very basal crustacean is belonging to the order of branchiopoda, the sistergroup of all other crustaceans. Based on its genome, three CSPs, two “antennal” IRs and an extended set of GRs were found, but no ORs (Penalva-Arana, Lynch and Robertson 2009; Croset et al. 2010; Vieira and Rozas 2011). Additionally, 83 IRs of the divergent subtype were identified. Nevertheless, no details about the expression of those genes are available and the functional characterization of the *D. pulex* antennulae is incomplete. An EAG study revealed that L-Arginine and to a far lesser amount L-Aspartic acid but not L-Alanine, L-Leucine or L-Serine elicit responses (Simbeya et al. 2012). In contrast to this, a large screen of putative olfactory cues in *C. clypeatus* identified a panel of carboxylic acids, aldehydes, amines and alcohols to be olfactory active but did not reveal responses to any

aminoacids (Krång et al. 2012). A recent study identified two IRs of the divergent subtype along with IR25a and IR93a homologues in the antennules of the spiny lobster *P. argus* (Corey et al. 2013). Together with the knowledge about ligand spectra of insect olfactory IRs and the chemical classes of active odors in *Coenobita*, IRs are likely candidate chemoreceptors.

2. Aims and objectives of the thesis

This thesis aims at understanding the peripheral mechanisms of chemosensing in crustaceans and how it was adapted during the transition from water to land. A special focus is put on the identification of chemosensory relevant receptors as they are the first step in chemical detection and odor discrimination. Therefore it is based on next generation sequencing techniques to investigate and compare the genetic makeup of a terrestrial *(Coenobita clypeatus)* and an aquatic hermit crab *(Pagurus bernhardus)*. Furthermore this study employs methods of electron and light microscopy as well as fluorescent RNA in situ hybridization and immunohistochemistry to investigate and describe the inner and outer structures of the hermit crab olfactory organ. In addition the soluble proteome of the *C. clypeatus* antennules is characterized in order to gain insights into the function and involvement of glandular structures found in the antennules.

3. Overview of manuscripts

Manuscript 1

The hermit crab's nose – antennal transcriptomics

Katrin C. Groh, Heiko Vogel, Marcus C. Stensmyr, Ewald Grosse-Wilde, Bill S. Hansson

Published in: *Frontiers in Evolutionary Psychology and Neuroscience* (doi:10.3389)

In the first manuscript we investigate the molecular biology of the hermit crab olfactory organ based on transcriptomes. Hermit crabs successfully conquered terrestrial habitats and are able to sense and evaluate odor cues transported in air just as well as their aquatic relatives. Since molecular biology data about crustacean olfaction is elusive, we investigated and compared antennal transcriptomes from an aquatic, *Pagurus bernhardus*, and a terrestrial hermit crab, *Coenobita clypeatus,* to investigate the differences in the genetic makeup of their olfactory organs. However, the set of expressed genes in both species is highly similar. The likely candidates for chemosensory receptors are similar sets of ionotropic receptors in both species.

Built on an idea conceived by all authors.

Experimental design: EGW, HV, MCS, BSH

Experimental performance: KCG (85%), EGW

Data analysis: KCG (70%), EGW, HV

Manuscript writing: KCG (60%), EGW, BSH

Manuscript 2

Expression of ionotropic receptors in terrestrial hermit crab's olfactory sensory neurons

Katrin C. Groh, Marcus C. Stensmyr, Ewald Grosse-Wilde, Bill S. Hansson

In preparation for *Frontiers in cellular Neuroscience*

In the second manuscript we aim for a deeper understanding of the repertoire and expression of *Coenobita clypeatus* chemosensory receptors by investigating an RNAseq based transcriptome and employing fluorescent RNA in *situ* hybridizations. The overall genetic makeup is highly similar to out earlier findings except the average statistical data confirming a better coverage and higher length of the identified transcripts. We identified additional ionotropic receptors in the antennules and show their expression and distribution in the cell bodies of the olfactory sensory neurons.

Built on an idea conceived by all authors.

Experimental design: EGW, KCG (25%), MCS, BSH

Experimental performance: KCG (100%)

Data analysis: KCG (80%), EGW

Manuscript writing: KG (80%), EGW, BSH

Manuscript 3

Morphology and histochemistry of the aesthetasc-associated epidermal glands in terrestrial hermit crabs of the genus Coenobita (Decapoda: Paguroidea)

Oksana Tuchina, Katrin C. Groh, Giovanni Talarico, Carsten H.G. Müller, Natalie Wielsch, Yvonne Hupfer, Aleš Svatoš, Ewald Grosse-Wilde and Bill S. Hansson

Published in: *PLOS ONE* (doi: 10.1371)

The third manuscript describes the structure and presumable functions of *Coenobita* antennular glands which are exocrine organ of the rectocanal type, derived from rosette-type tegumental glands and produce the mucosal substance covering the aesthetascs. This mucus might be crucial for the functionality of the olfactory organ of terrestrial hermit crabs. Two microscopically distinguishable cell types are associated with the glandular complex, forming two types of glands. A subpopulation of the glandular cells produces a CUB-serine protease which is likely excreted to the surface and might play a role in the antimicrobial defense. Data derived from a proteomic approach further support an immunological function of those glandular complexes.

Built on an idea conceived by all authors.

Experimental design: OT, KG (25%), GT, NW, AS, EGW, BSH

Experimental performance: OT, KG (25%), GT, YH

Data analysis: OT, KG (28%), GT, NW, YH

Manuscript writing: OT, KG (5%), GT, CM, NW

frontiers in NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 21 January 2014
doi: 10.3389/fnins.2013.00266

The hermit crab's nose—antennal transcriptomics

Katrin C. Groh[1], Heiko Vogel[2], Marcus C. Stensmyr[1], Ewald Grosse-Wilde[1]* and Bill S. Hansson[1]

[1] *Department of Evolutionary Neuroethology, Max Planck Institute for Chemical Ecology, Jena, Germany*
[2] *Department of Entomology, Max Planck Institute for Chemical Ecology, Jena, Germany*

Edited by:
Karol Osipowicz, Jefferson Neuroscience Hospital, USA

Reviewed by:
Don A. Baldwin, Pathonomics, LLC, USA
Hans-Georg Breitinger, The German University in Cairo, Egypt
Jawaid Ahsan, Central University of Bihar, India

****Correspondence:***
Ewald Grosse-Wilde, Department of Evolutionary Neuroethology, Max Planck Institute for Chemical Ecology, Hans-Knoell-Str. 8, 07745 Jena, Germany
e-mail: grosse-wilde@ice.mpg.de

In the course of evolution, crustaceans adapted to a large variety of habitats. Probably the most extreme habitat shift was the transition from water to land, which occurred independently in at least five crustacean lineages. This substantial change in life style required adaptations in sensory organs, as the medium conveying stimuli changed in both chemical and physical properties. One important sensory organ in crustaceans is the first pair of antennae, housing their sense of smell. Previous studies on the crustacean transition from water to land focused on morphological, behavioral, and physiological aspects but did not analyze gene expression. Our goal was to scrutinize the molecular makeup of the crustacean antennulae, comparing the terrestrial *Coenobita clypeatus* and the marine *Pagurus bernhardus*. We sequenced and analyzed the antennal transcriptomes of two hermit crab species. Comparison to previously published datasets of similar tissues revealed a comparable quality and GO annotation confirmed a highly similar set of expressed genes in both datasets. The chemosensory gene repertoire of both species displayed a similar set of ionotropic receptors (IRs), most of them belonging to the divergent IR subtype. No binding proteins, gustatory receptors (GRs) or insect-like olfactory receptors (ORs) were present. Additionally to their olfactory function, the antennules were equipped with a variety of pathogen defense mechanisms, producing relevant substances on site. The overall similarity of both transcriptomes is high and does not indicate a general shift in genetic makeup connected to the change in habitat. IRs seem to perform the task of olfactory detection in both hermit crab species studied.

Keywords: crustacea, antennules, olfaction, molecular evolution, genomics

INTRODUCTION

Crustaceans successfully conquered a variety of habitats, including backwater and freshwater as well as marine and terrestrial habitats. At least five lineages of crustaceans independently succeeded in the transition from water to land (Bliss and Mantel, 1968; Powers and Bliss, 1983). Such substantial changes in environment require extensive adaptation, for example regarding metabolism, water and ion balance, and behavior. Another obviously affected area are the acoustic-, visual and chemical senses, due to the differences in chemical and physical properties of the medium conveying the respective stimuli. In crustaceans antennules and antennae are important organs for sensory tasks; both can for example detect mechanical stimuli. Additionally, the antennules are the main chemosensory organ (Bush and Laverack, 1982; Cate and Derby, 2002). In terms of chemosensation, lobsters are the primary decapod models. Lobster antennulae are the olfactory organs, while the sense of taste is located on their walking legs (Atema, 1977; Derby and Atema, 1982; Johnson et al., 1988). The most numerous sensillum type on the antennules are aesthetascs, unimodal chemosensory sensilla located on the last antennular segment (Ghiradella et al., 1968a). Aesthetascs each house approximately 300 olfactory sensory neurons (OSNs) and are the place of odor detection (Gleeson, 1982). In the marine hermit crab *Pagurus bernhardus* the aesthetascs are long and slender (Hansson et al., 2011) while in its terrestrial relative *Coenobita clypeatus* they are short and blunt (Krång et al., 2012). This change is probably an adaptation to function in air rather than water, comparable to its terrestrial relatives *Coenobita compressus* (Ghiradella et al., 1968a,b) and *Birgus latro* (Stensmyr et al., 2005). The OSNs from all aesthetascs converge into one antennal nerve leading to the olfactory lobe (OL), the primary central processing center of olfactory input. The OL is composed of column-like glomeruli, approximately 560 in *P. bernhardus* and 1000 in *C. clypeatus* (Koczan, 2012). In the latter the OL is enlarged, consists of two sublobes and dominates the brain in size, while the volume of visual or mechanosensory processing centers is similar to their aquatic relatives (Harzsch and Hansson, 2008). Investigations of the terrestrial hermit crab *Birgus latro*, a close relative of *C. clypeatus*, suggested that the adaptation of the olfactory organs is insect-like (Rittschof and Sutherland, 1986; Greenaway, 2003; Stensmyr et al., 2005).

Contrary to insects, information on the molecular biology of the crustacean olfactory system is sparse. In contrast, the insect chemosensory system is very well-described. Here, relevant genes include two classes of soluble proteins found in the sensillum lymph; odorant binding proteins (OBPs) and chemosensory proteins (CSPs); while CSPs have been reported in all groups of Arthropoda, OBPs are limited to Hexapoda (Vieira and Rozas, 2011). As regards the relevant receptors, one chemoreceptor superfamily containing olfactory receptors (ORs) and gustatory receptors (GRs), and a receptor class named ionotropic receptors

(IRs) are known [for reviews see Vieira and Rozas (2011), Silbering et al. (2011)]. GRs which also have been identified in all major groups of Arthropoda are considered to be ancestral to the Hexapoda specific ORs (Robertson et al., 2003). GRs are expressed in insect antennae but also in other head and body appendages (Montell, 2009); GR-expressing neurons of insects respond to tastes like sugar or bitter compounds, but also to CO_2 (Kwon et al., 2006; Robertson and Kent, 2009). IRs are currently considered the most ancient chemoreceptors. They derived from the ionotropic glutamate receptors early in the protostomian lineage and are comprised of two groups based on expression data (Benton et al., 2009). While the antennal IRs of insects are involved in the detection of e.g., acids, amines, aldehydes, and alcohols, the divergent IRs are not expressed in insect olfactory organs and their function is so far unknown (Benton et al., 2009; Ai et al., 2010; Croset et al., 2010; Silbering et al., 2011). The most ancient IRs known to date are IR25a and IR93a, both belonging to the antennal IR subgroup. Insect odor specific IRs are always expressed together with IR25a or its homolog IR8a, which functions as a coreceptor in IR-expressing dendrites of *Drosophila melanogaster* (Croset et al., 2010).

Daphnia pulex is the only crustacean with near-complete genome data available. The data contains CSPs, GRs, and IRs, with most of the latter belonging to the divergent IR subgroup (Penalva-Arana et al., 2009; Croset et al., 2010; Vieira and Rozas, 2011; Kulmuni and Havukainen, 2013). Further indication of IR based olfaction in crustaceans was obtained from lobsters, identifying the homologs of IR25a, IR93a, and IR8a and showing their expression in OSN clusters (Hollins et al., 2003; Corey et al., 2013).

As mentioned, the transition to land has consequences for other functions beyond the chemical sense. The exposed position of crustacean antennae necessitates protection from injury and pathogens. While the cuticle, mechanosensors and withdrawal reflex account for the first, the second has not been studied in this tissue yet. Nevertheless, besides fruit flies, the immunosystem of decapod crustaceans is the most intensely studied among arthropods. The first step of pathogen defense is non-self recognition; in invertebrates mediated for example through pattern recognition proteins such as C-type lectins. Lectins recognize and bind sugars to agglutinate cells which are in turn recognized and destroyed by the innate immune system, involving the prophenoloxidase (proPO) activating system and apoptosis (Soederhall and Cerenius, 1998; Cerenius et al., 2010). Another important factor of the arthropod immune system are antimicrobial peptides (AMPs). The expression of insect AMPs is in some cases upregulated upon microbial challenge and subsequently negatively affects microbial growth. In crustaceans many kinds of AMPs have been described, including both constitutively expressed and inducible ones (Han-Ching Wang et al., 2010). There is indication that their regulation involves the Toll or the imd pathway, similar to the situation found in insects (Cerenius et al., 2010; Han-Ching Wang et al., 2010). Possible components of both pathways have been identified in shrimp species, including *Toll, imd, spätzle, relish, and dorsal* (Cerenius et al., 2010), but details about the extent and specificity of AMP regulation by either pathway remain to be investigated.

Previous studies on the crustacean transition from water to land considered morphological, behavioral and physiological aspects, but did not include analysis of gene expression (Bliss and Mantel, 1968; Powers and Bliss, 1983; Greenaway, 2003). Our goal was to scrutinize molecular adaptation in the antennae of terrestrial hermit crabs. Therefore, we selected adults of the hermit crab species *Coenobita clypeatus* (Herbst 1791; Paguroidea, Coenobitidae) as a representative of the genus Coenobitidae, which has planktonic larvae and a fully terrestrial adulthood for comparison of their molecular makeup with adults of the marine species *Pagurus bernhardus* (L. 1758, Paguroidea, Paguridae). According to recent investigations based on a combined analysis of molecular, morphological, and fossile data, the ancestral lineages of both species split approximately 173 Mya ago (Bracken-Grissom et al., 2013). We generated antennal transcriptome data using 454 and Solexa/Illumina technology. The data was assembled and investigated from the abstract level of GO annotation. Furthermore, we employed BLAST searches based on homology comparison and HMM profile prediction to screen selectively for genes involved in chemo- and mechanosensing, neuronal signaling, and immune response, and interpret the results in the context of known and prospective abilities of crustaceans.

MATERIALS AND METHODS

ANIMAL COLLECTION AND DISSECTION

Pagurus bernhardus, Coenobita clypeatus

Specimens of *P. bernhardus* were procured from the Biologische Anstalt Helgoland (Germany) and *C. clypeatus* specimens were ordered from Peter Hoch Import—Export Waldkirch (Germany). Both examined species are neither endangered nor protected. Adult animals of both sexes were dissected at the SLU Alnarp (Sweden); specimens were cold anesthetized and antennulae cut off using scissors pooling the antennules of each species. Dissected antennulae were shipped to Jena (Germany) on dry ice for RNA extraction.

SEQUENCING, READ CLEANUP, AND DE NOVO ASSEMBLY

For RNA extraction tissues were cooled further over liquid nitrogen. The frozen tissue was transferred to a liquid nitrogen cooled mortar and ground. The homogenate was covered with 1 ml TRI reagent (Sigma-Aldrich, St. Louis, MIS). Further steps were performed according to the manufacturer's instructions, replacing choloroform with 1-bromo-3-chloro-propane as recommended by Sigma-Aldrich to lower toxicity and for better phase-separation. For ethical reasons we restricted the number of specimens according to the minimal amount of RNA required. Including both sexes this meant 10 pairs of antennules for *C. clypeatus* and 5 pairs for *P. bernhardus*. Two micrograms of total RNA preparations were sent to Evrogen, Moscow (Russia) for production of normalized cDNA. The normalized cDNA was afterwards sequenced by AGOWA, Berlin (Germany) by 454 sequencing on an GS FLX+ sequencer, for 400 bp single read length. An additional RNA sample was not normalized, reverse transcribed and sequenced by Solexa/Illumina for 76 bp single end reads on an Illumina Genome Analyzer II. Bustard 1.3.2 was used for basecalling and Firecrest was employed for offline image analysis. Raw sequencing data is available for download

(EBI/ENA Study acc. number ERP002374, http://www.ebi.ac.uk/ena/data/view/ERP002374). Data was screened for contaminants and linker/adapter sequences, followed by *de novo* hybrid assembly of the sanitized reads in CLC genomics workbench V 5.5. Reads from both sequencing techniques were used simultaneously without removal of duplicates. Thorough contaminant cleanup was made necessary because of high contamination of raw 454 data by reads originating from various environmental taxa. Bioinformatic statistics were performed using scripts provided at http://manuals.bioinformatics.ucr.edu by Thomas Girke updated 28th of January 2010 and a script by Joseph Fass revised 2010 (c) 2009 The Regents of University of California, Davis Campus.

Manduca sexta

As a reference for arthropod antennal transcriptome data we chose the 454 sequenced antennal transcriptome of *M. sexta* published by Grosse-Wilde (Grosse-Wilde et al., 2011). We performed a new assembly of the data using CLC 5.5. For better comparativeness we used the same parameters as for the hermit crab datasets.

ANNOTATION

We created BLAST2GO databases based on each dataset and performed homology searches after dynamic translation (BLASTX) against non-redundant databases (National Center for Biotechnology Information, NCBI), using the default cutoff parameters of BLAST expectation value (1.0E-3), InterProScan and assigned KEGG maps. Annotation was based on an *E*-Value—Hit-Filter of 1.0E-6. Additionally we carried out BLAST searches based on HMM profile prediction (*E*-value cutoff 1.0E-3) and homology comparison (*E*-value cutoff 1.0E-1) of selected reference gene sequences against the datasets. In the following, "EST" will refer to annotated contigs.

BIOINFORMATICS

To compare the transcriptomes, B2GO combined graphs were drawn for all datasets and in all the three categories: cellular component, molecular function, and biological process. The cutoff of sequence count was set according to the maximum of possible nodes. If a term turned up in at least one species but did not in one or both of the others, annotation was checked manually and the number of sequences was added as necessary. We calculated the percentage of ESTs assigned to one term to the total count of ESTs assigned to terms on annotation level 2 or 3, respectively, by the following formula: (number of ESTs assigned to the term of interest $*$ 100)/(number of ESTs assigned to terms on level x). For further comparison we used R 2.15.2 (http://cran.r-project.org/bin/windows/base/old/2.15.2/) to plot the respective charts, test, and compare the categories in the species and the ratios or presence and absence of terms. Dendrograms were compiled using the MUSCLE alignment tool (Edgar, 2004) followed by FastTree2 dendrogram calculation (Price et al., 2010). Adobe Illustrator CS5 was employed to compile the figures.

RACE PCR

Selected candidate gene fragments were extended using SMARTer and Marathon RACE-PCR kits (Clontech) following the respective manufacturer's instructions and specifications. Gene specific primers were designed with the online input version of primer3 (http://frodo.wi.mit.edu/).

RESULTS

SEQUENCING AND ASSEMBLY

For transcriptome data generation, antennal DNA samples from *Coenobita clypeatus* and *Pagurus bernhardus* were sequenced using Roche 454 and Solexa/Illumina technologies. Reads were sanitized and 454 data was additionally cleaned from reads associated to contaminants like algae, fungi, and bacteria, the cutoff being set at $E < 10–20$ (**Table 1**).

Approximately 30% of the raw 454 sequencing data of *P. bernhardus* was removed, leaving 270,565 reads to proceed. The raw 454 sequencing data from *C. clypeatus* additionally contained internal linker sequences and was therefore split and sanitized. Five reads were split into 4 parts, 567 into 3 parts and 31,766 into 2 parts. After sanitization and contaminant cleanup 432,240 single reads remained. Assembly resulted in 34,227 contigs with an average length of 464 nucleotides for *C. clypeatus* and 9788 contigs with an average length of 415 for *P. bernhardus* (**Figure 1**). A higher total contig count in *C. clypeatus* was assembled from a smaller number of sequencing reads than in *P. bernhardus* (**Figure 1A**). Nevertheless, the N50 value is similar and the distribution of assembly coverage shown in **Figure 1B** is almost identical. The N50 values as well as the contig distribution in both hermit crab species are highly similar to *M. sexta* transcriptome (**Figures 1B,C**). Summary assembly statistics of the two datasets confirmed a quality similar to previously published datasets on insect olfactory tissues, allowing comparison. To eliminate problems caused by underrepresentation of critical sequences, contigs with an average coverage of less than 5 were ignored in the BLAST2GO annotation but included for survey of specific gene families, performing manual curation. The datasets for the GO comparison of both transcriptomes finally yielded 13,942 unique contigs for *C. clypeatus* and 7649 contigs for *P. bernhardus*, respectively (**Figure 1C**). For quality assessment we included a reassembly of a previously published antennal transcriptome dataset from the insect *Manduca sexta* containing contigs and single reads above a cutoff of 300 nucleotides.

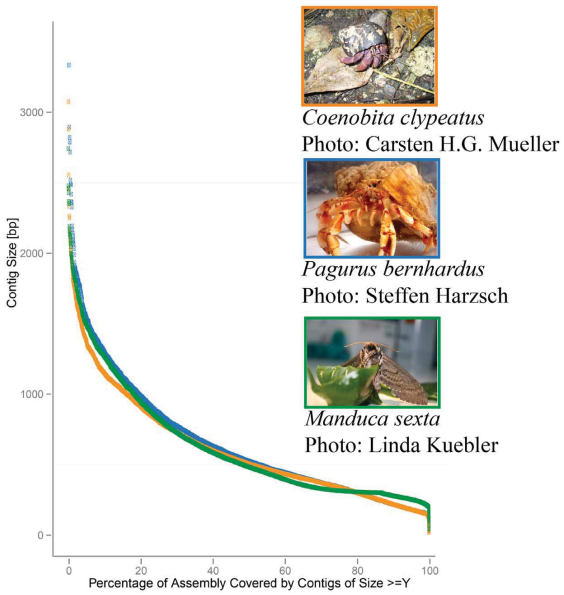
Table 1 | Number of contaminating sequences detected in and removed from 454 sequencing reads of both species; separated by taxids (ncbi).

Taxid	Taxon name	*P. bernhardus*	*C. clypeatus*
2	Bacteria	1316	357
2157	Archaea	231	10
2736	*Verrucomicrobium spinosum*	2057	1234
3041	Chlorophyta	2469	2241
4751	Fungi	15,108	1592
33,634	Stramenopiles	294	458
38,254	Glaucocystophyceae	10,637	2742
554,915	Amoebozoa	1126	536
Remaining reads		270,565	432,240

A

	Reads 454	Reads Illumina	Reads assembled (%)	Contigs	Average length	N50
Ccly	432,240	8,729,050	4,963,596 (54%)	34,227	464	484
Pber	270,565	10,686,804	5,621,889 (51%)	12,409	415	463
Msex	156,922	-	100,328 (64%)	10,752	506	573

B



C

	ESTs	N50	%GC
Ccly	13,942	507	41
Pber	7,649	527	42.5
Msex	14,437	485	35.8

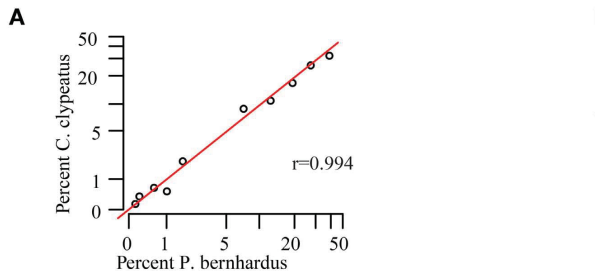
FIGURE 1 | Summary statistics sequence assembly. (A) Read numbers, contig numbers, and average values of CLC assembly **(B)** contig length distribution, orange: *C. clypeatus*, blue: *P. bernhardus*, green: *M. sexta* **(C)** assembly statistics final datasets, *P. bernhardus*/*C. clypeatus*: contigs with an average coverage above 5, *M. sexta*: contigs above 200 and single reads above 300 nucleotides.

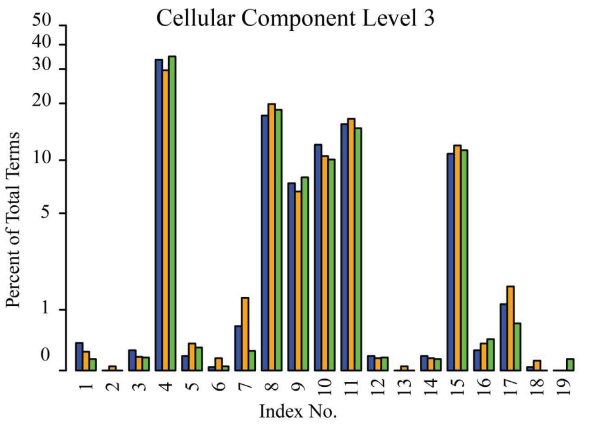
GO ANNOTATION

GO classification provides a general and transferable vocabulary for abstract description of functions and processes present within a given dataset (Ashburner et al., 2000). We performed GO annotation of the transcriptome data using BLAST2GO as described in Conesa et al. (2005). Again we included *M. sexta* data for comparison, performing a reanalysis. A total of 3453 contigs (25%) in the *C. clypeatus* dataset and 2505 contigs (33%) in *P. bernhardus* yielded a suitable BLAST result. Domain identification using InterProScan (IPS) yielded matches for 8369 contigs (60%) of *C. clypeatus* and 4717 contigs (62%) of *P. bernhardus*. Taken together this indicates a large number of sequences with no clear ortholog, but identifiable functional domains. For comparison, reanalysis of *M. sexta* yielded 8771 contigs with BLAST hit (61%) and 7943 contigs (55%) with IPS matches.

For distinction purposes we refer to the dataset generated from the *Pagurus bernhardus* transcriptome as "Pber," to the *Coenobita clypeatus* dataset as "Ccly" and to the *Manduca sexta* dataset as "Msex."

Since the function of the antennal tissues is probably very similar but not identical in all three species, we assigned KEGG maps to compare the representation of their main molecular functions. As expected they are equally well-represented in the three datasets (data not shown). Based on the GO terms assigned to contigs we compared the species regarding different categories and on different annotation levels (full list in Supplementary Table S1). One potential problem of GO annotation affecting comparisons is an over- or underrepresentation of ESTs resulting in a shift of the EST ratio assigned to specific terms on a distinct level. To address this issue we compared the representation of terms between the datasets and calculated the Pearson coefficient as a measure of similarity. If annotation of the two compared datasets is not adversely affected by this issue the correlation coefficient should be close to 1 for the category of "Cellular Component," since it is reasonable to assume similar term distribution here. Comparing Ccly and Pber the coefficient is 0.994 on level 2 and 0.991 on level 3, respectively (**Figures 2A,B**), confirming comparability. Therefore, we assume that there is no overall shift in





D

Index	GO.ID	Term	Pber	Ccly	Msex
1	GO:0070161	anchoring junction	0.37	0.24	0.14
2	GO:0048046	apoplast	0.00	0.05	0.00
3	GO:0005911	cell-cell junction	0.26	0.17	0.16
4	GO:0044464	cell part	33.49	29.62	34.79
5	GO:0048475	coated membrane	0.18	0.36	0.30
6	GO:0044420	extracellular matrix part	0.04	0.15	0.05
7	GO:0044421	extracellular region part	0.66	1.29	0.25
8	GO:0043227	membrane-bounded organelle	17.26	19.84	18.52
9	GO:0044425	membrane part	7.46	6.69	8.03
10	GO:0043228	non-membrane-bounded organelle	12.10	10.53	10.08
11	GO:0044422	organelle part	15.58	16.61	14.81
12	GO:0019867	outer membrane	0.18	0.15	0.16
13	GO:0032992	protein-carbohydrate complex	0.00	0.05	0.00
14	GO:0032993	protein-DNA complex	0.18	0.15	0.14
15	GO:0043234	protein complex	10.82	12.01	11.31
16	GO:0044456	synapse part	0.26	0.36	0.43
17	GO:0031982	vesicle	1.13	1.61	0.71
18	GO:0044423	virion part	0.04	0.12	0.00
19	GO:0031594	neuromuscular junction	0.00	0.00	0.14

FIGURE 2 | Distribution of GO terms in the category Cellular Component. (A) Cellular Component level 2, **(B)** Cellular Component level 3. Red line: predicted linear slope. *r* = Pearson coefficient, **(C)** Barplot Cellular Component level 3, all terms; blue: Pber, orange: Ccly, green: Msex; **(D)** chart index table.

the category representation caused by over- or underrepresentation of terms. In contrast to "Cellular Component," "Biological Process," and "Molecular Function" could be affected severely by the differences in lifestyle. However, correlation coefficients for the same analysis of "Biological Process" and "Molecular Function" on level 2 indicate no difference while on level 3 there is a Pearson coefficient of 0.989 for "Biological Process" and 0.970 for "Molecular Function," indicating slight different abundance and a potential shift in function. Based on the distribution of GO terms and their proportions we designed comparative graphs depicting the comparison of single terms with respect to their percentages. At the compelling level 3 of "Cellular Component" we find slight differences for example in the highest abundant term "cell part" (33.5 and 29.6%) (**Figures 2C, D**). While Ccly has a higher portion of "membrane bounded organelle" (19.8%) compared to Pber (17.3%), "non-membrane-bounded organelle" associated sequences represent a larger part of the Pber dataset (12.1%) than it does in Ccly (10.5%).

To gain deeper insights into the species-specificity of terms we compared the number of independent occurrences of terms in the categories "Biological Process" and "Molecular Function" on level 3. Both categories contained terms exclusively associated with Msex but as we want to focus on the crustaceans we will not discuss them further.

Figure 3A depicts Venn diagrams representing the counts of level 3 "Biological Process" terms which are either common to both or unique to one hermit crab species. The vast majority (88%) of terms is common to both datasets. Eight terms are limited to Ccly. While only few ESTs are assigned to most of them, the term "response to other organism" represents 0.3% of the total level 3 terms. Pber showed no unique terms on level 3 "Biological Process." Additionally we categorized all terms observed in at least one species, relative to the proportion of the term in direct comparison between species (**Figure 3B**). The majority of terms differ only slightly in representation, exhibiting less than 0.5% difference. Two terms differ by 0.5–1% and 3 are in the category above 1%. A majority of terms (77.5%) is common to both datasets regarding third level terms in the category "Molecular Function" (**Figure 3C**). Pber shows 5 unique terms all of which have a low number of assigned ESTs. In Ccly 11 terms were unique, each with only very few ESTs assigned. The category of differences between 0.5 and 1% (**Figure 3D**) contained nine terms while three terms exhibited a difference bigger than 1%.

BIOLOGICAL PROCESS

As already described in the chemosensory tissues of insects (Legeai et al., 2011), terms assigned to basic cell functions like cellular and metabolic processes had the highest representation in all three species. We also identified well-represented terms typically expected in sensory tissue like "response to stimulus" and "signaling" as initially described for *M. sexta* (Grosse-Wilde et al., 2011) also in the hermit crab datasets. The distribution of terms

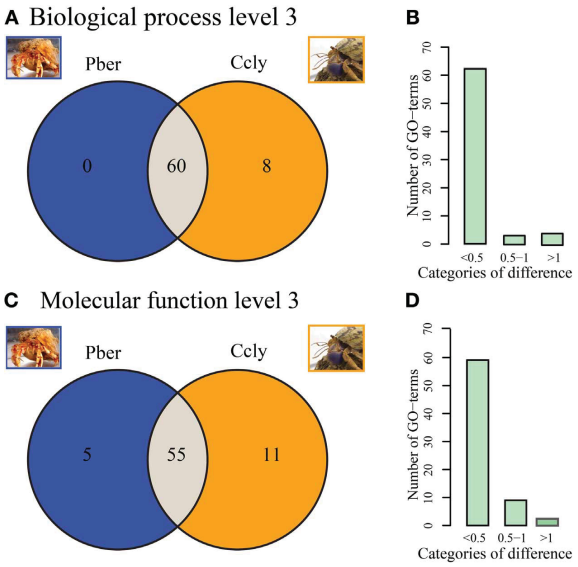


FIGURE 3 | (A,B) Biological Process. **(A)** Distribution of assigned terms in the category of "Biological Process" between Pber and Ccly, 60 terms shared between Pber and Ccly, 8 terms unique to Ccly. **(B)** Distribution of term ratios compared between in the categories "<0.5% difference," "0.5–1% difference," and ">1% difference." **(C,D)** Molecular function **(C)** Distribution of terms between Pber and Ccly, 55 terms shared, 5 terms unique to Pber, 11 terms unique to Ccly. **(D)** Distribution of term ratios in the categories "<0.5% difference," "0.5–1% difference," and ">1% difference." Percentage difference has been calculated as described in Bioinformatics.

in most of the categories showed visible differences, so we selected the most obvious ones with a difference above 0.5% between the hermit crabs to go into further detail (**Figure 4**). At level 2 the highest difference in abundance was observed concerning the term "cellular process" with 23% Pber and 19% Ccly (**Figure 4A**). We found 18 of its child terms on level 3, where only "cellular component biogenesis" and "cellular metabolic process" differed according to the selected range of more than 0.5%. On level 2 the 2nd highest differentially expressed term was "metabolic process" with 19.2% in Pber and 17.3% in Ccly. Besides the term "cellular metabolic process" being child term of "cellular process" as well, of its 6 level 3 child terms matched by our data 2 were also differentially represented in our dataset, "macromolecule metabolic process" and "primary metabolic process" (see table in **Figure 4B**).

MOLECULAR FUNCTION

The highest representation in all three species was connected to functions in the enzymatic activity and transport terms "binding" and "catalytic activity" confirming the observations in chemosensory tissues of other species (Legeai et al., 2011; Jacquin-Joly et al., 2012). Again we selected the most obvious terms differing between species (**Figure 5**). At the 2nd level the most pronounced differences occurred in the categories "structural molecule activity" (8.2% Pber, 4.4% Ccly) and "binding" with 40.6 and 43.7%, respectively (**Figure 5A**). This was also reflected by the higher levels, where one child term of "structural molecule activity" and 4 child terms of "binding" display differing ratio between the species. The highest was observed in the term "protein binding" with 16.3% Pber and 21.3% Ccly (**Figure 5B**). Looking at further annotation levels only a few terms show an obviously higher EST count in Ccly compared to Pber, most differences are small and finally sum up to the observed value. Regarding level 4 child terms of "protein binding" having higher proportions in Ccly compared to Pber include "cation binding," "hydrolase activity, acting on acid anhydrites," "nucleotide binding" and terms connected to transferase activity.

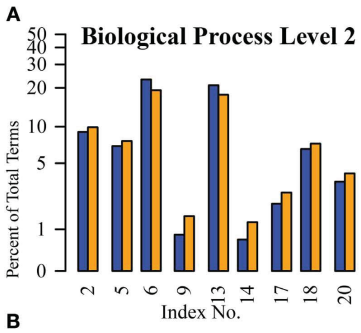
SELECTED GENE FAMILIES OF SENSING AND DEFENSE

Beyond the abstract level of GO terms we focused on olfaction, the main sensory task of the antennules. One advantage of using the GO annotation in terms of comparison and transferability is the abstract level of its vocabulary. However, for a more conclusive analysis we scrutinized annotation of single ESTs, manually revising annotations where necessary. This involved homology search on nucleic acid and translated RNA level, including contigs that were not annotated by BLAST2GO.

Chemosensation

First we looked at the gene families of ORs, GRs, OBPs, CSPs, and SNMPs, all of which are known to be main components of chemosensory detection in insects. While we were able to mostly replicate and expand on the previously published results for *M. sexta* using our new assembly, we were not able to identify any members of these families in the crustaceans. CSPs are binding proteins implicated in chemosensing that in contrast to OBPs are present not only in insects but also in genomic data of other arthropods (Vieira and Rozas, 2011; Kulmuni and Havukainen, 2013). However, the antennal transcriptomes of both hermit crabs did not contain any contigs potentially encoding CSPs.

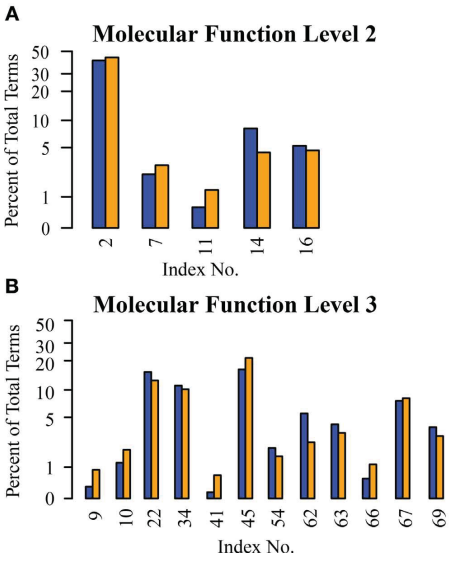
Second we looked at IRs, ancient receptors present throughout the protostomian lineage (Croset et al., 2010). The *C. clypeatus* transcriptome yielded 20 IR candidates identified by homology comparison, including the conserved IR25a and IR93a. In *P. bernhardus* 18 IR candidates are present, also including homologs of IR25a and IR93a. Both IR25a and IR93a of *C. clypeatus* were extended by RACE PCR although not to full length. Due to RNA material limitations this was not possible for *P. bernhardus* candidates. For higher reliability only putative IR coding contigs spanning at least two of the three transmembrane domains (Benton et al., 2009) without in frame stop codons or ambiguities were selected for the calculation of a multiple sequence alignment and the dendrogram shown in **Figure 6** (*C. clypeatus*: 7 candidates, *P. bernhardus*: 15 candidates). We decided to name them CclyIRX and PberIRX with X being ascending arabic numerals for discrimination where no clear homolog IR was available. Most of the hermit crab IR candidates group together with the non-coreceptor IRs from *P. argus* among the divergent IRs. PberIR9 seems to be close to the conserved antennal IR40a, a receptor with so far unknown ligand spectrum. The coreceptor IR8a was absent in both species.



Index	GO.ID	Term	Pber	Ccly
2	GO:0065007	biological regulation	9.08	9.89
5	GO:0071840	cellular component organiz. or biogenesis	6.97	7.67
6	GO:0009987	cellular process	23.12	19.2
9	GO:0040007	growth	0.83	1.49
13	GO:0008152	metabolic process	20.9	17.7
14	GO:0051704	multi-organism process	0.69	1.25
17	GO:0000003	reproduction	2.06	2.69
18	GO:0050896	response to stimulus	6.59	7.3
20	GO:0023052	signaling	3.42	4.06

Index	GO.ID	Term	Pber	Ccly
16	GO:0044085	cellular component biogenesis	1.82	2.32
24	GO:0044237	cellular metabolic process	10.31	8.73
38	GO:0043170	macromolecule metabolic process	7.71	6.46
50	GO:0044238	primary metabolic process	9.62	8.03
64	GO:0044281	small molecule metabolic process	3.11	3.91

FIGURE 4 | Selected terms of the category "Biological Process" with ratio differing by at least 0.5% between the species. (A) Level 2, **(B)** level 3; Blue bars: Pber, orange bars: Ccly. Numbering of bars indexed in the chart, proportions in the species given in per cent.



Index	GO.ID	Term	Pber	Ccly
2	GO:0005488	binding	40.6	43.7
7	GO:0030234	enzyme regulator activity	2.3	3.04
11	GO:0000988	protein binding transcription factor act.	0.59	1.33
14	GO:0005198	structural molecule activity	8.17	4.37
16	GO:0005215	transporter activity	5.24	4.62

Index	GO.ID	Term	Pber	Ccly
9	GO:0003682	chromatin binding	0.3	0.88
10	GO:0048037	cofactor binding	1.2	1.93
22	GO:0016787	hydrolase activity	15.33	12.6
34	GO:0003676	nucleic acid binding	11.13	10.2
41	GO:0042277	peptide binding	0.15	0.67
45	GO:0005515	protein binding	16.33	21.3
54	GO:0004871	signal transducer activity	2.05	1.54
62	GO:0003735	structural constituent of ribosome	5.54	2.46
63	GO:0022892	substrate-specific transporter activity	4.14	3.26
66	GO:0000989	transcr. factor binding transcr. factor act.	0.55	1.12
67	GO:0016740	transferase activity	7.64	8.18
69	GO:0022857	transmembrane transporter activity	3.84	2.98

FIGURE 5 | Selected terms of the category "Molecular function" with ratio differing by at least 0.5% between the species. (A) Level 2, **(B)** level 3; Blue bars: Pber, orange bars: Ccly. Numbering of bars indexed in the chart, proportions in the species given in per cent.

General GPCRs

We searched for G-Protein coupled receptors (GPCR) as this ancient and highly conserved protein family is involved in a broad variety of signal transduction processes. Any influence in genetic makeup between the Pber and Cclyp should be reflected here. In the *C. clypeatus* dataset we could identify a tyramine/octopamine receptor and several contigs assigned to GPCRs, including a rhodopsine like GPCR, relaxin receptors, one neuropeptide FF receptor2 and a homolog of methuselah. The *P. bernharus* dataset yielded a relaxin receptor and a tyramine-octopamine receptor.

Neuronal modulation and signaling

While our methods do not allow analysis of short neuropeptides, we could scrutinize receptors and enzymes involved in neuronal processes. The *Pagurus* dataset contained ESTs

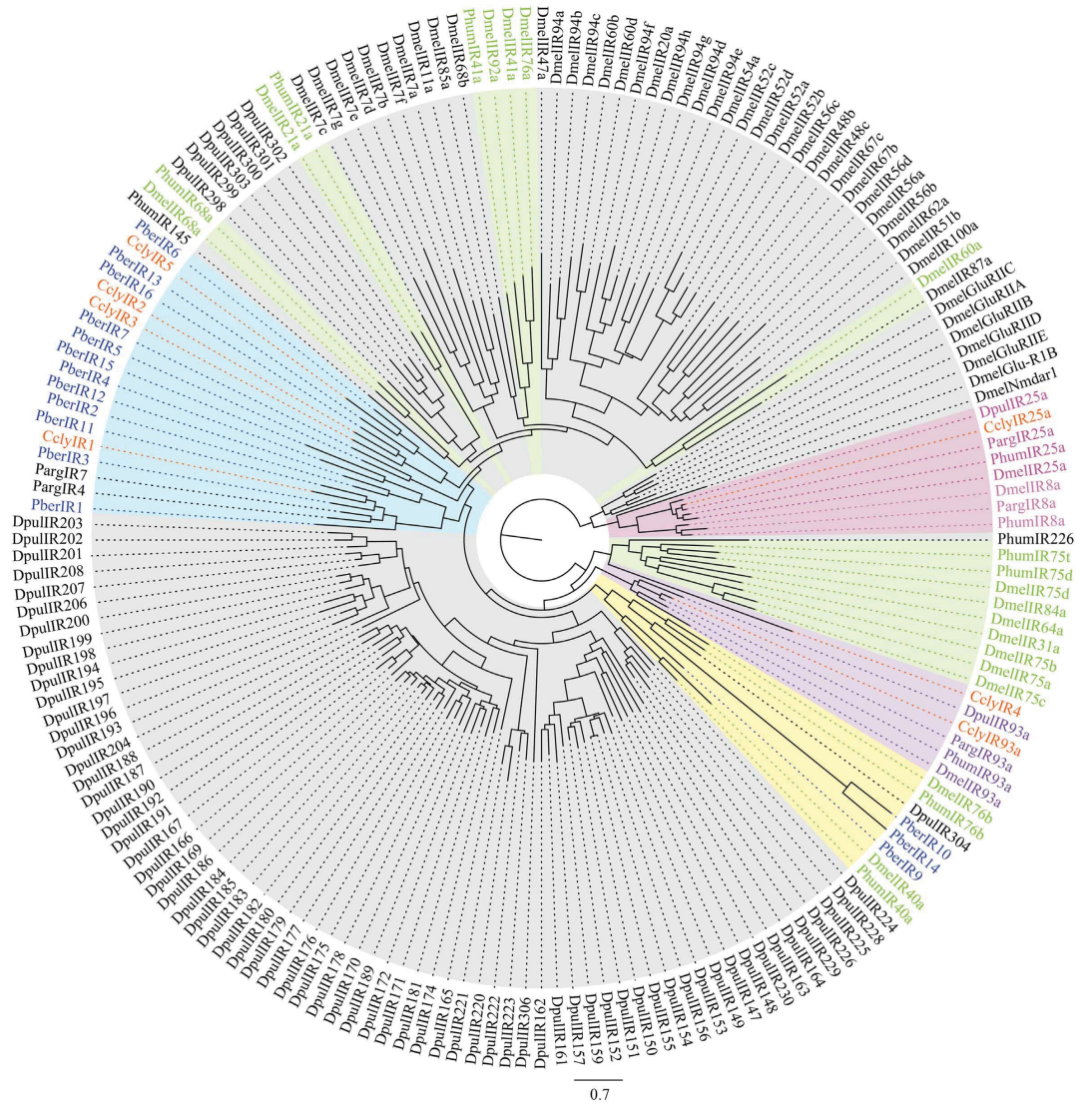


FIGURE 6 | Dendrogram of candidate IRs in *C. clypeatus* (orange) and *P. bernhardus* (blue) including reference sequences from *D. pulex* and *D. melanogaster*; highlighted in pink: ancient IRs and coreceptors, green and highlighted in green: antennal (olfactory) IRs, highlighted in blue: clusters of hermit crab candidate IRs, highlighted in yellow: cluster of both, antennal and divergent IRs containing hermit crab candidate IRs; modified after Croset et al. (2010).

annotated as dopamine beta hydroxylases, the already mentioned tyramine/octopamine receptor and a histamine gated chloride channel. In the *Coenobita* dataset several contigs were annotated as dopamine beta hydroxylases or tyrosine/dopamine beta hydroxylases; one dopamine transporter and a dopamine receptor interacting protein belonging to the heat shock protein (hsp) family could be assigned. Furthermore, a guanylyl cyclase was identified. While neither GABAa nor GABAb receptors could be found in the transcriptomes, several GABAa associated proteins and modulators were present.

Mechanosensation, Thermosensation

Bi- or even multimodal sensilla are widely distributed among arthropods and can be found on almost all body surfaces. Among decapods, bimodal chemo-mechanosensilla are common and morphologically well-described, playing a role in close-range sensing (Cate and Derby, 2001). While the numerous cilia sensitive to mechanical stimuli on the crustacean antennules have been studied (Atema, 1977), molecular information is lacking. A recent study in *Drosophila* larvae screened for ESTs which are necessary to maintain the ability to sense touch events (Tsubouchi et al., 2012). According to BLAST searches based on these

sequences the *Coenobita* dataset contained one EST with similarity to nompC, one chloride channel b and iGluRs of the NMDA2 type. In *P. bernhardus* only homologs to NMDA2 receptors could be found, confirming the presence of receptor transcripts with similar features. Thermosensing in *D. melanogaster* is mainly based on the TRP channels painless and TRPA1 (Tracey et al., 2003; Rosenzweig et al., 2005). Both are expressed in the arista, making it likely to expect similar receptors also in the crustacean antennules (Montell, 2009). However, neither homologs to painless or TRPA1 nor other TRP channels were present in our datasets.

Immune response and antimicrobial defense

An organ or a tissue exposed to the outside likely needs protection against potential pathogens. As the antennae have not been characterized as central production organs of immune-relevant factors we were looking broadly for expression of genes that have been considered to contribute to immune responses in crustaceans [for review see Cerenius et al. (2010)]. We were able to identify a variety of AMPs, a potent group of immune effectors [reviewed in Rosa and Barracco (2010)] and several ESTs annotated to genes which are of putative relevance in the crustacean immune system. A summary list is given in **Table 2**.

Table 2 | Potentially immune relevant ESTs, categories selected after Robalino et al. (2007).

	P. bernhardus	*C. clypeatus*
ANTIMICROBIAL		
C-type Lectin	10	7
Crustin	12	4
Carcinin	0	1
Anti-LPS factor	5	3
CELL ADHESION		
Cadherin 23	0	1
Integrin alpha	0	2
Integrin beta	1	0
Integrin beta binding protein	0	1
Peroxinectin	4	2
Tetraspanin	3	4
CELL DEATH		
Autophagy protein 9	0	3
OXIDATIVE STRESS		
Glutathione S-transferase	10	12
Peroxiredoxin	7	4
Thioredoxin	2	2
PROTEASES		
Aminopeptidase	1	8
Cathepsin A	0	2
Cathepsin B	2	2
Cathepsin L	1	1
Cubilin protease	1	2
ProPO CASCADE		
Prophenoloxidase	3	0
PO activating factor	7	6
PROTEASE INHIBITORS		
Serine protease inhibitor	10	17
RNAi		
Armitage	0	1
Tudor staphylococcal nuclease	1	3

DISCUSSION

Here we present the first transcriptome analysis of crustacean olfactory tissues, providing a comparison between an aquatic and a fully terrestrial species; *Pagurus bernhardus* and *Coenobita clypeatus*.

Summary assembly statistics of the two datasets confirmed a quality similar to previously published datasets on insect olfactory tissues and our reference transcriptome of *M. sexta*, allowing comparison. In agreement with previous non-olfactory transcriptome studies on crustacea, only a small number of ESTs (Ccly 25%, Pber 33%) was significantly similar to known genes in general homology searches, while InterproScan analysis assigned functional domains to ca. 60% of the ESTs in both datasets. This discrepancy most likely originates from a lack of information on crustacean molecular biology, amplified further by a lack of analysis of antennally expressed genes in general. Comparability of the crustacean datasets is further confirmed by the equality of term distribution in the GO term category of "Cellular Component." Differences observed in the category "Biological Process" were overall small and mainly caused by a differing number of subtypes or variants of factors and regulators. However, meaningful differences were observed in the category of "Molecular Function." For example, the 3rd level term "protein binding" is associated to nearly twice as many ESTs in *C. clypeatus* than in *P. bernhardus*. Nevertheless, the term is still very abstract, and the diversity among the associated terms is high. EST count ratios on lower annotation levels shift toward *C. clypeatus* mostly in terms of enzymatic activity on proteins and transcription machinery. The likely reason is the genetic heterogeneity of the *C. clypeatus* RNA sample. While all the *P. bernhardus* individuals originated from a comparatively small area around the island of Helgoland (Germany), the supplier for *C. clypeatus* could not provide us with more precise information about the region the animals were caught, but confirmed the possibility of different origins for the animals.

Beyond the abstract level of GO terms we focused on olfaction, the main sensory task of the antennules. As all previously described chemosensory genes (Grosse-Wilde et al., 2011) were present in the new assembly of our reference *M. sexta* dataset we assume the parameters set for assembly are reliable. Furthermore, it was possible to identify even lowly expressed genes such as IRs from the *M. sexta* data, which is based on substantially lower number of reads and total data. This indicates that the probability of missing entire gene families in the hermit crab datasets is negligible. The absence of receptors belonging to the OR family in the hermit crabs is consistent with findings in *Daphnia pulex* and the spiny lobster (Penalva-Arana et al., 2009; Corey et al., 2013) and supports the assumption that the OR superfamily developed only in the insect lineage of arthropods (Sanchez-Gracia et al., 2009). GRs, however, are present in the *Daphnia* genome but were not found to be antennaly expressed in the spiny lobster (Penalva-Arana et al., 2009; Corey et al., 2013). If the crustacean sense of

taste is exclusively located on the walking legs as indicated in lobsters (Johnson et al., 1988), this could explain the absence of GRs from the antennal transcriptomes. The identification of the putative coreceptors IR25a and IR93a confirms the presence of the ancient IR nose in hermit crabs. Both IRs are conserved within the tetraconata, while the function and ligands of IR93a are unknown (Silbering et al., 2011). Nevertheless, its conservation indicates an importance across lineages, which seems to be independent from terrestrial or aquatic lifestyle. Most other candidate IRs cluster together with IRs from *P. argus* as decapod IR subgroups with no clear insect or *Daphnia* homolog. As little is known on the function of divergent IRs, it needs to be noted that their presence in hermit crab antennal transcriptomes confirms studies of the lobster antennae (Corey et al., 2013) indicating the existence of a new IR subgroup, as these IRs do not belong to the insect antennal IRs while divergent IRs are not expressed in the antennae of any insect studied so far. As clusters contain IR candidates from both hermit crabs and the aquatic *P. argus* it is rather likely that they have a function that is not depending on lifestyle. Contradictory to the expectation of specific olfactory-relevant receptors emerging after the transition from water to land, both hermit crabs seem to have a similar set of IRs expressed in the first antenna pair. Although no major screening of olfactory sensitivity including a large set of chemically different odors has been performed in aquatic hermit crabs, a recent electroantennographic study by Krång et al. showed that *C. clypeatus* detects almost exclusively water soluble odorants (Krång et al., 2012). These chemicals are putative cues available in the environment of both species. While genomic evidence of OBPs is present for most insect genera from the pea aphid (Zhou et al., 2010) to the silkmoth (Gong et al., 2009) but restricted to hexapods, CSPs were found in the *Daphnia* genome and could thus be expected in the hermit crabs as well (Vieira and Rozas, 2011). However, no CSPs are expressed in the antennal transcriptomes of either hermit crab. Binding proteins have been shown to be not restricted to olfactory tissues in insects and are likely involved in other functions as well (Pelosi et al., 2006). The absence of known families of binding proteins from the antennal transcriptomes indicates an independence of chemosensory detection from carrier proteins. This also fits the assumption that binding proteins are needed to transfer lipophilic odorants through the sensillum lymph, a function that might be dispensable for water soluble odorants (Kaissling, 2001; Leal et al., 2005). The EAG responses of *C. clypeatus* indicate a partial overlap with the odor spectrum evoking responses in IR expressing OSNs of coeloconic sensilla of *Drosophila* (Silbering et al., 2011; Krång et al., 2012). This raises the question if hermit crabs use the IRs related to the divergent IRs in insects, or altogether different receptors for the detection of the same chemicals. Since the divergent IRs are prime candidates, studies on expression and functionality of these receptors are needed to clarify whether or not they are involved in the detection of specific odorants. Following the described olfactory pathway in lobster we confirm details of the olfactory transduction pathway, indicating histamine-dependent regulation and a putative involvement of G-protein mediated signaling. Consequently, we propose a hermit crab olfactory transduction mechanism comparable to the one found in lobster (Hatt and Ache, 1994).

While the numerous cilia sensitive to mechanical stimuli on the crustacean antennules have been studied (Keil, 2012), molecular information is lacking. A recent study in *Drosophila* larvae screened for ESTs which are necessary to maintain the ability to sense touch events (Tsubouchi et al., 2012). Transcripts of receptors with similar features are also present in the hermit crab transcriptomes, including nompC and a ClCb channel in *C. clypeatus* and NMDA receptors in both datasets. The sense of touch might thus be mediated by related receptors in hermit crabs as in *Drosophila*.

The immune system of crustaceans is a subject of intense research, while the role of pathogen defense in the antennules has not been studied. We identified ESTs representing both more unspecific as well as pathogen-specific innate immune mechanisms. The presence of crustins, carcinins, C-type lectins, and anti-lipopolysaccharide factors (ALFs) in our datasets confirm the presence of pattern recognition mechanisms as well as production of AMPs in hermit crabs. Furthermore, *P. bernhardus* exhibits a higher diversity of AMPs than *C. clypeatus*, presumably due to either a larger number of potential pathogens, or a higher rate of exposure to small organisms recognized as non-self (i.e., algae) in the aquatic environment. The different types of cathepsins present might play a role in immune response, as cathepsin expression is upregulated upon viral challenge in shrimp (Robalino et al., 2007). The RNAi pathway, another defense mechanism against viral pathogens, is represented by homologs of the *Drosophila* RNA helicase Armitage and the conserved Tudor staphylococcal nuclease. Our findings thus demonstrate expression of various immune relevant genes in the antennae, with most of the known crustacean defense mechanisms present. This indicates the antennulae as a site of production and not only action of pathogen defense. A potential reason for this is the structure of the antennular cuticle. To allow the passage of odor molecules, the aesthetasc cuticle is comparatively thin at the exposed side (Stensmyr et al., 2005) and can thus easily be penetrated by invading microorganisms. Additionally the aesthetascs are immersed in a moist layer of a mucus-like substance (unpublished data). This could putatively promote the growth of microbes and therefore affect olfaction by production of odorants or clogging the entry sites of odorants with microorganisms. The importance of antimicrobial defense is also indicated by the raw sequencing data containing a high number of reads originating from bacteria, fungi, and algae.

Both hermit crab species are well-adapted to their habitats. The substantial changes in lifestyle by becoming terrestrial led to various morphological changes in the periphery of antennules (Ghiradella et al., 1968b) and an enlargement and reorganization of olfactory brain centers in *C. clypeatus* (Harzsch and Hansson, 2008). The changes regarding the OL of terrestrial hermit crabs are considered to be an adaptation to the aerial sense of smell and as an indication of fundamental changes in olfaction in general (Rittschof and Sutherland, 1986; Stensmyr et al., 2005). However, the molecular differences between the antennal transcriptomes of the marine *Pagurus bernhardus* and the terrestrial *Coenobita clypeatus* are overall small, indicating that the exhibited changes in function and morphology are mainly founded on changes in

small numbers of genes. With respect to the main antennular function of chemosensing they display a similar set of IR candidates in count and sequence similarity. The number and full length of the receptor candidates, their ligand specificity, their expression pattern in the OSN populations and possible combinatorial effects need to be investigated further. No other chemosensory receptor candidates were identified in either of the species.

AUTHORS CONTRIBUTIONS

Katrin C. Groh and Ewald Grosse-Wilde carried out the bioinformatic experiments, data analysis and drafted the manuscript. Heiko Vogel participated in evaluation of data analysis. Heiko Vogel, Ewald Grosse-Wilde, Marcus C. Stensmyr, and Bill S. Hansson participated in study design, coordination and drafting of the manuscript. Katrin C. Groh wrote the paper. Reagents and analytic tools were provided by Bill S. Hansson. All authors read and approve the final manuscript.

ACKNOWLEDGMENTS

We want to acknowledge Regina Stieber for RNA sample preparation, Steffi Gebauer-Jung for invaluable help with sequence cleanup, the Max Planck Institute for Genetics in Berlin for sequencing, Steffen Harzsch, Carsten Mueller, and Linda Kuebler for kindly providing photos, Michael Thoma and Sarah Koch for help with statistical analysis using R and the Max Planck Society for funding.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/journal/10.3389/fnins.2013.00266/abstract

(a) Sequencing data (EBI study acc. nr. ERP002374, http://www.ebi.ac.uk/ena/data/view/ERP002374).

(b) ESTs have been deposited on ENA, accession numbers *P. bernhardus*: HABX01000001 to HABX01000085, *C. clypeatus*: HABY01000001 to HABY01000089.

REFERENCES

Ai, M., Min, S., Grosjean, Y., Leblanc, C., Bell, R., Benton, R., et al. (2010). Acid sensing by the Drosophila olfactory system. *Nature (London)* 468, 691. doi: 10.1038/nature09537

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Atema, J. (1977). "Functional separation of smell and taste in fish and crustaceans," in *Olfaction and Taste*, Vol. VI, eds P. Le Magnen and J. MacLeod (London: Information Retrieval Ltd.), 165–174.

Benton, R., Vannice, K. S., Gomez-Diaz, C., and Vosshall, L. B. (2009). Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. *Cell* 136, 149–162. doi: 10.1016/j.cell.2008.12.001

Bliss, D. E., and Mantel, L. H. (1968). Adaptations of crustaceans to land - a summary and analysis of new findings. *Am. Zool.* 8, 673–685. doi: 10.1093/icb/8.3.673

Bracken-Grissom, H., Cannon M. E., Cabezas, P., Feldmann, R. M., Schweitzer, C. E., Ahyong, S. T., et al. (2013). A comprehensive and integrative reconstruction of evolutionary history for Anomura (Crustacea: Decapoda). *BMC Evol. Biol.* 13:128. doi: 10.1186/1471-2148-13-128

Bush, B. M. H., and Laverack, M. S. (1982). "Mechanoreceptors," in *Biology of Crustacea*, Vol. 3, eds D. H. Bliss, H. L. Atwood, and D. C. Sandeman (New York, NY: Academic Press), 369–468.

Cate, H. S., and Derby, C. D. (2001). Morphology and distribution of setae on the antennules of the caribbean spiny lobster *Panulirus argus* reveal new types of bimodal chemo-mechanosensilla. *Cell Tissue Res.* 304, 439–454. doi: 10.1007/s004410100377

Cate, H. S., and Derby, C. D. (2002). Hooded Sensilla homologues: structural variations of a widely distributed bimodal chemomechanosensillum. *J. Comp. Neurol.* 444, 345–357. doi: 10.1002/cne.10153

Cerenius, L., Jiravanichpaisal, P., Liu, H. P., and Soederhall, I. (2010). "Crustacean immunity," in *Invertebrate Immunity*, Vol. 708, ed K. Soderhall (New York, NY: Springer Science + Business Media), 239–259.

Conesa, A., Goetz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., and Robles, M. (2005). BLAST2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610

Corey, E., Bobkov, Y., Ukhanov, K., and Ache, B. W. (2013). Ionotropic crustacean olfactory receptors. *PLoS ONE* 8:e60551. doi: 10.1371/journal.pone.0060551

Croset, V., Rytz, R., Cummins, S. F., Budd, A., Brawand, D., Kaessmann, H., et al. (2010). Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste. *PLoS Genet.* 6:e1001064. doi: 10.1371/journal.pgen.1001064

Derby, C. D., and Atema, J. (1982). Chemosensitivity of walking legs of the lobster *Homarus americanus* - neurophysiological response spectrum and thresholds. *J. Exp. Biol.* 98, 303–315.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Ghiradella, H., Case, J., and Cronshaw, J. (1968a). Structure of aesthetascs in selected marine and terrestrial decapods - chemoreceptor morphology and environment. *Am. Zool.* 8, 603. doi: 10.1093/icb/8.3.603

Ghiradella, H., Case, J., and Cronshaw, J. (1968b). Fine structure of aesthetasc hairs of *Coenobita compressus* Edwards. *J. Morphol.* 124, 361–385. doi: 10.1002/jmor.1051240309

Gleeson, R. A. (1982). Morphological and behavioral identification of the sensory structures mediating pheromone reception in the blue crab, *Callinectes sapidus*. *Biol. Bull.* 163, 162–171. doi: 10.2307/1541506

Gong, D. P., Zhang, H. J., Zhao, P., Xia, Q. Y., and Xiang, Z. H. (2009). The odorant binding protein gene family from the genome of silkworm, *Bombyx mori*. *BMC Genomics* 10:332. doi: 10.1186/1471-2164-10-332

Greenaway, P. (2003). Terrestrial adaptations in the Anomura (Crustacea: Decapoda). *Mem. Mus. Vic.* 60, 13–26.

Grosse-Wilde, E., Kuebler, L. S., Bucks, S., Vogel, H., Wicher, D., and Hansson, B. S. (2011). Antennal transcriptome of *Manduca sexta*. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7449–7454. doi: 10.1073/pnas.1017963108

Han-Ching Wang, K., Tseng, C., Lin, H., Chen, I. T., Chen, Y. H., Chen, Y. I., et al. (2010). RNAi Knock-down of the *Litopenaeus vannamei* Toll Gene (LvToll) significantly increases mortality and reduces bacterial clearance after challenge with *Vibrio harveyi*. *Dev. Comp. Immunol.* 34, 49–58. doi: 10.1016/j.dci.2009.08.003

Hansson, B. S., Harzsch, S., Knaden, M., and Stensmyr, M. (2011). "The neural and behavioral basis of chemical communication in terrestrial crustaceans," in *Chemical Communication in Crustaceans*, eds T. Breithaupt and M. Thiel (New York, NY: Springer Science + Business Media), 149–173. doi: 10.1007/978-0-387-77101-4_8

Harzsch, S., and Hansson, B. S. (2008). Brain architecture in the terrestrial hermit crab *Coenobita clypeatus* (Anomura, Coenobitidae), a crustacean with a good aerial sense of smell. *BMC Neurosci.* 9:58. doi: 10.1186/1471-2202-9-58

Hatt, H., and Ache, B. W. (1994). Cyclic nucleotide and inositol phosphate gated ion channels in lobster olfactory receptor neurons. *Proc. Natl. Acad. Sci. U.S.A.* 91, 6264–6268. doi: 10.1073/pnas.91.14.6264

Hollins, B., Hardin, D., Gimelbrant, A. A., and McClintock, T. S. (2003). Olfactory-enriched transcripts are cell-specific markers in the lobster olfactory organ. *J. Comp. Neurol.* 455, 125–138. doi: 10.1002/cne.10489

Jacquin-Joly, E., Legeai, F., Montagne, N., Monsempes, C., Francois, M. C., Poulain, J., et al. (2012). Candidate chemosensory genes in female antennae of the noctuid moth *Spodoptera littoralis*. *Int. J. Biol. Sci.* 8, 1036–1050. doi: 10.7150/ijbs.4469

Johnson, B. R., Merrill, C. L., Ogle, R. C., and Atema, J. (1988). Response properties of lobster chemoreceptors - tuning of olfactory neurons sensitive to hydroxyproline. *J. Comp. Physiol. A* 162, 201–211. doi: 10.1007/bf00606085

Kaissling, K. E. (2001). Olfactory perireceptor and receptor events in moths: a kinetic model. *Chem. Senses* 26, 125–150. doi: 10.1093/chemse/26.2.125

Keil, T. (2012). Sensory cilia in arthropods. *Arthropod Struct. Dev*. 41, 515–534. doi: 10.1016/j.asd.2012.07.001

Koczan, S. (2012). *3D Reconstruction of the Olfactory Pathway from the Aesthetasc to the Olfactory Lobe in the Hermit Crab, Coenobita clypeatus*. Diploma Thesis, Friedrich-Schiller University—Faculty of Biology and Pharmacy, Jena.

Krång, A. S., Knaden, M., Steck, K., and Hansson, B. S. (2012). Transition from sea to land: olfactory function and constraints in the terrestrial hermit crab *Coenobita clypeatus*. *Proc. Biol. Sci*. 279, 3510–3519. doi: 10.1098/rspb.2012.0596

Kulmuni, J., and Havukainen, H. (2013). Insights into the evolution of the CSP gene family through the integration of evolutionary analysis and comparative protein modeling. *PLoS ONE* 8:e63688. doi: 10.1371/journal.pone.0063688

Kwon, H. W., Lu, T., Rutzler, M., and Zwiebel, L. J. (2006). Olfactory responses in a gustatory organ of the malaria vector mosquito *Anopheles gambiae*. *Proc. Natl. Acad. Sci. U.S.A*. 103, 13526–13531. doi: 10.1073/pnas.0601107103

Leal, W. S., Chen, A. M., Ishida, Y., Chiang, V. P., Erickson, M. L., Morgan, T. I., et al. (2005). Kinetics and molecular properties of pheromone binding and release. *Proc. Natl. Acad. Sci. U.S.A*. 102, 5386–5391. doi: 10.1073/pnas.0501447102

Legeai, F., Malpel, S., Montagne, N., Monsempes, C., Cousserans, F., Merlin, C., et al. (2011). An expressed sequence tag collection from the male antennae of the noctuid moth *Spodoptera littoralis*: a resource for olfactory and pheromone detection research. *BMC Genomics* 12:86. doi: 10.1186/1471-2164-12-86

Montell, C. (2009). A taste of the Drosophila gustatory receptors. *Curr. Opin. Neurobiol*. 19, 345–353. doi: 10.1016/j.conb.2009.07.001

Pelosi, P., Zhou, J.-J., Ban, L. P., and Calvello, M. (2006). Soluble proteins in insect chemical communication. *Cell. Mol. Life Sci*. 63, 1658–1676. doi: 10.1007/s00018-005-5607-0

Penalva-Arana, D. C., Lynch, M., and Robertson, H. M. (2009). The chemoreceptor genes of the waterflea *Daphnia pulex*: many Grs but no Ors. *BMC Evol. Biol*. 9:79. doi: 10.1186/1471-2148-9-79

Powers, L. W., and Bliss, D. E. (1983). "Terrestrial adaptations," in *The Biology of Crustacea*. Vol. 8. *Environmental Adaptations,* eds F. J. Vernberg and W. B. Vernberg (New York, London, etc.: Academic Press), 271–333.

Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5:e9490. doi: 10.1371/journal.pone.0009490

Rittschof, D., and Sutherland, J. P. (1986). Field studies on chemically mediated behavior in land hermit crabs *Coenobita-rugosus* volatile and nonvolatile odors. *J. Chem. Ecol*. 12, 1273–1284. doi: 10.1007/BF01012347

Robalino, J., Almeida, J. S., McKillen, D., Colglazier, J., Trent, H. F., Chen, Y. A., et al. (2007). Insights into the immune transcriptome of the shrimp *Litopenaeus vannamei*: tissue-specific expression profiles and transcriptomic responses to immune challenge. *Physiol. Genomics* 29, 44–56. doi: 10.1152/physiolgenomics.00165.2006

Robertson, H. M., and Kent, L. B. (2009). Evolution of the gene lineage encoding the carbon dioxide receptor in insects. *J. Insect Sci*. 9, 19. doi: 10.1673/031.009.1901

Robertson, H. M., Warr, C. G., and Carlson, J. R. (2003). Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A*. 100, 14537–14542. doi: 10.1073/pnas.2335847100

Rosa, R. D., and Barracco, M. A. (2010). Antimicrobial peptides in crustaceans. *Invertebr. Surv. J*. 7, 262–284.

Rosenzweig, M., Brennan, K. M., Tayler, T. D., Phelps, P. O., Patapoutian, A., and Garrity, P. A. (2005). The Drosophila ortholog of vertebrate TRPA1 regulates thermotaxis. *Genes Dev*. 19, 419–424. doi: 10.1101/gad.1278205.fied

Sanchez-Gracia, A., Vieira, F. G., and Rozas, J. (2009). Molecular evolution of the major chemosensory gene families in insects. *Heredity* 103, 208–216. doi: 10.1038/hdy.2009.55

Silbering, A. F., Rytz, R., Grosjean, Y., Abuin, L., Ramdya, P., Jefferis, G. S. X. E., et al. (2011). Complementary function and integrated wiring of the evolutionarily distinct Drosophila olfactory subsystems. *J. Neurosci*. 31, 13357–13375. doi: 10.1523/JNEUROSCI.2360-11.2011

Soederhall, K., and Cerenius, L. (1998). Role of the prophenoloxidase-activating system in invertebrate immunity. *Curr. Opin. Immunol*. 10, 23–28. doi: 10.1016/S0952-7915(98)80026-5

Stensmyr, M. C., Erland, S., Hallberg, E., Wallen, R., Greenaway, P., and Hansson, B. S. (2005). Insect-like olfactory adaptations in the terrestrial giant robber crab. *Curr. Biol*. 15, 116–121. doi: 10.1016/j.cub.2004.12.069

Tracey, W. D., Wilson, R. I., Laurent, G., and Benzer, S. (2003). Painless, a Drosophila gene essential for nociception. *Cell* 113, 261–273. doi: 10.1016/S0092-8674(03)00272-1

Tsubouchi, A., Caldwell, J. C., and Tracey, W. D. (2012). Dendritic filopodia, ripped pocket, NOMPC, and NMDARs contribute to the sense of touch in Drosophila larvae. *Curr. Biol*. 22, 2124–2134. doi: 10.1016/j.cub.2012.09.019

Vieira, F. G., and Rozas, J. (2011). Comparative genomics of the odorant-binding and chemosensory protein gene families across the arthropoda: origin and evolutionary history of the chemosensory system. *Genome Biol. Evol*. 3, 476–490. doi: 10.1093/gbe/evr033

Zhou, J. J., Vieira, F. G., He, X. L., Smadja, C., Liu, R., Rozas, J., et al. (2010). Genome annotation and comparative analyses of the odorant-binding proteins and chemosensory proteins in the pea aphid *Acyrthosiphon pisum*. *Insect Mol. Biol*. 19, 113–122. doi: 10.1111/j.1365-2583.2009.00919.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 July 2013; accepted: 18 December 2013; published online: 21 January 2014.

Citation: Groh KC, Vogel H, Stensmyr MC, Grosse-Wilde E and Hansson BS (2014) The hermit crab's nose—antennal transcriptomics. Front. Neurosci. 7:266. doi: 10.3389/fnins.2013.00266

This article was submitted to Evolutionary Psychology and Neuroscience, a section of the journal Frontiers in Neuroscience.

5. Manuscript 2

Expression of ionotropic receptors in terrestrial hermit crab's olfactory sensory neurons

Katrin C. Groh[1], Marcus C. Stensmyr[2], Ewald Grosse-Wilde[1*], Bill S. Hansson[1]

Affiliations:

1 Max Planck Institute for Chemical Ecology, Department of Evolutionary Neuroethology, Germany
2 Lund University, Department of Biology – Functional Zoology, Sweden

Correspondence:

Dr. Ewald Grosse-Wilde
Max Planck Institute for Chemical Ecology
Department of Evolutionary Neuroethology
Hans-Knoell-Str. 8
07745 Jena, Germany

grosse-wilde@ice.mpg.de

Word count: 3356

number of figures: 4

Abstract

Coenobitidae are one out of at least five crustacean lineages which independently succeeded in the transition from water to land. This change in lifestyle required various adaptations of their olfactory organs in order to sense chemical cues in the new conveying medium. The crustacean sense of smell is situated in their first pair of antennae, also called antennules, and more precisely in the olfactory aesthetascs that are arranged in a field on the distal segment of the antennular flagellum. Aesthetascs house approximately 300 dendrites each, with their neurons cell bodies arranged in spindle-like complexes of ca. 150 each. Aesthetascs of aquatic crustaceans are the place of odor uptake and IR co-receptors IR25a and IR93a have been shown to be expressed in olfactory sensory neurons (OSNs) of the antennules. Our goal was to reveal the expression of non-co-receptor IRs in the OSNs of *Coenobita clypeatus*, a terrestrial hermit crab, with RNA in *situ* hybridization and their distribution pattern in the antennules. We investigated an RNAseq dataset revealing 22 novel divergent IRs in the Coenobita antennules and designed RNA probes for three exemplary IRs displaying their expression in cells of the OSN cell body complexes.

Keywords: crustacea, antennules, olfaction, ionotropic receptors, in *situ* hybridization

1. Introduction

One of the biggest challenges in animal evolution was the successful transition from water to land. Among arthropods, this step was successfully and independently accomplished by several taxa, including at least five crustacean lineages (Bliss and Mantel 1968; Powers and Bliss 1983). One of these are the hermit crabs (Coenobitidae), who transitioned approximately 20 mya according to fossil records (Glaessner 1969). Modern species with a predominantly or exclusively terrestrial lifestyle belong to the monotypic genus *Birgus* and the genus *Coenobita*, which contains approximately 15 recent species. Substantial changes were required to adapt to the new environment in terms of water and ion balance, metabolism and modulation of sensory organs in order to receive stimuli from a changed conveying medium. The crustacean antennae are important sensory organs, with the second pair of antennae being primary mechanosensors and the first pair of antennae, also called antennules, primary olfactory organs (Koehl et al. 2001; Eder and Atema 1978). Hermit crabs are known to rely on their chemical sense in many ways, including predator avoidance, searching for food, fresh- and salt water and resources like empty snail shells used to protect their soft abdomen (Rosen, Schwarz and Palmer 2009; Thacker 1997; Small et al. 1994; Krång et al. 2012).

The crustacean sister phylum, the insects, succeeded in the transition from water to land approximately 600 mya (Richter, Moller and Wirkner 2009; Little 2009). Insects adapted to terrestrial olfaction by employing two different receptor classes: the Olfactory Receptors (ORs), a gene family most likely derived from the older Gustatory Receptors (GRs), and the antennal Ionotropic Receptors (IRs) (Robertson, Warr and Carlson 2003; Benton et al. 2009). GRs but no ORs are present in the genome of *Daphnia pulex*, the only complete crustacean genome available to date, indicating that ORs developed in the insect lineage with their oldest known representatives in the primary wingless insect *Thermobia domestica* (Penalva-Arana, Lynch and Robertson 2009; Missbach et al. 2014). IRs, however, are derived from the ancient family of ionotropic glutamate receptors, with the earliest occurrence of the antennal IR co-receptor, IR25a, in the early protostomian lineage (Croset et al. 2010). Insect IRs are subdivided into two groups based on their expression: the “antennal IRs”, expressed in sensilla of the insect antennae, and “divergent IRs” which are not antennaly expressed and have not been functionally characterized yet (Croset et al. 2010; Benton et al. 2009; Silbering et al. 2011). The *Daphnia* genome yielded a large set of GRs, two antennal IRs (the ancient co-receptor IR25a and IR93a) and a high number of divergent IRs (Croset et al. 2010). The

lobster *Panulirus argus* expresses the co-receptor IR25a, the antennal IR93a and two divergent IRs in antennules (Corey et al. 2013). In terrestrial hermit crabs, IR25a, IR93a and seven divergent IRs were identified from the antennal transcriptome of *C. clypeatus* (Groh et al. 2014). IRs remain the only putative chemoreceptors in the crustacean antennules known, as neither GRs nor any other chemosensory receptor is expressed in this tissue (Corey et al. 2013; Groh et al. 2014). However, it is yet unknown if the divergent IRs identified from transcriptomic data are expressed in olfactory sensory neurons. The crustacean olfactory sensilla, called aesthetascs, are arranged in rows on the last antennular segment (Ghiradella, Case and Cronshaw 1968). Each aesthetasc houses approximately 300 OSN dendrites which are responsive to odor contact (Gleeson 1982). While in aquatic decapods the aesthetascs are long and slender with OSN cell body clusters lying shallow underneath the cuticle, the aesthetascs of *C. clypeatus* are short and blunt, with OSN cell bodies arranged in spindle like complexes of approx. 150 cell bodies each, which are withdrawn deeper into the last antennular segment (Krång et al. 2012; Stensmyr et al. 2005; Hansson et al. 2011; Koczan 2012).

We expanded on our previous investigation of the antennal transcriptome of C. clypeatus, generating a dataset with more depth for a more complete assessment of the IR repertoire. Based on these data we cloned key receptors and performed probes for fluorescent RNA in situ hybridization to demonstrate expression of “divergent” IRs in OSN cell bodies. We demonstrate for the first time that this orphan receptor subgroup with likely non-olfactory function in insects is with high probability the basis of olfaction in terrestrial hermit crabs.

2. Material and Methods

2.1 Animals and sample preparation

Specimens of *Coenobita clypeatus* were ordered from Peter Hoch Import – Export Waldkirch (Germany). Cold anesthetization of specimen was followed by antennulae dissection by scissors. Dissected antennules were either pooled for RNA extraction or transferred into fixation buffer for in situ experiments. *C. clypeatus* is a species neither endangered nor protected. All experiments were carried out in accordance with the national ethical guidelines ("genehmigungsfreien Versuchsvorhabens nach § 8a Abs.1 und 2 des Tierschutzgesetzes Deutschland vom 18. Mai 2006 BGBl. I S. 1206"), including notification and consent of the responsible administrative authorities.

2.2 RNA preparation and RNAseq

Antennules of 10 specimens were homogenized in a tissue-lyzer (Invitrogen) in 1ml TRI reagent (Sigma) at 50 hz for 3 min. The following steps were carried out according to the manufacturer's instructions, replacing choloroform with 1-bromo-3-chloro-propane. Total RNA preparations were sent to the Max Planck-Genome-Centre Cologne for TruSeq RNA sequencing of 15,000,000 reads. Raw sequencng data is available for download (EBI/ENA Study acc.number ERP005273). Data was screened for contaminants and linker-/adapter sequences followed by de novo assembly of the sanitized reads in CLC genomics workbench V 6.

2.3 Annotation

A BLAST2GO database was created based on the contigs with homology searches after dynamic translation (BLASTX) against non-redundant databases (National Center for Biotechnology Information, NCBI) using the default cutoff parameters (1.0E-3) and InterProScan (Conesa et al. 2005). GO graphs were calculated with a cutoff set to 10. Additionally the contigs were used to carry out BLAST searches of selected genes of interest.

2.4 Probe synthesis for *in situ* hybridization

Total RNA preparations from 35 Pairs of antennules were carried out as described for RNAseq above. RNA solution was cleaned up using the Poly(A)Purist^{TM} MAG Kit (Ambion,

Kaufungen, Germany) according to the manufacturer's instructions. Clean poly-A-RNA was transcribed to cDNA using SMARTer^{TM} RACE cDNA Amplification Kit (Takara Bio Europe/Clontech, Saint-Germain-en-Laye, France) generating both, 5'RACE-Ready and 3'RACE-Ready cDNA. Based on the contigs from de novo assembled RNAseq data, primers were designed using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/) and reviewed by Oligo Calc (http://www.basic.northwestern.edu/biotools/OligoCalc.html). PCR was carried out using Advantage® 2 PCR kit (Takara Bio Europe/Clontech, Saint-Germain-en-Laye, France) according to manufacturer's instructions. Obtained PCR fragments were cloned and sequenced. Clones matching the sequence of the respective contig were selected and linearized following recommended protocols for generating in situ hybridization probes using the SP6/T7 RNA transcription system with DIG (digoxygenin) labeling (Roche Diagnostics, Risch, Switzerland). Probes were subsequently shortened to 600-800 nucleotides length using a carbonate buffer (80mM $NaHCO_3$, 120 mM Na_2CO_3, pH 10.2) following the protocol of (Angerer and Angerer 1992). All probes were synthesized in sense- and antisense direction and used in parallel in all in *situ* experiments.

2.5 in *situ* hybridization

Antennules were cut as described above and transferred to 4% paraformaldehyde in 0.1 M $NaCO_3$, pH 9.5 for 24 h. Fixed antennules were individually dissected in PBS (phosphate-buffered saline = 0.85% NaCl, 1.4 mM KH_2PO_4 , 8 mM Na_2 HPO_4 , pH 7.1, 0.03% Triton X 100) removing one side of the lateral cuticle. Afterwards we proceeded as described by (Schymura et al. 2010), replacing WM-HBL by Hybridization Buffer (50% formamide, 2× SSC, 10% dextran sulphate, 20 mg/ml yeast t-RNA, 0.2 mg/ml herring sperm DNA) due to better performance and DAP buffer by Detection Buffer (0.1 M TRIS pH 9.5, 0.1M NaCl, 0.01 M $MgCl_2$, pH 8) according to manufacturer's instructions to the HNPP Fluorescent Detection Set (Roche Diagnostics, Risch, Switzerland). Antennules were subsequently stained 10 min in SYTOX blue solution (1:2000 in PBS) and washed 3 times 10 min with PBS. After the last washing, preparations were mounted on glass slides in PBS-Glycerol (1:3), covered with cover-slips and sealed with nail polish. Signals were visualized using an LSM 510 Meta confocal microscope (Zeiss, Jena, Germany) and Zeiss image browser. Settings for laser intensity and detection were not changed between scanning samples with sense and antisense probes of the same IR.

153 **2.6 Statistics**

154 In situ signals of at least three independent runs were counted in at least five spindle like
155 complexes of OSNs per antennular region. The regions were defined as “base” (proximal
156 third of the last antennular segment), “mid” (medial third of the last antennular segment) and
157 “tip” (distal third of the last antennular segment). Regions were compared in a one-way
158 ANOVA using R studio version 2.12.2 (2011-02-25) (Copyright (C) 2011 The R Foundation
159 for Statistical Computing).

160 **2.7 Immunohistochemistry**

161 Antennules were cut and prepared as described above. After treatment with HCl as described
162 in (Schymura et al. 2010) preparations were washed in PBS, blocked overnight in blocking
163 solution (PBS with 2% BSA, 0.3% Triton X 100, 0.05% Na-azide), incubated shaking for 6 h
164 at room temperature and 2 days at 8 degrees Celsius with primary antibody against IR25a
165 kindly provided by Elizabeth Corey (Stepanyan et al. 2004; Corey et al. 2013). After washing
166 in PBS the secondary antibody coupled to fluorophore Alexa 488 was applied overnight.
167 Preparations were mounted and visualized as described above. For control, preparations were
168 incubated without primary antibody.

169

170 **3. Results**

171 **3.1 Sequencing and assembly**

172 RNAseq based transcriptomes allow access to expression information of genes. A high read
173 count results in a so called deep dataset, representing expression of most or all genes of a
174 given tissue sample. Here, Solexa/Illumina sequencing of antennular RNA generated
175 29,987,467 sequencing reads of 96 nucleotides length each. Assembly by CLC genomics
176 workbench 6, after Vector/Adapter scan and cleanup, resulted in a total of 73,235 contigs
177 above 200 nucleotides length with an N50 of 744 bases. All contigs were included in the
178 subsequent analysis.

179

3.2 GO annotation

Access to the complexity of a given tissue is made easier by applying a generalized vocabulary, connecting knowledge of processes and functions of particular molecules to a transferable network of terms. Such a network is provided by GO classification (Ashburner et al. 2000) and was applied to the data using BLAST2GO as described in (Conesa et al. 2005). 18,583 out of 73,235 contigs (25.37%) yielded a BLAST result after dynamic translation (BLASTX) while InterProScan identified known domains in 28,593 contigs (39.04%). This confirms our earlier investigations of the same tissue (Groh et al. 2014), extending the total number of transcripts but pointing out the lack of identified orthologues and functional data for their majority. **Figure 1** surveys the general distribution of assigned terms on the second and third level of the three categories "Cellular Component", "Molecular Function" and "Biological Process" (full list of terms and proportions in **supplementary table 1**). Comparable to literature data of *C. clypeatus* antennules and functionally similar tissues (Grosse-Wilde et al. 2011; Legeai et al. 2011; Groh et al. 2014), most of the contigs were assigned the term "cell" (44%) followed by "organelle" (29%) in the category of "Cellular Component" on 2nd level **(1A)**. On level 3, the most abundant term was "cell part" (39%) followed by "membrane bounded organelle" (22%) **(1B)**. Term distribution in the category of "Molecular Function" largely confirmed the trend of term distribution found in our previous study [Groh et al. 2014] but varied to small extent in the proportions. Most ESTs on level 2 were assigned to terms of enzymatic functions, like "binding" (46%) and "catalytic activity" (36%) **(1C)**. On level 3 a total of 44 terms was assigned to the data. Figure **1D** therefore depicts only terms with an abundance of at least 0.5% of total terms on that level, easing visibility. A full list of terms is given in **supplementary table 1**. The highest proportion of ESTs was found in the term "protein binding" (15%), followed by "hydrolase activity" (11%) and "nucleotide binding" (10%). Expectedly, the highest proportion of ESTs on level 2 "Biological Process" is assigned to cellular processes (24%) and metabolism (20%) in general **(1E)**. As the antennules main function is chemoreception, terms associated to sensory tasks like "signaling" (5%) and "response to stimulus" (5%) were present as described in sensory tissues of other arthropods (Grosse-Wilde et al. 2011). A total of 92 terms was assigned to the data in the category "Biological Process" on level 3. Again only terms with an abundance of at least 0.5% of total terms on that level are included in the depiction in **figure 1F**. The highest proportion of terms included ESTs involved in basic processes of metabolism

and its regulation, but also more sensory specific processes, like “signaling process” (2%) and “signaling pathway” (2%), as well as “response to abiotic stimulus” (0.7%).

3.3 Olfactory pathway

Studies in lobster indicate an involvement of G-protein coupled receptor (GPCR) signaling in the peripheral signal processing. **Supplementary table 2** displays a list of contigs connected to the GO term “GPCR signaling” with their respective similarity to database entry homologues. The identified contigs include GPCRs, subunits of the GPCR signaling cascade as well as regulators of this signaling process.

3.4 Identification of IRs

Recent investigations of the chemosensory gene repertoire in crustaceans indicated the ionotropic receptors (IRs) to be the only chemosensory receptors present in crustacean antennules (Corey et al. 2013; Groh et al. 2014). We used IR sequences from both studies to screen our RNAseq based dataset for homologues, revealing 46 candidate sequences, including the previously identified CclyIR1, IR2, IR3, IR4, IR5, IR93a and IR25a. Using RACE-PCR some of the contigs were verified and extended towards both ends; only IR93a was successfully extended this way to full length. Nevertheless, 22 of the novel IR candidate sequences are spanning at least two of the three characteristic transmembrane domains of IRs and can therefore be considered as unigenes. The dendrogram depicted in **figure 2** is based on multiple sequence alignments of the ion channel pore region of our 29 known and novel candidate IRs and reference sequences of IRs retrieved from NCBI, including *Daphnia pulex, Panulirus argus* and *Drosophila melanogaster*. Besides the IR-coreceptor IR25a and IR93a, an IR known to be broadly expressed in crustacean OSN Populations (Corey et al. 2013), all IRs of *C. clypeatus* formed distinct clusters apart from the antennal IRs of *D. melanogaster*. The largest cluster consisted of 14 *CclyIR* candidates and was neighboring a cluster of *D. pulex* divergent IRs. The second largest cluster was next to DmelIR40a and included both known non-coreceptor *P. argus* IRs and 9 CclyIR candidates. A group of 4 *CclyIR* candidates was situated between a cluster of *D. melanogaster* antennal IRs and *D. melanogaster* divergent IRs.

3.5 Expression of IRs in OSNs

Immunohistochemical experiments show expression of the IR coreceptor IR25a in virtually all OSN cell bodies of the last antennular segment (**figure 3A**). Dendrites are continuously labeled from the spindle-like OSN cell body complexes to the aesthetasc patch (**figure 3D**) and inside the aesthetasc cuticle (**figure 3B**). The control shows no labelling of any cells of the last antennular segment (**figure 3C**).

As most of the non-coreceptor IR candidates belong to the group containing the divergent and not the insect olfactory antennal IRs, we designed RNA probes to verify expression of the respective receptors in OSN cell bodies. Antisense probes of four IRs (IR25a, IR1, IR6 and IR26) repeatedly produced specific signals in distinct cells of several OSN clusters. Sense control probes never produced signals in any cells of the antennules (**figure 4,** small captions to the lower left in **A-C**). To evaluate the distribution of IR expressing cells, we counted signals in OSN clusters in the basal, medial and distal antennular region. IR1 was expressed in all OSN clusters seen in the last antennular segment (**figure 4Aa, b**). In the elongated cell body complexes of the distal third of the antennules only 2-4 IR1 expressing cells could be found, which is significantly less than in both other regions (**figure 4Ac**). Signals of cells expressing IR6 were only detected in the distal third, with numbers varying between one and 5 cells per cluster (**figure 4B**). IR26 was not expressed in each OSN cell body complex but in most of the clusters of the proximal two third of the antennules with 1 to 2 cells per cluster. No IR26 expressing cells could be detected in the distal third (**figure 4C**).

4. Discussion

This study shows for the first time that the terrestrial hermit crab *Coenobita clypeatus* expresses a subclass of Ionotropic Receptors in its olfactory sensory neurons that is not expressed in insect OSNs, but presumably has the same function as the antennal IRs of insects.

The current expansion of our earlier transcriptome analysis confirmed the representation of genes in the tissue, allowing annotation by BLAST for only 25% of the contigs (for comparison see [Groh et al, 2014]). Nevertheless, the total contig count and average contig length were substantially higher than in our previous study, likely due to increased read

number. For general comparison with our earlier data we again performed BLAST2GO based annotation with parameters set identical to our earlier study. While the number of contigs assigned to one term was expectedly increased, the proportional distribution of contigs to terms on a given level was highly similar although not identical. With few exceptions, the contig proportions assigned to a term on any given level was smaller than 2%, though the total number of assigned contigs was three- to tenfold higher [see representative data of the category "Biological Process" in **supplementary table 3**, Groh et al 2014]. This finding confirms the assumption that a higher read number and thus better coverage of a transcriptome leads to a better resolution concerning single ESTs and gets closer to representing the complete genetic makeup of a given tissue.

As the antennules are the main olfactory organ of crustaceans we searched for representatives of genes involved in olfaction and olfactory processing in the hermit crabs antennules. IRs are the only known putative chemosensory receptors expressed in crustacean antennules (Corey et al. 2013; Groh et al. 2014). Until now, there was indication that the so called "divergent" subgroup of this receptor family is expressed in crustacean antennules, while only "antennal" IRs are olfactory in insects (Croset et al. 2010; Benton et al. 2009). Furthermore, the expression of the two antennal IR co-receptors IR25a and IR93a has been demonstrated in OSNs of the lobster *Panulirus argus* (Corey et al. 2013). Our study demonstrates that the IR co-receptor IR25a is expressed also at least in apparently all *Coenobita clypeatus* OSN cell bodies and is further present along the OSN dendrites to the aesthetascs. This makes it even more likely, that olfaction in terrestrial hermit crabs is based on IRs as only chemosensory receptors. Homology searches based on known crustacean IRs led to identification of 22 novel divergent IR candidates expressed in the *C. clypeatus* antennules, rising the known number of antennaly expressed IRs to 29 and accentuating their importance. Our whole mount in *situ* hybridization experiments showed not only the expression of three exemplary IRs in the cell bodies of OSNs of *C. clypeatus* but also the distribution pattern of expressing cells along the last antennular segment. Based on olfactory responses of isolated annuli it has been assumed that crustacean antennules are compound noses (Steullet et al. 2000). At least the herein investigated IRs are not homogeneously distributed along the aesthetasc bearing segment, but differ in numbers per OSN cell body complex. Nevertheless, the ratio of aesthetascs innervated by dendrites from one OSN cell body complex is not known. It might as well vary between the basal-, medial- and distal region of the last antennular segment, though the total number of cell bodies per complex is

304 similar. Furthermore, the aesthetascs at the antennular tip are considered to be senescent, as
305 the tip annuli of the lobster antennulae did not respond to any odor tested (Steullet et al.
306 2000). The expression of the IR coreceptor IR25a and presumably odor specific IRs in the
307 OSN cell body clusters connected to the tip aesthetascs indicates a probable functionality of
308 those in *C. clypeatus*. We therefore state that crustaceans, independently from insects,
309 recruited own subgroups of the Ionotropic Receptor family to serve as only chemosensory
310 receptors in their antennules.

311 Angerer, L. M., and R. C. Angerer. 1992. “In Situ Hybridization to Cellular RNA with
312 Radiolabeled RNA Probes.” In *In Situ Hybridization*, edited by D. G. Wilkinson, 17th ed.
313 Oxford: IRL Press.

314 Ashburner, M, C A Ball, J A Blake, D Botstein, H Butler, J M Cherry, A P Davis, et al. 2000.
315 “Gene Ontology: Tool for the Unification of Biology.” *Nature Genetics* 25 (1): 25–29.
316 doi:doi:10.1038/75556.

317 Benton, Richard, K S Vannice, C Gomez-Diaz, and L B Vosshall. 2009. “Variant Ionotropic
318 Glutamate Receptors as Chemosensory Receptors in Drosophila.” *Cell* 136 (1): 149–162.
319 doi:DOI 10.1016/j.cell.2008.12.001.

320 Bliss, D E, and L H Mantel. 1968. “Adaptations of Crustaceans to Land - a Summary and
321 Analysis of New Findings.” *American Zoologist* 8 (3): 673–685.

322 Conesa, A, S Gotz, J M Garcia-Gomez, J Terol, M Talon, and M Robles. 2005. “Blast2GO:
323 A Universal Tool for Annotation, Visualization and Analysis in Functional Genomics
324 Research.” *Bioinformatics* 21 (18): 3674–3676. doi:10.1093/bioinformatics/bti610.

325 Corey, Elizabeth a, Yuriy Bobkov, Kirill Ukhanov, and Barry W Ache. 2013. “Ionotropic
326 Crustacean Olfactory Receptors.” *PloS One* 8 (4) (January): e60551.
327 doi:10.1371/journal.pone.0060551.

328 Croset, V, R Rytz, S F Cummins, A Budd, D Brawand, H Kaessmann, T J Gibson, and R
329 Benton. 2010. “Ancient Protostome Origin of Chemosensory Ionotropic Glutamate Receptors
330 and the Evolution of Insect Taste.” *PLoS Genetics* 6 (8). doi:e1001064
331 10.1371/journal.pgen.1001064.

332 Eder, S, and J Atema. 1978. “Behavioral Evidence for Functional Separation of Taste and
333 Smell in Lobster, Homarus Americanus.” *Biological Bulletin* 155 (2): 435.

334 Ghiradella, H T, J F Case, and J Cronshaw. 1968. “Structure of Aesthetascs in Selected
335 Marine and Terrestrial Decapods - Chemoreceptor Morphology and Environment.” *American*
336 *Zoologist* 8 (3): 603–&.

337 Glaessner, M. F. 1969. “Decapoda.” In *Treatise on Invertebrate Palaeontology*, edited by K.
338 S. Lawrence, 400–533. Geological Society of America and Kansas University Press.

339 Gleeson, R. A. 1982. “Morphological and Behavioral Identification of the Sensory Structures
340 Mediating Pheromone Reception in the Blue Crab , Callinectes Sapidus.” *Biological Bulletin*
341 163 (1): 162–171.

342 Groh, Katrin C, Heiko Vogel, Marcus C Stensmyr, Ewald Grosse-Wilde, and Bill S Hansson.
343 2014. “The Hermit Crab’s Nose – Antennal Transcriptomics.” *Frontiers in Neuroscience* 7.
344 doi:10.3389/fnins.2013.00266.

345 Grosse-Wilde, E, L S Kuebler, S Bucks, H Vogel, D Wicher, and B S Hansson. 2011.
346 “Antennal Transcriptome of Manduca Sexta.” *Proceedings of the National Academy of*
347 *Sciences* 108 (18): 7449–7454. doi:10.1073/pnas.1017963108.

Hansson, Bill S, Steffen Harzsch, Markus Knaden, and Marcus Stensmyr. 2011. "The Neural and Behavioral Basis of Chemical Communication in Terrestrial Crustaceans." In *Chemical Communication in Crustaceans*, edited by T Breithaupt and M Thiel, 149–173. Springer Science + Business Media. doi:10.1007/978-0-387-77101-4_8.

Koczan, Stefan. 2012. "Diploma Thesis: 3D Reconstruction of the Olfactory Pathway from the Aesthetasc to the Olfactory Lobe in the Hermit Crab, *Coenobita clypeatus*". Friedrich-Schiller University Jena - Faculty of Biology and Pharmacy.

Koehl, M A R, Jeffrey R Koseff, John P Crimaldi, Michael G McCay, Tim Cooper, Megan B Wiley, and Paul A Moore. 2001. "Lobster Sniffing: Antennule Design and Hydrodynamic Filtering of Information in an Odor Plume." *Science* (Washington D C) 294 (5548): 1948–1951.

Krång, Anna-Sara, Markus Knaden, Kathrin Steck, and Bill S Hansson. 2012. "Transition from Sea to Land: Olfactory Function and Constraints in the Terrestrial Hermit Crab *Coenobita clypeatus*." *Proceedings of the Royal Society Biological Sciences* 279 (1744) (September 7): 3510–9. doi:10.1098/rspb.2012.0596.

Legeai, F, S Malpel, N Montagne, C Monsempes, F Cousserans, C Merlin, M C Francois, et al. 2011. "An Expressed Sequence Tag Collection from the Male Antennae of the Noctuid Moth *Spodoptera Littoralis*: A Resource for Olfactory and Pheromone Detection Research." *BMC Genomics* 12 (86). doi:10.1186/1471-2164-12-86.

Little, Colin. 2009. "The Terrestrial Invasion." Cambridge University Press.

Missbach, C, H K M Dweck, H Vogel, A Vilcinskas, M C Stensmyr, B S Hansson, and E Grosse-Wilde. 2014. "Evolution of Insect Olfactory Receptors." *eLife. (accepted for publication*

Penalva-Arana, D C, M Lynch, and H M Robertson. 2009. "The Chemoreceptor Genes of the Waterflea *Daphnia pulex*: Many Grs but No Ors." BMC Evolutionary Biology 9 (79). doi:79 10.1186/1471-2148-9-79.

Powers, L W, and D E Bliss. 1983. "Terrestrial Adaptations." In *The Biology of Crustacea*. Volume 8. Environmental Adaptations., edited by F J Vernberg and W B Vernberg, 271–333. Academic Press, New York, London etc.

Richter, Stefan, Ole S Moller, and Christian S Wirkner. 2009. "Advances in Crustacean Phylogenetics." Arthropod Systematics & Phylogeny 67 (2): 275–286.

Robertson, H M, C G Warr, and J R Carlson. 2003. "Molecular Evolution of the Insect Chemoreceptor Gene Superfamily in Drosophila Melanogaster." *Proceedings of the National Academy of Sciences of the United States of America* 100: 14537–14542. doi:10.1073/pnas.2335847100.

Rosen, Elyssa, Birgit Schwarz, and A Richard Palmer. 2009. "Smelling the Difference: Hermit Crab Responses to Predatory and Nonpredatory Crabs." *Animal Behaviour* 78 (3): 691–695. doi:10.1016/j.anbehav.2009.05.035.

386 Schymura, D., M. Forstner, A. Schultze, T. Kröber, L. Swevers, K. Iatrou, and J. Krieger.
387 2010. "Antennal Expression Pattern of Two Olfactory Receptors and an Odorant Binding
388 Protein Implicated in Host Odor Detection by the Malaria Vector *Anopheles gambiae*."
389 *International Journal of Biological Sciences* 6 (7) (January): 614–26.

390 Silbering, A. F, R. Rytz, Y. Grosjean, L. Abuin, P. Ramdya, G. S X E Jefferis, and R.
391 Benton. 2011. "Complementary Function and Integrated Wiring of the Evolutionarily
392 Distinct Drosophila Olfactory Subsystems." *The Journal of Neuroscience* : The Official
393 Journal of the Society for Neuroscience 31 (38) (September 21): 13357–13375.
394 doi:10.1523/JNEUROSCI.2360-11.2011.

395 Small M P, and Robert W Thacker. 1994. "Land Hermit Crabs Use Odors of Dead
396 Conspecifics to Locate Shells." *Journal of Experimental Marine Biology and Ecology* 182
397 (2): 169–182.

398 Stensmyr, M C, S Erland, E Hallberg, R Wallen, P Greenaway, and B S Hansson. 2005.
399 "Insect-like Olfactory Adaptations in the Terrestrial Giant Robber Crab." *Current Biology* 15
400 (2): 116–121. doi:10.1016/j.cub.2004.12.069.

401 Stepanyan, R, B Hollins, S E Brock, and T S McClintock. 2004. "Primary Culture of Lobster
402 (Homarus Americanus) Olfactory Sensory Neurons." *Chemical Senses* 29 (3): 179–187.
403 doi:10.1093/chemse/bjh023.

404 Steullet, P, H S Cate, W C Michel, and C D Derby. 2000. "Functional Units of a Compound
405 Nose: Aesthetasc Sensilla House Similar Populations of Olfactory Receptor Neurons on the
406 Crustacean Antennule." *Journal of Comparative Neurology* 418 (3): 270–280.

407 Thacker, Robert W. 1997. "Avoidance of Recently Eaten Food by Land Hermit Crabs,
408 Coenobita Compressus." *Animal Behaviour* 55: 485–496.

409

410 Figure legends

411 **Figure 1**: GO term distribution, A: Cellular component level 2, B: Cellular component level
412 3, C: Molecular Function level 2, D: Molecular function level 3, E: Biological Process level
413 2, F: Biological Process level 3

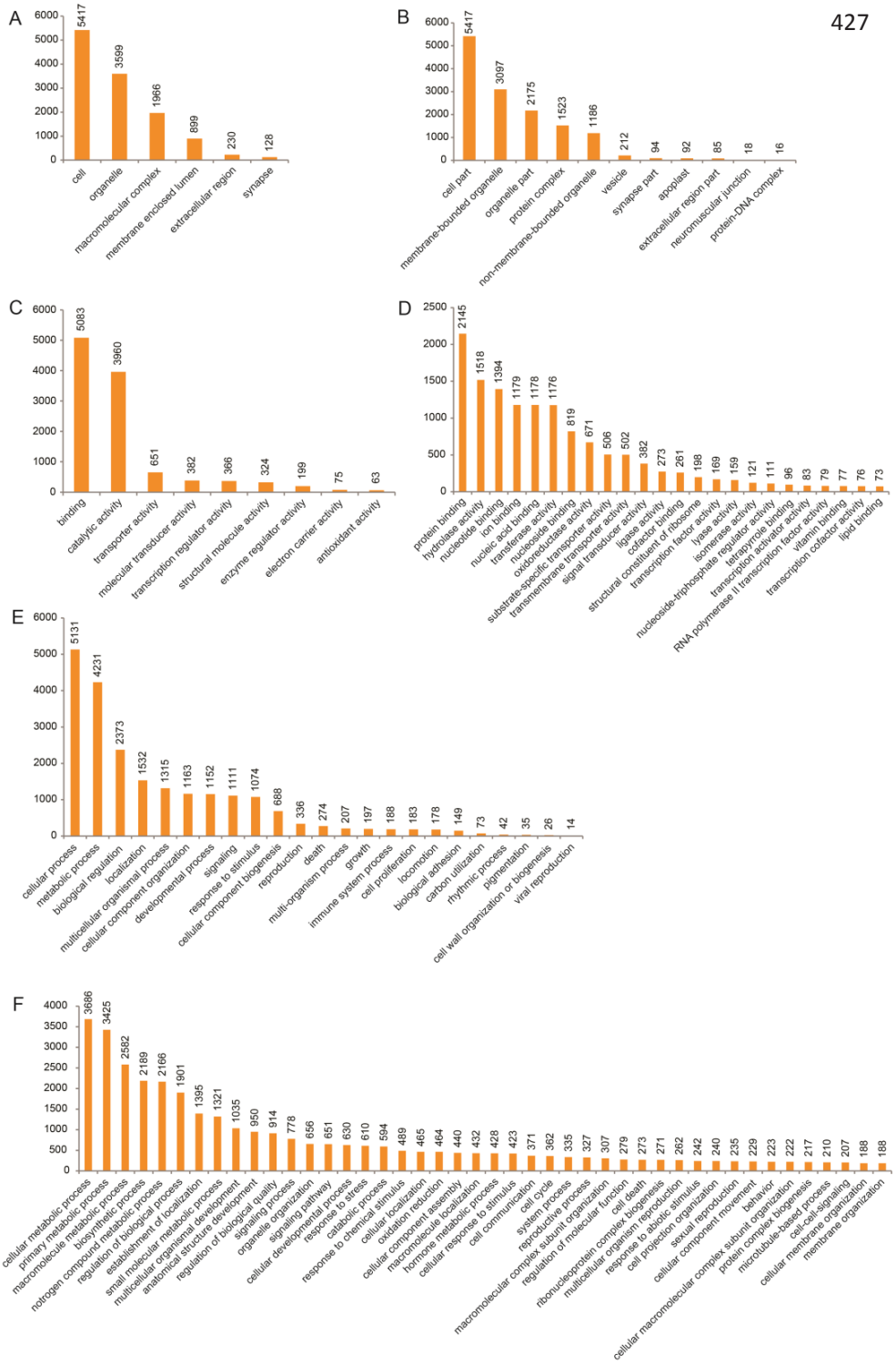
414 **Figure 2**:Dendrogram of ionotropic receptors, colored: IRs evidently expressed in olfactory
415 tissues/ organs green: insect antennal IRs, organge: *C. clypeatus* IRs, blue and pink: IR co-
416 receptors, grey: divergent IRs

417 **Figure 3**: LSM scan of whole mount immunohistochemical assay using IR25a antibody, A:
418 overview of OSN cell bodies expressing IR25a, B: IR25a in the aesthetasc dendrites, C:
419 closeup of an exemplary aesthetasc, D: Control without primary antibody, E: IR25a in
420 dendrites from cell bodies to the aesthetasc pad

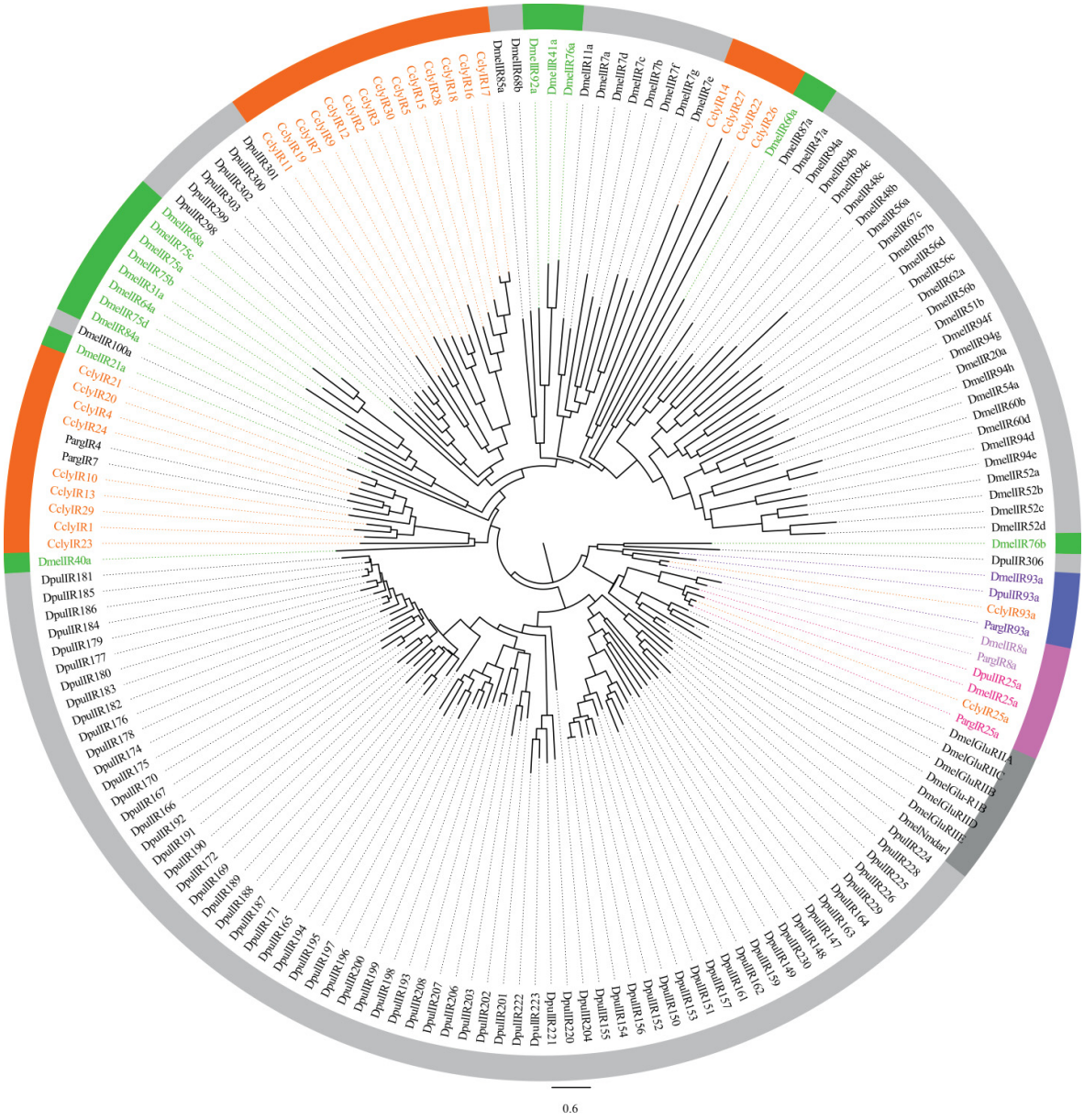
421 **Figure 4**: LSM scan of whole mount fluorescence RNA *in situ* hybridization, nuclear stain:
422 SyTox blue, Small letters refer to the respective region of higher magnification scan., signals
423 indicated by arrowheads; Boxplot: Signal count statistics and distribution significance of the
424 respective regions (one way ANOVA and Tukey's honestly significant difference test),
425 ***: $p<0.001$, .A: IR1, B: IR6, C: IR26

426 Figure 1

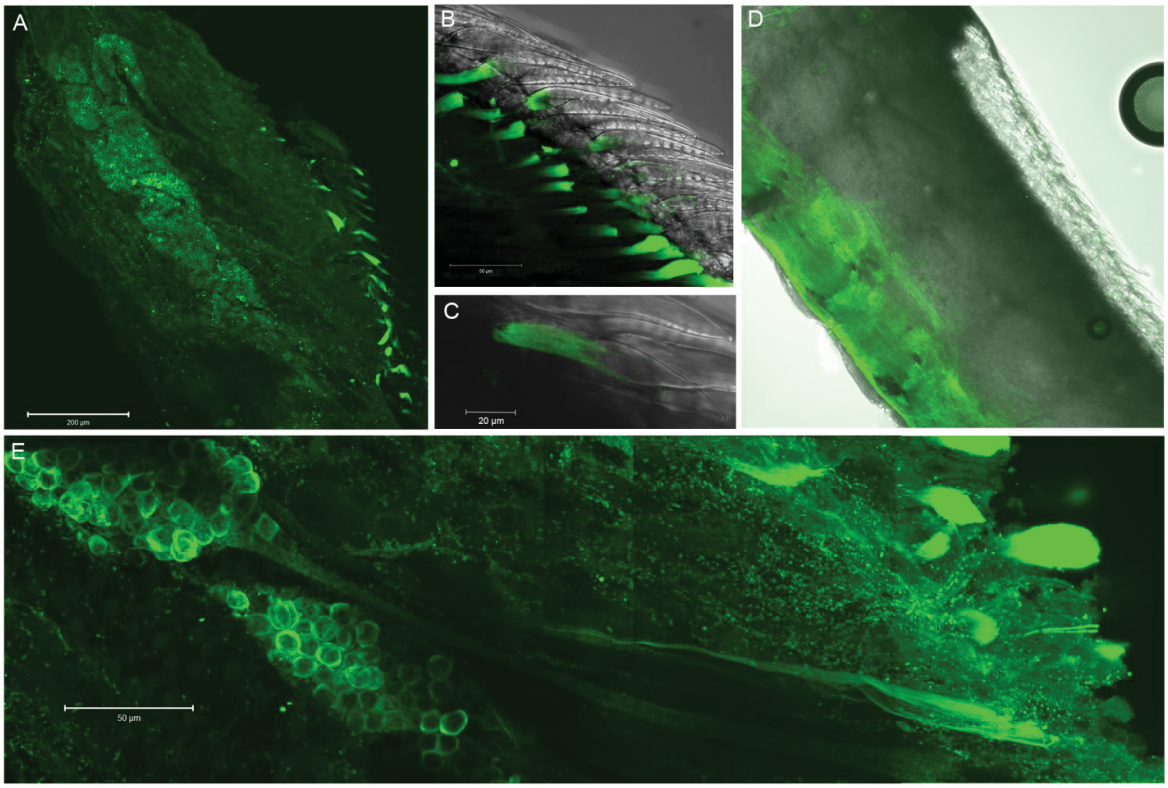
427



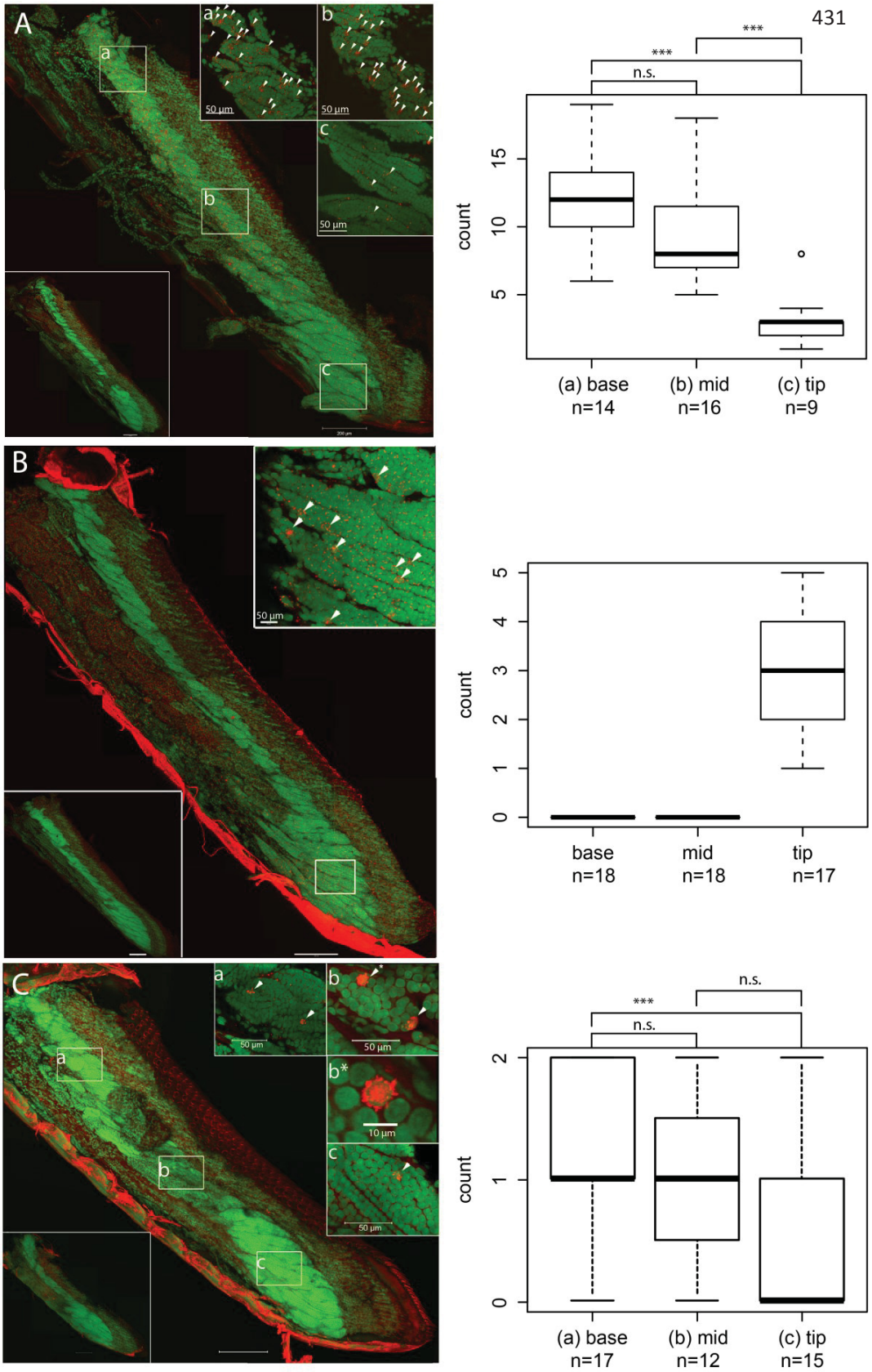
428 Figure 2


DmelIR68b
DmelIR85a
DmelIR92a
DmelIR41a
DmelIR76a
DmelIR11a
DmelIR7a
DmelIR7d
DmelIR7c
DmelIR7b
DmelIR7f
DmelIR7g
DmelIR7e
CclyIR14
CclyIR27
CclyIR22
CclyIR26
DmelIR60a
DmelIR47a
DmelIR94a
DmelIR94b
DmelIR94c
DmelIR48c
DmelIR48b
DmelIR56a
DmelIR67c
DmelIR67b
DmelIR56d
DmelIR56c
DmelIR62a
DmelIR56b
DmelIR51b
DmelIR94f
DmelIR94g
DmelIR20a
DmelIR94h
DmelIR54a
DmelIR60b
DmelIR60d
DmelIR94d
DmelIR94e
DmelIR52a
DmelIR52b
DmelIR52c
DmelIR52d
DmelIR76b
DpulIR306
DmelIR93a
DpulIR93a
CclyIR93a
PargIR93a
DmelIR8a
PargIR8a
DpulIR25a
DmelIR25a
CclyIR25a
PargIR25a
DmelGluRIIA
DmelGluRIIC
DmelGluRIIB
DmelGlu-RIIB
DmelGluRIID
DmelNmdar1
DpulIR224
DpulIR228
DpulIR225
DpulIR226
DpulIR229
DpulIR164
DpulIR163
DpulIR147
DpulIR148
DpulIR230
DpulIR149
DpulIR159
DpulIR162
DpulIR161
DpulIR157
DpulIR151
DpulIR153
DpulIR150
DpulIR152
DpulIR156
DpulIR154
DpulIR155
DpulIR204
DpulIR220
DpulIR221
DpulIR223
DpulIR222
DpulIR201
DpulIR202
DpulIR203
DpulIR206
DpulIR207
DpulIR208
DpulIR193
DpulIR199
DpulIR198
DpulIR200
DpulIR196
DpulIR197
DpulIR195
DpulIR194
DpulIR165
DpulIR171
DpulIR187
DpulIR188
DpulIR189
DpulIR169
DpulIR172
DpulIR190
DpulIR191
DpulIR192
DpulIR166
DpulIR167
DpulIR170
DpulIR175
DpulIR174
DpulIR178
DpulIR176
DpulIR182
DpulIR183
DpulIR180
DpulIR177
DpulIR179
DpulIR184
DpulIR186
DpulIR185
DpulIR181
DmelIR40a
CclyIR23
CclyIR1
CclyIR29
CclyIR13
CclyIR10
PargIR7
PargIR4
CclyIR24
CclyIR4
CclyIR20
CclyIR21
DmelIR21a
DmelIR100a
DmelIR84a
DmelIR75d
DmelIR64a
DmelIR31a
DmelIR75b
DmelIR75c
DmelIR75a
DmelIR68a
DpulIR298
DpulIR299
DpulIR303
DpulIR302
DpulIR300
DpulIR301
CclyIR11
CclyIR19
CclyIR7
CclyIR9
CclyIR12
CclyIR2
CclyIR3
CclyIR30
CclyIR5
CclyIR15
CclyIR28
CclyIR18
CclyIR16
CclyIR17
0.6

429 Figure 3



430 Figure 4



431

OPEN ACCESS Freely available online

PLOS ONE

Morphology and Histochemistry of the Aesthetasc-Associated Epidermal Glands in Terrestrial Hermit Crabs of the Genus *Coenobita* (Decapoda: Paguroidea)

Oksana Tuchina[1*9], Katrin C. Groh[1 9], Giovanni Talarico[2 9], Carsten H. G. Müller[3,4], Natalie Wielsch[5], Yvonne Hupfer[5], Aleš Svatoš[5], Ewald Grosse-Wilde[1], Bill S. Hansson[1]

1 Department of Evolutionary Neuroethology, Max Planck Institute for Chemical Ecology, Jena, Germany, **2** Department of Forensic Toxicology, Institute of Legal Medicine, University of Greifswald, Greifswald, Germany, **3** Department of Cytology and Evolutionary Biology, Ernst-Moritz-Arndt-University, Zoological Institute and Museum, Greifswald, Germany, **4** Department of Neuroscience, University of Arizona, Tucson, United States of America, **5** Research Group of Mass Spectrometry and Proteomics, Max Planck Institute for Chemical Ecology, Jena, Germany

Abstract

Crustaceans have successfully adapted to a variety of environments including fresh- and saltwater as well as land. Transition from an aquatic to a terrestrial lifestyle required adaptations of the sensory equipment of an animal, particularly in olfaction, where the stimulus itself changes from hydrophilic to mainly hydrophobic, air-borne molecules. Hermit crabs *Coenobita* spp. (Anomura, Coenobitidae) have adapted to a fully terrestrial lifestyle as adults and have been shown to rely on olfaction in order to detect distant food items. We observed that the specialized olfactory sensilla in *Coenobita*, named aesthetascs, are immersed in a layer of mucous-like substance. We hypothesized that the mucous is produced by antennal glands and affects functioning of the aesthetascs. Using various microscopic and histochemical techniques we proved that the mucous is produced by aesthetasc-associated epidermal glands, which we consider to be modified rosette-type aesthetasc tegumental glands known from aquatic decapods. These epidermal glands in *Coenobita* are multicellular exocrine organs of the recto-canal type with tubulo-acinar arrangement of the secretory cells. Two distinct populations of secretory cells were clearly distinguishable with light and electron microscopy. At least part of the secretory cells contains specific enzymes, CUB-serine proteases, which are likely to be secreted on the surface of the aesthetasc pad and take part in antimicrobial defense. Proteomic analysis of the glandular tissue corroborates the idea that the secretions of the aesthetasc-associated epidermal glands are involved in immune responses. We propose that the mucous covering the aesthetascs in *Coenobita* takes part in antimicrobial defense and at the same time provides the moisture essential for odor perception in terrestrial hermit crabs. We conclude that the morphological modifications of the aesthetasc-associated epidermal glands as well as the functional characteristics of their secretions are important adaptations to a terrestrial lifestyle.

Citation: Tuchina O, Groh KC, Talarico G, Müller CHG, Wielsch N, et al. (2014) Morphology and Histochemistry of the Aesthetasc-Associated Epidermal Glands in Terrestrial Hermit Crabs of the Genus *Coenobita* (Decapoda: Paguroidea). PLoS ONE 9(5): e96430. doi:10.1371/journal.pone.0096430

Editor: Fernando de Castro, Hospital Nacional de Parapléjicos – SESCAM, Spain

Received January 15, 2014; **Accepted** April 4, 2014; **Published** May 7, 2014

Funding: This study was supported by the Max Planck Society. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: otuchina@ice.mpg.de

9 These authors contributed equally to this work.

Introduction

Among hermit crabs (Decapoda, Anomura), Coenobitidae represents a perfect model to study the impact of terrestrialization on the structure and function of the olfactory system in invertebrates. From their earliest appearance in coastal habitats, which was 20 Mya according to the fossil record [1,2], coenobitids have adapted to a fully terrestrial lifestyle as adults [3]. The ability to permanently live ashore and locate distant food items or shell sources relies heavily on olfaction [4–6]. An essential prerequisite of good olfactory sense is the development of proper olfactory centers, which in coenobitids dominate the brain [7]. Long-distance detection of odors takes place on the external distal ramus of the antennules, called flagellum, which bears several rows of specialized olfactory sensilla, the aesthetascs [8]. In *Coenobita clypeatus*, each sensillum is peg-shaped and houses multiple ramified dendrites of approximately 300 olfactory sensory neurons [8,9]. The sensory cell somata are arranged in spindle-like complexes [8]. The thin cuticle of the exposed surface of the aesthetascs does not display any obvious pores, as is well-known for some types of olfactory sensilla of various myriapods and insects [10–12]. Thus, the question how airborne odor molecules penetrate the aesthetasc cuticle in crustaceans is not answered yet. However, numerous openings in the cuticle closely associated with the aesthetascs have been found in marine decapods [13]. These pores were identified as openings of epidermal exocrine glands, also referred to as aesthetasc tegumental glands in *Panulirus argus* (Latreille, 1804), here abbreviated as ATGs [13]. The secretory cells in the ATGs of *P. argus* were shown to contain a specific type of proteases – CUB-serine proteases (Csp) [13]. Serine proteases are a diverse family of trypsin- and chymotrypsin-like enzymes, which contain two amino acid residues, histidine and

aspartate, to stabilize the third active center, serine; together they form a catalytic triad acting on a substrate [14]. CUB is an extracellular protein domain, named after three proteins from which it was first identified: complement subcomponents Clr/Cls, embryonic sea urchin protein, Uegf, or fibropellin, and bone morphogenetic protein1, Bmpl [15]. CUB possesses four conserved cysteine residues, which form two disulfide bridges [15,16].

We wondered whether hermit crabs of the genus *Coenobita* also possess aesthetasc-associated epidermal glands and corresponding porous structures in their antennules.

Arthropod epidermal exocrine glands are well-known in crustaceans and may consist of a single secretory cell (class-I-glands according to Quennedey, [17]) or several specialized cell types, such as secretory cells and canal cells which guide the secretion outside via a cuticularized duct (class-III-glands [17]). Class-III-glands may remain solitary or may form glandular organs of higher complexity [12,18]. Most recently, Müller et al. [12,18,19] distinguished two subclasses of class-III-glands, flexo-canal and recto-canal epidermal glands; both types share an intermediary cell linking the secretory cell(s) and canal cell(s). Flexo-canal glands are characterized by meandering (flexuous) part of the conducting canal [19], while the recto-canal glands have straight-running conducting canals which locally widen to form huge reservoirs. Recto-canal glands are widely distributed among crustacean taxa and show great structural and functional diversity (reviewed in [18,20]). Glandular units may contain two, three or numerous cells. In the latter case, secretory, intermediary, and canal cells are often clustered in rosette-like formation; syncytial and lobed glandular complexes have also been reported reaching deeply below the epidermis of the head, mouth parts, pleopods, eyestalks, midgut and antennae in many different crustacean species (see [20,21] for details). The ubiquity of epidermal glands suggests that they play a pivotal role in the interaction of an organism with the environment. However, the exact function of investigated epidermal glands remains unknown in most cases.

Here, we set out to analyze the antennular glands of *Coenobita* species in a broad methodological approach ranging from an ultrastructural investigation comprising SEM and TEM techniques to standard methylene blue histology, histochemistry with nuclear marker sytox green, application of F-actin marker Phalloidin, cLSM imaging of fluorescent dyes around the conducting canals of the glands, applying backfilling techniques, immunohistochemistry with CUB-serine protease antibodies, and proteomic analysis of the glandular tissue. Besides providing a thorough description of the anatomy of the antennular glands, we also intend to gain insights into possible function of these glands in *Coenobita*.

Material and Methods

Animals

This study combines observations on the first antennae (antennules) of *Coenobita clypeatus* (Herbst, 1791), *C. scaevola* (Forskål, 1775) and *C. compressus* (H. Milne-Edwards, 1836). *C. clypeatus* was obtained from the "Zoologischer Groβhandel Peter Hoch" (August Jeanmaire Str. 12, 79183 Waldkirch, Germany; http://www.hoch-rep.com/). *C. compressus* was collected nearby Playa Naranjo in Santa Rosa Nacional Parque, Costa Rica, in May 2008. Permits were given by SINAC and MINAE, respectively ("Resolucion de investigacion cientifica", document: ACG-Pl-021-2008). *C. scaevola* was collected on the beach of Dahab (Sinai Peninsula, Egypt). Permits were provided by EEAA. Prior to all dissections, animals were anesthetized on ice for 10–30 min and removed from their shells.

Scanning electron microscopy (SEM)

SEM was performed on the large flagella of two specimens of *C. clypeatus*. After rinsing animals with tap water, the antennules were removed; their large flagella were cut off and were immediately placed in 50% ethanol. All samples were dehydrated in a graded series of ethanol (60, 70, 80% two times 10 min each; 90%, 96% for 10 min each; absolute EtOH overnight), critical-point-dried using a BAL-TEC CPD 030, mounted on aluminum stubs with adhesive film, and sputter-coated with gold on a BAL-TEC SCD005 prior to examination with a LEO 1450 VP scanning electron microscope, operated at 10 kV (Carl Zeiss AG, Oberkochen, Germany).

Transmission electron microscopy (TEM)

TEM was performed on the large flagella of *C. clypeatus*, *C. scaevola* and *C. compressus*. Two antennules of *C. clypeatus* were cut off and their large flagella were dissected to smaller subunits in cold fixative (2.5% glutaraldehyde either in phosphate or cacodylate buffer) and prefixed for 12 h at 4°C in the same solution. Samples were rinsed three times for 10 min with chilled buffer and postfixed in solutions containing either 2% osmium tetroxide (OsO_4) in phosphate buffer or 1% OsO_4 in cacodylate buffer for 2 h at 4°C. After rinsing three times for 10 min with chilled buffer, the samples were dehydrated in graded series of ethanol (see above). Dehydrated samples were embedded in epoxy resin (Durcupan) and polymerized for 20 h at 65°C. 30 antennules of *C. compressus* and 12 of *C. scaevola* were fixed in a cold solution modified after Karnovsky [22] containing 2.5% glutaraldehyde, 2.5% paraformaldehyde, 1.5% NaOH, and 1.2 g D-glucose dissolved in 2.25% sodium phosphate buffer (adjusted to pH 7.4) for 6 h (*C. compressus*) and 7d (*C. scaevola*), afterwards antennules were washed several times in the same buffer solution, dehydrated in a graded series of ethanol, and finally embedded in epoxy resin (Araldite). Ultrathin sections at thickness of 50 to 70 nm were cut with a Diatome diamond knife (Ultra 35°) on a Reichert Ultracut microtome. Sections were collected on Pioloform-coated single slot copper grids and examined without additional staining with a Zeiss CEM 902A (with a TVIPS FastScan digital camera) transmission electron microscope, operating at 10 kV (Carl Zeiss AG, Oberkochen, Germany).

Light microscopy (LM)

For orientation, semithin sections (150 to 400 nm in thickness) were transferred to glass-slides and stained with methylene blue-borax/Azure II according to Richardson et al., [23]. Sections were observed either on a Zeiss AxioImager Z.1 with an AxioCam HRc (Carl Zeiss AG, Jena, Germany) or on a Zeiss Axioscope 50 (Carl Zeiss AG, Jena, Germany) with a PixeLINK PL-B623CF-KIT 3.0 MP FireWire Camera (PixeLINK, Ottawa, Canada).

Confocal laser-scanning microscopy (cLSM)

In order to fully visualize the general distribution of the glandular ducts a crystal of micro-ruby (dextran tetramethylrhodamine-biotin, MW 3000, lysine fixable, Invitrogen) was inserted into the third annulus of the antennular flagellum using a glass capillary. After at least 3 h of exposition to the dye, antennules were cut and immediately fixed with 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS, pH 7.3), for no less than 2 h at room temperature on a shaker or overnight at 4°C. After fixation, tissues were washed in several changes of PBS for 2 h in total, the

cuticle was entirely or partially removed, and then tissues were dehydrated in a graded series of ethanol (50%, 70%, 80%, 2×99%, 15 min each), finally immersed in methyl salicylate and observed under the fluorescence stereomicroscope Leica MZ16FA (Leica Microsystems GmbH, Wetzlar, Germany) and the laser-scanning microscope Zeiss LSM 510 Meta (Carl Zeiss GmbH, Jena, Germany). Pictures were obtained using LSM image browser.

Histochemistry

We used phalloidin for selective staining of actin in the canal cells of the glands of 15 specimens of *C. clypeatus*, either alone or combined with nuclear stain (i.e. double-staining). Antennules were first fixed in 4% PFA in PBS for 2 h at room temperature on a shaker, then washed in several changes of PBS with 0.5% Triton X (PBST) for no less than 2 h and finally incubated in phalloidin Alexa 546 (1:50; Molecular Probes) overnight at 4°C on a shaker. After incubation, tissues were washed again in PBST and in case of double-staining incubated with sytox green (1:1000; Invitrogen). During incubation with sytox, tissues were checked time to time under the fluorescence stereomicroscope Leica MZ16FA (Leica Microsystems, Wetzlar, Germany) until the degree of staining of cell bodies was considered sufficient (usually approx. 20 min). Sometimes, a partial removing of the antennular cuticle was needed for better penetration of the dyes. Then, tissues were washed in PBS for at least 30 min and further processed in a graded series of ethanol and methyl salicylate as described above.

Csp immunoreactivity

The polyclonal Csp antibodies (5th bleed serum [99-5]) generated against the CUB-domain of *Panulirus argus* serine protease) were kindly provided by Prof. Dr. Manfred Schmidt and Prof. Dr. Charles D. Derby (see [13,16]). To test for Csp immunoreactivity, antennules of 5 specimens of *C. clypeatus* were first fixed in fresh 4% PFA as described above, and washed in several changes of PBS. The cuticle was removed completely and antennular tissues (whole mounts) were processed according to standard protocol for immunohistochemical labeling (see [7]). Tissues were incubated overnight in 4% NGS (Normal goat serum, Invitrogen) on PBS-Tx (0.5% TritonX, Sigma-Aldrich), washed several times in changes of 0.5% NGS on PBS-Tx and then incubated with primary antibody (5th bleed serum [99-5] rabbit, 1:400) for 3d at 4°C. After antibody incubation the tissues were rinsed several times in changes of 0.5% NGS on PBS-Tx and then incubated with the secondary antibody (donkey anti-rabbit Cy3, Jackson ImmunoResearch Lab, 1:1000) overnight. Subsequently, the tissues were washed in PBS for at least 2 h and incubated with nuclear marker sytox green (Invitrogen, 1:500) for 30 min, then washed in PBS for at least 1 h, dehydrated in a graded series of ethanol (70, 80, 90, 2×100%, 15 min each), cleared in methyl salicylate and studied under the confocal laser-scanning microscope. Simultaneously, we performed a control experiment by incubating the tissues for 3d in 0.5% NGS PBS-Tx, but omitted incubation with primary antibodies. Several antennules prepared for Csp-immunoreactivity experiments were left with their cuticle attached. These samples were cut in pieces of 2-3 annuli width and incubated with 10% EDTA for 1 week, with daily changes of the incubation solution (according to [13]), to test whether the incubation with EDTA affects staining. However, no differences have been observed between the samples incubated with EDTA and those which cuticle was removed mechanically by forceps.

Bioinformatics – BLAST searches and alignments

BLAST searches were carried out after dynamic translation (BLASTX) with a default E-value cutoff of 1.0E-6. Alignments were compiled using the MUSCLE alignment tool [24].

Sample preparation for proteomic analysis

The antennules of 30 anesthetized specimens of *C. clypeatus* were dissected in Lysisbuffer (7 M Urea, 2 M Thiourea, 2% 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate [CHAPS], 20 mM Tris(hydroxymethyl)-aminomethan [Tris]) containing 1 mM protease inhibitor (4-(2-Aminoethyl)benzensulfonylfluorid [AEBSF]). Two kinds of samples were taken: 1) the glandular tissue, 2) the antennular tissue containing olfactory sensory neurons and no glandular complexes (control sample). Both samples were homogenized in 1 ml Lysisbuffer for 3×20 sec with 10 sec intervals with Precellys Tubes Ceramic Beads (Precellys, 1.4 mm). The obtained lysate was transferred to polycarbonate-centrifuge vials (BECKMAN), and ultracentrifuged for 45 min at 100,000 g at 4°C. The supernatant was collected, concentrated using speed vac and 30 ul and used for SDS-PAGE.

In-gel tryptic digestion and peptide extraction

Purified proteins were separated by 1-D SDS-PAGE and visualized by staining with Coomassie Blue (R250). Protein bands were excised from the gel, cut into small slices, washed several times with 25 mM ammonium bicarbonate and destained with 50% ACN/25 mM ammonium bicarbonate. Disulfide bonds were reduced with 10 mM DTT at 50°C for 1 h and alkylated with 55 mM IAA at room temperature in the dark for 45 min. Following tryptic digestion in 0.5 µM solution (in 25 mM ammonium bicarbonate) of porcine trypsin (Promega) overnight at 37°C, the peptides were extracted from the gel pieces using 75% ACN/5% formic acid (FA), and dried down in a vacuum centrifuge [25]. For mass spectrometric analysis samples were reconstructed in 10 µL aqueous 0.1% FA.

Nano-UPLC-MS/MS analysis

The peptide mixtures (1 to 8 µL) were initially concentrated and desalted on a Symmetry C18 trap-column (20×0.18 mm, 5 µm particle size, Waters) using 0.1% FA as mobile phase at a flow rate of 15 µL/min. Then, peptides were separated on a nanoAcquity C18 analytical column (200 mm×75 µm ID, C18 BEH 130 material, 1.7 µm particle size, Waters) using an increasing acetonitrile gradient in 0.1% FA at a flow rate of 350 nl/min. The applied LC-gradient was: 1–30% B over 13 min, 30–50% B over 5 min, 50–95% B over 5 min, isocratic at 95% B for 4 min, and a return to 1% B over 1 min (phases A and B composed of 0.1%FA and 100% ACN in 0.1% FA, respectively); the analytical column was re-equilibrated for 9 min prior to the next injection.

The eluted peptides were on-line transferred via a NanoLock-Spray ion source into a Synapt HDMS tandem mass spectrometer (Waters). The source temperature was set to 80°C, cone gas flow 30 L/h, and the nanoelectrospray voltage was 3.2 kV. For all measurements, the mass spectrometer was operated in V-mode with a resolution power of at least 10,000 FWHM. All analyses were performed in positive ESI mode. The lockmass calibrant standard, human Glu-Fibrinopeptide B (650 fmol/mL in 0.1% FA/ACN (1:1 v/v)), was infused into the NanoLockSpray electrospray source at a flow rate of 500 nL/min through the reference sprayer every 30 sec to compensate for mass shifts in MS and MS/MS fragmentation mode.

LC-MS data were collected using MassLynx v4.1 software under data-dependent (DDA) and data-independent (DIA)/LC-

MS^E acquisition. For DDA, the acquisition cycle consisted of a survey scan covering the range of m/z 400–1700 Da followed by MS/MS fragmentation of the four most intense precursor ions collected at 1 sec intervals in the range of 50–1700 m/z. Dynamic exclusion was applied to minimize multiple fragmentations for the same precursor ions. For LC-MS^E analyses, full-scan LC-MS data were collected using alternating mode of acquisition: low energy (MS) and elevated energy (MS^E) mode at 1.5 sec intervals with a 0.2 sec inter-scan delay in the range m/z of 300–1900 and 50–1700, respectively. The collision energy of low energy MS mode and high-energy mode (MS^E) were set to 4 eV and 15–40 eV energy ramping, respectively.

Data processing and protein identification

ProteinLynx Global Server (PLGS) version 2.5.2 (Waters) was used for processing of raw files and for database searching. DDA raw files were initially baseline subtracted, smoothed, deisotoped, lock-mass corrected, and pkl-files were generated. Processed MS/MS spectra (pkl-files) were searched against the NCBI-nr database (updated on December 5, 2012, containing 21,786,050 sequence entries) combined with the *Coenobita* protein subdatabase (containing 300,210 entries, constructed from an in-house created EST database by its translation from all six reading frames [26] using MASCOT v2.4 software installed on a local server. Trypsin was set as the primary digest reagent, and one missed trypsin cleavage site was allowed. Mass tolerances for precursor and fragment ions were 15 ppm and 0.03 Da, respectively. A fixed modification of carbamidomethyl-Cys was specified, and oxidation-Met was set as a variable modification. Proteins matched by at least three peptides with ion scores above 30 or by one peptide with protein score of higher than 55 were considered as correct assignments.

In parallel, MS/MS spectra were searched using PLGS software as a search engine against a subdatabase containing common background proteins (human keratins and trypsin) and the unassigned spectra were sequenced *de novo*. *De Novo* sequencing was performed using following parameters: mass deviation, 0.002 Da, and ladder score>40. Originated peptide sequence proposals were subjected to sequence-similarity searching using the MS BLAST program installed on an in-house server. MS BLAST searches were performed against the complete NCBI-nr database (updated on December 5, 2012) as well as three subdatabases: *arthropoda*, *bacteria* (both downloaded from NCBI-nr, containing 853,629 and 13,530,307 sequence entries, respectively) and *Coenobita* (constructed as described above) using adopted settings parameters [27]. Statistical significance of the hits was evaluated according to the MS BLAST scoring scheme [27].

The continuum LC-MS^E data were lock-mass corrected, smoothed, background subtracted, centered, deisotoped, and charge state reduced using PLGS software. Product ion spectra were generated using following thresholds for low/high energy scan ions and peptide intensity: 150, 30, and 750 counts, respectively. Time-based alignment of precursor and fragment ions was done using Ion Accounting Algorithm as described [28]. Processed data were searched against the constructed Coenobita database (as described above) combined with the Swissprot database (downloaded on July 27, 2011 from http://www.uniprot.org/). Database searching was restricted to tryptic peptides with a fixed carbamidomethyl modification for Cys residues, along with variable oxidation of Met. Further, default searching parameter specifying mass measurement accuracy were used, minimum number of product ion matches per peptide (5), minimum number of product ion matches per protein (7), minimum number of peptide matches (2), and maximum number of missed tryptic cleavage sites (1). Maximum false positive rate was set to 2% and all peptides matched under the 2% FDR were considered as correct assignments.

All identified proteins of each dataset were combined in one list, removing duplicates. The list of proteins identified in the control section was subtracted from the list of proteins from the gland section; resulting in a list of proteins only present in the sample containing glandular complexes. Entries where contigs but no homologue was assigned based on peptides were searched against non-redundant databases of the National Center of Biotechnology Information (NCBI) after dynamic translation (BLASTX). BLAST results were added where applicable (E-Value$<10^{-6}$) (full list in Additional file 3). Entries were assigned to Gene Ontology terms in the category "Molecular Function" by manual search (http://www.geneontology.org/).

Results

Antennules, large flagella and aesthetasc pads

The antennules of decapods terminate in two morphologically distinct flagella. The larger one represents the lateral and the smaller one represents the medial flagellum of aquatic species. Terrestrial hermit crabs (Coenobitidae) predominantly hold and move their antennules in a way that the large flagellum is located dorsally, while the small flagellum is oriented ventrally. The large flagellum carries a distal pad equipped with peg-shaped olfactory sensilla, the aesthetascs, sitting on its ventral surface (Figure 1A). The aesthetascs are arranged in 5–7 rows (Figure 1B). The total number of aesthetascs per flagellum depends on the age and the size of an animal. The fine tips of the slightly bent aesthetascs predominantly point towards the tip of the flagellum (Figure 1B). Slender, minute setae occur at the lateral margins of the aesthetasc pad (Figure 1B–C, arrowheads) and are also sparsely distributed among the rows of aesthetascs (Figure 1B, arrows).

The aesthetascs are immersed in a layer of clear liquid, further referred to as "mucous-like substance"; the layer is clearly visible in observations using high power stereo microscopy. Ethanol treatment for SEM preparation fully removes this substance. Then, the surfaces of the observed SEM samples are apparently clean without displaying any conspicuous residues (compare Figure 1A–B, D–F). Applying the fixation protocol for TEM apparently supported better preservation of the mucous-like substance or components of it, because some of the observed samples reveal the presence of an osmiophilic substance surrounding and covering the aesthetascs (Figure 1C).

Organization and ultrastructure of aesthetasc-associated epidermal glands

Gland pores, canal cells and distal part of conducting canal. High-power scanning electron micrographs demonstrate the presence of numerous pores in the cuticle between the aesthetascs (Figure 1D, arrowheads) and on the lateral margins of the pad (Figure 1E). Albeit displaying different shapes, each gland pore possesses a collar-like fold surrounding the approx. 1 μm wide opening (Figure 1D–F). The pores occur predominantly alone (Figure 1D, E), but also paired arrangements are present (Figure 1F). The collar-like fold is formed by the cuticle protruding into the pore lumen (Figure 2A).

The thin, strongly electron-dense epicuticle extends deeply into the pore and lines the distal part of the conducting canal, or distal duct (Figure 2A, C). In cross-sections, distal ducts show a polymorphic outline (Figure 2C). The distal ducts pass through the cuticle of the aesthetasc pad (Figure 2B, arrowheads). Distal protrusions of the canal cells join the ducts in the middle portion of the cuticle (Figure 2D). At basal section level of the cuticle and

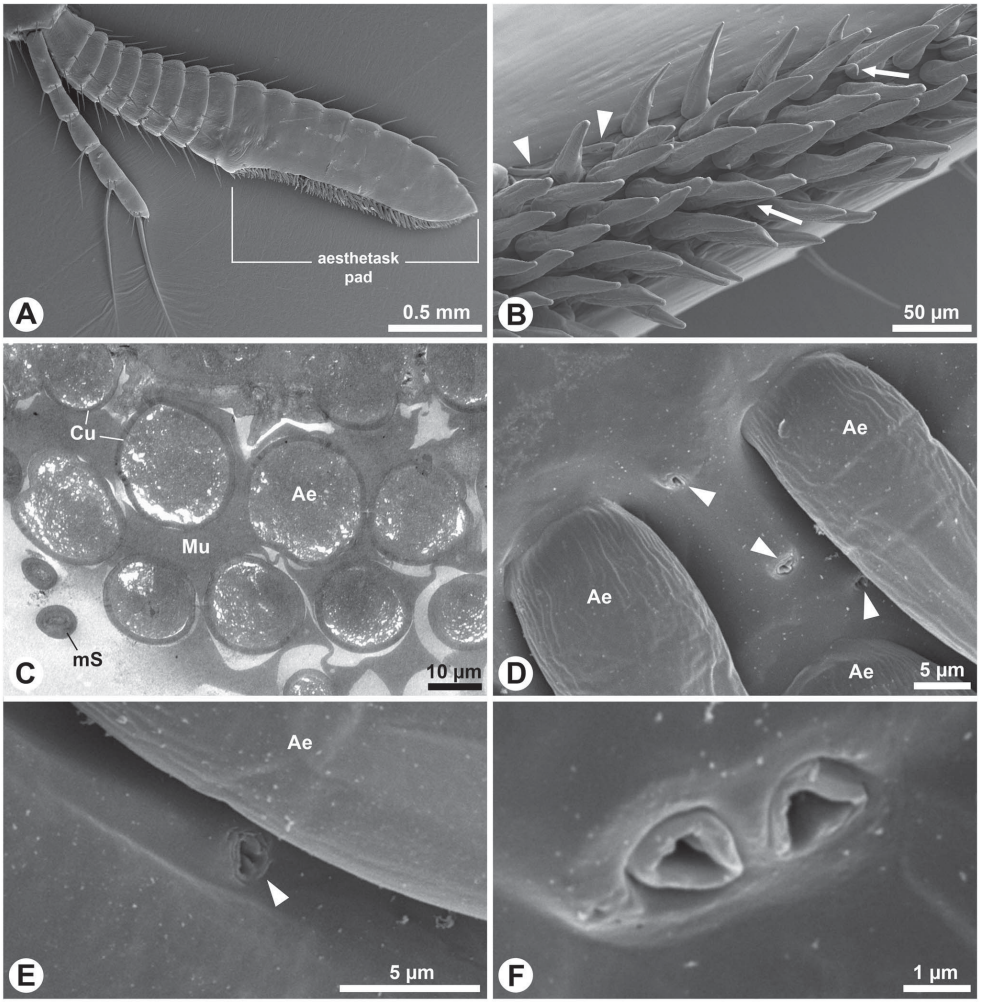


Figure 1. External (A–B, D–F) and internal (C) morphology of flagella, aesthetascs and antennal glands of *Coenobita*. A–B, D–F: SEM; **C**: TEM. **A**: Lateral view on the pair of flagella of an antennule of *C. clypeatus*. The larger dorsal flagellum carries the aesthetasc pad. **B**: Ventral view of the median part of the aesthetasc pad showing six rows of slightly bent, peg-shaped aesthetascs. Note minute setae at the lateral margin of the aesthetasc pad (arrowheads) and between the aesthetascs (arrows). **C**: Transverse section of the aesthetasc pad of *C. scaevola* not treated with fixative solutions. Mucous is present in the interspace of the aesthetascs. **D**: Detail of proximal margin of the aesthetasc pad in *C. clypeatus*. Note scattered gland pores (arrowheads) between adjacent aesthetascs. **E**: Solitary gland pore (arrowhead) from the lateral margin of the aesthetasc pad. **F**: Paired gland pore, note collar-like fold surrounding the opening. Ae, aesthetasc; Cu, cuticle; mS, minute seta; Mu, mucous.
doi:10.1371/journal.pone.0096430.g001

further below, each duct is fully surrounded by at least one canal cell (Figure 2E–F). The ducts pass diagonally through the hemolymphatic space of the flagellum (Figure 2G–I).

Secretory cells, intermediary cells and proximal part of conducting canal. The proximal part of the conducting canal (here also abbreviated as proximal duct) is formed by an intermediary cell and surrounded by numerous secretory cells. The secretory cells are clustered in a rosette-type formation, forming an elongated tubular acinus. Hence, the aesthetasc-associated epidermal glands in *Coenobita* species are of the compound tubulo-acinar type (Figure 3A–C). The acini are located above the layer of the olfactory sensory neuron cell somata; the dendrites of the latter innervate the aesthetascs (Figures 3B; 4A). The proximal ducts terminate as blind tubes inside the acini, and in different acini the proximal ducts are either unbranched (Figure 3E) or branched (Figure 3F). In high-power magnifications, outlets of the secretory cells are observed (insert Figure 3F). Notably, phalloidin staining was almost completely destroyed when incubated together with nuclear stain sytox green. The secretory cells release their products into the proximal duct (Figure 3G, arrowheads), which lacks an epicuticular intima, while the distal part of the conducting canal, lined with a thin epicuticular intima, does not have any connections to the secretory cells (Figure 3G).

Methylene blue and azure II histology (according to Richardson et al. [23]) reveals two distinct types of secretory cells (Figure 4A–C). Approximately two thirds of the acini consist of secretory cells with light cytoplasm and nuclei (further referred to as acini of type 1) and one third of the acini are stained distinctly darker (type 2). We observed the acini to be monotypic, i.e. comprising the

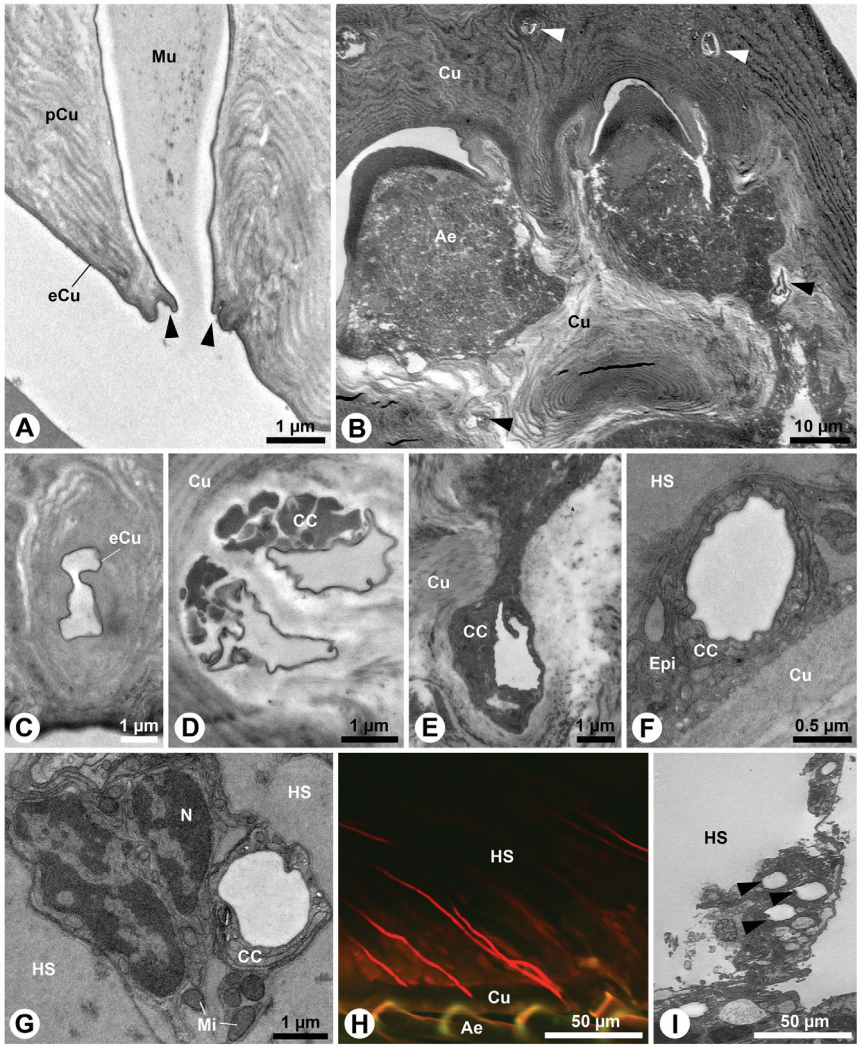


Figure 2. Aesthetasc-associated epidermal glands in *Coenobita*: pores, canal cells and distal duct. A–G, I: TEM; **H**: cLSM. **A**: Longitudinal section of a glandular pore of *C. compressus*. Note the collar-like fold (arrowheads). **B**: Horizontal section of aesthetasc pad of *C. clypeatus* showing two aesthetasc sockets and several cross-cut gland part of the conducting canal (arrowheads). **C**–**G**: Series of transverse sections showing the distal part of conducting canal on different section levels (from distal to proximal). **C**: One single duct passing through the cuticle of *C. clypeatus* closely below the glandular pore. **D**: Paired glandular ducts accompanied by cytoplasmic protrusions of the canal cells, on proximal level of the flagellar cuticle. *C. clypeatus*. **E**: Gland duct below the flagellar cuticle in *C. clypeatus*. **F**: Gland duct of *C. compressus* passing through the thin epidermis below the flagellar cuticle. The duct lumen is lined by thin epicuticle (cuticular intima). **G**: Duct passing through the hemolymphatic space in *C. compressus*. **H**: Optical sagittal section of the large flagellum of *C. clypeatus* showing several gland ducts (red; backfilling with tetramethylrhodamine dextran) in parallel orientation. **I**: Gland ducts of *C. clypeatus* (arrowheads). Ae, aesthetasc; CC, canal cell; Cu, cuticle; eCu, epicuticle; Epi, epidermis; HS, hemolymphatic space; Mi, mitochondrion; Mu, mucous; N, nucleus; pCu, procuticle.
doi:10.1371/journal.pone.0096430.g002

secretory cells of one type only. In transverse sections close to the aesthetasc pad, the glandular tissue occupies nearly 50% of the lumen of the large flagellum (Figure 4A). The glands can be recognized by their proximal ducts and the tubulo-acinar arrangements of the cells (Figure 4A–C). Acini have transversal diameters between 50–100 µm (Figure 4B), while they can be several hundreds of micrometers long (Figure 3B, D). Counting the total number of glandular acini in 15 antennules of *C. clypeatus* reveals enormous individual differences without any obvious relation to body size, number of annuli per flagellum or differences between females and males (Figure 5).

Although the morphological organization of both types of acini is generally similar, several ultrastructural differences exist and at least two of them are elucidated in the different staining results of the secretory cells visible in histological sections. Secretory cells with light cytoplasm have large, spherical nuclei with light caryoplasm and prominent nucleoli (Figure 6A). The nuclei are situated approximately in the centre of the cells (Figures 4B–C, 6A). The cytoplasm is densely packed with relatively small vesicles filled with an electron-lucent substance, most probably secretion (Figure 6A). The nuclei of the secretory cells with stronger osmiophilic cytoplasm are somewhat smaller (Figure 6A–B). They

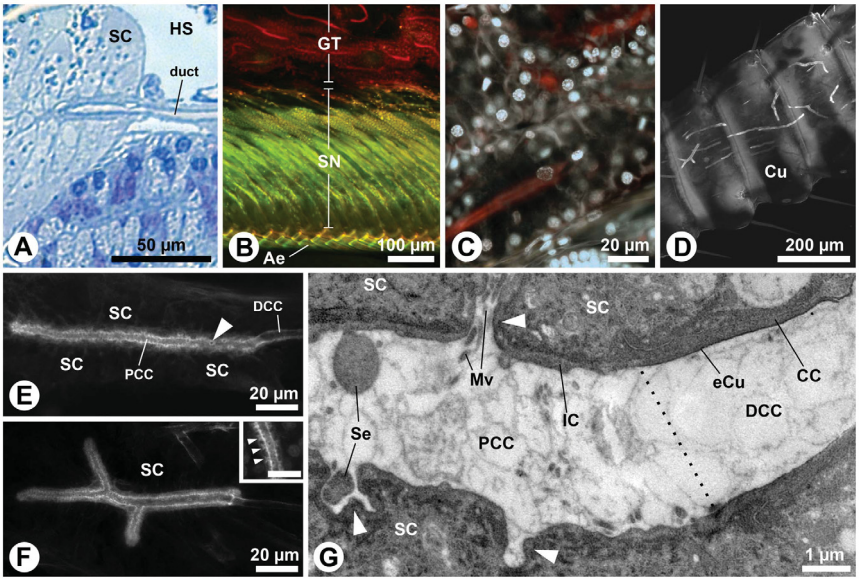


Figure 3. Aesthetasc-associated epidermal glands in *Coenobita clypeatus*: secretory cells, intermediary cells and the proximal duct. **A**: LM, **B-F**: cLSM, **G**: TEM. **A**: Oblique section of a duct entering the acinus. **B**: Optical sagittal section of the large flagellum with numerous proximal ducts of the aesthetasc-associated epidermal glands (red lines; phalloidin Alexa 546) located above the layer of sensory neurons cell bodies (green; sytox green). **C**: Phalloidin-labeled proximal ducts in higher detail (red; phalloidin Alexa 546), surrounded by the nuclei (white; sytox green) of the secretory cells. **D**: Antennomeres (annuli) of, proximal to the aesthetasc pad. Several proximal ducts (white lines; phalloidin Alexa 546) are visible through the cuticle. **E**: Unbranched proximal duct. **F**: Branched proximal duct; inset: higher magnification of the proximal duct showing outlets of the secretory cells (arrowheads, scale 10 µm). **G**: Longitudinal section of the proximal duct (left part of micrograph) not lined by a cuticle at the transition zone (indicated by dotted line) to the distal part of the conducting canal (right part of the micrograph) which is cuticle-lined. Note the orifices of the secretory cells opening into the proximal duct (arrowheads). Ae, aesthetasc; CC, canal cell; Cu, cuticle; DCC, distal part of the conducting canal; eCu, epicuticle; GT, glandular tissue; HS, hemolymphatic space; IC, intermediary cell; Mv, microvilli; PCC, proximal part of the conducting canal; SC, secretory cell; Se, secretion; SN, sensory neurons.
doi:10.1371/journal.pone.0096430.g003

are predominantly located in the basal portion of the cells (Figures 4B; 6A). Their nucleoli are inconspicuous and hardly visible. The cytoplasm is packed with relatively large vesicles filled with a slightly less electron-lucent secretion and with dense endoplasmic reticulum (ER), especially close to the nuclei (Figure 6A–B). All secretory cells have a conical, columnar shape with heights of 30–50 µm, basal diameters of 10–20 µm and apical diameters of 1–2 µm. Because of their alternating arrangement around the central intermediary cell, 10–15 cell bodies can be counted per transverse section (Figure 6A). The intermediary cells have average diameters of 10–20 µm (Figure 6A). The narrow apices of the secretory cells pass through the intermediary cells (Figure 5C). Since the proximal ducts are only about 5 µm wide, a maximum of 6 secretory cells can release their products into the duct lumen at the same transverse section level of the lumen (see Figure 6C). The secretory cells possess typical apical microvilli borders but no other peculiar apical structures like reservoirs, fibrous sieves or valves (Figure 6C). The microvilli are several micrometers long and can reach the centre of the lumen (Figure 6C).

In both types of acini, the secretion is first collected in the proximal duct and then transferred to the cuticle-lined distal part of the conducting canal (Figure 3G moderately osmiophilic (type 1) secretory cells; Figure 6D strongly osmiophilic (type 2) secretory cells). Histological evidence show that distal ducts of type 2 acini approach and merge into the distal ducts originating from type 1 acini (Figure 6D, E).

CUB-serine protease immunoreactivity

Treatment of the antennules with the antibody against the CUB domain of *Panulirus argus* serine protease (Csp) resulted in labeling of the glandular tissue in *Coenobita clypeatus* (Figure 7). CUB-immunoreactivity was observed within the secretory cells in the form of asymmetrical aggregations measuring 0.5–2 µm in diameter (Figure 7A, arrowheads). Not all secretory cells within a single acinus were CUB-immunoreactive, and it was not possible to differentiate between the two types of secretory cells using counterstaining with nuclear marker sytox green. Most of the CUB-immunoreactivity was concentrated close to the nucleus (Figure 7Aa). Notably, the proximal ducts lack CUB-immunoreactivity (Figure 7A), while at least parts of the distal ducts were labeled distinctly (Figure 7B, double arrowhead). In addition to the glandular tissue, CUB-immunoreactivity was detected at the level of the cilia formation, where the sheath cells surrounding the dendrites of the olfactory sensory neurons form an enlarged receptor lymph space around the cilia (Figure 7C, Ca; morphology described in e.g., [8,29]). No immunoreactivity was detected in the auxiliary cells (sheath cells surrounding the inner dendritic segments of the olfactory sensory neurons) of *Coenobita*, or any other antennular tissue. Negative controls, treated with 0.5% NGS PBS-Tx only, did not show any specific fluorescent signal.

To test whether or not CUB-serine protease is expressed in the antennules of *C. clypeatus* BLAST searches against the translated nucleotides of the *C. clypeatus* antennal transcriptome were performed [26]. We identified four contigs similar to the CUB-domain and several contigs matching the trypsin-domain of the reference sequence of *Panulirus argus*. However, the contigs were

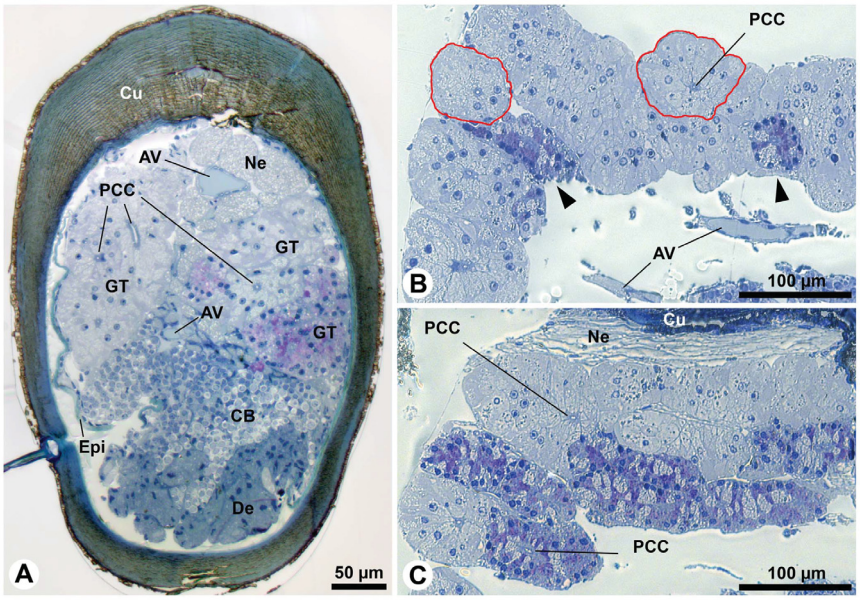


Figure 4. Aesthetasc-associated epidermal glands in *Coenobita clypeatus* stained with methylene blue and azur II (after Richardson et al., 1960). A–C: LM. **A**: Transverse section of large flagellum proximal to the aesthetasc pad. The epidermis has locally detached from the cuticle (in left half of the flagellum) due to a fixation artefact. The glandular tissue is located in the dorsal half of the flagellum. The ventral half is occupied by sensory neurons and associated sheath cells. **B**: Parasagittal section of glandular tissue. Note areas showing rosettes of darker stained secretory cells (arrowheads) and the nearly circular cross-sections of acini (red lines). **C**: Parasagittal section of acini. The population of darker stained secretory cells is more prominent. AV, arterial vessel; CB, sensory cell bodies; Cu, cuticle; De, dendrites of the olfactory sensory neurons; Epi, epidermis; GT, glandular tissue; Ne, antennal nerve; PCC, proximal part of the conducting canal.
doi:10.1371/journal.pone.0096430.g004

not spanning the entire sequence of the CUB domain serine protease. As several trypsin-like serine proteases are to be expected in the tissue we focused only on those ones matching the CUB domain. Figure 8A displays a multiple sequence alignment of the 4 identified putative CUB domain containing contigs. The amino acid similarity to the reference Csp of *P. argus* varied between

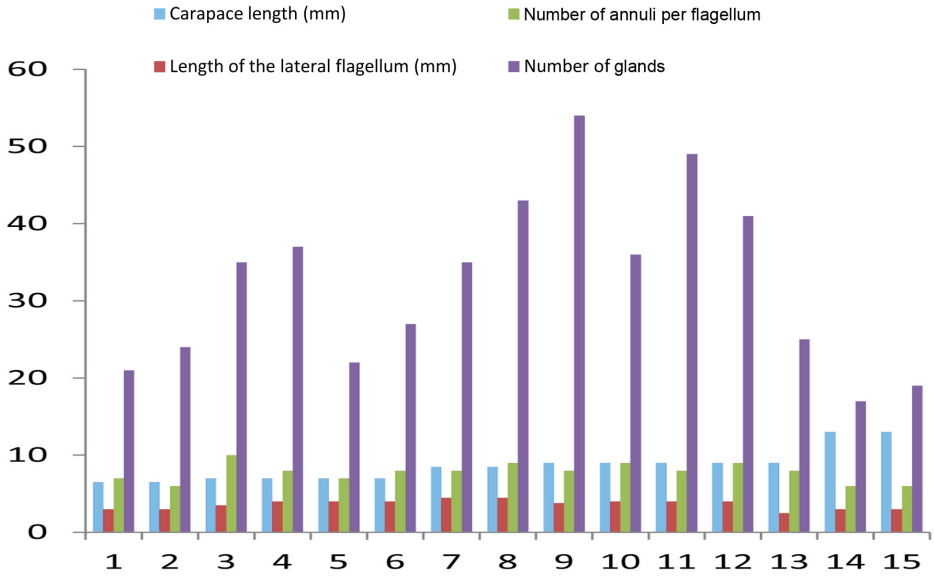


Figure 5. Comparative morphology and morphometry of the antennular-flagellar components in *Coenobita clypeatus*. Graph compiles length of the carapace (in mm, blue column), length of the lateral flagellum (in mm, red column), number of the annuli per flagellum (green column), and number of the proximal glandular ducts labeled with phalloidin (magenta column), compared among 15 antennules.
doi:10.1371/journal.pone.0096430.g005

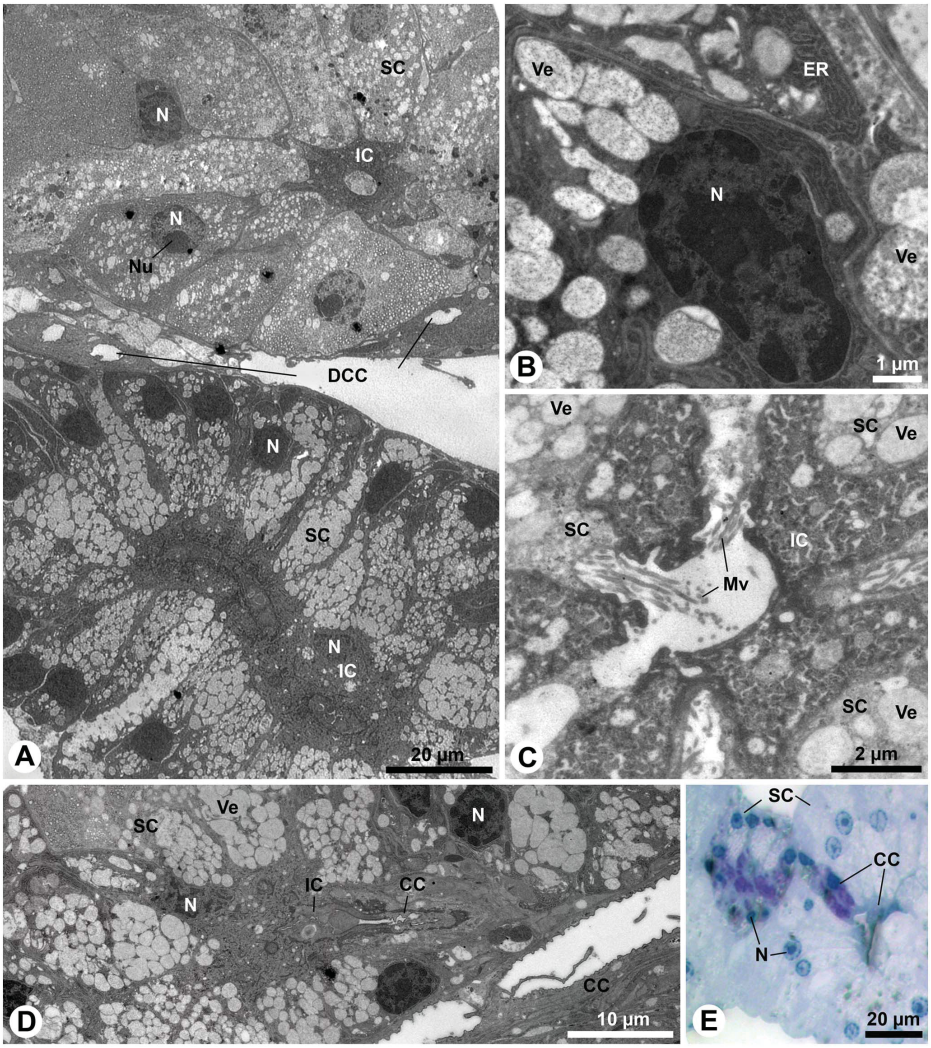


Figure 6. Acini in *Coenobita clypeatus*: Ultrastructure (A–D) and histological anatomy (E). A–D: TEM, **E**: LM. **A**: Oblique section showing acini of moderate (above) or strong (below) osmiophily. In both types, the secretory cells are circularly arranged around the central intermediary cell. **B**: Detail of a more osmiophilic secretory cell. Note densely aligned cisternae of the rough endoplasmatic reticulum. **C**: Transverse section of a proximal duct showing cytoplasmic details of the intermediary cell as well as apices of five surrounding, weaker electron-dense secretory cells in longitudinal section piercing the intermediary cell. Three of the secretory cells open into the proximal duct; slightly invaginated apices of the secretory cells form numerous microvilli, the latter project into the proximal duct. **D**: Oblique section showing cuticle-lined distal part of the conducting canal of a strongly osmiophilic acinus (narrow because cut tangentially), approaching another cuticle-lined distal duct. **E**: Detail showing a duct of a darker stained acinus merging into a duct of a lighter staining acinus. CC, canal cells; DCC, distal part of the conducting canal; ER, endoplasmatic reticulum; IC, intermediary cell; Mv, microvilli; N, nucleus; Nu, nucleolus; SC, secretory cell; Ve, vesicle.
doi:10.1371/journal.pone.0096430.g006

16.5% and 31.8% and was highest in the putative *C. clypeatus* CUB4. The similarity of putative *C. clypeatus* CUBs varied between 23.9% and 77.5% (see similarity matrix, Figure 8).

Protein identification

To perform proteomic analysis the soluble proteins of glandular complexes-containing tissue and antennular tissue containing olfactory sensory neurons were subjected for LC-MS/MS analysis. In order to identify protein candidates characterizing the gland secretions we compared the proteome of glandular complexes-containing tissue with those of the antennular tissue containing olfactory sensory neurons. As the genome of *C. clypeatus* is not available we used an in-house constructed transcriptome of *C. clypeatus* antennules and processed the acquired tandem mass spectra using combined proteomic strategy. The first step was to search them against available protein databases to identify proteins from the *C. clypeatus* protein subdatabase or to match peptides from highly conserved protein domains of closely related species (stringent database searching) followed by homology-based protein identification that relies on *de novo* sequencing of the acquired MS/MS spectra and searching them against available databases using mass spectrometry-driven BLAST [30] to identify proteins by homology (error-tolerant searching). To increase sequence coverage of the analyzed proteins we additionally applied data-

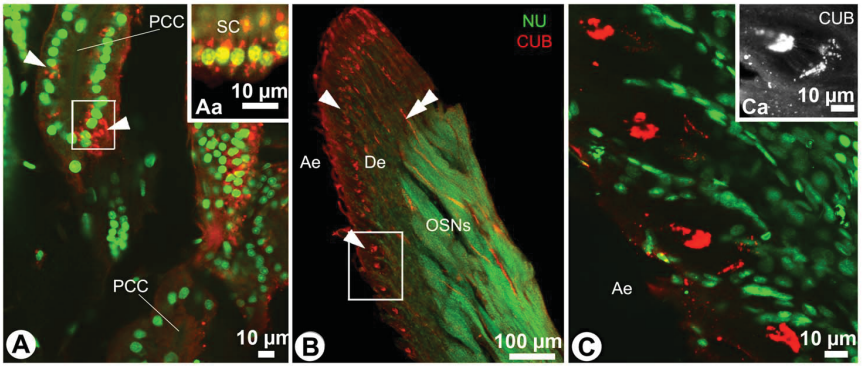
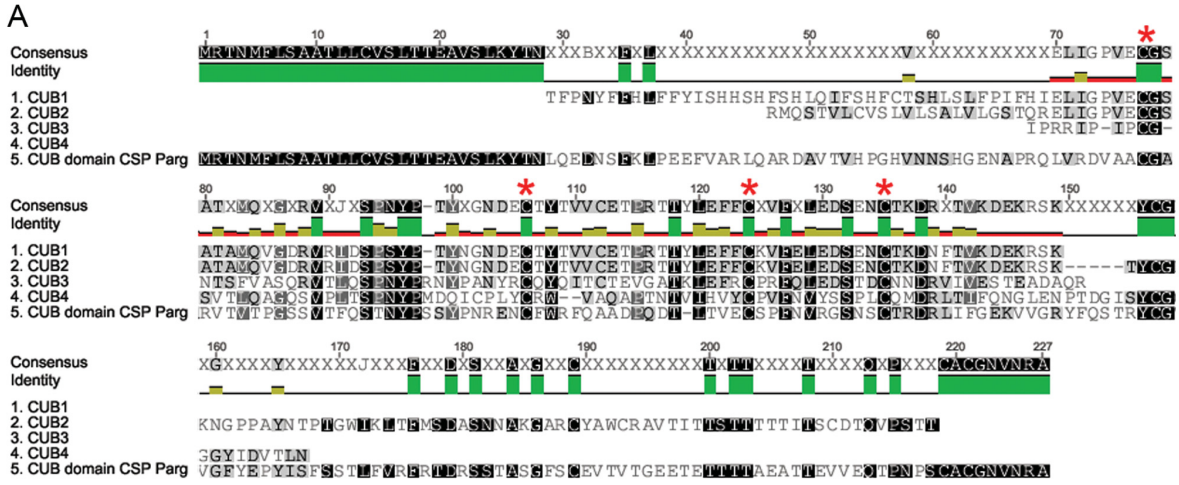


Figure 7. CUB-immunoreactivity in the antennules of *Coenobita clypeatus*. Flagellar tissues labeled with antibodies against the CUB domain of serine protease identified in the spiny lobster *Panulirus argus* (CUB, red) and stained with the nuclear marker sytox green (NU, green). cLSM. **A**: CUB-immunoreactivity in the secretory cells (arrowheads), **Aa**: some secretory cells from A (white frame) at higher magnification. **B**: An overview of the whole flagellum. The arrowheads mark the enlargements formed by sheath cells surrounding the dendrites of olfactory sensory neurons, the double arrowhead points towards labeling in the distal part of the conducting canal. **C**: CUB-immunoreactivity in the enlargements formed by sheath cells surrounding the dendrites of olfactory sensory neurons, **Ca**: details of CUB-immunoreactivity visualized in single (red) channel. Ae, aesthetasc field; De, dendrites of olfactory sensory neurons (OSNs); PCC, proximal parts of the conducting canal.
doi:10.1371/journal.pone.0096430.g007

independent acquisition referred to as MSE [30] which complements conventional data-dependent acquisition providing independent means of data validation.

The complete lists of the proteins identified in glandular complexes-containing tissue (GT) and antennular tissue containing olfactory sensory neurons (OSN) are depicted in (Tables S1, S2). Major bands for both samples ranked between 10 and 66 kDa, with the first strong band appearing at ca. 66 kDa and corresponding to ubiquitous proteins such as hemocyanin and heat shock proteins (Figure 9). All these hits along with other high abundant proteins such as beta-actin, alpha and beta tubulin, histone complex subunits as well as related contigs from the *C.*

A



B

	CUB1	CUB2	CUB3	CUB4	CUB domain CSP Parg
CUB1		77.5	32.5	25.7	16.5
CUB2	77.5		32.9	23.9	20.3
CUB3	32.5	32.9		25.0	24.7
CUB4	25.7	23.9	25.0		31.8
CUB domain CSP Parg	16.5	20.3	24.7	31.8	

Figure 8. CUB domain in *Coenobita clypeatus* ESTs. A: MUSCLE multiple sequence alignment of four putative CUB domain containing ESTs after translation into amino acids. Reference sequence: CUB domain of the Csp of *Panulirus argus* (Schmidt et al., 2006). Conserved cysteine residues are indicated by red asterisks. **B**: Similarity matrix of sequences, values are given in percent.
doi:10.1371/journal.pone.0096430.g008

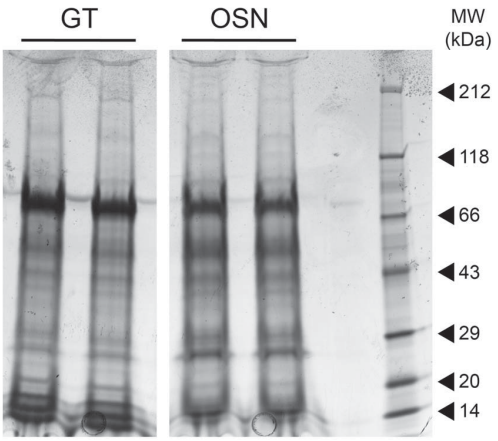


Figure 9. 1-D SDS-PAGE profiles of the proteomes of two samples tested. Protein bands were visualized with Coomassie Blue (R250). GT, glandular tissue; OSN, olfactory sensory neurons containing tissue.
doi:10.1371/journal.pone.0096430.g009

clypeatus antennal transcriptome [26] were mostly matched by all applied protein identification workflows. The major differences in the protein pattern were observed in the mass range from 14 to 20 kDa. Table 1 summarizes the annotated proteins only present in the glandular tissue (full list of proteins only present in glandular tissue in (Table S3). Most of these proteins indicate involvement in the secretory pathway, i.e. taking part in intracellular transport between the endoplasmatic reticulum (ER) and the Golgi apparatus, within the Golgi cisternae and the specific addressing and processing of vesicles. A second group is putatively involved in immune responses, including serine and aspartate proteases (coagulation factor XI, cathepsin D, hydrolase), protease inhibitors (alpha 2 macroglobulin), lectins, enzymes involved in responses to reactive oxygen species (such as glutathione-S-transferases) and an enzyme activated by interferon gamma (gamma-interferon inducible lysosomal thiol reductase). One identified protein is homologous to an uncharacterized protein produced in the salivary glands of *Ixodes scapularis* (Say, 1821). No CUB-serine proteases or CUB domain alone were identified in the proteome, neither in the glandular tissue, nor in the control sample.

Discussion

Characterization of aesthetasc-associated epidermal glands

Epidermal (tegumental) glands of crustaceans exhibit a wide range of structural complexities and are widespread in the crustacean integument. However, literature does only cover a fraction of the morphological diversity of epidermal glands in this group. Some crustacean taxa received attention with respect to their epidermal glands (e.g., Notostraca: [31]; Decapoda: [13,21]; Peracarida: [32,33]), while others remained rather disregarded, even though considered key groups in the debate on arthropod and crustacean phylogeny (e.g., Remipedia). Moreover, previous accounts addressed epidermal glands of certain body regions, in particular the head and appendages. For instance, previous studies gave insights into the anatomy of intra- and subepidermal glands of the head and mouth parts in land isopods [34], as well as pleopods [35], eyestalk [36], foregut, hindgut and the gill chamber [37] and antennae [13,38] in aquatic decapods. The epidermal glands are ubiquitous in some species, whereas in others they are often sparse and restricted to certain locations [39,20]. Epidermal glands in isopods have been shown to be especially abundant and structurally variable in terrestrial species compared to aquatic ones [37], and it has been suggested that ubiquity of epidermal glands might be connected to terrestrial adaptation in these animals [34,40,41].

Several reports mention pore structures potentially associated with different types of antennal sensilla of decapod crustaceans (summarized by [13]), but few studies linked external morphology of pore structures to electron microscopic data on cellular apparatus associated with these kinds of pores. Examples for these rare comprehensive approaches are the investigations on unicellular epidermal glands associated with guard setae in *Homarus gammarus* (Linnaeus, 1758) (see [42]),"rosette-type tegumental glands" associated with olfactory sensilla (aesthetascs) in *Panulirus argus* [13] and tegumental glands in the olfactory organ of *Homarus americanus* [38]. Even less information is available on the composition of the secretory product and the possible function of these glands in antennules.

Aesthetascs of the spiny lobster *P. argus* are numerous, long and slender setae [29], accompanied by other, non-olfactory sensilla: guard, companion and asymmetric setae [43]. Aesthetascs in *Coenobita* appear to have undergone specific adaptations to the terrestrial habitat: they are short and blunt [8;44], with slender and minute non-olfactory setae occurring mainly on the margins of the aesthetasc pad. The pores surrounded by characteristic collar-like folds, associated with the aesthetascs of *Coenobita*, however, resemble the "peg pores" of the aesthetasc tegumental glands (ATGs) identified in the spiny lobster *P. argus* [13]. In contrast, no structural analog to the more sparsely arranged minute "depression pores" [13] could be identified in the aesthetasc pad of *Coenobita*. Neither could unicellular glands like those known for *Homarus* (e.g. [42]) be observed in the area of the *Coenobita* aesthetasc pad. However, solitary minute pores which may resemble the "depression pores" [13] could be observed outside the aesthetasc pad in *C. clypeatus*. ATGs in *P. argus* lay directly beneath the cuticle of the large flagellum [13], while the aesthetasc-associated epidermal glands of *Coenobita* spp. are sunk much deeper into the inner regions of the flagellum. This substantial shift of the glands results in a considerable elongation (several hundreds of micrometers) of the distal glandular ducts. Organs being this much distanced from the cuticular surface and the subjacent epidermis, may have been an important prerequisite to live on land, as deeply sunk organs are prevented from desiccation and/or overheating (compare [45]). The aesthetasc-associated epidermal glands in *Coenobita* are bigger in diameter and longer (more flask-like rather than rosette-like), with more secretory cells per acinus. Bigger glands suggest that higher volume of secretion need to be produced in a short period of time, i.e. to impregnate the delicate cuticular shafts of the aesthetascs, and to react quickly to changes in air humidity.

The results of this study reveal that the general architecture of epidermal glands described for many different crustacean taxa (see review [20]) applies to hermit crabs (Paguroidea) as well. Epidermal glands of *Coenobita* spp. share typical basic set-up (class-III-glands acc. to [17], subclass of recto-canal epidermal glands acc. to [18]), comprising three cell types: secretory cells, which are organized in a rosette around the non-cuticularized proximal duct, intermediary cells (usually one intermediary cell per acinus), and canal cells, the latter forming the distal cuticularized part of the conducting canal. However, epidermal glands of *Coenobita* spp. also differ from those of other crustaceans by their higher complexity. The secretory cells comprise two types,

Table 1. Proteins exclusively present in the tissue containing aesthetasc-associated epidermal glands in *C. clypeatus*.

Protein	Mascot				MS BLAST				MSE				GO association
	Accession	Species	Score	Peptide hits	Accession	Species	Score	Peptide hits	Accession	Species	Score	Peptide hits	
alpha 2 macroglobulin (ConsensusfromContig31628_ full_rev_frame_1)						*Coenobita*	112	2		*Coenobita*	2516	3	GO:0019731: antibacterial humoral response
alpha 2-macroglobulin					ABD6146	*Scylla serrata*	278	6					
alpha spectrin	EFX88672	*Daphnia pulex*	296	6	XP_002083280	*Drosophila simulans*	617	13					
aspartate protease cathepsin D									AEO94539	*Triatoma infestans*	520	2	GO:0008233: peptidase activity
coagulation factor XI					XP_001845410	*Culex quinquefasciatus*	72	1					GO:0006508:proteolysis
coatomer protein complex					NP_001166195	*Bombyx mori*	71	1					GO:0016192:vesicle-mediated transport
coatomer protein complex (ConsensusfromContig810 full rev frame 0)										*Coenobita*	639	2	
coatomer protein delta, isoform B									NP_001162642	*Drosophila melanogaster*	723	5	
coatomer subunit delta-like									XP_003700942	*Megachile rotundata*	738	6	
C-type lectin (ConsensusfromContig 110279_full_fwd_frame_0)		*Coenobita*	82	1						*Coenobita*	1730	1	GO:0002682:regulation of immune system process
emp24 cargo transport protein (ConsensusfromContig 108988_full_fwd_frame_0)										*Coenobita*	307	2	GO:0051050:positive regulation of transport
gamma-interferon inducible lysosomal thiol reductase					XP_002400600	*Ixodes scapularis*	65	1					GO:0007166:cell surface receptor signaling pathway
GH15863 gene product from transcript GH15863-RA, Peptidase_C26					XP_001983370	*Drosophila grimshawi*	105	2					GO:0016787: hydrolase activity
GL17639 (ConsensusfromContig35747 full rev frame 1)										*Coenobita*	321	3	GO:0000042:protein targeting to Golgi
glutathione-S-transferase (ConsensusfromContig 523_full_rev_frame_0)						*Coenobita*	77	1					GO:0006979:response to oxidative stress
glutathione-S-transferase (ConsensusfromContig 80526_full_rev_frame_0)										*Coenobita*	426	3	
glutathione-S-transferase (ConsensusfromContig 83104_full_rev_frame_0)										*Coenobita*	765	1	

Table 1. Cont.

Protein	Mascot				MS BLAST				MSE				GO association
	Accession	Species	Score	Peptide hits	Accession	Species	Score	Peptide hits	Accession	Species	Score	Peptide hits	
GST_N_Metaxin									XP_002430330	*Pediculus humanus corporis*	489	4	GO:0070122: isopeptidase activity
S-formylglutathione hydrolase (ConsensusfromContig31027 full rev frame 1)										*Coenobita*	392	1	GO:0018738:S-formylglutathione hydrolase activity
signal peptidase complex subunit 3-like					XP_003739905	*Metaseiulus occidentalis*	65	1					GO:0009003: signal peptidase activity
sorting nexin-12									XP_002432737.1	*Pediculus humanus corporis*	841	3	GO:0015031:protein transport
trafficking protein particle complex subunit 10									XP_003742750	*Metaseiulus occidentalis*	671	2	GO:0016192: vesicle-mediated transport
transient receptor potential-gamma protein									EFN78599	*Harpegnathos saltator*	439	1	GO:0005261: cation channel activity
uncharacterized protein									XP_003748091	*Metaseiulus occidentalis*	531	2	homology "secreted salivary gland peptide" Isca XP_002414539

doi:10.1371/journal.pone.0096430.t001

which stain differently according to the protocol of Richardson et al. [23]. This effect, known as metachromasia, can be explained by different binding patterns of the dyes to substances in the tissue characterized by different chemical properties. In case of epidermal glands the reasons for metachromasia are the chemistry of the secretory material, as well as ultrastructural differences, such as variations in the density of rough endoplasmic reticulum. The fusion of the distal ducts from the acini with these two different secretory cell types indicates that their secretory products get mixed before being released onto the surface of the aesthetasc pad. Notably, the proximal ducts in *Coenobita* spp. being narrow and often branched are quite different from the proximal ducts of ATGs of their aquatic relative *P. argus* [13]. In the latter species, the proximal ducts resemble the local enlargements of the conducting canal, i.e. reservoirs, and never branch.

Investigation of proteomes in non-model organisms using similarity-based searching (MS BLAST) has been successfully used in many studies [29]. However, it still remains a challenging task characterizing species with significant phylogenetic diversity and the success rate of identification dramatically decreases with evolutionary distance to its closest relative represented in the databases [30]. While data from insecta clade within Pancrustacea [46], are well represented in the available databases, molecular data from crustaceans is sparse and lead to a high number of proteins found in the antennal transcriptome and antennal proteome but lacking matches to described homologues in other species. The uncharacterized transcripts listed in Table S3 might play important roles in the function of the epidermal glands but need further characterization. The peptides identified in the epidermal glands of *Coenobita* indicate a classic merocrine pathway from precursors produced in cisternae of rough ER, transportation to and processing in Golgi bodies, vesicular budding, collection in fusion granules, and final extrusion of secretion at the apex of the secretory cell (see Table1). The abundance of rough ER cisternae in epidermal glands suggests that glycoproteins might be a major secretory product [20]. The proximal ducts of the epidermal glands can be visualized with phalloidin, and thus the intermediary cell forming the lumen and/or secretory cells contain high concentrations of F-actin filaments. A similar phalloidin-positive staining was described in pleopod tegumental glands of the American lobster *Homarus americanus* (H. Milne Edwards, 1873) [21] and ATGs of *P. argus* [13]. The absence of phalloidin staining after incubation of the tissues with nuclear stain sytox green is probably due to an interaction with dimethyl sulfoxide. There are cases known when the intensity of phalloidin labeling is also affected by heat [13]. Rich networks of actin filaments can provide mechanical support for the proximal duct, help to maintain its shape and prevent the lumen from collapsing. It is also possible that filamentous actin together with myosin can function in order to push secretory products into the distal part of the conducting canal of the gland; we however did not see any evidence for myosin filaments or microtubules inside the intermediary cells forming the proximal duct, so it might be sufficient if the secretory products simply flow through the system. The distal parts of the conducting canal cannot be visualized with phalloidin. This makes us suggest that some additional mechanisms might be involved in guiding the secretion through the distal duct system and facilitating the final release of the glandular products onto the surface. It is known that secretion is released during antennular grooming when the first antennae are clasped and smeared over repeatedly by maxillipedes in the spiny lobster [47]. In some arthropods, muscles have been found in association with the epidermal glands, however not in the antennae [48]. Anyway, pulsating hemolymphs, propelled by the circulatory system, might also influence transport and release of secretions due to compressing the acini and the ducts. Since the secretory cells of the aesthetasc-associated epidermal glands of *Coenobita* neither possess any storage compartments or reservoirs larger than regular secretory granules, nor any kind of valve-like structure regulating the secretion flow, it is likely that as long as the glands deliver their product into the ducts, a constant flow of highly fluid secretion will reach the surface.

The mechanism that regulates and activates secretion in epidermal glands is poorly understood. It is assumed that in general crustacean epidermal glands are not innervated [20], although some exceptions have been reported, like the rosette-type glands in the gills of the grass shrimp *Palaemonetes pugio* (Holthuis, 1949), where synapses have been observed on the secretory cells [49]. Treatment of *Coenobita* antennules with synapsin antibodies did not reveal any positive immunoreactivity (Tuchina et al., unpublished data), suggesting that antennal epidermal glands in *Coenobita* are not innervated and the secretion is triggered and regulated by some other mechanism. Many epidermal glands have been shown to synchronize their secretion activity with the molting cycle [50,51], which means they are probably hormonally regulated. However, no such hormones have been identified to be produced in epidermal glands of coenobitids so far.

Functions of aesthetasc-associated epidermal glands and serine proteases

Although epidermal glands can be structurally similar at different body regions such as mouth parts and antennae, their presumed function, i.e. histochemistry of the secretory product, is likely different. Epidermal glands in the legs of the terrestrial isopod *Armadillium vulgare* (Latreille, 1804) produce polyphenol oxidase, an enzyme that has been implicated in the tanning of the cuticle after molting [52]. Epidermal glands from the gut of the lobster *Homarus gammarus* (Linnaeus, 1758) appear to contain phosphatases, ATPases and mucopolysaccharides [53], glands found beneath the cuticle of the head and mouth parts are often referred to as salivary glands or thought to be involved in digestion [20], while in glands of the amphipod *Gammarus pulex* (Linnaeus, 1758) catecholamines, like for example dopamine, seem to be the major secretory product [54].

The antennal epidermal glands of the spiny lobster *Panulirus argus* have been shown to contain specific type of proteases, CUB-serine proteases (Csp) [16,55]. In the following studies by Schmidt et al. [13] it was confirmed that Csps are found exclusively in the secretory cells of aesthetasc tegumental glands, ATGs, and are likely secreted onto the surface of the aesthetascs. A serine protease with low sequence similarity to Csp but with similar distribution pattern as in *P. argus* was also found in the antennae of *H. americanus* [56] with one olfactory enriched transcript (OET-03), identified as chymotrypsin-like serine protease and exclusively expressed in the secretory cells of aesthetasc-associated tegumental glands [38]. By homology searches we identified four putative CUB domains in the antennal transcriptome of *C. clypeatus*. The ESTs were short and did not allow a final characterization; however we find the four cysteine residues at their fixed positions as reported by Levine et al. [16] for *P. argus*. The reason why we did not find Csp or CUB domain alone in the antennal gland proteome might be because the proteins were not separated perfectly, and it is especially likely to be the case for membrane-associated proteins.

Serine proteases are known to have manifold functions. Thus, their role in the crustacean olfactory system is uncertain and a matter of ongoing discussion. It has been shown that serine proteases and their inhibitors are involved in the development of

the nervous system, including cellular migration, differentiation, process extension, as well as repairing and apoptosis of neurons and glial cells in vertebrates [57,58]. The pattern of activity of Csps in the antennae of the spiny lobster varies along the developmental axis of the olfactory organ, suggesting a possible role of Csps in development, maturation and/or degradation of olfactory sensory neurons [55], although the Csp-immunoreactivity was detected exclusively in glandular tissue [13]. In *Coenobita*, we identified CUB-immunoreactivity in the enlargements formed by the sheath cells surrounding the dendrites of the olfactory sensory neurons (not to be mixed with the report by Levine et al. [16] about Csp-immunoreactivity in auxillary cells in *P. argus*). The enlargements surrounding olfactory cilia in *Coenobita* spp. have been previously reported by Ghiradella et al., [59] and assumed to help in maintaining the proper osmotic conditions around cilia and also probably serve as mechanical shock absorbers. As reported by Schmidt et al., [13] for *P. argus*, the distribution of immunoreactive material was seemingly identical for two antisera used in their study: no. 99-6 raised against the CUB domain and 102-6 raised against the protease domain. In this study, only 5th bleed serum (99-5) against the CUB domain of the Csp was tested. It can be extrapolated that distribution of the Csp within the glandular cells most likely also corresponds to the detected CUB-immunoreactivity. The function of Csp in association with the dendrites of olfactory sensory neurons is unclear, but the possibility that serine proteases might be involved in development, apoptosis or maintenance of OSN dendritic function in *Coenobita* cannot be excluded and needs further investigation.

Other functions of serine proteases include perireception, regulation of other proteins and hormones, blood clotting, and immune response [60,61]. The serine proteases from ATGs of the spiny lobster *P. argus* have been suggested to be involved in enzymatic degradation of odorant molecules and/or production of active odorants from inactive precursors [60,62]. Peptide-mediated behaviors are well documented in aquatic animals [60]. Short peptides, free amino acids and their binary mixtures are known to trigger specific behaviors in crustaceans, such as attraction of anomuran crabs to the new shells [63,64], and simultaneous release of larvae and induction of larval settlement behavior in brachyuran crabs [65]. Non-volatile peptide cues from a dead snail have been shown to attract marine hermit crabs within minutes [66]. Females of the estuarine mud crab *Rhithropanopeus harrisii* (Gould, 1841) perform rhythmic and highly synchronized patterns of larval release ('abdominal pumping') in a response to carboxyl terminal arginine peptides ('pumping factor'), the latter are generated by the action of trypsine-like serine proteases on the membranes of the hatching eggs [67]. A pumping response can also be evoked, to a different degree, by treatment with exogenous trypsin, trypsin inhibitors [67] or a mixture of arginine and glycine, the two amino acids that are most prevalent after hydrolysis of the pumping factor [68]. As for hermit crabs, several amino acids, including arginine, glycine and leucine, were screened in *Pagurus bernhardus* (Linnaeus, 1758) and *C. clypeatus*, but did not elicit any response in electro-antennographic recordings [6], and thus were not tested behaviorally.

Aesthetasc-associated epidermal glands in *Coenobita* might play a role in chemosensing, the main function of crustacean antennules, however, no odorant binding proteins were found in proteomic analysis, or in the antennal transcriptome of *C. clypeatus* [26]. We also observed that epidermal glands in *Coenobita* are present in similar numbers in both sexes so we consider it unlikely that the secretions have a role in sexual communication. Schmidt and coauthors [13] proposed that epidermal glands secretion in *P. argus* have anti-fouling and/or friction-reducing properties, and would thus be important for the proper functioning of the aesthetascs. This hypothesis is corroborated by our data for *Coenobita* spp. Crustaceans, like other invertebrates do not possess acquired immunity and have to rely solely on the innate immune system as a defense against invading microorganisms. Aesthetascs with their thin cuticle are likely to be easy targets for microbial invaders, and in fact the antennular tissue in *C. clypeatus* was identified as a place of production of antimicrobial substances indicating the presence of a highly diverse microbial community [26]. Moreover, Krång et al. [6] demonstrated that high humidity is critical for odor perception in *C. clypeatus* and putatively also enhances microbial growth. The innate immune responses in crustaceans include melanization by activation of prophenoloxidases, a clotting process, phagocytosis, encapsulation and cell agglutination [61]. The components of the innate immune system, such as coagulation factor XI, cathepsin D and alpha 2 macroglobulin were identified in the proteome of the gland secretions, as well as diverse enzymes, involved in the responses to oxidative stress (catalases, glutathione reductases, glutathione-S-transferases, superoxide dismutases and peroxiredoxins). Opaque granules of melanin pigments were frequently observed in the antennular tissue of *C. clypeatus* (Tuchina et al., unpublished data) – an indication of a melanization process, a common reaction to pathogens among arthropods. The innate immune system is usually represented by its humoral (antimicrobial factors circulating in hemolymph) and cellular (hemocytes) components, so we find the need of a specialized structure, such as antennular epidermal glands producing a particular type of serine proteases in several crustacean species, is very remarkable.

We propose that the aesthetasc-associated epidermal glands in *Coenobita* spp. produce a mucous-like substance covering the aesthetascs and providing both moist interface essential for odor perception as well as mechanisms for antimicrobial defense, such as CUB-serine proteases, and probably other components of innate immunity. Together with the morphological characteristics of the glands, the proposed functions of the mucous are important adaptations for the functionality of the aesthetasc sensory system within terrestrial environments.

Conclusion

Olfactory sensilla in *Coenobita* spp. are covered by mucous-like substance, which is produced by aesthetasc-associated epidermal glands of high morphological complexity. The mucous takes part in antimicrobial defense and at the same time provides moisture which is known to be critical for odor perception in terrestrial hermit crabs. The morphological modifications of the aesthetasc-associated epidermal glands as well as the functional characteristics of their secretions are important adaptations to a terrestrial lifestyle.

Supporting Information

Table S1 The complete list of proteins identified in the glandular complexes-containing tissue, GT.
(XLSX)

Table S2 The complete list of proteins identified in the antennular tissue containing olfactory sensory neurons, OSN (control sample).
(XLSX)

Table S3 The complete list of proteins present exclusively in the glandular tissue.
(XLSX)

Acknowledgments

For support and use of lab facilities and microscopes we thank PD Dr. Martin Westermann and Renate Kaiser (both Electron Microscopy Center, Jena), Prof. Dr. Hans-Jürgen Agricola (Centre for Molecular Biomedicine, Jena) as well as PD Dr. Michael Nickel and Dr. Jörg Hammel (both Zoological Institute, Jena). Sincere thanks are given to Prof. Dr. Manfred Schmidt and Prof. Dr. Charles Derby (Neuroscience Institute, Georgia State University, USA) for providing the CUB-serine protease antibodies, and to Róger Blanco Segura, the former director of the National Park in Santa Rosa, and Ing. Javier Guevara Sequeira (SINAC) for sampling and export permissions for material of *C. compressus*. Likewise, we thank Mustafa Foudas (EEAAg) for permissions concerning *C. scaevola*. We would like to thank Prof. Dr. Hans-Peter Saluz (Hans Knoell Institute, Jena) for advice and support in experimental procedures of protein sample preparation, as well as our reviewers for their helpful comments.

Author Contributions

Conceived and designed the experiments: OT GT KCG BSH EG AS. Performed the experiments: OT GT KCG NW YH. Analyzed the data: OT GT KCG NW YH. Contributed reagents/materials/analysis tools: CHGM NW YH. Wrote the paper: OT GT CHGM. Obtaining permissions for *C. compressus* and *C. scaevola*: CHGM. Sample preparation for TEM, SEM: OT GT KCG CHGM. Ultramicrotomy and transmission electron microscopy: GT. Scanning electron microscopy: GT OT KCG. Laser scanning microscopy and histochemistry: OT. Proteomics experiments: NW YH. Proteomics data analysis: NW OT KCG. Critical revision of the manuscript: BSH EG AS.

References

1. Glaessner MF (1969) Decapoda. In: Lawrence KS, editor. Treatise Invertebrate Palaeontology Kansas: Geological Society of America and Kansas University Press. pp. 400–533.
2. Bliss DE, Mantel LH (1968) Adaptations of crustaceans to land - a summary and analysis of new findings. Am Zool 8: 673–685.
3. McLaughlin P, Lemaitre R, Sorhannus U (2007) Hermit crab phylogeny: a reappraisal and its "fall-out." J Crustacean Biol 27: 97–115.
4. Rittschof D, Sutherland JP (1986) Field studies on chemically mediated behavior in land hermit crabs *Coenobita rugosus* volatile and nonvolatile odors. J Chem Ecol 12: 1273–1284.
5. Stensmyr MC, Erland S, Hallberg E, Wallen R, Greenaway P, et al. (2005) Insect-like olfactory adaptations in the terrestrial giant robber crab. Curr Biol 15: 116–121.
6. Krång A-S, Knaden M, Steck K, Hansson BS (2012) Transition from sea to land: olfactory function and constraints in the terrestrial hermit crab *Coenobita clypeatus*. Proc R Soc Lond B Biol Sci 279(1742): 3510–3519.
7. Harzsch S, Hansson BS (2008) Brain architecture in the terrestrial hermit crab *Coenobita clypeatus* (Anomura, Coenobitidae), a crustacean with a good aerial sense of smell. BMC Neurosci 9: 58.
8. Ghiradella H, Case J, Cronshaw J (1968) Fine structure of aesthetasc hairs of *Coenobita compressus* Edwards. J Morphol 124: 361–385.
9. Koczan S (2012) 3D Reconstruction of the olfactory pathway from the aesthetasc to the olfactory lobe in the hermit crab, *Coenobita clypeatus*. Friedrich-Schiller-Universität Jena, Biologisch-Pharmazeutische Fakultät, Diploma Thesis.
10. Steinbrecht RA, Stankiewicz BA (1999) Molecular composition of the wall of insect olfactory sensilla - the chitin question. J Insect Physiol 45: 785–790.
11. Ache BW, Young JM (2005) Olfaction: diverse species, conserved principles. Neuron 48: 417–30.
12. Müller C, Sombke A, Hilken G, J R (2011) Sense organs. In: Minelli A, editor. Treatise on Zoology - Anatomy, Taxonomy, Biology. The Myriapoda. Leiden: Brill. pp. 235–278.
13. Schmidt M, Chien H, Tadesse T, Johns ME, Derby CD (2006) Rosette-type tegumental glands associated with aesthetasc sensilla in the olfactory organ of the Caribbean spiny lobster, *Panulirus argus*. Cell Tissue Res 325: 369–395.
14. Barrett AJ (1994) Classification of peptidases. Volume 244, methods in enzymology: proteolytic enzymes: serine and cysteine peptidases. California, USA: Academic Press. pp. 1–15.
15. Bork P, Beckmann G (1993) The CUB domain. A widespread module in developmentally regulated proteins. J Mol Biol 231: 539–545.
16. Levine MZ, Harrison PJ, Walthall WW, Tai PC, Derby CD (2001) A CUB-serine protease in the olfactory organ of the spiny lobster *Panulirus argus*. J Neurobiol 49: 277–302.
17. Quennedey A (1998) Insect epidermal gland cells: ultrastructure and morphogenesis. In: Harrison F, Locke M, editors. Microscopic Anatomy of Invertebrates. New York: Wiley-Liss pp.177–207.
18. Müller CHG, Rosenberg J, Hilken G (2013) Ultrastructure, functional morphology and evolution of recto-canal epidermal glands in Myriapoda. Arthropod Struct Dev, in press.
19. Müller C, Rosenberg J, Hilken G (2009) Fine structure and phylogenetic significance of "flexo-canal epidermal glands" in Chilopoda. In: Xylander W, Voigtländer K, editors. Proc. 14th International Congress of Myriapodology and Soil Organisms: pp. 269–294.
20. Talbot P, Demers D (1993) Tegumental glands of crustacea. In: Horst MN, Freeman JA, editors. Crustacean Integument. Morphol. Biochem Boca Raton, Florida: CRC Press. pp. 151–191.
21. Talbot P, Al-Hajj H, Demers D, Howard D (1991) Distribution of microfilaments and microtubules in *Homarus* pleopod tegumental glands (Crustacea: Decapoda). Zoomorphologie 110: 329–338.
22. Karnovsky MJ (1965) A formaldehyde-glutaraldehyde fixative of high osmolarity for use in electron microscopy. J Cell Biol 27: 137A.
23. Richardson KC, Jarett L, Finke EH (1960) Embedding in epoxy resins for ultrathin sectioning. Stain Technol. 35: 313–323.
24. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32 (5): 1792–1797.
25. Shevchenko A, Tomas H, Havlis J, Olsen J V, Mann M (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat Protoc 1: 2856–60.
26. Groh K, Vogel H, Stensmyr M, Grosse-Wilde E, Hansson B. The hermit crab's nose – antennal transcriptomics. Front Evol Psychol Neurosci, accepted.
27. Shevchenko A, Sunyaev S, Loboda A, Shevchenko A, Bork P, et al. (2001) Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. Anal Chem 73: 1917–26.
28. Li G-Z, Vissers JPC, Silva JC, Golick D, Gorenstein M V, et al. (2009) Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures. Proteomics 9: 1696–1719.
29. Gruenert U, Ache BW (1988) Ultrastructure of the aesthetasc (olfactory) sensilla of the spiny lobster *Panulirus argus*. Cell Tissue Res 251: 95–103.
30. Habermann B, Oegema J, Sunyaev S, Shevchenko A (2004) The power and the limitations of cross-species protein identification by mass spectrometry-driven sequence similarity searches. Mol Cell Proteomics 3: 238–249.
31. Rieder N (1977) Ultrastruktur und Funktion der Hautdrtisen von *Triops cancriformis* Bosc. (Crustacea, Notostraca). Zoomorphologie 88: 133–143.
32. Weirich D, Ziegler A (1997) Uropod and lateral plate glands of the terrestrial isopod *Porcellio scaber* Latr. (Oniscidae, Crustacea): An ultrastructural study J Morphol 233(2): 183–193.
33. Juberthie-Jupeau L, Crouau Y (1977/78) The tegumental glands of a troglobitic crustacean. Int J Speleol 9: 309–319.
34. Gorvett H (1956) Tegumental glands and the terrestrial life in wood lice. Proc Zool Soc London 126: 291–314.
35. Johnson B, Talbot P (1987) Ultrastructural analysis of the pleopod tegumental glands in the male and female lobsters, *Homarus americanus*. J Crustac Biol 7: 288–301.
36. Arseault A, Clatenberg R, Aiken D (1979) The morphology and secretory transporting mechanism of the tegumental glands of the lobster (*Homarus americanus*) as related to the molting cycle. J Submicros Cytol 11: 193–207.
37. Yonge C (1932) On the nature and permeability of chitin. I. The chitin lining the foregut of decapod Crustacea and the function of the tegumental glands. Proc R Soc Lond B Biol Sci B 111: 298–329.
38. Stepanyan R, Haley SB, McClintock TS (2005) Olfactory specific chymotrypsin-like serine protease from the aesthetasc tegumental gland of the lobster, *Homarus americanus*. Cell Tissue Res 322: 321–330.
39. Brunet M, Cuoc C, Arnaud J, Mazza J (1991) Tegumental glands in a copepod *Hemidiaptomus ingens*: structural, ultrastructural and cytochemical aspects. Tissue Cell 23: 733–743.
40. Hornung E (2011) Evolutionary adaptation of oniscidean isopods to terrestrial life: Structure, physiology and behavior. Terr Arthropod Rev 4: 95–130.
41. Gorvett H (1951) The tegumental glands in the land Isopoda. B. The lobed glands: structure and distribution. QJ Micros Sci 92: 275–296.
42. Gnatzy W (1984) "Campaniform" structures on lobster antennae are dermal glands. Cell Tissue Res 236: 729–731.
43. Schmidt M, Derby CD (2005) Non-olfactory chemoreceptors in asymmetric setae activate antennular grooming behavior in the Caribbean spiny lobster *Panulirus argus*. J Exp Biol 208: 233–248.
44. Ghiradella H, Cronshaw J, Case J (1970) Surface of the cuticle on the aesthetascs of *Cancer*. *Protoplasma* 69: 145–150.
45. Greenaway P, Morris S, McMahon BR (1988) Adaptations to a terrestrial existence by the robber crab *Birgus latro*: II. *In vivo* respiratory gas exchange and transport. J Exp Biol 140: 493–509.
46. Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, et al. (2010) Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. Nature 463(7284):1079–83.

47. Barbato JC, Daniel PC, Barbato JC, Daniel PC (1997) Chemosensory activation of an antennular grooming behavior in the spiny lobster *Panulirus argus* is tuned narrowly to L-glutamate. Biol Bull 193: 107–115.
48. Arnaud J, Brunet M, Mazza J (1988) Labral glands in *Centropages typicus* (Copepoda, Calanoida). 1. Sites of synthesis. J Morphol 197 (1): 21–32.
49. Doughtie D, Rao K (1982) Rosette glands in the gills of the grass shrimp *Palaemonetes pugio*. J Morphol 171: 41–67
50. Gorvett H (1946) The tegumental glands in the land Isopoda. A. The rosette glands. QJ Micros Sci 87: 209–234.
51. Tomaschko K (1995) Autoradiographic and morphological investigations of the defensive glands in adult *Pycnogonum litorale* (Arthropoda, Pantopoda). Eur J Entomol 92: 105–112.
52. Stevenson J (1961) Polyphenol oxidase in the tegumental glands in the relation to the moulting cycle of the isopod crustacean *Armadillium vulgare*. Biol Bull 121: 554.
53. Backer P, Gibson R (1977) Observations on the feeding mechanism structure of the gut and digestive physiology of the European lobster *Homarus gammarus* (L.) (Decapoda: Nephropidae). J Exp Mar Biol Ecol 26: 297–324.
54. Elofsson R, Aramant R, Lindvall O, Falck B (1978) Catecholaminergic salivary glands in *Gammarus pulex* (Crustacea, Amphipoda): an electron microscopic and microspectrofluorometric study general appearance and fluorescence. J Ultrastruct Res 64: 14–22.
55. Johns ME, Tai PC, Derby CD (2004) Serine proteases in the spiny lobster olfactory organ: their functional expression along a developmental axis, and the contribution of a CUB-serine protease. J Neurobiol 61: 377–391.
56. Hollins B, Hardin D, Gimelbrant AA, McClintock TS (2003) Olfactory-enriched transcripts are cell-specific markers in the lobster olfactory organ. J Comp Neurol 455: 125–138.
57. Hogan BL (1996) Bone morphogenetic proteins: multifunctional regulators of vertebrate development. Genes Dev 10: 1580–1594.
58. Festoff BW (1990) Serine proteases and their serpin inhibitors in the nervous system. 191. Pros NATO ASI Series A191. New York: Premium Press. 359 p.
59. Ghiradella HT, Case JF, Cronshaw J (1968b) Structure of aesthetascs in selected marine and terrestrial decapods: chemoreceptor morphology and environment. *Am Zool* 8: 603–621.
60. Rittschof D (1990) Peptide-mediated behaviors in marine organisms: evidence for a common theme. J Chem Ecol 16: 261–272.
61. Sritunyalucksana K, Soderhall K (2000) The proPO and clotting system in crustaceans. Aquaculture 191: 53–69.
62. Rittschof D (1993) Body odors and neutral-basic peptide mimics: a review of responses by marine organisms. Am Zool 33: 487–493.
63. Rittschof D (1980) Chemical attraction of hermit crabs and other attendants to simulated gastropod predation sites. J Chem Ecol 6: 103–118.
64. Small MP, Thacker RW (1994) Land hermit crabs use odors of dead conspecifics to locate shells. J Exp Mar Bio Ecol 182: 169–182.
65. Rittschof D, Barlow L, Schmidt AR (1985) Laboratory studies of olfaction and taste in land hermit crabs *Coenobita rugosus*. Am Zool 25: 113A.
66. Rittschof D (1980) Enzymatic production of small molecules attracting hermit crabs to simulated gastropod predation sites. J Chem Ecol 6: 665–675.
67. Rittschof D, Forward RB, Erickson BW (1990) Larval release in brachyuran crustaceans. Functional similarity of peptide pheromone receptor and catalytic site of trypsin. J Chem Ecol 16: 1359–1370.
68. Rittschof D, Forward RB, Mott DD (1985) Larval release in the crab *Rhithropanopeus harrisii* (Gould): chemical cues from hatching eggs. Chem Senses 10: 567–577.

7. General discussion

The present thesis extends on our knowledge of the molecular biology of the peripheral olfactory organs of crustaceans and is the first study to shed light on its adaptation during the transition from water to land.

Crustacean terrestrialization

Environmental changes lead to adaptations in organisms inhabiting those environments, as the naturally occurring genetic variation in a population allows selection for the best suitable genetic makeups to match particular changes. Random genetic variants may also allow animals to inhabit a new environment that cannot be conquered by their conspecifics. A combination of many changes is necessary to allow animals to succeed in a step as fundamental as the transition from water to land. One big challenge connected to successful performance on land is the avoidance of desiccation. While aquatic animals are constantly surrounded by aqueous solution, terrestrial hermit crabs take advantage of the mollusk shells their aquatic ancestors employed to protect their uncuticularized abdomen from predators. Small amounts of water are stored in the shell spiral to keep the abdomen moist (McMahon and Burggren 1979). Limbs requiring more exposed positions for functioning have been adapted morphologically, which applies for example to head appendages like the olfactory organs. The aquatic hermit crab *P. bernhardus* has short antennulae with long, hair like aesthetascs while the terrestrial hermit crab *C. clypeatus* has extended basal antennular segments but short and blunt aesthetascs, with the latter exposing far less of the thin cuticle to the outside (Koczan 2012). Regardless of morphological changes undergone in the transition from aquatic to terrestrial habitats, the general genetic makeup and receptors employed to detect chemicals in hermit crab antennules did not change to a great extent (manuscript 1 (Groh et al. 2014)). The high number of genes expressed in a given tissue makes it difficult to compare complex datasets like whole-organ transcriptomes. If exemplary genes are chosen to represent certain metabolic or signaling pathways, they might not be annotated correctly in all datasets, or might be found in both species while the pathways differ substantially or are partly substituted by other paths in one of the species. Additionally, the selection of distinct genes does not suffice to compare the totality of genes expressed as unexpected differences might be overlooked. The gene ontology terminology

allows us to apply a transferable vocabulary that can be categorized and compared between different datasets (Ashburner et al. 2000). The vocabulary includes a hierarchical structure, so a specific function or process is represented by a certain portion of the respective term on a given level. If a pathway is massively extended in one species, more ESTs[1] are contained in the respective term leading to a higher portion of this term in one dataset when compared to another. Concerning the antennal transcriptome data of *P. bernhardus* and *C. clypeatus* all observed differences are rather small in all categories (manuscript 1 (Groh et al. 2014)). That means that in approximately 170 million years of independent development and besides the transition from water to land in the Coenobita lineage, the repertoire of expressed genes in the olfactory organs is largely unchanged. It has to be mentioned that the lack of knowledge about antennaly expressed genes and the molecular genetics of crustaceans in general did only allow annotation of 25% of the contigs from the antennal transcriptome of *C. clypeatus* and 33% of *P. bernhardus*, which is comparable to datasets from other studies of non-olfactory crustacean transcriptomes. This obvious gap in knowledge raises questions yet to be answered. What is the function of those genes with no homologue known in any species? Are they crustacean specific and present in other tissues or are they involved in olfaction? Do they differ between aquatic crustaceans compared with terrestrial ones? Additional knowledge from investigations and annotations of decapod crustacean genomes and antennal transcriptomes might lead to deeper insights in a future re-annotation of the present data.

Chemosensory receptors

So called “olfactory receptors” are described in many animal taxa including mammals (Buck and Axel 1991), nematodes (Thomas and Robertson 2008) and insects (Clyne et al. 1999). The main peripheral receptor families thereby have very low sequence similarity in all three taxa and are thus considered genetically unrelated and developed independently in each lineage. The general concept of an ion channel gated by ligands and coupled to a G-protein mediated signal cascade is common to all three receptor families (Dryer 2000).

Recent investigations of protostomian genomes identified ionotropic receptors (IRs), a receptor family derived from ionotropic glutamate receptors, as ancient chemosensory

[1] EST: Expressed Sequence Tag, here: transcript of a gene in a certain tissue.

receptors involved in insect olfaction (Benton et al. 2009; Croset et al. 2010). The crustacean antennaly expressed IRs belong to the same receptor family like the IRs of insects but form distinct subgroups separated from insect antennal IRs (manuscript 1 (Groh et al. 2014) and manuscript 2). Contrary to what is assumed concerning insects, which are the phylogenetic sister group of crustaceans, terrestrial hermit crabs did not derive an additional novel class of receptors to account for the needs of olfaction in air but instead use the same subclass of IRs like their aquatic relatives (manuscript 1, Groh et al. 2014). Moreover, these are expressed in olfactory sensory neurons (OSNs) and likely serve as olfactory receptors (manuscript 2). This raises the question of why IRs did not suffice for insects but were complemented with a novel receptor class. Molecular data of the olfactory system of basal insects is sparse compared to the well studied neopteran insects and has been completely unknown until recent investigations of *Lepismachilis y-signata* and *Thermobia domestica* (Missbach et al. 2014). *L. y-signata* also has an exclusively IR based olfactory system and performs well in sensing olfactory cues (Missbach et al. 2014), while the first occurrence of ORCo dates back earlier than the last common ancestor of “hemi-” and holometabolan insects, at least around 310 Mya (Krieger et al. 2003; Pitts, Fox and Zwiebel 2004; Smadja et al. 2009; Missbach et al. 2014). Thereafter the receptor count expanded, ranging from few OR genes in the human body louse *Pediculus humanus* (Kirkness et al. 2010), followed by an expansion of the family in more derived flying insects reaching their highest known number with more than 400 genes in social hymenopterans like ants (Zhou et al. 2012). The likely explanation of those findings is that not the terrestrialization but the ability to fly raised the need for a receptor class more suitable for the available odor information.

How to transfer an aquatic nose to air

A system that is suited to pick up chemical cues from aqueous solution is not easily transferred to air, just as the opposite is true. The example of *C. clypeatus* already gives hints to its remaining dependence on water in development and behavior. Though adult *Coenobita* are strictly terrestrial, the entire larval development is planctonic until the megalopa finally emerges to land (Brodie 2005). Larval chemosensing might thus be similar or almost identical with chemosensing in relatives that remain aquatic their entire life. Natural populations of adults usually do not live far from water sources, either at the coastal regions or at freshwater supplies further inland (Hartnoll 1988; Greenaway 2003). At any

occasion, *Coenobita* goes for fresh- and saltwater for drinking and prefers high humidity around 80% in its surrounding. No crabs are roaming their natural habitat if humidity drops below approx. 60%, as they dig holes in the moist substrate and hide until conditions reach their preferred range. Besides the general problem of preventing desiccation, the importance of humidity for the functionality of the hermit crab olfactory system has been documented (Krång et al. 2012). Humidity proved crucial to elicit responses in antennular electrical measurements following chemical stimulation and successful performance in trap assays providing food with and without water (Krång et al. 2012). Additional indications for the importance of water are properties of the aesthetasc cuticle. The cuticle is entirely poreless and like in all other Coenobitids it is thinner at the environmentally exposed side compared to the unexposed side (Stensmyr et al. 2005). Nevertheless the cuticle has been shown to be permeable to aqueous solutions (Ghiradella, Case and Cronshaw 1968a). Together with the fact that all compounds evidently detected by the *Coenobita* antennules are water soluble, it is conclusive to assume that the odor uptake by the aesthetascs is depending on water. Presumably in order to keep the olfactory system functional, Coenobitids adapted glands in the antennules to produce a mucosal substance that is excreted to the aesthetasc pad and keeps it covered in moisture to allow odorants to dissolve and diffuse through the aesthetasc cuticle on their way to the dendrites (manuscript 3). This aqueous milieu might be stable only in a certain range of humidity, explaining the observation of anosmia in dry air (Krång et al. 2012). Moist surfaces, on the other hand, tend to be colonized by microorganisms. This microflora can be harmful and additionally impair the functionality of an olfactory organ by producing smelly metabolites. Therefore aquatic and terrestrial hermit crabs produce substances to reduce microbial growth on the antennules (manuscript 1 (Groh et al. 2014)). The fact that no immunoreactive peptides identified from the antennal transcriptome were found in the proteomic approach might be due to peptide size. Average antimicrobial peptides of the crustin family, for example, have a size between 56 and 201 aminoacids (Sperstad et al. 2011) which equals a weight of 0.6 to 2 kDa. The acrylamide gel used for protein separation resolved proteins with a size of 10 kDa and larger, thus not allowing the analysis of such small peptides (manuscript 3). To further investigate the properties of the surface mucus, it might be suitable to especially focus on small compounds. Additionally it might be

interesting to assess the suspected antimicrobial properties by testing the putative inhibitory influence of antennules on a solid plate bacteria culture.

Another interesting aspect requiring further investigations was found in the CUB domain containing serine proteases (Csp) and their localization in the antennular glandular complexes and around the basal bodies of OSN dendrites (manuscript 3). The antibody used was originally raised against the Csp CUB domain in P. argus (Schmidt et al. 2006). Lobsters, as all aquatic decapods studied so far, do not have the vacuole around the basal body as it was exclusively described in the terrestrial Coenobitids so far (H. T. Ghiradella, Case and Cronshaw 1968). Whether the same CUB domain containing enzyme or group of enzymes are expressed in both tissues or each tissue has an own set of Csps remains unclear. The expression of a Csp in the glandular complexes and the likely excretion to the aesthetasc pad points to a putative involvement in immune response and/or to odor molecule modification (D Rittschof 1990; D Rittschof 1993; Sritunyalucksana and Soderhall 2000). With respect to the differing odorant spectra eliciting responses in lobsters and hermit crabs, the question arises, as to whether or not the activity of specific Csps determines the ability to detect odorants such as aminoacids.

Olfactory processing

The detection of chemical cues at the periphery is the crucial first step of olfaction. Nevertheless, successful performance needs integration and processing in order to evaluate this olfactory input. The data presented in manuscript 2 reveal that exemplary IRs are expressed in certain cells of a spindle like complex of OSN cell bodies. Besides the co-receptor IR25a and IR93a, which are broadly expressed for example in *P. argus* OSNs (Corey et al. 2013), the number of expressing cells depends on the IR and the antennular region. If in *Coenobita*, similar to findings in lobsters (Boekhoff et al. 1994), a single odor can elicit both inhibitory and excitatory responses in different subpopulations of OSNs, the innervation of the olfactory lobe can be assumed to reflect an integration pattern. In the crayfish brain, even a small group of aesthetasc OSNs projects to the entire OL and also in *Coenobita* at least many OSNs are multiglomerular, and glomeruli receive input from several OSNs ((Mellon and Munger 1990) Tuchina et al., in preparation). The glomeruli of the OL are additionally interconnected with thousands of local interneurons of many different

neurotransmitter profiles in a pattern that is not yet resolved and far from understood (R.E. Sandeman and Sandeman 1987; R E Sandeman, Sandeman and Watson 1990; Orona and Ache 1992; Langworthy et al. 1997; Schmidt and Ache 1997; Johansson et al. 1999; Yasuda-Kamatani and Yasuda 2006; Harzsch and Hansson 2008; Polanska et al. 2012).

Behavioral assays with visually covered putative food items suggest that *C. clypeatus* can not only localize food sources but differentiate between suitable (ripe banana) and unsuitable (overripe banana) food sources (Krång et al. 2012). Surprisingly none of the single compounds eliciting responses in EAG measurements attracted hermit crabs to the odor baits (Krång et al. 2012). This indicates that the recognition of a food item based on its odor information alone requires a blend to be detected and processed in the hermit crab brain. The data presented in this thesis (manuscript 1 (Groh et al. 2014) and manuscript 2), together with the morphology and wiring of the OL, indicate a capacity for complex processing mechanisms that are necessary to compute recognizable odor objects[2] in the brain of *Coenobita*. To what extent this wiring can be explained by odor coding at the periphery remains to be investigated. The yet unknown respective ligand spectra of the IRs identified in this thesis (manuscripts 1 and 2) together with future studies on whether there is a combinatorial fashion of IR expression in the OSNs, and the representation of single odors and blends in the OL, might lead to a better understanding of odor coding in the *Coenobita* olfactory pathway.

Olfaction and communication

An important exertion of the olfactory sense is intraspecific communication. Male lobsters that had lost a fight recognize the dominant male in a second exposure and display distinct subordinate behavior. However, if the inferior lobster was made anosmic, then the initial fight to establish dominance was repeated in the second encounter (Karavanich and Atema 1991). This behavior is independent from vision, as blindfolded lobsters fight for an equally long time and with little difference in performance, and instead seems to be mediated by a substance released with the urine of the lobsters (Breithaupt and Atema 1993; Kaplan et al. 1993). Hermit crabs are considered eusocial, as they tend to live in groups where, for

[2] An odor object is generated by those components of a blend that are required to recognize their source. An example is Isoamyl alcohol, Isoamyl acetate, Butyl acetate and Elemicine as defining odors out of the banana bouquet (Schubert, Hansson, and Sachse 2014).

example, snail shells are exchanged among individuals (Haas 1950; Dan Rittschof and Sutherland 1986). Aquatic hermit crabs, and most likely their terrestrial relatives, are able to differentiate between their own odor and the odor of conspecifics. Moreover, hermit crabs can tell apart familiar from unfamiliar conspecifics (Gherardi, Tricarico and Atema 2005). In most but not all crustaceans, the chemical information about the individuals metabolism, state of health and emotional status is transported via metabolites in urine, an excretion and distribution medium for hormones and pheromones (Breithaupt and Thiel 2011). While a huge body of evidence from decades of experiments points to the existence of crustacean pheromones, their understanding in terms of chemical structure remains elusive (Rittschof and Cohen 2004). Studies on crustacean chemical communication point towards biogenic amines, such as serotonin and dopamine, to influence crustacean behavior but data remain inconclusive (Beltz 1999). Additionally it is not known if intraspecific chemical communication is a feature of long or short distances, or if it requires physical contact. If a long distance communication is present in hermit crabs, then the antennules and their receptor repertoire could be a basis for future approaches. One indication for such a mechanism might be the annual event of terrestrial hermit crab migration. In the beginning of August coenobitid hermit crabs of both genders meet by the thousands to mate and later release their fertilized eggs into the ocean (Hazlett 1981). While the general initiation of this event might be due to an internally triggered rhythm, it might still be coordinated also by aggregation pheromones released by the individuals. In this thesis, antennules of both sexes were pooled for subsequent analysis so no gender differences between the olfactory systems of hermit crabs could be observed. RNA in *situ* hybridization experiments were also carried out regardless of sex but did not reveal differences between individual antennules (manuscript 2). Most decapods have no sexual dimorphism concerning sensilla on their antennules, although *Callinectes sapidus* male aesthetascs are sensitive to a yet unknown pheromone elicited by the females (Gleeson 1982; Hallberg, Johansson and Wallen 1997). Future studies might aim at investigating if hermit crabs possess a sex specific pheromone system, and if so might further aim to identify which receptors are responsible for this detection.

8. Summary

The question how objects are detected by the individual via the senses was already debated by pre-Socratic philosophers like Democritus and Theophrastus. However, the first studies based on modern scientific approaches and experimental methods were carried out far later. For example Jean Henri Fabre studied insects in this context in the late 19th and early 20th century. Among other ideas he investigated the attraction of male moths to female secretions, a behavior today called pheromone communication. Today, insect olfaction is well understood, as novel methods and techniques has allowed a shift of focus from behavioral observations to approaches combining behavior with underlying chemical and physical mechanisms of odor detection, as well as peripheral chemoreception events combined with the computation of the olfactory input to the brain. A more and more prominent question therfore asks for the evolutionary background and history. A lot of knowledge about the insect sense of smell was added after various insect genomes and transcriptomes were analyzed. In stark contrast hereto, the genetic understanding of the crustacean olfactory sense is still lacking. Although lobsters have been studied for decadeswith respect to morphology, mechanistics, signal transduction and modulation, behavior and learning capacities, there is still a lack of molecular data. Studies based on the enrichment of expressed genes comparing different tissues have been carried out successfully and have led to important insights into peripheral olfactory events. Nevertheless, the *Daphnia* genome is still the only annotated crustacean genome so far and currently no comprehensive study of decapod genetics is available.

My work was aimed at filling this gap in understanding of crustaceans by adding genetic information concerning the totality of expressed genes in the antennules of decapod crustaceans as well as investigating the adaptations required to transfer an aquatic nose to air. This thesis thereby provides not only the first in depth genetic analysis of crustacean antennules, but also the first comparative study of aquatic and terrestrial species with the aquatic *Pagurus bernhardus* and the terrestrial *Coenobita clypeatus* (manuscripts 1 and 2). Even given an independent evolution during the last 173 million years and a terrestrial history of 20 million years in *C. clypeatus*, we nevertheless revealed very subtle differences between their antennal transcriptomes, implying that a fully functional nose in air requires

small genetic adaptations from its ancestral aquatic organ. This was somewhat unexpected, as keeping the functionality of this nose, transferred from one medium to another, led to a plurality of morphological adaptations in structures supporting the mechanism. In the case of terrestrial hermit crabs, crucial support is presumably given by glandular structures. Glandular complexes consisting of two structurally different cell types produce and excrete a mucosal substance to keep the aesthetasc surface moist and the microbial colonization at bay (manuscript 3). While glandular structures in general have been described from different crustacean organs in various aquatic species, these particular glands seem to be specific adaptations of terrestrial Coenobitids. They are adapted to the special needs of an environmentally exposed organ sensitive to changes in humidity.

The only putatively chemosensory relevant receptors expressed in the antennules of *Pagurus* and *Coenobita* are ionotropic receptors, which belong to distinct subgroups separate from the insect olfactory IRs. Nevertheless, response profiles recorded from *Coenobita* antennules indicate similar ligand spectra for crustacean IRs, much like the IR based nose of *Drosophilids*. Besides the identification of hermit crab olfactory IRs, this thesis also provides insights into the distribution of these receptors in the *Coenobita* olfactory organ (manuscript 2). Crustacean antennules have been considered to be compound noses, with a repetitive pattern of aesthetascs tuned to certain odors. The data presented in this thesis indicates the contrary, as specific IRs are expressed in certain areas but not homogeneously throughout the aesthetasc pad. This could mean that the *Coenobita* nose is organized in a similar fashion to the *Drosophila melanogaster* antennae, with dedicated areas for the detection of different olfactory qualities.

In summary, this thesis is a first step towards the comprehensive understanding of the molecular biology of the crustacean olfactory organ. It provides a basis for investigating the ligand spectrum of individual IRs, their combinatorial expression and provides the framework for future work on odor representation in the crustacean brain.

9. Zusammenfassung

Schon zu prä-Sokratischer Zeit haben sich die griechischen Philosophen Demokrit und Theophrast mit der Wahrnehmung von Objekten durch Individuen befasst. Weit später hat Jean Henri Fabre, ein Pionier moderner Studien, mit wissenschaftlichen Ansätzen und experimentellen Methoden, dieses Thema an Insekten untersucht. Ende des 19. und zu Beginn des 20 Jahrhunderts führte er umfassende Studien durch und untersuchte zum Beispiel die Wirkung von Sekretionen weiblicher Motten auf Mottenmännchen, heute bekannt als pheromonbasierte Kommunikation. Dank der Etablierung neuer Methoden und Techniken ist der Geruchssinn von Insekten heute gut untersucht. Neben der reinen Verhaltensbeobachtung wurde auch die zusammenhängende Betrachtung des Verhaltens und der zugrundeliegenden chemischen und physiologischen Prozesse der Duftwahrnehmung, sowie die Verarbeitung der peripheren Wahrnehmung im Gehirn ermöglicht. Auch der Frage nach dem evolutionären Hintergrund konnte sich entscheidend angenähert werden, als eine Vielzahl von Insektengenomen und antennalen Transkriptomen untersucht und der Wissenschaftswelt zugänglich gemacht wurde. Im Kontrast dazu ist der Geruchssinn von Krebsen weit weniger gut verstanden. In Hummern wird zwar seit Jahrzehnten an diesem Thema aus morphologischer und funktioneller Sicht, sowie in Bezug auf Verhalten und Lernen geforscht, umfassende molekulargenetische Daten allerdings sind nicht verfügbar. Einzig durch den Vergleich der Anreicherung exprimierter Gene konnten wichtige Einblicke in das olfaktorische Geschehen in den Antennen gewonnen werden. Dennoch gibt es derzeit abgesehen von *Daphnia* keine umfassenden Genomstudien in Krebsen allgemein und erst recht nicht in Dekapoden.

Meine Arbeit hat das Ziel diese Lücke zu schließen, indem sie zum einen die Gesamtheit der transkribierten Gene in Dekapodenantennen und zum anderen die Anpassung dieser Genexpression in Abhängigkeit vom umgebenden Medium untersucht. Sie ist dabei nicht nur die erste tiefgreifende molekulargenetische Studie der der Krebsantenne, sondern auch die erste vergleichende Studie die mit *Pagurus bernhardus* aquatische und mit *Coenobita clypeatus* landlebende Dekapoden einbezieht (Manuskripte 1 und 2). Dabei hat sich gezeigt, dass 173 Millionen Jahre insgesamt und bei *Coenobita* 20 Millionen Jahre davon an Land, nur wenig an der genetischen Ausstattung der Antennen verändert haben. Mit anderen Worten ausgedrückt: Es braucht nicht viel, um die Funktion eines ursprünglich an Wasser

angepassten Organs auf das Landleben einzustellen. In Anbetracht der zahlreichen morphologischen Veränderungen der Antennen in der Folge des Landganges, war dies ein unerwarteter Befund. So wurde in *Coenobita* ein Drüsenkomplex gefunden, der sehr wahrscheinlich entscheidenden Einfluss auf die Funktionsfähigkeit hat. Diese Drüsen bestehen aus zwei unterscheidbaren Zelltypen und produzieren ein schleimiges Sekret, welches durch Poren zwischen den Aesthetasken ausgeschieden wird. Dieses Sekret hat vermutlich die Aufgabe die Aesthetasken feucht zu halten und eine mikrobielle Besiedelung der Oberfläche zu verhindern (Manuskript 3). Drüsen im Allgemeinen sind aus verschiedenen aquatischen Krebsspezies bekannt, wobei dieser spezielle Typ eine spezifische Anpassung landlebender Coenobitiden zu sein scheint. Er unterstützt die Funktion eines Organs, das Umwelteinflüssen ausgesetzt ist und sensibel auf Veränderungen der relativen Luftfeuchtigkeit reagiert.

Die einzigen mutmaßlich chemosensorisch relevanten Rezeptoren, die in den Antennen von sowohl *Pagurus* als auch *Coenobita* exprimiert werden, sind ionotrope Rezeptoren. Sie gehören dabei zu einer Untergruppe der IRs, die sich deutlich von den bekannten antennalen IRs der Insekten abgrenzt. Trotzdem ähneln die Duftantworten von *Coenobita*-Antennen den Ligandenspektren der IRs von *Drosophila*. Neben der Identifikation der Einsiedlerkrebs-IRs konnte in dieser Arbeit auch ihre Verteilung in den Antennen von *Coenobita* gezeigt werden (Manuskript 2). Krebsantennen wurden als „Komplexnasen“ berichtet, die ein sich wiederholendes Muster an Aesthetasken mit ihren entsprechenden Duftprofilen präsentieren. Die Befunde dieser Arbeit deuten auf das Gegenteil hin, da die untersuchten IRs nicht gleichmäßig über die Länge des letzten Antennensegmentes verteilt, sondern in distinkten Bereichen des Aesthetaskenfeldes vertreten sind. Das könnte bedeuten, dass die Nase von *Coenobita* ähnlich organisiert ist wie die Antenne von *Drosophila*, mit Arealen die auf die Detektion bestimmter Düfte oder Duftgruppen spezialisiert sind.

Insgesamt ist diese Arbeit ein erster Schritt hin zu einem umfassenden molekularbiologischen Verständnis des olfaktorischen Organs von Krebsen. Es stellt eine Basis für weitere Untersuchungen dar, die das Ligandenspektrum der einzelnen IRs sowie

ihre Kombinatorik entschlüsseln können, sowie zukünftige Arbeiten an der Repräsentation des Geruchssinnes im Gehirn der Krebse ermöglichen.

10. References

Ache B W and Derby C D. 1985. “Functional organization of olfaction in crustaceans.” Trends Neurosci 8 (8): 356–364. doi:10.1016/0166-2236(85)90122-5.

Ache B W and Young J M. 2005. “Olfaction: diverse species, conserved principles.” Neuron 48: 417–30. doi: 10.1016/j.neuron.2005.10.022

Ai ., Min S, Grosjean Y, Leblanc C, Bell R, Benton R and Suh G. 2010. “Acid sensing by the Drosophila olfactory system.” Nature (London) 468 (7324): 691. doi:10.1038/nature09537.

Allmann S, Späthe A, Bisch-Knaden S, Kallenbach M, Reinecke A, Sachse S, Baldwin I T and Hansson B S. 2013. “Feeding-induced rearrangement of green leaf volatiles reduces moth oviposition.” eLife 2 (January): e00421. doi:10.7554/eLife.00421.

Angerer L M and Angerer R C. 1992. “In situ hybridization to cellular RNA with radiolabeled RNA probes.” In In Situ Hybridization, edited by D. G. Wilkinson, 17th ed. Oxford: IRL Press.

Ashburner M, Ball C A, Blake J A, Botstein D, Butler H, Cherry J M, Davis A P, et al. 2000. “Gene Ontology: tool for the unification of biology.” Nature Genetics 25 (1): 25–29. doi:10.1038/75556

Arnaud J, Brunet M and Mazza J. 1988. “Labral glands in *Centropages typicus* (Copepoda, Calanoida). 1. Sites of Synthesis.” J Morphol. 197 (1): 21–32.

Arseault A, Clatenberg R and Aiken D. 1979. “The morphology and secretory transporting mechanism of the tegumental glands of the lobster (*Homarus americanus*) as related to the molting cycle.” J Submicroscopic Cytology 11: 193–207.

Attygalle A B and Morgan E D. 1985. “Ant trail pheromones.” In Advances in Insect Physiology, edited by Michael J. Berridge, J. E. Treherne, and Vincent Brian Wigglesworth, Vol 18, 1–30. Academic Press, New York.

Atema J. 1977. “Functional separation of smell and taste in fish and crustaceans.” In Olfaction and Taste, ed. P. Le Magnen, J. MacLeod, 165–174. VI. London: Information Retrieval Ltd.

Backer P and Gibson R. 1977. “Observations on the feeding mechanism structure of the gut and digestive physiology of the European lobster *Homarus gammarus* (L.) (Decapoda: Nephropidae)”. J Exp Mar Biol Ecol 26: 297–324.

Barbato J C and Daniel P C. 1997. “Chemosensory activation of an antennular grooming behavior in the spiny lobster *Panulirus argus* is tuned narrowly to L-Glutamate.” Marine Biology Laboratory 193: 107–115.

Barrett A J. 1994. “Classification of peptidases.” Volume 244, methods in enzymology: proteolytic enzymes: serine and cysteine peptidases. California, USA: Academic Press. pp. 1–15.

Bayer T A, McClintock T S, Grunert U and Ache B W. 1989. “Histamine-induced modulation of olfactory receptor neurons in 2 species of lobster, *Panulirus-argus* and *Homarus-americanus*.” J Exp Biol 145: 133–146.

Bell J C. 1906. “The reactions of crayfish to chemical stimuli.” J Comp Neurol Psycho 16 (4): 299–326.

Beltz B S. 1999. “Distribution and functional anatomy of amine-containing neurons in decapod crustaceans.” Microsc. Res. Tech. 44 (2-3): 105–20. doi:10.1002/(SICI)1097-0029(19990115/01)44:2/3

Beltz B S, Kordas K, Lee M M, Long J B, Benton J L and Sandeman D C. 2003. “Ecological, evolutionary, and functional correlates of sensilla number and glomerular density in the olfactory system of decapod crustaceans.” J Comp Neurol 455 (2): 260–269. doi:10.1002/cne.10474.

Benton R, Vannice K S, Gomez-Diaz C and Vosshall L B. 2009. “Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*.” Cell 136 (1): 149–162. doi:DOI 10.1016/j.cell.2008.12.001.

Bliss D E and Mantel L H. 1968. “Adaptations of crustaceans to land - a summary and analysis of new findings.” Am Zool 8 (3): 673–685.

Boekhoff I, Michel W C, Breer H and Ache B W. 1994. “Single odors differentially stimulate dual 2nd messenger pathways in lobster olfactory receptor-cells.” J Neuro 14 (5): 3304–3309.

Bork P and Beckmann G. 1993. “The CUB domain. A widespread module in developmentally regulated proteins.” J Mol Biol 231: 539–545.

Bracken-Grissom H D, Cannon M E, Cabezas P, Feldmann R M, Schweitzer C E, Ahyong S T, Felder D L, Lemaitre R and Crandall K. 2013. “A comprehensive and integrative reconstruction of evolutionary history for anomura (Crustacea: Decapoda).” BMC Evol Biol 13 (1) (January): 128. doi:10.1186/1471-2148-13-128.

Breithaupt T and Atema J. 1993. “Evidence for the use of urine signals in agonistic interactions of the american lobster” Biol Bull185 (2): 318. <Go to

Breithaupt T and Thiel M. 2011. “Chemical communication in crustaceans.” Edited by Thomas Breithaupt and Martin Thiel. New York: Springer Science + Business Media.

Brodie R J. 2005. “Desiccation resistance in megalopae of the terrestrial hermit crab *Coenobita compressus*: water loss and the role of the shell.” Invertebr Biol 124 (3) (September): 265–272. doi:10.1111/j.1744-7410.2005.00024.x.

Brunet M, Cuoc C, Arnaud J and Mazza J. 1991. “Tegumental glands in a copepod *Hemidiaptomus ingens*: structural, ultrastructural and cytochemical aspects.” Tissue Cell 23: 733–743. doi: http://dx.doi.org/10.1016/0040-8166(91)90026-P

Buck L and Axel R. 1991. “A novel multigene family may encode odorant receptors - a molecular-basis for odor recognition.” Cell 65 (1): 175–187.

Buehlmann, C., Hansson B. S. and Knaden M. “Screening the desert for food: olfactory-driven foraging strategies in desert ants.” In Preparation.

Bush B M H and Laverack M S. 1982. “Mechanoreceptors.” In Biology of Crustacea., ed. D. H. Bliss, H. L. Atwood and D. C. Sandeman, 369–468. Volume 3. Academic Press, New York.

Carr W E S, Ache B W, and Gleeson R A. 1987. “Chemoreceptors of crustaceans - similarities to receptors for neuroactive substances in internal tissues.” Environ Health Persp71: 31–46. doi:10.2307/3430411.

Cate H S and Derby C D. 2001. “Morphology and distribution of setae on the antennules of the caribbean spiny lobster *Panulirus Argus* reveal new types of bimodal chemo-mechanosensilla.” Cell Tissue Res 304 (3): 439–454. doi:10.1007/s004410100377.

Cate H S and Derby C D 2002. “Hooded sensilla homologues: structural variations of a widely distributed bimodal chemomechanosensillum.” J Comp Neurol 444 (4): 345–357. doi:10.1002/cne.10153.

Cerenius L, Jiravanichpaisal P, Liu H P and Soederhall I. 2010. “Crustacean Immunity.” In Invertebrate Immunity, ed. K Soderhall, 708:239–259.

Chapman A D. 2009. “Numbers of Living Species in Australia and the World.” 2nd edition. Australian Biodiversity Information Services, Toowoomba, Australia.

Clyne P J, Warr C G, Freeman M R., Lessing D, Kim J and Carlson J R. 1999. “A novel family of divergent seven-transmembrane proteins.” Neuron 22 (2): 327–338.

Colbourne J K, Pfrender M E, Gilbert D, Thomas W K, Tucker A, Oakley T H, Tokishita S, et al. 2011. “The ecoresponsive genome of *Daphnia pulex*.” Science 331 (6017): 555–561. doi:10.1126/science.1197761.

Conesa A, Goetz S, Garcia-Gomez J M, Terol J, Talon M and Robles M. 2005. “Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.” Bioinformatics 21 (18): 3674–3676. doi:10.1093/bioinformatics/bti610.

Corey E A, Bobkov Y, Ukhanov K and Ache B W. 2013. “Ionotropic Crustacean Olfactory Receptors.” PloS One 8 (4) (January): e60551. doi:10.1371/journal.pone.0060551.

Croset V, Rytz R, Cummins S F, Budd A, Brawand D, Kaessmann H, Gibson T J and Benton R. 2010. “Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste.” PLoS Genetics 6 (8). doi:e1001064 10.1371/journal.pgen.1001064.

Derby C D and Atema J. 1982. “Chemosensitivity of walking legs of the lobster *Homarus americanus* - neurophysiological response spectrum and thresholds.” J Exp Biol 98 (JUN): 303–315.

Dionne V E and Dubin A E. 1994. “Transduction diversity in olfaction.” J Exp Biol 21: 1–21.

Doughtie D and Rao K. 1982. “Rosette glands in the gills of the grass shrimp *Palaemonetes pugio*.” J Morph 171: 41–67.

Dryer L. 2000. “Evolution of odorant receptors.” Bioessays 22 (9): 803–810.

Eder S and Atema J. 1978. “Behavioral evidence for functional separation of taste and smell in lobster, *Homarus americanus*.” Biol Bull155 (2): 435.

Edgar R C. 2004. “MUSCLE: multiple sequence alignment with high accuracy and high throughput.” Nucleic Acids Res 32 (5): 1792-1797.

Elofsson R, Aramant R, Lindvall O and Falck B. 1978. “Catecholaminergic salivary glands in *Gammarus pulex* (Crustacea, Amphipoda): an electron microscopic and microspectrofluorometric study general appearance and fluorescence.” J Ultrastruct Res 64: 14–22.

Festoff B W. 1990. “Serine proteases and their serpin inhibitors in the nervous system.” 191. Pros NATO ASI Series A191. New York: Premium Press. 359 p.

Gao Q, Yuan B and Chess A. 2000. “Convergent projections of Drosophila olfactory neurons to specific glomeruli in the antennal lobe.” Nat Neurosci 3 (8) (August): 780–5. doi:10.1038/77680.

Gherardi F, Tricarico E and Atema J. 2005. “Unraveling the nature of individual recognition by odor in hermit crabs.” J Chem Ecol 31 (12): 2877–2896. doi:10.1007/s10886-005-8400-5.

Ghiradella H, Case J and Cronshaw J. 1968a. “Fine structure of aesthetasc hairs of *Coenobita compressus* Edwards.” J Morph 124 (3) (March): 361–385. doi:10.1002/jmor.1051240309. h

Ghiradella H T, Case J F and Cronshaw J. 1968b. “Structure of aesthetascs in selected marine and terrestrial decapods - chemoreceptor morphology and environment.” Am Zool 8 (3): 603–&.

Glaessner M F. 1969. “Decapoda.” In Treatise on Invertebrate Palaeontology, edited by K. S. Lawrence, 400–533. Geological Society of America and Kansas University Press.

Gleeson R A. 1982. “Morphological and behavioral identification of the sensory structures mediating pheromone reception in the blue crab , *Callinectes sapidus*.” Biol Bull163 (1): 162–171.

Gnatzy W. 1984. “Campaniform structures on lobster antennae are dermal glands.” Cell Tissue Res 236: 729–731.

Gong D P, Zhang H J, Zhao P, Xia Q Y and Xiang Z H. 2009. “The odorant binding protein gene family from the genome of silkworm, *Bombyx mori*.” BMC Genomics, 10. doi:332 10.1186/1471-2164-10-332

Gorvett H. 1946. “The tegumental glands in the land Isopoda. A. The rosette glands.” QJ Micros Sci 87: 209–234.

Gorvett H. 1951. “The tegumental glands in the land Isopoda. B. The lobed glands: structure and distribution. QJ Micros Sci 92: 275–296.

Gorvett H. 1956. “Tegumental glands and the terrestrial life in wood lice.” Proc Zool Soc Lond 126: 291–314.

Greenaway P, Morris S and McMahon BR. 1988. “Adaptations to a terrestrial existence by the robber crab Birgus latro: II. In Vivo Respiratory Gas Exchange and Transport.” J Exp Biol 140: 493–509.

Greenaway, P. 2003. “Terrestrial adaptations in the Anomura (Crustacea: Decapoda).” Memoirs of Museum Victoria 60 (1).

Grimaldi, D A. 2010. “Fossil record and phylogeny of the Arthropoda introduction.” Arthropod Struct Dev 39 (2-3): 72–73. doi:10.1016/j.asd.2010.01.001.

Groh K C, Vogel H, Stensmyr M C, Grosse-Wilde E and Hansson B S. 2014. “The hermit crab’s nose – antennal transcriptomics.” Front Neurosci 7. doi:10.3389/fnins.2013.00266.

Grosse-Wilde, E, Kuebler L S, Bucks S, Vogel H, Wicher D and Hansson B S. 2011. “Antennal transcriptome of *Manduca sexta*.” PNAS 108 (18): 7449–7454. doi:10.1073/pnas.1017963108.

Grunert U and Ache B W. 1988. „Ultrastructure of the aesthetasc (olfactory) sensilla of the spiny lobster *Panulirus argus*.” Cell Tissue Res 251: 95–103.

Guenther C M, Miller H A, Basil J A and Atema J. 1996. “Orientation behavior of the lobster: responses to drectional chemical and hydrodynamic stimulation of the antennules.” Biol Bull191 (2): 310–311.

Haas F. 1950. “Hermit crabs in fossil snail shells in bermuda.” Ecology 31 ((1)): 152.

Habermann B, Oegema J, Sunyaev S and Shevchenko A. 2004. “The power and the limitations of cross-species protein identification by mass spectrometry-driven sequence similarity searches.” Mol Cell Proteomics 3: 238–249.

Hallberg E, Johansson K U I and Wallen R. 1997. “Olfactory sensilla in crustaceans: morphology, sexual dimorphism, and distribution patterns.” Int J Insect Morphol 26 (3-4): 173–180.

Han-Ching Wang K, Tseng C, Lin H, Chen I T, Chen Y H, Chen Y I, Chen T Y and Yang H L. 2010. “RNAi knock-down of the *Litopenaeus vannamei* Toll gene (LvToll) significantly increases mortality and reduces bacterial clearance after challenge with *Vibrio harveyi*.” Dev Com Immunol 34 (1) (January): 49–58. doi:10.1016/j.dci.2009.08.003.

Hansson B S, Harzsch S, Knaden M and Stensmyr M. 2011. “The neural and behavioral basis of chemical communication in terrestrial crustaceans.” In Chemical Communication in Crustaceans, edited by T Breithaupt and M Thiel, 149–173. Springer Science + Business Media. doi:10.1007/978-0-387-77101-4_8.

Hartnoll RG. 1988. “Evolution, systematics and geographical distribution.” In Biology of the Land Crabs, edited by McMahon B R and Burggren W W, 6–54. Cambridge, UK: Cambridge University Press.

Harzsch S and Bill S Hansson. 2008. “Brain architecture in the terrestrial hermit crab *Coenobita clypeatus* (Anomura, Coenobitidae), a crustacean with a good aerial sense of smell.” BMC Neurosci 9 (58) (January): 58. doi:10.1186/1471-2202-9-58.

Hatt H and Ache B W. 1994. “Cyclic nucleotide and inositol phosphate gated ion channels in lobster olfactory receptor neurons.” P Natl Acad Sci USA 91 (14): 6264–6268. doi:10.1073/pnas.91.14.6264.

Hazlett B A. 1981. “The behavioral ecology of hermit crabs.” Annu Rev Ecol Syst 12: pp. 1–22.

Hogan B L. 1996. “Bone morphogenetic proteins: multifunctional regulators of vertebrate development.” Gene Dev 10(13), pp.1580–1594.

Hollins B, Hardin D, Gimelbrant A A and McClintock T S. 2003. “Olfactory-enriched transcripts are cell-specific markers in the lobster olfactory organ.” J Comp Neurol, 455(1), pp.125–138.

Hornung E. 2011. “Evolutionary adaptation of oniscidean isopods to terrestrial life: Structure, physiology and behavior.” Terrestrial Arthropod Reviews, 4(2), pp.95–130.

Hurley D E. 1968. “Transition from water to land in amphipod crustaceans.” Am Zool 353 (1905): 327–353.

Jacquin-Joly E, Legeai F, Montagne N, Monsempes C, Francois M C, Poulain J, Gavory F, Walker Iii W B, Hansson B S and Larsson M C. 2012. “Candidate chemosensory genes in female antennae of the noctuid moth *Spodoptera littoralis*.” Int J Biol Sci 8 (7): 1036–1050. doi:10.7150/ijbs.4469.

Johansson K U I, Lundquist C T, Hallberg E and Nässel D R. 1999. “Tachykinin-related neuropeptide in the crayfish olfactory midbrain.” Cell Tissue Res 296 (2) (April 21): 405–415. doi:10.1007/s004410051300.

Johns M E, Tai P C and Derby C D. 2004. „Serine proteases in the spiny lobster olfactory organ: their functional expression along a developmental axis, and the contribution of a CUB-serine protease.” J Neurobiol, 61(3), pp.377–91. Doi: 10.1002/neu.20056

Johnson B R and Atema J. 1983. “Narrow-spectrum chemoreceptor cells in the antennules of the american lobster, *Homarus Americanus*.” Neuroscience Letters 41 (1-2): 145–150. doi:10.1016/0304-3940(83)90237-9.

Johnson B R, Merrill C L, Ogle R C and Atema J. 1988. “Response properties of lobster chemoreceptors - tuning of olfactory neurons sensitive to hydroxy-proline.” J Comp Physiol A 162 (2): 201–211. doi:10.1007/bf00606085.

Johnson B and Talbot P. 1987. “Ultrastructural analysis of the pleopod tegumental glands in the male and female lobsters, *Homarus americanus*.” J Crust Biol, 7, pp.288–301.

Juberthie-Jupeau L and Crouau Y. 1978. “The tegumental glands of a troglobitic crustacean.” Int. J. Speleol., 9, pp.309–319.

Kaissling K E. 2001. “Olfactory perireceptor and receptor events in moths: a kinetic model.” Chem. Senses. 26: 125-150

Kaplan L J, Lowrance C, Basil J and Atema J. 1993. “The role of chemical and visual cues in agonistic interactions of the american lobster.” Biol Bull 185 (2): 320–321.

Karavanich C and Atema J. 1991. “Role of olfaction in recognition of dominance in the american lobster *(Homarus americanus)*.” Biol Bull 181 (2): 359–360.

Keil T. 2012. “Sensory Cilia in Arthropods.” Arthropod Struct Dev 41 (6) (July 17): 515–534. doi:10.1016/j.asd.2012.07.001.

Kirkness E F, Haas B J, Sun W, Braig H R, Perotti M A, Clark J M, Lee S H, et al. 2010. “Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle.” Proc Natl Acad Sci USA 107 (27) (July 6): 12168–73. doi:10.1073/pnas.1003379107.

Knudsen J T and Gershenzon J. 2006. „Biology of Floral Scent." Edited by Natalia Dudareva and Eran Pichersky. Taylor & Francis Group. LLC.

Koczan S. 2012. "Diploma thesis: 3D reconstruction of the olfactory pathway from the aesthetasc to the olfactory lobe in the hermit crab, *Coenobita clypeatus*". Friedrich-Schiller University Jena - Faculty of Biology and Pharmacy.

Koehl M A R, Koseff J R, Crimaldi J P, McCay M G, Cooper T, Wiley M B and Moore P A. 2001. "Lobster sniffing: antennule design and hydrodynamic filtering of information in an odor plume." Science (Washington D C) 294 (5548): 1948–1951. doi: 10.1126/science.1063724

Krång A S, Knaden M, Steck K and Hansson B S. 2012. "Transition from sea to land: olfactory function and constraints in the terrestrial hermit crab *Coenobita clypeatus*." Proc R Soc B 279 (1744) (September 7): 3510–9. doi:10.1098/rspb.2012.0596.

Kulmuni J and Havukainen H. 2013. "Insights into the evolution of the CSP gene family through the integration of evolutionary analysis and comparative protein modeling." PloS one 8 (5) e63688. doi: 10.1371/journal.pone.0063688

Krieger J, Klink O, Mohl C, Raming K and Breer H. 2003. "A candidate olfactory receptor subtype highly conserved across different insect orders." J Comp Phys A 189 (7): 519–526. doi: 10.1007/s00359-003-0427-x

Kwon H W, Lu T, Rutzler M and Zwiebel L J. 2006. "Olfactory responses in a gustatory organ of the malaria vector mosquito *Anopheles gambiae*." P Natl Acad Sci USA 103 (36): 13526–13531. doi: 10.1073/pnas.0601107103

Langworthy K, Helluy S, Benton J and Beltz B. 1997. "Amines and peptides in the brain of the american lobster: immunocytochemical localization patterns and implications for brain function." Cell Tiss Res 288 (1) (March 11): 191–206. doi:10.1007/s004410050806.

Leal W S, Chen A M, Ishida Y, Chiang V P, Erickson M L, Morgan T I and Tsuruda J M. 2005: "Kinetics and molecular properties of pheromone binding and release." P Natl Acad Sci 102 (15): 5386-5391. doi: 10.1073/pnas.0501447102

Legeai F, Malpel S, Montagne N, Monsempes C, Cousserans F, Merlin C, Francois M C, et al. 2011. "An expressed sequence tag collection from the male antennae of the noctuid moth *Spodoptera littoralis*: a resource for olfactory and pheromone detection research." BMC Genomics 12 (86). doi:10.1186/1471-2164-12-86.

Levine M and Harrison P, 2001. "A CUB-serine protease in the olfactory organ of the spiny lobster *Panulirus argus*." J Neurobiol, 49(4), pp.277–302.

Li G Z, Vissers J P C, Silva J C, Golick D, Gorenstein M V and Geromanos S J. 2009. "Database searching and accounting of multiplexed precursor and product ion spectra from the data

independent analysis of simple and complex peptide mixtures." PROTEOMICS, 9(6), pp.1696–1719. doi: 10.1002/pmic.200800564

Lindqvist O V. 1972. "Components of water loss in terrestrial isopods." Phys Zool 45 (4) (October 1): 316–324; doi:10.2307/30155588.

Little C. 2009. "The Terrestrial Invasion." Cambridge University Press.

McClintock T S and Ache B W. 1989. "Histamine directly gates a chloride channel in lobster olfactory receptor neurons." P Natl Acad Sci USA 86 (20): 8137–8141. doi:10.1073/pnas.86.20.8137.

McClintock T S, Xu F, Quintero J, Gress A M and Landers T M. 1997. "Molecular cloning of a lobster G alpha(q) protein expressed in neurons of olfactory organ and brain." J Neurochem 68 (6) (June): 2248–54.

McLaughlin P A, Lemaitre R and Sorhannus U. 2007. "Hermit crab phylogeny: a reappraisal and its 'fall-out'." J Crust Bio 27 (1) (January 1): 97–115. doi:10.1651/S-2675.1.

McMahon B R and Burggren W W. 1979. "Respiration and adaptation to the terrestrial habitat in the land hermit crab *Coenobita clypeatus*." J Exp Biol 79 (1) (April 1): 265–281.

Mellon D and Munger S D. 1990. "Nontopographic projection of olfactory sensory neurons in the crayfish brain." J Comp Neuro 296 (2) (June 8): 253–62. doi:10.1002/cne.902960205.

Missbach, C, H K M Dweck, H Vogel, A Vilcinskas, M C Stensmyr, B S Hansson, and E Grosse-Wilde. 2014. "Evolution of Insect Olfactory Receptors." *eLife 2014;3:e02115* doi: 10.7554/eLife.02115

Montell, C. 2009. "A Taste of the Drosophila Gustatory Receptors." Curr Opin Neurobiol 19 (4): 345–353. doi:10.1016/j.conb.2009.07.001.

Müller C H G, Rosenberg J and Hilken G. 2009. „Fine structure and phylogenetic significance of "flexo-canal epidermal glands" in Chilopoda." In Xylander W and Voigtländer K, Proceedings of the 14th International Congress of Myriapodology, Soil Organisms. pp. 269–294.

Müller C H G, Sombke A, Hilken G and Rosenberg J. 2011. „Sense organs." In Minelli A., Treatise on Zoology - Anatomy, Taxonomy, Biology - The Myriapoda. Brill, Leiden, pp. 235–278.

Müller C H G, Rosenberg J and Hilken G. 2014. „Ultrastructure, functional morphology and evolution of recto-canal epidermal glands in Myriapoda." Arthropod Struct Dev 43(1):43-61, doi: 10.1016/j.asd.2013.08.001

Orona E and Ache B W. 1992. “Physiological and pharmacological evidence for histamine as a neurotransmitter in the olfactory CNS of the spiny lobster.” Brain Res 590 (1-2): 136–143. doi:10.1016/0006-8993(92)91089-w.

Penalva-Arana D C, Lynch M and Robertson H M. 2009. “The chemoreceptor genes of the waterflea *Daphnia pulex*: many GRs but no ORs.” BMC Evol Biol 9 (79). doi:79 10.1186/1471-2148-9-79.

Pelosi P, Zhou J J, Ban L P and Calvello M. 2006. “Soluble proteins in insect chemical communication.” Cell mol life sci 63(14), 1658–76. doi:10.1007/s00018-005-5607-0

Pitts R J, Fox A N and Laurence J Z. 2004. “A highly conserved candidate chemoreceptor expressed in both olfactory and gustatory tissues in the malaria vector *Anopheles gambiae*.” Proc N A Sc USA 101 (14): 5058–5063.

Polanska M A, Tuchina O, Agricola H, Hansson B S and Harzsch S. 2012. “Neuropeptide complexity in the crustacean central olfactory pathway: Immunolocalization of A-type allatostatins and RFamide-like peptides in the brain of a terrestrial hermit crab.” Mol Brain 5 (1) (January): 29. doi:10.1186/1756-6606-5-29.

Powers L W and Bliss D E. 1983. “Terrestrial adaptations.” In The Biology of Crustacea. Volume 8. Environmental Adaptations., edited by F J Vernberg and W B Vernberg, 271–333. Academic Press, New York, London etc.

Price M N, Dehal P S and Arkin A P 2010. “FastTree 2--approximately maximum-likelihood trees for large alignments.” Ed. Art F. Y. Poon. PloS One 5 (3) (January). doi:10.1371/journal.pone.0009490.

Quennedey A, 1998. “Insect epidermal gland cells: ultrastructure and morphogenesis.” In F. Harrison & M. Locke, eds. Microscopic anatomy of Invertebrates. Wiley-Liss, New York, pp. 177–207.

Regier J C, Shultz J W and Kambic R E. 2005. “Pancrustacean phylogeny: Hexapods are terrestrial crustaceans and maxillopods are not monophyletic.” P Roy Soc B-Biol Sci 272 (1561): 395–401. doi:10.1098/rspb.2004.2917.

Richardson A M M 2007. “Behavioral Ecology of Semiterrestrial Crayfish.” In Evolutionary Ecology of Social and Sexual Systems : Crustaceans as Model Organisms, edited by J. E. Duffy and C. M. Thiel, 319–339. Oxford University Press.

Richardson K, Jarrett L and Finke E. 1960. “Embedding in epoxy resins for ultrathin sectioning in electron microscopy.” Stain Technol., 35, pp.313–323.

Richter S. 2002. “The tetraconata concept: Hexapod-crustacean relationships and the phylogeny of crustacea.” Org Divers Evol 2 (3): 217–237. doi: 10.1078/1439-6092-00048

Richter S, Moller O S and Wirkner C S. 2009. “Advances in crustacean phylogenetics.” Arthropod Systematics & Phylogeny 67 (2): 275–286.

Rieder N. 1977. „ Ultrastruktur und Funktion der Hautdruesen von *Triops cancriformis* Bosc. (Crustacea, Notostraca) “, Zoomorphologie 88, pp.133–143.

Rittschof D. 1980. “Chemical attraction of hermit crabs and other attendants to simulated gastropod predation sites.” J Chem Ecol 6(1), pp.103–118.

Rittschof D. 1980. “Enzymatic production of small molecules attracting hermit crabs to simulated gastropod predation sites.” J Chem Ecol 6(3), pp.665–675.

Rittschof D, Barlow L and Schmidt AR. 1985. “Laboratory studies of olfaction and taste in land hermit crabs coenobita-rugosus.” Am Zool, 25(4), p.113A.

Rittschof D, Forward R B and Mott D D. 1985. “Larval release in the crab *Rhithropanopeus harrisii* (Gould): chemical cues from hatching eggs.” Chemical Senses, 10(4), pp.567–577.

Rittschof D and Sutherland J P. 1986. “Field studies on chemically mediated behavior in land hermit crabs *Coenobita-rugosus* volatile and nonvolatile odors.” J Chem Ecol 12(6), pp.1273–1284.

Rittschof D. 1990. “Peptide-mediated behaviors in marine organisms evidence for a common theme.” J Chem Ecol 16 (1) (January): 261–72. doi:10.1007/BF01021283.

Rittschof D, Forward R B and Erickson BW. 1990. “Larval release in brachyuran crustaceans functional similarity of peptide pheromone receptor and catalytic site of trypsin.” J Chem Ecol 16(4), pp.1359–70.

Rittschof D. 1993. “Body odors and neutral-basic peptide mimics : A review of responses by marine organisms” Author (s): Dan Rittschof Published by : Oxford University 33 (6): 487–493.

Rittschof D and Cohen J H. 2004. “Crustacean peptide and peptide-Like pheromones and kairomones.” Peptides 25 (9): 1503–1516. doi: 10.1016/j.peptides.2003.10.024

Robalino, J., Almeida, J. S., McKillen, D., Colglazier, J., Trent, H. F., Chen, Y. A., Peck, M. E. T. et al. 2007. “Insights into the immune transcriptome of the shrimp *Litopenaeus vannamei*: Tissue-specific expression profiles and transcriptomic responses to immune challenge.” Physiological Genomics 29 (1): 44–56. doi:10.1152/physiolgenomics.00165.2006

Robertson H M and Kent L B. 2009. “Evolution of the gene lineage encoding the carbon dioxide receptor in insects.” J Insect Sci 9 (19). doi: 10.1673/031.009.1901

Robertson H M, Warr C G and Carlson J R. 2003. “Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*.” Proc Nat Acad Sci USA 100: 14537–14542. doi:10.1073/pnas.2335847100.

Rosa, R. D., and Barracco, M. A. 2010. “Antimicrobial peptides in crustaceans.” Invertebrate Survival Journal 7 (2): 262–284.

Rosen E, Schwarz B and PalmerA R. 2009. “Smelling the difference: Hermit crab responses to predatory and nonpredatory crabs.” Animal Behaviour 78 (3): 691–695. doi:10.1016/j.anbehav.2009.05.035.

Rosenzweig, M., Brennan, K. M., Tayler, T. D., Phelps, P. O., Patapoutian, A., and Garrity, P. A. 2005. “The Drosophila Ortholog of Vertebrate TRPA1 Regulates Thermotaxis.” Gene Dev 19 (4): 419–424. doi:10.1101/gad.1278205.fied.

Sanchez-Gracia A, Vieira F G and Rozas J. 2009. “Molecular evolution of the major chemosensory gene families in insects.” Heredity, 103(3), 208–216. doi:10.1038/hdy.2009.55

Sandeman R E, Sandeman D Z and Watson A H. 1990. “Substance P antibody reveals homologous neurons with axon terminals among somata in the crayfish and crab brain.” J Comp Neurol 294 (4) (April 22): 569–82. doi:10.1002/cne.902940405.

Sandeman R E and Sandeman D C. 1987. “Serotonin-like immunoreactivity of giant olfactory interneurons in the crayfish brain.” Brain Research 403(2): 371-374.

Schmidt M and Ache B W. 1996. “Processing of antennular input in the brain of the spiny lobster, *Panulirus argus*: II. The olfactory pathway.” J Comp Physiol A 178 (5): 605–628.

Schmidt M and Ache B W. 1997. “Immunocytochemical analysis of glomerular regionalization and neuronal diversity in the olfactory deutocerebrum of the spiny lobster.” Cell Tissue Res 287 (3) (February 20): 541–563. doi:10.1007/s004410050778.

Schmidt M and Derby C D. 2005. „Non-olfactory chemoreceptors in asymmetric setae activate antennular grooming behavior in the caribbean spiny lobster *Panulirus argus*.” J Exp Biol, 208(2), pp.233–248.

Schmidt M, Chien H, Tadesse T, Johns M E and Derby C D. 2006. “Rosette-type tegumental glands associated with aesthetasc sensilla in the olfactory organ of the caribbean spiny lobster, *Panulirus argus*.” Cell Tissue Res 325 (2) (August): 369–395. doi:10.1007/s00441-006-0163-y.

Schubert M, Hansson B S and Sachse S. 2014. “The banana code – Natural blend processing in the olfactory circuitry of *Drosophila melanogaster*.” Frontiers in Physiology 59. doi:10.3389/fphys.2014.00059.

Schymura D, Forstner M, Schultze A, Kröber T, Swevers L, Iatrou K and Krieger J. 2010. “Antennal expression pattern of two olfactory receptors and an odorant binding protein implicated in host odor detection by the malaria vector *Anopheles gambiae*.” International Journal of Biological Sciences 6 (7) (January): 614–26.

Scott K, Brady R, Cravchik A, Morozov P, Rzhetsky A, Zuker C and Axel R. 2001. “A chemosensory gene family encoding candidate gustatory and olfactory receptors in Drosophila.” Cell 104 (5) (March 9): 661–73. doi: 10.1016/S0092-8674(01)00263-X

Shevchenko A, Sunyaev S, Loboda A, Bork P, Ens W and Standing K G. 2001. “Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching.” Analytical chemistry, 73(9), pp.1917–26. doi: 10.1021/ac0013709

Shevchenko A, Tomas H, Havlis J, Olsen J V and Mann M. 2006. “In-gel digestion for mass spectrometric characterization of proteins and proteomes.” Nature protocols, 1(6), pp.2856–60.

Silbering A F, Rytz R, Grosjean Y, Abuin L, Ramdya P, Jefferis G S X E and Benton R. 2011. “Complementary function and integrated wiring of the evolutionarily distinct Drosophila olfactory subsystems.” J Neurosci 31 (38) (September 21): 13357–13375. doi:10.1523/JNEUROSCI.2360-11.2011.

Simbeya C K, Csuzdi C E, Dew W A and Pyle G G. 2012. “Electroantennogram measurement of the olfactory response of *Daphnia spp.* and its impairment by waterborne copper.” Ecotoxicology and Environmental Safety 82: 80–84. doi:10.1016/j.ecoenv.2012.05.011.

Smadja C, Shi P, Butlin R K and Robertson H M. 2009. “Large gene family expansions and adaptive evolution for odorant and gustatory receptors in the pea aphid, *Acyrthosiphon pisum*.” Molecular Biology and Evolution 26 (9): 2073–2086. doi:10.1093/molbev/msp116.

Small M P and Thacker R W. 1994. “Land Hermit Crabs Use Odors of Dead Conspecifics to Locate Shells.” J Exp Mar Biol Ecol 182 (2): 169–182.

Soederhall, K., and Cerenius, L. 1998. “Role of the prophenoloxidase-activating system in invertebrate immunity.” Current Opinion in Immunology 10 (1) (February): 23–28. doi:10.1016/S0952-7915(98)80026-5.

Sperstad S V, Haug T, Blencke H M, Styrvold O B, Li C and Stensvag K. 2011. “Antimicrobial peptides from marine invertebrates: Challenges and perspectives in marine antimicrobial peptide discovery.” Biotechnology Advances 29 (5): 519–530. doi:10.1016/j.biotechadv.2011.05.021.

Sritunyalucksana K and Soderhall K. 2000. “The proPO and Clotting System in Crustaceans.” Aquaculture 191, pp. 53–69. doi: 10.1016/S0044-8486(00)00411-7

Steinbrecht R A. 1997. “Pore structures in insect olfactory sensilla: A review of data and concepts.” Int Journal Insect Morph Embryo 26 (3-4) (July): 229–245. doi:10.1016/S0020-7322(97)00024-X.

Steinbrecht R A 1998. “Odorant-binding proteins: Expression and function.” Annals of the New York Academy of Sciences 855 (1) (November): 323–332. doi:10.1111/j.1749-6632.1998.tb10591.x.

Steinbrecht R A and Stankiewicz B A. 1999. „Molecular composition of the wall of insect olfactory sensilla - the chitin question.” J Insect Physiol, 45(8), pp.785–790. doi: 10.1016/S0022-1910(99)00066-9

Stensmyr M C. 2003. “Novel natural ligands for Drosophila olfactory receptor neurons.” J Exp Biol 206 (4) (February 15): 715–724. doi:10.1242/jeb.00143.

Stensmyr M C, Erland S, Hallberg E, Wallen R, Greenaway P and Hansson B S. 2005. “Insect-like olfactory adaptations in the terrestrial giant robber crab.” Current Biology 15 (2): 116–121. doi:10.1016/j.cub.2004.12.069.

Stepanyan, R, B Hollins, S E Brock, and T S McClintock. 2004. “Primary culture of lobster (*Homarus americanus*) olfactory sensory neurons.” Chemical Senses 29 (3): 179–187. doi:10.1093/chemse/bjh023.

Steullet P, Cate H S, Michel W C and Derby C D. 2000. “Functional units of a compound nose: Aesthetasc sensilla house similar populations of olfactory receptor neurons on the crustacean antennule.” J Comp Neurology 418 (3): 270–280. doi: 10.1002/(SICI)1096-9861(20000313)418:3<270::AID-CNE3>3.0.CO;2-G

Stevenson J. 1961. “Polyphenol oxidase in the tegumental glands in the relation to the moulting cycle of the isopod crustacean *Armadillium vulgare*.” Biological Bulletin1, 121, p.554.

Sullivan J M and Beltz B S. 2001. “Development and connectivity of olfactory pathways in the brain of the lobster *Homarus americanus*.” J Comp Neurology 441 (1): 23–43.

Talbot P and Demers D. 1993. “Tegumental glands of crustacea.” In M. N. Horst & J. A. Freeman, eds. The Crustacean Integument. Morphology and Biochemistry. CRC Press, Boca Raton, Florida, pp. 151–191.

Talbot P, Al-Hajj H, Demers D and Howard D. 1991. ”Distribution of microfilaments and microtubules in Homarus pleopod tegumental glands (Crustacea, Decapoda)” Zoomorphology 110 pp.329–338.

Thacker R W. 1997. “Avoidance of recently eaten food by land hermit crabs, *Coenobita compressus*.” Animal Behaviour 55: 485–496.

Thomas J H and Robertson H M. 2008. “The Caenorhabditis chemoreceptor gene families.” BMC Biology 6. doi:42 10.1186/1741-7007-6-42.

Tomaschko K. 1995. “Autoradiographic and morphological investigations of the defensive ecdysteroid glands in adult *Pycnogonum litorale* (Arthropoda: Pantopoda)”. Eur J Entomol, 92, pp.105–112.

Tracey W D, Wilson R I, Laurent G and Benzer S. 2003. “Painless, a Drosophila gene essential for nociception.” Cell 113 (2) (April 18): 261–73. doi: 10.1016/S0092-8674(03)00272-1

Tsubouchi A, Caldwell J C and Tracey W D. 2012. “Dendritic Filopodia, Ripped Pocket, NOMPC, and NMDARs contribute to the sense of touch in Drosophila larvae.” Current Biology : CB 22 (22) (October 23): 2124–2134. doi:10.1016/j.cub.2012.09.019.

Vannini M and Ferretti J. 1997. “Chemoreception in two species of terrestrial hermit crabs (Decapoda: Coenobitidae).” J Crustacean Biol 17 (1) (February 1): 33–37. doi:10.2307/1549459.

Vieira F G and Rozas J. 2011. “Comparative genomics of the odorant-binding and chemosensory protein gene families across the arthropoda: Origin and evolutionary history of the chemosensory system.” Genome Biology and Evolution 3: 476–490. doi:10.1093/gbe/evr033.

Vosshall L B, Wong A M and Axel R. 2000. “An olfactory sensory map in the fly brain.” Cell 102 (2) (July 21): 147–59.

Weirich D and Ziegler A. 1997. „Uropod and lateral plate glands of the terrestrial isopod.” J Morphol 233(2): 183-193.

Yasuda-Kamatani Y and Yasuda A. 2006. “Characteristic expression patterns of allatostatin-like peptide, FMRFamide-related peptide, orcokinin, tachykinin-related peptide, and SIFamide in the olfactory system of crayfish *Procambarus clarkii*.” The Journal of Comparative Neurology 496 (1) (May 1): 135–47. doi:10.1002/cne.20903.

Yonge C. 1932. “On the nature and permeability of chitin. I. The chitin lining the foregut of decapod crustacea and the function of the tegumental glands.” Proc. Roy. Soc. Lond., B, 111(772), pp.298–329.

Zhou J J, Vieira F G, He X L, Smadja C, Liu R, Rozas J and Field L M. 2010. “Genome annotation and comparative analyses of the odorant-binding proteins and chemosensory proteins in the pea aphid *Acyrthosiphon pisum*.” Insect Molecular Biology, 19(Suppl. 2), 113–122. doi:10.1111/j.1365-2583.2009.00919.x

Zhou X, Slone J D, Rokas A, Berger S L, Liebig J, Ray A, Reinberg D and Zwiebel L J. 2012. “Phylogenetic and transcriptomic analysis of chemosensory receptors in a pair of divergent

ant species reveals sex-specific signatures of odor coding.” Edited by Nancy A. Moran. PLoS Genetics 8 (8) (January): e1002930. doi:10.1371/journal.pgen.1002930.

11. Declaration of independent assignment

I declare in accordance with the conferral of the degree of doctor from the School of Biology and Pharmacy of Friedrich-Schiller-University Jena that the submitted thesis was written only with the assistance and literature cited in the text.

People who assisted in experiments, data analysis and writing of the manuscripts are listed as co-authors of the respective manuscripts. I was not assisted by a consultant for doctorate theses.

The thesis has not been previously submitted whether to the Friedrich-Schiller-University, Jena or to any other university.

Jena,

_____ (Katrin Groh-Lunow)_____

Ort, Datum Unterschrift

12. Acknowledgements

This work would not have been possible without the support and help of many people I want to thank. To many of you I am deeply in debt and hope to be able to return the favor one day.

First my thanks go to Bill Hansson, director of the department of Evolutionary Neuroethology, for giving me the opportunity to work on this great and challenging project. You supported the projects progress in every possible way and build the department to be a family, making people feel welcome from their first day on. I am deeply grateful for being part of this family.

Many thanks also to my “Doktorvater” Ewald Grosse-Wilde. You guided me through the jungle of methods, techniques, politics and administrational issues; read and discussed countless versions of manuscripts, posters and presentations and taught me how to survive in the scientific world. I might not have met so many fascinating scientists and learned so much about the scientific spirit if it wasn`t for you.

Scientific collaborations are the spice of scientific life, so I want to thank Marcus Stensmyr for his input on the hermits work, Heiko Vogel and Steffi Gebauer-Jung for sharing their bioinformatic experience, Aleš Svatoš, Natalie Wielsch and Yvonne Hupfer for their expertise in proteomic sample preparation and evaluation and Steffen Harzsch, member of my PhD committee and expert on crustaceans, for lively discussions and ideas. Many thanks also to the other HAN group leaders for keeping my eyes open for views from differing ankles. Shannon – special thanks for improving my English language skills and “the word”. Thank you, Ian, for polishing the writing of my thesis.

The main contributors to everyday life happiness are the people one meets the most, so this one goes to my office members: Oksana, Alex and Christopher: I will miss the discussions, arts and shared cookie jars in our office!

Federica, when I needed it the most you helped me to find and keep my inner and outer balance with our weekly yoga. Many thanks for this and for introducing me to the “italian way of life”. I’m looking forward to meet you in your own ashram in the mountains.

It is rare to meet a person and feel instantly connected. I cannot even list every occasion when Christine made my day, but having you to share and discuss literally every thought and idea – scientific or not – is something I value highly. This includes gin-and-tonic at hotel bars to cheer up even the longest and most stressful conference day.

One big “thank You” I want to address to all former and present PhD students of the HAN department. You are the PhD seminars, journal clubs and last but not least the everyday coffee breaks that improved scientific thinking and exchange. Keep that up!

Regi, ein großes Dankeschön dafür, dass du die Zügel in der Molekularbiologie in der Hand hältst und für das reibungslose Arbeiten in die Bresche springst. Silke, Sylke, Sabine und Kerstin – danke für die vielen Male als etwas dringend Benötigtes noch ganz schnell organisiert wurde. Danke Sascha, für deine Unterstützung im Labor. Domenica möchte ich für zahllose Sequenzierungen danken sowie Jette für die Hilfe bei RNA Chips und Microarrays. Jerrit, danke für deine Hilfe bei der Herstellung der gekoppelten Peptide für die Antikörpersynthese. Danke auch an das Team des Gewächshauses, die immer zu Helfen bereit waren und an die Mitarbeiter der Bibliothek, deren Hilfe und Verbindungen die Recherche der Literatur entscheidend vereinfacht haben.

Ein ganz großes Dankeschön geht auch an die früheren und jetzigen Mitarbeiter der IT Abteilung. Ich wäre nicht sehr weit gekommen, wenn euer unermüdlicher Einsatz an den Computern, Scripten und Servern nicht alles am Laufen gehalten hätte. Danke auch für eure Geduld, wenn ich zum hundertsten Male mit einem „könnte bitte mal einer...“ in der Tür stand.

Swetlana, deine Hilfe und deinen Einsatz gebührend zu würdigen ist nicht einfach. Erst in deiner Abwesenheit wird deutlich, wie viel Organisation du uns eigentlich abnimmst und wie sehr wir es dir zu verdanken haben, wenn alles reibungslos läuft. Ein ganz großes Dankeschön dafür!

Besonderer Dank gilt auch meinen Eltern Ina und Andreas, die mich immer unterstützt haben und auf die ich immer zählen kann, egal was kommt.

Abschließend möchte ich noch einer ganz besonders wichtigen Person danken. Michael - danke für die Geduld und Nachsicht mit mir und meinen Launen. Danke auch für die vielen Male, an denen du mich für eine gewisse Zeit von der Arbeit losgerissen und mir die Möglichkeit gegeben hast wieder aufzutanken. Danke für deine Kraft mein Fels zu sein.

13. Curriculum vitae

Personal Data:

Name: Katrin Christine Groh-Lunow
(birth name: Katrin C. Groh)
Address: Am Heidehang 10
07973 Greiz-Obergrochlitz
Phone 03641 571460
Mobile 0171 1587119
Birthday: 21.03.1985 Greiz
married
German

Education

Jan. 2010 to date: PhD student at the Max-Planck-Institute of Chemical Ecology department Evolutionary Neuroethology in Jena, Thesis title: Analysis of antennal transcriptomes of hermit crabs comparing aquatic (*Pagurus bernhardus*) and terrestrial (*Coenobita clypeatus*) species

Oct. 2003 to Oct. 2009 Friedrich-Schiller-University Jena; studies in biology, Degree: Diploma, main subject: microbiology; minor subjects: genetics and medical microbiology

Diploma thesis: „Molekulare Ansätze zum Studium von Shrimp - (*Penaeus monodon*) - Pathogen - Interaktionen"; performed at the Hans Knöll Institute, department Cell- und Molecular Biology, Jena

Publications: Talks

Groh K.: The hermit crab's nose – antennal transcriptomics; IMPRS Evaluation Symposium 2013; MPI for Chemical Ecology, IMPRS, Jena, Germany, Apr 2013

Groh K.: The hermit crab's nose – antennal transcriptomics; 12th IMPRS Symposium / Max Planck Institute for Chemical Ecology, Jena, Germany, Apr 2013

Groh K.: Molecular insights into hermit crab olfaction; Chemical Ecology 2012 Mini-Symposium, Rothamsted, United Kingdom, Sep 2012

Groh K.: Molecular Insights into Hermit Crab Olfaction; Workshop „Molecular Tools to Study Olfaction in Arthropods" / Universität Konstanz, Konstanz, Germany, Mar 2012

Groh K.: How transition from water to land affects odor perception in hermit crabs; Chemical Ecology 2010 Mini-Symposium, Jena, Germany, Sep 2010

Publications: Posters

Groh K.*, Tuchina O., Talarico G., Koczan S., Rybak J., Stensmyr M.C., Grosse-Wilde E., Hansson B.S.: Olfaction in the hermit crab Coenobita clypeatus; 13th IMPRS Symposium / MPI for Chemical Ecology, Dornburg, Germany, Feb 2014

Groh K.*, Tuchina O., Talarico G., Koczan S., Rybak J., Stensmyr M.C., Grosse-Wilde E., Hansson B.S.: Olfaction in the hermit crab Coenobita clypeatus; 13th European Symposium for Insect Taste and Olfaction (ESITO), Villasimius, Italy, Sep 2013

Groh K.*, Grosse-Wilde E., Vogel H., Stensmyr M.C., Hansson B.S.: Comparative transcriptomics of arthropod antennae; 10th Göttingen Meeting of the German Neuroscience Society, Göttingen, Germany, Mar 2013

Groh K.: Olfaction in hermit crab Coenobita clypeatus; SAB Meeting 2012 / MPI for Chemical Ecology, Jena, Germany, Oct 2012

Groh K.*, Stensmyr M.C., Olsson S., Hansson B.S., Grosse-Wilde E.: Molecular insights into hermit crab olfaction; XVI International Symposium on Olfaction and Taste (ISOT), Stockholm, Sweden, Jun 2012

Tuchina O.*, Talarico G., Müller C., Groh K., Koczan S., Grosse-Wilde E., Hansson B.S.: Antennular morphology in Coenobita terrestrial hermit crabs; XVI International Symposium on Olfaction and Taste (ISOT), Stockholm, Sweden, Jun 2012

Groh K.: Molecular insights into hermit crab olfaction; 11th IMPRS Symposium / MPI for Chemical Ecology, Dornburg, Germany, Feb 2012

Groh K.*, Stensmyr M.C., Grosse-Wilde E., Hansson B.S.: The evolution of olfaction in hermit crabs; 12th European Symposium for Insect Taste and Olfaction (ESITO), St. Petersburg, Russian Federation, Sep 2011

Groh, K*.; Stensmyr, M.; Koczan, S.; Rybak, J.; Grosse-Wilde, E.; Hansson, B.: „The evolution of olfaction in hermit crabs", 7th Ecology and Behavior meeting (SERL), Rennes, France, May 2011

Groh K., Stensmyr M.C., Grosse-Wilde E., Hansson B.S.: The Evolution of Olfaction in Hermit Crabs; 9th Göttingen Meeting of the German Neuroscience Society, Göttingen, Germany, Mar 2011

Awards

SPP ISOT student travel grant (DFG Schwerpunktprogramm (SPP) 1392: Integrative Analysis of Olfaction) (2012)

PubMan

Tuchina, O., Groh, KC., Talarico, G., Mueller, CHG., Wielsch, N., Hupfer, Y., Svatoš, A.,Grosse-Wilde, E., Hansson, BS. (2014). Morphology and Histochemistry of the Aesthetasc-Associated Epidermal Glands in Terrestrial Hermit Crabs of the Genus *Coenobita* (Decapoda: Paguroidea). PLoS One, 9(5): e96430. doi: 10.1371/journal.pone.0096430

Groh, KC., Vogel, H., Stensmyr, MC., Grosse-Wilde, E., Hansson, BS. (2014). The hermit crab's nose – antennal transcriptomics. Frontiers in Evolutionary Psychology and Neuroscience. doi:10.3389/fnins.2013.00266.

Koch, SI., Groh, K., Vogel, H., Hansson, BS., Kleineidam, CJ., Grosse-Wilde, E. (2013). Caste-specific expression patterns of immune response and chemosensory related genes in the leaf-cutting ant, Atta vollenweideri. PLoS One, 8(11): e81518. doi:10.1371/journal.pone.0081518.

14. Appendix

Supplementary to manuscript 1: Supplementary Table S1

CClvl2

Index	GO.ID	Term	NoSeq Pber	Percent Pber	NoSeq Ccly	Percent Ccly	NoSeq Msex	Percent Msex
1	GO:0005623	cell	916	35.91	1218	32.65	1529	36.98
2	GO:0030054	cell junction	22	0.86	20	0.54	15	0.36
3	GO:0031012	extracellular matrix	3	0.12	14	0.38	3	0.07
4	GO:0005576	extracellular region	38	1.49	77	2.06	60	1.45
5	GO:0032991	macromolecular complex	453	17.76	641	17.18	686	16.59
6	GO:0016020	membrane	294	11.52	415	11.12	522	12.62
7	GO:0031974	membrane-enclosed lumen	169	6.62	340	9.11	290	7.01
8	GO:0043226	organelle	643	25.21	975	26.13	1003	24.26
9	GO:0045202	synapse	12	0.47	25	0.67	27	0.65
10	GO:0019012	virion	1	0.04	6	0.16	0	0

CClvl3

Index	GO.ID	Term	NoSeq Pber	Percent Pber	NoSeq Ccly	Percent Ccly	NoSeq Msex	Percent Msex
1	GO:0070161	anchoring junction	10	0.37	10	0.24	6	0.14
2	GO:0048046	apoplast	0	0.00	2	0.05	0	0.00
3	GO:0005911	cell-cell junction	7	0.26	7	0.17	7	0.16
4	GO:0044464	cell part	916	33.49	1218	29.62	1529	34.79
5	GO:0048475	coated membrane	5	0.18	15	0.36	13	0.30
6	GO:0044420	extracellular matrix part	1	0.04	6	0.15	2	0.05
7	GO:0044421	extracellular region part	18	0.66	53	1.29	11	0.25
8	GO:0043227	membrane-bounded organelle	472	17.26	816	19.84	814	18.52
9	GO:0044425	membrane part	204	7.46	275	6.69	353	8.03
10	GO:0043228	non-membrane-bounded organelle	331	12.10	433	10.53	443	10.08
11	GO:0044422	organelle part	426	15.58	683	16.61	651	14.81
12	GO:0019867	outer membrane	5	0.18	6	0.15	7	0.16
13	GO:0032992	protein-carbohydrate complex	0	0.00	2	0.05	0	0.00
14	GO:0032993	protein-DNA complex	5	0.18	6	0.15	6	0.14
15	GO:0043234	protein complex	296	10.82	494	12.01	497	11.31
16	GO:0044456	synapse part	7	0.26	15	0.36	19	0.43
17	GO:0031982	vesicle	31	1.13	66	1.61	31	0.71
18	GO:0044423	virion part	1	0.04	5	0.12	0	0.00
19	GO:0031594	neuromuscular junction	0	0.00	0	0.00	6	0.14

BP lvl2

Index	GO.ID	Term	SeqsPber	PercentPber	SeqsCcly	PercentCcly
1	GO:0022610	biological adhesion	24	0.64	37	0.6
2	GO:0065007	biological regulation	340	9.08	611	9.89
3	GO:0015976	carbon utilization	1	0.03	13	0.21
4	GO:0008283	cell proliferation	41	1.09	84	1.36
5	GO:0071840	cellular component organization or biogenesis	261	6.97	474	7.67
6	GO:0009987	cellular process	866	23.12	1185	19.19
7	GO:0016265	death	78	2.08	138	2.23
8	GO:0032502	developmental process	231	6.17	410	6.64
9	GO:0040007	growth	31	0.83	92	1.49
10	GO:0002376	immune system process	30	0.8	78	1.26
11	GO:0051179	localization	238	6.35	384	6.22
12	GO:0040011	locomotion	51	1.36	75	1.21
13	GO:0008152	metabolic process	783	20.9	1094	17.71
14	GO:0051704	multi-organism process	26	0.69	77	1.25
15	GO:0032501	multicellular organismal process	243	6.49	431	6.98
16	GO:0043473	pigmentation	2	0.05	14	0.23
17	GO:0000003	reproduction	77	2.06	166	2.69
18	GO:0050896	response to stimulus	247	6.59	451	7.3
19	GO:0048511	rhythmic process	0	0	13	0.21
20	GO:0023052	signaling	128	3.42	251	4.06
21	GO:0016032	viral reproduction	48	1.28	98	1.59

BP lvl3

Index	GO.ID	Term	Seqs Pber	Percent Pber	Seqs Ccly	Percent Ccly
1	GO:0030029	actin filament-based process	38	0.6	33	0.3
2	GO:0030534	adult behavior	0	0	11	0.1
3	GO:0007568	aging	19	0.3	47	0.42
4	GO:0048856	anatomical structure development	188	2.95	330	2.96
5	GO:0007610	behavior	25	0.39	45	0.4
6	GO:0009058	biosynthetic process	341	5.35	558	5.01
7	GO:0009056	catabolic process	171	2.68	305	2.74
8	GO:0007267	cell-cell signaling	25	0.39	53	0.48
9	GO:0001775	cell activation	16	0.25	22	0.2
10	GO:0007155	cell adhesion	24	0.38	37	0.33
11	GO:0007154	cell communication	51	0.8	95	0.85
12	GO:0007049	cell cycle	90	1.41	197	1.77
13	GO:0008219	cell death	77	1.21	138	1.24
14	GO:0051301	cell division	31	0.49	38	0.34
15	GO:0016049	cell growth	1	0.02	30	0.27
16	GO:0044085	cellular component biogenesis	116	1.82	258	2.32
17	GO:0006928	cellular component movement	30	0.47	41	0.37
18	GO:0016043	cellular component organization	243	3.81	425	3.82
19	GO:0071841	cellular component organization or biogenesis at cellular level	237	3.72	411	3.69
20	GO:0048869	cellular developmental process	118	1.85	194	1.74
21	GO:0019725	cellular homeostasis	38	0.6	54	0.48
22	GO:0051641	cellular localization	102	1.6	173	1.55
23	GO:0016044	cellular membrane organization	41	0.64	79	0.71
24	GO:0044237	cellular metabolic process	657	10.31	972	8.73
25	GO:0051716	cellular response to stimulus	146	2.29	299	2.68
26	GO:0007059	chromosome segregation	2	0.03	24	0.22
27	GO:0050817	coagulation	14	0.22	17	0.15
28	GO:0048589	developmental growth	11	0.17	21	0.19
29	GO:0021700	developmental maturation	12	0.19	18	0.16

30	GO:0051234	establishment of localization	216	3.39	353	3.17
31	GO:0007163	establishment or maintenance of cell polarity	20	0.31	22	0.2
32	GO:0002252	immune effector process	0	0	12	0.11
33	GO:0006955	immune response	19	0.3	42	0.38
34	GO:0044419	interspecies interaction between organisms	6	0.09	33	0.3
35	GO:0045321	leukocyte activation	0	0	16	0.14
36	GO:0051674	localization of cell	28	0.44	33	0.3
37	GO:0033036	macromolecule localization	90	1.41	153	1.37
38	GO:0043170	macromolecule metabolic process	491	7.71	719	6.46
39	GO:0032259	methylation	12	0.19	40	0.36
40	GO:0007017	microtubule-based process	49	0.77	78	0.7
41	GO:0035264	multicellular organism growth	1	0.02	13	0.12
42	GO:0032504	multicellular organism reproduction	37	0.58	83	0.75
43	GO:0007275	multicellular organismal development	206	3.23	374	3.36
44	GO:0035637	multicellular organismal signaling	25	0.39	39	0.35
45	GO:0006807	nitrogen compound metabolic process	328	5.15	622	5.58
46	GO:0071704	organic substance metabolic process	0	0	15	0.13
47	GO:0019637	organophosphate metabolic process	0	0	23	0.21
48	GO:0055114	oxidation-reduction process	130	2.04	192	1.72
49	GO:0042440	pigment metabolic process	0	0	14	0.13
50	GO:0044238	primary metabolic process	613	9.62	894	8.03
51	GO:0050789	regulation of biological process	302	4.74	576	5.17
52	GO:0065008	regulation of biological quality	108	1.7	200	1.8
53	GO:0050878	regulation of body fluid levels	14	0.22	20	0.18
54	GO:0065009	regulation of molecular function	64	1	133	1.19
55	GO:0022414	reproductive process	73	1.15	153	1.37
56	GO:0009628	response to abiotic stimulus	30	0.47	40	0.36
57	GO:0009607	response to biotic stimulus	19	0.3	44	0.4
58	GO:0042221	response to chemical stimulus	95	1.49	209	1.88
59	GO:0009719	response to endogenous stimulus	23	0.36	52	0.47
60	GO:0009605	response to external stimulus	44	0.69	60	0.54
61	GO:0051707	response to other organism	0	0	33	0.3
62	GO:0006950	response to stress	128	2.01	229	2.06
63	GO:0019953	sexual reproduction	33	0.52	75	0.67
64	GO:0044281	small molecule metabolic process	198	3.11	436	3.91
65	GO:0003008	system process	49	0.77	85	0.76
66	GO:0042330	taxis	24	0.38	34	0.31
67	GO:0016271	tissue death	0	0	11	0.1
68	GO:0022415	viral reproductive process	32	0.5	52	0.47
69	GO:0009605	response to external stimulus	43.00	0.60	57.00	0.47
70	GO:0051707	response to other organism	0.00	0.00	32.00	0.27
71	GO:0006950	response to stress	134.00	1.88	209.00	1.73
72	GO:0022613	ribonucleoprotein complex biogenesis	134.00	1.88	53.00	0.44
73	GO:0019953	sexual reproduction	33.00	0.46	71.00	0.59
74	GO:0044281	small molecule metabolic process	287.00	4.03	297.00	2.46
75	GO:0003008	system process	54.00	0.76	78.00	0.65
76	GO:0042330	taxis	23.00	0.32	33.00	0.27
77	GO:0016271	tissue death	0.00	0.00	11.00	0.09
78	GO:0016032	viral reproduction	48.00	0.67	92.00	0.76
79	GO:0030534	adult behavior	0.00	0.00	0.00	0.00
80	GO:0051705	behavioral interaction between organisms	0.00	0.00	0.00	0.00
81	GO:0071554	cell wall organization or biogenesis	0.00	0.00	0.00	0.00
82	GO:0019748	secondary metabolic process	0.00	0.00	0.00	0.00

MF lvl2

Index	GO.ID	Term	SeqsPber	PercentPber	SeqsCcly	PercentCcly
1	GO:0016209	antioxidant activity	18	0.96	22	0.91
2	GO:0005488	binding	760	40.6	1050	43.66
3	GO:0003824	catalytic activity	676	36.11	860	35.76
4	GO:0016247	channel regulator activity	0	0	3	0.12
5	GO:0042056	chemoattractant activity	0	0	1	0.04
6	GO:0009055	electron carrier activity	23	1.23	39	1.62
7	GO:0030234	enzyme regulator activity	43	2.3	73	3.04
8	GO:0060089	molecular transducer activity	41	2.19	44	1.83
9	GO:0016015	morphogen activity	0	0	1	0.04
10	GO:0001071	nucleic acid binding transcription factor activity	13	0.69	26	1.08
11	GO:0000988	protein binding transcription factor activity	11	0.59	32	1.33
12	GO:0004872	receptor activity	32	1.71	33	1.37
13	GO:0030545	receptor regulator activity	2	0.11	1	0.04
14	GO:0005198	structural molecule activity	153	8.17	105	4.37
15	GO:0045182	translation regulator activity	2	0.11	4	0.17
16	GO:0005215	transporter activity	98	5.24	111	4.62

MF lvl3

Index	GO.ID	Term	Seqs Pber	Percent Pber	Seqs Ccly	Percent Ccly
1	GO:0043176	amine binding	4	0.2	13	0.46
2	GO:0003823	antigen binding	2	0.1	0	0
3	GO:0060090	binding, bridging	4	0.2	8	0.28
4	GO:0030246	carbohydrate binding	13	0.65	17	0.6
5	GO:0031406	carboxylic acid binding	11	0.55	23	0.81
6	GO:0043028	caspase regulator activity	2	0.1	6	0.21
7	GO:0043498	cell surface binding	4	0.2	3	0.11
8	GO:0016248	channel inhibitor activity	0	0	2	0.07
9	GO:0003682	chromatin binding	6	0.3	25	0.88
10	GO:0048037	cofactor binding	24	1.2	55	1.93
11	GO:0051184	cofactor transporter activity	0	0	2	0.07
12	GO:0009975	cyclase activity	1	0.05	5	0.18
13	GO:0019239	deaminase activity	2	0.1	6	0.21
14	GO:0030337	DNA polymerase processivity factor activity	1	0.05	3	0.11
15	GO:0008144	drug binding	3	0.15	5	0.18
16	GO:0045155	electron transporter, transferring electrons from CoQH2-cytochrome c reductase complex and cytochrome c oxidase complex activity	1	0.05	1	0.04
17	GO:0008047	enzyme activator activity	6	0.3	19	0.67
18	GO:0004857	enzyme inhibitor activity	16	0.8	24	0.84
19	GO:0050840	extracellular matrix binding	0	0	1	0.04
20	GO:0005201	extracellular matrix structural constituent	0	0	1	0.04
21	GO:0042562	hormone binding	2	0.1	2	0.07
22	GO:0016787	hydrolase activity	307	15.33	358	12.56
23	GO:0043167	ion binding	152	7.59	225	7.89
24	GO:0016853	isomerase activity	34	1.7	35	1.23
25	GO:0019207	kinase regulator activity	3	0.15	8	0.28
26	GO:0016874	ligase activity	39	1.95	60	2.11
27	GO:0008289	lipid binding	21	1.05	23	0.81
28	GO:0001530	lipopolysaccharide binding	0	0	2	0.07
29	GO:0016829	lyase activity	26	1.3	34	1.19
30	GO:0051540	metal cluster binding	8	0.4	16	0.56
31	GO:0048270	methionine adenosyltransferase regulator activity	0	0	1	0.04
32	GO:0042165	neurotransmitter binding	0	0	1	0.04
33	GO:0030235	nitric-oxide synthase regulator activity	4	0.2	2	0.07

34	GO:0003676	nucleic acid binding	223	11.13	290	10.18
35	GO:0060589	nucleoside-triphosphatase regulator activity	7	0.35	20	0.7
36	GO:0042979	ornithine decarboxylase regulator activity	1	0.05	1	0.04
37	GO:0016491	oxidoreductase activity	148	7.39	216	7.58
38	GO:0019825	oxygen binding	1	0.05	1	0.04
39	GO:0001871	pattern binding	8	0.4	8	0.28
40	GO:0061134	peptidase regulator activity	9	0.45	13	0.46
41	GO:0042277	peptide binding	3	0.15	19	0.67
42	GO:0004601	peroxidase activity	14	0.7	16	0.56
43	GO:0019208	phosphatase regulator activity	11	0.55	6	0.21
44	GO:0031409	pigment binding	1	0.05	5	0.18
45	GO:0005515	protein binding	327	16.33	607	21.3
46	GO:0032947	protein complex scaffold	2	0.1	4	0.14
47	GO:0071814	protein-lipid complex binding	4	0.2	1	0.04
48	GO:0030546	receptor activator activity	2	0.1	0	0
49	GO:0030547	receptor inhibitor activity	0	0	1	0.04
50	GO:0043021	ribonucleoprotein binding	12	0.6	19	0.67
51	GO:0000386	second spliceosomal transesterification activity	1	0.05	2	0.07
52	GO:0008430	selenium binding	1	0.05	4	0.14
53	GO:0003700	sequence-specific DNA binding transcription factor activity	13	0.65	26	0.91
54	GO:0004871	signal transducer activity	41	2.05	44	1.54
55	GO:0008641	small protein activating enzyme activity	0	0	5	0.18
56	GO:0017080	sodium channel regulator activity	0	0	2	0.07
57	GO:0042302	structural constituent of cuticle	8	0.4	4	0.14
58	GO:0005200	structural constituent of cytoskeleton	6	0.3	10	0.35
59	GO:0008307	structural constituent of muscle	8	0.4	0	0
60	GO:0017056	structural constituent of nuclear pore	1	0.05	0	0
61	GO:0016490	structural constituent of peritrophic membrane	1	0.05	0	0
62	GO:0003735	structural constituent of ribosome	111	5.54	70	2.46
63	GO:0022892	substrate-specific transporter activity	83	4.14	93	3.26
64	GO:0004784	superoxide dismutase activity	3	0.15	5	0.18
65	GO:0046906	tetrapyrrole binding	13	0.65	18	0.63
66	GO:0000989	transcription factor binding transcription factor activity	11	0.55	32	1.12
67	GO:0016740	transferase activity	153	7.64	233	8.18
68	GO:0090079	translation regulator activity, nucleic acid binding	1	0.05	2	0.07
69	GO:0022857	transmembrane transporter activity	77	3.84	85	2.98
70	GO:0004803	transposase activity	0	0	1	0.04
71	GO:0046790	virion binding	1	0.05	1	0.04
72	GO:0003735	structural constituent of ribosome	111.00	4.50	18.00	0.61
73	GO:0022892	substrate-specific transporter activity	83.00	3.37	10.00	0.34
74	GO:0032542	sulfiredoxin activity	0.00	0.00	0.00	0.00
75	GO:0004784	superoxide dismutase activity	3.00	0.12	0.00	0.00
76	GO:0046906	tetrapyrrole binding	13.00	0.53	2.00	0.07
77	GO:0004791	thioredoxin-disulfide reductase activity	0.00	0.00	0.00	0.00
78	GO:0016563	transcription activator activity	1.00	0.04	0.00	0.00
79	GO:0003712	transcription cofactor activity	11.00	0.45	6.00	0.20
80	GO:0003711	transcription elongation regulator activity	1.00	0.04	0.00	0.00
81	GO:0000989	transcription factor binding transcription factor activity	0.00	0.00	0.00	0.00
82	GO:0016986	transcription initiation factor activity	1.00	0.04	0.00	0.00
83	GO:0016564	transcription repressor activity	0.00	0.00	1.00	0.03
84	GO:0016740	transferase activity	154.00	6.25	43.00	1.46
85	GO:0090079	translation regulator activity, nucleic acid binding	1.00	0.04	0.00	0.00
86	GO:0030371	translation repressor activity	0.00	0.00	0.00	0.00
87	GO:0022857	transmembrane transporter activity	77.00	3.12	10.00	0.34
88	GO:0046790	virion binding	1.00	0.04	0.00	0.00
89	GO:0019842	vitamin binding	6.00	0.24	2.00	0.07

Supplementary to manuscript 2: Supplementary Table S1

CC lvl2

GO ID	Term	#Seqs	%
GO:0005623	cell	5417	44.26
GO:0043226	organelle	3599	29.41
GO:0032991	macromolecular complex	1966	16.06
GO:0031974	membrane-enclosed lumen	899	7.35
GO:0005576	extracellular region	230	1.88
GO:0045202	synapse	128	1.05

CC lvl3

GO ID	Term	#Seqs	%
GO:0044464	cell part	5417	38.93
GO:0043227	membrane-bounded organelle	3097	22.26
GO:0044422	organelle part	2175	15.63
GO:0043234	protein complex	1523	10.95
GO:0043228	non-membrane-bounded organelle	1186	8.52
GO:0031982	vesicle	212	1.52
GO:0044456	synapse part	94	0.68
GO:0048046	apoplast	92	0.66
GO:0044421	extracellular region part	85	0.61
GO:0031594	neuromuscular junction	18	0.13
GO:0032993	protein-DNA complex	16	0.11

BP lvl2

GO ID	Term	#Seqs	%
GO:0009987	cellular process	5131	23.68
GO:0008152	metabolic process	4231	19.52
GO:0065007	biological regulation	2373	10.95
GO:0051179	localization	1532	7.07
GO:0032501	multicellular organismal process	1315	6.07
GO:0016043	cellular component organization	1163	5.37
GO:0032502	developmental process	1152	5.32
GO:0023052	signaling	1111	5.13
GO:0050896	response to stimulus	1074	4.96
GO:0044085	cellular component biogenesis	688	3.17
GO:0000003	reproduction	336	1.55
GO:0016265	death	274	1.26
GO:0051704	multi-organism process	207	0.96
GO:0040007	growth	197	0.91
GO:0002376	immune system process	188	0.87

GO:0008283	cell proliferation	183	0.84
GO:0040011	locomotion	178	0.82
GO:0022610	biological adhesion	149	0.69
GO:0015976	carbon utilization	73	0.34
GO:0048511	rhythmic process	42	0.19
GO:0043473	pigmentation	35	0.16
GO:0071554	cell wall organization or biogenesis	26	0.12
GO:0016032	viral reproduction	14	0.06

BP lvl3

GO ID	Term	#Seqs	%
GO:0044237	cellular metabolic process	3686	10.10
GO:0044238	primary metabolic process	3425	9.38
GO:0043170	macromolecule metabolic process	2582	7.07
GO:0009058	biosynthetic process	2189	6.00
GO:0006807	nitrogen compound metabolic process	2166	5.93
GO:0050789	regulation of biological process	1901	5.21
GO:0051234	establishment of localization	1395	3.82
GO:0044281	small molecule metabolic process	1321	3.62
GO:0007275	multicellular organismal development	1035	2.83
GO:0048856	anatomical structure development	950	2.60
GO:0065008	regulation of biological quality	914	2.50
GO:0023046	signaling process	778	2.13
GO:0006996	organelle organization	656	1.80
GO:0023033	signaling pathway	651	1.78
GO:0048869	cellular developmental process	630	1.73
GO:0006950	response to stress	610	1.67
GO:0009056	catabolic process	594	1.63
GO:0042221	response to chemical stimulus	489	1.34
GO:0051641	cellular localization	465	1.27
GO:0055114	oxidation reduction	464	1.27
GO:0022607	cellular component assembly	440	1.21
GO:0033036	macromolecule localization	432	1.18
GO:0042445	hormone metabolic process	428	1.17
GO:0051716	cellular response to stimulus	423	1.16
GO:0007154	cell communication	371	1.02
GO:0007049	cell cycle	362	0.99
GO:0003008	system process	335	0.92
GO:0022414	reproductive process	327	0.90
GO:0043933	macromolecular complex subunit organization	307	0.84
GO:0065009	regulation of molecular function	279	0.76
GO:0008219	cell death	273	0.75
GO:0022613	ribonucleoprotein complex biogenesis	271	0.74
GO:0032504	multicellular organism reproduction	262	0.72
GO:0009628	response to abiotic stimulus	242	0.66

GO:0030030	cell projection organization	240	0.66
GO:0019953	sexual reproduction	235	0.64
GO:0006928	cellular component movement	229	0.63
GO:0007610	behavior	223	0.61
GO:0034621	cellular macromolecular complex subunit organization	222	0.61
GO:0070271	protein complex biogenesis	217	0.59
GO:0007017	microtubule-based process	210	0.58
GO:0007267	cell-cell signaling	207	0.57
GO:0016044	cellular membrane organization	188	0.51
GO:0061024	membrane organization	188	0.51
GO:0030029	actin filament-based process	151	0.41
GO:0007155	cell adhesion	149	0.41
GO:0009607	response to biotic stimulus	147	0.40
GO:0051674	localization of cell	136	0.37
GO:0019725	cellular homeostasis	136	0.37
GO:0051707	response to other organism	123	0.34
GO:0006955	immune response	123	0.34
GO:0009605	response to external stimulus	110	0.30
GO:0048589	developmental growth	101	0.28
GO:0051301	cell division	100	0.27
GO:0009719	response to endogenous stimulus	97	0.27
GO:0019637	organophosphate metabolic process	88	0.24
GO:0043062	extracellular structure organization	81	0.22
GO:0007568	aging	69	0.19
GO:0021700	developmental maturation	60	0.16
GO:0071103	DNA conformation change	55	0.15
GO:0030534	adult behavior	55	0.15
GO:0071704	organic substance metabolic process	53	0.15
GO:0008037	cell recognition	53	0.15
GO:0007059	chromosome segregation	51	0.14
GO:0042440	pigment metabolic process	48	0.13
GO:0045321	leukocyte activation	45	0.12
GO:0001775	cell activation	45	0.12
GO:0051705	behavioral interaction between organisms	44	0.12
GO:0019748	secondary metabolic process	44	0.12
GO:0006323	DNA packaging	43	0.12
GO:0022411	cellular component disassembly	40	0.11
GO:0044419	interspecies interaction between organisms	35	0.10
GO:0035264	multicellular organism growth	33	0.09
GO:0034330	cell junction organization	33	0.09
GO:0007623	circadian rhythm	33	0.09
GO:0051606	detection of stimulus	33	0.09
GO:0042303	molting cycle	31	0.08
GO:0042330	taxis	23	0.06
GO:0048871	multicellular organismal homeostasis	22	0.06

GO:0048066	developmental pigmentation	20	0.05
GO:0007028	cytoplasm organization	20	0.05
GO:0001816	cytokine production	18	0.05
GO:0007622	rhythmic behavior	18	0.05
GO:0016271	tissue death	18	0.05
GO:0090130	tissue migration	17	0.05
GO:0070882	cellular cell wall organization or biogenesis	16	0.04
GO:0035265	organ growth	16	0.04
GO:0050673	epithelial cell proliferation	15	0.04
GO:0043954	cellular component maintenance	14	0.04
GO:0022415	viral reproductive process	14	0.04
GO:0042546	cell wall biogenesis	13	0.04
GO:0022406	membrane docking	11	0.03

MF lvl2

GO ID	Term	#Seqs	%
GO:0005488	binding	5083	45.78
GO:0003824	catalytic activity	3960	35.67
GO:0005215	transporter activity	651	5.86
GO:0060089	molecular transducer activity	382	3.44
GO:0030528	transcription regulator activity	366	3.30
GO:0005198	structural molecule activity	324	2.92
GO:0030234	enzyme regulator activity	199	1.79
GO:0009055	electron carrier activity	75	0.68
GO:0016209	antioxidant activity	63	0.57

MF lvl3

GO ID	Term	#Seqs	%
GO:0005515	protein binding	2145	15.44
GO:0016787	hydrolase activity	1518	10.93
GO:0000166	nucleotide binding	1394	10.03
GO:0043167	ion binding	1179	8.49
GO:0003676	nucleic acid binding	1178	8.48
GO:0016740	transferase activity	1176	8.46
GO:0001882	nucleoside binding	819	5.89
GO:0016491	oxidoreductase activity	671	4.83
GO:0022892	substrate-specific transporter activity	506	3.64
GO:0022857	transmembrane transporter activity	502	3.61
GO:0004871	signal transducer activity	382	2.75
GO:0016874	ligase activity	273	1.96
GO:0048037	cofactor binding	261	1.88
GO:0003735	structural constituent of ribosome	198	1.43
GO:0003700	transcription factor activity	169	1.22

GO:0016829	lyase activity	159	1.14
GO:0016853	isomerase activity	121	0.87
GO:0060589	nucleoside-triphosphatase regulator activity	111	0.80
GO:0046906	tetrapyrrole binding	96	0.69
GO:0016563	transcription activator activity	83	0.60
GO:0003702	RNA polymerase II transcription factor activity	79	0.57
GO:0019842	vitamin binding	77	0.55
GO:0003712	transcription cofactor activity	76	0.55
GO:0008289	lipid binding	73	0.53
GO:0003682	chromatin binding	67	0.48
GO:0051540	metal cluster binding	66	0.48
GO:0008047	enzyme activator activity	63	0.45
GO:0016564	transcription repressor activity	63	0.45
GO:0030246	carbohydrate binding	51	0.37
GO:0004601	peroxidase activity	48	0.35
GO:0031406	carboxylic acid binding	37	0.27
GO:0004857	enzyme inhibitor activity	37	0.27
GO:0005200	structural constituent of cytoskeleton	28	0.20
GO:0042277	peptide binding	28	0.20
GO:0019207	kinase regulator activity	24	0.17
GO:0008307	structural constituent of muscle	23	0.17
GO:0043176	amine binding	19	0.14
GO:0043021	ribonucleoprotein binding	16	0.12
GO:0008144	drug binding	16	0.12
GO:0001871	pattern binding	14	0.10
GO:0003711	transcription elongation regulator activity	13	0.09
GO:0019208	phosphatase regulator activity	12	0.09
GO:0051184	cofactor transporter activity	12	0.09
GO:0009975	cyclase activity	11	0.08

Supplementary to manuscript 2: Supplementary Table S2

Seq. Name	Seq. Description	Seq. Length	#Hits	min. eValue	mean Similarity
Contig_9514	tachykinin-like peptides receptor 99d-like	2165	20	1.26E-139	74.55%
Contig_70368	rac serine threonine-protein kinase-like isoform 1	398	20	1.97E-58	79.15%
Contig_66414	inositol -trisphosphate receptor type 1 isoform 2	402	20	7.00E-47	76.60%
Contig_63563	guanine nucleotide-binding protein subunit alpha-11	256	20	3.36E-21	67.35%
Contig_63173	breast cancer anti-estrogen resistance protein 1-like	716	20	9.40E-71	63.30%
Contig_62858	frizzled-8	312	20	4.56E-43	81.70%
Contig_60748	inositol -trisphosphate receptor type 1-like	348	20	1.87E-35	69.35%
Contig_59969	metabotropic glutamate receptor	296	20	1.49E-07	76.85%
Contig_59278	calcitonin receptor	336	20	4.38E-44	67.45%
Contig_58014	5-hydroxytryptamine receptor 2a	1032	20	1.39E-53	88.00%
Contig_57812	5-hydroxytryptamine receptor 2b	254	20	1.19E-36	79.55%
Contig_56143	glutathione synthetase	341	20	1.70E-30	71.25%
Contig_54444	octopamine receptor beta-3r-like	248	20	5.96E-26	80.00%
Contig_5148	guanine nucleotide-binding protein subunit gamma-e-like	850	20	6.70E-35	94.40%
Contig_49713	leucine-rich transmembrane protein	253	20	2.34E-15	61.15%
Contig_4913	sex peptide receptor	549	20	1.32E-15	79.20%
Contig_4648	g protein alpha subunit	4065	20	2.56E-97	98.05%
Contig_44596	histamine h1 receptor	329	20	7.09E-28	67.00%
Contig_42665	guanine nucleotide-binding protein subunit beta-5-like	324	20	1.33E-43	87.15%
Contig_41748	guanine nucleotide binding protein (g protein) alpha 14	370	20	1.96E-26	69.70%
Contig_4013	regulator of g protein signaling	1086	20	6.03E-170	77.25%
Contig_39270	regulator of g-protein signaling 7-like	1125	20	2.57E-134	87.40%
Contig_38998	latrophilin cirl-like isoform 1	399	20	9.76E-15	71.05%
Contig_29010	protocadherin-like wing polarity protein stan-like	341	20	6.18E-40	75.90%
Contig_28204	g-protein coupled receptor grl101-like	293	20	1.39E-23	68.10%
Contig_2709	gtp-binding protein alpha gna	3639	20	2.65E-100	98.45%
Contig_2366	guanine nucleotide-binding protein g subunit alpha-like	1799	20	0	91.50%
Contig_20793	rho guanine nucleotide exchange factor 12	1056	20	1.14E-104	68.40%
Contig_20191	guanine nucleotide-binding protein g subunit alpha-like	3320	20	0	92.40%
Contig_19255	regulator of g protein signaling	272	19	4.07E-11	77.68%
Contig_17240	regulator of g-protein signaling 2	419	20	7.28E-23	73.85%
Contig_17228	heterotrimeric gtp-binding protein alpha subunit g-alpha-q	847	20	9.35E-98	94.10%
Contig_16804	tyramine receptor	2374	20	0	69.75%
Contig_13983	guanine nucleotide-binding protein subunit gamma-1-like	2033	20	6.45E-23	85.60%
Contig_13313	rho guanine nucleotide exchange factor 12	526	20	5.53E-21	71.75%

Contig_11124	guanine nucleotide-binding protein g subunit alpha	720	20	2.06E-142	94.80%

Supplementary to manuscript 2: Supplementary Table S3

BP lvl2

		Groh et al. 2014		new data		
GO.ID	Term	SeqsCcly	PercentCcly	SeqsCcly	PercentCcly	Difference
GO:0022610	biological adhesion	37	0.60	149	0.69	0.09
GO:0065007	biological regulation	611	9.89	2373	10.95	1.06
GO:0015976	carbon utilization	13	0.21	73	0.34	0.13
GO:0008283	cell proliferation	84	1.36	183	0.84	0.52
GO:0071840	cellular component organization or biogenesis	474	7.67	1851	8.54	0.87
GO:0009987	cellular process	1185	19.19	5131	23.68	4.49
GO:0016265	death	138	2.23	274	1.26	0.97
GO:0032502	developmental process	410	6.64	1152	5.32	1.32
GO:0040007	growth	92	1.49	197	0.91	0.58
GO:0002376	immune system process	78	1.26	188	0.87	0.39
GO:0051179	localization	384	6.22	1532	7.07	0.85
GO:0040011	locomotion	75	1.21	178	0.82	0.39
GO:0008152	metabolic process	1094	17.71	4231	19.52	1.81
GO:0051704	multi-organism process	77	1.25	207	0.96	0.29
GO:0032501	multicellular organismal process	431	6.98	1315	6.07	0.91
GO:0043473	pigmentation	14	0.23	35	0.16	0.07
GO:0000003	reproduction	166	2.69	336	1.55	1.14
GO:0050896	response to stimulus	451	7.30	1074	4.96	2.34
GO:0048511	rhythmic process	13	0.21	42	0.19	0.02
GO:0023052	signaling	251	4.06	1111	5.13	1.07
GO:0016032	viral reproduction	98	1.59	14	0.06	1.53

Supplementary to manuscript 3: Additional file 1_1D_Glandular tissue

Spot number	Mascot searching					MS BLAST searching					MSE analysis				
	Protein name	Accession	*Species*	Score	Peptide hits	Protein name	Accession	*Species*	Score	Peptide hits	Protein name	Accession	*Species*	Score	Peptide hits
1A	Histone 2A	XP_002170269	*Hydra magnipapillata*	162	4	Histone 2A	XP_003702149	*Megachile rotundata*	175	3	Polyubiquitin	P0CG69	*Drosophila melanogaster*	48814	28
	ConsensusfromContig80707_full_rev_frame_1		*Coenobita*	79	1	ConsensusfromContig78959_full_rev_frame_0			115	2	histone H2A	XP_003702149.1_	*Megachile rotundata*	8140	4
	ConsensusfromContig28270_full_fwd_frame_0		*Coenobita*	77	1	RND efflux system, NodT	ABM18429	*Marinobacter aquaeolei VT8*	110	2	40S ribosomal protein	EEB18567 1	*Pediculus humanus corporis*	2788	3
	ConsensusfromContig115053_full_rev_frame_0		*Coenobita*	73	1	ConsensusfromContig42378_full_rev_frame_1			84	1	ConsensusfromContig80707_full_rev_frame_1		*Coenobita*	3580	2
						ConsensusfromContig115053_full_rev_frame_0			79	1	histone H4	GF20391	*Drosophila ananassae*	3498	4
						CDGSH-type zinc finger	ACG76244	*Amblyomma americanum*	73	1	ConsensusfromContig42378 full rev frame 1		*Coenobita*	1794	2
						coagulation factor XI	XP_001845410	*Culex quinquefasciatus*	72	1	ConsensusfromContig69829 full rev frame 2		*Coenobita*	1222	3
						ubiquitin (ConsensusfromContig81144_full_rev_frame_0; ConsensusfromContig27852_full_rev_frame_0)	AAB00498	*Homarus americanus*	71	1	ConsensusfromContig27957 full rev frame 0		*Coenobita*	1280	2
						ribosomal protein S28	ABO47825	*Heliconius melpomene*	71	1	ConsensusfromContig46924 full rev frame 0		*Coenobita*	1060	2
											ConsensusfromContig33937 full rev frame 1		*Coenobita*	1040	2

						ConsensusfromContig78959 full rev frame 0							*Coenobita*	806	1
2A	histone H2A.2.2	XP_002169186	*Hydra magnipapillata*	136	3	ConsensusfromContig78959_full_rev_frame_0 (cuticle protein)			249	4	Histone 2A	XP_003702149	*Megachile rotundata*	11447	6
	ConsensusfromContig8357_full_fwd_frame_0			72	2	ConsensusfromContig8357_full_fwd_frame_0			234	3	Histone H4	Q43083	*Pyrenomonas salina*	3490	5
	ConsensusfromContig27957_full_rev_frame_0			65	1	ConsensusfromContig46299_full_rev_frame_2			205	4	Sarcoplasmic calcium binding protein	P86909.1	*Chionoecetes opilio*	7310	1
	ConsensusfromContig78959_full_rev_frame_0			61	1	Histone 2A	XP_003702149	*Megachile rotundata*	177	3	ConsensusfromContig27957 full rev frame 0		*coenobita*	5583	1
						Glycosyltransferase	EEZ97542	*Tribolium castaneum*	150	3	histone H4	GF20391	*Drosophila ananassae*	5077	5
						ConsensusfromContig42378_full_rev_frame_1			84	1	ConsensusfromContig8357 full fwd frame 0		*coenobita*	3503	4
						ConsensusfromContig78959_full_rev_frame_0			81	1	ConsensusfromContig46299 full rev frame 2		*coenobita*	1487	1
						unknown	YP_004188519	*Vibrio vulnificus MO6-24/O*	76	1	ConsensusfromContig77970_full_fwd_frame_2		*coenobita*	347	3
						peptidase propeptide/YPEB domain protein	EIG27258	*Neisseria sicca VK64*	70	1	ConsensusfromContig69829_full_rev_frame_2		*coenobita*	248	2
3A	ConsensusfromContig80655_full_rev_frame_0 (Lipocalin)		*Coenobita*	78	2	cuticle proprotein proCP5.2	ABR27687	*Callinectes sapidus*	260	6	ConsensusfromContig80655_full_rev_frame_0 (intracellular fatty acid binding protein)		*Coenobita*	3954	3
	ConsensusfromContig28270_full_fwd_frame_0		*Coenobita*	64	1	ConsensusfromContig27957_full_rev_frame_0 (cuticle protein)		*Coenobita*	219	3	ConsensusfromContig82521 full rev frame 1 (nucleoside diphosphate kinase)		*Coenobita*	2372	2
	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	60	1	ConsensusfromContig80908_full_fwd_frame_1		*Coenobita*	116	2	ConsensusfromContig80908_full_fwd_frame_1		*Coenobita*	2076	4

	ConsensusfromContig78959_full_rev_frame_0		*Coenobita*	58	1	ConsensusfromContig78959_full_rev_frame_0		*Coenobita*	81	1	ConsensusfromContig111234_full_fwd_frame_0		*Coenobita*	1989	3
	metallo-beta-lactamase domain protein	ZP_09385291	*Flavonifractor plautii ATCC 29863*	58	1	CGNR zinc finger	ZP_10051040	*Amycolatopsis sp. ATCC 39116*	163	3	ConsensusfromContig110279_full_fwd_frame_0		*Coenobita*	1730	1
						Diguanylate-cyclase (DGC)	YP_341105	*Pseudoalteromonas haloplanktis TAC125*	75	1	ConsensusfromContig8357_full_fwd_frame_0		*Coenobita*	1261	2
											ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	1184	1
											ConsensusfromContig43579 full rev frame 1 (peptidyl prolyl cis trans isomerase)		*Coenobita*	622	2
											ConsensusfromContig110538_full_fwd_frame_0 (putative 40S ribosomal protein)		*Coenobita*	605	3
											ConsensusfromContig96050_full_fwd_frame_0		*Coenobita*	448	3
4A	ConsensusfromContig110279_full_fwd_frame_0		*Coenobita*	82	1	ConsensusfromContig27957_full_rev_frame_0 (probably cuticle protein)		*Coenobita*	173	2	Sarcoplasmic calcium binding protein	P86909.1	*Chionoecetes opilio*	7268	1
	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	69	1	serine-aspartate repeat-containing protein D	EIK10257	*Staphylococcus aureus subsp. aureus VRS1*	158	3	histone H3	XP_003746454.1	*Metaseiulus occidentalis*	2077	6
	ConsensusfromContig28576_full_rev_frame_1		*Coenobita*	69	2	farnesoic acid O-methyltransferase	ACL26692	*Nilaparvata lugens*	78	1	ConsensusfromContig110279_full_fwd_frame_0		*Coenobita*	1490	2
	ConsensusfromContig80557_full_rev_frame_0		*Coenobita*	56	1	ConsensusfromContig112781_full_rev_frame_2		*Coenobita*	72	1	ConsensusfromContig80987 full rev frame 0		*Coenobita*	1042	1
	ConsensusfromContig77555_full_rev_frame_1		*Coenobita*	55	1	unknown	EIJ45050	*Herbaspirillum sp. GW103*	72	1	ribosomal protein S20	AAX38500.1	*Palaemonetes pugio*	552	2
											ConsensusfromContig80557_full_rev_frame_0		*Coenobita*	427	1

											ConsensusfromContig77555_full_rev_frame_1		*Coenobita*	377	2
											cuticle protein	XP_311430.4	*Anopheles gambiae*	265	2
											serine protease K12H4.7	ACO15310.1	*Caligus clemensi*	220	1
5A	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	105	2	ConsensusfromContig52405_full_rev_frame_0 (translation initiation factor 5A)		*Coenobita*	155	2	Sarcoplasmic calcium binding protein	P86909.1	*Chionoecetes opilio*	6640	1
	ConsensusfromContig28270_full_fwd_frame_0		*Coenobita*	78	1	eukaryotic translation initiation factor 5A	ABI30653	*Penaeus monodon*	143	2	ConsensusfromContig52405_full_rev_frame_0 (eukaryotic translation initiation factor 5A)		*Coenibita*	1420	1
	ConsensusfromContig27839_full_rev_frame_0 (ADP-ribosylation factor)		*Coenobita*	54	1	Cuticle protein AMP1A	P81384.1	*Homarus americanus*	135	3	ConsensusfromContig28270_full_fwd_frame_0		*Coenibita*	883	1
						Long-chain-fatty-acid--CoA ligase	ABQ37149	*Bradyrhizobium sp. BTAi1*	114	2	ConsensusfromContig27957_full_rev_frame_0		*Coenibita*	703	3
						MacB-like periplasmic core domain	YP_003913374	*Ferrimonas balearica DSM 9799*	112	2	ConsensusfromContig27589_full_fwd_frame_2		*Coenibita*	340	3
						ConsensusfromContig98685_full_fwd_frame_0		*Coenobita*	113	2	??	AEM92665.1	*Drosophila melanogaster*	314	2
						ConsensusfromContig42378_full_rev_frame_1 (cuticle protein)		*Coenobita*	94	2	??	EFA13199.1	*Tribolium castaneum*	303	1
											serine protease K12H4.7	ACO15310.1	*Caligus clemensi*	181	1
6A	ConsensusfromContig80801_full_fwd_frame_0		*Coenobita*	223	4	calcified cuticle protein CP19.0	ABB91674	*Callinectes sapidus*	281	6	Sarcoplasmic calcium binding protein	P86909.1	*Chionoecetes opilio*	9614	1
	ConsensusfromContig28493_full_rev_frame_1		*Coenobita*	84	2	ConsensusfromContig80801_full_fwd_frame_0		*Coenobita*	271	4	ConsensusfromContig80801_full_fwd_frame_0		*Coenobita*	4628	5

	ConsensusfromContig79899_full_rev_frame_1		*Coenobita*	82	2	ConsensusfromContig28493_full_rev_frame_1		*Coenobita*	197	3	ConsensusfromContig78322_full_rev_frame_2		*Coenobita*	2773	3
	ConsensusfromContig28270_full_fwd_frame_0		*Coenobita*	64	1	ConsensusfromContig67413_full_fwd_frame_0		*Coenobita*	133	2	ConsensusfromContig67413_full_fwd_frame_0		*Coenobita*	2580	1
	ConsensusfromContig29101_full_rev_frame_2 (ADP-ribosylation factor)		*Coenobita*	56	1	ConsensusfromContig43977_full_rev_frame_1		*Coenobita*	132	2	ConsensusfromContig114956_full_rev_frame_1		*Coenobita*	1581	2
						ATP synthase delta (OSCP) subunit	XP_002072797	*Drosophila willistoni*	82	1	ConsensusfromContig108330_full_fwd_frame_0		*Coenobita*	1392	3
						ConsensusfromContig78322_full_rev_frame_2		*Coenobita*	74	1	ConsensusfromContig29101_full_rev_frame_2 (GTPase SAR1 and related small G protein)		*Coenobita*	1386	3
						transgelin	EHJ75654	*Danaus plexippus*	68	1	ConsensusfromContig109225_full_fwd_frame_2		*Coenobita*	1190	4
						ADP-ribosylation factor	XP_003704732	*Megachile rotundata*	66	1	ConsensusfromContig109676_full_fwd_frame_0		*Coenobita*	1044	1
						signal peptidase complex subunit 3-like	XP_003739905	*Metaseiulus occidentalis*	65	1	ConsensusfromContig28493_full_rev_frame_1		*Coenobita*	992	5
											ConsensusfromContig27965_full_rev_frame_1		*Coenobita*	643	4
											Sorting nexin-12	XP_002432737.1	*Pediculus humanus corporis*	841	3
											ConsensusfromContig60539_full_rev_frame_0		*Coenobita*	797	1
											ConsensusfromContig28270_full_fwd_frame_0		*Coenobita*	757	1
											hypothetical protein KGM_19876	EHJ65678	*Danaus plexippus*	676	1
											Sorting nexin 12	CAY54164.1	*Heliconius melpomene*	479	3

7A	ConsensusfromContig109750_full_rev_frame_0		*Coenobita*	190	3	ConsensusfromContig109750_full_rev_frame_0 (triosephosphate isomerase)		*Coenobita*	352	4	Sarcoplasmic calcium binding protein	P86909.1	*Chionoecetes opilio*	1644 1	1
	ConsensusfromContig114956_full_rev_frame_1		*Coenobita*	185	3	Rab GTPase family	EFA09670	*Tribolium castaneum*	229	5	ConsensusfromContig56178_full_fwd_frame_1		*Coenobita*	4854	2
	ConsensusfromContig110093_full_fwd_frame_2		*Coenobita*	108	3	calcified cuticle protein CP19.0	ABB91674	*Callinectes sapidus*	141	2	ConsensusfromContig114956_full_rev_frame_1		*Coenobita*	4756	3
	Rab GTPase family	XP_002973059	*Selaginella moellendorffii*	77	2	ConsensusfromContig112901_full_fwd_frame_2		*Coenobita*	128	2	Rab GTPase family	AGC10822.1	*Eriocheir sinensis*	4308	7
	ConsensusfromContig56178_full_fwd_frame_1		*Coenobita*	74	1	ConsensusfromContig78322_full_rev_frame_2		*Coenobita*	121	2	ConsensusfromContig84076_full_fwd_frame_1; ConsensusfromContig48845_full_rev_frame_2		*Coenobita*	3513	6
	ConsensusfromContig28613_full_rev_frame_1		*Coenobita*	73	1	ConsensusfromContig77989_full_rev_frame_2		*Coenobita*	115	2	ConsensusfromContig78322_full_rev_frame_2		*Coenobita*	3019	4
	ConsensusfromContig78322_full_rev_frame_2		*Coenobita*	71	2	ConsensusfromContig14210_full_rev_frame_1		*Coenobita*	95	1	ConsensusfromContig32512_full_rev_frame_2		*Coenobita*	2887	5
	20S proteasome alpha subunit	AAB93421	*Drosophila melanogaster*	64	1	ConsensusfromContig28613_full_rev_frame_1		*Coenobita*	92	1	triosephosphate isomerase	ABB81879.1	*Fenneropenaeus chinensis*	2801	4
	ConsensusfromContig32512_full_rev_frame_2 (glutathione S-transferase)		*Coenobita*	64	1	ConsensusfromContig56178_full_fwd_frame_1			84	1	ConsensusfromContig110093_full_fwd_frame_2 (peroxiredoxin)		*Coenobita*	2264	7
	ConsensusfromContig81037_full_rev_frame_0		*Coenobita*	59	1	ConsensusfromContig110093_full_fwd_frame_2		*Coenobita*	77	1	ConsensusfromContig77737_full_rev_frame_2 (peroxiredoxin)		*Coenobita*	2131	5
	ConsensusfromContig14210_full_rev_frame_1		*Coenobita*	57	1	peroxiredoxin	YP_001188	*Leptospira interrogans*	68	1	ConsensusfromContig82277 full rev frame 2 (Ras related protein Rab 1A)		*Coenobita*	2080	3
											Peroxiredoxin (PRX) family	XP_002057785.1	*Drosophila virilis*	1509	3
											ConsensusfromContig109750_full_rev_frame_0 (triosephosphate isomerase)		*Coenobita*	1444	7
											ConsensusfromContig76889_full_fwd_frame_1		*Coenobita*	1413	1

											ConsensusfromContig80793_full_rev_frame_0		*Coenobita*	1380	5
											ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	1320	3
											ConsensusfromContig80510_full_rev_frame_0		*Coenobita*	1264	3
											ConsensusfromContig34210_full_fwd_frame_0		*Coenobita*	1225	1
											ConsensusfromContig76889_full_fwd_frame_2		*Coenobita*	1142	2
											ConsensusfromContig67413_full_fwd_frame_0		*Coenobita*	1086	1
											ConsensusfromContig28646_full_rev_frame_0		*Coenobita*	1048	1
											Crustacean calcium-binding protein 23	P80364.1	*Homarus americanus*	1006	4
											ConsensusfromContig78603_full_rev_frame_0		*Coenobita*	980	5
											Ras related protein Rab 11B	EFN87269.1	*Harpegnathos saltator*	939	4
											ConsensusfromContig109090_full_rev_frame_1		*Coenobita*	833	3
											ConsensusfromContig83104_full_rev_frame_0		*Coenobita*	765	1
											ConsensusfromContig80732_full_rev_frame_1		*Coenobita*	749	2
											Glycosyl hydrolase family 1	AEW46885	*Chilo suppressalis*	703	2
											ConsensusfromContig28270_full_fwd_frame_0		*Coenobita*	656	1
											ConsensusfromContig64301_full_rev_frame_0		*Coenobita*	589	1
											proteasome subunit alpha	XP_002413831.1	*Ixodes scapularis*	429	1
											ConsensusfromContig80801_full_fwd_frame_0		*Coenobita*	543	4

											ConsensusfromContig80526_full_rev_frame_0		*Coenobita*	426	3
											ConsensusfromContig44642_full_rev_frame_2		*Coenobita*	313	2
											ConsensusfromContig108988_full_fwd_frame_0		*Coenobita*	307	2
											serine protease K12H4.7	ACO15310.1	*Caligus clemensi*	252	3
8A	cytosolic manganese superoxide dismutase	AEK77429	*Macrobrachium nipponense*	282	5	cytoplasmic manganese superoxide dismutase	CAR85665	*Cancer pagurus*	579	10	14-3-3 protein	AFD33362	*Scylla paramamosain*	13060	8
	14-3-3 protein	AFD28274	*Scylla paramamosain*	141	3	14-3-3 protein	AFD33362	*Scylla paramamosain*	300	6	cytosolic manganese superoxide dismutase	AFD61666	*Cherax quadricarinatus*	10428	6
	ConsensusfromContig56178_full_fwd_frame_1		*Coenobita*	70	1	ConsensusfromContig85153_full_rev_frame_0 (cytoplasmic manganese superoxide dismutase)		*Coenobita*	288	4	ConsensusfromContig85153 full rev frame 0		*Coenobita*	6208	4
	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	69	1	ConsensusfromContig35858_full_rev_frame_0 (14-3-3 protein)		*Coenobita*	112	2	ConsensusfromContig32822 full rev frame 0		*Coenobita*	3800	2
						ConsensusfromContig110086_full_rev_frame_1		*Coenobita*	108	1	heat shock 70 kDa protein	EHJ73638	*Danaus plexippus*	2857	6
						ConsensusfromContig109750_full_rev_frame_0 (triosephosphate isomerase)			87	1	ConsensusfromContig110093 full fwd frame 2 (Thioredoxin peroxidase)		*Coenobita*	1918	2
						ConsensusfromContig56178_full_fwd_frame_1		*Coenobita*	84	1	ConsensusfromContig108678_full_fwd_frame_0 (Prohibitin)		*Coenobita*	1713	9
						ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	71	1	ConsensusfromContig81048 full rev frame 0		*Coenobita*	1596	2
						general secretion pathway protein N	YP_004791128	*Stenotrophomonas maltophilia JV3*	71	1	ConsensusfromContig80585 full rev frame 1		*Coenobita*	1366	2
						Glycerol kinase	XP_001352729	*Drosophila pseudoobscura pseudoobscura*	66	1	ConsensusfromContig110086 full rev frame 1		*Coenobita*	1293	1

											ConsensusfromContig28270 full fwd frame 0		*Coenobita*	1249	1
											Phosphoglycerate mutase 2		*Culex quinquefasciatus*	1208	2
											Beta-ketoacyl-ACP reductase	EEC12925	*Ixodes scapularis*	1132	1
											elongation factor	ADJ36967	*Heterophlebus versabilis*	963	2
											ConsensusfromContig101727_full_rev_frame_0		*Coenobita*	955	1
											Ran protein	AEO23975	*Penaeus monodon*	806	3
											Triosephosphate isomerase	ABB81879	*Fenneropenaeus chinensis*	665	3
											ConsensusfromContig84140 full rev frame 1 (Phosphoglycerate mutase 2)		*Coenobita*	581	2
											ConsensusfromContig50037 full rev frame 0		*Coenobita*	536	3
											actin	ACR54118	*Palaemonetes varians*	531	1
											ConsensusfromContig78116 full rev frame 1		*Coenobita*	481	3
											ConsensusfromContig55748 full fwd frame 1 (actin)		*Coenobita*	433	1
											ConsensusfromContig12456 full rev frame 1		*Coenobita*	386	1
											ConsensusfromContig109750 full rev frame 0 (Triosephosphate isomerase)		*Coenobita*	358	4
											Peroxiredoxin-6	ACO15034	*Caligus clemensi*	330	2
											ConsensusfromContig35747 full rev frame 1		*Coenobita*	321	3

9A	14-3-3 protein	AFD28274	*Scylla paramamosain*	90	2	14-3-3 protein	AFD28274	*Scylla paramamosain*	366	6	Sarcoplasmic calcium binding protein	P86909.1	*Chionoecetes opilio*	8360	1
	ConsensusfromContig28984_full_fwd_frame_0		*Coenobita*	75	1	cytosolic manganese superoxide dismutase	AFD61666	*Cherax quadricarinatus*	243	4	14-3-3 protein	AFC36878.1	*Litopenaeus vannamei*	6232	9
	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	71	1	ConsensusfromContig85153_full_rev_frame_0 (cytosolic manganese superoxide dismutase)		*Coenobita*	153	2	cytosolic manganese superoxide dismutase	AFD61666.1	*Cherax quadricarinatus*	6017	5
	ConsensusfromContig85153_full_rev_frame_0 (cytosolic manganese superoxide dismutase)		*Coenobita*	69	1	voltage-dependent anion-selective channel	ADJ94951	*Eriocheir sinensis*	126	2	ConsensusfromContig85153 full rev frame 0 (cytosolic manganese superoxide dismutase)		*Coenobita*	2576	3
	voltage-dependent anion-selective channel	ADJ94951	*Eriocheir sinensis*	61	1	ConsensusfromContig44775_full_rev_frame_2 (14-3-3 protein)		*Coenobita*	113	2	ConsensusfromContig81048 full rev frame 0		*Coenobita*	1736	2
						ConsensusfromContig28984_full_fwd_frame_0 (14-3-3 zeta protein)		*Coenobita*	81	1	voltage-dependent anion-selective channel	ADJ94951	*Eriocheir sinensis*	980	2
						ConsensusfromContig523_full_rev_frame_0			77	1	ConsensusfromContig28984_full_fwd_frame_0		*Coenobita*	889	1
						coatomer protein complex	NP_001166195	*Bombyx mori*	71	1	tropomyosin	BAF95205 1	*Euphausia superba*	599	7
						tropomyosin	AAQ57241	*Aleuroglyphus ovatus*	62	1	ConsensusfromContig78175_full_rev_frame_1		*Coenobita*	576	3
						gamma-interferon inducible lysosomal thiol reductase	XP_002400600	*Ixodes scapularis*	65	1	electron transfer flavoprotein	EDV37366 1	*Drosophila ananassae*	534	2
						Farnesoic acid 0-methyl transferase	EFZ11646	*Solenopsis invicta*	65	1	glyceraldehyde 3 phosphate dehydrogenase	ACB99696	*Orconectes cf. virilis LMM-2008*	482	4
						related to structure-specific endonuclease	NP_690964	*Oceanobacillus iheyensis HTE831*	65	1	ConsensusfromContig78277 full fwd frame 0		*Coenobita*	447	2
											ConsensusfromContig28984_full_fwd_frame_2 (14-3-3 zeta protein)		*Coenobita*	430	3

											putative glucosamine 6 phosphate isomerase	ABV81865.1	*putative glucosamine 6 phosphate isomerase*	409	2
											ConsensusfromContig110093_full_fwd_frame_2		*Coenobita*	348	2
10A	glyceraldehyde-3-phosphate dehydrogenase	671058A	*Homarus sp.*	266	4	glyceraldehyde-3-phosphate dehydrogenase	ACK56132	*Portunus trituberculatus*	561	10	glyceraldehyde-3-phosphate dehydrogenase	ADU87475.1	*Neopetrolisthes maculatus*	13806	7
	Tropomyosin	AAC48288	*Homarus americanus*	177	3	ConsensusfromContig78044_full_rev_frame_1		*Coenobita*	226	3	Sarcoplasmic calcium binding protein	P86909.1	*Chionoecetes opilio*	10785	1
	ConsensusfromContig78044_full_rev_frame_1		*Coenobita*	153	3	ConsensusfromContig28115_full_fwd_frame_0		*Coenobita*	135	2	acidic ribosomal protein P0	ACY66540.1	*Scylla paramamosain*	5082	3
	lactate dehydrogenase A	AEK84338	*Daphnia pulex*	100	1	ConsensusfromContig52669_full_fwd_frame_1 (Glyceraldehyde 3 phosphate dehydrogenase)		*Coenobita*	114	1	ConsensusfromContig78044_full_rev_frame_1 (malate dehydrogenase)		*Coenobita*	4876	5
	receptor for activated protein kinase c1	ABU49887	*Penaeus monodon*	92	2	ConsensusfromContig44336_full_rev_frame_1		*Coenobita*	75	1	ConsensusfromContig9100 full rev frame 2 (glyceraldehyde 3 phosphate dehydrogenase)		*Coenobita*	4789	3
	60S acidic ribosomal protein	ACY66540	*Scylla paramamosain*	77	1	ConsensusfromContig47491_full_rev_frame_2 (acidic p0 ribosomal protein)		*Coenobita*	74	1	tropomyosin	BAG30719.1	*Papilio xuthus*	2790	4
	ConsensusfromContig110182_full_rev_frame_2		*Coenobita*	68	1						lactate dehydrogenase A	AEK84381.1	*Daphnia pulex*	2583	5
											ConsensusfromContig47491_full_rev_frame_2 (acidic p0 ribosomal protein)		*Coenobita*	2373	3
											malate dehydrogenase	ACY66481.1	*Scylla paramamosain*	2350	4
											ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	2235	2
											receptor for activated protein kinase c1	ABU49887.1	*Penaeus monodon*	1733	9

											ConsensusfromContig52669_full_fwd_frame_1 (Glyceraldehyde 3 phosphate dehydrogenase)		*Coenobita*	1642	1
											ConsensusfromContig81485 full fwd frame 2 (cytosolic malate dehydrogenase)		*Coenobita*	1403	3
											ConsensusfromContig31896_full_rev_frame_2		*Coenobita*	1224	2
											alpha 1 tubulin	AAC47522.1	*alpha 1 tubulin*	885	5
											ConsensusfromContig44336 full rev frame 1		*Coenobita*	796	5
											ConsensusfromContig31520 full rev frame 1 (receptor for activated protein kinase c1)		*Coenobita*	775	3
											ConsensusfromContig28115 full fwd frame 0		*Coenobita*	757	2
											G protein beta 1	AFD33363.1	*Scylla parama mosain*	612	6
											elongation factor 1	AAY56031.1	*Uroleucon sonchi*	475	3
											ConsensusfromContig31496 full fwd frame 1 (alpha tubulin)		*Coenobita*	490	4
											ConsensusfromContig31496 full fwd frame 1 (alpha 1 tubulin)		*Coenobita*	490	4
											Guanine nucleotide-binding protein	EFN78454	*Harpegnathos saltator*	432	5
											ConsensusfromContig31027 full rev frame 1		*Coenobita*	392	1
											ConsensusfromContig28107_full_rev_frame_1		*Coenobita*	365	1
											transaldolase-like protein	AEE62945.1	*Dendroctonus ponderosae*	337	2

11A	Arginine kinase	NP_001141078	*Zea mays*	207	4	arginine kinase	AEY84969	*Scylla paramamosain*	641	10	Arginine kinase	AAC39040	*Apis mellifera*	16720	10
	glyceraldehyde-3-phosphate dehydrogenase	671058A	*Homarus sp.*	151	3	ructose-1,6-bisphosphate aldolase	EFA04932	*Tribolium castaneum*	331	7	Glyceraldehyde 3 phosphate dehydrogenase	P00357	*Homarus americanus*	5048	6
	beta-actin	ABD48797	*Culex pipiens pipiens*	145	3	glyceraldehyde-3-phosphate dehydrogenase	ABA89862	*Pelobacter carbinolicus DSM 2380*	333	6	elongation factor 1-alpha	AFP75580	*Micracidini gen. sp. 2 BHJ-2012*	2680	3
	ConsensusfromContig37432_full_rev_frame_1		*Coenobita*	73	1	ConsensusfromContig52669_full_fwd_frame_1 (glyceraldehyde-3-phosphate dehydrogenase)		*Coenobita*	114	1	ConsensusfromContig9100_full_rev_frame_2 (Glyceraldehyde 3 phosphate dehydrogenase)		*Coenobita*	2315	1
	ConsensusfromContig27836_full_fwd_frame_2		*Coenobita*	53	1	ConsensusfromContig42813_full_fwd_frame_1		*Coenobita*	88	1	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	2299	4
						ConsensusfromContig27836_full_fwd_frame_2		*Coenobita*	88	1	ConsensusfromContig37432_full_rev_frame_1		*Coenobita*	1826	2
						ConsensusfromContig89471_full_rev_frame_0 (arginine kinase)		*Coenobita*	82	1	RNA binding protein, squid	NP_731825	*Drosophila melanogaster*	1208	5
											Actin cytoplasmic	EGW08753	*Cricetulus griseus*	1094	4
											ConsensusfromContig29513_full_rev_frame_0 (RNA binding protein)		*Coenobita*	798	2
											ConsensusfromContig633_full_fwd_frame_1 (RNA binding protein)		*Coenobita*	798	2
											ConsensusfromContig52669_full_fwd_frame_1 (glyceraldehyde-3-phosphate dehydrogenase)		*Coenobita*	545	1
											L lactate dehydrogenase	EFR27453	*Anopheles darlingi*	523	2

											Aspartate protease cathepsin D	AEO94539	*Triatoma infestans*	520	2
											ConsensusfromContig55748_full_fwd_frame_1 (actin)			473	1
											ConsensusfromContig30565 full rev frame 1			448	5
											Hemocyanin subunit 6	AEG64817	*Eriocheir sinensis*	373	3
											ConsensusfromContig27788 full rev frame 0			369	2
											Fructose bisphosphate aldolase	D2A6E8	*Tribolium castaneum*	334	4
12A	cytoplasmic type actin 2	ACI23574	*Homarus americanus*	277	5	actin	ADG45302	*Marsupenaeus japonicus*	750	12	Actin cytoplasmic	AAC47432.1 AAC47432.1 AAC47432.1 AAC47432.1	*Bombyx mori*	35937	19
	ribosome-associated protein P40	BAB78527	*Bombyx mori*	74	1	isocitrate dehydrogenase	YP_003088100	*Dyadobacter fermentans DSM 18053*	260	5	arginine kinase	CAC85911.1	*Plodia interpunctella*	3197	4
	hemocyanin	AAL27460	*Penaeus monodon*	71	1	ConsensusfromContig55748_full_fwd_frame_1; ConsensusfromContig33236_full_rev_frame_2 (actin)		*Coenobita*	107	1	Isocitrate dehydrogenase	XP_970446.1 XP_970446.1	*Tribolium castaneu*	3035	4
	ConsensusfromContig68773_full_rev_frame_1		*Coenobita*	68	1	ConsensusfromContig32340_full_fwd_frame_0		*Coenobita*	101	1	Elongation factor 1-alpha	AAD21853	*Cypridopsis vidua*	2862	5
	isocitrate dehydrogenase (NADP)	XP_390523	*Gibberella zeae PH-1*	67	1						ConsensusfromContig32340 full fwd frame 0 (Isocitrate dehydrogenase)		*Coenobita*	2437	2

	ConsensusfromContig32340_full_fwd_frame_0 (isocitrate dehydrogenase (NADP))		*Coenobita*	66	1						ConsensusfromContig78371 full rev frame 2		*Coenobita*	1253	2
	Arginine kinase	XP_003695937	*Apis florea*	65	1						ConsensusfromContig81048 full rev frame 0		*Coenobita*	1187	2
	ConsensusfromContig80831_full_fwd_frame_1		*Coenobita*	61	1						Ribosomal protein S2		*Litopenaeus vannamei*	1128	2
	ConsensusfromContig27671_full_rev_frame_2		*Coenobita*	54	1						ConsensusfromContig80831 full fwd frame 1		*Coenobita*	1020	2
											ConsensusfromContig28270 full fwd frame 0		*Coenobita*	877	2
											ConsensusfromContig80547 full fwd frame 1 (actin)		*Coenobita*	837	4
											ConsensusfromContig83931 full rev frame 1 (Fructose bisphosphate aldolase)		*Coenobita*	819	2
											Glyceraldehyde 3 phosphate dehydrogenase	AAR96458.1 AAR96458.1 AAR96458.1	*Cherax quadricarinatus*	814	4
											Hemocyanin like protein	ABV58636	*Metapenaeus ensis*	795	2
											Fructose bisphosphate aldolase		*Tribolium castaneum*	753	5
											ConsensusfromContig27671 full rev frame 2		*Coenobita*	741	2
											Tropomyosin		*Euphausia superba*	548	4
											ConsensusfromContig112805 full fwd frame 0 (Fructose bisphosphate aldolase)		*Coenobita*	467	1

											ConsensusfromContig75555_full_fwd_frame_1		*Coenobita*	392	1
											ConsensusfromContig110333_full_rev_frame_0		*Coenobita*	376	1
											ConsensusfromContig46635_full_rev_frame_2		*Coenobita*	349	1
											ConsensusfromContig78006_full_fwd_frame_2 (ConsensusfromContig78006_full_fwd_frame_2)		*Coenobita*	345	1
											ConsensusfromContig30363 full rev frame 0		*Coenobita*	335	2
											Fumarylacetoacetase	AAF47833	*Drosophila melanogaster*	263	4
											Collagen alpha-1(IV)	AAA28404.1	*Drosophila melanogaster*	219	14
											serine protease K12H4.7	ACO15310.1	*Caligus clemensi*	206	1
1C	F1-ATP synthase beta subunit	ABI34071	*Pacifastacus leniusculus*	136	3	Tubulin alpha 1	AAC47522	*Gecarcinus lateralis*	344	5	F0F1 ATP synthase subunit beta	AEV89780	*Schistocerca gregaria*	11775	9
	actin	AAA28321	*Drosophila melanogaster*	127	2	actin	XP_001601504.1	*Nasonia vitripennis*	291	6	Tubulin alpha 1	EDW38107	*Drosophila persimilis*	10924	11
	elongation factor 1 alpha	ABP57624	*Janatella hera*	110	1	skeletal muscle actin 2	ACI23565	*Homarus americanus*	283	6	elongation factor 1 alpha	EFA84695	*Polyspho ndylium pallidum*	8847	7
	Tubulin alpha 1	AAC47522	*Gecarcinus lateralis*	71	2	F0F1 ATP synthase subunit beta	EHK58393	*Mesorhizobium alhagi CCNWXJ12-2*	213	4	ConsensusfromContig31496_full_fwd_frame_1 (alpha-1-tubulin)		*Coenobita*	7093	5
	adenosylhomocysteinase	ACY66467	*Scylla paramamosain*	64	1	ConsensusfromContig51584_full_rev_frame_0 (enolase)		*Coenobita*	187	3	Tubulin beta 1	EFZ14304	*Solenopsis invicta*	4851	7

	ConsensusfromContig51584_full_rev_frame_0 (enolase)		*Coenobita*	50	1	ConsensusfromContig31496_full_fwd_frame_1 (alpha-1-tubulin)		*Coenobita*	104	1	Actin	Bombyx mori	*P04829*	3261	5
						ConsensusfromContig80831_full_fwd_frame_1		*Coenobita*	88	1	ConsensusfromContig80758 full rev frame 1		*Coenobita*	3005	3
						ConsensusfromContig82928_full_rev_frame_0; ConsensusfromContig89315_full_rev_frame_0 (beta-tubulin)		*Coenobita*	71	1	ConsensusfromContig110113 full rev frame 2		*Coenobita*	2066	4
											ConsensusfromContig77793 full rev frame 2 (ATP synthase subunit beta)		*Coenobita*	1927	2
											Enolase	ADU87515	*Hippa adactyla*	1873	2
											ConsensusfromContig31598 full rev frame 1 (ATP synthase subunit beta)		*Coenobita*	1867	2
											ConsensusfromContig81048 full rev frame 0		*Coenobita*	1421	2
											Adenosylhomocysteinase	ACY66467	*Scylla paramamosain*	1112	1
											ConsensusfromContig82928_full_rev_frame_0 (Tubulin beta 1)		*Coenobita*	1100	1
											40S ribosomal protein	Q0PXX8	*Diaphorina citri*	975	3
											ConsensusfromContig78006 full fwd frame 2 (40S ribosomal protein)		*Coenobita*	565	1
											Hemocyanin subunit 6	AEG64817	*Eriocheir sinensis*	220	1
											serine protease K12H4.7	ACO15310.1	*Caligus clemensi*	166	1
2C	Tubulin beta 1	Q25009	*Homarus americanus*	458	9	Tubulin beta	EFZ14304	*Solenopsis invicta*	1391	24	Tubulin alpha 1	EGI61892	*Acromyrmex echinatior*	21812	16

	F0F1 ATP synthase subunit beta	ABI34071	*Pacifastacus leniusculus*	256	4	Tubulin alpha 1	AEF33840.1	*Cherax quadricarinatus*	730	10	F0F1 ATP synthase subunit beta	ADC55252.1	*Litopenaeus vannamei*	17531	9
	Tubulin alpha 1	AAC47522	*Gecarcinus lateralis*	198	4	F0F1 ATP synthase subunit beta	YP_001261129	*Sphingomonas wittichii RW1*	159	3	Tubulin beta 2	Q94571	*Homarus americanus*	11717	22
	ConsensusfromContig80548_full_rev_frame_1 (calreticulin)		*Coenobita*	64	1	ConsensusfromContig31496_full_fwd_frame_1 (Tubulin alpha 1)		*Coenobita*	442	6	ConsensusfromContig25049_full_rev_frame_2		*Coenobita*	8816	2
	ConsensusfromContig2683_full_rev_frame_0 (26S protease regulatory subunit 6B-like protein)		*Coenobita*	56	1	ConsensusfromContig110800_full_fwd_frame_2ConsensusfromContig82928_full_rev_frame_0		*Coenobita*	234	3	ConsensusfromContig33429_full_fwd_frame_0 (Protein disulfide isomerase)		*Coenobita*	5154	2
						ConsensusfromContig82928_full_rev_frame_0 (tubulin beta)		*Coenobita*	215	3	ConsensusfromContig88167_full_rev_frame_0		*Coenobita*	4972	4
						ConsensusfromContig25049_full_rev_frame_2, ConsensusfromContig26577_full_rev_frame_1		*Coenobita*	165	3	ConsensusfromContig79010_full_fwd_frame_0 (protein disulfide isomerase)		*Coenobita*	4843	4
						ConsensusfromContig109711_full_rev_frame_1		*Coenobita*	96	1	ConsensusfromContig80548_full_rev_frame_1		*Coenobita*	3830	6
						ConsensusfromContig15227_full_rev_frame_1		*Coenobita*	86	1	Catalase	AET34916	*Macrobrachium rosenbergii*	3638	5
						ConsensusfromContig11887_full_rev_frame_1; ConsensusfromContig38994_full_rev_frame_0		*Coenobita*	76	1	Protein disulfide isomerase	ACN89260	*Litopenaeus vannamei*	3231	4
						ConsensusfromContig33019_full_rev_frame_0 (Tubulin beta)		*Coenobita*	165	3	Elongation factor 1 alpha	ABY59402	*Leptarthrus brevirostris*	3098	5
						ConsensusfromContig80548_full_rev_fram		*Coenobita*	73	1	ConsensusfromContig77793_full_rev_frame_2 (F0F1 ATP synthase subunit beta)		*Coenobita*	2462	1
											ConsensusfromContig112167_full_fwd_frame_1 (catalase)		*Coenobita*	2095	1

											ConsensusfromContig89338_full_fwd_frame_2 (Tubulin alpha 1)		*Coenobita*	2036	1
											ConsensusfromContig31598_full_rev_frame_1 (F0F1 ATP synthase subunit beta)		*Coenobita*	1933	3
											Calsequestrin-1	AEI88118	*Scylla paramamosain*	1880	3
											ATP synthase subunit alpha	AAF46903 AAF46903	*Drosophila melanogaster*	1904	8
											ConsensusfromContig81048 full rev frame 0		*Coenobita*	1713	1
											ConsensusfromContig89315_full_rev_frame_0 (Tubulin beta 2)		*Coenobita*	1688	2
											ConsensusfromContig37063 full fwd frame 2 (Protein disulfide isomerase)		*Coenobita*	1620	2
											ConsensusfromContig1407 full fwd frame 0		*Coenobita*	1348	3
											Calreticulin	ABC50166	*Fenneropenaeus chinensis*	1321	6
											ConsensusfromContig26577 full rev frame 1		*Coenobita*	1258	1
											ConsensusfromContig15227 full rev frame 1		*Coenobita*	1211	1
											ConsensusfromContig109711 full rev frame 1 (Protein disulfide isomerase)		*Coenobita*	1056	2
											ConsensusfromContig26380_full_fwd_frame_2		*Coenobita*	1049	3
											ConsensusfromContig80758 full rev frame 1 (enolase)		*Coenobita*	940	2
											ConsensusfromContig78240 full fwd frame 2 (Calreticulin)		*Coenobita*	905	4

											ConsensusfromContig29785_full_rev_frame_1		*Coenobita*	706	3
											ConsensusfromContig80872 full rev frame 1 (translation initiation factor)		*Coenobita*	673	5
											Enolase	1PDZ_A	*Homarus gammarus*	598	5
											Transient receptor potential-gamma protein	EFN78599	*Harpegnathos saltator*	439	1
											Pyruvate kinase	NP_524448.3	*Drosophila melanogaster*	486	4
											Protein disulfide isomerase	EFN85405.1	*Harpegnathos saltator*	359	3
											Rbm25 protein	Q17BR8	*Aedes aegypti*	350	12
											ConsensusfromContig48214_full_fwd_frame_1		*Coenobita*	338	1
											sulfotransferase	XP_001600316	*Nasonia vitripennis*	208	4
											Hemocyanin	AAM81357	*Pacifastacus leniusculus*	198	3
											serine protease K12H4.7	ACO15310.1	*Caligus clemensi*	166	3
3C	beta-tubulin	ACY66637	*Scylla paramamosain*	178	4	protein disulfide isomerase 2	AEE36486	*Fenneropenaeus chinensis*	704	13	ConsensusfromContig79010_full_fwd_frame_0 (protein disulfide isomerase)		*Coenobita*	9283	6
	heat shock protein 60	ACN30235	*Litopenaeus vannamei*	100	2	hemocyanin subunit 1	AAW57889	*Metacarcinus magister*	406	8	ConsensusfromContig33429_full_fwd_frame_0		*Coenobita*	7090	2
	ConsensusfromContig88167_full_rev_frame_0		*Coenobita*	81	1	catalase	AET34916	*Macrobrachium rosenbergi*	277	6	ConsensusfromContig25049_full_rev_frame_2		*Coenobita*	4304	2

								i							
	hemocyanin	AAL27460	*Penaeus monodon*	70	1	Calsequestrin-1	XP_002425249	*Pediculus humanus corporis*	242	5	ConsensusfromContig109711_full_rev_frame_1 (protein disulfide isomerase)		*Coenobita*	3119	3
	ConsensusfromContig87774_full_fwd_frame_1		*Coenobita*	54	1	ConsensusfromContig109711_full_rev_frame_1 (protein disulfide isomerase)		*Coenobita*	263	4	catalase	AET34916	*Macrobrachium rosenbergii*	2836	6
	pyruvate kinase	ABO21408	*Litopenaeus vannamei*	53	1	Tubulin	EFR28757	*Anopheles darlingi*	225	5	Tubulin alpha 1	Q25008.1	*Homarus americanus*	2504	7
						heat shock protein 60	ACN30235	*Litopenaeus vannamei*	201	4	disulfide isomerase	XP_001605359	*Nasonia vitripennis*	2348	2
						pyruvate kinase	ABY66598	*Litopenaeus vannamei*	150	3	ConsensusfromContig88167_full_rev_frame_0		*Coenobita*	1788	2
						ConsensusfromContig89315_full_rev_frame_0; ConsensusfromContig82928_full_rev_frame_0 (Tubulin)		*Coenobita*	110	2	ConsensusfromContig31496 full fwd frame 1 (Tubulin alpha)		*Coenobita*	1601	6
						ConsensusfromContig88167_full_rev_frame_0		*Coenobita*	103	1	hemocyanin	XP_003740838	*Coenobita*	1521	4
						ConsensusfromContig104681_full_fwd_frame_0 (pyruvate kinase)		*Coenobita*	84	1	Tubulin beta 2	Q94571	*Homarus americanus*	1393	4
						ConsensusfromContig1407_full_fwd_frame_0		*Coenobita*	73	1	ConsensusfromContig91554_full_rev_frame_2		*Coenobita*	1362	2
						ConsensusfromContig33429_full_fwd_frame_0 (protein disulfide isomerase)		*Coenobita*			ConsensusfromContig112167_full_fwd_frame_1 (catalase)		*Coenobita*	1333	2
											ConsensusfromContig84349 full rev frame 0 (catalase)		*Coenobita*	995	5
											heat shock protein 70	CAJ41162	*Xantho incisus*	962	6
											ConsensusfromContig110800 full fwd frame 2 (Tubulin beta)		*Coenobita*	815	1
											coatomer subunit delta-like	XP_003700942	*Megachile rotundata*	738	6

											delta-coatomer protein, isoform B	NP_001162642	*Drosophila melanogaster*	723	5
											Trafficking protein particle complex subunit 10	XP_003742750	*Metaseiulus occidentalis*	671	2
											ConsensusfromContig810 full rev frame 0 (coatomer protein complex)		*Coenobita*	639	2
											ConsensusfromContig99141 full fwd frame 0		*Coenobita*	580	4
											GST_N_Metaxin	XP_002430330	*Pediculus humanus corporis*	489	4
											pyruvate kinase	ABY66598	*Litopenaeus vannamei*	490	4
											ATP synthase subunit beta	ADC55252	*Litopenaeus vannamei*	452	8
											ConsensusfromContig82267 full rev frame 1		*Coenobita*	423	1
											heat shock protein 60	EGI60183	*Acromyrmex echinatior*	419	6
											hemocyanin	AF522504	*Pacifastacus leniusculus*	357	5
4C	hemocyanin	AAL27460	*Penaeus monodon*	175	3	hemocyanin	AAF64305	*Callinectes sapidus*	1780	36	heat shock protein 70	XP_002415926	*Ixodes scapularis*	4313	7
	heat shock protein 70	AAN86047	*Spodoptera frugiper*	96	2	heat shock protein 70	NP_524474	*Drosophila melanoga*	161	3	heat shock protein 70	AFQ62791	*Litopenaeus vannam*	3022	6

			da					*ster*					*ei*		
						ConsensusfromContig82003_full_rev_frame_2; ConsensusfromContig80527_full_fwd_frame_1 (heat shock protein 70)		*Coenobita*	77	1	heat shock protein 70	ABM92447	*Fennero penaeus chinensis*	2348	5
						ConsensusfromContig81783_full_fwd_frame_1		*Coenobita*	71	1	ConsensusfromContig80527_full_fwd_frame_1 (heat shock protein 70)		*Coenobita*	1881	1
						ConsensusfromContig43493_full_fwd_frame_0		*Coenobita*	71	1	haemocyanin	AF193793	*Nephrops norvegicus*	954	2
											Adenosine/AMP deaminase N-terminal	XP_001954181	*Drosophila ananassae*	466	2
5C	elongation factor 2	AF331798	*Aedes aegypti*	114	2	hemocyanin	AAW57893	*Metacarcinus magister*	111	20	heat shock protein 90	AGC54636	*Scylla paramamosain*	3729	11
	hemocyanin	AAL27460	*Penaeus monodon*	104	2	elongation factor 2	XP_002400289	*Ixodes scapularis*	235	5	translation elongation factor	AAL83698	*Spodoptera exigua*	2561	6
	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	93	1	endoplasmin	ACS75351	*Locusta migratoria*	77	1	ConsensusfromContig81844_full_rev_frame_0 (heat shock protein 90)		*Coenobita*	1444	2
	heat shock protein 90	NP_001038538 XP_688963 XP_709210 XP_709211	*Danio rerio*	82	2	heat shock protein 90	XP_001990119	*Drosophila grimshawi*	74	1	ConsensusfromContig15629_full_fwd_frame_0 (82 kDa heat shock protein)		*Coenobita*	1112	3
	heat shock protein 83	AK402420	*apilio polytes*	75	2	similar to ephrin	XP_967625	*Tribolium castaneum*	108	2	hemocyanin subunit Y	ABR14694	*Marsupenaeus japonicus*	1099	6
	ConsensusfromContig81844_full_rev_frame_0		*Coenobita*	58	1						ConsensusfromContig81048 full rev frame 0		*Coenobita*	836	2
											heat shock protein 83	AAB58358	*Drosophila auraria*	834	5

											elongation factor 2	AF331798	*elongation factor 2*	754	6
											ConsensusfromContig47227_full_rev_frame_2		*Coenobita*	745	3
											ConsensusfromContig78138_full_rev_frame_0		*Coenobita*	613	2
											nucleotidyl transferase superfamily	EFA00794	*Tribolium castaneum*	342	1
6C	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	157	2	hemocyanin	AEB77775	*Penaeus monodon*	257	5	ConsensusfromContig57845_full_fwd_frame_1		*Coenobita*	1473	1
	heat shock protein 90	NP_651601	*Drosophila melanogaster*	110	2	5,10-methylene-tetrahydrofolate dehydrogenase	XP_003700207	*Megachile rotundata*	131	2	heat shock protein 90	AGC54636	*Scylla paramamosain*	235	7
	ConsensusfromContig28270_full_fwd_frame_0		*Coenobita*	51	1	GH15863 gene product from transcript GH15863-RA, peptidesa like	XP_001983370	*Drosophila grimshawi*	105	2					
						Adenosine deaminase	YP_003547409	*Coraliomargarita akajimensis DSM 45221*	151	3					
7C	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	137	2	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	139	2	ConsensusfromContig31628_full_rev_frame_1 (alpha 2 macroglobulin)		*Coenobita*	2516	3
						ConsensusfromContig57845_full_fwd_frame_1		*Coenobita*	107	2	ConsensusfromContig57845_full_fwd_frame_1		*Coenobita*	2296	1
						outer membrane lipoprotein LolB	YP_004421191	*Gallibacterium anatis UMN179*	79*	1	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	2113	2
						transmembrane efflux protein	ZP_07313062	*Streptomyces griseoflavus Tu4000*	74	1	alpha 2 macroglobulin	ABP97431	*Fenneropenaeus chinensis*	1162	2
						conserved protein of unknown function	YP_006368848	*Modestobacter*	74	1	ConsensusfromContig108788_full_fwd_frame_0		*Coenobita*	779	1

								marinus							
						enoyl-CoA hydratase	ZP_10068016	*Streptomyces tsukubaensis NRRL18488*	74	1	ConsensusfromContig41055_full_fwd_frame_2		*Coenobita*	646	1
						xylanase b	ZP_10240077	*Paenibacillus peoriae KCTC 3763*	74	1					
						ConsensusfromContig108788_full_fwd_frame_0		*Coenobita*	68	1					
						ConsensusfromContig28034_full_fwd_frame_1		*Coenobita*	65	1					
8C	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	155	2	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	192	3	Sarcoplasmic calcium binding protein	P86909.1	*Chionoecetes opilio*	13219	1
						O-6-methylguanine DNA methyltransferase	ZP_10749522	*Caulobacter sp. AP07*	77	1	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	2579	5
						Ribulose-5-phosphate 4-epimerase	YP_003958511	*Eubacterium limosum KIST612*	76	1	Polyubiquitin	EDS39770	*Culex quinquefasciatus*	2474	14
						phosphoadenosine phosphosulphate (PAPS) reductase	XP_001483941	*Meyerozyma guilliermondii ATCC 6260*	74	1	ConsensusfromContig64225_full_rev_frame_0		*Coenobita*	1471	3
						hypothetical protein	EIW58564	*Trametes versicolor FP-101664 SS1*	73	1	ConsensusfromContig28270_full_fwd_frame_0		*Coenobita*	1342	2
						putative replication protein	YP_004215668	*Rahnella sp. Y9602*	72	1	tyrosine protein phosphatase	XP_003485392	*Bombus impatiens*	599	4
						hypothetical protein	YP_002908295	*Burkholderia glumae BGR1*	70	1	ConsensusfromContig75555_full_fwd_frame_1		*Coenobita*	569	1
						hypothetical protein	YP_004408136	*Verrucosispora maris AB-*	70	1	ConsensusfromContig108788_full_fwd_frame_0		*Coenobita*	543	1

								18-032							
						E3 ubiquitin-protein ligase RING1	NP_001188361	*Bombyx mori*	66	1	ConsensusfromContig810_full_fwd_frame_2		*Coenobita*	483	2
						hypothetical protein	EFX72439	*Daphnia pulex*	66	1	ConsensusfromContig111261_full_rev_frame_0		*Coenobita*	438	1
						ConsensusfromContig76092_full_fwd_frame_2; ConsensusfromContig85752_full_rev_frame_0		*Coenobita*	60	1	GI11496 gene product from transcript GI11496-RA	EDW19490	*Drosophila mojavensis*	437	1
						ConsensusfromContig68377_full_fwd_frame_1		*Coenobita*	66	1					
						GL18119 gene product from transcript GL18119-RA	XP_002028397	*Drosophila persimilis*	65	1					
						ConsensusfromContig15411_full_fwd_frame_0; ConsensusfromContig37868_full_rev_frame_2		*Coenobita*	63	1					
						ConsensusfromContig24388_full_fwd_frame_1		*Coenobita*	61	1					
9C	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	82	1	alpha 2-macroglobulin	ABD61456	*Scylla serrata*	278	6	polyubiquitin	XP_001845257	*Culex quinquefasciatus*	18534	29
	ankyrin	YP_001974936	*Culex quinquefasciatus*	66	1	cuticle protein	XP_002007010	*Drosophila mojavensis*	166	3	sarcoplasmic calcium binding protein	P86909.1	*Chionoecetes opilio*	2343	1
						ConsensusfromContig57845_full_fwd_frame_1 (cuticle protein)		*Coenobita*	137	2	ribosomal protein L40	NP_476776	*Drosophila melanogaster*	1341	3
						ConsensusfromContig31628_full_rev_frame_1		*Coenobita*	112	2	alpha 2 macroglobulin	AEC50080	*Pacifastacus leniusculus*	754	7
						hypothetical protein	YP_196487	*Ehrlichia ruminantium str. Gardel*	76	1	GF13947 gene product from transcript GF13947-RA	XP_001965778	*Drosophila ananassae*	603	2
						GI22644 gene product from transcript GI22644-RA	XP_002000219	*Drosophila mojavensis*	70	1	uncharacterized protein	XP_003748091	*Metaseiulus occidentalis*	531	2

						hypothetical protein SentesTyp_27217	ZP_03373770	*Salmonella enterica subsp. enterica serovar Typhi str. E98-2068*	70	1	hemocyanin subunit 6	AEG64817	*Eriocheir sinensis*	456	4
						hypothetical protein	EFX67436	*Daphnia pulex*	70	1	serine protease K12H4.7	ACO15310.1	*Caligus clemensi*	447	3
						glutamyl-tRNA synthetase	XP_002085142	*Drosophila simulans*	68	1	ConsensusfromContig37368_full_fwd_frame_0		*Coenobita*	380	3
						ConsensusfromContig38890_full_rev_frame_1		*Coenobita*	66	1					
10C	alpha spectrin	EFX88672	*Daphnia pulex*	296	6	alpha spectrin	XP_002083280	*Drosophila simulans*	617	13	polyubiquitin	XP_550846	*Anopheles gambiae*	12109	30
	ConsensusfromContig87576_full_rev_frame_1		*Coenobita*	66	1	ConsensusfromContig57845_full_fwd_frame_1 (cuticle protein)		*Coenobita*	137	2	alpha spectrin	EEZ99233	*Tribolium castaneum*	4685	32
	ankyrin	AAA85852	*Caenorhabditis elegans*	63	1	hemocyanin subunit 6	AEG64817	*Eriocheir sinensis*	80	1	ConsensusfromContig57845_full_fwd_frame_1		*Coenobita*	1624	2
						UDP-N-acetylglucosamine 2-epimerase	YP_002536919	*Geobacter daltonii FRC-32*	77	1	ConsensusfromContig81144_full_rev_frame_0 (polyubiqitin)		*Coenobita*	1335	3
						cuticle protein	XP_002061935	*Drosophila willistoni*	67	1	ribosomal protein L40	NP_476776	*Drosophila melanogaster*	1335	4
						ConsensusfromContig111333_full_rev_frame_2		*Coenobita*	65	1	ConsensusfromContig89277_full_rev_frame_0 (spectrin)		*Coenobita*	1200	2
						dehydrogenase/reductase SDR family member 11-like	XP_001604944	*Nasonia vitripennis*	64	1	ConsensusfromContig115516_full_rev_frame_1		*Coenobita*	477	1
						ConsensusfromContig38334_full_fwd_frame_2		*Coenobita*	62	1					

Supplementary to manuscript 3: Additional file 2_1D_Control tissue

Spot number	Mascot searching					MS BLAST searching					MSE analysis				
	Protein name	Accession	*Species*	Score	Peptide hits	Protein name	Accession	*Species*	Score	Peptide hits	Protein name	Accession	*Species*	Score	Peptide hits
1E	ribosomal protein rps27	ADY18400	*Glycera tridactyla*	125	4	Ubiquitin-63E	NP_728908	*Drosophila melanogaster*	646	14	Ubiquitin	AAM50562	*Drosophila melanogaster*	264320	50
	ConsensusfromContig27852_full_rev_frame_0 (polyubiquitin)		*Coenobita*	123	4	ConsensusfromContig46299_full_rev_frame_2		*Coenobita*	202	4	ConsensusfromContig27957 full rev frame 0; ConsensusfromContig27090_full_fwd_frame_0		*Coenobita*	20963	6
	ConsensusfromContig46299_full_rev_frame_2		*Coenobita*	101	2	cuticle protein CUT9	ABM54460	*Portunus pelagicus*	195	4	ribosomal protein L40	NP_476776	*Drosophila melanogaster*	18771	6
	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	67	1	ConsensusfromContig27090_full_fwd_frame_0; ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	151	2	40S ribosomal protein	XP_001966954	*Drosophila ananassae*	5158	1
	Histone H2A.x	XP_002170269	*Hydra magnipapillata*	94	2	60S ribosomal protein	ADV40075	*Latrodectus hesperus*	164	3	ConsensusfromContig110270_full_rev_frame_0; ConsensusfromContig31888_full_rev_frame_1		*Coenobita*	3092	3
	ConsensusfromContig28270_full_fwd_frame_0		*Coenobita*	77	1	ConsensusfromContig27852_full_rev_frame_0; ConsensusfromContig81144_full_rev_frame_0 (60S ribosomal protein)		*Coenobita*	164	3	ConsensusfromContig46299_full_rev_frame_2; ConsensusfromContig89274_full_fwd_frame_0		*Coenobita*	2790	2
	ConsensusfromContig110085_full_rev_frame_1 (small nuclear ribonucleoprotein polypeptide G)		*Coenobita*	64	1	ConsensusfromContig89274_full_fwd_frame_0		*Coenobita*	99	2	ConsensusfromContig115053_full_rev_frame_0		*Coenobita*	2178	3
	ConsensusfromContig115053_full_rev_frame_0		*Coenobita*	63	1	ConsensusfromContig80945_full_rev_frame_2		*Coenobita*	82	1					
	ribosomal protein S28	CAA10101	*Prunus persica*	61	1	ConsensusfromContig115053_full_rev_frame_0		*Coenobita*	79	1					
	integral membrane protein	XP_001933871	*Pyrenophora tritici-repentis*	57	1	40S ribosomal protein	EFN78565	*Harpegnathos saltator*	79	1					

	ConsensusfromContig48891_full_fwd_frame_1; ConsensusfromContig54921_full_fwd_frame_2; ConsensusfromContig73166_full_fwd_frame_2; ConsensusfromContig101421_full_fwd_frame_1; ConsensusfromContig89438_full_rev_frame_2; ConsensusfromContig102248_full_rev_frame_2		*Coenobita*	52	1	Succinate-semialdehyde dehydrogenase	ACO11208	*Caligus rogercresseyi*	63	1					
						ConsensusfromContig78130_full_fwd_frame_2; ConsensusfromContig81297_full_rev_frame_2		*Coenobita*	63	1					
2E	Histone 2A	CAF98839	*Tetraodon nigroviridis*	253	4	Histone H4	XP_001967756	*Drosophila ananassae*	464	8	Histone 2A	XP_002077543	*Drosophila simulans*	32801	4
	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	120	2	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	449	9	Ubiquitin	AAM50562	*Drosophila melanogaster*	20583	15
	ConsensusfromContig108410_full_rev_frame_2		*Coenobita*	85	2	cuticle protein CUT5	ABM54458	*Portunus pelagicus*	398	8	ConsensusfromContig108974_full_fwd_frame_1 (Histone 2A)		*Coenobita*	18068	2
	ConsensusfromContig31518_full_rev_frame_2		*Coenobita*	80	1	Histone H2B	XP_001604919	*Nasonia vitripennis*	207	4	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	10787	3
	ConsensusfromContig46299_full_rev_frame_2		*Coenobita*	72	1	ConsensusfromContig46299_full_rev_frame_2		*Coenobita*	226	3	Sarcoplasmic calcium binding protein	P86909	*Chionoecetes opilio*	6273	1
	ConsensusfromContig43708_full_rev_frame_2 (Cytochrome c oxidase subunit 5A)		*Coenobita*	65	1	ConsensusfromContig78959_full_rev_frame_0		*Coenobita*	189	4	ConsensusfromContig69829_full_rev_frame_2		*Coenobita*	2703	5
	ConsensusfromContig69829_full_rev_frame_2		*Coenobita*	65	2	ConsensusfromContig31518_full_rev_frame_2		*Coenobita*	162	3	ConsensusfromContig80655_full_rev_frame_0 (cellular retinoic acid retinol binding protein)		*Coenobita*	2416	3
	ConsensusfromContig78959_full_rev_frame_0		*Coenobita*	60	1	ConsensusfromContig42378_full_rev_frame_1		*Coenobita*	118	2	Histone H2B	P83863	*Litopenaeus vannamei*	2111	2
	ConsensusfromContig90850_full_rev_frame_1		*Coenobita*	60	1	ConsensusfromContig54479_full_fwd_frame_1		*Coenobita*	97	1	fatty acid binding protein	ADM64456	*Eriocheir sinensis*	1752	2

	ConsensusfromContig78130_full_fwd_frame_2		*Coenobita*	56	1	ConsensusfromContig90850_full_rev_frame_1		*Coenobita*	82	1	ConsensusfromContig46924_full_rev_frame_0		*Coenobita*	1711	1
	ConsensusfromContig46924_full_rev_frame_0		*Coenobita*	51	1	ConsensusfromContig8357_full_fwd_frame_0		*Coenobita*	82	1	ConsensusfromContig78130_full_fwd_frame_2		*Coenobita*	1498	1
											ribosomal protein L40	NP_476776	*Drosophila melanogaster*	1522	2
											ConsensusfromContig27852 full rev frame 0 (ribosomal protein L40)		*Coenobita*	1522	2
											ConsensusfromContig81144 full rev frame 0		*Coenobita*	1428	1
											ConsensusfromContig31518_full_rev_frame_2		*Coenobita*	1307	1
											ConsensusfromContig81109_full_fwd_frame_1		*Coenobita*	1123	1
											Insect cuticle protein	P81384	*Homarus americanus*	952	2
											ConsensusfromContig42378_full_rev_frame_1		*Coenobita*	889	2
											ConsensusfromContig43708_full_rev_frame_2 (Cytochrome c oxidase subunit 5A mitochondrial)		*Coenobita*	852	4
											ConsensusfromContig78959_full_rev_frame_0		*Coenobita*	614	1
											ConsensusfromContig80818_full_fwd_frame_0		*Coenobita*	584	1
											ConsensusfromContig110061_full_rev_frame_1		*Coenobita*	534	2
											ConsensusfromContig108788_full_fwd_frame_0		*Coenobita*	482	1
											PaaI_thioesterase	XP_002067844	*Drosophila willistoni*	476	3
											ConsensusfromContig80581_full_fwd_frame_1		*Coenobita*	430	4
											Glycosyltransferase_GTB_type	AAY55589	*Drosophila melanogaster*	331	3
											Cytochrome c oxidase subunit 5A, mitochondrial	EFN75078	*Harpegnathos saltator*	315	4

											UDP-glucoronosyl and UDP-glucosyl transferase	NP_652629	*Drosophila melanogaster*	300	3
3E	ConsensusfromContig80655_full_rev_frame_0		*Coenobita*	190	3	Histone 2A	XP_003746454	*Metaseiulus occidentalis*	210	4	histone H3	XP_003746454	*Metaseiulus occidentalis*	26042	20
	40S ribosomal protein	BAB78484	*Marsupenaeus japonicus*	92	2	tRNA (guanine-N7-)-methyltransferase	YP_002931468	*Eubacterium eligens ATCC 27750*	186	4	ConsensusfromContig80655_full_rev_frame_0 (fatty acid binding protein)		*Coenobita*	7721	7
	ConsensusfromContig81323_full_fwd_frame_1		*Coenobita*	79	1	Nucleoside diphosphate kinase Group I	EIE86357	*Rhizopus delemar RA 99-880*	189	4	Sarcoplasmic calcium binding protein	P86909	*Chionoecetes opilio*	6433	1
	ConsensusfromContig80987_full_rev_frame_0		*Coenobita*	76	1	ConsensusfromContig81323_full_fwd_frame_1		*Coenobita*	163	2	histone H2B	XP 003746448	*Metaseiulus occidentalis*	6368	6
	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	73	1	cuticle proprotein proCP2.1	ABR27679	*Callinectes sapidus*	152	3	ConsensusfromContig54204_full_fwd_frame_2		*Coenobita*	5371	2
	formate dehydrogenase subunit alpha	YP_001681325	*Heliobacterium modesticaldum Ice1*	60	1	ConsensusfromContig80655_full_rev_frame_0; ConsensusfromContig114621_full_fwd_frame_0 (Lipocalin / cytosolic fatty-acid binding protein)		*Coenobita*	146	2	cellular retinoic acid/retinol binding protein	ADY80038	*Procambarus clarkii*	5330	3
	Nucleoside diphosphate kinase	P27950	*Ginglymostoma cirratum*	59	1	Lipocalin / cytosolic fatty-acid binding protein	AFN08746	*Scylla paramamosain*	130	2	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	5163	4
	ConsensusfromContig80988_full_rev_frame_2; ConsensusfromContig110638_full_rev_frame_1 (urocanate hydratase)		*Coenobita*	59	1	ConsensusfromContig75335_full_fwd_frame_2		*Coenobita*	122	2	ConsensusfromContig82521_full_rev_frame_1 (nucleoside diphosphate kinase)		*Coenobita*	3587	2
						ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	109	1	ConsensusfromContig81323_full_fwd_frame_1		*Coenobita*	2199	3
						ConsensusfromContig43579_full_rev_frame_1; ConsensusfromContig66356_full_rev_frame_0; ConsensusfromContig88482_full_rev_frame_1		*Coenobita*	80	1	40S ribosomal protein	P48855	*Procambarus clarkii*	2223	4

						ConsensusfromContig54204_full_fwd_frame_2		*Coenobita*	103	2	Nucleoside diphosphate kinase	AFL02665	*Penaeus monodon*	2107	3
						hypothetical protein	YP_001884687	*Clostridium botulinum*	77	1	ConsensusfromContig108788 full fwd frame 0		*Coenobita*	1651	1
						hypothetical protein	YP_006261405	*Deinococcus gobiensis*	76	1	ConsensusfromContig28014 full rev frame 2		*Coenobita*	1264	4
						hypothetical protein	YP_004858389	*Bacillus coagulans*	72	1	ConsensusfromContig28573 full fwd frame 2		*Coenobita*	1206	3
						ConsensusfromContig80987_full_rev_frame_0		*Coenobita*	69	1	ConsensusfromContig43579_full_rev_frame_1 (Cyclophilin_ABH_like)		*Coenobita*	1191	5
						ConsensusfromContig29660_full_fwd_frame_0		*Coenobita*	66	1	ConsensusfromContig47578 full fwd frame 1		*Coenobita*	994	1
						ConsensusfromContig28446_full_rev_frame_0		*Coenobita*	66	1	Cyclophilin_ABH_like	XP_966308	*Tribolium castaneum*	983	1
											ConsensusfromContig108505_full_rev_frame_2 (40S ribosomal protein)		*Coenobita*	978	2
											40S ribosomal protein	ACR78697	*Rimicaris exoculata*	893	2
											ConsensusfromContig80987_full_rev_frame_0		*Coenobita*	829	2
											Paal_thioesterase	XP_002067844	*Drosophila willistoni*	528	1
											ConsensusfromContig80557_full_rev_frame_0		*Coenobita*	512	2
											Translocon-associated protein, delta subunit	AEB54627	*Procambarus clarkii*	508	2
											ConsensusfromContig51846_full_fwd_frame_0		*Coenobita*	494	2
											phosphatidylinositol transfer protein SEC14	XP_002407592	*Ixodes scapularis*	475	2
											ConsensusfromContig27595_full_fwd_frame_0		*Coenobita*	430	1
											ConsensusfromContig43708_full_rev_frame_2		*Coenobita*	428	4
											ConsensusfromContig43674_full_rev_frame_1		*Coenobita*	400	1
											serine protease K12H4.7	ACO15310.1	*Caligus clemensi*	352	2
											Insect cuticle protein	XP_001653421	*Aedes aegypti*	302	2

4E	calcium binding protein P22	AAQ97760	*Danio rerio*	69	1	eukaryotic translation initiation factor 5A	ABI30653	*Penaeus monodon*	220	3	Sarcoplasmic calcium binding protein	P86909	*Chionoecetes opilio*	11184	1
	ConsensusfromContig77555_full_rev_frame_1		*Coenobita*	66	1	ConsensusfromContig27957_full_rev_frame_0 (cuticle protein)		*Coenobita*	187	2	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	2174	3
	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	60	1	ConsensusfromContig52405_full_rev_frame_(eukaryotic translation initiation factor 5A)		*Coenobita*	184	2	ConsensusfromContig89286_full_rev_frame_1		*Coenobita*	1570	2
	ConsensusfromContig80655_full_rev_frame_0 (cellular retinoic acid/retinol binding protein)		*Coenobita*	55	1	ConsensusfromContig30597_full_rev_frame_1		*Coenobita*	168	1	GTPase SAR1 and related small G proteins	XP_968387	*Tribolium castaneum*	1512	3
	cellular retinoic acid/retinol binding protein	AAL68638	*Metapenaeus ensis*	55	1	cuticle protein CUT5	ABM54458	*Portunus pelagicus*	162	3	ConsensusfromContig51846_full_fwd_frame_0		*Coenobita*	807	2
	ConsensusfromContig89294_full_fwd_frame_1 (40S ribosomal protein)		*Coenobita*	54	1	ferritin	ADM26622	*Scylla paramamosain*	142	3	ConsensusfromContig27839_full_rev_frame_0 (GTPase SAR1 and related small G proteins)		*Coenobita*	800	3
	ConsensusfromContig28308_full_fwd_frame_1		*Coenobita*	54	1	Cuticle protein CP1246	P81582	*Cancer pagurus*	127	2	histone H3	XP_003746454	*Metaseiulus occidentalis*	735	4
	ConsensusfromContig27839_full_rev_frame_0 (ADP-ribosylation factor)		*Coenobita*	54	1	ConsensusfromContig80655_full_rev_frame_0 (Lipocalin / cytosolic fatty-acid binding protein)		*Coenobita*	113	2	Ubx protein	AAK58677	*Procambarus clarkii*	728	1
						Lipocalin / cytosolic fatty-acid binding protein	ADM64456	*Eriocheir sinensis*	113	2	ConsensusfromContig54204_full_fwd_frame_2		*Coenobita*	659	1
						ConsensusfromContig27589_full_fwd_frame_2		*Coenobita*	97	1	ConsensusfromContig27589_full_fwd_frame_2		*Coenobita*	521	3
						40S ribosomal protein	ACY66492	*Scylla paramamosain*	78	1	ConsensusfromContig78759_full_rev_frame_0		*Coenobita*	474	3
						ConsensusfromContig89294_full_fwd_frame_1 (40S ribosomal protein)	ACY66492	*Coenobita*	78	1	Pleckstrin homology-like domain	XP_002042404	*Drosophila sechellia*	393	3
						ConsensusfromContig46299_full_rev_frame_2 (cuticle protein)		*Coenobita*	72	1	ConsensusfromContig108478_full_rev_frame_1 (arp2/3 complex 20 kd subunit)		*Coenobita*	385	2
						ConsensusfromContig78130_full_fwd_frame_2		*Coenobita*	64	1	arp2/3 complex 20 kd subunit	ACY66433	*Scylla paramamosain*	320	3
											peptidyl-prolyl cis-trans isomerase H-like	XP_003742389	*Metaseiulus occidentalis*	363	4

5E	ConsensusfromContig80801_full_fwd_frame_0		*Coenobita*	219	4	ConsensusfromContig28493_full_rev_frame_1		*Coenobita*	239	3	Sarcoplasmic calcium binding protein	P86909	*Chionoecetes opilio*	10155	1
	ConsensusfromContig28493_full_rev_frame_1		*Coenobita*	73	1	ConsensusfromContig80801_full_fwd_frame_0		*Coenobita*	238	4	ConsensusfromContig80801 full fwd frame 0		*Coenobita*	3574	8
	ConsensusfromContig109225_full_fwd_frame_2		*Coenobita*	71	2	ConsensusfromContig84076_full_fwd_frame_1		*Coenobita*	203	3	ConsensusfromContig114956 full rev frame 1		*Coenobita*	3524	3
	ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	69	1	calcified cuticle protein CP19.0 isoform A	ABB91674	*Callinectes sapidus*	200	4	ConsensusfromContig67413 full fwd frame 0		*Coenobita*	3156	1
	ConsensusfromContig109676_full_fwd_frame_0		*Coenobita*	55	1	ConsensusfromContig27965_full_rev_frame_1		*Coenobita*	180	3	ConsensusfromContig28493 full rev frame 1		*Coenobita*	2576	6
	ConsensusfromContig114956_full_rev_frame_1		*Coenobita*	55	1	ConsensusfromContig67413_full_fwd_frame_0		*Coenobita*	166	3	ConsensusfromContig27957 full rev frame 0		*Coenobita*	1993	1
						ConsensusfromContig44642_full_rev_frame_2		*Coenobita*	159	3	ConsensusfromContig27965_full_rev_frame_1		*Coenobita*	1239	5
						ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	104	1	ConsensusfromContig80653_full_fwd_frame_0 (crustacyanin A1)		*Coenobita*	1188	2
						ConsensusfromContig109676_full_fwd_frame_0		*Coenobita*	81	1	ADP-ribosylation factor-like protein 3	EHJ74704	*Danaus plexippus*	1105	4
						vesicle coat complex COPII GTPase subunit SAR1	ACY69967	*Cimex lectularius*	72	1	GTPase SAR1 and related small G proteins	EFX85937	*Daphnia pulex*	870	4
						mitochondrial ATP synthase subunit alpha	ACK44331	*Drosophila silvestris*	66	1	ConsensusfromContig109225_full_fwd_frame_2		*Coenobita*	803	2
						Calponin homology domain	NP_001191867	*Acyrthosiphon pisum*	113	3	ConsensusfromContig89286_full_rev_frame_1		*Coenobita*	741	2
											ConsensusfromContig79881_full_fwd_frame_2		*Coenobita*	737	4
											ConsensusfromContig44642_full_rev_frame_2		*Coenobita*	664	2
											ConsensusfromContig82266_full_fwd_frame_0 (GTPase SAR1 and related small G proteins)		*Coenobita*	636	3
											ConsensusfromContig109676_full_fwd_frame_0		*Coenobita*	632	2
											ConsensusfromContig29101_full_rev_frame_2 (ADP-ribosylation factor)		*Coenobita*	615	2
											ConsensusfromContig108330_full_fwd_frame_0		*Coenobita*	596	3
											crustacyanin-A1 precursor	ACL37112	*Panulirus cygnus*	511	1
											Ubiquitin	XP_550846	*Anopheles gambiae*	568	15

											ConsensusfromContig84076_full_fwd_frame_1		*Coenobita*	475	2
											ADP ribosylation factor	XP_966753	*Tribolium castaneum*	470	2
											ConsensusfromContig78584 full rev frame 2		*Coenobita*	369	4
											Chromosome segregation ATPases	XP_002414430	*Ixodes scapularis*	228	2
6E	ConsensusfromContig109750_full_rev_frame_0		*Coenobita*	192	3	ConsensusfromContig80793_full_rev_frame_0		*Coenobita*	560	10	Sarcoplasmic calcium binding protein	P86909	*Chionoecetes opilio*	13521	1
	ConsensusfromContig77737_full_rev_frame_2 (peroxiredoxin)		*Coenobita*	157	3	glutathione S transferase D1	AEB54656	*Procambarus clarkii*	405	8	ConsensusfromContig84076_full_fwd_frame_1; ConsensusfromContig48845_full_rev_frame_2		*Coenobita*	10421	5
	Ras-related protein Rab-2	ACO10816	*Caligus rogercresseyi*	124	2	ConsensusfromContig84076_full_fwd_frame_1		*Coenobita*	339	5	ConsensusfromContig114956_full_rev_frame_1		*Coenobita*	6144	3
	ConsensusfromContig114956_full_rev_frame_1		*Coenobita*	110	2	triosephosphate isomerase	ADN52396	*Eriocheir sinensis*	236	4	ConsensusfromContig77737_full_rev_frame_2 (peroxiredoxin)		*Coenobita*	4615	6
	peroxiredoxin	AAL28430	*Drosophila melanogaster*	109	2	ConsensusfromContig109750_full_rev_frame_0 (triosephosphate isomerase)		*Coenobita*	258	4	ConsensusfromContig56178_full_fwd_frame_1		*Coenobita*	4582	2
	ConsensusfromContig110093_full_fwd_frame_2		*Coenobita*	105	2	Peroxiredoxin	ACF35639	*Eriocheir sinensis*	251	5	Ubiquitin	XP_001845257	*Culex quinquefasciatus*	4411	16
	ConsensusfromContig32512_full_rev_frame_2		*Coenobita*	91	2	ConsensusfromContig14210_full_rev_frame_1 (glutathione S-transferase)		*Coenobita*	246	5	ConsensusfromContig78322_full_rev_frame_2; ConsensusfromContig42540_full_fwd_frame_1; ConsensusfromContig93752_full_rev_frame_1		*Coenobita*	3463	5
	ConsensusfromContig78603_full_rev_frame_0		*Coenobita*	90	2	ConsensusfromContig48845_full_rev_frame_2		*Coenobita*	218	3	Rab GTPase family 1	AGC10822	*Eriocheir sinensis*	3181	4
	ConsensusfromContig56178_full_fwd_frame_1		*Coenobita*	84	1	ConsensusfromContig110093_full_fwd_frame_2 (Peroxiredoxin)		*Coenobita*	212	3	ConsensusfromContig80793_full_rev_frame_0		*Coenobita*	2944	9
	ConsensusfromContig80801_full_fwd_frame_0		*Coenobita*	78	2	ConsensusfromContig77737_full_rev_frame_2 (Peroxiredoxin)		*Coenobita*	197	3	Putative rab35 member ras oncoprotein family	JAA54907	*Rhipicephalus pulchellus*	2672	7
	ConsensusfromContig80793_full_rev_frame_0		*Coenobita*	74	2	ATP synthase delta	EFR22427	*Anopheles darlingi*	151	3	Peroxiredoxin	XP_002066945	*Drosophila willistoni*	2581	4

	ConsensusfromContig64301_full_rev_frame_0		*Coenobita*	73	1	ConsensusfromContig78022_full_rev_frame_1		*Coenobita*	146	3	ConsensusfromContig110093_full_fwd_frame_2 (cytosolic thioredoxin peroxidase)		*Coenobita*	2313	3
	ConsensusfromContig48845_full_rev_frame_2		*Coenobita*	62	1	ConsensusfromContig47371_full_rev_frame_1		*Coenobita*	136	2	ConsensusfromContig81907_full_rev_frame_2		*Coenobita*	2178	1
	ConsensusfromContig80518_full_fwd_frame_2		*Coenobita*	56	1	ConsensusfromContig29432_full_rev_frame_0 (Peroxiredoxin)		*Coenobita*	135	2	Ras-related protein Rab-10	ACO10653	*Caligus rogercresseyi*	2008	4
						ConsensusfromContig45170_full_fwd_frame_1 (glutathione S transferase)		*Coenobita*	118	2	ConsensusfromContig82277_full_rev_frame_2 (Rab family protein)		*Coenobita*	1917	3
						Ras-related protein Rab-2	ACO13165	*Lepeophtheirus salmonis*	99	1	ConsensusfromContig32512_full_rev_frame_2		*Coenobita*	1736	5
						ConsensusfromContig64301_full_rev_frame_0		*Coenobita*	90	1	ConsensusfromContig93752_full_rev_frame_1; ConsensusfromContig42540_full_fwd_frame_1; ConsensusfromContig78322_full_rev_frame_2		*Coenobita*	1618	3
						ConsensusfromContig83879_full_rev_frame_1 (Ras-related protein Rab-2)		*Coenobita*	86	1	ConsensusfromContig80801_full_fwd_frame_0		*Coenobita*	1381	7
						ConsensusfromContig56178_full_fwd_frame_1		*Coenobita*	84	1	ConsensusfromContig64301_full_rev_frame_0		*Coenobita*	1355	1
						ConsensusfromContig44642_full_rev_frame_2		*Coenobita*	84	1	ConsensusfromContig85384_full_rev_frame_2; ConsensusfromContig48845_full_rev_frame_2; ConsensusfromContig84076_full_fwd_frame_1		*Coenobita*	1336	2
						ConsensusfromContig85384_full_rev_frame_2		*Coenobita*	82	1	ConsensusfromContig28646_full_rev_frame_0		*Coenobita*	1277	3
						ConsensusfromContig112901_full_fwd_frame_2 (ATP synthase)		*Coenobita*	80	1	Crustacean calcium-binding protein 23	P80364	*Homarus americanus*	1261	3
						ConsensusfromContig114956_full_rev_frame_1		*Coenobita*	131	2	ConsensusfromContig109750_full_rev_frame_0 (triosephosphate isomerase)		*Coenobita*	1143	6
						ConsensusfromContig4764_full_rev_frame_0		*Coenobita*	74	1	ConsensusfromContig80510_full_rev_frame_0		*Coenobita*	1082	2
											triosephosphate isomerase	ABB81879	*Fenneropenaeus chinensis*	1022	2

											ConsensusfromContig29432_full_rev_frame_0		*Coenobita*	1014	3
											ConsensusfromContig28493_full_rev_frame_1		*Coenobita*	890	3
											ConsensusfromContig81037_full_rev_frame_0		*Coenobita*	860	1
											ConsensusfromContig44642_full_rev_frame_2		*Coenobita*	822	2
											proteasome subunit alpha	XP_314945	*Anopheles gambiae*	682	3
											ConsensusfromContig27957_full_rev_frame_0		*Coenobita*	651	2
											ConsensusfromContig45170_full_fwd_frame_1		*Coenobita*	636	1
											ConsensusfromContig83346_full_rev_frame_0; ConsensusfromContig80793_full_rev_frame_0		*Coenobita*	511	5
											glutathione S transferase	AAQ74442	*Rhipicephalus appendiculatus*	469	3
											ConsensusfromContig77926_full_rev_frame_2 (mitochondrial manganese superoxide dismutase)		*Coenobita*	420	2
											mitochondrial manganese superoxide dismutase	AAZ81617	*Macrobrachium rosenbergii*	375	2
											odorant binding protein 13 partial	CAX63070	*Acyrthosiphon pisum*	313	4
											ConsensusfromContig81144_full_rev_frame_0 (Ubiquitin)		*Coenobita*	305	3
											ConsensusfromContig35747_full_rev_frame_1		*Coenobita*	264	2
											ConsensusfromContig78015_full_fwd_frame_1		*Coenobita*	247	2
											serine protease K12H4 7	ACO15310	*Caligus clemensi*	196	1
7E	cytosolic manganese superoxide dismutase	AEK77429	*Macrobrachium nipponens*	210	3	EF-hand, calcium binding motif	EJY71683	*Oxytricha trifallax*	565	12	cytosolic manganese superoxide dismutase	AAO42751	*Palaemonetes pugio*	16828	4

			e												
	14-3-3 protein	AFD28274	*Scylla paramamosain*	187	3	cytoplasmic manganese superoxide dismutase	CAR85671	*Xantho poressa*	503	10	14 3 3 zeta	AFD33362	*Scylla paramamosain*	16740	6
	ConsensusfromContig109750_full_rev_frame_0 (triosephosphate isomerase)		*Coenobita*	119	2	14-3-3 protein	AFQ20813	*Portunus trituberculatus*	359	6	ConsensusfromContig85153_full_rev_frame_0 (cytosolic manganese superoxide dismutase)		*Coenobita*	11847	4
	ConsensusfromContig32822_full_rev_frame_0 (OET17, calcyphosine)		*Coenobita*	90	1	olfactory enriched transcipt 17.17	AAL04110	*Homarus americanus*	283	4	ConsensusfromContig32822_full_rev_frame_0		*Coenobita*	6600	2
	ConsensusfromContig29432_full_rev_frame_0 (peroxiredoxin)		*Coenobita*	82	2	ConsensusfromContig85153_full_rev_frame_0		*Coenobita*	223	3	ConsensusfromContig56178_full_fwd_frame_1		*Coenobita*	5046	3
	ConsensusfromContig114956_full_rev_frame_1		*Coenobita*	78	1	thioredoxin-dependent peroxide reductase	XP_001948034	*Acyrthosiphon pisum*	195	4	ConsensusfromContig28051_full_rev_frame_0		*Coenobita*	4095	1
	ConsensusfromContig80571_full_rev_frame_1 (phosphoglycerate mutase)		*Coenobita*	76	2	ConsensusfromContig32822_full_rev_frame_0 (EF-hand, calcium binding motif)		*Coenobita*	174	2	ConsensusfromContig77737_full_rev_frame_2 (Peroxiredoxin)		*Coenobita*	3965	6
	ConsensusfromContig56178_full_fwd_frame_1		*Coenobita*	74	1	Peroxiredoxin-6	EGI62739	*Acromyrmex echinatior*	166	3	ConsensusfromContig114956_full_rev_frame_1		*Coenobita*	3441	3
	oxygen-independent coproporphyrinogen III oxidase	ZP_01302555	*Sphingomonas sp.*	67	1	ConsensusfromContig17144_full_fwd_frame_1; ConsensusfromContig90974_full_rev_frame_2		*Coenobita*	156	2	ConsensusfromContig44775_full_rev_frame_2 (14-3-3 protein)		*Coenobita*	2616	1
	ConsensusfromContig35747_full_rev_frame_1 (Endoplasmic reticulum protein ERp29)		*Coenobita*	65	1	ConsensusfromContig56178_full_fwd_frame_1		*Coenobita*	150	2	ConsensusfromContig109750_full_rev_frame_0 (triosephosphate isomerase)		*Coenobita*	2575	5
	ConsensusfromContig80518_full_fwd_frame_2 (Glutathione-S-Transferase)		*Coenobita*	60	1	ConsensusfromContig28613_full_rev_frame_1 (EF-hand, calcium binding motif)		*Coenobita*	122	2	ConsensusfromContig35747_full_rev_frame_1; ConsensusfromContig26433_full_rev_frame_0		*Coenobita*	2379	5
	CoA-binding domain-containing protein	YP_001769462	*Methylobacterium sp.*	60	1	ConsensusfromContig109750_full_rev_frame_0 (triosephosphate isomerase)		*Coenobita*	121	2	ConsensusfromContig110093_full_fwd_frame_2 (Peroxiredoxin)		*Coenobita*	2144	3
	ConsensusfromContig77534_full_rev_frame_0 (NADH:ubiquinone oxidoreductase)		*Coenobita*	59	1	ConsensusfromContig80793_full_rev_frame_0		*Coenobita*	115	2	Peroxiredoxin	XP_002077958	*Drosophila simulans*	2148	5
	FK506-binding protein	EEH18592	*Paracoccidioides brasiliensis*	57	1	elongation factor 1 alpha	ADC97984	*Aulacorthum albimagnoliae*	114	2	ConsensusfromContig110086_full_rev_frame_1		*Coenobita*	1718	2

	ConsensusfromContig80793_full_rev_frame_0 (Glutathione-S-transferase)		*Coenobita*	56	1	ConsensusfromContig44775_full_rev_frame_2		*Coenobita*	111	2	GTP-binding nuclear protein Ran	XP_564524	*Anopheles gambiae*	1639	3
						ConsensusfromContig48845_full_rev_frame_2; ConsensusfromContig47371_full_rev_frame_1; ConsensusfromContig84076_full_fwd_frame_1		*Coenobita*	87	1	ConsensusfromContig108678_full_fwd_frame_0 (prohibtin protein)		*Coenobita*	1589	7
						triosephosphate isomerase	ADN52396	*Eriocheir sinensis*	79	1	ConsensusfromContig26433_full_rev_frame_0; ConsensusfromContig35747_full_rev_frame_1		*Coenobita*	1510	4
						ConsensusfromContig35747_full_rev_frame_1		*Coenobita*	78	1	ConsensusfromContig80518_full_fwd_frame_2		*Coenobita*	1408	5
						ConsensusfromContig35858_full_rev_frame_0		*Coenobita*	78	1	ConsensusfromContig29432_full_rev_frame_0		*Coenobita*	1203	2
						ConsensusfromContig41255_full_rev_frame_1		*Coenobita*	70	1	Histidine phosphatase	EFX86204	*Daphnia pulex*	1177	4
						tRNA 2-selenouridine synthase	ZP_08959355	*Halomonas sp.*	75	1	ConsensusfromContig80571_full_rev_frame_1		*Coenobita*	1086	1
						ConsensusfromContig50037_full_rev_frame_0; ConsensusfromContig3249_full_rev_frame_0		*Coenobita*	71	1	ConsensusfromContig84140_full_rev_frame_1 (Histidine phosphatase)		*Coenobita*	1067	2
											triosephosphate isomerase	ADN52396	*Eriocheir sinensis*	1034	4
											ConsensusfromContig84076_full_fwd_frame_1; ConsensusfromContig85384_full_rev_frame_2		*Coenobita*	1020	4
											ConsensusfromContig64301_full_rev_frame_0		*Coenobita*	894	1
											ConsensusfromContig3249_full_rev_frame_0;K ConsensusfromContig50037_full_rev_frame_0		*Coenobita*	886	1
											ConsensusfromContig79932_full_fwd_frame_1		*Coenobita*	820	3
											ConsensusfromContig81907_full_rev_frame_2		*Coenobita*	761	2
											voltage-dependent anion-selective channel	ADJ94951	*Eriocheir sinensis*	720	6
											ConsensusfromContig77534_full_rev_frame_0 (NADH ubiquinone oxidoreductase)		*Coenobita*	691	5

											ConsensusfromContig50037_full_rev_frame_0; ConsensusfromContig3249_full_rev_frame_0		*Coenobita*	552	1
											Protein l(2)37Cc	EFN69844	*Camponotus floridanus*	529	5
											Tropomyosin	EFZ16118	*Solenopsis invicta*	527	5
											ConsensusfromContig85384_full_rev_frame_2; ConsensusfromContig84076_full_fwd_frame_1		*Coenobita*	479	2
											Phosphoglycerate mutase 2	ACO14734	*Caligus clemensi*	450	3
											ConsensusfromContig28613_full_rev_frame_1		*Coenobita*	416	2
											NADH:ubiquinone oxidoreductase NDUFV2/24 kDa	EFX75440	*Daphnia pulex*	384	3
											ConsensusfromContig78116_full_rev_frame_1		*Coenobita*	383	2
											ConsensusfromContig43585_full_fwd_frame_1		*Coenobita*	309	2
											Putative 20s proteasome regulatory subunit alpha	JAA59228	*Rhipicephalus pulchellus*	247	3
8E	14-3-3 protein	AFD28274	*Scylla paramamosain*	205	3	14-3-3 zeta	ADI87601	*Fenneropenaeus merguiensis*	447	8	glyceraldehyde 3 phosphate dehydrogenase	ADU87475	*Neopetrolisthes maculatus*	20854	10
	cytosolic manganese superoxide dismutase	AEK77429	*Macrobrachium nipponense*	154	3	cytosolic manganese superoxide dismutase	ADB90400	*Marsupenaeus japonicus*	410	8	ConsensusfromContig9100_full_rev_frame_2 (glyceraldehyde 3 phosphate dehydrogenase)		*Coenobita*	10225	2
	voltage-dependent anion-selective channel	ADJ94951	*Eriocheir sinensis*	101	2	ConsensusfromContig85153_full_rev_frame_0		*Coenobita*	264	4	ConsensusfromContig78044_full_rev_frame_1 (malate dehydrogenase)		*Coenobita*	6940	6
	ConsensusfromContig29592_full_rev_frame_0 (carbonyl rductase)		*Coenobita*	98	2	tropomyosin	XP_001999696	*Drosophila mojavensis*	226	5	14 3 3 protein	AFD28274	*Scylla paramamosain*	3995	6
	ConsensusfromContig80571_full_rev_frame_1 (phosphoglycerate mutase)		*Coenobita*	84	2	ConsensusfromContig28984_full_fwd_frame_2 (14-3-3 zeta)		*Coenobita*	171	4	malate dehydrogenase	ACY66481	*Scylla paramamosain*	3903	4

	tropomyosin	AAC48288	*Homarus americanus*	77	1	Voltage-dependent anion channel	EFX77112	*Daphnia pulex*	172	3	arginine kinase	NP_001011603	*Megachile rotundata*	3644	5
	ConsensusfromContig28984_full_fwd_frame_0 (14-3-3 protein)		*Coenobita*	57	1	ConsensusfromContig35858_full_rev_frame_0 (14-3-3 zeta)		*Coenobita*	151	2	lactate dehydrogenase	XP_002007947	*Drosophila mojavensis*	3085	5
						NADH-ubiquinone dehydrogenase	XP_002402024	*Ixodes scapularis*	118	2	tropomyosin	AAT40866	*Tyrophagus putrescentiae*	2910	5
						triosephosphate isomerase	ADN52396	*Eriocheir sinensis*	113	2	ConsensusfromContig81485_full_fwd_frame_2 (cytosolic malate dehydrogenase)		*Coenobita*	2492	4
						malate dehydrogenase	XP_002006381	*Drosophila mojavensis*	106	2	Ecdysone-inducible gene L3	NP_476581	*Drosophila melanogaster*	2383	4
						ConsensusfromContig78044_full_rev_frame_1 (malate dehydrogenase)		*Coenobita*	106	2	Ribosomal protein L10 family	ACY66540	*Scylla paramamosain*	2218	3
						ConsensusfromContig109750_full_rev_frame_0 (triosephosphate isomerase)		*Coenobita*	87	1	ConsensusfromContig32852 full fwd frame 0		*Coenobita*	2088	2
						ConsensusfromContig44775_full_rev_frame_2 (14-3-3 zeta)		*Coenobita*	79	1	cytosolic manganese superoxide dismutase	AEK77429	*Macrobrachium nipponense*	1970	3
						ConsensusfromContig77534_full_rev_frame_0 (NADH-ubiquinone dehydrogenase)		*Coenobita*	77	1	ConsensusfromContig85153_full_rev_frame_0 (cytosolic manganese superoxide dismutase)		*Coenobita*	1915	2
						ConsensusfromContig84273_full_fwd_frame_2 (Voltage-dependent anion channel)		*Coenobita*	68	1	ribosomal protein S3	ACR78696	*Rimicaris exoculata*	1877	4
											ConsensusfromContig31896_full_rev_frame_2		*Coenobita*	1640	3
											receptor for activated protein kinase c1	ABU49887	*Penaeus monodon*	1199	6
											ConsensusfromContig35858_full_rev_frame_0 (14 3 3 like protein)		*Coenobita*	1182	2
											ribosomal phosphoprotein P0	DAA34132	*Amblyomma variegatu*	957	3

													m		
											60S acidic ribosomal protein P0	ADV40160	*Latrodectus hesperus*	922	2
											40S ribosomal protein S3	XP_001599651	*Nasonia vitripennis*	944	4
											ConsensusfromContig52669_full_fwd_frame_1 (glyceraldehyde 3 phosphate dehydrogenase)		*Coenobita*	826	1
											ConsensusfromContig31520_full_rev_frame_1 (receptor for activated protein kinase c1)		*Coenobita*	822	2
											ConsensusfromContig56178 full fwd frame 1		*Coenobita*	725	1
											ConsensusfromContig44775_full_rev_frame_2 (14 3 3 like protein)		*Coenobita*	716	2
											heat shock protein 90	AGC54636	*Scylla paramamosain*	655	6
											ConsensusfromContig44336_full_rev_frame_1		*Coenobita*	618	4
											ConsensusfromContig110086_full_rev_frame_1		*Coenobita*	590	1
											ConsensusfromContig2084_full_fwd_frame_1		*Coenobita*	549	5
											ConsensusfromContig78277_full_fwd_frame_0		*Coenobita*	548	2
											ConsensusfromContig81844_full_rev_frame_0 (heat shock protein 90)		*Coenobita*	508	1
											ConsensusfromContig23991_full_fwd_frame_0		*Coenobita*	444	1
											carotenoid binding protein	BAE44463	*Bombyx mori*	394	3
											Putative copper zinc superoxide dismutase	JAA55780	*Rhipicephalus pulchellus*	332	2
											3-hydroxyacyl-CoA dehydrogenase	EFX73881	*Daphnia pulex*	294	3
											serine protease K12H4 7	ACO15310	*Caligus clemensi*	230	3

											60S ribosomal protein l13	ABX75438	*Lycosa singoriensis*	214	3
9E	glyceraldehyde-3-phosphate dehydrogenase	ACK56132	*Portunus trituberculatu*	189	4	glyceraldehyde-3-phosphate dehydrogenase	YP_002430524	*Desulfatibacillum alkenivorans*	454	10	glyceraldehyde-3-phosphate dehydrogenase	ADU87475.1	*Neopetrolisthes maculatus*	20854	10
	ConsensusfromContig78044_full_rev_frame_1 (malate dehydrogenase)		*Coenobita*	161	3	Tropomyosin	XP_001999696	*Drosophila mojavensis*	347	7	ConsensusfromContig9100_full_rev_frame_2 (GAPDH)		*Coenobita*	10225	2
	receptor for activated protein kinase c1	ABU49887	*Penaeus monodon*	76	2	ConsensusfromContig78044_full_rev_frame_1 (malate dehydrogenase)		*Coenobita*	274	4	ConsensusfromContig78044_full_rev_frame_1 (malate dehydrogenase)		*Coenobita*	6940	6
	ConsensusfromContig85153_full_rev_frame_0 (manganese superoxide dismutase)		*Coenobita*	72	1	malate dehydrogenase	AEL23074	*Cherax quadricarinatus*	249	4	arginine kinase	XP_003695937.1	*Apis florea*	5881	6
	tropomyosin	AAA28966	*Drosophila melanogaster*	67	1	ConsensusfromContig85153_full_rev_frame_0		*Coenobita*	139	2	14-3-3 protein	AFD28274.1	*Scylla paramamosain*	3995	6
	ConsensusfromContig29592_full_rev_frame_0 (carbonyl reductase)		*Coenobita*	64	1	ConsensusfromContig110404_full_fwd_frame_0		*Coenobita*	116	2	manganese superoxide dismutase	AFD61666.1	*Cherax quadricarinatus*	3293	3
	ConsensusfromContig45714_full_rev_frame_0 (sodium potassium-transporting ATPase)		*Coenobita*	57	1	ConsensusfromContig52669_full_fwd_frame_1 (glyceraldehyde-3-phosphate dehydrogenase)		Coenobita	114	1	lactate dehydrogenase	ABB36437.1	*Drosophila melanogaster*	3137	3
	ConsensusfromContig110087_full_rev_frame_1 (thioredoxin)		*Coenobita*	56	1	ConsensusfromContig45714_full_rev_frame_0		*Coenobita*	107	1	tropomyosin	AAT40866.1	*Tyrophagus putrescentiae*	2910	5
	ConsensusfromContig32852_full_fwd_frame_0 (4-hydroxy-2-oxoglutarate aldolase)		*Coenobita*	55	1	ConsensusfromContig44336_full_rev_frame_1		*Coenobita*	82	1	ConsensusfromContig81485_full_fwd_frame_2 (malate dehydrogenase)		*Coenobita*	2492	4
						ConsensusfromContig110087_full_rev_frame_1		*Coenobita*	69	1	ConsensusfromContig32852_full_fwd_frame_0		*Coenobita*	2087	2
											ConsensusfromContig85153_full_rev_frame_0 (manganese superoxide dismutase)		*Coenobita*	1915	2
											ribosomal protein S3	ACR78696.1	*Rimicaris exoculata*	1877	4

											ConsensusfromContig47491_full_rev_frame_2 (60S acidic ribosomal protein)		*Coenobita*	1694	3
											ConsensusfromContig31896_full_rev_frame_2		*Coenobita*	1640	3
											14 3 3 like protein	AAY56092.1	*Penaeus monodon*	1468	4
											ConsensusfromContig35858_full_rev_frame_0 (14-3-3 protein)		*Coenobita*	1181	2
											60S acidic ribosomal protein	DAA34132.1	*Amblyomma variegatum*	957	3
											ConsensusfromContig52669_full_fwd_frame_1 (GAPDH)		*Coenobita*	825	1
											ConsensusfromContig31520_full_rev_frame_1 (receptor for activated protein kinase C)		*Coenobita*	822	2
											receptor for activated protein kinase C	EFX70343.1	*Daphnia pulex*	729	5
											ConsensusfromContig56178_full_fwd_frame_1		*Coenobita*	724	1
											ConsensusfromContig44775_full_rev_frame_2 (14-3-3 protein)		*Coenobita*	715	2
											hsp90	AGC54636.1	*Scylla paramamosain*	655	6
											ConsensusfromContig44336_full_rev_frame_1		*Coenobita*	618	4
											malate dehydrogenase	AEP03089.1	*Miletus boisduvali*	594	2
											ConsensusfromContig110086_full_rev_frame_1		*Coenobita*	589	1
											ConsensusfromContig78277_full_fwd_frame_0		*Coenobita*	548	2
											ConsensusfromContig81844_full_rev_frame_0 (hsp90)		*Coenobita*	507	1
											ConsensusfromContig23991_full_fwd_frame_0		*Coenobita*	444	1
											carotenoid binding protein	NP_001036998.1	*Bombyx mori*	427	4
10E	arginine kinase	NP_0010	*Apis*	22	4	GAPDH	EIE81131.	*Rhizopus*	42	9	arginine kinase		*Apis*	333	11

		11603	*mellifera*	8			1	*delemar*	9			AAC39040.1	*mellifera*	42	
	glyceraldehyde-3-phosphate dehydrogenase	ACK56132	*Portunus trituberculatus*	185	3	arginine kinase	XP_002046272.1	*Drosophila virilis*	368	7	GAPDH	P00357	*Homarus americanus*	12601	7
	ConsensusfromContig30565_full_rev_frame_1 (hydroxyacid oxidase)		*Coenobita*	147	3	ConsensusfromContig52669_full_fwd_frame_1 (GAPDH)		*Coenobita*	114	1	ConsensusfromContig9100_full_rev_frame_2 (GAPDH)		*Coenobita*	5946	2
	tropomyosin	AAC48288	*Homarus americanus*	129	2	ConsensusfromContig2084_full_fwd_frame_1		*Coenobita*	114	2	ConsensusfromContig108959_full_rev_frame_0		*Coenobita*	5182	4
	ConsensusfromContig78044_full_rev_frame_1 (malate dehydrogenase)		*Coenobita*	75	1	ConsensusfromContig30669_full_fwd_frame_2		*Coenobita*	97	2	ConsensusfromContig78044_full_rev_frame_1 (malate dehydrogenase)		*Coenobita*	3341	7
	ConsensusfromContig28277_full_fwd_frame_0 (translation initiation factor 2)		*Coenobita*	67	1	ConsensusfromContig27788_full_rev_frame_0 (translation initiation factor 3)		*Coenobita*	85	1	lactate dehydrogenase	EFX89380.1	*Daphnia pulex*	2666	2
	ConsensusfromContig27788_full_rev_frame_0 (translation initiation factor 3)		*Coenobita*	64	1	ConsensusfromContig35401_full_rev_frame_0		*Coenobita*	74	1	ConsensusfromContig27836_full_fwd_frame_2		*Coenobita*	2512	4
	G protein alpha subunit	AAA28585	*Drosophila melanogaster*	57	2	ConsensusfromContig78044_full_rev_frame_1		*Coenobita*	74	1	ConsensusfromContig30565_full_rev_frame_1		*Coenobita*	1950	5
						ConsensusfromContig81379_full_fwd_frame_2		*Coenobita*	69	1	ConsensusfromContig25049_full_rev_frame_2		*Coenobita*	1685	2
						ConsensusfromContig37036_full_fwd_frame_2		*Coenobita*	60	1	tropomyosin	EFN72212.1	*Camponotus floridanus*	1613	8
						ConsensusfromContig37064_full_fwd_frame_0; ConsensusfromContig65291_full_rev_frame_1; ConsensusfromContig44572_full_fwd_frame_1		*Coenobita*	60	1	RNA binding protein squid	NM_001260159.2	*Drosophila melanogaster*	1592	4
						ConsensusfromContig28277_full_fwd_frame_0		*Coenobita*	60	1	ConsensusfromContig29513_full_rev_frame_0 (RNA binding protein squid)		*Coenobita*	1568	2
						ConsensusfromContig31221_full_fwd_frame_1		*Coenobita*	60	1	ConsensusfromContig633_full_fwd_frame_1 (RNA binding protein squid)		*Coenobita*	1568	2
											ConsensusfromContig2084_full_fwd_frame_1		*Coenobita*	1230	6
											ConsensusfromContig52669_full_fwd_frame_1 (GAPDH)		*Coenobita*	991	1
											ConsensusfromContig79931_full_fwd_frame_2		*Coenobita*	909	3

											laminin receptor	ABH10628.1	*Litopenaeus vannamei*	801	2
											ConsensusfromContig30221_full_rev_frame_0		*Coenobita*	761	5
											transaldolase	H9HGA6	*Atta cephalotes*	645	4
											ConsensusfromContig88167_full_rev_frame_0		*Coenobita*	613	2
											ConsensusfromContig83931_full_rev_frame_1		*Coenobita*	604	2
											Short chain specific acyl CoA dehydrogenase	EGI57190.1	*Acromyrmex echinatior*	570	4
											malate dehydrogenase	AAY63978.1	*Lysiphlebus testaceipes*	533	2
											hsp20	ACR53995.1	*Drosophila melanogaster*	458	3
											ConsensusfromContig78006_full_fwd_frame_2 (laminin receptor)		*Coenobita*	386	3
											ConsensusfromContig110329_full_rev_frame_1		*Coenobita*	280	2
11E	actin	ACD37723	*Cicer arietinum*	210	4	actin	EIW81961.1	*Coniophora puteana*	1059	18	actin	ADG45302.1	*Marsupenaeus japonicus*	52381	24
	arginine kinase	NP_001011603	*Apis mellifera*	200	3	Isocitrate DH	XP_003399046.1	*Bombus terrestris*	458	9	ConsensusfromContig35530_full_rev_frame_2		*Coenobita*	5446	3
	ConsensusfromContig86526_full_rev_frame_0 (glutamine synthetase)		*Coenobita*	120	2	aldolase	XP_001851198.1	*Culex quinquefasciatus*	422	9	arginine kinase	AAF43437.1	*Eriocheir sinensis*	5100	12
	ribosome-associated protein	BAB78527	*Bombyx mori*	99	1	ConsensusfromContig86526_full_rev_frame_0		*Coenobita*	205	4	ConsensusfromContig49799_full_rev_frame_2		*Coenobita*	4805	2
	ConsensusfromContig32340_full_fwd_frame_0 (isocitrate dehydrogenase)		*Coenobita*	98	1	ConsensusfromContig36749_full_rev_frame_0 (isocitrate DH)		*Coenobita*	162	3	glutamine synthetase	AGA83299.1	*Penaeus monodon*	3592	5
	ConsensusfromContig35530_full_rev_frame_2; ConsensusfromContig57240_full_rev_frame_0 (juvenile hormone inducible protein)		*Coenobita*	87	2	ConsensusfromContig32340_full_fwd_frame_0 (isocitrate DH)		*Coenobita*	146	2	EF2	XP_001354162.1	*Drosophila pseudoobscura pseudoob*	3411	6

													scura		
	ConsensusfromContig112392_full_fwd_frame_2 (Ecdysteroid kinase)		*Coenobita*	56	1	ConsensusfromContig43934_full_rev_frame_0		*Coenobita*	105	2	ConsensusfromContig86526_full_rev_frame_0 (glutamine synthetase)		*Coenobita*	2429	5
	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	55	1	ConsensusfromContig80547_full_fwd_frame_1 (aldolase)		*Coenobita*	103	2	ConsensusfromContig57240_full_rev_frame_0		*Coenobita*	2352	1
						ConsensusfromContig826_full_rev_frame_1 (aldolase)		*Coenobita*	95	2	ConsensusfromContig112392_full_fwd_frame_2		*Coenobita*	2232	3
						ConsensusfromContig57240_full_rev_frame_0		*Coenobita*	76	1	EF1	CBE70970.1	*Oleria didymaea ramona*	2147	10
						ConsensusfromContig35401_full_rev_frame_0		*Coenobita*	74	1	isocitrate dehydrogenase	XP_001845030.1	*Culex quinquefasciatus*	2112	4
						ConsensusfromContig80831_full_fwd_frame_1		*Coenobita*	73	1	GAPDH	ADU87477.1	*Oncopagurus orientalis*	2072	5
						ConsensusfromContig86298_full_rev_frame_0		*Coenobita*	70	1	ConsensusfromContig32340_full_fwd_frame_0 (isocitrate dehydrogenase)		*Coenobita*	2027	2
						ConsensusfromContig35530_full_rev_frame_2		*Coenobita*	65	1	ConsensusfromContig61511_full_rev_frame_1		*Coenobita*	1646	2
											ConsensusfromContig9100_full_rev_frame_2 (GAPDH)		*Coenobita*	1636	2
											ConsensusfromContig83931_full_rev_frame_1 (aldolase)		*Coenobita*	1544	4
											Short chain specific acyl CoA dehydrogenase	AAO01112.1	*Drosophila pseudoobscura*	1502	6
											ConsensusfromContig82947_full_fwd_frame_0		*Coenobita*	1399	3
											ConsensusfromContig36749_full_rev_frame_0 (isocitrate dehydrogenase)		*Coenobita*	1180	3
											ConsensusfromContig82892_full_rev_frame_2		*Coenobita*	1042	1
											ribosome associated protein P40	BAB78527.1	*Bombyx mori*	880	5
											citrate synthase	XP_003739390.1	*Metaseiulus occidentalis*	856	3

											ConsensusfromContig15557_full_rev_frame_2 (glutamine synthetase)		*Coenobita*	850	1
											ConsensusfromContig43934_full_rev_frame_0 (glutamine synthetase)		*Coenobita*	850	1
											aldolase	EFA04932.1	*Tribolium castaneum*	833	5
											ConsensusfromContig10046_full_fwd_frame_0		*Coenobita*	806	1
											ConsensusfromContig34532_full_fwd_frame_2		*Coenobita*	765	3
											ConsensusfromContig62694_full_fwd_frame_2		*Coenobita*	754	1
											ConsensusfromContig27671_full_rev_frame_2		*Coenobita*	696	4
											ConsensusfromContig112805_full_fwd_frame_0 (aldolase)		*Coenobita*	591	2
											hexokinase	ABA63173.1	*Anopheles arabiensis*	501	2
											ConsensusfromContig88167_full_rev_frame_0		*Coenobita*	322	2
											ConsensusfromContig28763_full_rev_frame_0		*Coenobita*	247	3
12E	F1-ATP synthase beta subunit	ABI34071	*Pacifastacus leniusculus*	447	7	tubulin alpha	CAG03982.1	*Tetraodon nigroviridis*	1157	19	tubulin alpha	EDW38107.1	*Drosophila persimilis*	23482	18
	alpha-1-tubulin	AAC47522	*Gecarcinus lateralis*	314	5	tubulin beta	EFZ14304.1	*Solenopsis invicta*	1008	19	ATP synthase	ACU31112.1	*Childia sp.*	17272	13
	beta-1 tubulin	EF095147	*Homarus americanus*	301	6	ATP synthase	AES62833.1	*Medicago truncatula*	767	15	ConsensusfromContig31496_full_fwd_frame_1 (tubulin alpha)		*Coenobita*	13497	11
	eukaryotic initiation factor 4A	ABG67961	*Callinectes sapidus*	119	2	protein disulfide isomerase	ADY90107.1	*Penaeus monodon*	280	6	tubulin beta	AAB07482.1	*Homarus americanus*	6948	12
	ConsensusfromContig80552_full_rev_frame_1 (enolase)		*Coenobita*	87	2	thioredoxin	XP_003424375.1	*Nasonia vitripennis*	269	5	EF1	AAA28526.1	*Drosophila melanogaster*	6028	10
	ConsensusfromContig88167_full_rev_frame_0 (catalase)		*Coenobita*	67	1	ConsensusfromContig31496_full_fwd_frame_1 (tubulin alpha)		*Coenobita*	235	3	ConsensusfromContig88167_full_rev_frame_0		*Coenobita*	4291	4

	ConsensusfromContig80548_full_rev_frame_1 (calraticulin)		*Coenobita*	64	1	ConsensusfromContig110800_full_fwd_frame_2		*Coenobita*	161	2	actin	EFZ09249.1	*Solenopsis invicta*	4091	7
	protein disulfide isomerase	ADY90107	*Penaeus monodon*	60	1	ConsensusfromContig113644_full_fwd_frame_2		*Coenobita*	106	2	ConsensusfromContig80758_full_rev_frame_1 (enolase)		*Coenobita*	3280	3
	glutathione reductase	ZP_08416680	*Weissella cibaria*	58	1	ConsensusfromContig82928_full_rev_frame_0; ConsensusfromContig89315_full_rev_frame_0 (tubulin beta)		*Coenobita*	105	2	ConsensusfromContig77793_full_rev_frame_2 (ATP synthase)		*Coenobita*	3174	1
						ConsensusfromContig80552_full_rev_frame_1 (enolase)		*Coenobita*	103	1	ConsensusfromContig110113_full_rev_frame_2		*Coenobita*	2979	4
						ConsensusfromContig80872_full_rev_frame_1		*Coenobita*	99	2	ConsensusfromContig89315_full_rev_frame_0 (tubulin beta)		*Coenobita*	2029	2
						ConsensusfromContig112915_full_rev_frame_0		*Coenobita*	83	1	ConsensusfromContig82928_full_rev_frame_0 (tubulin beta)		*Coenobita*	2029	2
						ConsensusfromContig80758_full_rev_frame_1		*Coenobita*	79	1	ConsensusfromContig80548_full_rev_frame_1 (calreticulin)		*Coenobita*	2027	3
						ConsensusfromContig80548_full_rev_frame_1 (calreticulin)		*Coenobita*	73	1	ConsensusfromContig31598_full_rev_frame_1 (ATP synthase)		*Coenobita*	2017	2
						ConsensusfromContig28310_full_fwd_frame_0 (ATP synthase)		*Coenobita*	69	1	enolase	P56252	*Homarus gammarus*	1576	6
						ConsensusfromContig88167_full_rev_frame_0		*Coenobita*	66	1	ConsensusfromContig26577_full_rev_frame_1		*Coenobita*	1482	2
						ConsensusfromContig29457_full_rev_frame_0		*Coenobita*	64	1	Rab GDP dissociation inhibitor alpha	EFX87450.1	*Daphnia pulex*	1404	6
											ConsensusfromContig110800_full_fwd_frame_2 (tubulin beta)		*Coenobita*	1403	2
											TIF 4A	ABG67961.1	*Callinectes sapidus*	1245	7
											ConsensusfromContig33019_full_rev_frame_0 (tubulin beta)		*Coenobita*	1227	3
											calreticulin	ABC50166.1	*Fenneropenaeus chinensis*	943	3
											ConsensusfromContig80552_full_rev_frame_1		*Coenobita*	741	1
											ConsensusfromContig31496_full_fwd_frame_0 (tubulin alpha)		*Coenobita*	714	2

											catalase	AET34916.1	*Macrobrachium rosenbergii*	654	3
											hexokinase	ADD20426.1	*Glossina morsitans morsitans*	588	5
											ConsensusfromContig57240_full_rev_frame_0		*Coenobita*	571	1
											ConsensusfromContig32677_full_fwd_frame_1		*Coenobita*	461	1
											ConsensusfromContig84349_full_rev_frame_0 (catalase)		*Coenobita*	442	1
											ConsensusfromContig5363_full_fwd_frame_0 (tubulin beta)		*Coenobita*	363	1
											Laminin receptor	ABH10628.1	*Litopenaeus vannamei*	328	3
											ConsensusfromContig78006_full_fwd_frame_2 (Laminin receptor)		*Coenobita*	317	1
											thioredoxin	EFA09317.1	*Tribolium castaneum*	261	1
											ConsensusfromContig19315_full_rev_frame_2		*Coenobita*	244	1
											ATP dependent RNA helicase	CAA56197.1	*Drosophila melanogaster*	175	4
1G	beta-1 tubulin	EF095147	*Homarus americanus*	387	7	catalase	ACI13850.2	*Portunus trituberculatus*	1087	22	ATP synthase	ACB36913.1	*Litopenaeus vannamei*	21500	11
	alpha-1-tubulin	AAC47522	*Gecarcinus lateralis*	312	7	tubulin beta	EFZ14304.1	*Solenopsis invicta*	1070	23	ConsensusfromContig25049_full_rev_frame_2 (ATP synthase)		*Coenobita*	12615	2
	ConsensusfromContig79010_full_fwd_frame_0 (calsequestrin)		*Coenobita*	165	3	tubulin alpha	EHB00535.1	*Heterocephalus glaber*	824	15	ConsensusfromContig33429_full_fwd_frame_0 (tubulin alpha)		*Coenobita*	10511	3
	protein-disulfide isomerase	ACD44938	*Scylla paramamosain*	91	2	protein disulfide isomerase	ACN89260.1	*Litopenaeus vannamei*	628	10	ConsensusfromContig31496_full_fwd_frame_1 (ATP synthase)		*Coenobita*	10413	11
	ConsensusfromContig80548_full_rev_frame_1		*Coenobita*	78	1	ConsensusfromContig31496_full_fwd_frame_1 (tubulin alpha)		*Coenobita*	392	7	tubulin alpha	GB10009-PB	*Apis mellifera*	10364	10

	(calreticulin)														
	ConsensusfromContig25049_full_rev_frame_2 (catalase)		*Coenobita*	70	1	ConsensusfromContig25049_full_rev_frame_2		*Coenobita*	335	5	Protein disulfide isomerase	ACN89260.1	*Litopenaeus vannamei*	10064	9
						ConsensusfromContig26577_full_rev_frame_1		*Coenobita*	273	4	tubulin beta	AAB07482.1	*Homarus americanus*	7333	11
						ConsensusfromContig79010_full_fwd_frame_0		*Coenobita*	227	3	ConsensusfromContig88167_full_rev_frame_0		*Coenobita*	6915	5
						ATP synthase	YP_006716665.1	*Pelobacter carbinolicus*	223	4	ConsensusfromContig79010_full_fwd_frame_0 (protein disulfid isomerase)		*Coenobita*	6858	4
						ConsensusfromContig78240_full_fwd_frame_2		*Coenobita*	204	2	ConsensusfromContig80548_full_rev_frame_1 (calreticulin)		*Coenobita*	5657	8
						ConsensusfromContig81228_full_rev_frame_1		*Coenobita*	142	3	ConsensusfromContig81309_full_fwd_frame_0		*Coenobita*	5045	4
						ConsensusfromContig1407_full_fwd_frame_0		*Coenobita*	122	2	Catalase	AET34916.1	*Macrobrachium rosenbergii*	3785	8
						ConsensusfromContig82928_full_rev_frame_0; ConsensusfromContig89315_full_rev_frame_0 (tubulin beta)		*Coenobita*	107	2	ConsensusfromContig109711_full_rev_frame_1		*Coenobita*	3727	2
						transglutaminase	YP_001624686.1	*Renibacterium salmoninarum*	77	1	ConsensusfromContig26577_full_rev_frame_1		*Coenobita*	3693	1
						ConsensusfromContig80548_full_rev_frame_1		*Coenobita*	73	1	ConsensusfromContig37063_full_fwd_frame_2		*Coenobita*	2877	1
						ConsensusfromContig77986_full_fwd_frame_0		*Coenobita*	71	1	ConsensusfromContig15227_full_rev_frame_1		*Coenobita*	2580	1
											ConsensusfromContig31598_full_rev_frame_1 (ATP synthase)		*Coenobita*	2568	3
											ConsensusfromContig26380_full_fwd_frame_2		*Coenobita*	2491	5
											ConsensusfromContig81228_full_rev_frame_1		*Coenobita*	2395	3
											ConsensusfromContig77793_full_rev_frame_2 (ATP synthase)		*Coenobita*	2170	1
											ConsensusfromContig37063_full_fwd_frame_1		*Coenobita*	1940	3

											ConsensusfromContig110800_full_fwd_frame_2 (tubulin beta)		*Coenobita*	1787	3
											ConsensusfromContig112167_full_fwd_frame_1 (catalase)		*Coenobita*	1478	1
											ConsensusfromContig115046_full_fwd_frame_1		*Coenobita*	1389	1
											ConsensusfromContig33019_full_rev_frame_0 (tubulin beta)		*Coenobita*	1271	4
											Glucose 6 phosphate isomerase	AF310780.1	*Automate gardineri*	1126	2
											Pyruvate kinase	ABO21408.1	*Litopenaeus vannamei*	1093	7
											ConsensusfromContig78240_full_fwd_frame_2 (calreticulin)		*Coenobita*	1086	5
											ConsensusfromContig59947_full_fwd_frame_0		*Coenobita*	963	1
											ConsensusfromContig28310_full_fwd_frame_0		*Coenobita*	958	2
											ConsensusfromContig84349_full_rev_frame_0 (catalase)		*Coenobita*	901	2
											ConsensusfromContig89315_full_rev_frame_0 (tubulin beta)		*Coenobita*	427	1
											ConsensusfromContig82928_full_rev_frame_0 (tubulin beta)		*Coenobita*	427	1
											Elongation factor 1 alpha	AAD21849.1	*Heteromysis formosa*	379	1
											ConsensusfromContig36027_full_fwd_frame_1		*Coenobita*	316	2
											polyubiquitin	NP_001261383	*Drosophila melanogaster*	162	10
2G	putative chaperonin GroL	XP_003371438	*Trichinella spiralis*	219	4	HSP60	AFA36427.1	*Portunus trituberculatus*	597	12	hsp60	EGI60183.1	*Acromyrmex echinatior*	6793	11

	HSP 70	ABM90803	*Dendrolimus punctatus*	194	4	protein disulfide isomerase	AEE36486.1	*Fenneropenaeus chinensis*	509	9	ConsensusfromContig33429_full_fwd_frame_0		*Coenobita*	6327	4
	protein disulfide isomerase	ACN89260	*Litopenaeus vannamei*	142	4	hsp70	ADV59560.1	*Paracyclopina nana*	212	4	ConsensusfromContig91554_full_rev_frame_2		*Coenobita*	6034	2
	ConsensusfromContig81309_full_fwd_frame_0 (luciferase)		*Coenobita*	79	1	ConsensusfromContig82003_full_rev_frame_2 (hsp70)		*Coenobita*	174	3	ConsensusfromContig79010_full_fwd_frame_0 (protein disulfide isomerase)		*Coenobita*	5923	4
	ConsensusfromContig33429_full_fwd_frame_0 (thioredoxin)		*Coenobita*	61	1	ConsensusfromContig812289_full_rev_frame_1		*Coenobita*	148	2	ConsensusfromContig25049_full_rev_frame_2		*Coenobita*	5733	2
						ConsensusfromContig109711_full_rev_frame_1 (protein disulfide isomerase)		*Coenobita*	80	1	hsp70	ABF18258.1	*Aedes aegypti*	4037	6
						ConsensusfromContig81309_full_fwd_frame_0		*Coenobita*	78	1	ConsensusfromContig81309_full_fwd_frame_0		*Coenobita*	3917	3
						ConsensusfromContig80527_full_fwd_frame_1 (hsp70)		*Coenobita*	77	1	ConsensusfromContig88167_full_rev_frame_0		*Coenobita*	3800	4
											ATP synthase	ABI34071.1	*Pacifastacus leniusculus*	2772	9
											ConsensusfromContig109711_full_rev_frame_1 (protein disulfide isomerase)		*Coenobita*	2650	3
											ConsensusfromContig115046_full_fwd_frame_1		*Coenobita*	2056	1
											chaperonin	EAA13612	*Anopheles gambiae*	1948	8
											ConsensusfromContig77817_full_fwd_frame_1		*Coenobita*	1925	2
											ConsensusfromContig37063_full_fwd_frame_2		*Coenobita*	1914	1
											hsp90	AGC54636.1	*Scylla paramamosain*	1573	5
											ConsensusfromContig31046_full_rev_frame_0		*Coenobita*	1442	3
											ubiquitin	XP_550846.3	*Anopheles gambiae*	1402	14
											ConsensusfromContig26577_full_rev_frame_1		*Coenobita*	1399	1
											ConsensusfromContig15629_full_fwd_frame_0 (hsp90)		*Coenobita*	1363	4

											ConsensusfromContig81228_full_rev_frame_1		*Coenobita*	1339	3
											catalase	AET34916.1	*Macrobrachium rosenbergii*	1295	7
											ConsensusfromContig31496_full_fwd_frame_1 (tubulin alpha)		*Coenobita*	1219	5
											tubulin alpha	NP_476772.1	*Drosophila melanogaster*	1180	5
											hemocyanin	ABR14694.1	*Marsupenaeus japonicus*	1086	6
											ConsensusfromContig84349_full_rev_frame_0 (catalase)		*Coenobita*	1017	3
											ConsensusfromContig82267_full_rev_frame_1 (hsp60)		*Coenobita*	951	3
											phosphoglucose isomerase	ACS27516.1	*Colias eurytheme*	743	5
											pyruvate kinase	ABY66598.1	*Litopenaeus vannamei*	404	5
											conserved protein	GK15917	*Drosophila willistoni*	355	6
											ConsensusfromContig5520_full_fwd_frame_2		*Coenobita*	343	1
											succinate dehydrogenase	JAA72360.1	*Ixodes ricinus*	246	5
3G	Bip	AFQ62791	*Litopenaeus vannamei*	301	6	hemocyanin	ADZ15149.1	*Litopenaeus vannamei*	1084	23	hsp70	ABF18258.1	*Aedes aegypti*	21465	10
	hemocyanin	AEG64817	*Eriocheir sinensis*	70	2	hsp70	EFN61604.1	*Camponotus floridanus*	627	12	ConsensusfromContig80527_full_fwd_frame_1 (hsp70)		*Coenobita*	11580	2
	N-acetyltransferase	YP_003812257	*gamma proteobacterium HdN1*	60	1	ConsensusfromContig82003_full_rev_frame_2 (hsp70)		*Coenobita*	167	3	ConsensusfromContig82003_full_rev_frame_2 (hsp70)		*Coenobita*	9400	2
						ConsensusfromContig80551_full_fwd_frame_0		*Coenobita*	136	2	ConsensusfromContig80551_full_fwd_frame_0 (hsp70)		*Coenobita*	4231	4

						ConsensusfromContig80527_full_fwd_frame_1 (hsp70)		*Coenobita*	114	2	hemocyanin	AAL27460.1	*Penaeus monodon*	2236	5
						ConsensusfromContig85252_full_rev_frame_2		*Coenobita*	73	1	ConsensusfromContig81168_full_rev_frame_2		*Coenobita*	1929	3
						ConsensusfromContig18312_full_fwd_frame_2		*Coenobita*	64	1	ConsensusfromContig81047_full_rev_frame_2 (hsp70)		*Coenobita*	1678	2
						ConsensusfromContig113739_full_rev_frame_2		*Coenobita*	63	1	glu pro tRNA synthetase	AGB92285.1	*Scardiella approximatella*	476	2
						ConsensusfromContig69220_full_rev_frame_0; ConsensusfromContig3066_full_rev_frame_2; ConsensusfromContig33600_full_rev_frame_2		*Coenobita*	62	1	ConsensusfromContig109695_full_rev_frame_2		*Coenobita*	327	2
						ConsensusfromContig40556_full_fwd_frame_2		*Coenobita*	62	1					
4G	HSP 90	ACJ01642	*Eriocheir sinensis*	220	4	hemocyanin	AEG64817.1	*Eriocheir sinensis*	648	13	hsp70	BAJ78982.1	*Marsupenaeus japonicus*	4067	9
	molecular chaperone HtpG	GAA47176	*Clonorchis sinensis*	143	3	hsp90	gb|ACQ90225.1|	*Portunus trituberculatus*	159	3	hsp90	AFX84559.1	*Lygus hesperus*	3683	7
	HSP 70	ABF18258	*Aedes aegypti*	139	3	ConsensusfromContig40556_full_fwd_frame_2		*Coenobita*	133	2	ConsensusfromContig82003_full_rev_frame_2 (hsp70)		*Coenobita*	2537	3
	endoplasmin	ACS75351	*Locusta migratoria*	87	3	hsp70	ACA84007.1	*Haemaphysalis longicornis*	116	2	ConsensusfromContig80527_full_fwd_frame_1 (hsp70)		*Coenobita*	1810	2
	ConsensusfromContig85811_full_fwd_frame_0 (Acyl-CoA dehydrogenase)		*Coenobita*	71	1	carboxylesterase	YP_003165471.1	*Accumulibacter phosphatis*	84	1	hemocyanin	ABV58636.1	*Metapenaeus ensis*	1458	3
	hemocyanin	AEG64817	*Eriocheir sinensis*	58	1	ConsensusfromContig82003_full_rev_frame_2; ConsensusfromContig80527_full_fwd_frame_1 (hsp70)		*Coenobita*	77	1	ConsensusfromContig15629_full_fwd_frame_0 (hsp90)		*Coenobita*	1245	2
											ConsensusfromContig81844_full_rev_frame_0 (hsp90)		*Coenobita*	1136	2
											ConsensusfromContig47227_full_rev_frame_2		*Coenobita*	682	3
											ConsensusfromContig27127_full_fwd_frame_1		*Coenobita*	572	2

											ConsensusfromContig80551_full_fwd_frame_0 (hsp70)		*Coenobita*	561	2
											hsp80	CAD59258.1	*Darwinula stevensoni*	482	2
											ConsensusfromContig78138_full_rev_frame_0		*Coenobita*	438	3
											ConsensusfromContig30298_full_rev_frame_0		*Coenobita*	345	1
											serine protease	ACO15310.1	*Caligus clemensi*	292	1
5G	HSP90 (ConsensusfromContig27768_full_fwd_frame_1; ConsensusfromContig81844_full_rev_frame_0; ConsensusfromContig15629_full_fwd_frame_0)		*Coenobita*	164	3	endoplasmin	XP_001844128.1	*Culex quinquefasciatus*	894	17	ConsensusfromContig15629_full_fwd_frame_0 (hsp90)		*Coenobita*	10470	2
						Histidine kinase-like ATPase	XP_321706.5	*Anopheles gambiae*	843	16	hsp90	CAJ28987.1	*Ceratitis capitata*	8163	11
						tumor rejection antigen	XP_002413149.1	*Ixodes scapularis*	599	11	ConsensusfromContig81844_full_rev_frame_0 (hsp90)		*Coenobita*	6467	2
	endoplasmin	ACS75351	*Locusta migratoria*	113	3	hemocyanin	AEG64817.1	*Eriocheir sinensis*	486	10	EF2	XP_975635.1	*Tribolium castaneum*	3340	12
	gelsolin	CCI71879	*Homarus americanus*	97	2	hsp90	XP_002063871.1	*Drosophila willistoni*	391	8	ConsensusfromContig103621_full_rev_frame_0 (EF)		*Coenobita*	2116	2
	EF2b	NP_525105	*Drosophila melanogaster*	94	2	ConsensusfromContig27768_full_fwd_frame_1		*Coenobita*	128	2	ConsensusfromContig27768_full_fwd_frame_1 (hsp90)		*Coenobita*	1623	3
	hemocyanin	CAD87762	*Perla marginata*	61	1	ConsensusfromContig81844_full_rev_frame_0 (hsp90)		*Coenobita*	110	2	chromobox protein homolog 3	XP_003243341.1	*Acyrthosiphon pisum*	1560	1
	ATPase CDC48 subfamily	EFX72171	*Daphnia pulex*	61	1	ConsensusfromContig86082_full_fwd_frame_0		*Coenobita*	105	2	ATPase	AEE61431.1	*Apis florea*	1385	13
						ConsensusfromContig15629_full_fwd_frame_0 (hsp90)		*Coenobita*	80	1	hemocyanin	ABV58636.1	*Metapenaeus ensis*	1174	3
											gelsolin	CCI71880.1	*Homarus americanus*	1034	8
											ConsensusfromContig59083_full_fwd_frame_1 (hsp90)		*Coenobita*	985	2

											ConsensusfromContig110317_full_fwd_frame_0 (hsp90)		*Coenobita*	726	1
											ConsensusfromContig34115_full_fwd_frame_0		*Coenobita*	668	2
											glycogen phosphorylase	BAJ23879.1	*Marsupenaeus japonicus*	607	8
											serine protease	ACO15310.1	*Caligus clemensi*	470	1
											ConsensusfromContig77694_full_rev_frame_0		*Coenobita*	457	3
											endoplasmic reticulum membrane fusion protein	AAF17568.1	*Drosophila melanogaster*	378	5
											glycosyl transferase	NP_001163490.1	*Drosophila melanogaster*	362	3
											ConsensusfromContig103948_full_rev_frame_1 (EF2)		*Coenobita*	195	2
6G	glycoprotein 93	NP_651601	*Drosophila melanogaster*	83	1	hsp70	XP_002401154.1	*Ixodes scapularis*	277	6	ubiquitin	XP_550846.3	*Anopheles gambiae*	28184	29
	26S proteasome regulatory subunit	XP_002407373	*Ixodes scapularis*	83	1	26S proteasome regulatory subunit	XP_002407373.1	*Ixodes scapularis*	199	4	hsp90	ACS75351.1	*Locusta migratoria*	3495	7
						endoplasmin	ACS75351.1	*Locusta migratoria*	115	2	ConsensusfromContig15629_full_fwd_frame_0 (hsp90)		*Coenobita*	2988	3
						ConsensusfromContig59543_full_rev_frame_2		*Coenobita*	96	2	ConsensusfromContig88167_full_rev_frame_0		*Coenobita*	2410	4
	Tudor staphylococcal nuclease	AEK49107	*Penaeus monodon*	82	2	methyltransferase	EKD45392.1	*uncultured bacterium*	74	1	EF2	BAM18989.1	*Papilio polytes*	2117	7
											ConsensusfromContig27852_full_rev_frame_0 (ubiquitin)		*Coenobita*	2087	3
											ConsensusfromContig81844_full_rev_frame_0 (hsp90)		*Coenobita*	1799	1
	ATP-citrate synthase	XP_623083	*Apis mellifera*	74	1						ConsensusfromContig57845_full_fwd_frame_1		*Coenobita*	1750	2
											ConsensusfromContig34474_full_fwd_frame_1		*Coenobita*	1587	3

											ConsensusfromContig31888_full_rev_frame_1 (ubiquitin)		*Coenobita*	1324	2
											ConsensusfromContig110270_full_rev_frame_0 (ubiquitin)		*Coenobita*	1324	2
											ConsensusfromContig73466_full_fwd_frame_0		*Coenobita*	1279	1
											ConsensusfromContig44495_full_rev_frame_2		*Coenobita*	814	2
											ConsensusfromContig27768_full_fwd_frame_1 (hsp90)		*Coenobita*	649	3
											ConsensusfromContig2325_full_rev_frame_1		*Coenobita*	605	1
											Preprotein translocase Sec61	XP_002005983.1	*Drosophila mojavensis*	598	2
											ConsensusfromContig3775_full_rev_frame_1		*Coenobita*	555	1
											gelsolin	CCI71879.1	*Homarus americanus*	347	4
											NADH-ubiquinone oxidoreductase	XP_001356709.2	*Drosophila pseudoobscura pseudoobscura*	340	5
											serine protease	ACO15310.1	*Caligus clemensi*	236	2
7G	ConsensusfromContig28051_full_rev_frame_0 (xanthine dehydrogenase/oxidase)		*Coenobita*	63	1	ConsensusfromContig90974_full_rev_frame_2		*Coenobita*	144	3	ConsensusfromContig54204_full_fwd_frame_2		*Coenobita*	7809	4
	carbamoyl-phosphate synthase	YP_823133	*Candidatus Solibacter usitatus*	58	1	ConsensusfromContig25946_full_fwd_frame_2		*Coenobita*	137	3	ConsensusfromContig28051_full_rev_frame_0		*Coenobita*	2227	2
						ConsensusfromContig115875_full_fwd_frame_1		*Coenobita*	136	3	ConsensusfromContig57845_full_fwd_frame_1		*Coenobita*	1735	1
						major facilitator transporter	ABV99912.1	*Salinispora arenicola*	121	2	ConsensusfromContig108788_full_fwd_frame_0		*Coenobita*	1682	1
						ConsensusfromContig27902_full_fwd_frame_1		*Coenobita*	100	2	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	1574	2

						ConsensusfromContig57845_full_fwd_frame_1 (cuticle protein)		*Coenobita*	83	1					
						trehalose	XP_001985945	*Drosophila grimshawi*	76	1					
						ring-hydroxylating dioxygenase	YP_001260833	*Sphingomonas wittichii*	74	1					
						Alpha-tubulin suppressor	XP_001993649	*Drosophila grimshawi*	72	1					
						ConsensusfromContig108788_full_fwd_frame_0		*Coenobita*	68	1					
						cuticle protein	XP_002061935.1	*Drosophila willistoni*	67	1					
						cytochrome P450	EEC19243.1	*Ixodes scapularis*	67	1					
8G	clathrin	ACY44823	*Libinia emarginata*	73	2	ConsensusfromContig108788_full_fwd_frame_0		*Coenobita*	119	2	ubiquitin	XP_550846.3	*Anopheles gambiae*	7204	15
	ConsensusfromContig81048_full_rev_frame_0		*Coenobita*	56	1	ConsensusfromContig90974_full_rev_frame_2; ConsensusfromContig17144_full_fwd_frame_1		*Coenobita*	98	1	ConsensusfromContig108788_full_fwd_frame_0		*Coenobita*	1146	1
						fimbrial assembly protein	EFE52677.1	*Providencia rettgeri*	74	1	ConsensusfromContig54204_full_fwd_frame_2		*Coenobita*	1068	2
						trypsin	XP_001847026.1	*Culex quinquefasciatus*	64	1	ConsensusfromContig64225_full_rev_frame_0		*Coenobita*	993	1
											Vesicle coat complex COPI	EFX89146.1	*Daphnia pulex*	368	4
											ConsensusfromContig89626_full_fwd_frame_0 (Vesicle coat complex COPI)		*Coenobita*	338	1
											ConsensusfromContig110329_full_rev_frame_1		*Coenobita*	325	2
											clathrin	ACY44823.1	*Libinia emarginata*	116	6
9G	ConsensusfromContig57845_full_fwd_frame_1 (cuticle protein)		*Coenobita*	56	1	cationic trypsin	XP_003929837	*Saimiri boliviensis boliviensis*	213	4	ubiquitin	XP_001845257.1	*Culex quinquefasciatus*	21017	30
						ConsensusfromContig23420_full_		*Coenobita*	96	2	ConsensusfromContig57845_		*Coenobita*	157	3

						fwd_frame_2					full_fwd_frame_1			1	
						ConsensusfromContig57845_full_fwd_frame_1		*Coenobita*	83	1	ConsensusfromContig27852_full_rev_frame_0 (ubiquitin)		*Coenobita*	1328	4
						LamG domain-containing protein jellyroll fold domain-containing protein	YP_003509361.1	*Stackebrandtia nassauensis*	77	1	ribosomal protein L40	NP_476776.1	*Drosophila melanogaster*	1328	4
						2';5' RNA ligase	EFU60408.1	*Actinomyces sp.*	76	1	ConsensusfromContig88494_full_fwd_frame_2		*Coenobita*	615	1
						ConsensusfromContig73654_full_rev_frame_2 (cationic trypsin)		*Coenobita*	60	1					
10G	myosin	EHJ67857	*Danaus plexippus*	83	2	myosin	EFA07932	*Tribolium castaneum*	323	5	myosin	EHJ67857.1	*Danaus plexippus*	492	5
	sensory box/GGDEF family protein	YP_002540690	*Agrobacterium radiobacter*	68	1	Ribose-binding protein of ribose ABC transporter	EEP89808.1	*Yersinia kristenseni i*	114	2					
	laminin	EFN82670	*Harpegnathos saltator*	68	1	ConsensusfromContig78915_full_fwd_frame_2		*Coenobita*	70	1					
	ConsensusfromContig87576_full_rev_frame_1 (myosin)		*Coenobita*	67	1	DNA-directed RNA polymerase III	ACO15732.1	*Caligus clemensi*	64	1					
						ConsensusfromContig87576_full_rev_frame_1		*Coenobita*	61	1					
11G	spectrin	XP_002433506	*Ixodes scapularis*	379	8	Spectrin	EFX88672	*Daphnia pulex*	1431	28	polyubiquitin	XP_550846.3	*Anopheles gambiae*	8015	29
	ConsensusfromContig84462_full_fwd_frame_0 (ankyrin)		*Coenobita*	75	1	phosphate ABC transporter permease	ADL03004.1	*Clostridium saccharolyticum*	115	2	spectrin	EFX88672.1	*Daphnia pulex*	3875	22
						LuxR family transcriptional regulator	ADC57464.1	*Klebsiella variicola*	77	1	ConsensusfromContig89277_full_rev_frame_0 (spectrin)		*Coenobita*	1635	2
						cation efflux pump	EED33747.1	*gamma proteobacterium*	72	1	ConsensusfromContig84462_full_fwd_frame_0		*Coenobita*	697	2

Supplementary to manuscript 3: Additional file 3

	Mascot				MS BLAST				MSE				GO association
	Accession	***Species***	**Score**	**Peptide hits**	**Accession**	***Species***	**Score**	**Peptide hits**	**Accession**	***Species***	**Score**	**Peptide hits**	
alpha 2 macroglobulin (ConsensusfromContig31628_full_rev_frame_1)						*Coenobita*	112	2		*Coenobita*	2516	3	GO:0019731: antibacterial humoral response
alpha 2-macroglobulin					ABD61456	*Scylla serrata*	278	6					
alpha spectrin	EFX88672	*Daphnia pulex*	296	6	XP_002083280	*Drosophila simulans*	617	13					
Aspartate protease cathepsin D									AEO94539	*Triatoma infestans*	520	2	GO:0008233: peptidase activity
coagulation factor XI					XP_001845410	*Culex quinquefasciatus*	72	1					GO:0006508 : proteolysis
coatomer protein complex					NP_001166195	*Bombyx mori*	71	1					GO:0016192 : vesicle-mediated transport
coatomer protein complex (ConsensusfromContig810 full rev frame 0)										*Coenobita*	639	2	
coatomer protein delta, isoform B									NP_001162642	*Drosophila melanogaster*	723	5	
coatomer subunit delta-like									XP_003700942	*Megachile rotundata*	738	6	
C-type lectin (ConsensusfromContig110279_full_fwd_frame_0)		*Coenobita*	82	1						*Coenobita*	1730	1	GO:0002682 : regulation of immune system process
emp24 cargo transport protein (ConsensusfromContig108988_full_fwd_frame_0)										*Coenobita*	307	2	GO:0051050 : positive regulation of transport
gamma-interferon inducible lysosomal thiol reductase					XP_002400600	*Ixodes scapularis*	65	1					GO:0007166 : cell surface receptor signaling pathway
GH15863 gene product from transcript GH15863-RA, Peptidase_C26					XP_001983370	*Drosophila grimshawi*	105	2					GO:0016787: hydrolase activity
GL17639 (ConsensusfromContig35747 full rev frame 1)										*Coenobita*	321	3	GO:0000042 : protein targeting to Golgi

glutathione-S-transferase (ConsensusfromContig523_full_rev_frame_0)						*Coenobita*	77	1					GO:0006979 : response to oxidative stress
glutathione-S-transferase (ConsensusfromContig80526_full_rev_frame_0)										*Coenobita*	426	3	
glutathione-S-transferase (ConsensusfromContig83104_full_rev_frame_0)										*Coenobita*	765	1	
GST_N_Metaxin									XP_002430330	*Pediculus humanus corporis*	489	4	
S-formylglutathione hydrolase (ConsensusfromContig31027 full rev frame 1)										*Coenobita*	392	1	GO:0018738 : S-formylglutathione hydrolase activity
signal peptidase complex subunit 3-like					XP_003739905	*Metaseiulus occidentalis*	65	1					GO:0009003: signal peptidase activity
Sorting nexin-12									XP_002432737.1	*Pediculus humanus corporis*	841	3	GO:0015031 : protein transport
Trafficking protein particle complex subunit 10									XP_003742750	*Metaseiulus occidentalis*	671	2	GO:0016192: vesicle-mediated transport
Transient receptor potential-gamma protein									EFN78599	*Harpegnathos saltator*	439	1	GO:0005261: cation channel activity
uncharacterized protein									XP_003748091	*Metaseiulus occidentalis*	531	2	homology "secreted salivary gland peptide" Isca XP_002414539