Molecular biology of barnacle *Balanus improvisus* settlement

Anna Abramova

Department of Chemistry and Molecular Biology
Faculty of Science

UNIVERSITY OF GOTHENBURG

Cover illustration: Anna Abramova

Molecular Biology of Barnacle *Balanus improvisus* settlement

anna.abramova@marine.gu.se

ISBN 978-91-7833-614-2 (PRINT)
ISBN 978-91-7833-615-9 (PDF)

Printed in Gothenburg, Sweden 2019
Printed by BrandFactory

To my family and friends
who have supported me
through these years

ABSTRACT

The aim of this thesis was to investigate molecular mechanisms of various aspects of barnacle settlement, using the acorn barnacle *Balanus improvisus*. This barnacle is a common fouling species and a model organism for studies in settlement biology, in particular in relation to antifouling research.

In order to facilitate the development of genomic resources in this species, we conducted a pilot study for the sequencing of the *B. improvisus* genome and performed an initial genomic characterization. The analysis revealed that *B. improvisus* genome has an extremely high genetic diversity, with about 5% nucleotide diversity in coding regions. In addition, we experimentally estimated the *B. improvisus* genome size, based on DNA staining and flow cytometry measurements, resulting in a haploid genome size of 738 Mbp.

To investigate molecular changes during the settlement process, transcriptomes of four different settlement stages, ie free-swimming, close-search, attached and juvenile, were compared. We identified several key genes involved in the hormonal regulation of molting and metamorphosis, including the broad complex, ecdysone receptor and retinoid X receptor, adding a new level of insight to the molecular mechanisms involved in settlement. Furthermore, we used two types of surfaces with different wettability to test if differences in surface preferences are reflected in gene expression. The results revealed that exploration of the “favourable” hydrophobic surface induced more genes and with larger changes in expression than on hydrophilic suggesting a stronger transcriptional response.

We also investigated two specific aspects related to barnacle chemical communication during settlement - sensory receptors and pheromones. Analysis of the transcriptome of cyprid antennules resulted in the identification of two receptor classes, the chemosensory ionotropic receptors and mechanosensory receptors represented by several TRP subfamilies. We identified and characterized six homologs of the waterborne pheromone WSP in *B. improvisus* that showed differential expression during settlement. These results suggest the existence of a pheromone mix, where con-specificity might be determined by a combination of sequence characteristics and the concentration of the individual components.

With the aim to further establish *B. improvisus* into a potent marine model system, a detailed protocol was developed for an all-year-round culturing of *B. improvisus* and adapted at Tjärnö Marine Laboratory.

Finally, I summarise current knowledge on the molecular mechanisms of barnacle settlement and outline new research directions to further improve our understanding of the settlement biology of this species.

Keywords: barnacles, settlement, transcriptomics, ecdysone cascade, chemosensory receptors, waterborne pheromones

LIST OF PAPERS

This thesis is based on the following studies, referred to in the text by their Roman numbers:

I. Abramova, A., Rosenblad, M. A., Blomberg, A., & Larsson, T. A. (2019). Sensory receptor repertoire in cyprid antennules of the barnacle *Balanus improvisus*. *PloS One*, 14(5), e0216294.

II. Abramova, A., Lind, U., Blomberg, A., Rosenblad, M. A. (2019). The complex barnacle perfume: identification of waterborne pheromone homologues in *Balanus improvisus* and their differential expression during settlement. *Biofouling*, 1-13.

III. Jonsson, P. R., Wrange, A. L., Lind, U., Abramova, A., Ogemark, M., & Blomberg, A. (2018). The Barnacle *Balanus improvisus* as a Marine Model-Culturing and Gene Expression. *JoVE (Journal of Visualized Experiments)*, (138), e57825.

IV. Abramova, A., Rosenblad, M. A., Blomberg, A. Insights into the molecular mechanisms of *Balanus improvisus* settlement and discriminatory behaviour towards surfaces. (Manuscript)

V. Rosenblad, M. A., Abramova, A, Lind, U., Olasson, P., Giacomello, S., Nystedt, B., Blomberg, A. A pilot study for the sequencing of the barnacle *Balanus improvisus* genome reveals extreme genetic diversity. (Manuscript)

CONTENT

ABBREVIATIONS

20E	20-hydroxyecdysone
AC	Adenylate cyclase
AFM	Atomic force microscopy
BRC	Broad complex
cAMP	Cyclic AMP
CDS	Coding sequence
CEGMA	Core eukaryotic genes mapping approach
DAG	Diacylglycerol
DOPA	Dihydroxyphenylalanine
EcR	Ecdysone receptor
EGF	Epidermal growth factor
ER	Endoplasmic reticulum
FACS	Fluorescence-activated cell sorting
FSC-A	Forward light scatter
FTZ-F1	Fushi tarazu transcription factor 1
GGR	Glycyl-glycyl-arginine
GO	Gene ontology
GPCR	G-protein coupled receptor
GR	Gustatory receptor
HMM	Hidden Markov model
HR3	Hormone receptor 3
HR38	Hormone receptor 38
iGluR	Ionotropic glutamate receptor
IP_3	Inositol trisphosphate
IR	Ionotropic receptor
MF	Methyl farnesoate
MULTIFUNCin	Multifunctional protein cue
OctA	Octopamine
OR	Odorant receptor
PIP_2	Phosphatidylinositol bisphosphate
PKC	Protein kinase C
PSU	Practical salinity units
RXR	Retinoid X receptor
SIPC	Settlement-inducing protein complex
SNP	Single nucleotide polymorphism
TRP	Transient receptor potential channel
WSP	Waterborne settlement pheromone

1 INTRODUCTION

1.1 Historical note

Barnacles is a group of sessile crustaceans that has long been disliked by many as strongly as admired by some. For the most, these animals are known for their ability to stick permanently to the hulls of the boats and thus being a concern for humankind since the dawn of seafaring. However, barnacles have fascinated many scientists in the fields as diverse as evolution, taxonomy and material science.

Barnacle biology has long been a source of confusion for the early naturalists. These creatures once were thought to be the eggs of barnacle geese after numerous accounts of timber woods washed ashore carrying "multitudes of little shells; having within them little birds perfectly shaped, supposed to be barnacles" according to Moray's 1678 description, which today are recognised as goose necked barnacles, *Lepas sp.* (Anderson 1993) (Figure 1).


CHAP. 171.

Of the Goose tree, Barnacle tree, or the tree bearing Geese.

Britannicæ Conchæ anatiferæ.
The breed of Barnacles.

Figure 1. The legend of barnacles being eggs of the barnacle geese. Reproduction from Plantarum, Seu Stirpium Icones, Mathias de L'Obel (1581).

This legend persisted until the eighteenth century, when Linnaeus classified barnacles among Mollusca in the *Systema Naturae* due to their calcareous shell and presence of the mantle cavity. However, inconsistency of the internal morphology of these animals gave rise to the alternative views on the origin of barnacles. In 1829, J.V. Thompson, a British army surgeon, was the first one to demonstrate the presence of crustacean nauplii preceding the sessile adult stage (Winsor 1969), leading to the ultimate classification of barnacles within Crustacea.

Charles Darwin became fascinated by barnacles after a discovery of a burrowing barnacle during the Beagle voyage, followed by eight years of intensive studies of biology and taxonomy of the entire group. As a result, he published four monographs that firmly established barnacles as a subclass of the Crustacea: the Cirripedia, that remain the foundation of the cirriped biology till that day.

Since then, considerable advances have been made in many aspects of barnacle biology including distribution and taxonomy, fossil history and evolution, life history and physiology. One direction in barnacle research that got particular attention in the last decades is related to barnacles being one of the most common organisms in biofouling communities on marine constructions all over the world. Understanding of barnacle settlement and metamorphosis as well as adhesion has been the main focus of the studies in the last decades (Holm 2012).

1.2 Barnacle biology and life history

Barnacles is a diverse group of crustaceans containing more than 1,000 exclusively sessile species, inhabiting all kinds of marine environments from rocky shores to abyssal vents (Pérez-Losada et al. 2008). Due to their sessile lifestyle, the adult morphology is rather different compared to other crustaceans. As the current work mainly concerns the species belonging to Thoracica, the following description of the morphology and life cycle will be restricted to this group.

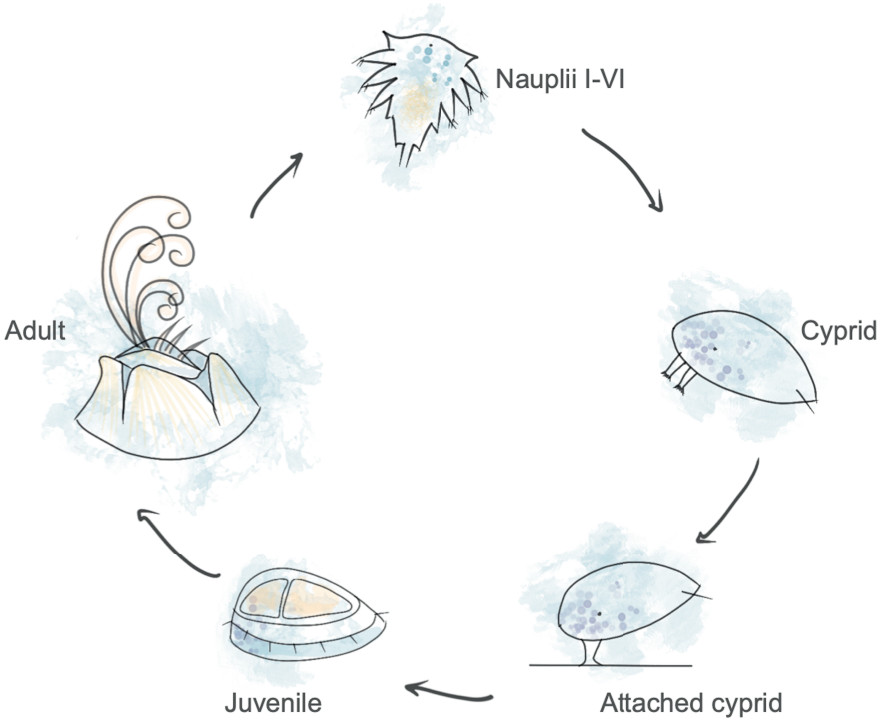


Figure 2. Barnacle life cycle includes six feeding nauplii stages and one non-feeding cyprid stage. Cyprid attaches to a substrate and undergoes metamorphosis into a juvenile.

Louis Agassiz described representatives of Thoracica in the beginning of the nineteenth century, as “nothing more than a little shrimp-like animal, standing on its head in a limestone house and kicking food into its mouth” (Glenner and Hebsgaard 2006). Indeed, an adult barnacle is surrounded by a calcareous shell firmly attached to a substrate and only cirri, the biramous thoracic limbs modified into feeding appendages, are extended out from the shell to filter food particles from the surrounding water.

Barnacles are predominantly cross-fertilizing hermaphrodites. Mating occurs by the means of an unfoldable penis that can reach up to eight times the body length and can also change morphology depending on the surrounding conditions such as wave exposure (Hoch 2008). After fertilization occurs, embryos are brooded within the mantle cavity of the adult until released as nauplii larvae. Nauplii undergo five subsequent molts until becoming a barnacle-specific larva called ‘cyprid’ (Figure 2). The cyprid has a critical role in the transition from the planktonic to the sessile phase. It is equipped with a comparably well-developed nervous system and characterized by complex behaviour. Cyprids select the settlement spot based on a range of biological and physico-chemical cues (Lang et al. 1979, Judge and Craig 1997, Matsumura et al. 1998, Dreanno et al. 2006b). If during exploration the site is considered favourable, the cyprid permanently attaches to the substrate with adhesives secreted from the cement glands, and then undergoes metamorphosis.

1.3 Cyprid morphology

Cyprid is the last larval stage in the barnacle life cycle and it is evolved for surface exploration, finding the favourable settlement spot and performing final adhesion prior to metamorphosis. The fact that the overall anatomy of the cypris larva is uniform across all cirripedes despite the diversity of the adult forms suggests that its morphology is supremely adapted and fine-tuned for this task (Høeg et al. 2009).

The cypris larva of *B. improvisus* is approximately 500 μm from the rostral to the caudal end of the carapace. It has a bivalved chitin shell, resembling an ostracod in shape (Figure 3, A). Cyprid body is divided into the cephalon and thorax enclosed in anterior and posterior mantle cavities. Six pairs of thoracic legs project from the ventral surface of the thorax and are used for swimming reaching an average speed of 3 cm/s (Maleschlijski et al. 2012). Anterior part includes a pair of compound eyes and a naupliar eye, as well as cement gland and oil cells that provide the nonfeeding cyprid with energy. The first pair of antennae (antennules) is located anteriorly and can be extended far beyond the carapace as well as fully retracted into the mantle cavity. Each antennule comprises four segments, from which the third one is modified into an attachment disc. The attachment disc is densely covered with cuticular villi and has a multitude of pores for the secretion of temporary adhesive. The temporary adhesive is a proteinaceous substance deposited from antennules with each step. It contributes to adhesion during surface exploration as well as has a pheromonal activity attracting other cyprids (Dreanno et al. 2006a).

1.4 Cyprid sensory organs

The cyprid has a well-developed brain and nervous system compared to other stages in the life cycle (Harrison and Sandeman 1999). Cyprids need it to sort and process input from various sensory organs, including antennules, lattice organs, sensory setae distributed all over the body and eyes, and to coordinate a relevant behavioural response.

It is believed that exploration of the substratum predominantly relies on the antennules (Figure 3) (Bielecki et al. 2009, Maruzzo et al. 2011). An antennule consist of four segments connected by joints allowing rotation and movements in different directions during surface exploration (Maruzzo et al. 2011). The third segment is short and carries a bell-shaped attachment disc. Cuticullar villi of the attachment disc were suggested to play role in the exploration of topographical features of the surface (Maruzzo et al. 2011). Indeed, hair-like mechanosensitive structures are ubiquitous in nature, spanning the inner ear of mammals for hearing and lateral line of fish for flow detection (Rizzi et al. 2015). The basic mechanism is believed to include a mechanical deflection of the hair-like structure that is converted into a neuronal signal transduced further to the nervous system (Rizzi et al. 2015). However, this has not yet been shown experimentally in cyprids.

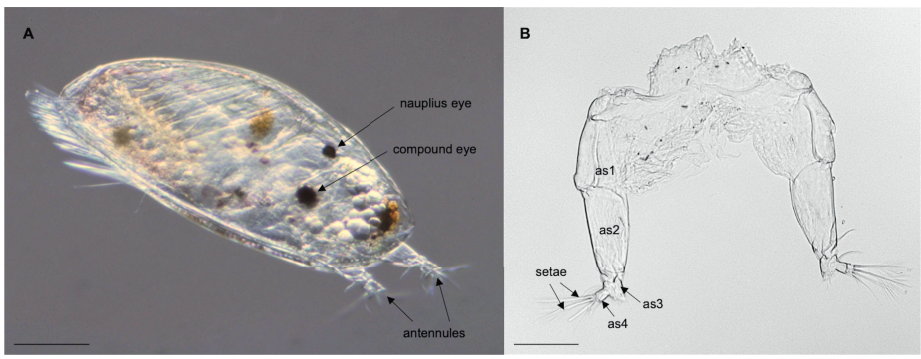


Figure 3. B. improvisus cyprid larva and a pair of antennules. (A) Lateral view of a cyprid with the pair of antennules and eyes indicated, scale bar = 100 µm; (B) a dissected pair of antennules with sensory setae and antennular segments (as1 - 4) indicated, scale bar = 50 µm.

The fourth segment has a number of setae of different length and curvature mediating both chemo- and mechanoreception (Bielecki et al. 2009). Some setae have a terminal pore and are suggested to perform contact chemoreception, while others are sac-shaped and thin-walled, called aesthetascs, that potentially can sense waterborne compounds (Maruzzo et al. 2011). During surface exploration the fourth segment was observed flicking - a behaviour typical in other crustaceans to facilitate sampling of chemical information in the surrounding environment (Breithaupt and Thiel 2010). It has been suggested that the majority of chemosensory setae, except the aesthetascs, are bimodal and also have a mechanoreceptive function (Bielecki et al. 2009). However, the nature of the receptors involved in sensing settlement cues has remained elusive. Identification of barnacle sensory receptors would be a considerable asset and a step forward, both for studying external factors affecting the settlement behaviour as well as for the development of new antifouling technologies. Apart from antennules, cyprids have several other sensory organs that can be important during settlement. Porous fields of cuticle located on the dorsal side of the carapace, called lattice organs, are assumed to have chemosensory function. While unlikely to be used for substrate exploration the lattice organs might be important during the pelagic phase of exploration, e.g sensing of waterborne pheromones (Høeg et al. 1998). Furthermore, cyprids also have a pair of frontal filaments located on the ventral side of the carapace. The exact function of these structures remains unknown, however they were suggested to have chemo- or photosensory function or being involved in orientation or pressure sensing (Walker 1974). The caudal rami are appendages projecting from the ventral surface of the thorax and bearing setae. It has been observed that caudal rami touch the surface during the inspection stage of surface exploration, and might play a mechanosensory role (Crisp and Barnes 1954). In particular, during inspection cyprid draws arcs with caudal rami on the surrounding surface potentially testing the microtexture of the substrate as well as checking for obstacles (Crisp 1974). Furthermore, cyprid has one nauplius eye and a pair of compound eyes (Figure 3, A). The nauplius eye is present from hatching to settlement while the compound eyes appear only in cyprid and therefore might be important during settlement (Matsumura and Qian 2014). A study of *B. amphitrite* settlement suggests that cyprids can locate conspecific adults based solely on vision. Shell of adult *B. amphitrite* have been shown to emit red fluorescence that might serve as a specific signal for cyprids (Matsumura and Qian 2014). Moreover, cyprids were able to differentiate between adult barnacle and similar-sized objects such as stones, suggesting that visual input also plays an important role during settlement.

1.5 Cyprid exploratory behaviour

Cypris larva developed a complex exploratory behaviour in order to assure settlement under favourable conditions. The original model of cyprid exploratory behaviour on surfaces includes three patterns: wide search, close search and inspection (Crisp 1974). First, the cyprid performs wide search by walking almost in a straight line on its pair of antennules investigating a substantial area of the substrate. This stage is followed by a close search when the cyprid makes short steps and turns exploring a more narrow area, being attached with only one antennule and probing the surroundings with the other one. During the inspection, the last stage preceding adhesion, the larva remains on one spot with both antennules on the surface performing testing right at the attachment site. Recently, this behavioural model was confirmed by video tracking and automated quantitative 3-dimensional analysis showing that this behavioural pattern is largely accurate and conserved between individual cyprids (Aldred et al. 2018). The analysis revealed that wide searching and close searching are discrete and consecutive behaviours, while inspection can occur at other time points. Important observation is that close searching leads directly to settlement in the absence of interrupting factors. Stereoscopic tracking followed by visual analysis of trajectories added several other motion patterns including spiralling, swimming and sinking, when the thoracopods stop beating and cyprid passively sinks in the water column (Maleschlijski et al. 2015). Sinking was observed as the main pattern leading to a contact with a horizontal substrate.

Several studies showed that duration and types of exploratory behaviour differ depending on the underlying surface properties. For instance, analysis of *B. improvisus* cyprid behaviour revealed that they spend more time exploring smooth surfaces and more time swimming when presented with micro-textured surfaces (Berntsson et al. 2000), whereas differences in step duration and length were found when surfaces with different chemistries were used (Chaw and Birch 2009).

1.6 Settlement cues

Life histories of many marine benthic invertebrates are rather complex and often include free-living planktonic larvae responsible for dispersal and subsequent settlement and metamorphosis. Various environmental cues - chemical, physical and biological - acting at different spatial scales, are used by marine larvae to locate a suitable substrate for survival and reproduction during the adult stage of their life cycle. Considerable experimental evidence exists that chemical cues are one of the main factors influencing settlement in many different species, including molluscs (Painter et al. 1998), polychaetes (Lam et al. 2003) and bryozoans (Dahms et al. 2004). Chemical cues originate from different sources including specific prey or host, conspecific individuals or biofilms, and can be waterborne or absorbed to the substrate. Despite considerable experimental evidence exists that chemical cues are very important for substrate selection by larvae, identity of only a handful of natural chemical cues was characterised. In many known cases of settlement induction and metamorphosis, the substances mimic the action of neurotransmitters with no evidence that the neurotransmitter or neuroactive substance is related to the settlement cue itself (Morse 1990, Rodriguez et al. 1993). These examples include choline, DOPA, dopamine, norepinephrine, epinephrine, and arachidonic acid that have been observed to induce settlement and/or metamorphosis, with no relation to a natural settlement cue (Morse 1990). Such experiments help to understand participation of the nervous system in the control of settlement and metamorphosis, however, they do not reveal the identity of the exogenous cues or its receptors.

During early days of barnacle research, settlement was considered as a purely random process governed predominantly by currents. However, considerable evidence has been accumulated suggesting that cyprids developed surface selectivity and are able to delay metamorphosis to ensure settlement under favourable conditions. The relative importance of settlement cues still remains unknown, probably because under natural conditions cyprids are affected by many factors simultaneously and it is hard to reproduce this interactive effect in the laboratory assays. It was suggested that cyprids used biological cues performing broad exploration, while looking for a clean site during close exploration and microtextured surface to enhance adhesion during the inspection phase (Le Tourneux and Bourget 1988, Berntsson et al. 2000).

1.6.1 Barnacle pheromones

Gregarious behaviour in barnacles, ie the attraction of conspecific individuals and the subsequent settling in dense communities, increases the probability of reproduction and is mediated by settlement pheromones (Clare and Matsumura 2000). It has been noticed long time ago that cyprids preferentially settle in the presence of conspecifics adults. Since Knight-Jones and Stevenson almost 70 years ago described barnacle gregarious settlement behaviour for the first time (Knight-Jones and Stevenson 1950), considerable advances have been made in the understanding of the mechanisms of this behaviour.

Early research of the gregarious settlement led to the discovery that arthropodin, a glycoprotein of arthropod cuticles, induced settlement of *Balanus balanoides* cyprids on treated surfaces (Crisp and Meadows 1962). Later studies resulted in the identification of the contact pheromone in *B. amphitrite*, called settlement inducing protein complex (SIPC), that was purified from homogenates of whole adult barnacles by ammonium sulfate precipitation, ion-exchange chromatography, gel filtration, and lectin-affinity chromatography on lentil lectin-Sepharose. It was shown that SIPC is a 171 kDa cuticular glycoprotein with similarity to the α2-macroglobulin protein family and produced by the epidermis (Dreanno et al. 2006b). SIPC induced cyprid settlement when absorbed on nitrocellulose membranes suggesting that it is active when bound to the surface (Matsumura et al. 1998), where cyprids could potentially detect it by their antennules while walking on the surface of adult barnacles. Furthermore, SIPC was detected in the cyprid temporary adhesive suggesting that it is thus involved in both adult-larval and larva-larva interactions at settlement (Dreanno et al. 2006a).

SIPC-like proteins were identified in the crude extracts of diverse barnacle species but not from other non-barnacles organisms. This suggests that SIPS are ubiquitous in and specific for barnacles (Kato-Yoshinaga et al. 2000). While cyprids preferentially settle in response to conspecific SIPC extracts, they also respond to a lesser extent to allospecific SIPC, suggesting that SIPC in closely-related barnacle species are rather similar to be recognised by cyprids (Kato-Yoshinaga et al. 2000, Dreanno et al. 2007). While the exact mechanism determining species-specificity remains unknown, it has been suggested that either variable regions of the SIPC or glycosylation patterns might provide the bases for conspecific recognition by cyprids (Yorisue et al. 2012). Further research showed that SIPC has in fact a dual role; it works as a settlement-inducing cue by attracting cyprids, but also as a settlement avoidance cue at higher concentrations informing of overcrowding and

increased reproductive competition (Kotsiri et al. 2018a). A SIPC homologue identified and characterised in *Balanus glandula* (called MULTIFUNCin) induces both gregariousness as well as predation by sea snails (Ferrier et al. 2016, Zimmer et al. 2016).

Apart from the SIPC that induces settlement when bound to the surface, there is also evidence for a waterborne settling factor (Clare and Matsumura 2000). The first evidence of a waterborne cue came from an assay that used seawater conditioned with adults of *Semibalanus balanoides* showing that it contains a factor that induces temporary attachment of cyprids (Rittschof 1985). However, there is still ambiguity regarding the nature of the waterborne cue with several studies reporting varying estimates of the size of the active component of the conditioned water, ranging from 3-5 kDa peptides to less than 500 Da peptides (Clare and Matsumura 2000). In addition, several synthetic di- and tripeptides were tested for settlement inducing activity, showing that peptides containing basic carboxy-terminal amino acids, in particular the glycyl-glycyl-arginine (GGR) peptide, had the tendency to enhance settlement of cyprids (Tegtmeyer and Rittschof 1988). However, this effect could not be reproduced by another study (Clare and Yamazaki 2000). Later, a protein corresponding to 32 kDa was purified from homogenized adult extracts of *B. amphitrite* and was shown to induce cyprid settlement. The protein diffuses into seawater when embedded in an agarose gel and induces settlement, suggesting a waterborne pheromone function (Endo et al. 2009). Interestingly, the obtained N-terminal sequence did not show any resemblance to SIPC or any other proteins in databases. The full sequence of this *B. amphitrite* waterborne pheromone was published in sequence databases in 2012 (called waterborne settlement pheromone “WSP”, BAM34601). The indications of the presence of several WSP homologues in some barnacles species led us to explore the presence of WSP homologues in *B. improvisus*, using transcriptome data from cyprids and adults. Six sequences that are homologous to the published *B. amphitrite* WSP were identified from *B. improvisus* (**Paper II**).

1.6.2 Surface properties

Surface texture is one of the most important factors influencing barnacle larvae settlement (Aldred and Clare 2008). It can affect settlement in many ways, including changes in hydrodynamics, where pits and crevices can protest cyprids from direct hydrodynamic stress, while microtextures may allow better adhesion through more efficient interlocking of the adhesives and

the substrate. Surface roughness in general promotes invertebrates larvae settlement (Crisp 1974). However, *B. improvisus* prefers smooth over rough surfaces to settle (Berntsson et al. 2000). Experiments with moulded surfaces revealed a narrow range of roughness that was inhibitory to settlement of *B. improvisus* larvae, in particularly a roughness height within the range 30–45 μm, an average roughness of 5–10 μm and a roughness width of 150–200 μm (Berntsson et al. 2000). Cyprids engaged less in exploratory behaviour on these surfaces suggesting that reduced settlement occurred through behavioural rejection of the microtextured surfaces (Berntsson et al. 2000). Aldred et al. (2010) suggested that a "likelihood of removal" from the chosen substratum could be a criterion by which cyprids assess potential settlement sites. Indeed, experimental evidence supports that cyprids preferentially settle on the surfaces from which subsequent removal is less likely (Aldred et al. 2010), however, the exact mechanisms remain unclear.

Surface chemistry has been noticed long ago to affect settlement preferences of barnacle larvae, however, it remains currently unclear if it does so via surface charge, surface free energy or surface chemistry. The difficulty of discerning which of the surface characteristics is of the most importance lays in the fact that in laboratory assays test surfaces often differ by more than one parameter, confusing interpretation (Di Fino et al. 2014). Initial studies showed apparent preference of *B. amphitrite* cyprids for hydrophilic surfaces whereas *B. improvisus* showed a tendency to settle more on hydrophobic (Dahlström et al. 2004, Finlay et al. 2010). Surface wettability is a proxy for surface free energy - a measure of the ability of a surface to interact with other materials - which is important when it comes to the binding of an adhesive to the substrate (Petrone et al. 2011). Interestingly, adult barnacles produce differently looking adhesives depending on the surface free energy of the substrate, with thicker and weakly bound adhesives on hydrophilic surfaces but flatter and tightly bound on hydrophobic (Berglin and Gatenholm 2003, Wiegemann and Watermann 2003, Wiegemann 2005).

Currently, it is not known which part of the attachment process is affected by surface wettability. Dahlström et al. (2004) suggested that either cyprids might recognise the chemical groups of the surface signalling the quality of the substrate or physic-chemical forces, eg electrostatic repulsion, might impede cyprids to make a contact with the surface. Since, *B. improvisus* cyprids settle on both surfaces despite preference for the hydrophobic one, a chemical recognition might be involved in the detection of the surface wettability.

1.6.3 Other settlement cues

During the settlement process cyprids also consider other environmental factors, such as light conditions (Lang et al. 1979), flow regime (Judge and Craig 1997) and tidal height (Olivier et al. 2000). Furthermore, presence of microbial biofilms also affects settling of barnacle cyprids. Composition of biofilm can serve as an indicator of local environmental conditions of substratum. Based on observation that cyprids prefer intertidal biofilms over unfilmed surfaces it was suggested that they may distinguish tidal height based on microbial composition (Qian et al. 2003). Apart from various exogenous factors, the physiological condition of larvae also affects attachment and metamorphosis. The condition is determined by the stored energy reserve and age of larva. Thus, when the energy reserve (mainly lipid droplets) drops below a critical threshold level cyprid attaches to a less favorable substratum even in the absence of settlement cues, ie the desperate larva hypothesis (Harder et al. 2001).

1.7 Molecular mechanisms behind settlement

Despite the fact that considerable experimental evidence exists that many marine invertebrate larvae rely on various environmental cues during settlement, molecular mechanisms that translates information from the environment into “developmental decision” to settle or metamorphose are far from fully known. Overall, it is assumed that a settlement cue is sensed by an external receptor that induces influx of ions resulting in the depolarization of membrane and creating signal propagating further to the nervous system. This signal activates a cascade of downstream targets initiating morphogenetic program and changes in behaviour leading to settlement and metamorphosis. Noteworthy there is a striking similarity with respect to the molecular mechanisms between a handful of characterised signalling pathways in marine invertebrate larvae (Clare 1996b). Subsequent sections will describe what is currently known about this in barnacles.

1.7.1 Receptors for settlement cues

It is believed that settlement cues are sensed by receptors on the cyprid antennular setae resulting in the activation of signal transduction and initiation of the larval settlement and metamorphosis. The nature of the receptors remains elusive, however, G-protein coupled receptors (GPCRs) have been thought to be likely candidates for the reception of settlement cues in a number of marine invertebrate species (Tran and Hadfield 2012). There is an evidence of the presence of GPCRs and associate signalling pathways in barnacles (Gohad et al. 2010, Gohad et al. 2012). There are several studies showing involvement of serotonin and octopamine receptors in settlement and metamorphosis (Kawahara et al. 1997, Lind et al. 2010). However, taking into account the diverse physiological roles of these recpetors, it remains unclear if they facilitate settlement through binding the external settlement cues or are mostly involved in the internal signaling transduction.

In the majority of invertebrates studied in details so far, a wide range of environmental stimuli is recognised by chemosensory receptors from two different classes: olfactory receptors (ORs) and gustatory receptors (GRs) (Derby et al. 2016). However, searches for these two types of receptors in crustaceans have so far been unsuccessful (Hollins et al. 2003, Corey et al. 2013, Groh et al. 2014, Groh-Lunow et al. 2015), except for in the *Daphnia* genome where several GRs were found (Penalva-Arana et al. 2009). An additional group of chemosensory receptors called ionotropic receptors (IRs)

was originally described in *Drosophila* (Benton et al. 2009) and later found in other arthropods. Importantly, antennular transcriptomics studies suggest that IRs are the only chemoreceptors found in crustacean antennules (Corey et al. 2013, Groh-Lunow et al. 2015, Derby et al. 2016), therefore making IRs the potential candidate for detection of chemical settlement cues during cyprid surface exploration.

When it comes to mechanoreception, in arthropoda it is accomplished through mechanosensitive ion channels (transporting Ca^{2+} or Mg^{2+}) that are gated upon mechanical stress and allow the influx of calcium ions resulting in a membrane potential (Schuler et al. 2015). Transient receptor potential (TRP) channels play major roles in various sensory modalities such as hearing, hygrosensation, vision and mechanosensation in a diverse set of animals (Peng et al. 2015). Among Arthropods, TRPs have been mainly studied in insects and several of them were functionally characterised, e.g. *Drosophila* TRPN and TRPV have been recently shown to be involved in hearing, hygro- and mechanosensation (Peng et al. 2015). Several types of TRPs have been identified in *D. pulex*, however, their exact function in this species is unknown (Peng et al. 2015). Based on behavioural studies and the use of chemical activators/inhibitors, it has been suggested that cyprid surface exploration might rely on mechanosensitive Ca^{2+} channels, in particular involving the *D. melanogaster* homologs of *painless* and TRPA1 (Kotsiri et al. 2018b).

Recently we identified and characterised several IRs and TRP channels in the cyprid antennules and showed their variable expression during the settlement, thus suggesting possible involvement in the sensing of settlement cues (**Paper I**).

1.7.2 Signal transduction pathways

Pharmacological approach has been widely used to study signalling pathways in larvae of several marine invertebrates, as well as barnacles. Despite there are several limitations in this approach, including off-target effects, it helped to discover several potential components of the signalling pathways.

Chemosensory signal transduction in general involves two main intracellular signalling pathways, one mediated through cAMP and another though phosphoinositide-derived signals (Ache and Young 2005). Both pathways target ion channels that upon activation allows calcium entry generating membrane potential that propagates through the nervous system.

Pharmacological studies suggest that barnacles also follow the common theme. In particular, there is a strong evidence for cyclic AMP (Clare et al. 1995), calcium (Clare 1996a) and protein kinase C (PKC) (Yamamoto et al. 1995) involvement in settlement and metamorphosis (Zhang et al. 2012).

According to the current hypothetical scheme of the signal transduction involved in barnacle settlement (Clare 1996a) (Figure 4), binding of the settlement cue(s) to external receptors, e.g. G-protein linked receptors, activates activates adenylate cyclase that produces cAMP and through the cascade of this secondary messenger induces influx of calcium. Change of membrane potential leads to efflux of chloride ions generating a membrane potential that further propagates into the nervous system. Based on studies of inhibitors and modulators of cAMP it was suggested that cAMP is involved in the *B. amphitrite* settlement acting proximal to the settlement cues receptors (Clare and Matsumura 2000).

From studies in other crustaceans there is an evidence that phosphoinositide pathway is involved in olfactory receptor cell chemosensory transduction (Ache and Young 2005) Interestingly, presumptive targets of the phosphoinositide signalling in lobster are representatives of TRP channels, in particular TRPV and TRPM subfamilies (Bobkov and Ache 2005), that we earlier identified in the cyprid antennules (**Paper I**). However, pharmacological studies in barnacles suggest that phosphoinositide pathway is distant to the settlement cues receptors and regulates initiation of metamorphosis (Clare and Matsumura 2000).

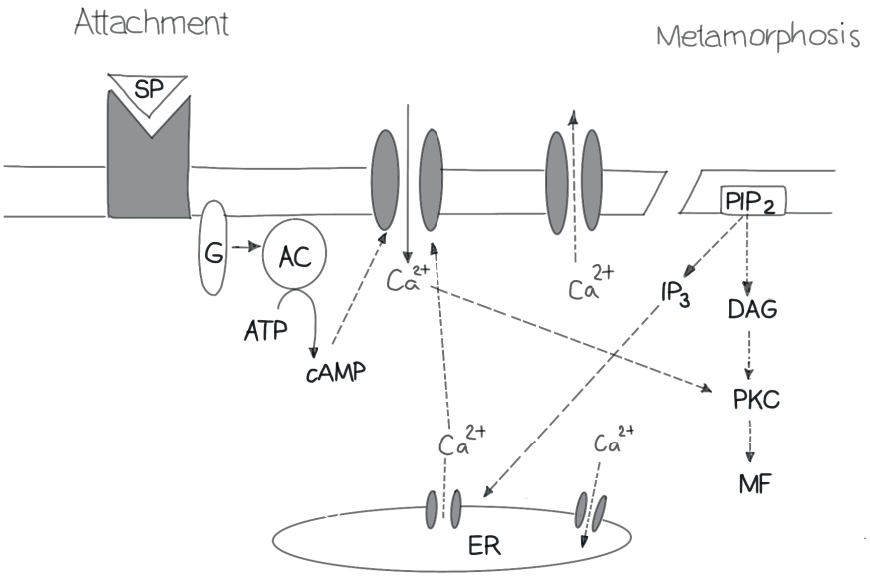


Figure 4. Hypothetical scheme of signalling transduction during barnacle settlement: binding of a settlement pheromone (SP) to an external G-protein (G) coupled receptor activates adenylate cyclase (AC) that produces cyclic AMP (cAMP); cAMP acts on ion channels inducing influx of calcium, resulting in attachment. Phosphatidylinositol pathway acts on protein kinase C (PKC) which regulates methyl farnesoate (MF) and induces metamorphosis. SP settlement pheromone, G G protein, AC adenylate cyclase, Reproduction of scheme from AS Clare 1996.

1.7.3 Endocrine control of molting and metamorphosis

Right after the attachment cyprid begins molting that includes apolysis, separation of old cuticle from epidermis, degeneration of muscles and new cuticle formation (Freeman and Costlow 1983). Metamorphosis can be regarded as a special molt event and much of the known mechanisms are informed from the studies on molting in the model organism *D. melanogaster* (Ventura et al. 2018). Regulation of molting associated with metamorphosis is well characterised in insects, where it is initiated by a sharp increase in the ecdysone titre that triggers a cascade of tissue-specific gene expression through a hierarchy of ecdysone-responsive genes. In comparison to insects, the ecdysteroids-induced cascade in crustaceans, and barnacles in particular, is less well characterised. The majority of gene homologs involved in biosynthetic and signalling pathways of ecdysteroids and sesquiterpenoids can be identified in non-insect arthropods implying that these two hormonal systems were present in the last common ancestor of arthropods (Qu et al. 2015). It has also been suggested before that the process of surface exploration and metamorphosis in barnacles have certain resemblance to wandering and metamorphosis in insects, with cyprids corresponding to the pupal stage (Kotsiri et al. 2018b). Altogether, this suggests a possibility of similar mechanisms of the control of molting and metamorphosis in barnacles and other arthropods and will be used as a framework and working model for further discussion.

Studies in *B. amphitrite* revealed the presence of the key arthropod hormones regulating molting and metamorphosis, 20-hydroxyecdysone (20E) and methyl farnesoate (MF) (Yamamoto et al. 1997a, Yamamoto et al. 1997b). Most of the previous studies showed increase in attachment and metamorphosis when 20E applied exogenously (Clare and Matsumura 2000). Methyl farnesoate was proposed to be the analog of juvenile hormone in insect (Smith et al. 2000). Application of MF has been shown to impede settlement and metamorphosis at low concentrations and induce precocious metamorphosis without attachment at high concentration leading to developmental abnormalities (Yamamoto et al. 1997a, Yamamoto et al. 1997b). Results of the previous studies suggested a link between MF and PKC-dependent signaling in cyprid metamorphosis (Clare et al. 1995, Clare 1996b). In *D. pulex*, PKC acts upstream of MF signalling during sex-determination (Toyota et al. 2015). Based on the *Daphnia* model and experiments, Kotsiri et al. (2018b) proposed a working model where signals from surface exploration are decoded at the site of MF synthesis modulated by PKC. The extent of the surface exploration is controlled by MF synthesis that suppress the surge of ecdysteroids titre required to proceed to

metamorphosis. This corresponds to the current view on regulation of crustacean metamorphosis where MF acts as a metamorphic inhibitor, despite that it still remains unknown how exactly it regulates the ecdysteroids synthesis (Hyde et al. 2019b).

Overall, previous studies suggest that MF and 20E regulate morphogenetic program that results in metamorphosis in barnacles, but currently nothing is known how the interplay of these hormones translates into the wide-scale transcriptomic shift observed.

2 OVERALL OBJECTIVES OF THE THESIS

The overall aim of this PhD project was to examine molecular mechanisms of different aspects of barnacle settlement. Specific aims are:

- To get a molecular description of the entire settlement process, including pre-settlement stages, we performed transcriptome profiling of four different stages, i.e free-swimming, close-search, attached and juveniles.

- To examine the gene expression on surfaces with different wettabilities, hydrophobic and hydrophilic, with the aim to understand events involved in the cyprids' decision-making.

- To get a better resolution of the repertoire of receptors involved in settlement cues sensing by performing antennule transcriptomics.

- To further investigate barnacle chemical communication by examining WSP homologues in *B. improvisus* and investigating their expression during settlement.

- To bring forward first insights into the complexity of the barnacle genome by characterisation of the draft genome assembly of *B. improvisus.*

3 METHODOLOGICAL CONSIDERATIONS

This chapter's aim is to give an overview of different methodologies and techniques used in the papers comprising this thesis. The attached papers contains more detailed descriptions of the materials and methods used.

3.1 Study organism *Balanus improvisus*

The bay barnacle *B. improvisus* Darwin, 1854 is a cosmopolitan species that is one of the most common biofouling organisms in tropical and temperate seas (Kerckhof 2002, Berntsson and Jonsson 2003). It is believed to originate from the east coast of the American continent and there are indications that Argentina could be part of its native region (Wrange et al. 2016). The species was probably introduced to Scandinavian waters in the 19th century by shipping (Blom and Nyholm 1962). *B. improvisus* prefers brackish conditions but is capable of living in waters with salinities ranging from 1.6 psu up to 40 psu, and it is the only barnacle species surviving the brackish conditions of the Baltic Sea (Wrange et al. 2014, Blomberg et al. 2019). This barnacle is commonly found on rocks, jetties, and boat hulls, and is therefore of general interest for understanding the mechanisms of biofouling in marine and brackish waters. It has become a model organism for the investigations of settling biology, in particular in relation to antifouling research and the mechanisms involved (Holm 2012).

Similar to most other barnacles, *B. improvisus* is predominantly hermaphroditic with cross-fertilization, however, cases of self-fertilization were observed both in the laboratory and under natural conditions (Furman and Yule 1990). The reproductive period of *B. improvisus* is continuous through June to September with distinct spawning peaks in late July and late August (Berntsson and Jonsson 2003). *B. improvisus* is a useful model for studying settling biology, however, natural seasonal spawning yields an unpredictable supply of cyprid larvae for studies. A protocol for the all-year-round culturing of *B. improvisus* at Tjärnö marine station has been developed and a detailed description of all steps in the production line has been outlined (i.e., the establishment of adult cultures on panels, the collection and rearing of barnacle larvae, and the administration of feed for adults and larvae) (**Paper III**). The *B. improvisus* cyprids reared in the culturing facility were used for the settlement assays (**Paper IV**), for the genome estimation (**Paper V**) as well as for the dissection of antennules (**Paper I**).

3.2 Genomic resources

Substantial amount of knowledge has been accumulated over the years of research regarding settlement biology of different barnacle species, including *B. improvisus*, however, there is an apparent lack of genomic data that could provide molecular resolution. Currently, there is no reference genome available for any barnacle species and even the genome size for most of the species remains unclear or information is missing.

Paper V provides the first characterisation of *B. improvisus* genome. This analysis revealed high genetic diversity in *B. improvisus* between two alleles within one single individual (roughly 5% in coding regions, even great in intergenic regions, **Paper V**). This makes the genome assembly process particularly challenging, resulting in a mosaic of arbitrarily alternating alleles and a highly fragmented assembly. To minimize the problems related to the high genetic diversity, it is desirable to use a single individual for extracting DNA as a base for the sequencing. However, *B. improvisus* is a relatively small barnacle, 5-12 mg dry weight per individual, making it difficult to obtain enough good quality DNA from one single individual. In addition, the DNA amount and quality from each individual are highly variable making DNA extraction for library preparation tedious and results quite unpredictable (Panova et al. 2016). The high genetic diversity and the low amount of DNA from a single individual have constituted great challenges when using short-read sequences for the generation of a reference genome (**Paper V**). Despite that, the draft genome assembly of the *B. improvisus* based on a short-read sequencing assembly (**Paper V**) can be useful in order to investigate specific genes and gene families (Lind et al. 2010, Lind et al. 2013, Lind et al. 2017).

There is a rapid increase in transcriptomic studies investigating different aspects of barnacle biology, including genes involved in cementation (Wang et al. 2015), hormonal (Yan et al. 2012) and osmotic regulation (Lind et al. 2013), as well as molting and settlement (Chen et al. 2011, Lin et al. 2014). Despite the fact that sequencing platforms generate enormous amounts of data at a reasonable price and constantly increasing read length, there are considerable challenges associated with analysis of the sequencing data from non-model organisms. Transcriptome profiling is limited by sampling only a fraction of transcripts expressed in a particular tissue, developmental stage or condition (Carninci et al. 2005). Furthermore, transcription levels of coding and non-coding regions will vary greatly across cell types and conditions, with some genes having only a single copy per cell while others millions of copies, eg rRNAs. Therfore, in order to detect lowely expressed

transcripts an appropriate transcriptome sequencing depth is required. The general guidlines provided by studies examining the optimal coverage for the differential gene expression analysis suggest that 10 to 30 M reads together with sufficient replication provides enough power to detect differentailly expressed genes between samples (Liu et al. 2013, Sims et al. 2014).

In the absence of a reference genome, analysis of the short read transcriptome data depends on the *de novo* transcriptome assembly, which is prone to errors, potentially resulting in artificial chimeric sequences representing mixed contigs; ie. a mix of alleles, paralogs and/or pseudogenes, as well as an overestimation of the total number of transcripts. As an example, a recent study in *B. amphitrite* (Wang et al. 2015) reported an assembled transcriptome size of ~114 Mbp which was more than four times greater than the number previously reported by Chen et al. (2011). In the case of *B. improvisus* that have an extremely high genetic diversity (**Paper V**), creating a good *de novo* transcriptome assembly is a challenge. Alleles that differ by almost 5% in their coding regions are kept separate by the assembly program, and can often be the cause of mis-assembled reads. This altogether, results in the portion of reads assigned to several places during the calculation of gene abundances leading to the erroneous gene expression levels. In the absence of the reference genome for the *B. improvisus*, we tried to overcome this by improving the assembly through clustering and filtering out low-quality and lowly expressed transcripts. In case of several selected genes of particular interest to us, we manually inserted curated sequences (and removed the corresponding de novo-assembled contigs) to avoid the complexity from multi-mapping of reads leading to better expression estimates.

As of 01/09/2019, NCBI's BioProject database returns 518 crustacean transcriptome entries comprising 3448 sequence read archives. Despite this growing amount of data on crustaceans and abundance of data on insects, often less than 30% of crustacean NGS assemblies being annotated by any protein database (Hyde et al. 2019a). Currently, the sequencing datasets are dominated by the transcripts that do not have any homology to publicly available reference genomes. As a result, the function of these putative or hypothetical transcripts remains unknown. These transcripts could represent novel species-specific genes that are important for a particular stage in development. Currently, analysis of these species-specific genes remain challenging but provide very interesting resources for further structural and functional studies.

3.3 Sample collection

Settlement bioassays is a widely used method in the laboratory, not only for more in-depth studies of barnacle settlement, but also to test biocide toxicity and surface/materials effects on larval settlement and metamorphosis (Holm 2012). Settlement assays with cyprids are usually performed in polystyrene petri dishes under static laboratory conditions with test condition being a single varying factor, e.g external substances, surface chemistry or surface structure, etc (Berntsson et al. 2000, Dahlström et al. 2004). Typically, a number of cyprids are divided into petri dishes containing sea water and kept over a certain amount of time. Settlement is evaluated using a microscope by counting settled cyprids (attached and metamorphosed) and dead or living cyprids. It is a quick and convenient method, however, there are several important things that should be taken into consideration.

First of all, settlement propensity of cyprids vary greatly in laboratory assays (Holm 2012). All larvae rarely respond in the same way and there will always be some cyprids that settle while others do not. This variability is thought to arise from genetic, maternal as well as environmental factors together and can have a confounding effect on the settlement assay results (Holm et al. 2000, Head et al. 2004). To avoid these effect and to make the results more generally applicable, Holm et al. (2000) proposed to use replicating experiments with larvae from different broodstocks for each replicate (i.e. making the replicates fully independent). Taking this into consideration, we used three independent batches (from different trays with different parents) for the repetition of the experiment in **Paper IV**. Briefly, nauplius larvae were collected from three independent trays of the culture with panels containing unique adult barnacles, and these were reared separately until reaching the cyprid stage. In this way we obtained three independent cyprid batches with different parental background. The advantage of this approach is more generalised and applicable results, however, it also introduces more variability in the data resulting in somewhat high standard deviation in the gene expression analysis. The abundance of a single gene could vary for up to 67% between replicates. To remove variation associated with batch effect we applied a normalization proceedure that reduced the variation to 30% and increased the statistical power for detecting more of the differentially expressed genes.

Furthermore, the density of cyprids in the assay can also affect settlement due to gregarious effects (Holm et al. 2000, Head et al. 2004). For instance, it has been shown that footprints deposited by *B. amphitrite* cyprids stimulate

the attachment of conspecific larvae (Clare et al. 1994). Therefore, a single-larva assays are considered the best option to avoid the density-dependent effects (Clare et al. 1994). However, we had to optimize the number of cyprids to meet the requirements for the amount of RNA needed for sequencing (~500ng) and used 20 cyprids per petri dish (**Paper III**).

It has been shown that physiological condition of cyprids changes with ageing, which affects their responsiveness to cues and willingness to settle (Dahlström et al. 2000, Harder et al. 2001). Newly hatched larvae of many marine invertebrates have a pre-competent state where they are not able to respond to settlement cues (Degnan and Degnan 2010). There is evidence suggesting that barnacle cyprids also have a pre-competent state, and newly metamorphosed day 0 cyprids are less prone to settle (Maréchal et al. 2012). For the experiment in the **Paper IV** we used one-to-two days old cyprids of *B. improvisus* that were settled for 4-5 days, therefore eliminating any strong influence from age-related differences.

For **Paper IV**, we collected individuals at different stages of settlement: free-swimming, close search, attached and juveniles. The classification of pre-settlement behaviours were based on earlier descriptions (Lagersson and Hoeg 2002). Free-swimming stage included all behaviours when cyprids were not laying passively on the bottom or antennules were not in contact with the surface. Close-search was characterized by larvae making short steps and turns while exploring the surface in the vicinity (not longer than the body length). We did not collect the inspection stage because it is extremely difficult to distinguish the stationary larva testing the substratum from the permanently attached cyprid (Maruzzo et al. 2011). We also omitted the wide-search stage based on the results from a pilot study (2015) showing that gene expression profile in the wide-search is very similar to the close-search, and the latter is most probably where the decision-making process takes place. Attached stage corresponds to a time period of 4 to 7 hour after cementation when the larval body is oriented parallel to the surface with the carapace tightly applied to the substratum, as described by Maruzzo et al. (2012). Juveniles were collected at the stage when the body shape resembles an adult barnacle with the ventral side in contact with the surface, and with the thoracopods continuously beating inside or extended outside. All stages were observed under stereomicroscope and collected with tweezers after at least 20s of observing the characteristic behaviour for each cyprid.

The aim of **Paper I** was to generated transcriptome data from cyprid antennules in order to enrich for sensory genes. Free-swimming cyprid

larvae, approximately 2-3 days old, were collected, placed in RNAlater (Ambion), and stored at 4°C for two days. Subsequently, individual larvae were placed on a glass slide and the pair of antennules was separated at the base (see Figure 3) of the proximal segment by means of stainless steel insect pins under a stereomicroscope. Each pair of antennules was then transferred into the sample buffer for RNA extraction. In total, three independent samples were collected each containing antennular pairs from 50 individual cyprids, and these samples were finally pooled to obtain enough RNA for sequencing.

3.4 Settlement on two types of surfaces: hydrophobic and hydrophilic

Two different types of surfaces were used in the **Paper IV** to examine effects of surface wettability on the gene expression during settlement. The choice of the surfaces was guided by the previous study by Dahlström et al. (2004) showing that settlement on hydrophilic dishes was reduced by 38% compared to hydrophobic.

For the settlement assay, we used exactly the same types of petri dishes except differences in wettability, ie. petri dishes with standard hydrophobic surface (Nunc No 150340, Ø 48 mm) and the same type of petri dishes but with hydrophilised surface (Nunclon^{tm} Delta, Nunc No 150326, Ø 48 mm) though a physical treatment with energy beams and oxygen making the molecules of the resin form carboxyl and hydroxyl groups. Contact angle measurements were used to determine wettability change during the course of the experiment of the test surfaces, with or without cyprids. The dishes were collected and stored for up to a week in humid conditions in the 4C prior to measurements were performed. Before the measurement the surfaces were rinsed with water. Three measurements per dish were taken and an average estimated. The drop-spread method was used to estimate the contact angle at the air-surface-liquid interface. This method is conceptually easy to apply for a flat, smooth and chemically homogeneous surfaces (Susana et al. 2012). In reality, surfaces often depart from these assumptions due to surface roughness or chemical heterogeneity. When measuring contact angles on the experimental surfaces, salt crystals were observed in some cases, even after rinsing with water, distorting the water droplet. Since we used exactly the same surfaces as previously been used by Dahlström et al. (2004), the discrepancy in the reported measurements of the initial wettability between the two studies could be explained by the above mentioned surface heterogeneities. Importantly, however, over the course of 5 days of the experiment, wettability of the two types of test surfaces remained distinct. It should be mentioned that the effect of surface wettability alone on settlement is difficult to measure (Hung et al. 2008). Changing the surface wettability entails changing other parameters of the surface, such as surface chemistry. Our reasoning was to use polystyrene petri dishes that are as identical as possible by most other parameters, e.g. rugosity, modulus, but differ in wettability. For testing our question, if settlement preferences and surface properties are reflected in gene expression, it is enough that the two types of surface have different attractiveness for cyprids independent of the underlying cause.

4 MAIN RESULTS AND DISCUSSION

4.1 Culturing of *Balanus improvisus*

In **Paper III** we present a detailed protocol for the all-year-round culturing of *B. improvisus*. Rittschof et al. (1992) earlier described a method for culturing barnacles from the release of nauplii to the settlement of cyprids for the species *B. amphitrite*. The protocol we presented has been adapted from the all-year-round culturing of *B. improvisus* at Tjärnö Marine Research Laboratory (Sweden), which has been designed over a long time. A detailed description of all steps in the production line is outlined, including the production and rearing of barnacle larvae, as well as the administration of feed for adults and larvae. The barnacle *B. improvisus* has several advantageous features as a model for studies of osmoregulation because of its extreme low-salinity tolerance (Blomberg et al. 2019), with a focus on molecular and physiological mechanisms as well as ecological interactions and evolutionary consequences. It is also widely used as a model for the investigations of settling biology of larvae, in particular in relation to antifouling research and the mechanisms involved. However, natural seasonal spawning yields an unpredictable supply of cyprid larvae for studies. The ability to culture this barnacle through its whole life cycle all year round is, therefore, a major asset to enable various types of molecular and mechanistic studies. In addition, its presence in marine/brackish waters worldwide allows for a combination of field and experimental studies. Controlled breeding can also produce families of known pedigrees for long-term culturing, and a generation time of a few months may allow long-term experimental evolution.

4.2 A pilot study for the sequencing of the *B. improvisus* genome

The aim of **Paper V** is to provide the first genomic characterization of the *B. improvisus* genome. Despite ecological and economical importance, there is currently no reference genome available for any of the barnacle species and even the available genome size estimates are unclear. In this paper we here provide an initial genomic characterization of the bay-barnacle *B. improvisus*, including genome size, repeat content and nucleotide diversity.

Conflicting estimates of genome sizes of 740 Mbp (Rheinsmith et al. 1974) or 1400 Mbp (Bachmann and Rheinsmith 1973) have been reported for *B. amphitrite*, with two other Balanidae species reported to have similar large sizes of 1260 Mbp (*Semibalanus cariosus*) and 1400 Mbp (*B. eburneus)* (Rheinsmith et al. 1974). Due to the fragmented nature of the draft genome assembly, it was not possible to reliably predict the genome size for *B. improvisus*. We therefore performed an experimental genome size estimation of *B. improvisus* based on flow cytometry of propidium iodide stained nuclei. Analysis resulted in an estimated haploid genome size of 738 Mbp (+- 9 Mbp) (Figure 5), supporting the previous lower-end genome size estimate of *B. amphitrite* (Rheinsmith et al. 1974).

Alignment to the repeat libraries available for the other species and a constructed species-specific library for *B. improvisus* resulted in 1% and 17% repeat content, respectively. The large difference in estimates between the previously known repeats from other species compared to species-specific repeats, indicates that the *B. improvisus* repeats are novel or very divergent compared to previously known repeats. An analysis of the raw genome sequencing reads, instead of the draft genome assembly, resulted even in a higher estimate of 40% repeat content. Since the detection methods can be expected to be underestimating rather than overestimating the repeat content, we conclude that at least 40% of the *B. improvisus* genome is represented by high-copy repeat families.

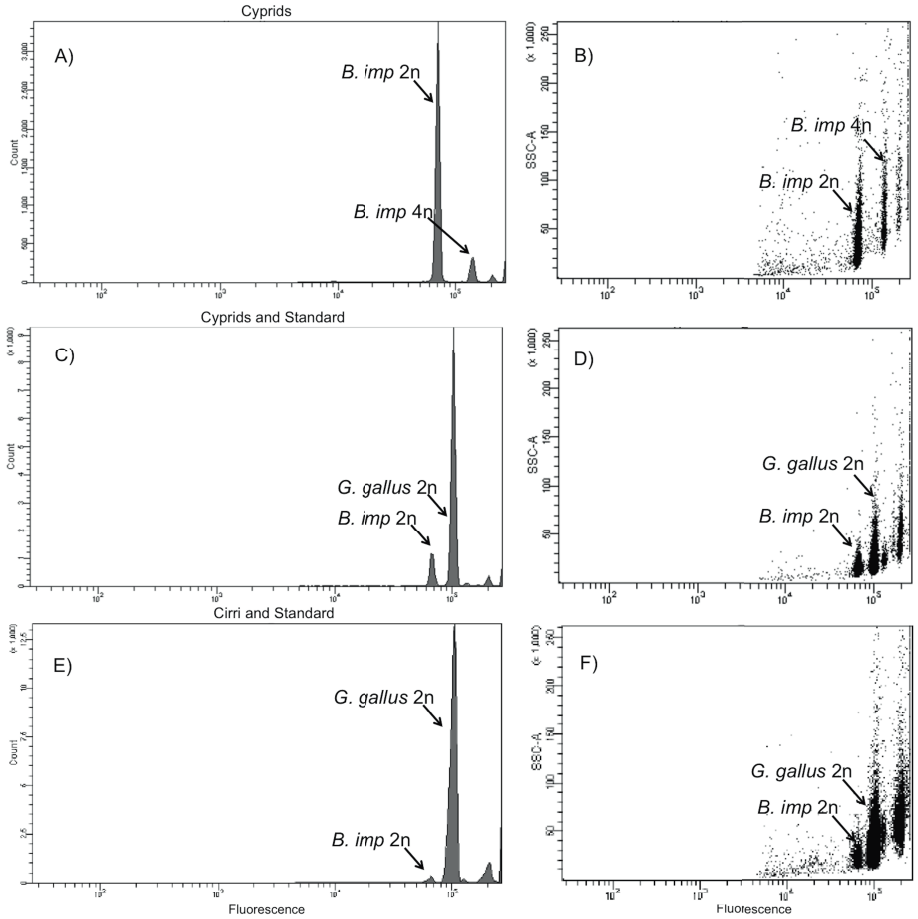


Figure 5. Genome size estimation for the B. improvisus using flow cytometry: flow cytometry of propidium iodide-stained nuclei from cyprids and cirri of Balanus improvisus and the chicken blood standard Gallus gallus (C-value of 1,14 Gbp) nuclei. (A-B) Histograms showing fluorescence peaks of B. improvisus 2n and 4n nuclei extracted from cyprid larvae and corresponding forward scatter light (FSC-A; reflecting size of nuclei/particles) distribution against fluorescence. (C-D) cyprid 2n nuclei and reference G. gallus 2n nuclei (E-F) adult barnacle cirri and chicken 2n nuclei.

The most striking feature of the *B. improvisus* genome is the extremely high genetic diversity, with about 5% variation in coding regions and much higher variation in the non-coding parts. This is among the highest nucleotide diversity described for any species so far, but in line with very high levels of genetic diversity observed also in other marine species (Leffler et al. 2012). Intronic and intergenic region as a rule are more variable than coding regions therefore, it is highly likely that the current estimate of the nucleotide diversity is an underestimate since it relies on the coding regions, which constitute only a small portion of the genome. Although draft genome assemblies as the one presented here can be useful in order to investigate specific genes and gene families (Lind et al. 2010, Lind et al. 2013, Lind et al. 2017), a reference barnacle genome is urgently needed to advance science in this important group of organisms. Based on the nucleotide diversity, in combination with a high repeat content, we advice to opt for long-read sequencing technologies when planning a barnacle genome project; this could then aim for allelic resolution and a complete diploid genome assembly.

4.3 Insights into the molecular mechanisms of *B. improvisus* settlement and discriminatory behaviour towards surfaces

Paper IV represent the results of a study that was undertaken to get a better understanding of the molecular mechanisms during barnacle settlement. The first hypothetical scheme of molecular mechanisms of barnacle settlement was proposed in Clare (1996a) and based on comparative studies from other marine larvae and pharmacological data from barnacles (Figure 4). According to the scheme, a settlement pheromone binds to G-protein coupled receptors and activates cAMP and increases intracellular calcium resulting in depolarisation of membrane and propagation of signal to the central nervous system. Metamorphosis on the other hand is suggested to be controlled through the phosphatidylinositol pathway activating protein kinase C (PKC) that regulates MF levels (Figure 4). Recently, Kotsiri et al. (2018b) added to the understanding of the molecular mechanisms, suggesting that the positive signals stimulate MF secretion that in turn suppress ecdysteroids synthesis eventually resulting in timely completion of metamorphosis (Kotsiri et al. 2018b). However, the downstream genes responding to these signals and responsible for unraveling the morphogenetic program in barnacles remain unknown. To get an overview and a better molecular description of the entire settlement process, we performed transcriptome profiling of four different stages, i.e. free-swimming, close-search, attached and juveniles.

Overall, the *B. improvisus* global transcriptional profiling during the different settlement stages revealed highly dynamic changes in gene expression as seen in a heatmap based on Pearson correlation coefficients. It is clear from the heatmap that the two cyprid stages, free-swimming and close-search, displayed very similar gene expression with in fact no significant changes (fold discovery rate 0.05, fold change 2). Cyprid stages were enriched for upregulated genes related to “ion channels activity” GO category, with TRPs and IRs among them. This suggests that these receptor genes expressed during the surface exploration could be involved in sensing of settlement cues and add new candidates to the hypothetical scheme of molecular events during barnacle settlement. Involvement of different types of receptors, i.e. mechanosensory and chemosensory, further suggests that combined input from various types of receptors acting through different signalling pathways could provide the basis for fine-tuned recognition of the suitable settling spot as well as amplify the signal from the favourable environment, as was earlier suggested by Morse (1993).

The transition from close-search to the attached stage displayed most of the changes, with a larger proportion of genes exhibit upregulation. Upregulated genes display an enrichment in processes related to molting and cuticle development as well peroxide metabolic processes. Interestingly, among the upregulated genes in the attached stage we found 37 transcription factors, among which several related to the ecdysone pathway, e.i. nuclear hormone receptors *FTZ-F1*, *HR3* and *HR38,* and these were also specifically upregulated in the attached stage. These genes together with ecdysone receptor (EcR), retinoid-X-receptor (RXR) and broad complex (BRC) are involved in the ecdysone cascade in the other arthropods. The presence of these genes and their differential expression during barnacle settlement adds another level to the understanding of molecular mechanisms during settlement and creates a framework for further investigations.

Additionally, the settlement experiments were carried out on two different surfaces with different chemical properties, hydrophobic and hydrophilic, to gather more knowledge on the cyprids' decision-making process and settling/metamorphosis in relation to distinct surface properties and surface choice. Studies in gastropod *Haliotis asinina* settlement suggesting that settlement cues differ in terms of their inducing effectiveness, with a more effective cue triggering direct and quick pathways leading to more rapid metamorphosis than for a less effective cue (Williams and Degnan 2009). In the case of barnacle cyprid larvae, experimental evidence revealed that cyprids settle faster on a hydrophobic than on a hydrophilic surface, suggesting that the former one is a more effective settlement cue (Dahlström et al. 2004) Our results showed five times more genes upregulated on the hydrophobic surface than on the hydrophilic in the close-search stage, and 10 times more in the attached stage, suggesting that exploration of the "favourable" surface might induce stronger and/or earlier transcriptomic response. This is further supported by the fact that over 30% of the annotated genes more highly expressed when exploring the favourable hydrophobic surface in the close-search stage are related to signaling, including G-protein coupled receptors, MFS-type transporters and phosphatidylinositol kinases. Furthermore, we observed roughly three-fold higher expression of the key gene in the ecdysone-regulated cascade on the hydrophobic surface i.e. *broad complex* gene, as well as genes involved in major developmental pathways, including Hedgehog transcription factors and components in EGF signalling.

4.4 Antennal transcriptomics

In **Paper I** we explored the repertoire of sensory receptors in *B. improvisus* cyprid antennules with the goal to get a better description of the sensory systems involved in the settling behaviour of this species. Barnacle settlement involves sensing of a variety of exogenous cues with a pair of antennules assumed to be the main sensory organ during surface exploration. Antennules are equipped with a number of setae that have both chemo- and mechanosensing function, however, no chemo- or mechanosensory receptors have been characterized to date in barnacles. We carried out transcriptome sequencing of *B. improvisus* cyprid antennules. The generated transcriptome assembly was used to search for ionotropic receptors (IRs) and transient receptor potential channels (TRPs), using our developed HMM models. Among potential chemosensory genes, we identified the ionotropic receptors IR25a, IR8a and IR93a to be expressed in the cyprid antennules. Furthermore, we identified dozens of sequences in the antennal dataset that had some sequence homology to IRs and iGluRs and might belong to the divergent IRs, however, the fragmented nature of the data does not allow to unambiguously classify them. The total size of the IR repertoire in *B.improvisus* is currently difficult to estimate without a complete genome as a reference. Our data shows that common IR subunits (IR25a, IR8a and IR93a) in *B. improvisus* were highly expressed during surface exploratoration and early attachment, suggesting possible importance in detecting settlement cues (Figure 6).

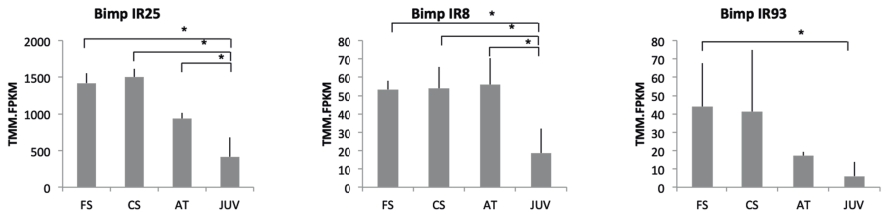


Figure 6. Expression of the B. improvisus ionotropic receptors during settlement. FS free-swimming, CS close-search, AT attached stage, JUV juvenile. Significant changes in gene expression (FDR 0.05) are indicated with an asterisk.

Similarly, these IR subunits are highly expressed in a settlement stage of the parasitic salmon louse *Lepeophtheirus salmonis* (Komisarczuk et al. 2017). In particular, IR25a and IR8b were shown to be crucial for the identification of the host species during the settlement (Komisarczuk et al. 2017). Moreover, studies in both *Drosophila* and the spiny lobster *P. argus* revealed that IRs preferebly recognise water-soluble, hydrophilic acids and amines, which are metabolic derivatives of amino acids (Silbering et al. 2011, Corey et al. 2013). Given that the currently known barnacle pheromones are proteins (Clare and Matsumura 2000, Endo et al. 2009), the here reported IRs potentially present targets to study pheromone-evoked settlement behaviour of cyprids.

Apart from chemical cues, cyprids distinguish surface structures by means of mechanosensory setae (Bielecki et al. 2009, Maruzzo et al. 2011). According to current knowledge, mechanoreception in arthropoda is accomplished through mechanosensitive ion channels that are gated upon stress and allow the influx of cations resulting in a receptor potential (Schuler et al. 2015). Based on the effect of specific agonists of calcium channels, it has been suggested that TRP channels mediate calcium entry into neurons when cyprid senses a favourable substratum (Kotsiri et al. 2018b). With HMM models based on the known arthropod TRP receptors, we detected 13 TRP candidates in the antennular transcriptomes of *B. improvisus* cyprids. This number corresponds roughly to the number of TRP genes found to be expressed in *D. pulex* transcriptome, with the exception that TRPP genes are absent in *Daphnia* (Dong et al. 2013). Several of the TRP genes were differentially expressed during the course of larval settlement. In particular, TRPV that showed expression patterns similar to the ionotropic receptors with high expression in the exploratory stages and decreased expression in the juvenile stage. The TRPV receptor is considered as a molecular integrator of noxious stimuli ranging from pungent natural products, such as capsaicin, to acidic environment and high temperatures (Szallasi et al. 2007). Interestingly, a recent study suggested that cyprid settlement could be akin to the behavioural responses of *Drosophila* to noxious stimuli mediated by TRPs (Kotsiri et al. 2018b). In particular, surface avoidance would be mediated by noxious stimuli-sensing channels, cancelling the signals that complete settlement. This study gives expanded knowledge about the sensory systems present in the barnacles, a taxonomic group for which only limited information is currently available. It furthermore serves as a starting point for more in-depth studies of how sensory signaling affects settling behaviour in barnacles with implications for preventing biofouling.

4.5 *B. improvisus* waterborne settlement pheromones

The aim of **Paper II** was to identify WSP homologues in *B. improvisus* and investigate their expression during settlement. Gregariousness in barnacles, ie the attraction of conspecific individuals and the subsequent settling and aggregation in dense communities, is crucial for reproduction, and is achieved through cyprid settlement in response to chemical cues produced by individuals of the same species (Clare and Matsumura 2000). Earlier studies resulted in the identification of the contact pheromone in *B. amphitrite,* called settlement inducing protein complex (SIPC), which is active when bound to a surface (Matsumura et al. 1998). Apart from the SIPC, there is also evidence for a waterborne settling factor (Clare and Matsumura 2000). According to the current model, cyprids detect the waterborne cue while swimming and respond to it by transition from water to a substrate where the surface-bound SIPC pheromone in turn induces permanent attachment (Elbourne and Clare 2010). A protein corresponding to 32 kDa was purified from homogenized adult extracts of *B. amphitrite* and was shown to induce cyprid settlement. The protein diffuses into seawater when embedded in an agarose gel and induce settlement, suggesting a waterborne pheromone function (Endo et al. 2009). The full sequence of this *B. amphitrite* waterborne pheromone was published in sequence databases in 2012 (called waterborne settlement pheromone “WSP”, BAM34601). Since the original publication of the WSP discovery in 2009 (Endo et al. 2009) there have been no follow-up studies to further elucidate the structure and function of the WSP. However, it was recently indicated that more than one WSP gene is present in *Tetraclita japonica formosana* (Lin et al. 2014) and *B. amphitrite* (So et al. 2016, Wang et al. 2018) but the available information is extremely scarce as only one of the three sequence was published (Lin et al. 2014) and none are deposited in public databases. Moreover, the study by (Endo et al. 2009) mentioned preliminary unpublished data showing that several proteins with molecular mass around 32 kDa detected in barnacle-conditioned seawater have settlement inducing activity. The indications of the presence of several WSP homologues in some barnacles species led us to explore the presence of WSP homologues in *B. improvisus*, using transcriptome data from cyprids and adults. Six sequences that are homologous to the published *B. amphitrite* WSP were identified from *B. improvisus*. Phylogenetic analysis (Figure 7) clearly showed that five of the six identified WSP sequences from *B. improvisus* proteins belong to the same clade as WSP homologues from *B. amphitrite*, suggesting that several WSPs were already present before the two species diverged.

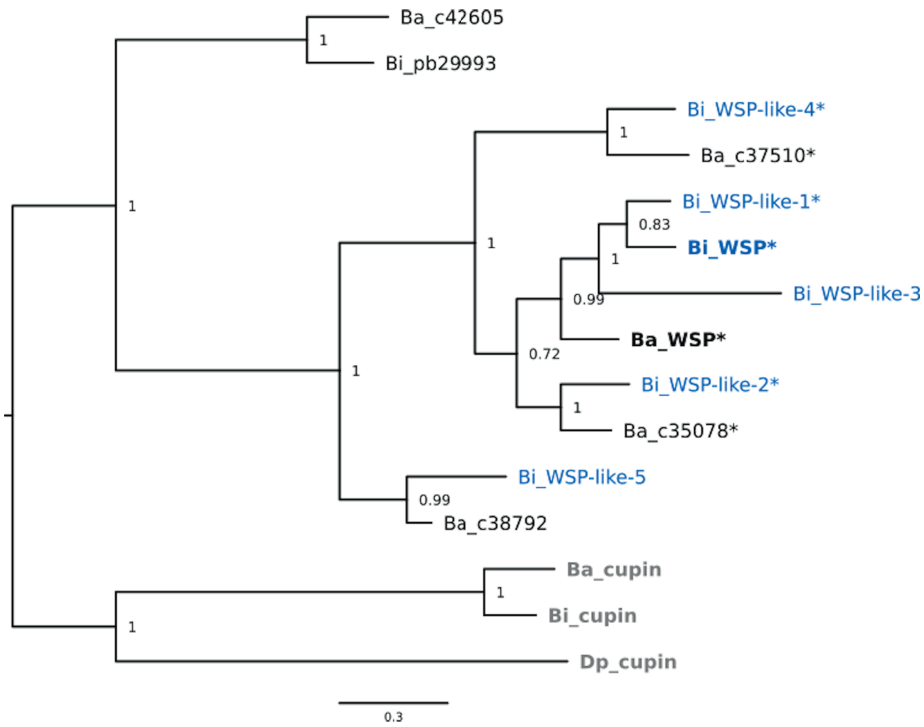


Figure 7. Bayesian phylogenetic tree of the WSP homologues in B. improvisus and B. amphitrite based on the alignment in Figure S1A. Cupin_5 domain proteins were added as an outgroup. All identified B. improvisus WSP homologues are marked in blue. Asterisks indicate WSP-like proteins with a long and charged C-terminus.

Despite that some of the sequences appear to form orthologous groups in the phylogenetic analysis, it is not presently possible to infer true orthologous relationships between the WSP-like candidates from *B. improvisus* and *B. amphitrite* due to the lack of complete genome-based proteomes for any of these barnacle species. The presence of several WSP-like candidates in barnacles is consistent with previously reported data on pheromones in some insects, nematodes and marine invertebrates that consist of a mix of several components. It has been suggested that the specificity of these pheromones is presumably determined by the relative concentrations of the different

components in the mixture (Linn et al. 1984, Roelofs 1995, Cummins et al. 2004). Moreover, evidence exists that the concentration of the different pheromone components depends on age, sex and habitat, thus providing additional information to the individuals receiving the pheromone signal (Coppée et al. 2011, Choe et al. 2012). One particularly well-studied example of pheromone mixtures among marine invertebrates is the pheromone blend in the sea slug *Aplysia* (Cummins et al. 2004). Mate attraction in these animals involves several waterborne pheromones, eg attractin, seductin and temptin, that in different combinations act to maintain mating aggregations (Cummins et al. 2004). Details on how the waterborne cue is released, distributed and sensed by cyprids are by large remain unknown. Studies on host location by cyprids of a parasitic barnacle *Heterosaccus dollfusi* (Pasternak et al. 2004) revealed that larvae use chemoreception to initiate the motion and rheoreception (sensing direction of the flow) to follow the host's odour plume upstream, until locating the organism releasing the odour (Pasternak et al. 2004). This study also suggested that larvae use not the concentration gradient of attractant to find the direction but rather its presence or absence in the flow. In the context of a pheromone mixture, it is therefore probable that cyprids first encounter the most abundant WSP that travels the longest distance from where it was released. Following the odour plume, cyprids will sense the other more lowly expressed WSP homologues that would serve as a confirmation of the direction as well as possibly trigger downward movement bringing cyprids closer to the surface where the cue is released by already established adults.

Further studies are needed to determine the function of all the WSP homologues, especially what features of WSPs are behind the settling-inducing effect, and whether there is a functional connection to the properties of the adhesives. Although we have shown the existence of several WSP-like sequences in both *B. improvisus* and *B. amphitrite*, a complete set of genes encoding WSP-like proteins in any barnacle species can only be obtained using a high-quality genome as reference. Despite that, the identification of a set of WSP homologues is an important discovery that will first of all promote further research into the great complexity of barnacle chemical communication. Furthermore, the protein sequences of identified WSP homologues are now available in the NCBI database and will aid in annotation of WSPs in future studies.

5 SUMMARY AND FUTURE PERSPECTIVES

The overall aim of this thesis was to gain a better understanding of the molecular mechanisms of the barnacle settlement process. Previous studies have substantially contributed to the knowledge on endocrine regulation (Yamamoto et al., 1997a; Yamamoto et al., 1997b), signaling pathways (Clare et. al., 1995; Clare 1996) and physiological mechanisms (Harder et al., 2001) regulating attachment and metamorphosis, and yet the process of barnacle settlement remains not fully understood.

The papers constituting this thesis investigated various aspects of barnacle settlement biology including settlement cues and candidate receptors involved in their perception, unfolding of the morphogenetic program and response to surface chemistry during settlement. The scheme on Figure 8 is an attempt to summarize the results and put them in the context of the entire settlement process. Settlement process begins with perception of a settlement cue. The diverse nature of the settlement cues suggests that larvae rely on at least three sensory systems: a mechanosensory, most likely represented by the numerous setae of the antennules; chemosensory, that includes lattice organs and antennular setae; and photosensory, represented by the pair of cyprid compound eyes (Matsumura and Qian 2014). The nature of the receptors remains illusive, however sensing of pheromones is thought to occur through G-protein coupled receptors and involve cAMP signaling pathway (Clare 1996a). Invertebrates also have an alternative, non-G-protein-based signaling pathway achieved through receptors that upon binding of ligands act as ion channels. We identified and characterised several novel classes of mechano- and chemoreceptors in the transcriptome from the cyprid antennules, i.e. ionotropic receptors and transient receptor potential channels. Therefore, presence of several different receptor classes and associated signaling pathways in cyprids further suggests that integration of signals from different settlement cues during barnacle settlement provides a basis for fine-tuned recognition of the suitable settlement spot (Morse 1993). As suggested by Kotsiri et al., (2018) and depicted on the Figure 8, the signals from settlement cues are decoded at the site of MF synthesis that in turn regulates the titre of ecdysone. When ecdysone reaches the target tissue it is converted into the active form of 20E that binds to a receptor complex EcR/RXR and initiate a vast transcriptional response leading to timely completion of settlement and metamorphosis. Here we identified and characterised several key genes from the ecdyson cascade to be regulated during the settlement, including both EcR and RXR, broad complex and nuclear hormone receptors HR3, HR38 and FTZ-F1.

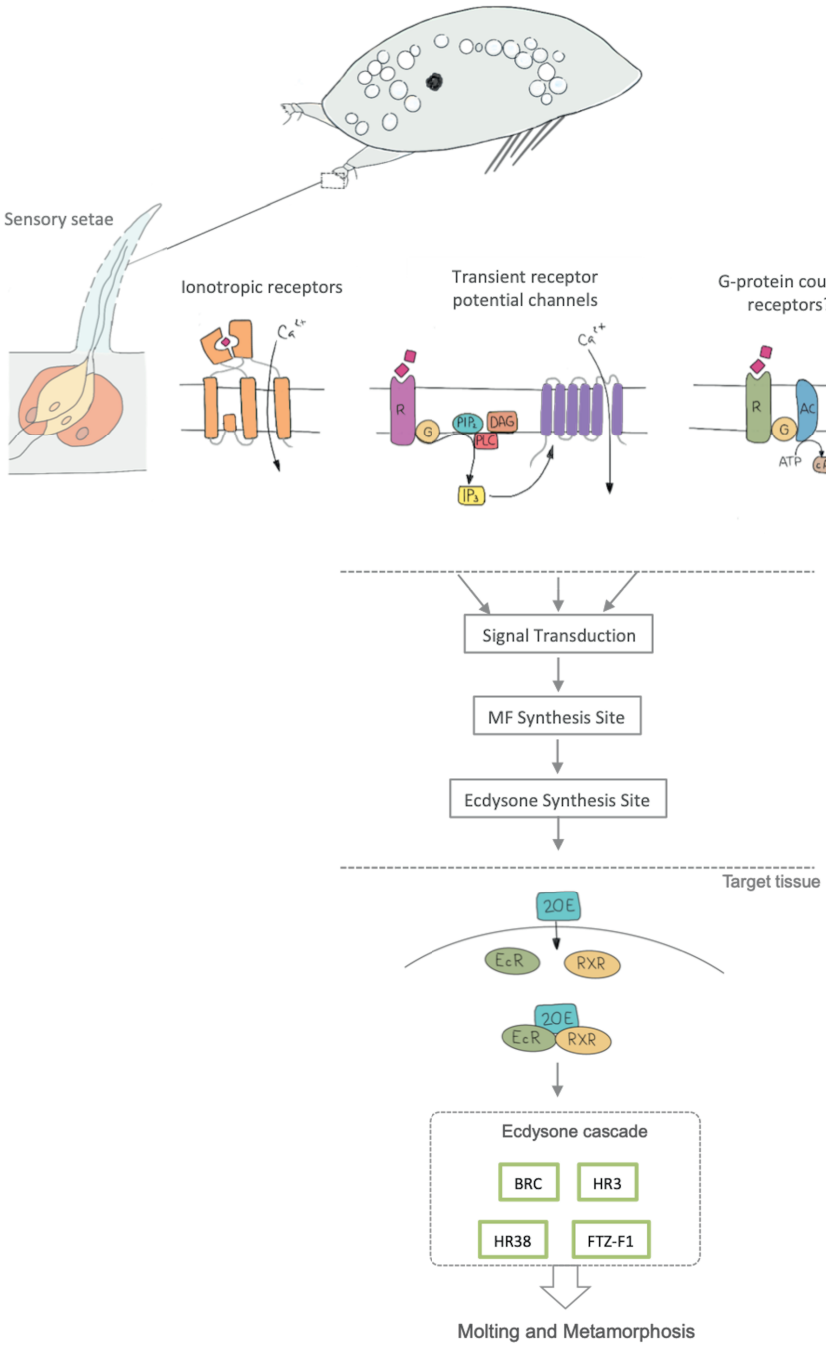

Sensory setae
Ionotropic receptors
Transient receptor potential channels
G-protein coupled receptors?
Ca2+
R
G
PIP2
DAG
PLC
IP3
AC
ATP
cAMP
Signal Transduction
MF Synthesis Site
Ecdysone Synthesis Site
Target tissue
20E
EcR
RXR
Ecdysone cascade
BRC
HR3
HR38
FTZ-F1
Molting and Metamorphosis

Figure 8. A working model of molecular mechanisms of settlement process: sensing of the settlement cues is mediated by several candidate receptor classes acting through different signalling pathways. The signals from settlement cues are decoded at the site of MF synthesis that in turn regulates the titre of ecdysone. When ecdysone reaches the target tissue it is converted into the active form of 20E that binds to a receptor complex EcR/RXR and initiate a vast transcriptional response leading to timely completion of settlement and metamorphosis. Ca calcium, R receptor, G G-protein, PIP_2 phosphatidylinositol bisphosphate, DAG diacylglycerol, PLC phospholipase C, IP3 inositol trisphosphate, AC adenylate cyclase, cAMP cyclic AMP, 20E 20-hydroxyecdysone, EcR ecdysone receptor, RXR retinoid R receptor, BRC broad complex, HR3 nuclear receptor 3, HR38 nuclear receptor 38, FTZ-F1 fushi tarazu transcription factor 1.

Overall, the aim of this project was to improve state-of-the-art understanding of barnacle settlement. The studies included in this thesis provide novel details on the molecular mechanisms involved in different aspects of the settlement process - from the chemical communication to genes involved in the regulation of metamorphosis. Thus, the aim has been to some extent achieved, though creating even more new questions to be answered. The identified candidates open up new opportunities for further functional research as well as potential development of new antifouling strategies.

As a follow up, more time-resolved studies have to be performed in order to elucidate the exact expression patterns of the ecdysone cascade genes in relation to each other to better define their roles in the control of barnacle settlement and metamorphosis. Identification of candidate receptors for settlement cues opens up endless possibilities for future studies. The first natural step would be to perform a validation of candidate receptors functions. RNA interference was successfully used to establish the role of IRs in host recognition in the sea louse *Lepeophtheirus salmonis* (Komisarczuk et al. 2017). Developing *in situ* probes for the identified receptors will allow identifying their expression location as well as investigate if they are also expressed in other sensory organs, such as frontal filament and lattice organs. *In vitro* cell-based assays can be used to test if known pheromones are binding to the identified receptor candidates. To this end, the identified herein *B. improvisus* WSP homologs should be heterologously expressed to test whether they have a settlement inducing activity. Further studies can use Y-maze or tracking in flow technique, as was successfully applied for the parasitic barnacle *Heterosaccus dollfusi* (Pasternak et al. 2004), to reveal if WSP homologs work as a pheromone mixture and if they trigger downward movement of cyprids.

6 ACKNOWLEDGEMENTS

This work would not be possible without the support and help of many people. To all of you, I would like to express my warm and sincere gratitude.

First of all, I would like to thank my supervisor, Anders Blomberg, for giving me the opportunity to work on this great and challenging project. You kindly guided me through the jungle of methods, techniques, administrative issues and read through endless versions of manuscripts. Your passion for science and music was very inspirational along this journey. Thank you for all your advice and support!

Ulrika, I am very thankful for sharing all the ups and downs of these five years, sitting back to back in our office. I admire your approach to solving any lab problems and making seemingly unrealistic projects to work. Thank you for all your help and advice! As well as for all the nice memories of singing in “Kom och Sjung”, performing barnacle play at the Vetenskaps festivalen and all the travels to Tjärnö.

I am extremely grateful to Magnus, none of this work would be possible without your expertise and guidance in the bioinformatics. Your constant aspiration for new discoveries was of a great inspiration to me. Thank you for all of your help, support and patience; for the ‘lunch therapy’ sessions and bubbel inventory, for keeping the spirit up and the last not least introducing me to the Swedish culture, I still have to try that semla with warm milk :)

I would like to extend my deepest gratitude to Tomas. This collaboration, brought up by your great idea of a shared project, taught me so much and largely formed my current scientific interests. It resulted in my first scientific publication and most importantly in a great friendship that I hope will last regardless of where life will bring us. I so much appreciate all of your support and help.

I would like to thank Martin Ogemark for all the efforts in keeping barnacle culture up and running. For being always so helpful with providing cyprids for my experiments, for all the conversations in Swedish we had and listening to Freak Kitchen!

I would like to thank my assistant supervisor Per Jonsson for giving invaluable input to my work and reviewing manuscripts, helping with designing experiments and advice with statistical analysis.

Special thanks to Hans Elwing, I so much appreciate your help and valuable advice. Thank you for the interesting discussions and sharing the ideas!
I would like to thank Mattias Berglin for help with wettability measurements, John Patrick Alao for the technical assistance with FACS, many thanks to Jonathan Havenhand for recommendations on statistical analysis, to Peter

Carlsson and Ali Moussavi Nik for advice on *in situ* hybridization, to Emil Karlsson for guidance with bioinformatics analysis, to Lars Nordvall for being the bästa IT tekniker. Thank you, Malte Hermansson, for taking time to discuss my PhD planning, for your help and support. Thanks should also go to the administration at CMB and Marine Departments for always being very helpful.

This research would not have been possible without the financial support from the EU SEAFRONT project. I was very fortunate to be a part of the Training Associate program within this project providing a great opportunity for multidisciplinary training, discussions and collaborations.

Sander, I am truly happy I had met you! Thank you for all these wonderful moments we shared during the project meetings, training weeks, ice-cold Portugal surfing, and even colder snorkeling in Sweden, learning some Dutch, Mallorca hiking, all the nice dinners and conversations. These years would not be as enjoyable without you!

One big “Thank You” to all of you Simon, Karl, Martin, Valeria, Sylvie, Vaskar, Josefine, Sveta, Peter for all the lunches, fikas, monthly tastings and just making everyday happier! Karl, special thanks for the inspiration for the cover and tasty coffee. Payam, thank you so much for the wonderful conversations, for cheering-up and a great advise.

James, I had to travel all the way to Mozambique to meet you, and I do not regret it. Thank you for all the sushi dates, for crazy jamming sessions and gigs!

Words cannot express how grateful I am to my mother. If not for all your love and constant support through out my life I would not be where I am now. I’m deeply indebted to you for everything you have done for me. Thank you for encouraging all of my pursuits and inspiring to follow my dreams. Deepest gratitude to my grandmother for all the kindness, unparalleled love and support.

Pål, I would like to thank barnacles. If it was not for their genuine love to your neptunkryssare and my love for them, we would probably have never met :) Words are not enough to express how much I do appreciate your support.

Thanks God for the lindy hop! Being a part of the dancing community changed my life in so many ways and let me meet many wonderful people. Vanessa, Alexander, Kalle, Jonathan, Mina, Ulf, Mokhtar.. I wish I could name all of you here. Without all the smiles we shared and your support this journey would be very hard.

Finally, thanks to all the cyprids that contributed to this project.

7 REFERENCES

Ache, B. W. and J. M. Young (2005). "Olfaction: diverse species, conserved principles." Neuron **48**(3): 417-430.

Aldred, N., A. Alsaab and A. S. Clare (2018). "Quantitative analysis of the complete larval settlement process confirms Crisp's model of surface selectivity by barnacles." Proceedings of the Royal Society B: Biological Sciences **285**(1872): 20171957.

Aldred, N. and A. S. Clare (2008). "The adhesive strategies of cyprids and development of barnacle-resistant marine coatings." Biofouling **24**(5): 351-363.

Aldred, N., A. Scardino, A. Cavaco, R. de Nys and A. S. Clare (2010). "Attachment strength is a key factor in the selection of surfaces by barnacle cyprids (Balanus amphitrite) during settlement." Biofouling **26**(3): 287-299.

Anderson, D. T. (1993). Barnacles: structure, function, development and evolution, Springer Science & Business Media.

Bachmann, K. and E. Rheinsmith (1973). "Nuclear DNA amounts in pacific Crustacea." Chromosoma **43**(3): 225-236.

Benton, R., K. S. Vannice, C. Gomez-Diaz and L. B. Vosshall (2009). "Variant Ionotropic Glutamate Receptors as Chemosensory Receptors in Drosophila." Cell **136**(1): 149-162.

Berglin, M. and P. Gatenholm (2003). "The barnacle adhesive plaque: morphological and chemical differences as a response to substrate properties." Colloids and Surfaces B: Biointerfaces **28**(2-3): 107-117.

Berntsson, K. M. and P. R. Jonsson (2003). "Temporal and spatial patterns in recruitment and succession of a temperate marine fouling assemblage: a comparison of static panels and boat hulls during the boating season." Biofueling **19**(3): 187-195.

Berntsson, K. M., P. R. Jonsson, M. Lejhall and P. Gatenholm (2000). "Analysis of behavioural rejection of micro-textured surfaces and implications for recruitment by the barnacle Balanus improvisus." Journal of Experimental Marine Biology and Ecology **251**(1): 59-83.

Bielecki, J., B. K. K. Chan, J. T. Hoeg and A. Sari (2009). "Antennular sensory organs in cyprids of balanomorphan cirripedes: standardizing terminology using Megabalanus rosa." Biofouling **25**(3): 203-214.

Blom, S.-E. and K.-G. Nyholm (1962). Settling times of Balanus balanoides (L.), Balanus crenatus Brug., and Balanus improvisus Darwin on the west coast of Sweden, Almqvist & Wiksell.

Blomberg, A., K. Sundell, P. R. Jonsson and A.-L. Wrange (2019). "Osmoregulation in barnacles: An evolutionary perspective of potential mechanisms and future research directions." Frontiers in Physiology **10**: 877.

Bobkov, Y. V. and B. W. Ache (2005). "Pharmacological properties and functional role of a TRP-related ion channel in lobster olfactory receptor neurons." Journal of neurophysiology **93**(3): 1372-1380.

Breithaupt, T. and M. Thiel (2010). Chemical communication in crustaceans, Springer Science & Business Media.

Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M. Frith, N. Maeda, R. Oyama, T. Ravasi, B. Lenhard and C. Wells (2005). "The transcriptional landscape of the mammalian genome." Science **309**(5740): 1559-1563.

Chaw, K. C. and W. R. Birch (2009). "Quantifying the exploratory behaviour of Amphibalanus amphitrite cyprids." Biofouling **25**(7): 611-619.

Chen, Z.-F., K. Matsumura, H. Wang, S. M. Arellano, X. Yan, I. Alam, J. A. Archer, V. B. Bajic and P.-Y. Qian (2011). "Toward an understanding of the molecular mechanisms of barnacle larval settlement: a comparative transcriptomic approach." PLoS One **6**(7): e22913.

Choe, A., T. Chuman, S. H. von Reuss, A. T. Dossey, J. J. Yim, R. Ajredini, A. A. Kolawa, F. Kaplan, H. T. Alborn, P. E. A. Teal, F. C. Schroeder, P. W. Sternberg and A. S. Edison (2012). "Sex-specific mating pheromones in the nematode Panagrellus redivivus." Proceedings of the National Academy of Sciences of the United States of America **109**(51): 20949-20954.

Clare, A. (1996a). "Signal transduction in barnacle settlement: calcium re-visited." Biofouling **10**(1-3): 141-159.

Clare, A., R. Thomas and D. Rittschof (1995). "Evidence for the involvement of cyclic AMP in the pheromonal modulation of barnacle settlement." Journal of Experimental Biology **198**(3): 655-664.

Clare, A. S. (1996b). "Marine natural product antifoulants: status and potential." Biofouling **9**(3): 211-229.

Clare, A. S., R. K. Freet and M. McClary (1994). "On the antennular secretion of the cyprid of Balanus amphitrite amphitrite, and its role as a settlement

pheromone." Journal of the Marine Biological Association of the United Kingdom **74**(1): 243-250.

Clare, A. S. and K. Matsumura (2000). "Nature and perception of barnacle settlement pheromones." Biofouling **15**(1-3): 57-71.

Clare, A. S. and M. Yamazaki (2000). "Inactivity of glycyl-glycyl-arginine and two putative (QSAR) peptide analogues of barnacle waterborne settlement pheromone." Journal of the Marine Biological Association of the United Kingdom **80**(5): 945-946.

Coppée, A., T. Mathy, M.-C. Cammaerts, F. J. Verheggen, M. Terzo, S. Iserbyt, I. Valterová and P. Rasmont (2011). "Age-dependent attractivity of males' sexual pheromones in Bombus terrestris (L.)[Hymenoptera, Apidae]." Chemoecology **21**(2): 75-82.

Corey, E. A., Y. Bobkov, K. Ukhanov and B. W. Ache (2013). "Ionotropic Crustacean Olfactory Receptors." Plos One **8**(4).

Crisp, D. (1974). "Factors influencing the settlement of marine invertebrate larvae." Chemoreception in marine organisms: 177-265.

Crisp, D. and H. Barnes (1954). "The orientation and distribution of barnacles at settlement with particular reference to surface contour." The Journal of Animal Ecology: 142-162.

Crisp, D. J. and P. Meadows (1962). "The chemical basis of gregariousness in cirripedes." Proceedings of the Royal Society of London. Series B. Biological Sciences **156**(965): 500-520.

Cummins, S. F., A. E. Nichols, A. Amare, A. B. Hummon, J. V. Sweedler and G. T. Nagle (2004). "Characterization of Aplysia enticin and temptin, two novel water-borne protein pheromones that act in concert with attractin to stimulate mate attraction." Journal of Biological Chemistry **279**(24): 25614-25622.

Dahlström, M., H. Jonsson, P. R. Jonsson and H. Elwing (2004). "Surface wettability as a determinant in the settlement of the barnacle Balanus improvisus (Darwin)." Journal of Experimental Marine Biology and Ecology **305**(2): 223-232.

Dahlström, M., L. G. Mårtensson, P. R. Jonsson, T. Arnebrant and H. Elwing (2000). "Surface active adrenoceptor compounds prevent the settlement of cyprid larvae of Balanus improvisus." Biofouling **16**(2-4): 191-203.

Dahms, H.-U., S. Dobretsov and P.-Y. Qian (2004). "The effect of bacterial and diatom biofilms on the settlement of the bryozoan Bugula neritina." Journal of Experimental Marine Biology and Ecology **313**(1): 191-209.

Degnan, S. M. and B. M. Degnan (2010). "The initiation of metamorphosis as an ancient polyphenic trait and its role in metazoan life-cycle evolution." Philosophical Transactions of the Royal Society B: Biological Sciences **365**(1540): 641-651.

Derby, C. D., M. T. Kozma, A. Senatore and M. Schmidt (2016). "Molecular Mechanisms of Reception and Perireception in Crustacean Chemoreception: A Comparative Review." Chemical Senses **41**(5): 381-398.

Di Fino, A., L. Petrone, N. Aldred, T. Ederth, B. Liedberg and A. Clare (2014). "Correlation between surface chemistry and settlement behaviour in barnacle cyprids (Balanus improvisus)." Biofouling **30**(2): 143-152.

Dong, C. H., A. M. Hu, Y. Ni, Y. X. Zuo and G. H. Li (2013). "Effects of midazolam, pentobarbital and ketamine on the mRNA expression of ion channels in a model organism Daphnia pulex." Bmc Anesthesiology **13**.

Dreanno, C., R. R. Kirby and A. S. Clare (2006a). "Smelly feet are not always a bad thing: the relationship between cyprid footprint protein and the barnacle settlement pheromone." Biology letters **2**(3): 423-425.

Dreanno, C., R. R. Kirby and A. S. Clare (2007). "Involvement of the barnacle settlement-inducing protein complex (SIPC) in species recognition at settlement." Journal of experimental marine biology and ecology **351**(1-2): 276-282.

Dreanno, C., K. Matsumura, N. Dohmae, K. Takio, H. Hirota, R. R. Kirby and A. S. Clare (2006b). "An alpha2-macroglobulin-like protein is the cue to gregarious settlement of the barnacle Balanus amphitrite." Proc Natl Acad Sci U S A **103**(39): 14396-14401.

Elbourne, P. D. and A. S. Clare (2010). "Ecological relevance of a conspecific, waterborne settlement cue in Balanus amphitrite (Cirripedia)." Journal of Experimental Marine Biology and Ecology **392**(1-2): 99-106.

Endo, N., Y. Nogata, E. Yoshimura and K. Matsumura (2009). "Purification and partial amino acid sequence analysis of the larval settlement-inducing pheromone from adult extracts of the barnacle, Balanus amphitrite (=Amphibalanus amphitrite)." Biofouling **25**(5): 429-434.

Ferrier, G. A., S. J. Kim, C. S. Kaddis, J. A. Loo, C. Ann Zimmer and R. K. Zimmer (2016). MULTIFUNCin: a multifunctional protein cue induces habitat selection by, and predation on, barnacles, Oxford University Press.

Finlay, J. A., S. M. Bennett, L. H. Brewer, A. Sokolova, G. Clay, N. Gunari, A. E. Meyer, G. C. Walker, D. E. Wendt and M. E. Callow (2010). "Barnacle settlement and the adhesion of protein and diatom microfouling to xerogel films with varying surface energy and water wettability." Biofouling **26**(6): 657-666.

Freeman, J. A. and J. D. Costlow (1983). "The cyprid molt cycle and its hormonal control in the barnacle Balanus amphitrite." Journal of Crustacean Biology **3**(2): 173-182.

Furman, E. R. and A. B. Yule (1990). "Self-fertilisation in Balanus improvisus Darwin." Journal of Experimental Marine Biology and Ecology **144**(2-3): 235-239.

Glenner, H. and M. B. Hebsgaard (2006). "Phylogeny and evolution of life history strategies of the parasitic barnacles (Crustacea, Cirripedia, Rhizocephala)." Molecular Phylogenetics and Evolution **41**(3): 528-538.

Gohad, N. V., N. Aldred, B. Orihuela, A. S. Clare, D. Rittschof and A. S. Mount (2012). "Observations on the settlement and cementation of barnacle (Balanus amphitrite) cyprid larvae after artificial exposure to noradrenaline and the locations of adrenergic-like receptors." Journal of Experimental Marine Biology and Ecology **416**: 153-161.

Gohad, N. V., N. M. Shah, A. T. Metters and A. S. Mount (2010). "Noradrenaline deters marine invertebrate biofouling when covalently bound in polymeric coatings." Journal of Experimental Marine Biology and Ecology **394**(1-2): 63-73.

Groh, K. C., H. Vogel, M. C. Stensmyr, E. Grosse-Wilde and B. S. Hansson (2014). "The hermit crab's nose-antennal transcriptomics." Frontiers in Neuroscience **7**.

Groh-Lunow, K. C., M. N. Getahun, E. Grosse-Wilde and B. S. Hansson (2015). "Expression of ionotropic receptors in terrestrial hermit crab's olfactory sensory neurons." Frontiers in Cellular Neuroscience **8**.

Harder, T., V. Thiyagarajan and P. Y. Qian (2001). "Combined effect of cyprid age and lipid content on larval attachment and metamorphosis of Balanus amphitrite Darwin."

Harrison, P. J. and D. C. Sandeman (1999). "Morphology of the nervous system of the barnacle cypris larva (Balanus amphitrite Darwin) revealed by light and electron microscopy." The Biological Bulletin **197**(2): 144-158.

Head, R. M., K. M. Berntsson, M. Dahlström, K. Overbeke and J. C. Thomason (2004). "Gregarious settlement in cypris larvae: the effects of cyprid age and assay duration." Biofouling **20**(2): 123-128.

Hoch, J. M. (2008). "Variation in penis morphology and mating ability in the acorn barnacle, Semibalanus balanoides." Journal of Experimental Marine Biology and Ecology **359**(2): 126-130.

Høeg, J., B. Hosfeld and P. G. Jensen (1998). "TEM studies on the lattice organs of cirripede cypris larvae (Crustacea, Thecostraca, Cirripedia)." Zoomorphology **118**(4): 195-205.

Høeg, J. T., M. Pérez-Losada, H. Glenner, G. Kolbasov and K. Crandall (2009). "Evolution of morphology, ontogeny and life cycles within the Crustacea Thecostraca." Arthropod Systematics and Phylogeny **67**(2): 199-217.

Hollins, B., D. Hardin, A. A. Gimelbrant and T. S. McClintock (2003). "Olfactory-enriched transcripts are cell-specific markers in the lobster olfactory organ." Journal of Comparative Neurology **455**(1): 125-138.

Holm, E. R. (2012). Barnacles and biofouling, Oxford University Press.

Holm, E. R., M. McClary Jr and D. Rittschof (2000). "Variation in attachment of the barnacle Balanus amphitrite: sensation or something else?" Marine Ecology Progress Series **202**: 153-162.

Hung, O., V. Thiyagarajan and P. Qian (2008). "Preferential attachment of barnacle larvae to natural multi-species biofilms: does surface wettability matter?" Journal of Experimental Marine Biology and Ecology **361**(1): 36-41.

Hyde, C. J., A. Elizur and T. Ventura (2019a). "The crustacean ecdysone cassette: a gatekeeper for molt and metamorphosis." The Journal of steroid biochemistry and molecular biology **185**: 172-183.

Hyde, C. J., Q. P. Fitzgibbon, A. Elizur, G. G. Smith and T. Ventura (2019b). "Transcriptional profiling of spiny lobster metamorphosis reveals three new additions to the nuclear receptor superfamily." BMC genomics **20**(1): 531.

Judge, M. L. and S. F. Craig (1997). "Positive flow dependence in the initial colonization of a fouling community: results from in situ water current

manipulations." Journal of Experimental Marine Biology and Ecology **210**(2): 209-222.

Kato-Yoshinaga, Y., M. Nagano, S. Mori, A. Clare, N. Fusetani and K. Matsumura (2000). "Species specificity of barnacle settlement-inducing proteins." Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology **125**(4): 511-516.

Kawahara, H., A. Isoai and Y. Shizuri (1997). "Molecular cloning of a putative serotonin receptor gene from barnacle, Balanus amphitrite." Gene **184**(2): 245-250.

Kerckhof, F. (2002). "Barnacles (Cirripedia, Balanomorpha) in Belgian waters, an overview of the species and recent evolutions, with emphasis on exotic species." Bulletin van het Koninklijk Belgisch Instituut voor Natuurwetenschappen. Biologie= Bulletin de l'Institut Royal des Sciences Naturelles de Belgique. Biologie(Suppl.).

Knight-Jones, E. and J. Stevenson (1950). "Gregariousness during settlement in the barnacle Elminius modestus Darwin." Journal of the Marine Biological Association of the United Kingdom **29**(2): 281-297.

Komisarczuk, A. Z., S. Grotmol and F. Nilsen (2017). "Ionotropic receptors signal host recognition in the salmon louse (Lepeophtheirus salmonis, Copepoda)." Plos One **12**(6).

Kotsiri, M., M. Protopapa, S. Mouratidis, M. Zachariadis, D. Vassilakos, I. Kleidas, M. Samiotaki and S. G. Dedos (2018a). "Should I stay or should I go? The settlement-inducing protein complex guides barnacle settlement decisions." Journal of Experimental Biology **221**(22): jeb185348.

Kotsiri, M., M. Protopapa, G. M. Roumelioti, A. Economou-Amilli, E. Efthimiadou and S. Dedos (2018b). "Probing the settlement signals of Amphibalanus amphitrite (vol 34, pg 492, 2018)." Biofouling **34**(5): X-X.

Lagersson, N. C. and J. T. Hoeg (2002). "Settlement behavior and antennulary biomechanics in cypris larvae of Balanus amphitrite (Crustacea : Thecostraca : Cirripedia)." Marine Biology **141**(3): 513-526.

Lam, C., T. Harder and P.-Y. Qian (2003). "Induction of larval settlement in the polychaete Hydroides elegans by surface-associated settlement cues of marine benthic diatoms." Marine Ecology Progress Series **263**: 83-92.

Lang, W. H., R. B. Forward Jr and D. C. Miller (1979). "Behavioral responses of Balanus improvisus nauplii to light intensity and spectrum." The Biological Bulletin **157**(1): 166-181.

Le Tourneux, F. and E. Bourget (1988). "Importance of physical and biological settlement cues used at different spatial scales by the larvae of Semibalanus balanoides." Marine Biology **97**(1): 57-66.

Leffler, E. M., K. Bullaughey, D. R. Matute, W. K. Meyer, L. Segurel, A. Venkat, P. Andolfatto and M. Przeworski (2012). "Revisiting an old riddle: what determines genetic diversity levels within species?" PLoS biology **10**(9): e1001388.

Lin, H. C., Y. H. Wong, L. M. Tsang, K. H. Chu, P. Y. Qian and B. K. K. Chan (2014). "First study on gene expression of cement proteins and potential adhesion-related genes of a membranous-based barnacle as revealed from Next-Generation Sequencing technology." Biofouling **30**(2): 169-181.

Lind, U., M. Alm Rosenblad, A. L. Wrange, K. S. Sundell, P. R. Jonsson, C. Andre, J. Havenhand and A. Blomberg (2013). "Molecular characterization of the alpha-subunit of Na(+)/K(+) ATPase from the euryhaline barnacle Balanus improvisus reveals multiple genes and differential expression of alternative splice variants." PLoS One **8**(10): e77069.

Lind, U., M. Jarva, M. A. Rosenblad, P. Pingitore, E. Karlsson, A. L. Wrange, E. Kamdal, K. Sundell, C. Andre, P. R. Jonsson, J. Havenhand, L. A. Eriksson, K. Hedfalk and A. Blomberg (2017). "Analysis of aquaporins from the euryhaline barnacle Balanus improvisus reveals differential expression in response to changes in salinity." Plos One **12**(7).

Lind, U., M. A. Rosenblad, L. H. Frank, S. Falkbring, L. Brive, J. M. Laurila, K. Pohjanoksa, A. Vuorenpää, J. P. Kukkonen and L. Gunnarsson (2010). "Octopamine receptors from the barnacle Balanus improvisus are activated by the α2-adrenoceptor agonist medetomidine." Molecular pharmacology **78**(2): 237-248.

Linn, C., L. Bjostad, J. Du and W. Roelofs (1984). "Redundancy in a chemical signal: Behavioral responses of maleTrichoplusia ni to a 6-component sex pheromone blend." Journal of chemical ecology **10**(11): 1635-1658.

Liu, Y., J. Zhou and K. P. White (2013). "RNA-seq differential expression studies: more sequence or more replication?" Bioinformatics **30**(3): 301-304.

Maleschlijski, S., S. Bauer, N. Aldred, A. S. Clare and A. Rosenhahn (2015). "Classification of the pre-settlement behaviour of barnacle cyprids." Journal of The Royal Society Interface **12**(102): 20141104.

Maleschlijski, S., G. Sendra, A. Di Fino, L. Leal-Taixé, I. Thome, A. Terfort, N. Aldred, M. Grunze, A. Clare and B. Rosenhahn (2012). "Three dimensional tracking of exploratory behavior of barnacle cyprids using stereoscopy." Biointerphases **7**(1): 50.

Maréchal, J.-P., K. Matsumura, S. Conlan and C. Hellio (2012). "Competence and discrimination during cyprid settlement in Amphibalanus amphitrite." International biodeterioration & biodegradation **72**: 59-66.

Maruzzo, D., N. Aldred, A. S. Clare and J. T. Høeg (2012). "Metamorphosis in the cirripede crustacean Balanus amphitrite." PloS one **7**(5): e37408.

Maruzzo, D., S. Conlan, N. Aldred, A. S. Clare and J. T. Hoeg (2011). "Video observation of surface exploration in cyprids of Balanus amphitrite: the movements of antennular sensory setae." Biofouling **27**(2): 225-239.

Matsumura, K., M. Nagano and N. Fusetani (1998). "Purification of a larval settlement-inducing protein complex (SIPC) of the barnacle, Balanus amphitrite." Journal of Experimental Zoology **281**(1): 12-20.

Matsumura, K. and P.-Y. Qian (2014). "Larval vision contributes to gregarious settlement in barnacles: adult red fluorescence as a possible visual signal." Journal of Experimental Biology **217**(5): 743-750.

Morse, D. E. (1990). "Recent progress in larval settlement and metamorphosis: closing the gaps between molecular biology and ecology." Bulletin of marine Science **46**(2): 465-483.

Morse, D. E. (1993). "Signalling in planktonic larvae." Nature **363**(6428): 406.

Olivier, F., R. Tremblay, E. Bourget and D. Rittschof (2000). "Barnacle settlement: field experiments on the influence of larval supply, tidal level, biofilm quality and age on Balanus amphitrite cyprids." Marine Ecology Progress Series **199**: 185-204.

Painter, S. D., B. Clough, R. W. Garden, J. V. Sweedler and G. T. Nagle (1998). "Characterization of Aplysia attractin, the first water-borne peptide pheromone in invertebrates." The Biological Bulletin **194**(2): 120-131.

Panova, M., H. Aronsson, R. A. Cameron, P. Dahl, A. Godhe, U. Lind, O. Ortega-Martinez, R. Pereyra, S. V. Tesson and A.-L. Wrange (2016). DNA

extraction protocols for whole-genome sequencing in marine organisms. Marine genomics, Springer**:** 13-44.

Pasternak, Z., B. Blasius and A. Abelson (2004). "Host location by larvae of a parasitic barnacle: larval chemotaxis and plume tracking in flow." Journal of Plankton Research **26**(4): 487-493.

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**.

Peng, G. D., X. Shi and T. Kadowaki (2015). "Evolution of TRP channels inferred by their classification in diverse animal species." Molecular Phylogenetics and Evolution **84**: 145-157.

Pérez-Losada, M., M. Harp, J. T. Høeg, Y. Achituv, D. Jones, H. Watanabe and K. A. Crandall (2008). "The tempo and mode of barnacle evolution." Molecular phylogenetics and evolution **46**(1): 328-346.

Petrone, L., A. Di Fino, N. Aldred, P. Sukkaew, T. Ederth, A. S. Clare and B. Liedberg (2011). "Effects of surface charge and Gibbs surface energy on the settlement behaviour of barnacle cyprids (Balanus amphitrite)." Biofouling **27**(9): 1043-1055.

Qian, P.-Y., V. Thiyagarajan, S. C. K. Lau and S. C. K. Cheung (2003). "Relationship between bacterial community profile in biofilm and attachment of the acorn barnacle Balanus amphitrite." Aquatic microbial ecology **33**(3): 225-237.

Qu, Z., N. J. Kenny, H. M. Lam, T. F. Chan, K. H. Chu, W. G. Bendena, S. S. Tobe and J. H. L. Hui (2015). "How did arthropod sesquiterpenoids and ecdysteroids arise? Comparison of hormonal pathway genes in noninsect arthropod genomes." Genome biology and evolution **7**(7): 1951-1959.

Rheinsmith, E., R. Hinegardner and K. Bachmann (1974). "Nuclear DNA amounts in Crustacea." Comparative Biochemistry and Physiology Part B: Comparative Biochemistry **48**(3): 343-348.

Rittschof, D. (1985). "Oyster drills and the frontiers of chemical ecology: unsettling ideas." Bull Am Mar Union **1**: 111-116.

Rittschof, D., A. Clare, D. Gerhart, S. A. Mary and J. Bonaventura (1992). "Barnacle in vitro assays for biologically active substances: toxicity and settlement inhibition assays using mass cultured Balanus amphitrite amphitrite Darwin." Biofouling **6**(2): 115-122.

Rizzi, F., A. Qualtieri, T. Dattoma, G. Epifani and M. De Vittorio (2015). "Biomimetics of underwater hair cell sensing." Microelectronic Engineering **132**: 90-97.

Rodriguez, S. R., F. P. Ojeda and N. C. Inestrosa (1993). "Settlement of benthic marine invertebrates." Marine ecology progress series. Oldendorf **97**(2): 193-207.

Roelofs, W. L. (1995). "Chemistry of Sex Attraction." Proceedings of the National Academy of Sciences of the United States of America **92**(1): 44-49.

Schuler, A., G. Schmitz, A. Reft, S. Ozbek, U. Thurm and E. Bornberg-Bauer (2015). "The Rise and Fall of TRP-N, an Ancient Family of Mechanogated Ion Channels, in Metazoa." Genome Biology and Evolution **7**(6): 1713-1727.

Silbering, A. F., R. Rytz, Y. Grosjean, L. Abuin, P. Ramdya, G. S. X. E. Jefferis and R. Benton (2011). "Complementary Function and Integrated Wiring of the Evolutionarily Distinct Drosophila Olfactory Subsystems." Journal of Neuroscience **31**(38): 13357-13375.

Sims, D., I. Sudbery, N. E. Ilott, A. Heger and C. P. Ponting (2014). "Sequencing depth and coverage: key considerations in genomic analyses." Nature Reviews Genetics **15**(2): 121.

Smith, P. A., A. S. Clare, H. H. Rees, M. C. Prescott, G. Wainwright and M. C. Thorndyke (2000). "Identification of methyl farnesoate in the cypris larva of the barnacle, Balanus amphitrite, and its role as a juvenile hormone." Insect biochemistry and molecular biology **30**(8-9): 885-890.

So, C. R., K. P. Fears, D. H. Leary, J. M. Scancella, Z. Wang, J. L. Liu, B. Orihuela, D. Rittschof, C. M. Spillmann and K. J. Wahl (2016). "Sequence basis of barnacle cement nanostructure is defined by proteins with silk homology." Scientific reports **6**: 36219.

Susana, L., F. Campaci and A. C. Santomaso (2012). "Wettability of mineral and metallic powders: applicability and limitations of sessile drop method and Washburn's technique." Powder technology **226**: 68-77.

Szallasi, A., D. N. Cortright, C. A. Blum and S. R. Eid (2007). "The vanilloid receptor TRPV1: 10 years from channel cloning to antagonist proof-of-concept." Nature reviews Drug discovery **6**(5): 357.

Tegtmeyer, K. and D. Rittschof (1988). "Synthetic Peptide Analogs to Barnacle Settlement Pheromone." Peptides **9**(6): 1403-1406.

Toyota, K., H. Miyakawa, C. Hiruta, K. Furuta, Y. Ogino, T. Shinoda, N. Tatarazako, S. Miyagawa, J. R. Shaw and T. Iguchi (2015). "Methyl farnesoate synthesis is necessary for the environmental sex determination in the water flea Daphnia pulex." Journal of insect physiology **80**: 22-30.

Tran, C. and M. G. Hadfield (2012). "Are G-protein-coupled receptors involved in mediating larval settlement and metamorphosis of coral planulae?" The Biological Bulletin **222**(2): 128-136.

Ventura, T., F. Palero, G. Rotllant and Q. P. Fitzgibbon (2018). "Crustacean metamorphosis: an omics perspective." Hydrobiologia **825**(1): 47-60.

Walker, G. (1974). "The fine structure of the frontal filament complex of barnacle larvae (Crustacea: Cirripedia)." Cell and tissue research **152**(4): 449-465.

Wang, C., J. N. Schultzhaus, C. R. Taitt, D. H. Leary, L. C. Shriver-Lake, D. Snellings, S. Sturiale, S. H. North, B. Orihuela and D. Rittschof (2018). "Characterization of longitudinal canal tissue in the acorn barnacle Amphibalanus amphitrite." PloS one **13**(12): e0208352.

Wang, Z., D. H. Leary, J. Liu, R. E. Settlage, K. P. Fears, S. H. North, A. Mostaghim, T. Essock-Burns, S. E. Haynes and K. J. Wahl (2015). "Molt-dependent transcriptomic analysis of cement proteins in the barnacle Amphibalanus amphitrite." BMC genomics **16**(1): 859.

Wiegemann, M. (2005). "Adhesion in blue mussels (Mytilus edulis) and barnacles (genus Balanus): mechanisms and technical applications." Aquatic Sciences-Research Across Boundaries **67**(2): 166-176.

Wiegemann, M. and B. Watermann (2003). "Peculiarities of barnacle adhesive cured on non-stick surfaces." Journal of Adhesion Science and Technology **17**(14): 1957-1977.

Williams, E. A. and S. M. Degnan (2009). "Carry-over effect of larval settlement cue on postlarval gene expression in the marine gastropod Haliotis asinina." Molecular ecology **18**(21): 4434-4449.

Winsor, M. P. (1969). "Barnacle Larvae in the Nineteenth CenturyA Case Study in Taxonomic Theory." Journal of the History of Medicine and Allied Sciences **24**(3): 294-309.

Wrange, A.-L., C. André, T. Lundh, U. Lind, A. Blomberg, P. J. Jonsson and J. N. Havenhand (2014). "Importance of plasticity and local adaptation for

coping with changing salinity in coastal areas: a test case with barnacles in the Baltic Sea." BMC evolutionary biology **14**(1): 156.

Wrange, A.-L., G. Charrier, A. Thonig, M. A. Rosenblad, A. Blomberg, J. N. Havenhand, P. R. Jonsson and C. André (2016). "The story of a hitchhiker: population genetic patterns in the invasive barnacle Balanus (Amphibalanus) improvisus Darwin 1854." PloS one **11**(1): e0147082.

Yamamoto, H., S. Kawaii, E. Yoshimura, A. Tachibana and N. Fusetani (1997a). "20-Hydroxyecdysone regulates larval metamorphosis of the barnacle, Balanus amphitrite." Zoological science **14**(6): 887-893.

Yamamoto, H., T. Okino, E. Yoshimura, A. Tachibana, K. Shimizu and N. Fusetani (1997b). "Methyl farnesoate induces larval metamorphosis of the barnacle, Balanus amphitrite via protein kinase C activation." Journal of Experimental Zoology **278**(6): 349-355.

Yamamoto, H., A. Tachibana, K. Matsumura and N. Fusetani (1995). "Protein kinase C (PKC) signal transduction system involved in larval metamorphosis of the barnacle, Balanus amphitrite." Zoological science **12**(4): 391-397.

Yan, X.-C., Z.-F. Chen, J. Sun, K. Matsumura, R. S. Wu and P.-Y. Qian (2012). "Transcriptomic analysis of neuropeptides and peptide hormones in the barnacle Balanus amphitrite: evidence of roles in larval settlement." PLoS One **7**(10): e46513.

Yorisue, T., K. Matsumura, H. Hirota, N. Dohmae and S. Kojima (2012). "Possible molecular mechanisms of species recognition by barnacle larvae inferred from multi-specific sequencing analysis of proteinaceous settlement-inducing pheromone." Biofouling **28**(6): 605-611.

Zhang, Y., L.-S. He, G. Zhang, Y. Xu, O.-O. Lee, K. Matsumura and P.-Y. Qian (2012). "The regulatory role of the NO/cGMP signal transduction cascade during larval attachment and metamorphosis of the barnacle Balanus (= Amphibalanus) amphitrite." Journal of Experimental Biology **215**(21): 3813-3822.

Zimmer, R. K., G. A. Ferrier, S. J. Kim, C. S. Kaddis, C. A. Zimmer and J. A. Loo (2016). "A multifunctional chemical cue drives opposing demographic processes and structures ecological communities." Ecology **97**(9): 2232-2239.

I

PLOS ONE

RESEARCH ARTICLE

Sensory receptor repertoire in cyprid antennules of the barnacle *Balanus improvisus*

Anna Abramova[1], **Magnus Alm Rosenblad**[1], **Anders Blomberg**[1]*, **Tomas Axel Larsson**[2]

1 Department of Chemistry and Molecular Biology, University of Gothenburg, Gothenburg, Sweden,
2 Department of Marine Sciences, University of Gothenburg, Gothenburg, Sweden

* anders.blomberg@cmb.gu.se

Citation: Abramova A, Alm Rosenblad M, Blomberg A, Larsson TA (2019) Sensory receptor repertoire in cyprid antennules of the barnacle *Balanus improvisus*. PLoS ONE 14(5): e0216294. https://doi.org/10.1371/journal.pone.0216294

Editor: Johannes Reisert, Monell Chemical Senses Center, UNITED STATES

Received: November 16, 2018

Accepted: April 17, 2019

Published: May 2, 2019

Data Availability Statement: All the sequences described in the manuscript are available from the GenBank database (accession numbers MK093193 – MK093208). Raw sequence reads and transcriptome assemblies were uploaded to the SRA and TSA repositories (BioProject ID PRJNA528777 and PRJNA528169).

Funding: AA and TAL were funded by a grant from Centre for Marine Evolutionary Biology at the University of Gothenburg, and AB received funding from Swedish research council (#2017-04559) and the EU Commission (SEAFRONT— Synergistic

Abstract

Barnacle settlement involves sensing of a variety of exogenous cues. A pair of antennules is the main sensory organ that the cyprid larva uses to explore the surface. Antennules are equipped with a number of setae that have both chemo- and mechanosensing function. The current study explores the repertoire of sensory receptors in *Balanus improvisus* cyprid antennules with the goal to better understand sensory systems involved in the settling behavior of this species. We carried out transcriptome sequencing of dissected *B. improvisus* cyprid antennules. The generated transcriptome assembly was used to search for sensory receptors using HMM models. Among potential chemosensory genes, we identified the ionotropic receptors IR25a, IR8a and IR93a, and several divergent IR candidates to be expressed in the cyprid antennules. We found one gustatory-like receptor but no odorant receptors, chemosensory or odorant-binding proteins. Apart from chemosensory receptors, we also identified 13 potential mechanosensory genes represented by several transient receptor potential channels (TRP) subfamilies. Furthermore, we analyzed changes in expression profiles of IRs and TRPs during the *B. improvisus* settling process. Several of the sensory genes were differentially expressed during the course of larval settlement. This study gives expanded knowledge about the sensory systems present in barnacles, a taxonomic group for which only limited information about receptors is currently available. It furthermore serves as a starting point for more in depth studies of how sensory signaling affects settling behavior in barnacles with implications for preventing biofouling.

Introduction

The barnacle *Balanus improvisus* Darwin 1854 *(= Amphibalanus improvisus)* is a common fouling species in temperate waters that, together with *Balanus amphitrite (= Amphibalanus amphitrite)*, has become a model organism for developing and testing antifouling coatings [1]. The barnacle life cycle involves sessile adults and planktonic larvae, including six naupliar larval stages followed by a cyprid larva that settles and metamorphoses into the sessile stage. Settling cyprids show a complex searching behavior that includes several distinct stages; e.g. free

Fouling Control Technologies; #614034). 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.

swimming, wide search, close search and inspection [2,3]. Surface exploration involves sensing of a variety of physical and chemical exogenous cues [4,5].

Adult barnacles have quite simplified sensory organs due to their sessile lifestyle. Apart from light-sensitive ocelli, the main sensory organs of adults are the thoracopods, called cirri, also used for suspension feeding. The cirri have distinct functionality depending on the various types of mechano-, thermo-, hygro- and chemosensory setae they carry [6]. In contrast to adults, cyprids that actively swim and explore the environment are equipped with more complex sensory organs such as a pair of compound eyes, frontal filaments, lattice organs, as well as the pair of antennules. It is believed that exploration of the substratum predominantly relies on the antennules. Cyprid antennules posses an attachment disc densely covered with cuticular villi and a number of setae that perform chemosensory and mechanosensory functions [7,8]. Some setae have a terminal pore and are suggested to perform contact chemoreception, while others are sac-shaped and thin-walled, called aesthetascs, that potentially can sense waterborne compounds [7,9]. It has been suggested that the majority of chemosensory setae, except the aesthetascs, are bimodal and also have a mechanoreceptive function [8]. However, the nature of the actual receptors involved in sensing settlement cues has remained elusive. Identification of barnacle sensory receptors would be a considerable asset and a step forward, both for studying external factors affecting the settlement behavior as well as for the development of new antifouling technologies.

Despite crustaceans being used as models to study olfaction, the repertoire of the chemosensory receptors remains largely unknown. In the majority of invertebrates studied in detail so far, chemoreception is most often mediated by three different gene families. These include two families of seven transmembrane receptors, namely the odorant receptors (ORs) and gustatory receptors (GRs), and a third family of ligand-gated ion channels, called the ionotropic receptors (IRs) [10]. Previous searches in arthropods revealed that ORs are only present in Hexapoda and seems to be lacking in all other groups [11–14]. GRs were suggested to be the most ancient family of chemosensory genes in arthropods, however currently only limited data are available for crustaceans with ten GRs found in *Eurytemora affinis* (Copepoda) [12], 58 in the *Daphnia* genome [14] and one candidate gene reported from the *B. amphitrite* transcriptome [12]. Initally discovered in *Drosophila* [15], IRs have thereafter been found in several Protostomes, suggesting that they are an ancestral chemosensory receptor family originating from the larger family of ionotropic glutamate receptors (iGluRs) [16]. IRs have a similar domain organisation to the iGluRs but are lacking one or more residues known to bind glutamate, suggesting that they have evolved to recognise other types of molecules. IRs act in complexes containing up to three subunits, including one or two common co-receptors, i.e. IR25a and IR8a, and individual odor-specific IRs [17]. Insect IRs are traditionally divided into two groups: "antennal" IRs, that were originally discovered in *Drosophila* antennal olfactory neurons, and species-specific "divergent" IRs, mainly expressed in the gustatory organs and involved in the detection of taste [16]. Importantly, earlier studies of antennal transcriptomes suggest that IRs are the only chemoreceptors found in crustacean antennas [10,11,18], therefore making IRs the potential candidate for detection of chemical cues during cyprid settlement.

Apart from chemical cues, it has been suggested that cyprids can distinguish surface structures by means of mechanosensory setae [8,9]. According to current knowledge, mechanoreception in arthropods is accomplished through mechanosensitive ion channels (transporting Ca^{2+} or Mg^{2+}) that are gated upon mechanical stress and allow the influx of calcium ions resulting in a receptor potential [19]. Transient receptor potential (TRP) channels play major roles in various sensory modalities such as hearing, hygrosensation, vision and mechanosensation in a diverse set of animals [20]. Based on the amino acid sequences and the presence of

specific domains, TRP channels have been divided into seven subfamilies—TRPA, TRPC, TRPM, TRPML, TRPN, TRPP and TRPV [20]. Among arthropods, TRPs have been mainly studied in insects and several of them were functionally characterised, e.g. *Drosophila* TRPN and TRPA have been shown to be involved in larval crawling behavior and thermosensation [20,21]. Several types of TRPs have been identified in *D. pulex*, however, their exact function in this species is unknown [20]. Based on behavioral studies and the use of chemical activators/inhibitors, it has been suggested that cyprid surface exploration rely on mechanosensitive Ca^{2+} channels, in particular involving the *D. melanogaster* homologs of *painless* and TRPA1 [22].

No chemo- or mechanosensory receptors have been characterized in barnacles, except one partial GR-like gene found in *B. amphitrite* [12]. There are several studies showing involvement of serotonin and octopamine receptors in settlement and metamorphosis, however, it is not certain if these receptors are involved in the sensing of external settling cues [23,24]. Enrichment for chemosensory tissue followed by transcriptome sequencing proved to be a successful strategy to identify sensory receptor candidates [18,25,26]. We therefore generated transcriptome data from cyprid antennules in order to enrich for sensory genes (Fig 1). We made curated HMM models for sensory genes and performed bioinformatics searches in *B. improvisus* antennular dataset. Furthermore, we characterized the expression patterns of these sensory genes at four time points during the settlement process. As a result, we identified several differentially expressed genes encoding potential chemo- and mechanoreceptors. This study forms a basis for a better understanding of the barnacle sensing during settling and provides a foundation for development of more advanced antifouling technologies.

Materials and methods

Dissection of antennules and RNA extraction

Cyprid larvae were reared in an all-year-round laboratory culture of *B. improvisus* as described in Jonsson et al., 2018 [27]. Free swimming cyprid larvae, approximately 2–3 days old, were collected, placed in RNAlater (Ambion), and stored at 4˚C for two days. Subsequently, individual larvae were placed on a glass slide and the pair of antennules was separated at the base of the proximal segment (see Fig 1) by means of stainless steel insect pins (Fine Science Tools,

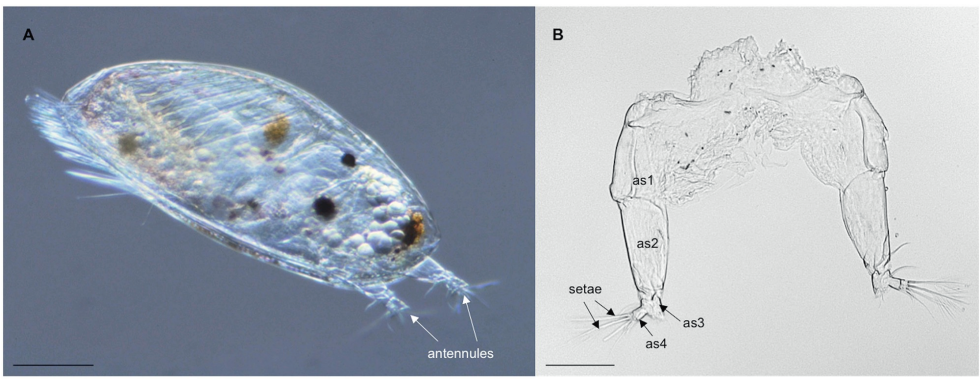


Fig 1. *B. improvisus* cyprid larva and a pair of antennules. (A) Lateral view of a cyprid with the pair of antennules indicated, scale bar = 100 μm. (B) a dissected pair of antennules with sensory setae and antennular segments (as1–4) indicated, scale bar = 50 μm.

https://doi.org/10.1371/journal.pone.0216294.g001

size 000, #26000–25) under a stereomicroscope (Bresser Researcher, WF 10X/20, 4X). Each pair of antennules was then transferred into the RLT buffer supplied with the RNeasy micro kit (Qiagen, GmbH, Hilden, Germany, #74004). In total, three independent samples were collected each containing antennular pairs from 50 individual cyprids. Samples were homogenized with Precellys Lysing Kit for hard tissue (CK28, Bertin Technologies, Montigny-le-Bretonneux, France, KT03961-1-002.2) and RNA extracted following the RNeasy micro kit (Qiagen, GmbH, Hilden, Germany, #74004) standard procedure.

RNA-library preparation protocols and sequencing

The three independent RNA samples were pooled together into one sample to obtain enough RNA material for RNA library preparation (we obtained in total ≈ 1.2 μg RNA) and sent to the Science for Life Laboratory in Stockholm for sequencing. The cDNA Library was prepared according to the Illumina TruSeq Stranded mRNA Library Prep Kit (RS-122-2103) protocol with poly-A selection, and then sequenced on one lane of Illumina HiSeq 2500 to generate 260 million read pairs (126bp paired-end reads; in total roughly 89Gbp). The raw sequence data was uploaded to sequence read archive (SRA) with accession number SRR8755479.

QC, assembly and Gene Ontology classification of sequencing data

The obtained antennular sequence data was initially quality checked using the software FastQC 0.8.0, Trim Galore! (http://www.bioinformatics.babraham.ac.uk/) and Cutadapt [28]. De novo transcriptome assembly was done using Trinity 2.0.6 [29] (with default settings) resulting in 239,215 contigs. The assembly was further clustered with cd-hit (v4.6.5) [30] with similarity threshold 0.95 to remove redundant transcripts resulting in 214,110 Trinity transcripts. The assembly was uploaded to the transcriptome shotgun assembly (TSA) database with BioProject ID PRJNA528169. We evaluated the completeness and duplication level of the assembly with arthropoda BUSCO v.2 [31], comprising 113 arthropod species and 1066 single copy orthologs. Transdecoder (v3.0.1) [32] was used to predict the coding regions in the transcripts, resulting in 15,931 complete ORFs and 35,284 internal (missing N- and C-terminal), while the rest were classified as 5' and 3' partials. GOSeq [33] inside the Trinity wrapper was applied to analyse gene ontology (GO) terms representation to score overrepresentation/underrepresentation of molecular function, biological process and cellular components. The GO terms were further passed to the web-based CateGOrizer (https://www.animalgenome.org/tools/catego/) software that batch analyses GO annotations by mapping them to the generic GO slim subset.

Construction of HMMs

Searching with Hidden Markov Models (HMMs) is a more sensitive approach for identification of distant homologs than ordinary BLAST searches. We made separate HMMs for eight classes of the sensory receptors to be analysed, including both chemo- and mechanoreceptors. In particular, individual models were made for the different subtypes of ionotropic receptors (IR7a, IR8a, IR25a and IR93a) as they are the main candidates for olfactory receptors in crustaceans [11]. To search for classical chemosensory insect receptors we also made models for gustatory and odorant receptors, as well as for odorant-binding proteins (OBP), while for chemosensory proteins we used an HMM model downloaded from the PFAM database (PF03392).

HMM models were initially based on crustacean sequences previously reported as sensory receptors. Currently there are a limited number of sensory genes reliably annotated and functionally characterised in crustaceans. Because of this lack of information, we also included

insect sensory receptors, and these are overrepresented in our final HMM models (see S1 Table for details). Alignments of sequences were performed with Mafft [34] and visualised with Jalview 2.10.1 [35]. Manually curated alignments for each receptor type were used to construct HMMs with HMMER v3.1b2 [36] to be used for searches against the predicted ORFs from the barnacle transcriptomes.

Sensory gene candidates identification and the filtering process

For the HMM searches we used the antennular transcriptome dataset generated in the current study, as well as RNA-seq datasets previously generated from either free swimming cyprids or an adult of *B. improvisus* [37]. Firstly, from each of these datasets we extracted ORFs from the transcriptomes using Transdecoder v3.0.1. The created HMM models were used to mine the obtained ORFs for sensory receptor candidates. Given the shared domains between several of the included receptor families, one expects the HMMs to pick up some false positive hits. To filter out the false positives and retain the best predictions for a specific receptor class, we concatenated all the results sorted based on the E-value, and the hit with the lowest E-value from each HMM search was kept. The resulting top candidates were then compared against the NCBI nr database using BLASTP to further confirm the receptor identity and to exclude any false positives, i.e. making sure the best hit from the HMM search indeed belonged to the right receptor class annotation. Furthermore, we used the SMART (http://smart.embl-heidelberg.de) [38] domain search web-server to predict the domain content of the candidate receptor sequences from *B. improvisus*.

The initial searches in the *B. improvisus* antennule transcriptome revealed several short hits that despite the short length could be identified as potential sensory receptors. In an attempt to extend the sequence of these partial candidates, we used available Illumina and PacBio RNA datasets from the adults and cyprids of *B. improvisus* [37,39]. These extended candidates were further checked against genomic data (unpublished) to verify sequences. The sequences have been deposited in GenBank with accession numbers MK093193 –MK093208.

Sequence and phylogenetic analysis of identified candidates

Identified candidate receptors were aligned with arthropod receptor sequences from NCBI nr database (see S2 File for the accession numbers). Phylogenetic trees based on the resulting alignments were made with PhyML method at phylogeny.fr (http://www.phylogeny.fr). The receptor candidates were named according to amino acid similarity to previously identified *Drosophila* and *Panulirus argus* IRs [11,15] as well as classification from the phylogenetic analysis results.

Expression analysis of candidate genes during settlement

The candidate sensory receptors were analysed during a developmental time series experiment to identify their expression changes during settlement. The details of the experimental set-up and the sampling of different settling stages are described in Abramova et al., 2018 (manuscript). The data is available at the NCBI under BioProject ID PRJNA528777. Briefly, 1–2 days old cyprid larvae were left to settle in petri dishes for 4–5 days. After that time-period, we collected several settling/developmental stages, including free swimming cyprids, cyprids during close search, early attached cyprids and early metamorphosed juveniles. Each sample contained 20 individuals and was collected in three independent replicates. RNA was extracted with previously optimized protocol [27] and sent for sequencing at the Science for Life Laboratory in Stockholm. Obtained RNA-seq data was assembled with Trinity software and used to quantify transcripts' expression levels. Subsequently, differential gene expression analysis was

performed with the edgeR package using the likelihood ratio test [40]. Genes that had adjusted p-values equal or smaller than 0.05 and a fold change more than 4 were reported.

Results

Antennular transcriptome

In order to enrich for antennual sensory genes we dissected and collected antennule pairs from 150 cyprid individuals (Fig 1) from which we extracted RNA. Sequencing of the obtained RNA generated more than 260 million good quality read pairs (in total roughly 89 Gbp). *De novo* assembly was perfomed using Trinity [29] and clustered with CD-HIT to reduce redundancy. The resulted assembly contained 214,110 Trinity transcripts representing 142,285 Trinity genes (Table 1). Average contig length and N50 value were 417 and 458, respectively. Transdecoder was used to identify candidate coding regions within the assembled contigs resulting in total of 90,896 ORFs from which 15,931 were complete ORFs.

To assess the quality and duplication level in the antennular transcriptome, we ran the BUSCO set comprised of arthropod single copy orthologs (S1 File). The analysis revealed a duplication level of 18%. Among all transcripts, 7.4% were predicted to have complete ORFs (Table 1).

Transcriptome annotation

We performed functional annotation of the antennular transcriptome using Trinotate, resulting in 15,953 annotated transcripts, which is similar to the number of functionally annotated contigs in the whole cyprid *de novo* transcriptome assembly (17,015) (S1 File; unbuplished data). GO enrichment analysis was done using the annotated transcripts. In the subset of genes that were at least ten times more abundant in the antennules than in the whole cyprid dataset, we observed several GO terms enriched in the Molecular Function categories related to the sensory function performed by antennules. In particular, "receptor activity", "molecular transducer activity" and "transferase activity" were among the top 15 enriched GO terms for the genes with higher expression in the antennules than in the whole body (Fig 2A). To get an overview of the differences between genes expressed in the antennules and whole body datasets, we also mapped GO terms from both datasets to the GO slim subset. Comparison revealed that the distribution of GO terms are overall very similar between the two datasets (S1 Fig). However, when we looked at the GO terms assigned to the GO slim category "signal transduction" (GO:0007165) we could see that the antennular dataset showed enrichment for signaling related genes compared to the whole cyprid transcriptome, e.g. genes involved in

Table 1. ***De novo*** **antennular transcriptome assembly and annotation statistics, nt = nucleotides.**

Total assembled bases	89 341,753
Total Trinity transcripts*	214,110
Total Trinity genes	142,285
Average contig length (nt)	417
Contig N50 (nt)	458
ORFs**	90,896
Annotated genes	15,953

*after CD hit clustering and reducing redundancy by removing duplicates at 95% identity level.

**all types of ORFs predicted by Transdecoder (15,931 complete, 35,284 internal, and 39,654 partial)

https://doi.org/10.1371/journal.pone.0216294.t001

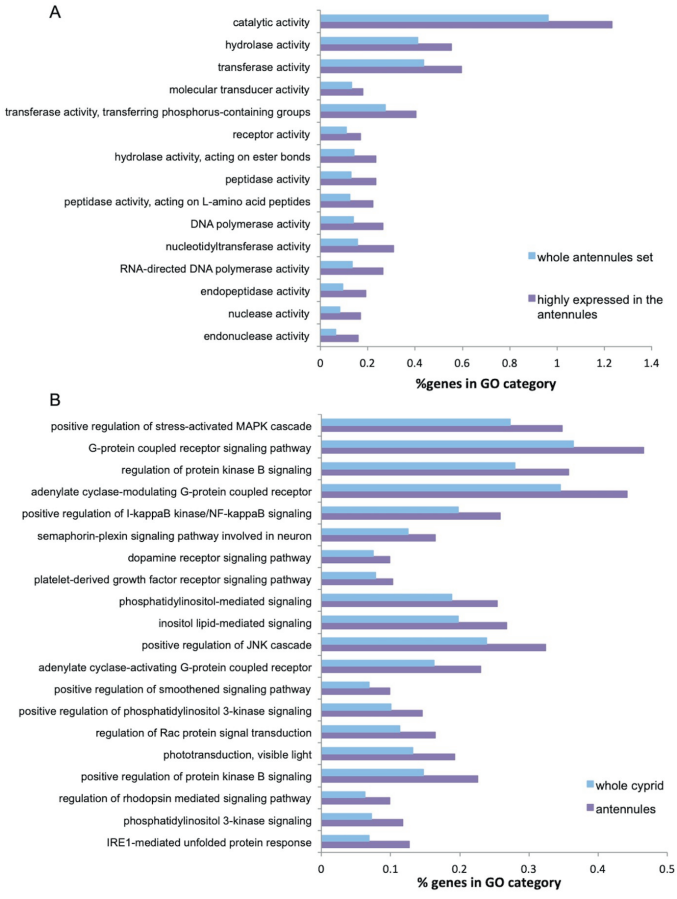


Fig 2. Functional characterisation of the antennular dataset. (A) GO enrichment for the top 15 Molecular Function categories of the subset of genes highly expressed in antennular dataset (at least 10 times more than in the cyprid) compared to the whole antennular dataset as a background. Each GO term has at least 15 genes assigned in the antennular dataset. (B) GO enrichment comparison of the genes in the Biological Process category 'signal transduction' (GO:0007165) between the antennules and whole cyprid datasets. Categories are sorted—left to right—based on the fold-change differences between the two dataset. Each GO term has at least 20 genes assigned in the antennular dataset.

https://doi.org/10.1371/journal.pone.0216294.g002

"phosphatidyl 3-kinase signalling" and "adenylate cyclase-activating G-protein coupled signalling" (Fig 2B).

HMM models and search results

In order to find sensory receptor genes that are expressed in the cyprid antennules, we constructed nine HMM models based on previously annotated sensory receptor sequences from arthropods. Our HMM searches of the antennular transcriptome identified dozens of sensory

receptor candidates (S5 File). However, a majority of these candidates were partial ORFs found on short contigs, too short to be unambiguously identified as belonging to a particular class of receptors. These partial candidate sequences were extended when possible, using transcriptomic contigs from whole cyprid and adults, and these extended candidates were confirmed using DNA data. This resulted in the identification of several putative IRs, including IR25a, IR8a, and IR93a as well as four additional divergent IRs. Search for TRP receptors revealed many false positives comprising fragments of other proteins containing ankyrin repeates also present in TRPs. The filtering resulted in identitifcation of 13 TRP candidates in the antennular dataset. In addition, we found one GR-like candidate expressed in antennules and five more were retrieved from the adult transcriptome assembly. Identified GR candidates were all partial and had a very low identity to the known GRs (26–38%) (S3 File). In particular, the identified GR-like receptor sequence from antennules was 171 aa long containing only two transmembrane domains out of seven normally present in GRs. It had 30% identity to the previously reported *B. amphitrite* partial (80aa) GR-like sequence [12] and 38% identity to *Drosophila hydei* putative gustatory receptor 98b (XP_023161714.1). Furthermore, the identified *B. improvisus* GR-like receptor had the signature amino acid domain of the GR family [41], the "TYxxxxxQF" motif (S4 File).

We also searched for the ORs, odorant-binding proteins and chemosensory proteins, three groups of proteins that are important for chemosensory function in several arthropod groups. We did not detect any clear candidates for these classes, in line with earlier results indicating that these proteins are specific to Hexapoda [16].

Ionotropic receptors

After filtering of the initial candidates scored by HMMs, we identified several putative IRs in our antennual RNA sequences (Table 2). One of the candidates, *Bimp*IR25a, revealed 56% identity to the spiny lobster *P. argus* subunit IR25a and 52% identity to the *D. melanogaster* IR25 at the amino acid level. Domain prediction with SMART showed the presence of the two characteristic IR domains: a conserved domain corresponding to the periplasmic substrate binding protein family (PBPe) (E-value $9.37e^{-60}$), overlapping with a Ligated ion channel L-glutamate- and glycine-binding site (E-value $2.46e^{-19}$) (Fig 3).

Furthermore, we found one IR8a candidate in *B. improvisus* corresponding to a protein of 834aa. Sequence alignment revealed 49% identity to the *P. argus* IR8a (AGJ51189.1) and 44% to *D. melanogaster* IR8a (NP_727328.1). Sequence analysis showed the presence of both IR-characteristic domains in *Bimp*IR8a (Fig 3). Furthermore, one gene encoding IR93a was found in the *B. improvisus* antennular dataset. The sequence corresponded to 853 aa, with 40% identity to *P. argus* IR93a protein sequence (AGJ51190.1), and contained the ion channel pore region and a part of the ligand-binding domain.

The remaining identified IR candidates had low similarity to the known IRs and iGluRs (26–35% ID) (Table 2). Despite the fact that the overall sequence identity to the known IRs and iGluRs was low, four of the candidates contained parts of the ligand-binding domain and ion channel pore region; we named these *Bimp*IR1 to 4. We found that two of these putative IR candidates, *Bimp*IR4 and *Bimp*IR5, were located tail to tail on different strands of the same genomic scaffold of 42,766 kb (S2 Fig). It has been previously observed that IRs are often organised in tandem arrays in the genome in *Drosophila*, as a result of recent local gene duplications [16]. We found only one such example in our *B. improvisus* contigs likely due to the fragmented nature of the genome assembly.

To classify the putative IRs a phylogenetic analysis was performed. The analysis showed that IRs formed two clusters in the phylogenetic tree of IRs (Fig 4). One cluster contained

Table 2. *B. improvisus* candidate IRs, BLASTp search was carried against the Arthropod nr NCBI database.

Gene name	Accession numbers	Length (aa)	Best blastp hit	E-value	Identity
*Bimp*IR25a	MK093206	922	olfactory ionotropic receptor IR25a, *Panulirus argus AGJ51188.1*	0.0	56%
*Bimp*IR8a	MK093207	860	Ionotropic receptor 8a *Blattella germanica, PSN54615.1*	0.0	50%
*Bimp*IR93a	MK093208	849	olfactory ionotropic receptor IR93a, *Panulirus argus AGJ51190.1*	0.0	41%
*Bimp*IR1	-	525	ionotropic receptor 129, *Blattella germanica PSN40732.1*	3e-23	30%
*Bimp*IR2		542	ionotropic receptor 76a *Daphnia magna, KZS15214.1*	2e-77	35%
*Bimp*IR3	-	469	ionotropic receptor 21a-like *Eurytemora affinis, XP_023320534.1*	1e-35	26%
*Bimp*IR4	-	357	variant ionotropic glutamate receptor *Coenobita clypeatus, CEF34384.1*	3e-29	27%

https://doi.org/10.1371/journal.pone.0216294.t002

*Bimp*IR25a and *Bimp*IR8a together with the IR25a/IR8a and ancestral iGluRs from *Drosophila* and several crustacean species used as a reference. Another cluster comprised species-specific/divergent IRs. This cluster contained *Bimp*IR93a that formed a clade with IR93a orthologs from the reference species, and other divergent IRs showing similarity to *B. germanica* IR214 and *D. magna* IR76a.

We compared the ligand-binding domains of the identified antennual *B. improvisus* IRs with the corresponding *D. melanogaster* IR25a and IR8a that retain most of the three conserved glutamate-interacting residues of iGluRs (Fig 5). *B. improvisus* IR25a has retained

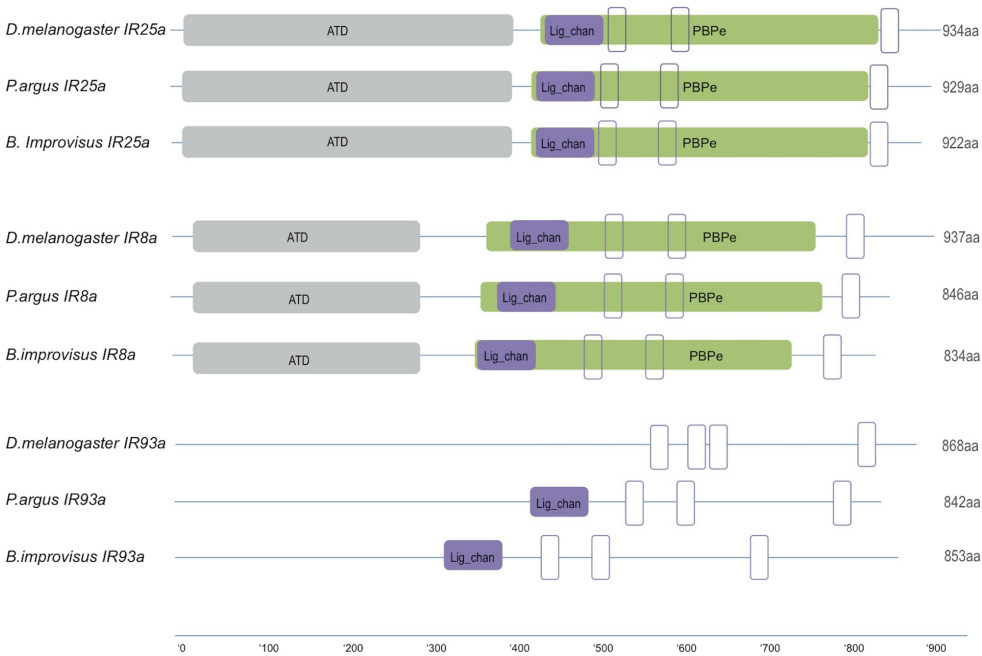


Fig 3. The comparison of protein domain organisation of IR25a, IR8a and IR93a from *D. melanogaster*, *P. argus* and *B. improvisus*. ATD amino-terminal domain (grey), Lig_chan Ligated ion channel L-glutamate- and glycine-binding site (PF10613) (violet) overlapping with PBPe (PF00060) domain (green). White boxes represent transmembrane domains predicted with TMHMM Server v. 2.0.

https://doi.org/10.1371/journal.pone.0216294.g003

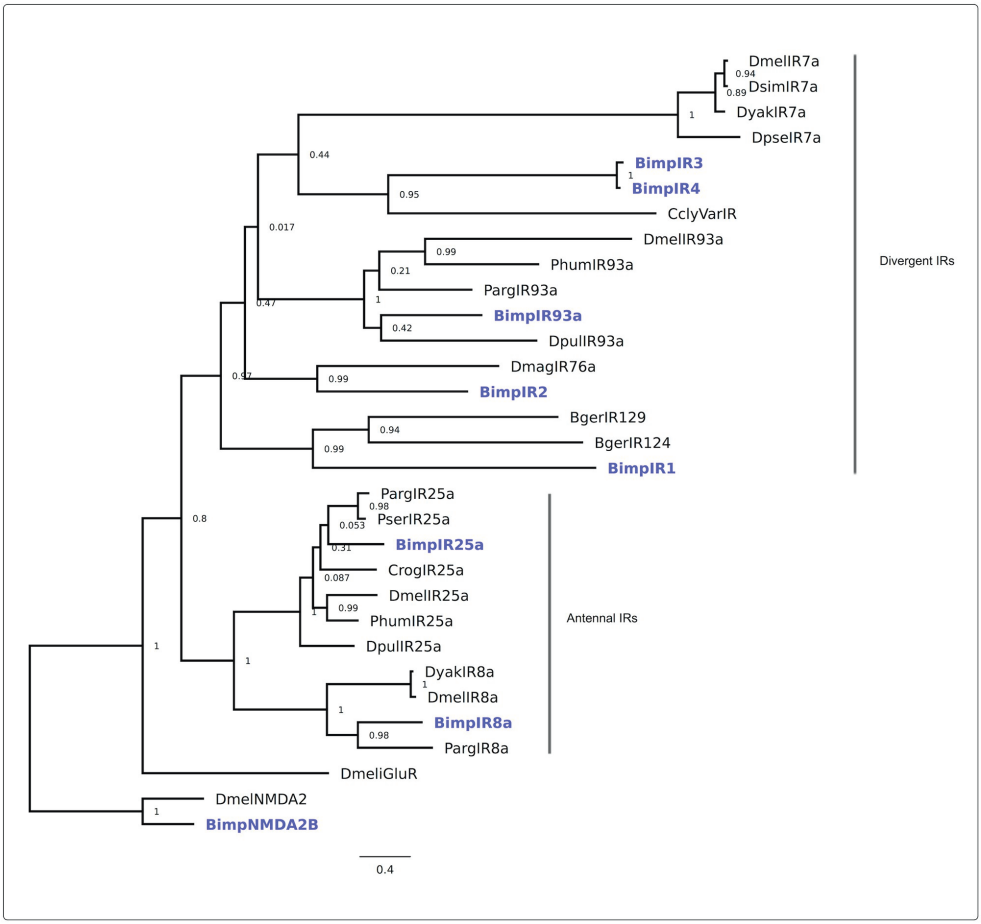


Fig 4. Maximum-likelihood phylogeny for *B. improvisus* candidate IRs including IR25a, IR8a, IR93a, and four divergent *Bimp*IR1-4. IR candidate amino acid sequences from *B. improvisus* were aligned with reference sequences from other arthropod species from NCBI: Parg *P. argus*, Dmel *D. melanogaster*, Phum *Pediculus humanus*, Dpul *D. pulex*, Ccly *Coenobita clypeatus*, Dsim *D. simulans*, Dyak *D. yakuba*, Dpse *D. pseudoobscura*, Dmag *D. magna*, Bger *B. germanica*, Crog *Caligus rogercresseyi*, Pser *Palaemon serratus*. See S2 File for the accession numbers. NMDA type receptors from *D. melanogaster* and *B. improvisus* were used as an outgroup. The numbers indicate bootstrap values for each branch. The scale bar indicates the inferred number of substitutions per site.

https://doi.org/10.1371/journal.pone.0216294.g004

arginine (R), threonine (T) and aspartate/glutamate (D/E) residues characteristic for the iGluRs, while IR8a has a threonine residue replaced by isoleucine (I) in the S2 domain. We observed that among our putative IR candidates a few, especially divergent IRs, lack one or more of these conserved residues suggesting that they are not binding glutamate like in the case of *Drosophila* IRs [15].

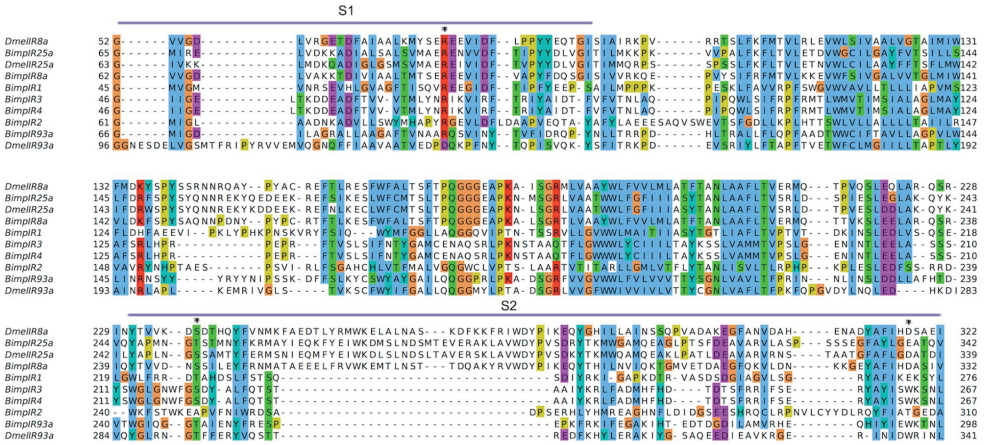


Fig 5. Alignment of PBPe domains from *D. melanogaster* IR25a, IR8a, IR93a and candidate *B. improvisus* IRs. The three glutamate binding positions characterised from Benton et al. (2009) are marked with asterisk. S1 and S2 stand for ligand-binding lobes of the "Venus flytrap" of the iGluRs.

https://doi.org/10.1371/journal.pone.0216294.g005

Transient receptor potential channels

Using the HMM models when searching the *B. improvisus* antennular dataset we also identified 13 sequences representing putative TRP receptors (Table 3). Further, to assign the candidates to TRPs subfamilies we made maximum-likelihood phylogeny including TRPs from *Drosophila* as well as from several crustacean species (Fig 6). This analysis revealed strong evidence that five of our 13 TRP candidates belong to the TRPA family, including TRPA1,

Table 3. Table of TRP candidates identified from cyprid antennules transcriptome.

Gene name	Accession numbers	Length (aa)	Best blast hit	E-value	Identity
*Bimp*_TRPA1	MK093202	1262	transient receptor potential cation channel subfamily A member 1 homolog isoform X1 *Limulus polyphemus*, *XP_022241898.1*	0.0	47%
*Bimp*_painless	MK093200	909	transient receptor potential cation channel protein painless-like *Eurytemora affinis*, *XP_023338859.1*	4e-130	33%
*Bimp*_pyrexia1	MK093194	1191	transient receptor potential channel pyrexia-like *Zootermopsis nevadensis*, *XP_021941613.1*	2e-177	37%
*Bimp*_pyrexia2	MK093195	881	transient receptor potential channel pyrexia *Microplitis demolitor*, *XP_008553590.1*	2e-73	32%
*Bimp*_pyrexia3	MK093205	923	transient receptor potential channel pyrexia-like *Hyalella azteca*, *XP_018007126.1*	9e-107	34%
*Bimp*_nompC	MK093193	986	ion channel nompc *Culex quinquefasciatus*, *XP_001842099.1*	0.0	46%
*Bimp*_TRPM1	MK093204	1002	transient receptor potential cation channel trpm-like *Eurytemora affinis*, *XP_023322531.1*	0.0	35%
*Bimp*_TRPM2	MK093196	1654	transient receptor potential cation channel trpm *Nilaparvata lugens*, *XP_022186364.1*	0.0	47%
*Bimp*_TRPV	MK093197	762	TRP channel protein inactive *Nephotettix cincticeps*, *ATU07274.1*	0.0	63%
*Bimp*_TRPgamma	MK093198	846	transient receptor potential-gamma protein isoform X2 *Vollenhovia emeryi*, *XP_011868062.1*	0.0	67%
*Bimp*_PKD1-like1	MK093199	976	polycystic kidney disease protein 1-like 2 *Limulus polyphemus*, *XP_022236079.1*	0.0	45%
*Bimp*_polycystin2	MK093201	735	Polycystin-2-like *Limulus polyphemus*, *XP_013778837.1*	0.0	60%
*Bimp*_PKD1-like2	MK093203	2422	polycystic kidney disease protein 1-like protein 2-like protein *Euroglyphus maynei*, *OTF78065.1*	0.0	32%

https://doi.org/10.1371/journal.pone.0216294.t003

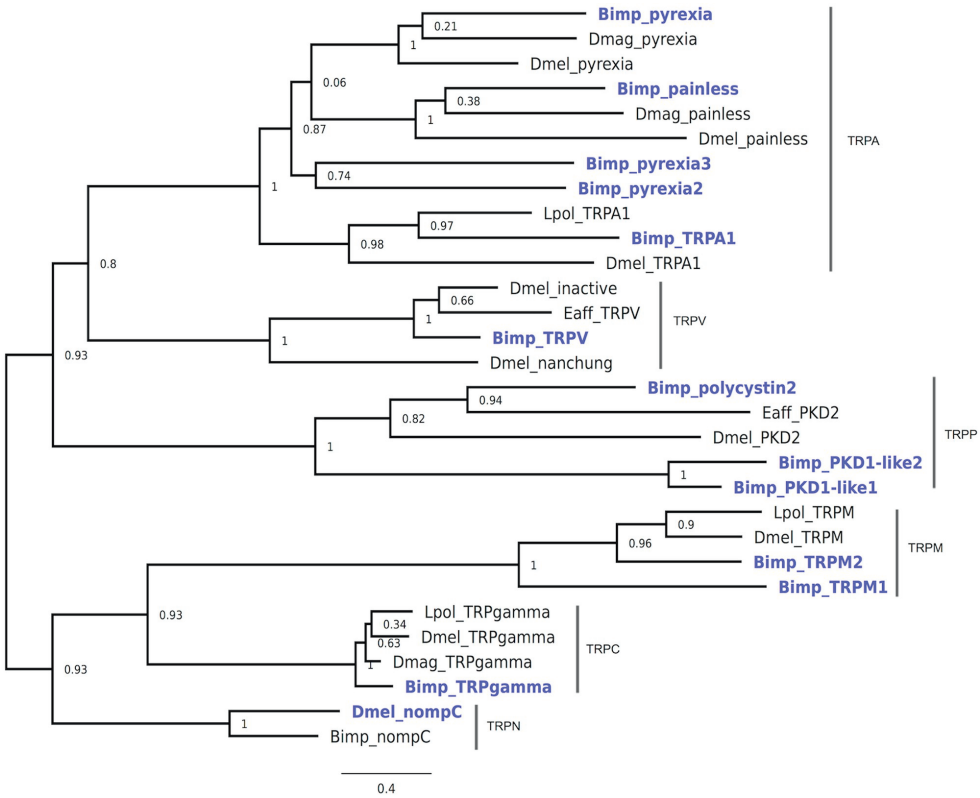


Fig 6. *B. improvisus* and other arthropod TRP channels. Amino acid sequences of candidates from *B. improvisus* (*Bimp*) and other arthropod species (accession numbers S2 File) were aligned using MUSCLE program following by PhyML for maximum likelihood analysis using online phylogeny.fr server. *B.* The numbers indicate bootstrap values for each branch. The scale bar indicates the inferred number of substitutions per site.

https://doi.org/10.1371/journal.pone.0216294.g006

painless and three *pyrexia*-like sequences. In addition, we found two TRPM-like transcripts, three *polycystic kidney disease*-like proteins belonging to TRPP subfamily, and transcripts corresponding to TRPV and TRPγ.

In addition, one contig was annotated as the ion channel nompC ("no mechanoreceptor potential C"), the first TRP receptor described from *Drosophila*. This protein showed 49% identity (E 0.0) with *D. melanogaster* nompC (NP_001097089.2). *Drosophila* nompC is 1,726 aa long and contains 29 ankyrin repeats, however, the *B. improvisus* sequence is partial (986 aa), lacking part of the N-terminus and containing only eleven ankyrin repeats.

The phylogenetic analysis revealed that *B. improvisus* TRPs cluster well together with corresponding representatives of well-studied classes from *Drosophila* and other arthropods with high support (Fig 6).

Expression of the sensory receptors during cyprid settlement

We performed differential gene expression analysis to investigate how the identified sensory receptor candidates were expressed during distinct settlement-stages, i.e. free swimming cyprids, cyprids during close search, early attached cyprids and early metamorphosed juveniles (Fig 7 and S2 Table). Expression of the ionotropic receptors appears to be highest in the free-swimming and close-search stages and decreases substantially after attachment of cyprids, with the exception for IR8a expression that remained high in the attached stage followed by a roughly 3-fold decrease in the juvenile. The overall considerably higher expression of the IR25 candidate (roughly 1,500 FPKM in comparison to about 50 FPKM for the other two IRs) is compatible with IR25 being the common co-receptor for many of the more specific IRs [11,15], thus being more highly expressed or/and in a greater number of cells compared to the other IR classes. In contrast, the divergent IRs demonstrated comparatively low and highly variable expression, consistent with their suggested cell-specific expression [17]. The divergent IR2 as well as GR-like receptor were not found in the settling assembly, probably due to extremely low expression levels.

For the TRP channels we observe quite variable expression during settlement (Fig 8). Similar to the IRs, the majority of the candidate TRPs are downregulated in the juvenile stage compared to the cyprid stages. However, nompC showed high expression in the attached stage, whereas pyrexia 2 and PKD1-like 1 levels did not change significantly during the settlement.

Discussion

A remarkable feature of barnacles is their complex settlement behavior allowing them to find a suitable attachment site for survival and reproduction. During settlement, barnacle larvae display a high degree of specificity towards surface chemistry and structure, as well as the ability to discriminate between conspecific and other barnacle species' cues [42]. While some of the cues triggering settlement have been characterised, the exact molecular mechanisms of how cyprids receive these cues remain largely unknown. In the present study we identified three types of sensory receptors in the barnacle *B. improvisus*, olfactory ionotropic receptors, gustatory-like receptors and mechanosensory transient receptor potential channels, present in the cyprid antennules that most likely are involved in the sensing of settlement cues. We could not find any signs of odorant receptors, chemosensory or odorant-binding proteins known to have chemosensory function in other invertebrate species. These results are consistent with the reported absence of these genes from other crustaceans, i.e. the hermit crab and lobster antennal transcriptomes [11,13]. Consistent with the number of GRs found in the antennular flagellum of spiny lobster *P. argus* [43], we identified only one GR-like receptor in cyprid antennules, and five additional in the adult transcriptome. GRs have a diverse set of functions ranging from detection of sugars and carbon dioxide to sensing of heat and light [41]. Being one of the most ancient receptors in Arthropoda, GRs expanded in some lineages, eg *D. pulex* has 58 GRs [14], and were completely lost in others. Unfortunately, the fragmented nature of our datasets did not allow for more in-depth analysis and most likely underestimates the true number of GR genes in *B. improvisus*.

Our transcriptomic data from antennules of cyprids show the presence of several transcripts from ionotropic receptors. Evidence from recent studies suggests that IRs are widely distributed among crustaceans and are the main candidates to initiate chemosensory signaling [11]. The total number of IRs varies greatly from species to species with the hermit crab *C. clypeatus* having 27 and *D. pulex* having 85 IRs [16,18]. Based on the protein sequences and phylogenetic analysis we identified and classified the *B. improvisus* IR25a, IR8a and IR93 subunits, as well as four divergent IRs. Furthermore, we identified dozens of sequences in the antennal

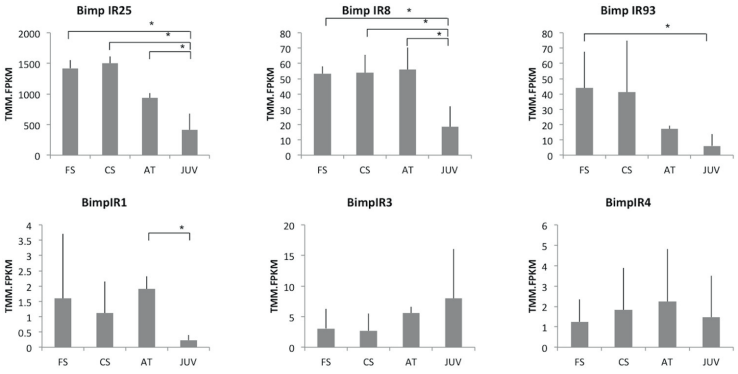


Fig 7. *B. improvisus* IRs expression during settlement. FS free-swimming, CS close search, AT early-attached stage, JUV juvenile. Significant changes in gene expression (FDR 0.05) are indicated with an asterisk.

https://doi.org/10.1371/journal.pone.0216294.g007

dataset, as well as in whole-cyprids and adults, that had sequence homology to IRs and iGluRs and might belong to the divergent IRs, however, the fragmented nature of the data does not allow to unambiguously classify them. The total size of the IR repertoire in *B. improvisus* is currently difficult to estimate without a better genome assembly as a reference.

According to the current insect model, the common IR25a and IR8a subunits are present in each olfactory receptor neuron and form heteromers with cell-specific divergent IRs rendering ligand specificity of individual sensory cells [17]. Indeed, we observed that *B. improvisus* IR25a showed considerably higher expression in cyprids compared to other identified IRs, while the expression level of the divergent IRs were considerably lower consistent with their cell-specific role [11]. Our data show that common IR subunits (IR25a, IR8a and IR93a) in *B. improvisus* were highly expressed during surface exploration and early attachment, suggesting possible importance in detecting settlement cues. Noticeably, these IR subunits are also highly expressed in the settlement stage of the parasitic salmon louse *Lepeophtheirus salmonis* [44] and crucial for the identification of the host species [44]. Furthermore, members of the IR20 clade in *Drosophila* are suggested to be involved in sensing pheromones and regulation of courtship behaviour [45]. It is thus likely that the IRs present in the cyprids antennules could be also involved in sensing barnacle pheromones and therefore are potential targets to study pheromone-evoked settlement behavior of cyprids.

Apart from chemical cues, cyprids are thought to percieve water flow and distinguish surface structures by means of mechanosensory setae [7–9,46]. Cyprids of *B. improvisus* prefer to settle on smooth rather than structured surfaces [47]. According to current knowledge, mechanoreception in arthropoda is accomplished through mechanosensitive ion channels that upon stress allow the influx of cations resulting in a receptor potential [19]. Based on the effect of specific agonists of calcium channels, it has been suggested that TRP channels mediate calcium entry in neurons when the cyprid senses a favourable substratum [22]. With HMM models based on the known arthropod TRP receptors, we detected 13 TRP candidates in the antennular transcriptome of *B. improvisus* cyprids (Table 3). This number corresponds roughly to the number of TRP genes found to be expressed in *D. pulex* transcriptome, with the exception that TRPP genes seems to be absent in *Daphnia* [48].

We identified five candidates belonging to TRPA receptors including TRPA1, *painless* and three *pyrexia* genes [20]. TRPA1 is involved in sensing of volatile molecules in *Drosophila* and

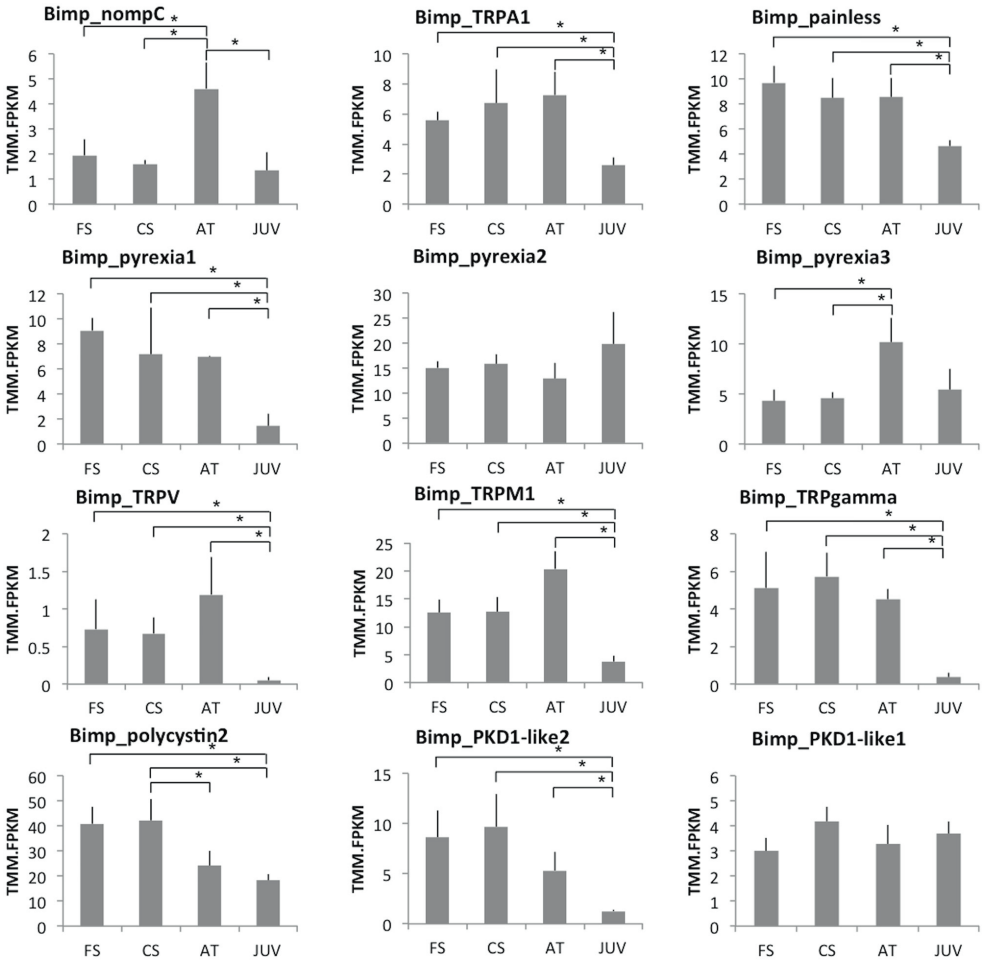


Fig 8. Expression of the *B. improvisus* TRP channels during settlement. TRPM2 shows exactly the same expression pattern as TRPM1 and was not shown. FS free-swimming, CS close search, AT early-attached stage, JUV juvenile. Asterisk shows significant change in gene expression FDR 0.05.

https://doi.org/10.1371/journal.pone.0216294.g008

nose-touch responses and foraging in *Caenorhabditis elegans* [49,50]. The *B. improvisus* homolg of TRPA1 was highly expressed during exploratory stages and early attachment. The *painless* gene in *Drosophila* larvae is involved in behavioral responses to thermal and mechanical stimuli as well as playing a role in reception of pheromones and sexual behavior in adults [51–53]. Interestingly, benzyl isothiocyanate that is an activator for both TRPA1 and *painless* channels had an effect on the settling of cyprids, inhibiting settlement at lower concentrations and stimulating at higher, suggesting that TRP channels activation generate Ca^{2+} signals that coordinate the settlement process [22]. Like in *D. pulex* [20], the *pyrexia* gene appears to be

expanded in *B. improvisus* being represented by three paralogues. While the role of *pyrexia* in crustaceans remains to be established, this gene detects a range of temperatures and negative geotaxis in *Drosophila* and contributes to host-seeking in *Anopheles gambiae* [54].

Furthermore, we found two TRPM-like transcripts among *B. improvisus* TRP candidates. In *Drosophila*, TRPM is mainly involved in ion homeostasis during larval development [20,49], whereas in mammals TRPM is expressed specifically in the olfactory epithelium and is involved in transducing pheromone signals [55]. The fourth segment of the cyprid antennule carries an aesthetascs-like setae that was suggested to sense waterborne chemical cues [7]. In this perspective, future studies on the localization of the expression of particular types of TRPs by, for example, *in situ* hybridization are of a great interest. Two of the other identified TRP channels, namely TRPgamma and TRPV, showed expression patterns similar to the TRPM with high expression in the exploratory stages and decreased expression in the juvenile stage. TRPgamma is localised in the proprioceptive cells and contribute to fine motor control in *Drosophila* and *C. elegans* [56]. Fine motor control in flies is required for challenging tasks, which rely on coordinating a repertoire of fine movements, including subtle changes in body angles and leg positions [56], which resemble the complex movements of cyprid antennules during the surface exploration [7]. The TRPV receptor is considered as a molecular integrator of noxious stimuli ranging from pungent natural products, such as capsaicin, to acidic environment and high temperatures [57]. A recent study suggested that cyprid settlement could be akin to the behavioral responses of *Drosophila* to noxious stimuli mediated by TRPs [22]. In particular, surface avoidance would be mediated by noxious stimuli-sensing channels, cancelling the signals that complete settlement. Therefore, further investigation of the antennual TRPs could bring new insights into the cyprid decision-making process.

The current study provides the first molecular evidence of the existence of IRs and TRP channels in barnacles cyprid antennules and their variable expression during the settling, thus suggesting possible roles for them in the process of barnacle settlement. The generated datasets bring new opportunities to further investigate settling behavior and provide possible new targets for development of antifouling agents with more specific effects on barnacles.

Supporting information

S1 File. Annotation and BUSCO assessment of the datasets.
(XLSX)

S2 File. Arthropod sequences used as a reference for the *B. improvisus* IRs phylogenetic analysis.
(XLSX)

S3 File. GR-like candidates identified from antennular and adult *B. improvisus* transcriptomes.
(PDF)

S4 File. *B. improvisus* GR-like fragment alignment.
(PDF)

S5 File. Results of the searches with constructed HMM models.
(PDF)

S6 File. Fasta file with amino acid sequences described in the text.
(TXT)

S1 Fig. Comparison of the distribution of GO slim categories between whole body cyprid and antennular dataset.
(PNG)

S2 Fig. Location of the putative IR candidates on the genomic scaffold.
(PNG)

S1 Table. The number and taxonomic representation of sequences used to construct HMM models.
(XLSX)

S2 Table. EdgeR results for the receptor candidates.
(XLSX)

Acknowledgments

We would like to acknowledge Science for Life Laboratory (SciLifeLab Stockholm), National Genomics Infrastructure (NGI–Sweden), for providing sequencing technical service. The authors would like to thank Martin Ogemark for providing and maintaining barnacle larvae culture, John Havenhand for advice on microscopy, and Ulrika Lind and Per Jonsson for valueable comments on the manuscipt.

Author Contributions

Conceptualization: Anna Abramova, Tomas Axel Larsson.

Data curation: Anna Abramova, Magnus Alm Rosenblad, Tomas Axel Larsson.

Formal analysis: Anna Abramova, Magnus Alm Rosenblad, Tomas Axel Larsson.

Funding acquisition: Anna Abramova, Anders Blomberg, Tomas Axel Larsson.

Investigation: Anna Abramova, Tomas Axel Larsson.

Methodology: Anna Abramova, Tomas Axel Larsson.

Project administration: Tomas Axel Larsson.

Resources: Anna Abramova, Tomas Axel Larsson.

Supervision: Anders Blomberg, Tomas Axel Larsson.

Validation: Magnus Alm Rosenblad, Anders Blomberg, Tomas Axel Larsson.

Visualization: Anna Abramova, Tomas Axel Larsson.

Writing – original draft: Anna Abramova, Anders Blomberg, Tomas Axel Larsson.

Writing – review & editing: Anna Abramova, Magnus Alm Rosenblad, Anders Blomberg, Tomas Axel Larsson.

References

1. Aldred N, Clare AS. The adhesive strategies of cyprids and development of barnacle-resistant marine coatings. Biofouling. 2008; 24(5):351–63. https://doi.org/10.1080/08927010802256117 PMID: 18597201
2. Lagersson NC, Hoeg JT. Settlement behavior and antennulary biomechanics in cypris larvae of Balanus amphitrite (Crustacea: Thecostraca: Cirripedia). Mar Biol. 2002; 141(3):513–26.

3. Aldred N, Alsaab A, Clare AS. Quantitative analysis of the complete larval settlement process confirms Crisp's model of surface selectivity by barnacles. Proc R Soc B. 2018; 285(1872):20171957. https://doi.org/10.1098/rspb.2017.1957 PMID: 29445024
4. Dreanno C, Matsumura K, Dohmae N, Takio K, Hirota H, Kirby RR, et al. An alpha2-macroglobulin-like protein is the cue to gregarious settlement of the barnacle Balanus amphitrite. Proc Natl Acad Sci U S A. 2006; 103(39):14396–401. https://doi.org/10.1073/pnas.0602763103 PMID: 16983086
5. Clare AS, Matsumura K. Nature and perception of barnacle settlement pheromones. Biofouling. 2000; 15(1–3):57–71. https://doi.org/10.1080/08927010009386298 PMID: 22115292
6. Chan BKK, Garm A, Hoeg JT. Setal morphology and cirral setation of thoracican barnacle cirri: adaptations and implications for thoracican evolution. J Zool. 2008; 275(3):294–306.
7. Maruzzo D, Conlan S, Aldred N, Clare AS, Hoeg JT. Video observation of surface exploration in cyprids of Balanus amphitrite: the movements of antennular sensory setae. Biofouling. 2011; 27(2):225–39. https://doi.org/10.1080/08927014.2011.555534 PMID: 21302160
8. Bielecki J, Chan BKK, Hoeg JT, Sari A. Antennular sensory organs in cyprids of balanomorphan cirripedes: standardizing terminology using Megabalanus rosa. Biofouling. 2009; 25(3):203–14. https://doi.org/10.1080/08927010802688087 PMID: 19169952
9. Lagersson NC, Garm A, Hoeg JT. Notes on the ultrastructure of the setae on the fourth antennulary segment of the Balanus amphitrite cyprid (Crustacea: Cirripedia: Thoracica). J Mar Biol Assoc Uk. 2003; 83(2):361–5.
10. Derby CD, Kozma MT, Senatore A, Schmidt M. Molecular Mechanisms of Reception and Perireception in Crustacean Chemoreception: A Comparative Review. Chem Senses. 2016; 41(5):381–98. https://doi.org/10.1093/chemse/bjw057 PMID: 27107425
11. Corey EA, Bobkov Y, Ukhanov K, Ache BW. Ionotropic Crustacean Olfactory Receptors. Plos One. 2013; 8(4).
12. Eyun S-i, Soh HY, Posavi M, Munro JB, Hughes DS, Murali SC, et al. Evolutionary history of chemosensory-related gene families across the Arthropoda. Mol Biol Evol. 2017; 34(8):1838–62. https://doi.org/10.1093/molbev/msx147 PMID: 28460028
13. Groh KC, Vogel H, Stensmyr MC, Grosse-Wilde E, Hansson BS. The hermit crab's nose-antennal transcriptomics. Front Neurosci-Switz. 2014; 7.
14. Penalva-Arana DC, Lynch M, Robertson HM. The chemoreceptor genes of the waterflea Daphnia pulex: many Grs but no Ors. Bmc Evol Biol. 2009; 9.
15. Benton R, Vannice KS, Gomez-Diaz C, Vosshall LB. Variant Ionotropic Glutamate Receptors as Chemosensory Receptors in Drosophila. Cell. 2009; 136(1):149–62. https://doi.org/10.1016/j.cell.2008.12.001 PMID: 19135896
16. Croset V, Cummins SF, Benton R. Ancient Protostome Origin of Chemosensory Ionotropic Glutamate Receptors and the Evolution of Insect Taste and Olfaction. J Neurogenet. 2010; 24:30–1.
17. Abuin L, Bargeton B, Ulbrich MH, Isacoff EY, Kellenberger S, Benton R. Functional Architecture of Olfactory Ionotropic Glutamate Receptors. Neuron. 2011; 69(1):44–60. https://doi.org/10.1016/j.neuron.2010.11.042 PMID: 21220098
18. Groh-Lunow KC, Getahun MN, Grosse-Wilde E, Hansson BS. Expression of ionotropic receptors in terrestrial hermit crab's olfactory sensory neurons. Front Cell Neurosci. 2015;8.
19. Schuler A, Schmitz G, Reft A, Ozbek S, Thurm U, Bornberg-Bauer E. The Rise and Fall of TRP-N, an Ancient Family of Mechanogated Ion Channels, in Metazoa. Genome Biol Evol. 2015; 7(6):1713–27. https://doi.org/10.1093/gbe/evv091 PMID: 26100409
20. Peng GD, Shi X, Kadowaki T. Evolution of TRP channels inferred by their classification in diverse animal species. Mol Phylogenet Evol. 2015; 84:145–57. https://doi.org/10.1016/j.ympev.2014.06.016 PMID: 24981559
21. Nilius B, Appendino G, Owsianik G. The transient receptor potential channel TRPA1: from gene to pathophysiology. Pflug Arch Eur J Phy. 2012; 464(5):425–58.
22. Kotsiri M, Protopapa M, Roumelioti GM, Economou-Amilli A, Efthimiadou E, Dedos S. Probing the settlement signals of Amphibalanus amphitrite (vol 34, pg 492, 2018). Biofouling. 2018; 34(5):X–X.
23. Kawahara H, Isoai A, Shizuri Y. Molecular cloning of a putative serotonin receptor gene from barnacle, Balanus amphitrite. Gene. 1997; 184(2):245–50. PMID: 9031635
24. Lind U, Alm Rosenblad M, Hasselberg Frank L, Falkbring S, Brive L, Laurila JM, et al. Octopamine receptors from the barnacle balanus improvisus are activated by the alpha2-adrenoceptor agonist medetomidine. Mol Pharmacol. 2010; 78(2):237–48. https://doi.org/10.1124/mol.110.063594 PMID: 20488921

25. Grosse-Wilde E, Kuebler LS, Bucks S, Vogel H, Wicher D, Hansson BS. Antennal transcriptome of Manduca sexta. P Natl Acad Sci USA. 2011; 108(18):7449–54.
26. Nunez-Acuna G, Valenzuela-Munoz V, Marambio JP, Wadsworth S, Gallardo-Escarate C. Insights into the olfactory system of the ectoparasite Caligus rogercresseyi: Molecular characterization and gene transcription analysis of novel ionotropic receptors. Exp Parasitol. 2014; 145:99–109. https://doi.org/10.1016/j.exppara.2014.08.003 PMID: 25131775
27. Jonsson PR, Wrange AL, Lind U, Abramova A, Ogemark M, Blomberg A. The Barnacle Balanus improvisus as a Marine Model—Culturing and Gene Expression. Jove-J Vis Exp. 2018(138).
28. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet journal. 2011; 17(1):pp. 10–2.
29. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011; 29(7):644–U130. https://doi.org/10.1038/nbt.1883 PMID: 21572440
30. Fu LM, Niu BF, Zhu ZW, Wu ST, Li WZ. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012; 28(23):3150–2. https://doi.org/10.1093/bioinformatics/bts565 PMID: 23060610
31. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015; 31(19):3210–2. https://doi.org/10.1093/bioinformatics/btv351 PMID: 26059717
32. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013; 8(8):1494–512. https://doi.org/10.1038/nprot.2013.084 PMID: 23845962
33. Young MD, Wakefield MJ, Smyth GK, Oshlack A. goseq: Gene Ontology testing for RNA-seq datasets. R Bioconductor. 2012.
34. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol. 2013; 30(4):772–80. https://doi.org/10.1093/molbev/mst010 PMID: 23329690
35. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009; 25(9):1189–91. https://doi.org/10.1093/bioinformatics/btp033 PMID: 19151095
36. Johnson LS, Eddy SR, Portugaly E. Hidden Markov model speed heuristic and iterative HMM search procedure. Bmc Bioinformatics. 2010; 11.
37. Lind U, Jarva M, Rosenblad MA, Pingitore P, Karlsson E, Wrange AL, et al. Analysis of aquaporins from the euryhaline barnacle Balanus improvisus reveals differential expression in response to changes in salinity. Plos One. 2017; 12(7).
38. Letunic I, Doerks T, Bork P. SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Research. 2012; 40(D1):D302–D5.
39. Lind U, Alm Rosenblad M, Wrange AL, Sundell KS, Jonsson PR, Andre C, et al. Molecular characterization of the alpha-subunit of Na(+)/K(+) ATPase from the euryhaline barnacle Balanus improvisus reveals multiple genes and differential expression of alternative splice variants. Plos One. 2013; 8(10):e77069. https://doi.org/10.1371/journal.pone.0077069 PMID: 24130836
40. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26(1):139–40. https://doi.org/10.1093/bioinformatics/btp616 PMID: 19910308
41. Robertson HM. The insect chemoreceptor superfamily is ancient in animals. Chem Senses. 2015; 40(9):609–14. https://doi.org/10.1093/chemse/bjv046 PMID: 26354932
42. Dreanno C, Kirby RR, Clare AS. Involvement of the barnacle settlement-inducing protein complex (SIPC) in species recognition at settlement. J Exp Mar Biol Ecol. 2007; 351(1–2):276–82.
43. Kozma MT, Schmidt M, Ngo-Vu H, Sparks SD, Senatore A, Derby CD. Chemoreceptor proteins in the Caribbean spiny lobster, Panulirus argus: Expression of Ionotropic Receptors, Gustatory Receptors, and TRP channels in two chemosensory organs and brain. Plos One. 2018; 13(9):e0203935. https://doi.org/10.1371/journal.pone.0203935 PMID: 30240423
44. Komisarczuk AZ, Grotmol S, Nilsen F. Ionotropic receptors signal host recognition in the salmon louse (Lepeophtheirus salmonis, Copepoda). Plos One. 2017; 12(6).
45. Koh T-W, He Z, Gorur-Shandilya S, Menuz K, Larter NK, Stewart S, et al. The Drosophila IR20a clade of ionotropic receptors are candidate taste and pheromone receptors. Neuron. 2014; 83(4):850–65. https://doi.org/10.1016/j.neuron.2014.07.012 PMID: 25123314
46. Larsson AI, Jonsson PR. Barnacle larvae actively select flow environments supporting post-settlement growth and survival. Ecology. 2006; 87(8):1960–6. PMID: 16937634

47. Berntsson KM, Jonsson PR, Lejhall M, Gatenholm P. Analysis of behavioural rejection of micro-textured surfaces and implications for recruitment by the barnacle Balanus improvisus. J Exp Mar Biol Ecol. 2000; 251(1):59–83. PMID: 10958901

48. Dong CH, Hu AM, Ni Y, Zuo YX, Li GH. Effects of midazolam, pentobarbital and ketamine on the mRNA expression of ion channels in a model organism Daphnia pulex. Bmc Anesthesiol. 2013; 13.

49. Fowler MA, Montell C. Drosophila TRP channels and animal behavior. Life Sci. 2013; 92(8–9):394–403. https://doi.org/10.1016/j.lfs.2012.07.029 PMID: 22877650

50. Kindt KS, Viswanath V, Macpherson L, Quast K, Hu HZ, Patapoutian A, et al. Caenorhabditis elegans TRPA-1 functions in mechanosensation. Nat Neurosci. 2007; 10(5):568–77. https://doi.org/10.1038/nn1886 PMID: 17450139

51. Al-Anzi B, Tracey WD, Benzer S. Response of Drosophila to Wasabi is mediated by painless, the fly homolog of mammalian TRPA1/ANKTM1. Curr Biol. 2006; 16(10):1034–40. https://doi.org/10.1016/j.cub.2006.04.002 PMID: 16647259

52. Sakai T, Kasuya J, Kitamoto T, Aigaki T. The Drosophila TRPA channel, Painless, regulates sexual receptivity in virgin females. Genes Brain Behav. 2009; 8(5):546–57. https://doi.org/10.1111/j.1601-183X.2009.00503.x PMID: 19531155

53. Tracey WD, Wilson RI, Laurent G, Benzer S. painless, a Drosophila gene essential for nociception. Cell. 2003; 113(2):261–73. PMID: 12705873

54. Wang GR, Qiu YT, Lu T, Kwon HW, Pitts RJ, Van Loon JJA, et al. Anopheles gambiae TRPA1 is a heat-activated channel expressed in thermosensitive sensilla of female antennae. Eur J Neurosci. 2009; 30(6):967–74. https://doi.org/10.1111/j.1460-9568.2009.06901.x PMID: 19735290

55. López F, Delgado R, López R, Bacigalupo J, Restrepo D. Transduction for pheromones in the main olfactory epithelium is mediated by the Ca2+-activated channel TRPM5. J Neurosci. 2014; 34(9):3268–78. https://doi.org/10.1523/JNEUROSCI.4903-13.2014 PMID: 24573286

56. Akitake B, Ren Q, Boiko N, Ni J, Sokabe T, Stockand JD, et al. Coordination and fine motor control depend on Drosophila TRPγ. Nature communications. 2015; 6:7288. https://doi.org/10.1038/ncomms8288 PMID: 26028119

57. Szallasi A, Cortright DN, Blum CA, Eid SR. The vanilloid receptor TRPV1: 10 years from channel cloning to antagonist proof-of-concept. Nature reviews Drug discovery. 2007; 6(5):357. https://doi.org/10.1038/nrd2280 PMID: 17464295

II

BIOFOULING
https://doi.org/10.1080/08927014.2019.1602123

The complex barnacle perfume: identification of waterborne pheromone homologues in *Balanus improvisus* and their differential expression during settlement

Anna Abramova, Ulrika Lind, Anders Blomberg and Magnus Alm Rosenblad

Department of Chemistry and Molecular Biology, University of Gothenburg, Gothenburg, Sweden

ABSTRACT

A key question in barnacle biology is the nature of cues that induce gregarious settlement. One of the characterised cues is the waterborne settlement pheromone (WSP). This study aimed to identify WSP homologues in *Balanus improvisus* and to investigate their expression during settlement. Six WSP homologues were identified, all containing an N-terminal signal peptide, a conserved core region, and a variable C-terminus comprising several -GR- and -HDDH- motifs. The *B. improvisus* WSP homologues were expressed in all settlement stages but showed different expression patterns. The homologue most similar to the *B. amphitrite* WSP was the most abundant and was constantly expressed during settlement. In contrast, several of the other WSP homologues showed the greatest expression in the juvenile stage. The presence of several WSP homologues suggests the existence of a pheromone mix, where con-specificity might be determined by a combination of sequence characteristics and the concentration of the individual components.

ARTICLE HISTORY
Received 12 December 2018
Accepted 24 March 2019

KEYWORDS
Barnacles; settlement cues; waterborne settlement pheromone; Balanus improvisus; gregarious settlement

Introduction

Barnacles are among the most important fouling organisms in the marine environment. During settlement, the barnacle cyprid larvae respond to chemical and physical cues in the surrounding environment while swimming in the water column as well as when exploring various surfaces. The surface exploration comprises wide search, close search, and inspection behaviours that precede the settling decision leading to permanent adhesion through the release of cement (Crisp 1976; Aldred et al. 2018). Because adult barnacles are permanently attached, cyprid substratum selection is the main factor determining the local abundance and distribution of barnacles in natural populations as well as on marine constructions (eg ships' hulls), and is therefore a key question in barnacle biology and ecology (Thiyagarajan 2010). Understanding the mechanisms underlying cyprid attraction and substratum selection will facilitate the tailoring of antifouling technologies resulting in the development of surfaces that are actively rejected by cyprids during the initial stages of the surface exploration, thus preventing attachment.

Gregariousness in barnacles, ie the attraction of conspecific individuals and the subsequent settling and aggregation in dense communities, is crucial for reproduction and is achieved through cyprid settlement in response to chemical cues produced by individuals of the same species (Clare and Matsumura 2000). Since Knight-Jones and Stevenson (1950) described gregarious barnacle settlement behaviour for the first time almost 70 years ago, considerable advances have been made in the understanding of the mechanisms of this behaviour. Early research in gregarious settlement led to the discovery that extracts of whole adults contain a 'settling factor' that induces conspecific settlement of cyprids (Crisp and Meadows 1962). Later studies resulted in the identification of the contact pheromone in *Balanus amphitrite* (=*Amphibalanus amphitrite*), called settlement-inducing protein complex (SIPC) (Matsumura et al. 1998), the cDNA sequence of which was cloned and sequenced (Dreanno et al. 2006). It was shown that SIPC is a 171 kDa cuticular glycoprotein with similarity to the α2-macroglobulin protein family. SIPC is active when bound to a surface (Matsumura et al.

CONTACT Anders Blomberg anders.blomberg@cmb.gu.se

Supplemental data for this article can be accessed at https://doi.org/10.1080/08927014.2019.1602123.

1998), and cyprids can potentially detect the SIPC with their antennules while walking on the surface of, or close to, adult barnacles. Cyprids are able to differentiate between conspecific and allospecific SIPC (Dreanno et al. 2007). While the exact mechanism remains unknown, it has been suggested that either variable regions of the SIPC or glycosylation patterns might provide species specificity (Yorisue et al. 2012). Further research showed that SIPC has a dual role; it works as a settlement-inducing cue by attracting cyprids, but also as a settlement avoidance cue at higher concentrations informing of overcrowding and increased reproductive competition (Kotsiri et al. 2018). A SIPC homologue has also been identified and characterised in *Balanus glandula* (called MULTIFUNCin) where it induces both gregariousness as well as predation by sea snails (Ferrier et al. 2016; Zimmer et al. 2016).

Apart from the SIPC that induces settlement when bound to the surface, there is also evidence for a waterborne settling factor (Clare and Matsumura 2000). According to the current model, cyprids detect the waterborne cue while swimming and respond to it by transitioning from water to a substratum where the surface-bound SIPC pheromone in turn induces permanent attachment (Elbourne and Clare 2010). The first evidence of a waterborne cue came from an assay that used seawater conditioned with adults of *Semibalanus balanoides* showing that it contains a factor that induces temporary attachment of cyprids (Rittschof 1985). However, there is still ambiguity regarding the nature of the waterborne cue with several studies reporting varying estimates of the size of the active component of the conditioned water, ranging from peptides of 3–5 kDa to peptides of $<$ 500 Da (Clare and Matsumura 2000). In addition, several synthetic di- and tripeptides were tested for settlement-inducing activity, showing that peptides containing basic carboxy-terminal amino acids, in particular the glycyl-glycyl-arginine (GGR) peptide, had the tendency to enhance settlement of cyprids (Tegtmeyer and Rittschof 1988). However, this effect could not be reproduced by another study (Clare and Yamazaki 2000). Later, a protein corresponding to 32 kDa was purified from homogenized adult extracts of *B. amphitrite* and was shown to induce cyprid settlement. The protein diffuses into seawater when embedded in an agarose gel and induces settlement, suggesting a waterborne pheromone function (Endo et al. 2009). Interestingly, the N-terminal sequence obtained did not show any resemblance to SIPC or to any other proteins in the available databases. The full sequence of this *B. amphitrite* waterborne pheromone was published in sequence databases in 2012 and was called waterborne settlement pheromone (WSP): BAM34601.

Since the original publication of the discovery of WSP in 2009 (Endo et al. 2009), there have been no follow-up studies to further elucidate the structure and function of the WSP. However, it was recently indicated that more than one WSP gene is present in *Tetraclita japonica formosana* (Lin et al. 2014) and *B. amphitrite* (So et al. 2017; Wang et al. 2018), but the available information is extremely scarce and only one of the sequences has been published (Lin et al. 2014) and none are deposited in public databases. Moreover, the study by Endo et al. (2009) mentioned preliminary unpublished data showing that several proteins with a molecular mass of around 32 kDa detected in barnacle-conditioned seawater have settlement-inducing activity. Altogether, this indicated that there appear to be more than one WSP homologue in barnacles. From the ecological point of view, this opens the possibility that a combination of WSP homologues might work as a pheromone blend. Pheromone mixtures are commonly used for chemical communication in various animals, including marine invertebrates (Kelly 1996; Cummins et al. 2004). In particular, the sea slug *Aplysia* releases a pheromone blend comprising more than three different types of waterborne pheromones to attract mates (Cummins et al. 2005).

The indications of the presence of several WSP homologues in some barnacle species led the present authors to explore the presence of WSP homologues in *B. improvisus* using transcriptome data from cyprids and adults. Several WSP candidates in *B. improvisus* were identified, cloned, and compared to homologues present in *B. amphitrite*. In addition, gene expression analysis of the WSP homologues in *B. improvisus* during the settling process was performed showing that the genes are clearly differentially expressed. This study adds yet another dimension to the complexity of barnacle pheromones forming a basis for a better understanding of the gregarious behaviour as well as providing essential information for developing highly barnacle-selective antifouing strategies.

Materials and methods

Identification and cloning of WSP homologues in B. improvisus

Blast searches of several different RNA datasets (unpublished data) from cyprids and adults of

Table 1. Pairwise identity and similarity (in parenthesis) of the *B. amphitrite* WSP (BAM34601) and the identified *B. improvisus* WSP homologues.

	Ba_WSP	Bi_WSP	Bi_WSP-like 1	Bi_WSP-like 2	Bi_WSP-like 3	Bi_WSP-like 4	Bi_WSP-like 5
Ba_WSP	100%(100%)	74%(91%)	70%(88%)	66%(86%)	51%(78%)	51%(79%)	39%(71%)
Bi_WSP	74%(91%)	100%(100%)	80%(93%)	64%(84%)	55%(81%)	48%(79%)	38%(71%)
Bi_WSP-like 1	70%(88%)	80%(93%)	100%(100%)	59%(82%)	56%(80%)	50%(78%)	42%(72%)
Bi_WSP-like 2	66%(86%)	64%(84%)	59%(82%)	100%(100%)	41%(70%)	54%(82%)	38%(72%)
Bi_WSP-like 3	51%(78%)	55%(81%)	56%(80%)	41%(70%)	100%(100%)	42%(74%)	36%(69%)
Bi_WSP-like 4	51%(79%)	48%(79%)	50%(78%)	54%(82%)	42%(74%)	100%(100%)	40%(75%)
Bi_WSP-like 5	39%(71%)	38%(71%)	42%(72%)	38%(72%)	36%(69%)	40%(75%)	100%(100%)

B. improvisus were performed using *B. amphitrite* WSP (BAM34601) as a query. Six sequences that constituted full or partial fragments of WSP-like sequences were found, five of which were cloned and confirmed by Sanger sequencing. The sixth sequence was identified from a single molecule PacBio RNA dataset obtained from an adult of *B. improvisus* (Lind et al. 2017). The sequences identified were named according to the level of identity to the published *B. amphitrite* WSP, eg the Bi_WSP clone was 74% identical and 91% similar at the protein level to the *B. amphitrite* WSP (Table 1), while the WSP-like 1–5 sequences had a gradually lower (in the range of 39%–70% identity and 71%–88% similarity) level of identity/similarity that followed the gene numbering.

Cloning of the cDNA containing the complete open reading frame (ORF) for *Bi_WSP* and *Bi_WSP-like 2–5* was done by first performing RACE using PCR primers based on sequences found in a Sanger-sequenced cDNA library prepared from *B. improvisus* cyprids (unpublished data). RNA preparations of cyprids and cDNA synthesis were carried out as described in Lind et al. (2010). 3′-RACE was performed for all five pheromones (GeneRacer kit; Invitrogen), and 5′-RACE was performed for *Bi_WSP-like 3*. As a first step, a touch-down PCR was performed using the following PCR program: an initial denaturing step of 98 °C for 2 min followed by 5 cycles of 98 °C for 30 s, 72 °C for 1 min, 5 cycles of 98 °C for 30 s, 70 °C for 1 min, and 25 cycles of 98 °C for 30 s, 65 °C for 30 s, and 72 °C for 1 min. A final elongation step of 72 °C for 7 min was added. A nested PCR was thereafter performed using 1 µl of the touch-down reaction applying the following PCR program: an initial denaturing step of 98 °C for 2 min, followed by 35 cycles of 98 °C for 30 s, 65 °C for 30 s, and 72 °C for 1 min and a final elongation step of 72 °C for 7 min. The polymerase PfuUltra (Stratagene) was used in all PCR reactions. Sequences obtained from RACE were used to design primers to clone the complete ORF for the genes using the same PCR program as for the nested PCR, but with an annealing temperature of 60 °C. For the primers used for the cloning, see Supplemental material Table S1. *Bi_WSP-like* 1 was not cloned. Instead, a full-length sequence corresponding to *Bi_WSP-like 1* was identified from an adult single-molecule PacBio RNA dataset (Lind et al. 2017), where the corresponding consensus contig covering the ORF was supported by more than 30 PacBio reads. The sequences have been deposited in GenBank with accession numbers MK275628–MK275633. To find the sequences that are most divergent from the *B. improvisus* WSP homologues but still having the *B. amphitrite* WSP as the best match in NCBI nr, the identified *B. improvisus* WSP and WSP-like 1–5 protein sequences were used as queries for searching in the authors' different datasets. The search resulted in the identification of one additional sequence showing 30% identity or less to any of the queries, and this was named Bi_pb29993 according to the PacBio contig.

Searching for WSP homologues in other species

To determine whether barnacles other than *B. improvisus* also have several WSP homologues, the available larval and adult *B. amphitrite* NCBI/SRA datasets (SRX120025/SRR426836, SRX1035030/31, SRX1035107/27) were downloaded and transcriptome assemblies were created using Trinity 2.0.6 (Haas et al. 2013). These *B. amphitrite* assemblies were searched using the six identified *B. improvisus* WSP homologues. To avoid using wrongly assembled sequences, only *B. amphitrite* WSP-like candidates that were found in at least two assemblies were used for further sequence analysis. All candidate *B. amphitrite* ORFs were checked *vs* NCBI nr to make sure they were most similar to the published *B. amphitrite* WSP and not more similar to other types of proteins. *B. amphitrite* sequences that had the best match to the *B. amphitrite* WSP in NCBI nr searches were regarded as WSP-like candidates and named according to their contig numbers.

All identified protein sequences from *B. improvisus* and *B. amphitrite* were analysed using the InterPro (https://www.ebi.ac.uk/interpro/) and Pfam (https://pfam.xfam.org) domain databases, as well as with

SignalP and TargetP available at http://www.cbs.dtu.dk/services/. All six *B. improvisus* WSP homologues as well as Bi_pb29993 had weak matches (see Supplemental material Table S2 for the details) to the Pfam cupin_5 domain (PF06172), referred to as the RmlC-like cupin domain in InterPro. Sequences with a better match to the cupin_5 protein domain were also identified in both *B. improvisus* and *B. amphitrite* and were named Bi_cupin and Ba_cupin, respectively.

To further investigate whether crustaceans other than barnacles have proteins similar to the WSP homologues, an alignment of the six *B. improvisus* WSP homologues with the signal peptide and C-terminus removed was used to create an HMM profile with the HMMER 3.0 package (Potter et al. 2018). A set of all available 867,743 crustacean proteins was downloaded from the NCBI protein database (2018-05-02) and searched using the obtained WSP-specific HMM profile. The HMM search identified 27 sequences including the *B. amphitrite* WSP and 26 sequences from either *Daphnia* or *Eurytemora*. Among these 26 sequences, none had the *B. amphitrite* WSP as the best Blastp match in NCBI nr nor were they more similar to the *B. improvisus* WSP homologues than to the barnacle cupin_5 representatives in the present authors' dataset. To identify any other domains than cupin_5 that were present in the identified sequences, the protein sequences were searched against the collection of profiles in Pfam with hmmscan HMMER 3.0 (Potter et al. 2018). The search showed that there were no other domains than cupin_5 in the identified sequences, with scores well above the trusted Pfam cut-off (37.2–98.3).

Protein alignments

The *B. improvisus* and *B. amphitrite* WSP homologue protein sequences were aligned using Clustalw 2.0.10 and visualised with Jalview 2.10.1. The alignment revealed the presence of a well-aligned central region of ~200 aa with a few gaps (Figure 1b) that is referred to as the 'core region' throughout the text. The fasta36 package (Pearson 2016) was used to create the identity table for the core region of the identified *B. improvisus* WSP homologues and the *B. amphitrite* WSP (Table 1). The Zappo sequence colour scheme was used to show the physico-chemical properties of amino acids in the C-terminal region (Figure 2 and Figure S1B in the Supplemental material).

Phylogenetic trees

The *B. improvisus* and *B. amphitrite* WSP homologues identified were used for phylogenetic analysis. The *Daphnia* cupin_5, the closest homologue to the *B. improvisus* WSPs found among crustaceans, was used as an outgroup together with *B. improvisus* and *B. amphitrite* cupin_5 representatives.

The two methods used for the phylogenetic analysis were MrBayes 3.2.6 and PhyML 3.0 at phylogeny.fr. The ClustalW alignment of the identified *B. improvisus* and *B. amphitrite* WSP homologues and cupin-5 proteins (Figure S1A) was used as input into MrBayes 3.2.6 with one cold and three hot chains running for 100,000 generations at which point the average standard deviation was < 0.01. Alignments of full-length sequences or core regions only were used as input to PhyML.

Expression analysis of the B. improvisus waterborne pheromones identified

Transcriptomic data on settlement stages were used to investigate the expression of the *B. improvisus* WSP homologues identified, and the details of the experimental set up and the sampling of the different settling stages are described in Abramova et al., 2018 (manuscript). Briefly, cyprid larvae of *B. improvisus* were reared in a laboratory culture as described in Jonsson et al. (2018). For the expression analysis, 1–2-day-old cyprids were settled on Petri dishes (Nunc No 150340, Ø 48 mm) for 4–5 days. Four different settling stages were collected: free-swimming, close-search, attached cyprid, and juvenile. Each stage was represented by three biological replicates to ensure the generality of the findings (three independent batches of cyprids each coming from different sets of *B. improvisus* parents) with 20 individuals in each. Extracted RNA was used to make sequencing libraries with the Illumina TruSeq Stranded mRNA protocol following poly-A selection, and these were multiplexed and sequenced on three Illumina lanes. Library preparation and Illumina sequencing were performed at the national genomics facility at the Science for Life Laboratory in Stockholm (https://www.scilifelab.se/platforms/ngi/). Obtained RNA-seq data were assembled with the Trinity 2.0.6 software (Haas et al. 2013), and the previously identified *B. improvisus* WSP homologues were used to identify the corresponding contigs in the transcriptome assembly. The sequences of *Bi_WSP* and *Bi_WSP-like 2* were complete in the assembly, whereas the other WSP homologues were either found as partial contigs or were split between different Trinity contigs. To make a statistical analysis possible, all the fragmented WSP-like contigs were removed and substituted with the

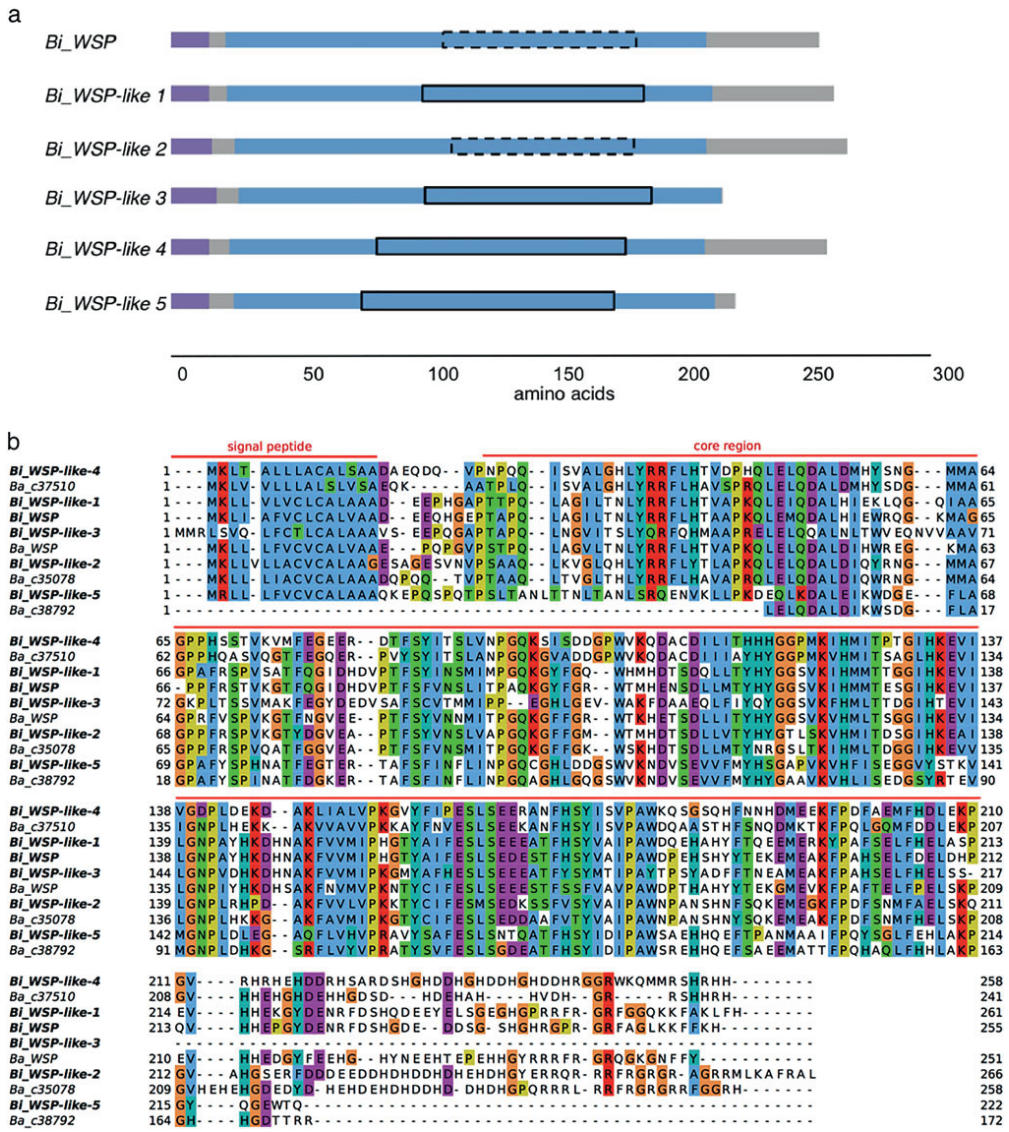


Figure 1. Structure and sequence analysis of the *B. improvisus* WSP homologues identified. (a) Structure of the *B. improvisus* WSP homologues showing the relative length and domain organisation. The signal peptide is in purple, the core region is in blue, and the cupin_5 domains are marked according to the Pfam domain score (Table S2). A solid line is used if the score is above the threshold of 25.0, and a dashed line is used if it is below the threshold. (b) Sequence alignment of the *B. improvisus* (in bold) and *B. amphitrite* WSP homologues showing conservation in the core region (marked) and high variability of the C-terminal region. The default Clustal X colour scheme in the Jalview program was used.

corresponding WSP-like sequences with complete ORFs that were previously identified. This edited assembly was used as the template for read mapping and to calculate the transcript abundances using the pseudoalignment-based quantification method kallisto (Bray et al. 2016) through the use of a Trinity 2.0.6 wrapper. The resulting abundances were used to generate a matrix of counts and a TMM (trimmed mean

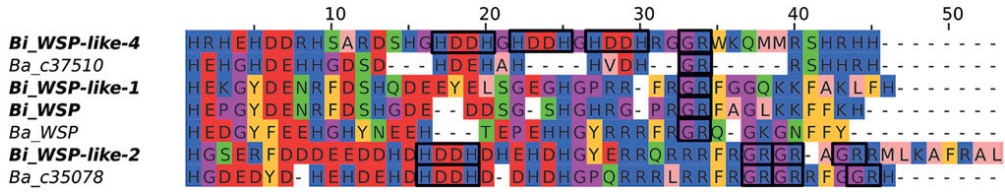


Figure 2. Characterisation of the C-terminus of the WSP homologues identified. Alignment of the C-terminus from the identified WSP homologues with the Zappo colouring scheme showing charged amino acids in red and blue. Examples of the -HDDH- and -GR- motifs are marked with black boxes (see Figure S1B in the Supplemental material for the full alignment).

Table 2. Description of the *B. improvisus* WSP homologues.

B. improvisus sequences	Accession number	Sequence type for ORF	Length (aa)	C-terminus	Best hit in NCBI nr Arthropoda
Bi_WSP	MK275628	Cloned	255	Long	BAM34601.1
Bi_WSP-like 1	MK275629	PacBio RNA	261	Long	BAM34601.1
Bi_WSP-like 2	MK275630	Cloned	266	Long	BAM34601.1
Bi_WSP-like 3	MK275631	Cloned	217	Short	BAM34601.1
Bi_WSP-like 4	MK275632	Cloned	258	Long	BAM34601.1
Bi_WSP-like 5	MK275633	Cloned	222	Short	BAM34601.1
Bi_pb29993	–	PacBio RNA	239	Short	BAM34601.1
Bi_cupin	–	Illumina RNA	232	Short	EFX90303.1

of M-values)-normalised matrix. Lowly-expressed genes were filtered out by keeping only genes with at least 2 CPM (corresponding to 60 counts on average) in at least two samples across the entire experiment. The preliminary data quality check indicated a pronounced batch effect reflecting the different parental backgrounds of the cyprid batches. To minimise the batch effect, the script removeBatchEffect from the edgeR 3.4.2 was applied to the count matrix (Robinson et al. 2010). The differential gene expression analysis was performed using the edgeR package. Genes with false discovery rate (FDR) 0.05 and a logFC $>$ 2 were considered to be significantly differentially expressed.

Results

Identification of WSP-like sequences in B. improvisus

To identify sequences in *B. improvisus* that are similar to the published WSP in *B. amphitrite*, tBlastn searches of RNA-seq datasets obtained from both cyprids and adults of *B. improvisus* were performed using the published *B. amphitrite* WSP as a query. Six unique *B. improvisus* RNA-contigs were found, of which five encoded partial WSP proteins. These sequences were used to obtain full-length variants of all WSP homologues in *B. improvisus*. The protein sequence identity between the six obtained *B. improvisus* homologues and the *B. amphitrite* WSP was 39%–74% (Table 1), with Bi_WSP being the most similar. The numbering of the WSP-like homologues in *B. improvisus* is related to their sequence identity to the Bi_WSP, with Bi_WSP-like 1 being the highest (80%). All *B. improvisus* protein sequences contained a predicted N-terminal signal peptide, a rather well-conserved central core region, and a highly variable C-terminal region (Figure 1). Four of the six identified *B. improvisus* WSP homologues were of almost the same length (255–266 aa) while Bi_WSP-like 3 and Bi_WSP-like 5 (217 and 222 aa, respectively) had a roughly 40 aa shorter C-terminus (Figure 1 and Table 2).

Except for the relatively high sequence identity to WSP from *B. amphitrite*, the WSP homologues in *B. improvisus* did not have any reliable matches to other proteins in the databases. However, all *B. improvisus* WSP homologues contained a region with weak resemblance to the cupin_5 domain (PF06172) (Table S2). Cupin_5 is a domain of unknown function, and proteins containing this domain belong to the cupin superfamily (http://www.ebi.ac.uk/interpro/entry/IPR009327). In the Pfam database, for metazoans there are currently only 17 sequences from 10 species indicated to have a cupin_5 domain. Among these sequences, the scores for the predicted cupin_5 domain varied from 41.3 to 139.2 with the only two crustacean representatives having scores of 70.5 or 78.7, both coming from *Daphnia* (see Table S3 in Supplemental material). The Pfam cupin_5 domain scores for the *B. improvisus* WSP homologues varied between 19.8 and 36.3, with some of them thus being below the Pfam cupin_5 profile threshold score of 25.0 (Table S2). However, proteins with higher cupin_5 Pfam scores in both

B. improvisus and *B. amphitrite* (with scores 45.4 and 46.0) were found, and these were named Bi_cupin and Ba_cupin, respectively. The Bi_cupin sequence was only 15% identical to Bi_WSP. Even if some of the WSP homologues in *B. improvisus* have a domain with weak similarity (slightly above the threshold score) to cupin_5, for the ease of the following discussion only proteins with high domain scores, ie Bi_cupin, Ba_cupin and *Daphnia* cupin_5, will be referred to as 'cupin_5 proteins' further in the text.

A search for more WSP-like sequences in *B. improvisus,* using the six initially identified *B. improvisus* WSP homologues as queries, resulted in the identification of one additional sequence showing a roughly 30% identity to the other WSPs and 15% identity to the Bi_cupin. This sequence was named Bi_pb29993 according to the name of the assembly contig. An alignment with the WSP-like sequences (Figure S1A) showed that the Bi_pb29993 protein was rather different from the six *B. improvisus* WSP homologues, but it still had the *B. amphitrite* WSP as the best match in the NCBI protein database.

Analysing the sequences of the *B. improvisus* WSP homologues showed that the long C-terminus found in most homologues was clearly distinct from the conserved core region by containing many charged residues (D/E, R/H/K) as well as glycines, with several -GR- and -HDDH- motifs or close variants of these (Figure 2). It should be noted that the frequent -HDDH- and -GR- motifs in the C-terminal regions of the identified WSP homologues vary greatly in their number, position, and composition, making it difficult to align this particular region of the sequences.

Additional WSP-like sequences in B. amphitrite

By using the identified WSP homologues from *B. improvisus* as queries for searching the published *B. amphitrite* RNA datasets, three more *B. amphitrite* WSP homologues in addition to the published Ba_WSP (Endo et al. 2009) were identified, of which two were encoding complete proteins (Figure 1b). Furthermore, two *B. amphitrite* sequences with the best match either to Bi_cupin or Bi_pb29993 were identified.

The alignment of the identified WSP homologues (Figure 1b) clearly shows that the *B. amphitrite* WSP-like sequences are structurally similar to the *B. improvisus* sequences, with an N-terminal signal peptide, a conserved core region, and the greatest difference found in the C-terminal region. The C-terminus of three of the *B. amphitrite* sequences, including the published Ba_WSP, had a prevalence of charged amino acids similar to what was found for the *B. improvisus* homologues Bi_WSP and Bi_WSP-like 2 and 4 (Figure 2 and Figure S1B in Supplemental material), while the fourth *B. amphitrite* sequence had a shorter C-terminus resembling *B. improvisus* WSP-like 3 and 5.

Phylogenetic analysis

Phylogenetic analyses showed that all *B. improvisus* WSP homologues were found in a well-supported core clade together with the *B. amphitrite* WSP, except for the Bi_WSP-like 5, which formed a sister group (Figure 3). The Bi_WSP-like 3 seems to be more closely related to Bi_WSP than to Bi_WSP-like 5, even though both Bi_WSP-like 3 and 5 lack the long, charged C-terminus (proteins with a long C-terminus are marked with an asterisk in Figure 3). The WSP homologues from *B. improvisus* and *B. amphitrite* were not divided into separate clades, thus suggesting that several WSP homologues were already present before these two species diverged. While the phylogenetic method used (either Bayesian or Maximum Likelihood) did not influence the overall topology of the phylogenetic tree, it did change the exact branching within the core clade indicating the difficulties in resolving the true topology in this part of the tree (Figure S2).

Overall, the phylogenetic analysis provided strong evidence that the *B. improvisus* WSP homologues identified are clearly more closely related to the *B. amphitrite* WSP than to the cupin_5 proteins (*Daphnia* cupin_5, Bi_cupin, and Ba_cupin).

Expression of B. improvisus *WSP candidates during settlement*

The WSP homologues identified from *B. improvisus* showed different expression patterns during settlement (Figure 4). The expression of *Bi_WSP* did not change significantly during settlement progression and was by far the highest, with roughly a hundred-fold higher expression compared to the other WSPs in the cyprid stages (Figure 4g). In contrast, *Bi_WSP-like 1* was overall the least-expressed candidate in any of the settling stages (note the different scales on the y-axes). *Bi_WSP-like 2*, *Bi_WSP-like 3*, and *Bi_WSP-like 5* showed similar expression patterns, being relatively lowly expressed in the cyprids and in the attachment stage and substantially upregulated when

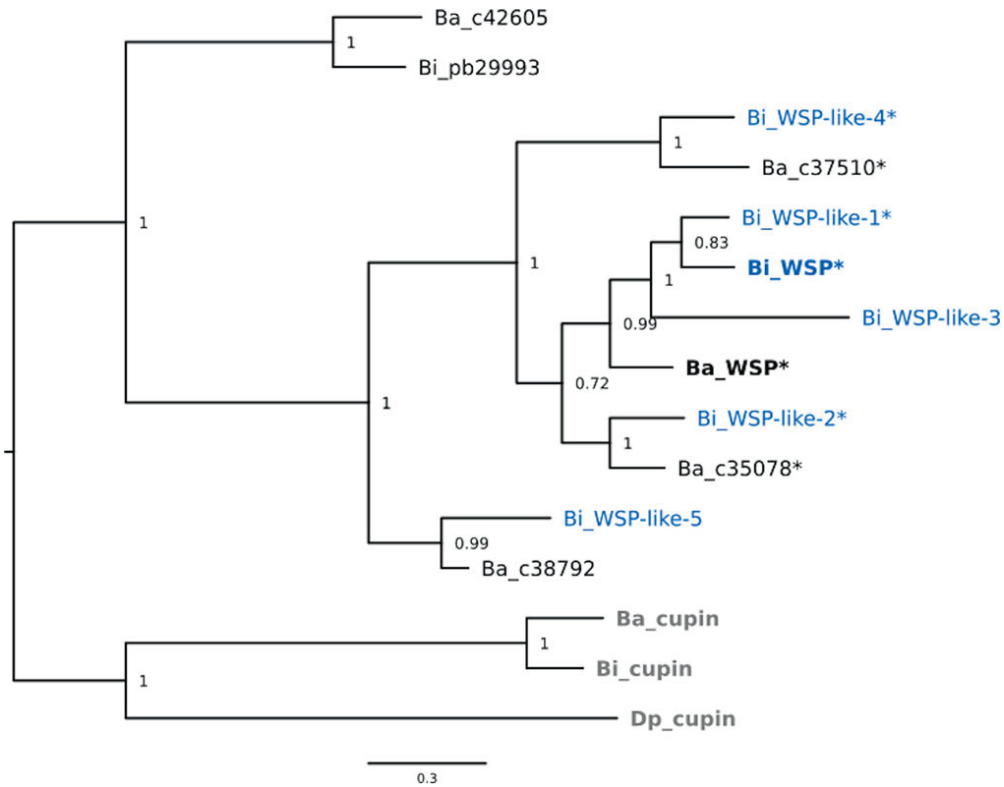


Figure 3. Bayesian phylogenetic tree of the WSP homologues in *B. improvisus* and *B. amphitrite* based on the alignment in Figure S1A. Cupin_5 domain proteins were added as an outgroup. All identified *B. improvisus* WSP homologues are marked in blue. Asterisks indicate WSP-like proteins with a long and charged C-terminus.

cyprids metamorphosed into juveniles; for example, *Bi_WSP-like 5* underwent about a 7-fold upregulation in the juvenile stage compared to the attached stage (Figure 4f). In summary, the metamorphosed juvenile stage of *B. improvisus* displayed substantial expression of all WSP homologues, with *Bi_WSP* having the highest expression.

Discussion

Since the initial publication of the discovery of WSP in 2009 (Endo et al. 2009), there have been several studies indicating that more than one WSP gene is present in some barnacle species (Lin et al. 2014; So et al. 2017; Wang et al. 2018). Originally, Endo et al. (2009) reported the results of a preliminary study that detected several proteins with weights around 32 kDa in *B. amphitrite*-conditioned seawater that exhibited settlement-inducing activity. In addition, several WSP sequences were found in the water-soluble fraction of proteins rinsed from the cement (So et al. 2017) and among the proteins from longitudinal canal and submantle tissues of *B. amphitrite* (Wang et al. 2018). In adults of *T. j. formosana,* Lin et al. (2014) reported the presence of two different WSP sequences displaying different expression in prosoma and basis, suggesting the existence of more than one WSP also in this species. However, only one of the sequences from these studies was published. The indications of the presence of several WSP homologues in some barnacle species led the present authors to explore the presence of WSP homologues in *B. improvisus* using transcriptome data from cyprids and adults. Six sequences that are homologous to the published *B. amphitrite* WSP were identified from *B. improvisus*. The similar size and isoelectric point observed for several of the WSP homologues identified here (data not shown) suggest that in the previous studies based

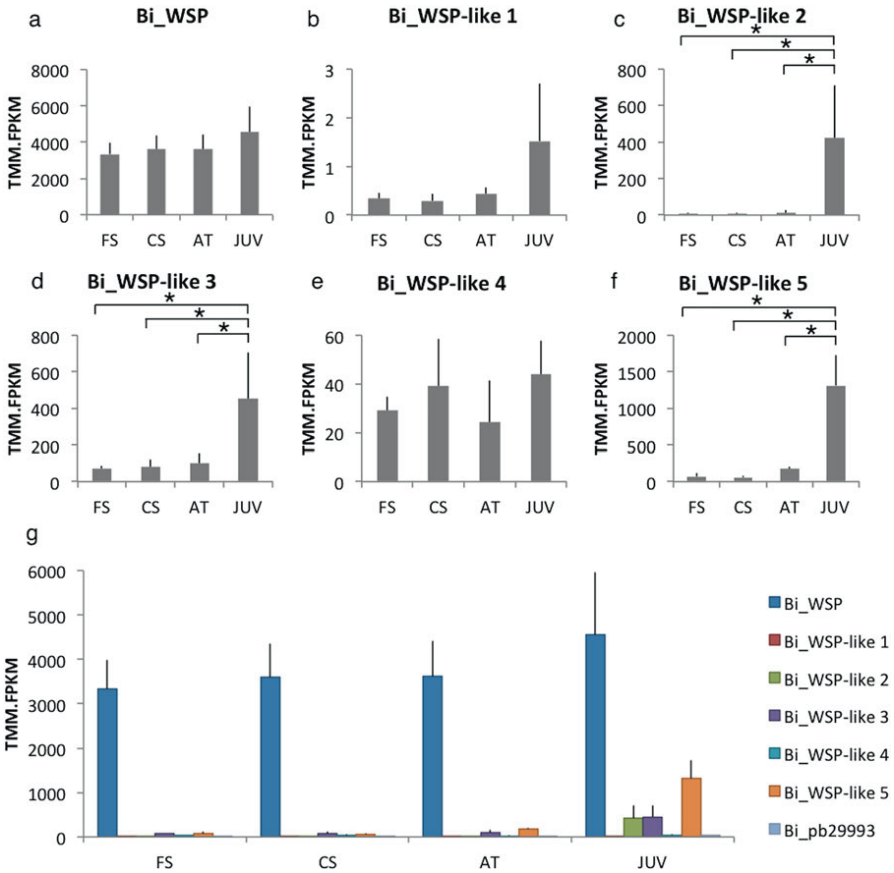


Figure 4. Gene expression of WSP homologues and Bi_pb29993 during settlement. (a–f) The relative expression of the WSP homologues; (g) The relative expression of the WSP homologues and Bi_pb29993. The order of the bars in the graph corresponds to the order of the names in the legend. FS: free-swimming, CS: close-search, AT: attached cyprid, JUV: juvenile.

on protein purifications (Endo et al. 2009; So et al. 2017) a single band on protein gels corresponding to 32 kDa could theoretically represent co-purification of several different proteins.

Phylogenetic analysis clearly showed that five of the six identified WSP sequences from *B. improvisus* proteins belong to the same clade as WSP homologues from *B. amphitrite*, suggesting that several WSPs were already present before the two species diverged. Despite some of the sequences appearing to form orthologous groups in the phylogenetic analysis (Figure 3), it is not presently possible to infer the true orthologous relationships between the WSP-like candidates from *B. improvisus* and *B. amphitrite* due to the lack of complete genome-based proteomes for any of these barnacle species. Furthermore, it should be pointed out that it is not yet possible to establish the true number of WSP family homologues in *B. improvisus* and *B. amphitrite* using only transcriptome data because some of the WSP genes might be expressed at extremely low levels at the life stages and tissues examined and thus not be detected. The phylogenetic analysis presented here should therefore be seen as a first step to aid in the classification of WSP-like sequences. Although an evolutionary relationship is not evidence of similar function, it is clear that the identification of several WSP-like sequences in these species suggests that WSP is a substantially larger protein family than previously thought and that multiple WSP paralogues also exist in other barnacle species.

There is currently no evidence for the presence of WSP homologues in any crustaceans other than

barnacles, suggesting that the WSPs are a relatively late evolutionary invention and thus a protein family that might be unique to barnacles. However, part of the core region of the WSP homologues has some similarity to the cupin_5 domain found in both prokaryotes and eukaryotes. Nevertheless, the sequence and phylogenetic analyses indicated that the WSPs are clearly distinct from other cupin_5 domain proteins found in *B. improvisus*, *B. amphitrite*, and *Daphnia*. Moreover, identification of the cupin_5 domain does not provide any information about function because the cupin_5 family is poorly characterised (Zhou et al. 2005; Gaowa and Zhang 2009; Du et al. 2010).

Analysis of the sequence-features of the identified WSP homologues revealed that all have quite conserved N-terminal signal peptides and core regions, with the greatest variability residing in the C-terminal regions. The importance of the C-terminus for water-soluble pheromones was noticed many years ago because digestion with carboxypeptidases that cleave carboxyl-terminal lysine and arginine residues eliminated pheromone activity (Tegtmeyer and Rittschof 1988). Furthermore, the C-terminus has a striking number of charged amino acids compared to the core region (Figure 2 and Figure S1B). A charged, basic C-terminus has previously been suggested to be present in a small barnacle waterborne cue of 3-5 kDa ($\approx$ 40 aa) (Tegtmeyer and Rittschof 1988). Interestingly, the sizes of the peptides that showed potent induction of settlement in the study by Tegtmeyer and Rittschof (1988) correspond to the average size of the C-termini in the WSP homologues reported here. This opens up the possibility that the short peptides reported (Rittschof 1985; Tegtmeyer and Rittschof 1988; Clare and Matsumura 2000) with settlement-inducing activity are derived from cleavage of the WSPs, releasing biologically active peptides from the C-termini as earlier proposed (Rittschof 1993). A similar mechanism was found in lobsters and hermit crabs, where olfactory serine proteases have been suggested to cleave after arginine and lysine residues in inactive odorant peptide molecules to generate active chemical cues (Rittschof 1990; Levine et al. 2001). Notably, the charged amino acids in the C-terminus of some of the identified WSP homologues are organised in repeated motifs. The most commonly observed were -HDDH- and -GR- motifs (Figure 2), but some of the WSP homologues also contained variations of these motifs, eg -HDEH-, -HDH-, and -GGR-. Interestingly, the latter motif was previously reported to be important for settlement-inducing activity in both barnacle and oyster larvae (Zimme-Faust and Tamburri 1994; Browne and Zimmer 2001).

In comparison to SIPC, WSP is largely under-investigated. An important piece of missing evidence is whether WSP is actually released into the water by adult organisms in nature and functions as a waterborne cue, as earlier suggested (Endo et al. 2009). Further experimental studies are required to establish whether the multiple WSP homologues identified here in both *B. amphitrite* and *B. improvisus* act as pheromones. However, assuming that the WSP homologues exhibit pheromone activity, the current study brings a new level of complexity to the mechanisms of barnacle chemical communication and barnacle settlement ecology. The presence of several WSP-like candidates in barnacles is consistent with previously reported data on pheromones in some insects, nematodes, and marine invertebrates that consist of a mix of several components. It has been suggested that the specificity of these pheromones is determined by the relative concentrations of the different components in the mixture (Linn et al. 1984; Roelofs 1995; Cummins et al. 2004). Moreover, evidence exists that the concentration of the different pheromone components depends on age, sex, and habitat, thus providing additional information to the individuals receiving the pheromone signal (Coppée et al. 2011; Choe et al. 2012). One particularly well-studied example of pheromone mixtures among marine invertebrates is the pheromone blend in the sea slug *Aplysia* (Cummins et al. 2005). Mate attraction in these animals involves several waterborne pheromones such as attractin, seductin, and temptin, that in different combinations act to maintain mating aggregations (Cummins et al. 2005). In barnacles, according to the current hypothesis, cyprids detect the waterborne cue released by adults while swimming and respond to it by transitioning from the water column to a substratum where the surface-bound SIPC pheromone in turn induces permanent attachment (Elbourne and Clare 2010).

Details on how the waterborne cue is released, distributed, and sensed by cyprids remain unknown. However, studies on host location by cyprids of the parasitic barnacle *Heterosaccus dollfusi* (Pasternak et al. 2004) showed that larvae use chemoreception to initiate the motion and rheoreception (sensing direction of the flow) to follow the host's odour plume upstream until locating the organism releasing the odour. This study also suggested that larvae do not use the concentration gradient of the attractant to find the direction but rather its presence or absence in the water flow. In the context of a pheromone mixture, it is therefore

probable that cyprids' first encounter the most abundant WSP that travels the longest distance from where it was released. Following the odour plume, cyprids will sense the other more lowly expressed WSP homologues that would serve as a confirmation of the direction towards the odour source as well as possibly trigger downward movement thus bringing cyprids closer to the surface where the cue is released by already established adults.

The relative abundance of the WSPs in combination with the sequence differences between species might be the basis for the species specificity of the cue. However, no obvious and consistent amino acid differences between the WSPs from *B. amphitrite* and *B. improvisus* were observed in the core region. Instead, the C-terminus with its intriguing repetitions of acidic and basic residues might contain species-unique combinations of motifs. Although there is some evidence that glycosylation patterns might be involved in the species specificity of the SIPC protein (Yorisue et al. 2012), nothing currently indicates that this type of post-translational modification plays a role in the activity of WSP. On the contrary, Endo et al. (2009) found that lentil lectin treatment (which binds to and inhibits sugars on proteins) did not influence the pheromone activity of their WSP fraction. On the other hand, the waterborne cue might not necessarily be species specific as has been shown for *Aplysia* attractins, which are the only peptide pheromone family in invertebrates known not to be species specific (Painter et al. 2003). Behavioural studies on *Aplysia* also showed that fewer individuals were attracted by a single component than by a pheromone blend, suggesting that pheromones act in concert (Painter et al. 2003). The present data show that different barnacle life stages express different combinations of WSPs. In particular, the cyprids expressed mostly *Bi_WSP*, whereas juveniles expressed all six identified WSP homologues. It should be pointed out that newly metamorphosed juvenile barnacles were found to be particularly attractive to cyprids (Crisp and Meadows 1962). Therefore, the blend of several WSPs produced by juveniles might signal to cyprids that it is a suitable place to successfully settle and metamorphose.

Except for the reported pheromone function of the published *B. amphitrite* WSP (Endo et al. 2009), neither pheromone activity nor any other function for the WSP homologues identified here from *B. amphitrite* or *B. improvisus* have been experimentally examined. In addition to inducing cyprid settlement, the WSP homologues might have completely different or dual functions. For instance, the identification of WSP among the proteins rinsed from cement collected from adult *B. amphitrite* (So et al. 2017) suggests that WSPs might be a part of the permanent adhesive from where they leak into the water and induce cyprid settlement. A patent has been issued for an adhesive 35 kDa protein in *B. improvisus*, *B. crenatus*, and *S. balanoides* extracted from the hemolymph or base plate (Kaplan et al. 2003) that contains an N-terminal fragment of about 20 amino acids that exactly matches the Bi_WSP downstream of the signal peptide region (starting at aa number 16). Furthermore, the *Semibalanus* sequence (File S1), which is a protein of 245 aa and found both in the hemolymph and the base plate, is clearly a WSP homologue with 67%, 65%, and 62% identity to Bi_WSP-like 2, Ba_WSP, and Bi_WSP, respectively. Additionally, the triad -HDDH- motif in the C-terminus of this 35 kDa *Semibalanus* sequence (mentioned in the same patent) is found in Bi_WSP-like 4 with a G separating the three -HDDH- motifs, whereas Bi_WSP-like 2 contains three repetitions of the alternative -HDH- motif. It should be noted that a repeated -HDH- motif in the C-terminus is also found in two of the *B. amphitrite* WSP sequences (Figure 2). It is interesting that these motifs consistently occur in all the barnacle species mentioned, albeit sometimes with small differences, especially because the C-terminus is the least conserved part of these WSP homologues. The presence and particular combinations of motifs might indicate differences in function for the different WSP homologues, whether acting as pheromones or as a part of the cement.

In the current study, several WSP homologues in *B. improvisus* were identified and characterised. Further experimental studies are needed to determine the function of all of the WSP homologues, especially what features of WSPs are behind the settling-inducing effect and whether there is a functional connection to the properties of the adhesives. Although several WSP-like sequences exist in both *B. improvisus* and *B. amphitrite*, a complete set of genes encoding WSP-like proteins in any barnacle species can only be obtained using a high-quality genome as a reference. Despite that, the identification of a set of WSP homologues is an important discovery that will first of all promote further research into the great complexity of barnacle chemical communication. Furthermore, it is clear from the present study that the WSP homologues seem to be barnacle-specific and thus might be used as targets for non-toxic antifouling agents analogous to pheromone-based applications in insect management

strategies (Witzgall et al. 2010). Finally, the protein sequences of the identified WSP homologues are now available in the NCBI database, and this will aid in the annotation of WSPs in future studies.

Acknowledgements

The authors would like to acknowledge support from the SciLifeLab National Genomics Infrastructure (NGI) and Uppmax for providing assistance in massive parallel sequencing and computational infrastructure, respectively. The authors would also like to thank Martin Ogemark for maintaining the *B. improvisus* culture and providing cyprid larvae, Johan Lagmo for technical assistance during the initial phase of the project in cloning some of the WSP-like genes from *B. improvisus*, and Per Jonsson for valuable comments on the manuscript.

Declaration of interest statement

No potential conflicts of interest were reported by the authors.

Funding

Funding for this project was provided to AB by the Swedish Research Council (VR, #2017-04559), the EU Commission (SEAFRONT—Synergistic Fouling Control Technologies; #614034) and Swedish Research Council for Sustainable Development (Formas; #210-2011-1733), as well as to MAR from Carl Tryggers Foundation (CTS11:14). This work was supported by the Centre for Marine Evolutionary Biology at the University of Gothenburg (http://www.cemeb.science.gu.se/). MAR was partially funded by National Bioinformatics Infrastructure Sweden (NBIS).

ORCID

Anna Abramova http://orcid.org/0000-0002-0493-7808
Ulrika Lind http://orcid.org/0000-0003-1746-9043
Anders Blomberg http://orcid.org/0000-0002-1260-3920
Magnus Alm Rosenblad http://orcid.org/0000-0002-4077-7821

References

Aldred N, Alsaab A, Clare AS. 2018. Quantitative analysis of the complete larval settlement process confirms Crisp's model of surface selectivity by barnacles. Proc R Soc B. 285:20171957. doi:10.1098/rspb.2017.1957

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 34:525–527. doi:10.1038/nbt.3519

Browne KA, Zimmer RK. 2001. Controlled field release of a waterborne chemical signal stimulates planktonic larvae to settle. Biol Bull. 200:87–91. doi:10.2307/1543088

Choe A, Chuman T, von Reuss SH, Dossey AT, Yim JJ, Ajredini R, Kolawa AA, Kaplan F, Alborn HT, Teal PEA et al. 2012. Sex-specific mating pheromones in the nematode *Panagrellus redivivus*. P Natl Acad Sci USA. 109: 20949–20954. doi:10.1073/pnas.1218302109

Clare AS, Matsumura K. 2000. Nature and perception of barnacle settlement pheromones. Biofouling. 15:57–71. doi:10.1080/08927010009386298

Clare AS, Yamazaki M. 2000. Inactivity of glycyl-glycyl-arginine and two putative (QSAR) peptide analogues of barnacle waterborne settlement pheromone. J Mar Biol Assoc UK. 80:945–946. doi:10.1017/S0025315400002952

Coppée A, Mathy T, Cammaerts M-C, Verheggen FJ, Terzo M, Iserbyt S, Valterová I, Rasmont P. 2011. Age-dependent attractivity of males' sexual pheromones in *Bombus terrestris* (L.)[Hymenoptera, Apidae]. Chemoecology. 21: 75–82. doi:10.1007/s00049-011-0070-x

Crisp D. 1976. Settlement responses in marine organisms. Butterworths, London.

Crisp DJ, Meadows P. 1962. The chemical basis of gregariousness in cirripedes. Proc R Soc Lond B Biol Sci. 156:500–520.

Cummins SF, Nichols AE, Amare A, Hummon AB, Sweedler JV, Nagle GT. 2004. Characterization of *Aplysia* enticin and temptin, two novel water-borne protein pheromones that act in concert with attractin to stimulate mate attraction. J Biol Chem. 279:25614–25622. doi:10.1074/jbc.M313585200

Cummins SF, Nichols AE, Warso CJ, Nagle GT. 2005. *Aplysia* seductin is a water-borne protein pheromone that acts in concert with attractin to stimulate mate attraction. Peptides. 26:351–359. doi:10.1016/j.peptides.2004.10.024

Dreanno C, Kirby RR, Clare AS. 2007. Involvement of the barnacle settlement-inducing protein complex (SIPC) in species recognition at settlement. J Exp Mar Biol Ecol. 351:276–282. doi:10.1016/j.jembe.2007.07.003

Dreanno C, Matsumura K, Dohmae N, Takio K, Hirota H, Kirby RR, Clare AS. 2006. An alpha2-macroglobulin-like protein is the cue to gregarious settlement of the barnacle *Balanus amphitrite*. Proc Natl Acad Sci U S A. 103: 14396–14401. doi:10.1073/pnas.0602763103

Du Y, He YX, Gaowa S, Zhang X, Chen Y, Zhang SC, Zhou CZ. 2010. Crystal structures of the apo and GDP-bound forms of a cupin-like protein BbDUF985 from *Branchiostoma belcheri tsingtauense*. Proteins. 78: 2714–2719.

Elbourne PD, Clare AS. 2010. Ecological relevance of a conspecific, waterborne settlement cue in *Balanus amphitrite* (Cirripedia). J Exp Mar Biol Ecol. 3922:99–106.

Endo N, Nogata Y, Yoshimura E, Matsumura K. 2009. Purification and partial amino acid sequence analysis of the larval settlement-inducing pheromone from adult extracts of the barnacle, *Balanus amphitrite* (=*Amphibalanus amphitrite*). Biofouling. 25:429–434. doi:10.1080/08927010902875113

Ferrier GA, Kim SJ, Kaddis CS, Loo JA, Ann Zimmer C, Zimmer RK. 2016. MULTIFUNCin: a multifunctional protein cue induces habitat selection by, and predation on, barnacles. Integr Comp Biol. 56:901–13. doi:10.1093/icb/icw076

Gaowa S, Zhang S. 2009. Identification, expression, function and localization of a DUF985 domain-containing hypothetical gene from amphioxus *Branchiostoma belcheri*. Comp Biochem Physiol B Biochem Mol Biol. 152:28–37. doi:10.1016/j.cbpb.2008.09.085

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M

et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 8:1494–1512. doi:10.1038/nprot.2013.084

Jonsson PR, Wrange AL, Lind U, Abramova A, Ogemark M, Blomberg A. 2018. The barnacle *Balanus improvisus* as a marine model - culturing and gene expression. Jove-J Vis Exp.

Kaplan D, Gatenholm P, Berglin K, Platko J, Pepper L, Ngangan A, inventors; Kaplan, D, assignee. 2003 30.04.2002 Barnacle adhesion proteins. United States patent No. 10/134,507.

Kelly DR. 1996. When is a butterfly like an elephant? Chem Biol. 3:595–602. doi:10.1016/S1074-5521(96)90125-8

Knight-Jones E, Stevenson J. 1950. Gregariousness during settlement in the barnacle *Elminius modestus* Darwin. J Mar Biol Assoc UK. 29:281–297. doi:10.1017/S0025315400055375

Kotsiri M, Protopapa M, Mouratidis S, Zachariadis M, Vassilakos D, Kleidas I, Samiotaki M, Dedos SG. 2018. Should I stay or should I go? The settlement-inducing protein complex guides barnacle settlement decisions. J Exp Biol. 221:jeb185348. doi:10.1242/jeb.185348

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. doi:10.1002/neu.10010

Lin HC, Wong YH, Tsang LM, Chu KH, Qian PY, Chan BKK. 2014. First study on gene expression of cement proteins and potential adhesion-related genes of a membranous-based barnacle as revealed from Next-Generation Sequencing technology. Biofouling. 30: 169–181. doi:10.1080/08927014.2013.853051

Lind U, Jarva M, Rosenblad MA, Pingitore P, Karlsson E, Wrange AL, Kamdal E, Sundell K, Andre C, Jonsson PR et al. 2017. Analysis of aquaporins from the euryhaline barnacle *Balanus improvisus* reveals differential expression in response to changes in salinity. Plos One. 12.

Lind U, Rosenblad MA, Frank LH, Falkbring S, Brive L, Laurila JM, Pohjanoksa K, Vuorenpää A, Kukkonen JP, Gunnarsson L. 2010. Octopamine receptors from the barnacle *Balanus improvisus* are activated by the α2-adrenoceptor agonist medetomidine. Mol Pharmacol. 78: 237–248. doi:10.1124/mol.110.063594

Linn C, Bjostad L, Du J, Roelofs W. 1984. Redundancy in a chemical signal: Behavioral responses of male *Trichoplusia ni* to a 6-component sex pheromone blend. J Chem Ecol. 10:1635–1658. doi:10.1007/BF00988431

Matsumura K, Nagano M, Fusetani N. 1998. Purification of a larval settlement-inducing protein complex (SIPC) of the barnacle, *Balanus amphitrite*. J Exp Zool. 281:12–20. doi:10.1002/(SICI)1097-010X(19980501)281:1<12::AID-JEZ3>3.0.CO;2-F

Painter SD, Clough B, Black S, Nagle GT. 2003. Behavioral characterization of attractin, a water-borne peptide pheromone in the genus *Aplysia*. Biol Bull. 205:16–25. doi:10.2307/1543441

Pasternak Z, Blasius B, Abelson A. 2004. Host location by larvae of a parasitic barnacle: larval chemotaxis and plume tracking in flow. J Plankton Res. 26:487–493. doi:10.1093/plankt/fbh040

Pearson WR. 2016. Finding protein and nucleotide similarities with FASTA. Curr Protoc Bioinformatics. 53:3–9.

Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. 2018. HMMER web server: 2018 update. Nucleic Acids Res. 46:200–204.

Rittschof D. 1985. Oyster drills and the frontiers of chemical ecology: unsettling ideas. Bull Am Mar Union. 1: 111–116.

Rittschof D. 1990. Peptide-mediated behaviors in marine organisms Evidence for a common theme. J Chem Ecol. 16:261–272. doi:10.1007/BF01021283

Rittschof D. 1993. Body odors and neutral-basic peptide mimics: a review of responses by marine organisms. Am Zool. 33:487–493.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 26: 139–140. doi:10.1093/bioinformatics/btp616

Roelofs WL. 1995. Chemistry of sex attraction. Proc Natl Acad Sci U S A. 92:44–49. doi:10.1073/pnas.92.1.44

So CR, Scancella JM, Fears KP, Essock-Burns T, Haynes SE, Leary DH, Diana Z, Wang CY, North S, Oh CS et al. 2017. Oxidase activity of the barnacle adhesive interface involves peroxide-dependent catechol oxidase and lysyl oxidase enzymes. Acs Appl Mater Inter. 9:11493–11505. doi:10.1021/acsami.7b01185

Tegtmeyer K, Rittschof D. 1988. Synthetic peptide analogs to barnacle settlement pheromone. Peptides. 9: 1403–1406. doi:10.1016/0196-9781(88)90209-4

Thiyagarajan V. 2010. A review on the role of chemical cues in habitat selection by barnacles: new insights from larval proteomics. J Exp Mar Biol Ecol. 392:22–36. doi:10.1016/j.jembe.2010.04.030

Wang C, Schultzhaus JN, Taitt CR, Leary DH, Shriver-Lake LC, Snellings D, Sturiale S, North SH, Orihuela B, Rittschof D. 2018. Characterization of longitudinal canal tissue in the acorn barnacle *Amphibalanus amphitrite*. Plos One. 13:e0208352. doi:10.1371/journal.pone.0208352

Witzgall P, Kirsch P, Cork A. 2010. Sex pheromones and their impact on pest management. J Chem Ecol. 36: 80–100. doi:10.1007/s10886-009-9737-y

Yorisue T, Matsumura K, Hirota H, Dohmae N, Kojima S. 2012. Possible molecular mechanisms of species recognition by barnacle larvae inferred from multi-specific sequencing analysis of proteinaceous settlement-inducing pheromone. Biofouling. 28:605–611. doi:10.1080/08927014.2012.695776

Zhou CZ, Meyer P, Quevillon-Cheruel S, De La Sierra-Gallay IL, Collinet B, Graille M, Blondeau K, François JM, Leulliot N, Sorel I. 2005. Crystal structure of the YML079w protein from *Saccharomyces cerevisiae* reveals a new sequence family of the jelly-roll fold. Protein Sci. 14:209–215.

Zimme-Faust RK, Tamburri MN. 1994. Chemical identity and ecological implications of a waterborne, larval settlement cue. Limnol Oceanogr. 39:1075–1087. doi:10.4319/lo.1994.39.5.1075

Zimmer RK, Ferrier GA, Kim SJ, Kaddis CS, Zimmer CA, Loo JA. 2016. A multifunctional chemical cue drives opposing demographic processes and structures ecological communities. Ecology. 97:2232–2239. doi:10.1002/ecy.1455

III

Video Article

The Barnacle *Balanus improvisus* as a Marine Model - Culturing and Gene Expression

Per R. Jonsson[1], Anna-Lisa Wrange[3], Ulrika Lind[2], Anna Abramova[2], Martin Ogemark[1], Anders Blomberg[2]

[1]Department of Marine Sciences, University of Gothenburg

[2]Department of Chemistry and Molecular Biology, University of Gothenburg

[3]IVL Swedish Environmental Research Institute

Correspondence to: Anders Blomberg at anders.blomberg@cmb.gu.se

URL: https://www.jove.com/video/57825
DOI: doi:10.3791/57825

Keywords: Environmental Sciences, Issue 138, Barnacle, Crustacean, *Balanus (Amphibalanus) improvisus*, culture, Nauplii, Cyprids, tissue dissection, RNA extraction, quantitative PCR, gene expression

Date Published: 8/8/2018

Citation: Jonsson, P.R., Wrange, A.L., Lind, U., Abramova, A., Ogemark, M., Blomberg, A. The Barnacle *Balanus improvisus* as a Marine Model - Culturing and Gene Expression. *J. Vis. Exp.* (138), e57825, doi:10.3791/57825 (2018).

Abstract

Barnacles are marine crustaceans with a sessile adult and free-swimming, planktonic larvae. The barnacle *Balanus (Amphibalanus) improvisus* is particularly relevant as a model for the studies of osmoregulatory mechanisms because of its extreme tolerance to low salinity. It is also widely used as a model of settling biology, in particular in relation to antifouling research. However, natural seasonal spawning yields an unpredictable supply of cyprid larvae for studies. A protocol for the all-year-round culturing of *B. improvisus* has been developed and a detailed description of all steps in the production line is outlined (*i.e.*, the establishment of adult cultures on panels, the collection and rearing of barnacle larvae, and the administration of feed for adults and larvae). The description also provides guidance on troubleshooting and discusses critical parameters (*e.g.*, the removal of contamination, the production of high-quality feed, the manpower needed, and the importance of high-quality seawater). Each batch from the culturing system maximally yields roughly 12,000 nauplii and can deliver four batches in a week, so up to almost 50,000 larvae per week can be produced. The method used to culture *B. improvisus* is, probably, to a large extent also applicable to other marine invertebrates with free-swimminglarvae. Protocols are presented for the dissection of various tissues from adults as well as the production of high-quality RNA for studies on gene expression. It is also described how cultured adults and reared cyprids can be utilized in a wide array of experimental designs for examining gene expression in relation to external factors. The use of cultured barnacles in gene expression is illustrated with studies of possible osmoregulatory roles of Na^+/K^+ ATPase and aquaporins.

Video Link

The video component of this article can be found at https://www.jove.com/video/57825/

Introduction

Barnacles are marine crustaceans with a sessile adult and free-swimming, planktonic larvae. Most of the 1,200 species of barnacles inhabit shallow water and many are often exposed to low salinity. One species, the bay barnacle *Balanus (Amphibalanus) improvises* (*B. improvisus*), can tolerate almost freshwater and Charles Darwin described this species from a small stream in the estuary of Rio de la Plata in Uruguay[1]. The extreme tolerance tolow salinity makes *B. improvisus* a particularly relevant model for the studies of osmoregulatory mechanisms[2,3]. This barnacle prefers brackish conditions but is capable of living in waters with salinities from around 1.6 psu to as high as 40 psu[4]. It is the only barnacle species found in the brackish Baltic Sea. *B. improvisus* is believed to originate from the east coast of the American continent but today is found worldwide due to dispersal by shipping[5]. It is a major fouling organism and is commonly found on rocks, jetties, and boat hulls, and is therefore of general interest for understanding the mechanisms of biofouling on constructions in marine and brackish waters[6,7].

Similar to most other barnacles, *B. improvisus* is hermaphroditic with cross-fertilization; reproduction occurs through mating between neighboring individuals using an elongated penis and internal fertilization. The reproductive period is mainly from May to September. *B. improvisus* has seven pelagic larval stages (six nauplii followed by one cyprid stage[8]). The fertilized egg hatches into a nauplius larva which is free-swimming and feeds in the water column for up to several weeks before molting into a non-feeding cyprid larva. The cyprid uses multiple cues to find a suitable site to settle and then undergoes metamorphosis into a sessile juvenile barnacle[9]. The species can be cultured in the laboratory and has a lifespan of 1–2 years in the sea (2–3 years in laboratory culture). On average, *B. improvisus* grows to 10 mm in diameter (with a maximum of around 20 mm) and reaches a maximum height of about 6 mm (although it can grow taller under crowded conditions). The species can be identified by its smooth calcareous even shell (white or greyish), the radially patterned calcareous base of the shell plate, and the shape of the tergal plates[1,10].

The barnacle *B. improvisus* has several advantageous features as a model for studies of osmoregulation, with a focus on molecular and physiological mechanisms as well as ecological interactions and evolutionary consequences. It is also widely used as a model for the investigations of settling biology, in particular in relation to antifouling research and the mechanisms involved[7,11,12,13]. However, natural seasonal spawning yields an unpredictable supply of cyprid larvae for studies. The ability to culture this barnacle through its whole lifecycle all year round

is, therefore, a major asset to enable various types of molecular and mechanistic studies. In addition, its presence in marine/brackish waters worldwide allows for a combination of field and experimental studies. Controlled breeding can also produce families of known pedigrees for long-term culturing[14], and a generation time of a few months may allow long-term experimental evolution. There is also a draft genome and several transcriptomes available, and these resources have been used for the cloning of several genes (*e.g.*, genes of importance in osmoregulation)[2,3].

The aim of this protocol is to describe how to establish and maintain a culture of the barnacle *B. improvisus* throughout the year in order to perform gene expression studies on adults or larvae of this organism. Rittschof *et al.*[15] briefly described a method for culturing barnacles from the release of nauplii to the settlement of cyprids for the species *Balanus amphitrite*. The protocol has been adapted for the all-year-round culturing of *B. improvisus* at Tjärnö Marine Research Laboratory (Sweden), and a detailed description of all steps in the production line is outlined, including the production and rearing of barnacle larvae, as well as the administration of feed for adults and larvae. For an overview of the complete procedure, see **Figure 1**. The use of the culturing system is exemplified with some common experimental set-ups and illustrated in functional genomics studies of Na^+/K^+ ATPase and aquaporins, elucidating their possible functions in osmoregulation[2,3]. It is sometimes essential to examine the gene expression in specific tissues, and some of the basics of barnacle dissection will be covered. With a good supply of high-quality seawater, the culturing of the barnacle *B. improvisus*, and potentially many other species, should be possible in marine laboratories throughout the world.

Protocol

1. Collection of Adult Barnacles in the Field to Start a New Broodstock

1. **Deploy thermoplastic (poly-methyl methacrylate) transparent panels (110 x 110 x 1.5 mm^3) at ≈ 1–3 m depth in calm waters in order to collect adult barnacles from the field. Use frames that can take several panels (Figure 2A), drill 2 holes (with a diameter of 6 mm) in the top corners of each panel and attachthepanelstotherackusingcableties. Thus, thepanelswillhangverticallyin the water.**
    1. When the settlement of barnacles has occurred (identified as small white hard dots on the panels), leave the panels for at least 2–3 weeks to make sure that the barnacles are large enough (5 mm) to be transferred to the lab without getting dried or damaged.
2. Bring the panels into the laboratory when they are covered with settled *B. improvisus* that have reached a size of *ca*. 5 mm indiameter.
   NOTE: The time for adults to develop to 5 mm in the field is highly dependent on the availability of food and the temperature. Normally, this takes about 3 weeks at the Tjärnö Marine Research Laboratory (58.87 °N, 11.14 °E). Most commonly, panels for the new broodstock are generally deployed in late June and collected at the end of August.
3. Before introducing the panels into the culture facility, clean other organisms from the panels. It is most important to remove other barnacle species. Please note that it is important to frequently clean the panels to get rid of contaminating species during the cultivation since this will significantly increase the chances of survival of the broodstock.
4. Place the panels vertically on edge in a stand with milled grooves in a polyethylene tray (400 x 400 x 20 mm^3) (**Figures 2C** and **2E**). The grooves in the stand are 15 mm apart.
5. In order to prepare the trays, make an entrance port at the base and an exit port at the top on the opposite side of each tray. Connect the entrance port of each tray to a 100 L reservoir containing about 30 L of water with an inlet of open seawater at 20 °C to allow a flow-through at a rate of *ca*. 2 L per min.
   NOTE: Up to 8 trays were connected to a single 100 L reservoir.
6. Connect the 100 L reservoir to temperature-controlled seawater (20 °C) (*e.g.,* provided by a heat pump exchanging heat from the incoming seawater).
   NOTE: The barnacle *B. improvisus* can grow and reproduce at a full marine salinity (30–35 psu); however, reducing the salinity to *ca*. 25 psu by adding freshwater to the 100 L reservoir increases the output of larvae from the culture.

2. Starting New Generations of Adults from Cultured Cyprids

1. Construct cubes of thermoplastic panels open at the top by taping 5 panels together using a non-toxic, water-resistant tape.
2. Place the cube in a small tray (in case of leakage) and fill it with seawater (25 psu). Keep the water at room temperature (*ca*. 20 °C). Add roughly 200 cyprids at the top of the cube.
3. Feed the juveniles with 100 mL of *Skeletonema marinoi* every other day (see step 5) while in the cube.
   NOTE: Juvenile barnacles can have difficulties with ingesting *Artemia salina* (see step 3) since they are big and, hence, difficult to swallow for a newly settled barnacle.
4. Change the water 1x a week.
5. After 2 weeks, when the panels contain enough established juveniles of at least 5 mm in diameter, take the cube apart and move the panels with the juveniles to trays with flow-throughseawater and, thereafter, feed the barnacles with *A. salina* nauplii.

3. Culture of *Artemia salina* Nauplii as Feed for Adult Barnacles

1. Set up a culturing system for *A. salina* by using a 1.5 L plastic bottle where the bottomis cut off and the bottle is placed upside-down in a stand and illuminated from the side. Fit the bottleneck to a silicon tube with a clamp and attach it to an aeration pump (**Figure 2D**).
2. In order to hatch the eggs, add about 15 mL of dry resting eggs of *Artemia* to 1 L ofseawater.
   NOTE: *Artemia* eggs hatch into nauplius larvae after 24–48 h. To delay the hatching of *A. salina* nauplii (*e.g.*, during weekends), the light can be turned off to reduce the temperature slightly.
3. To harvest *Artemia* nauplii, turn off the aeration and darken the upper part of the bottle with aluminium foil (or similar *e.g.*, a can).. Illuminate the lower part of the bottle for 10 min. Hatched *Artemia* nauplii will swim toward the light. Non-hatched cysts will sink to the bottom, and cyst shells will float at the surface.
4. Open the clamp at the bottom. Collect the intermediate fraction as a dense population of the swimming nauplii.

NOTE: Every day (except on weekends), 1 L of the dense *Artemia* nauplii suspension was manually added to the 100 L reservoir connecting to the trays with the adult barnacles. Avoid any feeding with empty cyst shells, since they do not provide nutrition and mainly result in needing to clean the culture at more frequent intervals.

4. Collection and Rearing of Barnacle Larvae

1. Clean the panels with adult barnacles by gently spraying them with freshwater and, if necessary, remove any fouling from the barnacle shells and panels using a soft toothbrush. Also, clean the trays (without barnacles) and tubes in hot (75 °C) freshwater.
2. Place a sieve made of a 90 µm plankton-net glued to the end of a cut PVC pipe (16 cm in diameter, 15 cm in height) into a polyethylene tray (30 x 20 x 10 cm^3) with an overflow port. Collect the *B. improvisus* larvae in the sieve overnight (**Figure 2C**).
   NOTE: The sieve was positioned just below the exit port of the tray of the barnacle panels to receive the outflowing water and to filter out the nauplii. The larvae remained in the sieve while the seawater overflowed the small tray. This small tray ensured that the sieve never driedout.
3. Place a 30 L bucket in a large water bath where the temperature is maintained at 26 °C using aquarium fish tank heaters (**Figure 2E**). Aeration and agitation are assured by air-bubbling through the system.
4. Fill the bucket with 20 L of filtered seawater (0.2 µm; 25 psu). Add 1 L to the bucket of a 60/40 mix of the 2 diatom microalgae, *S. marinoi* and *C. gracilis* (see step 5). This will give an initial density of about 5 x 10^4 diatoms per mL in the bucket.
   NOTE: Barnacle nauplii larvae are positively phototactic. The buckets were made of opaque white plastic with lids that let through some light. During the winter, the room was lit during the daytime (8:00–17:00), but dark during the nighttime. During the summer, the outside light came into the room during most of the day.
5. Transfer the collected *B. improvisus* nauplii to a crystallizing dish (300 mL; 90 mm in diameter; 50 mm in height).
6. Illuminate the dish from the side, which will attract the larvae to the light source. Collect the barnacle nauplii that gather at the incoming light with a pipette and transfer them to another crystallizing dish. Remove any remaining *Artemia* larvae.
7. Transfer the *B. improvisus* nauplii larvae to a beaker with 1 L of filtered seawater.
8. Count the number of nauplii by stirring the beaker gently to get an even suspension of larvae and then taking five 1 mL samples from different places in the beaker. Transfer each sample into a microplate for a visual inspection with a stereomicroscope. Count the number of nauplii in each of the five samples and add them up.
9. Multiply the counted number by 200 (1,000 mL/5 mL) to estimate how many thousands of larvae are in the whole beaker. If there are too many nauplius larvae in the beaker, dilute the sample until the density is at most 14,000 larvae/L (usually in the range of 11,000–12,000 larvae/L).
10. Add 1 L of *B. improvisus* nauplii to one bucket, thus adding roughly 11,000–12,000 larvae per bucket.
    NOTE: Make sure not to add more nauplii, since this will result in too little food in relation to the larvae and, hence, increase mortality.
11. After 3 days, collect barnacle nauplii on a 90 µm sieve. Then clean the bucket (with 75 °C freshwater), fill it with filtered seawater, add a new diatom feed (the same amount as at the start), and finally add the nauplii again. Cyprids start to appear after about 6 days (± 1 day).
12. Collect cyprids first with a 90 µm sieve (designed as in point 4.2 above). Then separate non-moulted barnacle nauplii and cyprids with a 320 µm sieve on top of a 160 µm sieve. *B. improvisus* nauplii will collect in the 320 µm sieve and cyprids in the 160 µm sieve.
    NOTE: The cyprid larvae may be stored at 10 °C in a crystallizing dish in darkness for later use, up to *ca*. 6 days. However, the storage can affect the quality and performance of the barnacle larvae, so experiments comparing different treatments should ideally use larvae from the same batch (and similar storage time) to avoid confounding effects[16].

5. Culture of Microalgae as Feed for Barnacle Nauplius Larvae

NOTE: Algae were grown in 3 different types of cultures: (i) stock cultures, which are for the long-term maintenance of the strains that were being used for the inoculation of the scaling-up; (ii) start cultures, which are the first step in the scale-up; and finally (iii) the production culture, which is the final production scale of large quantities of algae as barnacle feed.

1. Order diatom species from the Culture Collection of Algae and Protozoa (CCAP) to be used as feed for the barnacle nauplius larvae. The 2 species *Skeletonema marinoi* (CCAP strain 1077/5) and *Chaetoceros simplex* var. *gracilis* (CCAP strain 1085/3) both give good results as feed.
2. Filter all seawater to be used in the algal cultures, using a cartridge filter system with a nominal pore size of 0.2 µm. (The filtered seawater is also to be used for the hatching of *A. salina* eggs, and the culturing of barnacle nauplius larvae.) Autoclave the filtered seawater for the algal cultures (at 105 °C for 5 min).
3. Prepare a medium for the culture of microalgae, using autoclaved seawater enriched with Guillard's f/2 solution containing inorganic nutrients, trace metals, and vitamins (see Guillard 1975[17] for the detailed recipe). In addition, prepare a solution of silicate (Na_2SiO_3), but keep this separate from the f/2 enrichment to prevent a solid precipitation.
   NOTE: The concentrations of the enrichment stock solution and the stock solution of silicate were prepared to be used as a 1 mL addition of each to 1 L of seawater.
4. Autoclave all equipment used in the algal culture at 120 °C for 20 min. Autoclave screw-capped test tubes with 1 mL of f/2 enrichment and 1 mL of silicate solution. Add the enrichment and silicate solutions to theseawater when the autoclaved glassware and liquids have cooled down to room temperature.
5. **Grow stock cultures of microalgae in 40 mL test tubes with screw caps.**
   1. Fill the test tubes with about 30 mL of the enriched seawater. Inoculate the cultures with a sterile Pasteur pipette using about 1 mL of a 2 week old stock culture. The screw cap is then fitted but not tightened,to allow some gas exchange. If available, the inoculation should be carried out in a laminar flow cabinet to reduce the risk of contamination.
      NOTE: The aim of the stock cultures is to maintain algal cultures on a long-term basis and to serve as inoculum for start cultures. Stock cultures were re-inoculated every 2 weeks.
   2. Expose the new stock culture to a white light with an intensity of *ca*. 25–50 µmole $m^{-2}s^{-1}$ (with a light-dark cycle of 16:8 h). Move the old stock cultures to a low light intensity as a back-up. Discard any 2 week old stock cultures.

6. **For the scaling up to the production culture of algae, inoculate the start cultures from the stock culture and grow them in 500 mL Erlenmeyerflasks.**
    1. Autoclave 4 Erlenmeyer flasks, pipettes and cotton stoppers with 300 mL of filtered (0.2 µm) seawater, and 4 test tubes with 0.3 mL of f/2 enrichment and 4 with silicate solutions. Cool down the f/2 enrichment and the silicate to room temperature, and add the solutions to the Erlenmeyer flasks. Inoculate them with *ca*. 1 mL of the 2-week old stock culture.
    NOTE: Prepare 2 flasks with *S. marinoi* and 2 with *C. simplex*.
    2. Put the flasks on a shaking table in a light intensity of 50 µmole $m^{-2}s^{-1}$. When the start cultures have acquired dense algal populations (yellow-brown in color), they are ready to be used as inoculum for the production cultures that will provide the barnacle larvae with food.

7. **Grow production cultures in 4 L polycarbonate bottles with silicone stoppers with 2 drilled ports fitted with glass tubes.**
    1. Connect 1 glass tube reaching to the bottom of the bottle to an air pump *via* a silicone tube fitted with a 0.2 µm air filter. Make sure the second glass tube ends just below the silicone stopper and is filled with cotton to allow the exit of injected air.
    2. Each week, fill four 4 L bottles with filtered seawater and autoclave them together with the stoppers. In addition, autoclave 4 test tubes with 4 mL of f/2 enrichment and silicate solutions.
    3. After cooling, add the f/2 enrichment and silicate solutions to the bottles and then inoculate them with half of the volume of the start-up cultures in the Erlenmeyer flasks.
    4. Place the four 4 L bottles at a light intensity of 50–100 µmole $m^{-2}s^{-1}$.
    5. Harvest the production cultures after *ca*. 1 week; they are then sufficiently dense to be used for feeding the barnacle nauplius larvae.
    NOTE: Production cultures have a lifespan of about 2 weeks, which means that there are 8 active production bottles at any time.

6. Designing Experimental Studies Using Barnacles

1. **Place panels with juvenile or adult barnacles in controlled aquaria where they can be grown under identicalconditions. This is called a common-garden experiment, which, for instance, can be used to understand local adaptions or phenotypic plasticity[18], or to study gene expression changes in relation to external factors (*e.g.*, salinity, temperature, or pH).**
NOTE: It is also possible to use barnacles directly collected from the field on panels. An advantage of using laboratory-bred individuals is that maternal effects can be avoided when using the next generation of offspring.
    1. Expose barnacles to the specific environmental conditions during a chosentime interval, followed by the harvesting of adults (*e.g.*, for studies on gene expression)[2].

2. **For experiments with cyprid larvae to study gene expression[3], place the cyprids in controlled aquaria where larvae can be cultivated under identical conditions.**
    1. Harvest the cyprids after specified time periods by filtering with sieves (as described in step 4.7), and extract RNA according to the protocols below (step 8).

7. Dissection of Barnacles

1. Clean the lab space where the dissection and DNA sampling are performed, bothbefore and between individuals, including all dissection tools. This is done using chlorine for the bench (or 96% ethanol for the forceps).
NOTE: Labelled tubes containing fixation media (ethanol or an RNA stabilization solution) are prepared in advance.
2. Select large and short-term starved individuals (do not feed them for 2 days prior to the dissection). Clean the barnacle shell with a toothbrush to minimize the risk of contamination from other species (*e.g.*, bacteria, algae) and rinse it off with water.
NOTE: The reason for the short-term starvation of individuals before the dissection is to avoid any DNA contamination (*e.g.*, from *Artemia* cysts in the gut).
3. **Place the individual barnacles on an even surface, either attached to a panel or loose in a dish. Dissect the respective tissues from the adult. There are several ways to dissect barnacles, depending on the purpose of the study.**
    1. Dissection method A: fixing the whole barnacle as quickly as possible.
        1. Remove the whole barnacle from the panel by using a scalpel, inserting it under the barnacle and close to the panel surface.
        NOTE: In most cases, this manipulation leaves the basal plate almost intact, thus not affecting the barnacle inside.
        2. Carefully crack the outer shell on one side by inserting forceps (**Figure 3**). This is done to facilitate the intrusion of a fixating medium inside the shell.
        NOTE: If the shell is not broken at all before placing the barnacle in a test tube, it may lead to a slower fixation process and an inferior quality of DNA later.
        3. Put the broken barnacle in the test tube containing either an RNA storage solution or ethanol. Leave the barnacle in the solution for at least 24 h for its fixation. When the barnacle is fixed, the animal can be taken out of the fixation media and placed on a dissection tray. The cirri, mantle, and soma (body) can be separated from each other and placed in separate tubes for further extractions of DNA or RNA.
        NOTE: It is important to observe if egg lamellae with fertilized eggs are present inside the barnacle. If present, these should be removed before proceeding with the DNA extractions to avoid finding multiple genotypes in the sample.

    2. Dissection method B: removing the barnacle from the shell before its fixation.
    NOTE: This method may not always result in the mantle being sampled since it is attached to the inside of the calcareousshell. But it can be a quick method for sampling DNA from an individual.
        1. Carefully remove, by using dissection forceps, the tergal and scutal plates (**Figure 3**) by inserting the tip of the forceps between the tergal and scutal plates, grabbing hold of a plate, and pulling gently to remove them.

2. Grab hold of the cirri with the forceps and pull the barnacle straight out and place it directly in 96% ethanol or RNA stabilizing solution for fixation..
   NOTE: Sometimes, the mantle will also be sampled at the same time, as seen as a thin epithelium with dark pigmentation (in *B. improvisus*). Otherwise, the mantle can be removed from the inside of the shell using a scalpel and, thereafter, placed in ethanol or RNA stabilizing solution.
   NOTE: The tergal and scutal plates (if intact) can be dried and saved since they are useful for the species' identification[10]. A general recommendation, if there is any doubt of species, is to also photograph the whole barnacle prior to initializing the dissection.

8. RNA Extraction for Quantitative PCR

1. Put adults to be used for RNA extractions into an RNA stabilizing solution, incubate them overnight (up to 24 h) at 4 °C, and then store them at -80 °C.
   NOTE: Cyprid larvae are preferentially dry-frozen, without RNA later directly in -80 °C, by placing them in cryotubes and then submerging the tubes briefly (30 s) in liquid nitrogen before storing them at -80 °C.
2. At the time of the RNA extraction, thaw the barnacles on ice and use them intact or dissect out the tissues to be used (see step 7).
   NOTE: Due to high genetic diversity between individuals (≈ 3–5% variation in coding regions; Alm Rosenblad *et al.*, unpublished data) it is advisable to poola number of adult individuals, which minimizes the effect of sequence variation between individuals in the later qPCR step.
3. For the RNAextraction, add 350 µL of lysis buffer provided with an RNA preparation kit into homogenization tubes containing 2.8 mm of ceramic beads.
   NOTE: Ceramic beads provide a better yield of RNA compared to sonication (**Representative Results**).
4. Take an adult barnacle (whole or tissues) or collect a minimum of 20 cyprid larvae using forceps and put them into homogenization tubes.
5. Proceed directly to the disruption and homogenization with a bead mill.
6. Put the homogenization tubes with the sample and beads in the bead mill holder. Shake them with a frequency of 4.0 m/s for 20 s. Cool the sample for 1 min on ice. Repeat 2x.
7. Prepare the RNA according to the protocol of the commercially available RNA isolation kit.
   NOTE: A risk assessment (including a reading of the safety data sheets) should be performed prior to using RNA extraction methods, to identify hazardous chemicals that require the use of a fume hood and other protective clothing, including gloves (*e.g.*, the addition of β-mercaptoethanol to the lysis buffer during the RNA extraction should be performed in a fume hood).
8. Quantify the RNA. Prepare a working solution and add a standard and samples to a total volume of 200 µL. Vortex them for 2–3 s and incubate them for 2 min. Insert the tubes into a fluorometer and take readings.
9. Check the RNA samples for any protein contamination with a spectrophotometer, and check their RNA integrity with an automated electrophoresis system (**Representative Results**).
   NOTE: For good quality preparations, RNA is recommended to have a ratio of 260/280 nm of 2–2.2 and a DNA of around 1.8. Lower values indicate protein contaminations and that the extraction procedure has to be optimized.

9. Gene Expression: cDNA Synthesis and qPCR

1. **Treat the obtained RNAwith DNase to remove any remainingDNA before making the cDNA for qPCR.**
   1. Check for contaminating genomic DNA by adding an RNA sample but no reverse transcriptase when preparing the cDNA. If there is a PCR amplification of the selected gene on the cDNA made without reverse transcriptase, there is DNA contamination in the RNA.
      NOTE: For samples to be used for RNA-seq (RNA-sequencing using next-generation sequencing technology), the DNase step is less critical. In particular, for samples with low levels of RNA, the DNase step might be omitted in order not to lose too much of the RNA in the process.

2. Perform cDNA synthesis on DNase-treated RNA using an amount of RNA within the range specified in the commercial cDNA synthesis kit, but usually, use at least 50 ng.
3. Design qPCR primers so that they anneal to parts of the gene of interest where the sequence identity between individuals is as high as possible, but the sequence identity between gene paralogs is as low as possible. Please see the corresponding publications for the specific primer-pairs used for gene expression studies of Na^+/K^+ ATPase[3] and aquaporins[2] (**Figure 5**).
   NOTE: For this purpose, RNA-seq sequence data from hundreds of cyprids or several adults can be used.
4. Design primers for a gene to be used for the normalization of the expression levels (*e.g.*, actin). The primers used successfully for actin are, forward: 5'-CATCAAGATCAAGATCATCGC-3', and reverse: 5'-ATCTGCTGGAAGGTGGAC-3'.
   NOTE: Actin is a commonly-used control gene. However, others have also been suggested as a reference and tested out on barnacles[19]. Besides actin, the use of RPL8, 36B4, EF1, and NADHd1 as a reference has been tested[3]. It was concluded that there are no large differences in the expression levels of the respective reference genes between soma, cirri, adult, or cyprids.
5. Optimize the annealing temperature for the designed primers by adding increasing amounts of cDNA (typically in the range of 0.25–50 ng) to a mix containing SYBR Green, dNTP, polymerase, and 0.3 µM each of forward and reverse primer.
6. Run qPCR protocols at different annealing temperatures as follows: an initial denaturation temperature of 95 °C for 3 min, a denaturation step at 95 °C for 20 s, annealing temperatures of 55–63 °C for 20 s, and an elongation at 72 °C for 30 s. In total, run 40 PCR cycles.
   NOTE: Primer efficiencies should lie in the range of 90–105% and are calculated as E = (10^(-1/slope)-1) x 100 where the slope is the slope of the curve obtained when plotting the log value of the cDNA concentrations against their cycle threshold (Ct) values.
7. Perform qPCR using an appropriate amount of cDNA, usually 1–10 ng, using the PCR protocol described in step 3 with the optimum annealing temperature found.
   NOTE: The higher amount of cDNA is only used for genes that are expressed in extremely low amounts.

Representative Results

With the described procedure for the culturing of adult barnacles of *B. improvisus*, up to four batches of nauplius larvae can be produced per week. It would be possible to collect nauplius larvae almost every night, but this requires more people and infrastructure (with many barnacles in the broodstock, a culture will release larvae continuously). An additional limiting factor for the larvae production appears to be the availability of feed of high quality, in particular regarding the diatom *Skeletonema*. Maximally, each batch from the culturing system consists of roughly 12,000 nauplii, so up to 50,000 nauplii per week can be cultured. However, some weeks there may be up to tenfold fewer larvae produced. A single adult can produce up to 7,000 larvae per day[14], which means that 1–2 adults are releasing larvae for each batch. Within a week, about 70–90% of the collected nauplii will develop into cyprids (yielding roughly 30,000 cyprids per week, maximally) that can be used for settlement assays and molecular studies.

It should be stressed that there are variations in cyprid features between batches, and in general, there are larger variations *between* batches than *within* batches. For example, the settling success in settlement assays varies between 30 and 70% for different batches. Most likely, this is caused by the individual genetic variation between the specific pairs of adults releasing larvae during the different sampling periods. It is, of course, recommended that repeated experiments (biological replicates) should include cyprids from a number of batches if more general statements about the results are to be made. The batch-to-batch variation puts demands on the experimental design, where proper controls and normalizations in gene expression studies should be applied. However, even after several statistical normalization procedures have been implemented that considerably reduce between-batch variation, some effects of the batch are usually still apparent (unpublished data).

Following the provided protocol, it is possible to obtain, on average, 500 ng of high-quality RNA from as little as 20 cyprids, irrespective of the stage of barnacle settlement (**Table 1**). The quality of the RNA is usually measured as the ratio between the 18S and 28S peaks (the expected position of the two peaks are indicated in **Figure 4**). However, in the case of barnacles and many other arthropods, the 28S rRNA breaks down when heated (as part of the analysis method) and migrates together with the 18S peak[20]. This is why there is, in principle, one single rRNA peak in this type of analysis for barnacles. It is clear from this test (**Figure 4**) that a homogenization by ceramic beads provides the RNA with the highest integrity and is, therefore, the method of choice. The RNA is sufficient in amount and quality to generate high-quality sequencing libraries for sequencing, resulting in an average of 70 million reads per sample (the number of reads, of course, depends on the level of multiplexing during the sequencing). The amount of RNA is also sufficient for cDNA synthesis and qPCR expression analysis of a large number of genes.

Figure 5 shows the result from qPCR analyses of aquaporins and of Na^+/K^+ ATPase (NAK1) splice variants, where expression changes were investigated in response to changes in environmental cues[2,3]. A comparison of the relative expression of the long and short splice variants of NAK1 shows a twofold increase for the long NAK1 mRNA in low salinity in relation to the short NAK (**Figure 5A**). Thus, the data indicate that alternative splicing makes the long form predominant under low salinity conditions. In the case of aquaporins, it is apparent that the two water-transporting paralogs AQP1 and AQP2 display differential expression (**Figure 5B**). In particular, in the mantle tissue, it is apparent that the AQP1 is substantially down-regulated at lower salinities, which is not seen for AQP2. Instead, AQP2 shows a slightly increased expression at lower salinities, but in the soma. These findings provide a base for investigations of the functional roles of the different *B. improvisus* ion transporters and aquaporins in barnacle osmoregulation.

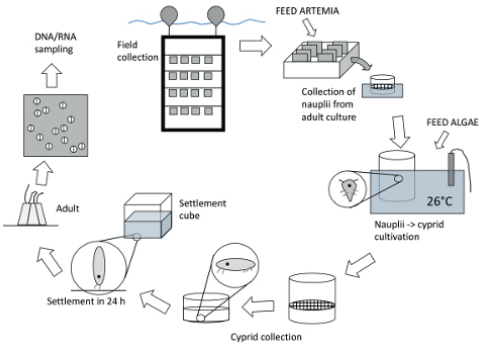


Figure 1: Overview of the whole culturing procedure and the RNA extraction for gene expression studies in adults. To initiate a new culture, panels are attached to a frame and deployed in the sea at 1 - 3 m depth. After several weeks, the panels with adults/juveniles are placed vertically in racks in trays in the laboratory. Each tray holds about 40 panels with adults. With roughly 100 adults per panel, in total ≈ 4,000 adult individuals are cultured per tray. The adult barnacles are fed with *Artemia* and can be kept year-around. Nauplius larvae are collected several times per week from the trays *via* a filtering through a sieve. The collected nauplii are transferred to buckets kept at 26 °C in a water bath and fed with microalgae. Nauplii are reared until they molt into non-feeding cyprids, which are collected by filtering. New panels can be established in the laboratory by the settling of cyprids on panels, either to provide new panels for the year-around culture or to be used for specific experimental set-ups with altered external conditions. RNA is then extracted from juveniles/adults at the end of the experiment or at specific time-points.
Please click here to view a larger version of this figure.

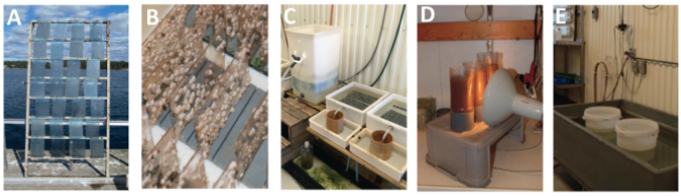


Figure 2: Images of some important steps in the culturing procedure. (**A**) This image shows the frames with panels for collecting new populations from the field. (**B**) This image shows the panels with adult barnacles of *B. improvisus* in racks that are placed in trays. The panels are placed about 2 cm apart. (**C**) This image shows the trays with the barnacle panels and the feeding tank to the left. From each tray, there is an outlet where the sieves are placed for the collection of nauplius larvae. (**D**) This image shows the production of *Artemia* feed for the adult barnacles. (**E**) This image shows the rearing of nauplii in buckets placed in a water bath set to 26 °C. Please click here to view a larger version of this figure.

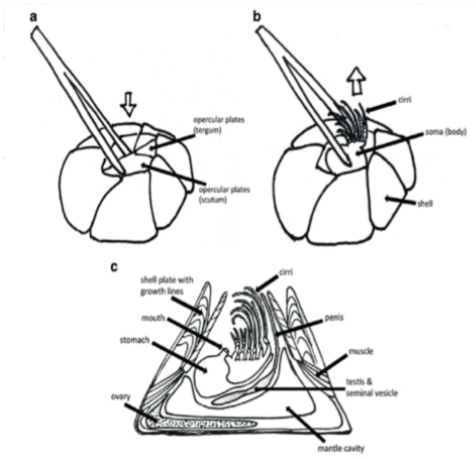


Figure 3: Description of the initial dissection steps of adult barnacles: removing the body from its shell. (**A**) Grab one of the opercular plates by inserting forceps gently through the aperture. Pull gently to remove the plate and expose the animal. (**B**) Pull out the animal by grabbing the soma part just below the cirri. (**C**) This panel shows the overall anatomy of acorn barnacles. The mantle and potentially fertilized eggs (in the ovary) stay in the shell cavity when the body is pulled out. A note on barnacle anatomy (for a more in-depth account, see Anderson)[9]: the wall plates of a barnacle slope inward, and together, they form a volcano-like cone. An opening, the aperture, is covered by the two opercular plates, which form a door, or operculum, to close the aperture. Acorn barnacles generally have a calcareous basal plate that is glued firmly to the substratum; however, some species of barnacles lack this calcareous plate (*e.g.*, *S. balanoides*). Barnacles secrete the exoskeleton from the darkly pigmented mantle (the carapace). The outer surface of the double-layered mantle is calcified to become rigid, while the inner surface of the mantle is not calcified and is therefore flexible. Inside the aperture, the cirri are present in a retracted position. These are the thoracic appendages that barnacles use for suspension feeding. The ovaries are located close to the base of the barnacle, while the testes lie in the soma. This figure has been adopted from Panova *et al.*[21] and has been published with permission from Springer. Please click here to view a larger version of this figure.

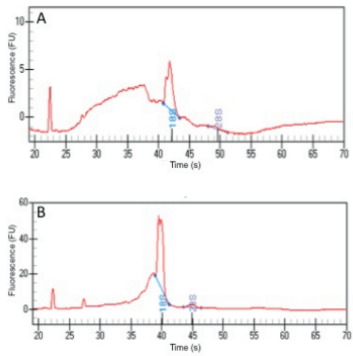


Figure 4: Determination of rRNA integrity by capillary gel electrophoresis. RNA was prepared by two different homogenization methods: (**A**) sonication and (**B**) ceramic beads. The unit on the y-axis, FU, stands for fluorescence units. Please click here to view a larger version of this figure.

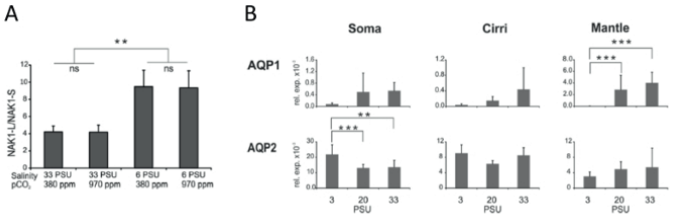


Figure 5: Gene expression results from two different studies of genes important for osmoregulation in *B. improvisus*. (**A**) This panel shows the differential expression as measured by qPCR of splice variants of the Na^+/K^+ ATPase *NAK1* in response to various salinities and pCO_2 levels[3]. In the low salinity treatment, the expression of the long isoform (NAK1-L) is increased relative to the short (ANOVA, $P < 0.001$). (**B**) This panel shows the expression of aquaporins in adults of *B. improvisus* during their exposure to various salinities[2]. qPCR was used to determine the aquaporin expression levels relative to actin. Adult individuals were incubated at 3 different salinities (3, 20, and 33 PSU) for 14 days. For the RNA preparation, the soma, cirri, and mantle of the adults were separated. In both figures, error bars indicate the standard deviation. The ** and *** indicate the level of significance (ANOVA), 0.01 and 0.001, respectively. These figures have been modified from Lind *et al.*[2,3]. Both figures are published with permission from PLoS ONE. Please click here to view a larger version of this figure.

Settlement Phase	Total amount of RNA (ng)
Free swimming	512
Exploration: Close search	518
Attached cyprid	550
Newly metamorphosed juvenile	832

Table 1: Yields of RNA. This table shows the RNA quantity extracted with an RNA preparation kit from pools of 20 cyprid individuals collected at different stages during the settlement process.

Discussion

The barnacle culture at Tjärnö Marine Research Laboratory (Sweden) has been running over 20 years and has been used for studies in many different research areas. Over 30 scientific papers have been published that have utilized the culturing system during the past years, including studies in antifouling[13,22], hydrodynamics[23], chemical ecology[24], climate change[16], evolutionary biology[5], and molecular biology[2].

To avoid the selection of certain individuals that are more adapted to the laboratory environment (individuals that might not be representative of the wild population), it is recommended to collect a new broodstock from the field each year. In addition, it is also good practice to rejuvenate the culture annually, since there is roughly 50 - 80% of mortality in adults during a normal year. However, if the aim is to produce inbred lines or to set up studies of experimental evolution, only laboratory-reared families are to be used.

A good time to collect *B. improvisus* on panels at the Tjärnö Marine Research Laboratory is in June–August because at that time, there is a good supply of cyprid larvae in the sea. Check the panels weekly to see when the barnacle settlement starts and manually remove other settled species than *B. improvisus* (*e.g.*, mussels, tunicates, bryozoans, hydroids, nemerteans/tubeworms, and other barnacle species) from the panels

(*e.g.*, with a toothbrush). Around Tjärnö, there are three shallow-water barnacle species present (*B. improvisus*, *Semibalanus balanoides*, and *Balanus crenatus*). However, *B. improvisus* is the dominant fouler of smooth hard surfaces during July–August. *S. balanoides* has its settlement period during early spring and prefers mainly natural substrates (*e.g.*, stones). *B. crenatus* can occur at low numbers on the panels during the summer.

It is also possible to start new adult barnacle generations from cultured cyprids, which would be essential if certain linages with specific traits have been established, or in studies of experimental evolution. The most convenient way to start new generations of adults is to settle cyprids on thermoplastic panels in the laboratory. These panels with newly settled cyprids could also be used in experimental treatments or for exposure in the field. In emergency cases, one can also use adults on boulders from a nearby site (*e.g.*, Idefjorden in the case of the Tjärnö Marine Research Laboratory) where *B. improvisus* is common. These already established adults are treated in the same way as the adults on the panels, thus being placed in trays and fed *via* the flow-through system. Flow cells can also be used to establish panels with barnacles[25]. These are flow-through chambers with plankton net on the sides on which the cyprids do not settle, with panels as the only settlement surface for the larvae.

There are several steps that are critical for setting up a long-term functioning barnacle culture including all life stages. The methods used to culture *B. improvisus* are probably, to a great extent, also applicable to other marine invertebrates with free-swimming, planktotrophic larvae. Culturing procedures for some species are already well described (*e.g.*, for blue mussels and different species of oysters)[26], while for other marine invertebrates, there are only a few examples of long-term cultures spanning their whole life cycle. One of the first successful attempts to culture barnacles (*B. amphitrite*) was done by Rittschof *et al.*[15]. Long-term financial and personal resources should be in place before considering setting up a barnacle culturing facility. The maintenance of this kind of year-around barnacle culture requires at least one person working half-time. There may be some potential for the future automation of some steps in the production line, mainly the culturing of microalgae[27]. In addition, in order to succeed, it is essential to have access to large quantities of high-quality seawater. The culturing of microalgae, *Artemia,* and barnacles does not involve any particular safety procedures. However, tests of some antifouling substances or toxic chemicals may need special precautions.

The panels were checked several times a week for contaminations. The seawater used in the culture was pumped up from a 40 m depth in the Koster Fjord outside the Tjärnö Marine Research Laboratory and was passed through two sand filters before entering the lab water system. If no filtering of the water had been done, there would be much more contamination in the culture. It is essential to regularly clean the panels in the culture from detritus and other invertebrates (*e.g.*, stolon-building hydroids and predatory nemerteans) that enter the system through the supply of seawater from the field. For example, if no larvae are produced despite the fact that the culture has been well-fed and otherwise seems to be in good condition, the problem might be the presence of nemerteans that appear to inhibit mating. Naturally, many of the contaminating organisms in the culture at the Tjärnö Marine Research Laboratory were specific for the Swedish west coast, and other types of contaminating organisms will be prevalent and be more of a challenge in other geographic areas. On the west coast of Sweden, it is unusual to find contamination by other barnacle species on the panels. Occasionally, the establishment of *S. balanoides* has been found, but this is a very marginal problem (at the most, one *S. balanoides* contaminant for 10,000 *B. improvisus* samples). The lack of contaminating species was most likely dependent on the regime to establish new cultures during the summer, when larvaefrom *B. improvisus* were highly dominant. In addition, there was also a clear enrichment of *B. improvisus* on the panels since this species is selective for smooth surfaces[13].

It is essential to remove dead adult barnacles. If the empty shells are left on the panels, they can become a shelter both for *Artemia* nauplii as well as for various contaminating species. In addition, it has been noticed that dead individuals influence the well-being of neighboring individuals, probably with the release of toxic compounds during the decomposition. An additional consequence of adult mortality is that some individuals will be left alone and too far away from any other adults to allow mating (even though barnacles have the longest penis in the animal world in relation to its size)[28]. These individuals will survive but are non-productive for larvae. However, these solitary adult individuals can gently be removed without harming the base-plate and be placed horizontally close to others to enable mating. Barnacles can also be mated by placing panels with one adult on each but close enough so that cross-fertilization can occur. In this way, genetic lines can be produced[14].

It is critical to produce a feed of high quality and to feed cultures almost every day. Even a few days without food can result in a decreasing release of larvae. Previous tests of diet composition have shown that diatoms are essential for the growth and survival of barnacle nauplii. Several diatom species seem adequate as feed, although small or solitary cells (less than 10 µm in diameter) may be necessary for the nauplii's ingestion. The species *S. marinoi*, *C. simplex*, and *T. pseudonana* have all proved to be adequate feed for *B. improvisus* nauplii, as well as easy to cultivate. In addition, the feed quality is generally higher for exponentially growing algae. It has also been reported that diatoms are essential for establishing productive cultures of *B. amphitrite*[15]. A theory of the importance of diatoms is that they have a unique fatty acid profile and are particularly rich in the highly polyunsaturated 20:5 fatty acid[29]. It has been shown that certain fatty acids are important for the successful development of oyster larvae[30].

Over the years, there have been no incidences of detrimental diseases in the barnacle culture. In many commercial invertebrate aquacultures, like oysters and mussels, diseases are rather common and may be very harmful. Detrimental effects of viruses have also been reported from wild populations. The native oyster in France was replaced by the Portuguese oyster *Crassostrea angulata* in 1925, but this species was wiped out by an iridovirus around 1970[31]. More recently, there have been massive mortality events in the Pacific oyster *Crassostrea gigas* in cultures worldwide, which appears to be associated with the ostreid herpesvirus 1[32]. No reports on pathogens, bacteria, or viruses on barnacles have been published so far. However, in the ongoing genome-project on *B. improvisus*, virus sequences were found (Alm Rosenblad *et al.*, unpublished data) but with no apparent link to symptoms of diseases. Mixtures of antibiotics have previously been applied to the cultures to minimize the risk of bacterial infections; however, this procedure is currently abandoned and so far, this has not caused any contamination problems.

If seawater is heated (as described above), overheating may be the most serious risk in the culture production line. It is, of course, difficult to safeguard against overheating, although sensors and appropriate alert systems may be used (*e.g.*, sending e-mail or text messages to responsible persons). Incidences of this kind in the past have resulted in the substantial killing of adults in the culture. This can, of course, be devastating and ruin long-term investments of time and money. In particular, this would be catastrophic if inbred genetic lines have been established. To ensure the longevity of such lines and secure them from accidental loss, it would be desirable to develop a cryopreservation methodology for barnacles. It has been reported that larvae from the Pacific oyster can be frozen down and revived with partial success[33]. Cryobanking has also been a valuable tool to preserve the genetic resources of a wide range of species[34]. Even nauplii from *B. amphitrite* are

reported to survive freezing[35], and it was found that 20% of the frozen-down individuals successfully metamorphosed into cyprids[36]. However, applying freezing for the long-term sustainability of cultures has so far not been adopted, but this would indeed be needed for the maintenance of selected lines; this would be an essential step to firmly establishing *B. improvisus* into a potent marine modelsystem.

Here, a protocol was presented for the dissection of various tissues from adults of *B. improvisus* (*i.e.*, cirri, soma, and mantel). However, it should be stressed that other tissues can also be extracted. For instance, the soft tissue between the exterior and internal mantle of the membranous-base species *Tetraclita japonica formosana* has been carefully isolated and used for an RNA extraction and RNA-seq analysis of gene expression[37]. The outlined optimized extraction protocol described here provides sufficient amounts of high-quality RNA for sequencing from a minimal amount of starting material. First, the collection of individual larvae directly into the homogenization tubes minimizes any loss during the transfer from one tube to another. Furthermore, among the different methods tested, the homogenization with ceramic beads proved to be the most efficient in terms of RNA yield and integrity, compared to sonication or pestle homogenization. When planning for gene expression or genomics experiments, one has to keep in mind the challenge of the high genetic variation in barnacles, at least for *B. improvisus*. Barnacle has a genetic diversity in the range 3–5%, even in coding regions (Alm Rosenblad *et al.*, unpublished data). This, of course, puts specific demands on the design of primers for qPCR analysis, where more conserved regions should be identified and used as templates for primers in order to get consistent expression results betweenbatches. Conserved regions for target genes, like aquaporins and Na+/K+ ATPases, can be identified by studying the sequence variability of these genes in RNA-seq data obtained from populations of cyprids containing hundreds of individuals. For a genome analysis, DNA will be sampled. However, obtaining high-quality DNA from *B. improvisus* can be challenging[21].

In conclusion, the established barnacle culture has proven to be instrumental in different kinds of experimental studies. In particular, the all-year-around larval production allows us to conduct experiments without being limited to the naturally occurring spawning period (for *B. improvisus*, this is during the summer). Obtained larvae can be used to perform a wide set of experimental studies, including settlement assays, behavior assays, expression studies on specific genes, as well as genome-wide transcriptome studies.

Disclosures

The authors have nothing to declare.

Acknowledgements

This research was supported by grant 2017-04559 from the Swedish Research Council (VR) and the EU supported project SEAFRONT to Anders Blomberg. In particular, the establishment of the culturing facility has, over the years, been supported by grants to Per R. Jonsson from the following funding agencies: SSF (Swedish Foundation for Strategic Research) through the program Marine Science and Technology and MISTRA through the program Marine Paint. Kent Berntsson was instrumental in the early phases of setting up the culturingfacility. Additional funding for establishing the culturing facility has come from the Centre for Marine Evolutionary Biology (www.cemeb.science.gu.se), which is supported by a Linnaeus grant from the Swedish Research Councils FORMAS and VR.

References

1. Darwin, C. *A monograph of the sub-class Cirripedia, with figures of all species. The Balanidae (or sessile cirripedes); the Verrucidae, etc., etc., etc.* Ray Society. London, UK (1854).
2. Lind, U. *et al.* Analysis of aquaporins from the euryhaline barnacle *Balanus improvisus.* reveals differential expression in response to changes in salinity. *PLoS ONE.* **12** (7), e0181192 (2017).
3. Lind, U. *et al.* Molecular characterization of the alpha-subunit of Na+/K+ ATPase from the euryhaline barnacle *Balanus improvisus.* reveals multiple genes and differential expression of alternative splice variants. *PLoS ONE.* **8**, e77069 (2013).
4. Fyhn, H. J. Holeuryhalinity and its mechanisms in a cirriped crustacean, *Balanus improvisus. Comparative Biochemistry and Physiology - Part A: Comparative Physiology.* **53** (1), 19-30 (1976).
5. Wrange, A. L. *et al.* The story of a hitchhiker: population genetic patterns in the invasive barnacle *Balanus.* (*Amphibalanus.*) *improvisus.* Darwin 1854. *PLoS ONE.* **11** (1), e0147082 (2016).
6. Berntsson, K. M., Jonsson, P. R. Temporal and spatial patterns in recruitment and succession of a temperate marine fouling assemblage: a comparison of static panels and boat hulls during the boating season. *Biofouling.* **19** (3), 187-195 (2003).
7. Lind, U. *et al.* Octopamine receptors from the barnacle *Balanus improvisus.* are activated by the alpha2-adrenoceptor agonist medetomidine. *Molecular Pharmacology.* **78** (2), 237-248 (2010).
8. Semmler, H., Hoeg, J. T., Scholtz, G., Wanninger, A. Three-dimensional reconstruction of the naupliar musculature and a scanning electron microscopy atlas of nauplius development of *Balanus improvisus.* (Crustacea: Cirripedia: Thoracica). *Arthropod Structure & Development.* **38** (2), 135-145 (2009).
9. Anderson, D. T. *Barnacles - structure, function, development and evolution.* Chapman & Hall. London, UK (1994).
10. Kennedy, V. S., DiCosimo, J. Subtidal distribution of barnacles in Chesapeake Bay, Maryland. *Estuaries.* **6**, 95-101 (1983).
11. Dreanno, C., Kirby, R. R., Clare, A. S. Locating the barnacle settlement pheromone: spatial and ontogenetic expression of the settlement-inducing protein complex of *Balanus amphitrite. Proceedings of the National Academy of Sciences.* **273**, 2721-2728 (2006).
12. Rittschoff, D. *et al.* Cues and context: larval responses to physical and chemical cues. *Biofouling.* **12** (1-3), 31-44 (1998).
13. Berntsson, K. M., Jonsson, P. R., Lejhall, M., Gatenholm, P. Analysis of behavioural rejection of micro-textured surfaces and implications for recruitment by the barnacle *Balanus improvisus. Journal of Experimental Marine BIology and Ecology.* **251** (1), 59-83 (2000).
14. Gamfeldt, L., Wallén, J., Jonsson, P. R., Berntsson, K. M., Havenhand, J. N. Increasing intraspecific diversity enhances settling success in a marine invertebrate. *Ecology.* **86**, 3219- 3224 (2005).
15. Rittschof, D., Clare, A. S., Gerhart, D. J., Sister Avelin, M., Bonaventura, J. Barnacle *in vitro.* assays for biologically active substances: toxicity and settlement inhibition assays using mass cultured *Balanus amphitrite amphitrite.* Darwin. *Biofouling.* **6** (2), 115-122 (1992).

16. Pansch, C., Schaub, I., Havenhand, J., Wahl, M. Habitat traits and food availability determine the response of marine invertebrates to ocean acidification. *Global Change Biology.* **20** (3), 765-777 (2014).
17. Guillard, R. R. L. Culture of phytoplankton for feeding marine invertebrates. In *Culture of marine invertebrate animals.* edited by Smith, W.L., Chanley, M.H., 22-60, Springer. Boston, MA (1975).
18. Wrange, A. L. *et al.* Importance of plasticity and local adaptation for coping with changing salinity in coastal areas: a test case with barnacles in the Baltic Sea. *BMC Evolutionary Biology.* **14**, 156 (2014).
19. Bacchetti De Gregoris, T. *et al.* Construction of an adult barnacle (*Balanus amphitrite.*) cDNA library and selection of reference genes for quantitative RT-PCR studies. *BMC Molecular Biology.* **10**, 62 (2009).
20. Winnebeck, E. C., Millar, C. D., Warman, G. R. Why does insect RNA look degraded? *Journal of Insect Science.* **10**, 159 (2010).
21. Panova, M. *et al.* DNA extraction protocols for whole-genome sequencing in marine organisms. *Methods in Molecular Biology.* **1452**, 13-44 (2016).
22. Berglin, M., Larsson, A., Jonsson, P., Paul Gatenholm, P. The adhesion of the barnacle, *Balanus improvisus.*, to poly(dimethylsiloxane) fouling-release coatings and poly(methyl methacrylate) panels: the effect of barnacle size on strength and failure mode. *Journal of Adhesion Science and Technology.* **15**, 1485-1502 (2001).
23. Larsson, A. I., Granhag, L. M., Jonsson, P. R. Instantaneous flow structures and opportunities for larval settlement: barnacle larvae swim to settle. *PLoS ONE.* **11**, e0158957 (2016).
24. Toth, G. B., Lindeborg, M. Water-soluble compounds from the breadcrumb sponge *Halichondria panicea.* deter attachment of the barnacle *Balanus improvisus. Marine Ecology Progress Series.* **354**, 125-132 (2008).
25. Pansch, C., Jonsson, P. R., Berglin, M., Pinori, E., Wrange, A. L. A new flow-through bioassay for testing low-emission antifouling coatings. *Biofouling.* **33** (8), 613-623 (2017).
26. Walne, P. Observation of food value of 7 species of algae to larvae of *Ostrea edulis.* 1. Feeding experiments. *Journal of the Marine Biological Associateion of the UK.* **43**, 767- 784 (1963).
27. Sandnes, J. M. *et al.* Real-time monitoring and automatic density control of large-scale microalgal cultures using near infrared (NIR) optical density sensors. *Journal of Biotechnology.* **122** (2), 209-215 (2006).
28. Neufeldt, C. J., Palmer, A. R. Precisely proportioned: intertidal barnacles alter penis form to suit coastal wave action. *Proceedings of the Royal Society B Biological Sciences.* **275**, 1081-1087 (2008).
29. Dunstan, G., Volkman, J., Barrett, S., Leroi, J., Jeffrey, S. Essential polyunsaturated fatty acids from 14 species of diatom (*Bacillariophyceae.*). *Phytochemistry.* **35**, 155-161 (1994).
30. Jonsson, P. R., Berntsson, K., André, C., Wängberg, S.-Å. Larval growth and settlement of the European oyster (*Ostrea edulis.*) as a function of food quality measured as fatty acid composition. *Marine Biology.* **134** (3), 559-570 (1999).
31. Comps, M., Bonami, J. R., Vago, C. Virus-disease of Portuguese oyster (*Crassostrea angulata.* lmk). *Comptes rendus hebdomadaires des séances de l'Académie des Sciences. Série D, Sciences naturelles.* **282**, 1991-1993 (1976).
32. Segarra, A. *et al.* Detection and description of a particular *Ostreid herpesvirus.* 1 genotype associated with massive mortality outbreaks of Pacific oysters, *Crassostrea gigas.*, in France in 2008. *Virus Research.* **153** (1), 92-99 (2010).
33. Suquet, M. *et al.* Survival, growth and reproduction of cryopreserved larvae from a marine invertebrate, the Pacific oyster (*Crassostrea gigas.*). *PLoS ONE.* **9** (4), e93486 (2014).
34. Martínez-Páramo, S. *et al.* Cryobanking of aquatic species. *Aquaculture.* **472**, 156-177 (2017).
35. Anil, A. C., Tulaskar, A. S., Khandeparkar, D. C., Wagh, A. B. Cryopreservation of *Balanus amphitrite.* nauplii. *Cryobiology.* **34**, 131-140 (1997).
36. Oo, K. *et al.* Cryopreservation of nauplius larvae of the barnacle, *Balanus amphitrite.* Darwin. *Fisheries Science.* **64**, 857-860 (1998).
37. Lin, H. C. *et al.* First study on gene expression of cement proteins and potential adhesion-related genes of a membranous-based barnacle as revealed from next-generation sequencing technology. *Biofouling.* **30** (2), 169-181 (2014).

IV

Insights into the molecular mechanisms of *Balanus improvisus* settlement and discriminatory behaviour towards surfaces

Anna Abramova, Magnus Alm Rosenblad, Anders Blomberg

Department of Chemistry and Molecular Biology, University of Gothenburg, Gothenburg, Sweden

Corresponding author e-mail: anders.blomberg@cmb.gu.se

Abstract

The barnacle *Balanus improvisus* is a common fouling species in temperate waters. Barnacle cyprid larvae have developed complex settlement behavior and surface selectivity to ensure survival and reproduction. However, the molecular mechanisms behind settlement as well as initiation of metamorphosis remain largely unknown. To investigate gene expression changes during the settlement process of the barnacle *B. improvisus,* we performed transcriptomics of four settlement stages: free-swimming, close-search, early-attached and juvenile. In order to understand the impact of surface properties the experiment was performed on both hydrophobic and hydrophilic surfaces. Comparison between exploratory stages, free-swimming and close-search, did not reveal any significantly regulated genes. However, large-scale gene expression changes occur during the transition from a pelagic cyprid to the early-attached stage. We identified several ecdysone cascade genes, including EcR, RXR, BRC, E75, H3 and H38. Gene expression on surfaces with different wettabilities further provides support that chemical recognition is involved in the surface wettability response during settlement. Furthermore, gene expression response on the hydrophobic surface suggests that exploration of the "favourable" surface might induce stronger and/or earlier transcriptomic response. Overall, transcriptomics analysis of settlement stages revealed several important genes that potentially play a key role in the regulation of adhesion, molting and metamorphosis in barnacles. Our results allow to add more details to the current understanding of the molecular mechanisms behind the settlement process and provide a basis for investigation of the key genes controlling this complex process.

Keywords

Transcriptomics, barnacles, settlement, metamorphosis, ecdyson cascade, hydrophobic surface

Background

Barnacles are common fouling organisms in communities growing on man-made underwater constructions. These sessile organisms have drastically influenced commercial enterprises related to shipping operators for millenia, and still are, making marine biofouling prevention a worldwide industry of five billion US dollars a year (Davidson et al. 2016). Barnacles constitute a diverse group of crustaceans called Cirripedia that is mainly represented by sessile suspension feeders. This group contains more than 1,000 species inhabiting all marine environments from rocky intertidal zones to abyssal vents (Pérez-Losada et al. 2008).

The barnacle *Balanus (Amphibalanus) improvisus* is one of the most common fouling organisms (Berntsson and Jonsson 2003). This barnacle species is a typical inhabitant of rocky shore communities, but is also often found on boat hulls, jetties and other underwater constructions. Similar to most other barnacles, *B. improvisus* is predominantly hermaphroditic with cross-fertilization, however, self-fertilization can also occur (Furman and Yule 1990). The natural reproductive period of this species in Swedish waters is continuous throughout June to September, with distinct spawning peaks in late July and late August (Berntsson and Jonsson 2003). However, culturing in the laboratory allows for a constant supply of larvae, making it an attractive model species for studying different aspects of settlement biology including the testing of antifouling coatings (Jonsson et al. 2018).

The barnacle life cycle involves sessile adults and planktonic larvae, including six naupliar stages followed by a non-feeding cyprid larva. Cyprids actively explore their surrounding environment for a suitable site for settlement and subsequent metamorphosis into adults. The original model of cyprid exploratory behavior includes three types of behaviours: wide searching, close searching and inspection (Crisp 1974). First, the cyprid performs wide search by walking almost in a straight line on its pair of antennules, investigating a substantial area of the substrate. Wide searching is followed by close search where the cyprid makes short steps and turns exploring a more narrow area, being attached with one antennule and probing the surroundings with the other one. During the inspection, the last stage preceding adhesion, the larva remains on one spot with both antennules attached on the surface performing testing of the attachment site. If during inspection the site is considered favourable for settlement, the cyprid permanently attaches to the substrate with adhesives secreted from the cement glands, and then undergoes metamorphosis. Recently, this behavioral model was confirmed by video recording and automated quantitative 3-dimensional analysis, also showing that the behaviour is largely conserved between individual cyprids (Aldred et al. 2018).

Cyprids actively select the best settlement site based on many different cues, such as light conditions, flow regime, surface topological structure, surface chemistry, and the presence of conspecific adults and larvae (Lang et al. 1979, Judge and Craig 1997, Dreanno et al. 2006a, Dreanno et al. 2006b, Matsumura and Qian 2014). In particular, previous settlement experiments with *B. improvisus* cyprids on surfaces with different wettabilities revealed that settlement inversely correlates with increased

wettability, and highly wettable surfaces totally inhibit cyprid settlement (Dahlström et al. 2004). Following this observation, two possible mechanisms on how wettability can affect settlement were proposed; either through chemical recognition of the molecular groups on the surface or that physical forces, such as electrostatic repulsion, impede cyprid contact with the surface (Dahlström et al. 2004). However, the exact mechanism still remains unknown.

It is believed that settlement cues are sensed by receptors on the cyprid antennular setae resulting in the activation of the signal transduction and initiation of larval attachment. The nature of the receptors remains elusive, however, G-protein coupled receptors (GPCRs) have been thought to be likely candidates for the reception of settlement cues in a number of marine invertebrate species (Tran and Hadfield 2012). Recently, we identified two receptor classes expressed in the cyprids' antennules, ionotropic receptors (IRs) and transient receptor potential channels (TRPs), that might mediate sensing of settlement cues (Abramova et al. 2019). TRPs were previously proposed to mediate tactile surface exploration in *B. amphitrite* (Kotsiri et al. 2018). There is evidence from pharmacological studies that two signaling pathways, e.i. cyclic AMP and phosphatidylinositol, are involved in transduction of signals from settlement cues (Clare et al. 1995, Yamamoto et al. 1995). Various biogenic amines and hormones such as serotonin, dopamine, methyl farnesoate (MF) and 20-Hydroxyecdysone (20E) were also reported to regulate larval attachment and metamorphosis (Yan et al. 2012).

The prevailing hypothesis about metamorphosis in crustaceans, heavily relying on comparison with what is known in insects, is that metamorphosis is inhibited by methyl farnesoate, an analog of juvenile hormone in insects, while a surge of ecdysone initiates metamorphosis (Hyde et al. 2019b). In insects, ecdysone binds to the ecdysone receptor and initiates an ecdyson cascade, including wide-scale transcriptional changes leading to shedding of old cuticle, tissue remodelling and other physiological changes accompanying metamorphosis. Knowledge about the ecdysone cascade is rather limited in crustaceans, however, recent studies reported the presence of the key-genes of the cascade in various crustacean species (Hyde et al. 2019b). Furthermore, it has been suggested that settlement of barnacles shows resemblance to the development of holometabolous insects (Smith et al. 2000, Kotsiri et al. 2018). Therefore, knowledge on insects can be used to construct working models to further advance our understanding of barnacle settlement and metamorphosis. Knowing the molecules orchestrating barnacle development will be essential to answer the key question on how the information from different settlement cues is integrated and translated into to the decision to attach and metamorphose.

The present study was undertaken to get a better understanding of the molecular mechanisms during barnacle settlement and metamorphosis. We used RNA-seq analysis based on short-read sequencing technology that is widely used for transcriptome profiling of both model and non-model organisms. To get an overview and a better molecular description of the entire settlement process, we selected four different stages for analysis, i.e. free-swimming, close-search, attached and juveniles.

Additionally, we performed settlement experiments comparing transcriptome profiles on two different surfaces with different chemical properties, hydrophobic and hydrophilic. In summary, this study provides new insights into molecular mechanisms involved in barnacle settlement and metamorphosis. Importantly, we identified key genes of ecdysone cascade, ie EcR, RXR, BRC and several nuclear receptors, adding another level of complexity to the established molecular mechanisms of settlement and metamorphosis control. Gene expression on surfaces with different wettabilities suggests that exploration of the “favourable” surface induced a stronger transcriptial response.

Methods

Cyprids culture

Cyprid larvae of *B. improvisus* were reared in an all-year-round laboratory culture as described earlier (Jonsson et al. 2018). Briefly, broodstock is established by collecting adults of *B. improvisus* on the thermoplastic panels from the field at the Tjärnö marine research station (N 58.88°, E 11.15°) in June-August each year. Panels are kept in trays with flow-through seawater system (25 ‰) at 20C. Adults of *B. improvisus* were fed with *Artemia salina* nauplii. Released barnacle nauplii were collected on sieves (90 µm) through the trays overflow and then transferred to 25-liter buckets and fed with a mixture of *S. marinoi* and *C. gracilis*. After about 6 days, when the cyprid larvae start to appear, the cultures were sieved (200 µm) and cyprids were removed for subsequent experiments.

Cyprid settlement

Nauplius larvae were collected from three independent trays with panels containing adult barnacles in the laboratory culture, and reared separately until reaching the cyprid stage. In this way we obtained three independent cyprid batches with different parental background. One-to-two days old cyprids were distributed between polystyrene petri dishes filled with 15 ml filtered seawater (25 ‰), 20 individuals in each petri dish, keeping larvae from the three different batches separate. We used two types of petri dishes: petri dishes with standard hydrophobic surface (Nunc No 150340, Ø 48 mm) and petri dishes with hydrophilised surface (Nunclon^{tm} Delta, Nunc No 150326, Ø 48 mm). Since standard hydrophobic petri dishes are commonly used in the settling experiments/assays this type of surface will be used in the initial characterisation of the developmental stages during settling. The settling experiment was repeated three times with three independent batches of cyprids.

Dishes were maintained at room temperature under natural light conditions. After 4-5 days, we collected individuals at different stages of settlement: free-swimming, close search, attached and juveniles. The classification of pre-settlement behaviors were based on earlier descriptions (Lagersson and Hoeg 2002). The free-swimming stage included all behaviors when cyprids were not laying passively on the bottom or antennules were not in a contact with the surface. Close-search was

characterized by larvae making short steps and turns while exploring the surface in the vicinity not longer than the body length. We did not collect the inspection stage because it is extremely difficult to distinguish the stationary larva testing the substratum from the permanently attached cyprid (Maruzzo et al. 2012). The attached stage corresponds to a time period of 4 to 7 hour after cementation when the larval body is oriented parallel to the surface with carapace tightly applied to the substratum, as described by Maruzzo et al. (2012). Juveniles were collected at the stage when the body shape resembles such of an adult barnacle with ventral side in contact with the surface, and thoracopods were continuously beating inside or extended outside. All stages were observed under stereomicroscope and collected with tweezers directly into the RLT buffer supplied with RNeasy micro kit (Qiagen) after at least 20s of observing the characteristic behavior for each cyprid.

RNA extraction and sequencing

Samples were homogenized with Precellys Lysing Kit for hard tissue (CK28) and extracted following the RNeasy micro kit protocol. We collected 4 stages from each surface and each batch resulting in 24 samples in total. Extracted RNA was made into sequencing libraries with Illumina TruSeq Stranded mRNA protocol following Poly-A selection. The libraries were multiplexed and fully mixed into one single pool to circumvent any sequencing bias. The multiplexed pool was sequenced on 3 lanes, yielding on average 30 mln reads per sample. Library preparation and Illumina sequencing was performed at the Science for Life Laboratory in Stockholm (https://www.scilifelab.se). RNA sequencing generated in total more than 695 million raw reads with on average 30±18 million reads per sample, which should be sufficient to obtain reliable measurement of most genes in the genome (Wang et al. 2015b). One sample, corresponding to the early-attached stage from the second batch of cyprids, failed to produce sequencing library, therefore, only 23 samples were sequenced.

Transcriptome assembly

The quality control program FastQC 0.8.0 (Andrews 2010) was used to assess read sequence quality and to detect and remove adapters. Subsequently, digital normalization was carried out using Trinity 2.0.6 (Haas et al. 2013). Digital normalization allows for reduction of the total number of reads considerably reducing computation time while retaining the complexity and ability to reconstruct transcripts. We applied the Trinity script for *in silico* normalization for all libraries separately with maximum coverage 100 and other parameters left in the default mode. De novo transcriptome assembly was done using Trinity 2.0.6 including all normalized libraries. Assembly with default parameters resulted in a highly fragmented assembly containing 1488124 trinity “genes”. To optimize the assembly we set minimum k-mer coverage to 4 and minimum contig length to 300. These changes resulted in considerable reduction of the number of trinity “genes” to 115268. Further, the assembly was clustered with CD-HIT 4.6.5 (Li and Godzik 2006) with similarity threshold 0.95 and run through Transrate 1.0.3 (Smith-Unna et al. 2016) program to perform assembly quality check and

select contigs with good reads support, reducing assembly further to 94910 trinity "genes". In the final step, the transcripts were filtered based on the expression values with Trinity script filter_low_expr_transcripts.pl, retaining only the most highly expressed isoform per gene with the --highest_iso_only option. This resulted in the final 72,407 number of contigs. We evaluated the completeness and duplication level of the assembly with arthropoda BUSCO v.2 (Simão et al. 2015), comprising 113 arthropod species and 1066 single copy orthologs. Transdecoder (v3.0.1) was applied to predict open reading frames (ORFs) in the assembled contigs; 48% were predicted by Transdecoder to have ORFs. Annotation of the transcripts was done with the Trinotate pipeline (Haas et al. 2013). Obtained annotation was subsequently used to perform GO enrichment analysis in GOseq 1.30.0 (Young et al. 2012).

To identify potential barnacle-specific genes not presently available in the public databases, we created a customised database encompassing UniRef90 proteins (19 september 2017) and predicted proteins from adult transcriptomes of three other barnacle species, namely *B. amphitrite* (SRS291871), *B. balanus* and *B. crenatus* from previously generated RNA-seq data using Illumina sequencing technology (data not published). We used the generated transcriptome reference assembly for *B. improvisus* and run Blastp against the local customised database containing other barnacles' transcriptomes.

Gene Expression Analysis

The resulting assembly was used to calculate the transcript abundances using the alignment-based quantification method RSEM (Bray et al. 2016) through the use of a Trinity 2.0.6 wrapper. Resulting abundances were passed to the Trinity script abundance_estimates_to_matrix.pl with --est_method RSEM and `--aln_method` bowtie2, to generate a matrix of counts and TMM (trimmed mean of M-values) normalized matrix. We filtered out and discarded lowly-expressed genes by keeping genes with at least 2 cpm (corresponding to 60 counts on average) in at least 2 samples across the entire experiment for the subsequent analysis. After preliminary data quality check we observed a pronounced batch effect related to the batch of cyprids (from which parental culturing tray the cyprids originated). To minimize the batch effect from the count matrix we applied the script removeBatchEffect from the edgeR 3.4.2 (Robinson et al. 2010). The batch-corrected count matrix was used for the hierarchical clustering and PCA analysis.

Subsequent differential gene expression analysis was performed using the edgeR package. To identify differentially expressed genes (DEGs) between different stages and account for the batch effect, we used limma package (Ritchie et al. 2015) with *voom* method, linear modelling and empirical Bayes moderation following guidelines in Law et al. (2016). For each pairwise comparison we only reported genes that had adjusted p-values equal or smaller than 0.05 and a fold change more than 4. We performed hierarchical clustering on all 23 samples from the four settling stages using complete linkage and Pearson correlation. Data from the individual independent samples are shown in the heat map, indicating a rather high level of repeatability between cyprid batches.

Results

De novo transcriptome of *B. improvisus* settlement stages

Cyprid larvae of *B. improvisus* were studied during settling onto two different type of surfaces, hydrophobic and hydrophilic. On each of the surfaces, four settlement stages were collected: free-swimming (FS), close-search (CS), attached (AT) and juvenile (JUV). The settling process was observed under the stereomicroscope and individuals at different stages of settling were collected with tweezers. Close-search was characterized by larvae making short steps and turns while exploring the surface. Attached stage corresponds to when cyprids have attached and the larval body is oriented parallel to the surface (roughly 4 to 7 hour after the initial cementation; (Maruzzo et al. 2012)). Juveniles were collected at the stage when the body shape starts to resemble an adult barnacle.

After optimisation of the extraction procedures enough high-quality RNA for sequencing was obtained from 20 individuals for each sample (Jonsson et al. 2018). RNA sequencing generated on average 30±18 (± standard deviation, SD) million reads per sample. This gives sufficient depth of sequencing to obtain measurements of most of the differentially expressed genes in the genome (but not enough to score differential alternative splicing) (Liu et al. 2013, Wang et al. 2015b), and corresponds to 100-fold greater depth of sequencing than earlier transcriptome studies on different settling stages of barnacles (Chen et al. 2011). Three independent biological replicates for each settling stage, allows for statistical analyses of differentially expressed genes between different settling stages.

Sequencing reads from all settlement stages were used to create a *de novo* RNA reference assembly for read mapping and quantification. After the clustering and filtering steps, the resulting reference assembly comprised of 72,407 genes (Table 1). To assess the level of completeness and duplication in our reference assembly, we ran BUSCO utilising the arthropod-set that contains 1,066 arthropods' single-copy orthologs (Table S1). More than 69% of complete orthologs and roughly 10% partial orthologs were present in the assembly; i. e. the *B. improvisus* reference assembly contained about 80% of the arthropod BUSCO set. The high nucleotide diversity of *B. improvisus* that is in the range 5% in coding regions (Alm Rosenblad et al., manuscript) would lead to individual alleles of genes to be assembled separately. The BUSCO analysis revealed a duplication level of 1.6% in our final reference assembly, where the last filtering step during assembly optimisation/filtering decreased the number of duplicated transcripts by almost 6-fold (9.3% to 1.6%; Table S1).

Around 48% of the transcripts in our reference assembly were predicted by Transdecoder to have ORFs, where 6,238 contained a complete ORF with an average length of 326 amino acids (Table 1). Blast hits (E-value < 1,00E-05) in the Swiss-Prot database was found for 10,600 transcripts and 1,716

of the remaining transcripts contains a Pfam domain; thus, in total we provided GO annotations for 13,010 (18%) of the transcripts. To identify potential barnacle-specific genes, we created a customised database encompassing proteins from adult transcriptomes of three other barnacle species, namely *B. amphitrite*, *B. balanus* and *B. crenatus (*Alm Rosenblad et al., unpublished data). We ran blastp of the generated reference assembly for *B. improvisus* against the local customised barnacle database; of the 72,407 *B. improvisus* transcripts 17,132 got Blast hits (E-value < 1,00E-05) only to the barnacles and not to any Uniref90 proteins, giving a current estimate of roughly 25% of barnacle-specific genes in our *B. improvisus* dataset.

Table 1 *De novo* transcriptome assembly statistics.

Assembly Statistics	
Total assembled bases	56,337,608
Total Trinity transcripts - reference assembly	72,407
% GC	53.7
Average contig length (nt)	778
Maximum length (nt)	17,992
Minimum length (nt)	300
Assembly Annotation	
Complete ORFs	6,238
Annotated in Swiss-Prot	10,600
Annotated only in Pfam	1,716
Genes with GO annotation	13,010
Estimated barnacle-specific genes*	18,020

*Contigs that did not find any match in UniRef90 but instead mapped only to the customised/local barnacle database were suggested to be barnacle-specific genes.

Transcriptome profiling of stage transitions during the settlement process

The hydrophobic surface is the preferred surface for settling of cyprids of *B. improvisus* (Dahlström et al. 2004), and we therefore first examined gene expression changes during the different stage transitions on this type of surface. Overall, the *B. improvisus* global transcriptional profiles during the different settlement stages revealed highly dynamic changes in gene expression as seen in

hierarchical clustering (heatmap) based on Pearson correlation coefficients (Figure 1A). In total 4,275 (≈6%) of all genes showed a statistically significant (FDR < 0.05, fold change (FC) > 2) change in expression during at least one of the stage transitions, with the highest number of genes regulated in the transition between the CS and AT stages (≈ 5% of all genes). It is clear from the heatmap that a larger proportion of genes are up-regulated (1106 genes) compared to down-regulated (296 genes) during the transitions from the cyprid stages to the attached stage. It is also clear that the two cyprid stages, FS and CS, displayed very similar gene expression with in fact no statistically significant changes (FDR < 0.05, FC > 2).

K-means clustering (k=10) of the differentially expressed genes also indicated several clear serial patterns of expression (Figure 1B). Among the induced genes in the attached stage, 164 genes displayed a transient expression profile with a drastic increased expression in attached satge and then roughly cyprid-level of expression again in the juvenile stage (clusters V; for a list with those with a functional description see Table S2). The other genes being regulated in the attached stage showed more or less a similar or higher level of expression in the juvenile stage, however, finally reaching different absolut levels of expression (clusters III, IV and X). A smaller number of genes (186) were slightly induced in the attached stage and then further substantially increased in expression in the juvenile stage (clusters VII and VIII) . In contrast, 376 genes did not display a change in gene expression until the juvenile stage (cluster XI). Once again we observe that induction is more common than repression, and only three out of ten clusters were characterised by the main response being down-regulation of genes (clusters II, VI, and X).

Overall, only 562 (12%) of the 4,725 genes regulated during at least one of the transitions were functionally annotated and these were used to examine the gene ontology (GO) enrichment in the different stages. We first compared the gene annotations for genes more highly expressed in the cyprids (FS plus CS) in relation to juveniles. The top 35 of the enriched GO terms in the cyprid in comparison to juveniles represent mainly two biological processes; signal transduction and muscle activity (Figure S3). Overall, out of the total 370 enriched GO molecular function terms for genes more highly expressed in cyprids, 31 were related to channel complex or channel activity (average enrichment factor 2), which is line with cyprids being designed for swimming and searching for environmental cues. Next we examined GO enrichment in the different stage transitions. As expected, the GO analysis of upregulated genes in the AT stage compared to the CS stage showed enrichment for functions related to molting and cuticle development (biological process; Figure 2). In addition, and maybe more surprising, we also found functional enrichment for upregulated genes in the AT stage that were involved in hydrogen peroxide metabolic process.

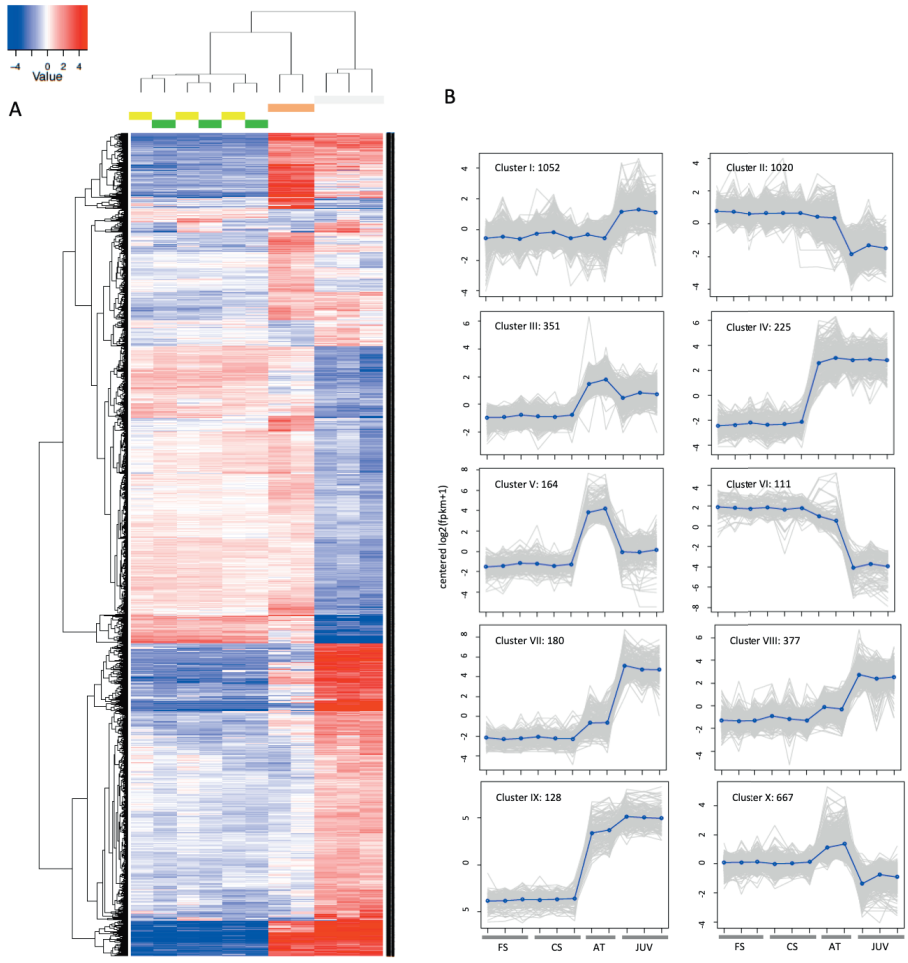


Figure 1 A) The heatmap represents hierarchical clustering (based on Pearson correlation coefficient) of the 4,275 differentially expressed genes (differentially expressed in at least one of the transitions; FDR < 0.05, FC > 2) during the serial transitions in the settlement process. Colour coding of the top dendrogram: yellow free-swimming stage, green close-search stage, orange attached and grey juvenile; B) k-means clustering (k = 10); free-swimming FS, close-search CS, attached AT, juvenile JUV.

Among the top-candidates (the most highly induced) in the AT stage we found several interesting genes involved in developemnt (Table 2), e.g. the relaxin receptor 2 (24-fold induced) that has been shown in *Drosophila* to effect developmental timing and growth coordination via regulation of the ecdysone pathway (Jaszczak et al. 2016), the pro-resilin (20-fold induced) that is an elastomeric protein found in arthropods that provides rubber-elasticity to mechanically active organs/tissue

(Michels et al. 2016), and juvenile hormone binding protein (18-fold induced) that is a haemolymph protein that has been shown to regulate larval development in insects (Jindra et al. 2013). Furthermore, in total 350 genes in our reference assembly were assigned to the DNA-binding transcription factor activity (GO category GO:0003700). About 160 (46%) of these transcription factors were significantly regulated during the settling process, with 37 regulated in the transition from close-search to the AT stage and 76 between AT and JUV stages (FDR < 0.05, FC > 2). The top GO categories of the upregulated genes in the transition to juveniles reflected active protein synthesis as well as metabolic synthesis (Table S5).

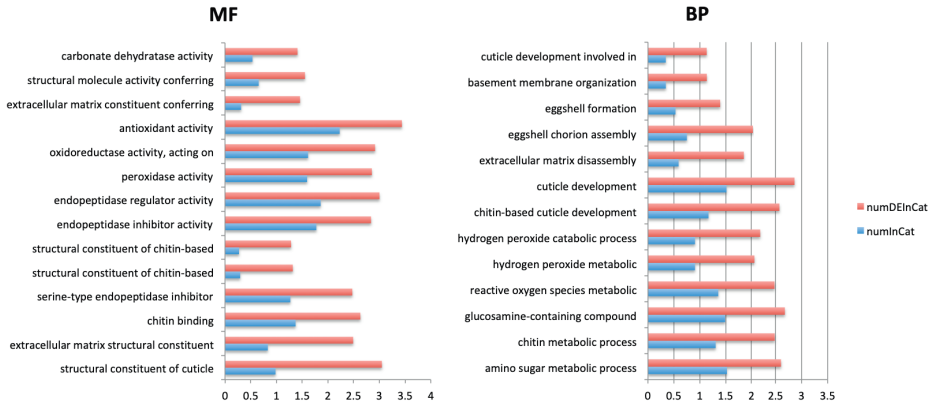


Figure 2. Gene ontology (GO) enrichment for the 1,260 genes upregulated in the AT stage compared to close search stage (enrichment factor >1.5; sorted by FDR). numInCat number of genes in the category, numDEInCat number of differentially expressed genes in the category; MF molecular function, BP biological process.

Table 2 Top 25 genes upregulated in the attached stage compared to close-search (FDR 0.05, FC 2), sorted by Fold change.

Gene	Annotation	Description	Fold Change	FDR
c188803_g1	RXFP2_MOUSE	Relaxin receptor 2	24.3	7.75E-06
c192671_g1	CUD4_LOCMI	Endocuticle structural glycoprotein ABD-4	23.3	2.19E-07
c178083_g1	CU16_MANSE	Larval cuticle protein 16/17	23.1	2.14E-05
c197858_g1	TENX_HUMAN	Tenascin-X	22.0	6.43E-06
c198520_g1	CAHZ_DANRE	Carbonic anhydrase	21.6	2.82E-06
c190658_g1	PERC_DROME	Chorion peroxidase	21.4	2.59E-06
c198871_g1	CAH2_RABIT	Carbonic anhydrase 2	21.0	5.22E-06
c189619_g2	PCPI_SABMA	Carboxypeptidase inhibitor SmCI	20.9	1.03E-05
c189619_g1	KCP_HALAI	BPTI/Kunitz domain-containing protein	20.4	2.83E-05
c196917_g1	CU22_BOMMO	Larval cuticle protein LCP-22	20.2	5.73E-06
c182182_g2	CUD1_SCHGR	Endocuticle structural glycoprotein SgAbd-1	20.2	5.67E-06
c198595_g1	CU1A_HOMAM	Cuticle protein AMP1A	20.1	3.04E-07
c183110_g1	CU20_TENMO	Adult-specific cuticular protein ACP-20	20.0	8.52E-06
c204253_g1	VPS62_YEAST	Vacuolar protein sorting-associated protein 62	20.0	6.50E-05
c193243_g1	RESIL_DROME	Pro-resilin	19.8	5.38E-05
c199490_g1	FRAS1_HUMAN	Extracellular matrix protein FRAS1	19.7	2.90E-06
c198991_g1	LDB3_MOUSE	LIM domain-binding protein 3	19.1	4.06E-06
c202872_g1	RPB1_PLAFD	DNA-directed RNA polymerase II subunit RPB1	19.0	4.81E-07
c207191_g1	ZN683_HUMAN	Tissue-resident T-cell transcription regulator protein ZNF683	18.8	9.70E-06
c178697_g1	CU27_MANSE	Pupal cuticle protein 27	18.7	0.000294851
c117818_g1	CUPA1_CANPG	Cuticle protein AM/CP1114	18.6	2.59E-06
c195607_g1	JHBP_MANSE	Juvenile hormone-binding protein	18.3	0.000155011
c175847_g2	CUP7_DROME	Pupal cuticle protein Edg-78E	18.1	6.48E-05
c201529_g2	CUD2_SCHGR	Endocuticle structural glycoprotein SgAbd-2	17.9	2.10E-05

Regulatory genes involved in settling and metamorphosis

As indicated above, we found 37 transcription factors highly induced in the transition to the attached stage. Several transcription factors involved in the ecdysone pathway were among these, such as the nuclear hormone receptors FTZ-F1, *HR3* and *HR38* (Figure 3). In this context, it is interesting to note that we also found a gene annotated as a member of the cytochrome P450 protein family to be strongly induced in the attached stage (Figure 3), which is a protein family known to be directly involved in ecdysone synthesis. All these genes were found to be transiently expressed in the attached stage and were in the juvenile stage back to roughly the levels in the cyprid stages (transient expression profile cluster V; Figures 1B and 3).

We also found in our dataset several other regulated genes to be key-genes the ecdysone cascade. These included the retinoid X receptor (Bimp_RXR), the ecdysone receptor (Bimp_EcR) and the broad complex protein (Bimp_BRC), all belonging to the nuclear receptors superfamily of transcription

factors. However, these were not highly induced during the AT stage but instead displayed some other type of regulation, i.e. Bimp_RXR were slightly but significantly upregulated in the AT stage, Bimp_BRC showed a drastic downregulation in AT while the expression of Bimp_EcR was downregulated in the JUV stage (Figure 3).

Since these are key-genes in the ecdysone cascade we wanted to firmly confirm their true homologies by analyses of their phylogeny. The Bimp_EcR protein had 50% identity to the *Daphnia magna* ecdysone receptor A2 (BAF49031.1), and contains both characteristic domains of this protein family, i.e. the ligand binding domain of the Ecdysone receptor (cd06938, E-value 5.59e-46, 62% identical to LBD in *D. magna*) and DNA-binding domain of the Ecdysone receptor family composed of two C4-type zink fingers (cd07161, E-value 2.26e-56, 90.1% identical to DBD in *D. magna*). In the phylogenetic tree, using Liver X Receptor (LXR), a member of nuclear receptor family from chicken *Gallus gallus* and rainbow trout *Oncorhynchus mykiss* as an outgroup, the here identified Bimp_EcR sequence clustered together with EcR from other crustaceans and insects (Figure 4).

The retinoid X receptor, RXR, in *B. improvisus* was 60% identical to the *D. magna* RXR protein (ABF74729.1) and contained both the DNA- and ligand-binding domains found in the other RXR receptor (3.51e-52 and 2.68e-98 NCBI CDD). The phylogenetic analysis including orthologs from other arthropod species, and using ecdysone receptors from *D. melanogaster* and *D. magna* as an outgroup, showed that Bimp_RXR truly belongs to the RXR family of the nuclear receptors (Figure S12-A).

We also identified a gene in *B. improvisus* containing both a BTB domain (E-value 1.78e-21) and several zinc finger domains and 65% identical (E-value 8e-49) to the broad-complex protein from Pacific beetle cockroach *Diploptera punctata* (AIM47237.1). Phylogenetic analysis including orthologous sequences from different crustaceans and insects representing different members of the BTB-ZF family, and using ecdysone receptors from *D. melanogaster*, *Daphnia* and kuruma shrimp *Penaeus japonicus* as an outgroup, showed that the *B. improvisus BR-C*-like gene clearly belong to the BTB-ZF family (Figure S12-B). We conclude that we have identified several members of the ecdysone cascade to be strongly to moderately regulated during the settlement stages.

In addition to genes involved in the ecdysone pathway, we observed that the key-enzyme involved in the biosynthesis of the methyl farnesoate hormone, was strongly regulated during settlement. The gene encoding the enzyme juvenile hormone acid methyltransferase (JHAMT), was highly expressed in the free-swimming and close-search stages and decreased after the attachment (Figure 3). In contrast, the juvenile hormone binding protein (JHBP) that resides in the hemolymph, where it binds to and transports the methyl farnesoate hormone (Hao et al. 2013, Sin et al. 2015), was specifically upregulated after the attachment and then strongly down-regulated in the juvenile stage.

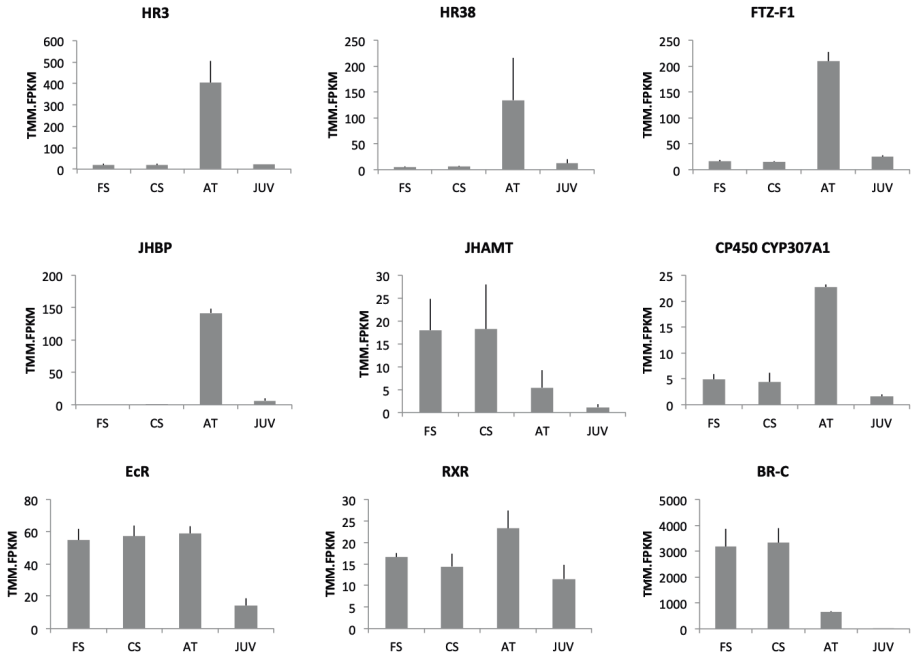


Figure 3 Expression of the *B. improvisus* molting and metamorphosis candidate genes: HR3 hormone receptor 3, HR38 hormone receptor 38, FTZ-F1 fushi tarazu transcription factor 1, JHBP juvenile hormone binding protein, JHAMT juvenile hormone acid methyltransferase; CP450 cytochrome P450, EcR ecdysone receptor, RXR retinoid X receptor, BR-C broad complex. TMM.FPKM: Trimmed Mean of M-values and Fragments per kilobase per million mapped fragments normalized counts. FS free-swimming, CS close-search, AT attached, JUV juvenile.

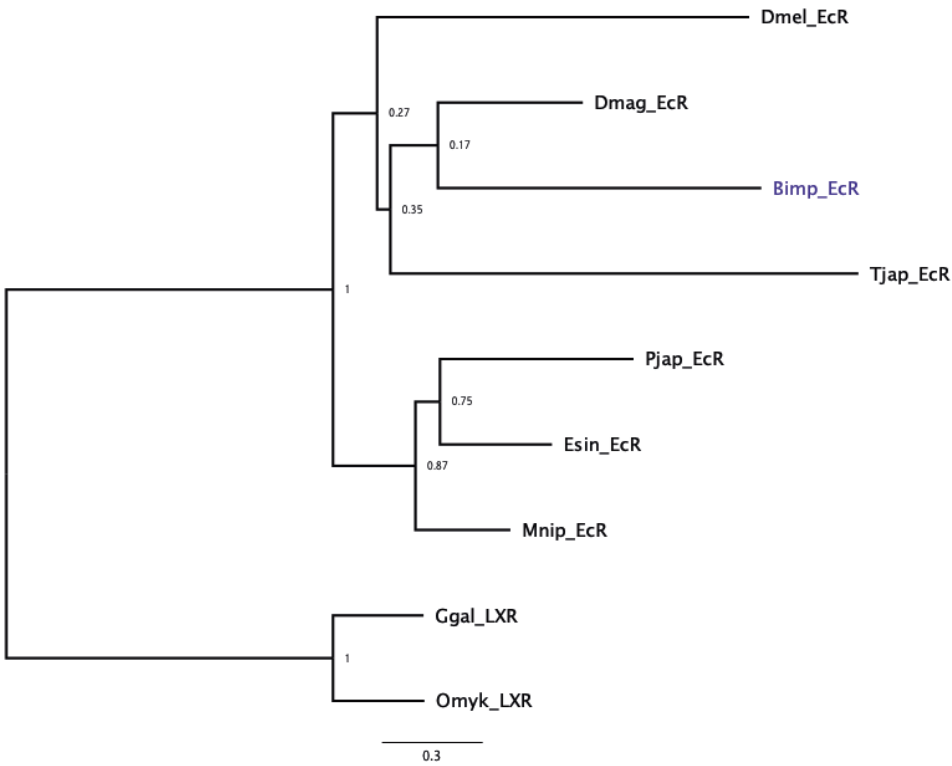


Figure 4 Maximum-likelihood phylogeny of *B. improvisus* candidate ecdyson receptor. Dmel *D. melanogaster*, Dmag *D. magna*, Tjap *Tigriopus japonicus*, Pjap *Penaeus japonicus,* Esin *Eriocheir sinensis*, Mnip *Macrobrachium nipponense*, Ggal *Gallus gallus*, Omyk *Oncorhynchus mykiss*. The numbers indicate bootstrap values for each branch.

Differential developmental gene expression patterns during settlement on surfaces with different wettability

To get a better mechanistic view on how cyprids perceive and respond to different surfaces, we examined differential gene expression on two type-surfaces, hydrophilic and hydrophobic. At the start of the experiment, the hydrophobic surface had a contact angle of 81.1° while the hydrophilic had a contact angle of 51.4° (Dahlström et al. 2004), and this difference in wettability was maintained during the course of the experiment (Figure S6). Even if this difference in wettability could be regarded as rather minor, cyprid settlement has earlier been shown to be significantly different between the two (Dahlström et al. 2004).

We observed a great number of significant differences in gene expression between the two surfaces in all settlement stages (Table 2). There is an almost equal number of genes being preferentially expressed (roughly 90) on either of the two surfaces in the free-swimming stage. The number of differentially expressed genes between the two surfaces is the highest in the close-search and attached stages and decreased again in the juvenile stage, e.g. 250 genes were upregulated on the hydrophobic compared to the hydrophilic surface in the close-search stage. This shows that most of the surface-specific responses occurred mainly around the time of attachment, with the gene expression differences being less pronounced when metamorphosis into juveniles was completed.

Table 2 The number of genes differentially expressed comparing separately each settlement stage between the two type surfaces (FDR < 0.05, no FC filtering) (all genes=ann+unann).

		Number of upregulated genes		
Settlement stage	Total number of DEGs	Hydrophobic	Hydrophilic	% non annotated (functionally unknown)
Free-swimming	179	94	85	46
Close-search	303	250	53	49
Early-attached	265	240	26	55
Juvenile	148	73	75	55

In addition, we also compared the serial transition profiles of gene expression on the two type of surfaces. First of all, it should be pointed out that for neither the hydrophobic nor the hydrophilic surface we found any significant changes (FDR < 0.05, FS > 2) when comparing free-swimming to close-stages. Comparison of the number of transition-specific genes between the other stages on the two surfaces revealed that the number of genes and their magnitude of response are considerably larger (both for increase and decrease in expression) for the hydrophobic surface than for the hydrophilic surface (Figure 5A and B). In particular we found this to be true for the transition to the attached stage; among the genes more highly expressed in the attached stage compared to close-search on the two surfaces, we saw that 27% of all regulated genes showed a greater than 2 fold upregulation on the hydrophobic surface while only 4% genes exhibited this level of regulation on the hydrophilic surface. We conclude that the wettability of the surface has a pronounced impact on the gene expression programmes during settling in *B. improvisus*.

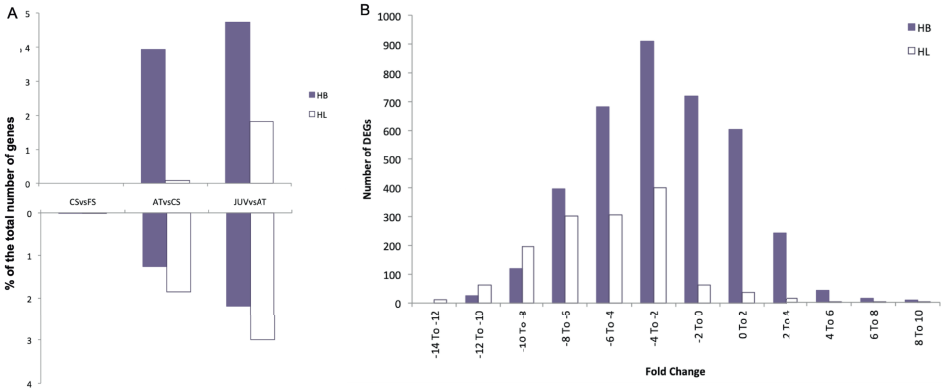


Figure 5 Comparison of gene expression on the two type-surfaces: A) The number of regulated genes (FDR < 0.05, FC > 2) between each stage transition (transition-specific genes) on hydrophobic (HB) and hydrophilic (HL) surfaces (percentage from the total number of differentially expressed genes, DEGs); B) Fold changes range observed on the two surfaces during close-search to attached stage transition (FDR < 0.05).

Roughly half of the differentially expressed genes between the two surfaces received no functional annotation, indicating that these settling-important genes are orphan genes not earlier characterised (Table 2). Among the surface type-specific genes that were annotated, we observed several genes related to signalling transduction, such as calmodulin, MFS-type transporter guanylate-cyclase and membrane-associated guanylate kinase to be more highly expressed on the hydrophobic surface (Table S8-11). Furthermore, genes related to the hormonal regulation of molting and metamorphosis

were among the highly regulated genes on both surfaces. In particular, the broad complex protein *BRC*, one of the key regulators in the ecdyson cascade, was downregulated 1.4 times on the hydrophilic surface in the free-swimming stage (Table S8). Proteins involved in the neuropeptide levels/activity control, i.e. juvenile hormone binding proteins and neprilysins, were more than four times upregulated on the hydrophilic surface. Neprilysins are metalloendopeptidases that were identified in diverse groups of invertebrates, including insects and annelids (Turner et al. 2001). Based on studies in mollusc *Aplysia californica*, it has been suggested that neprilysins are involved in metabolism of synaptically released neuropeptides, inactivating peptide transmitters and modulators. Recent transcriptomic study of neuropeptides and peptide hormones in *B. amphitrite* showed that some of them are specifically expressed in the cyprid stage and are important for settlement (Yan et al. 2012). Neprilysin-21 is 27 fold more expressed on the hydrophilic surface and might be involved in the peptidergic neurotransmission during settlement under these conditions.

Discussion

Gene expression during *B. improvisus* settlement

A persistent question in barnacle biology is how cyprids 'decide' to settle. Experimental studies show that cyprids actively reject unfavourable settlement spots and display different behavioural responses on different types of surfaces. Surface chemistry, surface structure and the presence of conspecific biogenic cues are among the most important settlement factors (Matsumura et al. 1998, Berntsson et al. 2000, Dahlström et al. 2000, Dreanno et al. 2006b). The diverse nature of the settlement cues suggests that during settlement larvae rely on the combined input from at least three sensory systems: a mechanosensory, most likely represented by the numerous setae of the antennules (Maruzzo et al. 2011); a chemosensory, that includes lattice organs and sensory setae of the antennules (Bielecki et al. 2009, Maruzzo et al. 2011); and a photosensory, represented by the pair of cyprid compound eyes (Matsumura and Qian 2014). However, how the signals are integrated and transformed into the cyprid's decision to settle at a specific spot still remains unknown.

The close-search and early attached stages are of particular interest when trying to understand early events in the cyprids' decision-making. In our experimental design, we therefore included both free-swimming and close-search stages to compare cyprids during no-contact with the surface and cyprid with the immediate contact with the surface and the closest time-point to the attachment. However, comparison of gene expression profiles between free-swimming and close-search stages did not detect any significantly regulated genes on any of the two types of surfaces. Quantitative analysis shows that cyprids often return to the water column in the transition between wide-search and close-search (Aldred et al. 2018). Therefore, rapid shifts between the free-swimming and close-search stages could possibly obscure any distinct difference in gene expression between these two stages, which might be the reason we could not score these stages to be distinct. A limitation in our study was the amount of RNA needed for sequencing-library preparation and we thus had to pool 20 individuals

for each stage. This made it practically challenging to follow and examine close-search individuals after certain times of exploration. With future technical developments where individual cyprid's RNA would suffice, more stage-specific unique expression profiles in these interesting surface-exploratory phases could be studied.

The first hypothetical scheme of molecular mechanisms of barnacle settlement was proposed by Clare (1996), based on comparative studies from other marine larvae and pharmacological data from barnacles. According to their scheme, external cues, i.e. pheromones, initiate attachment by binding to G-protein coupled receptors that activate cAMP and increase intracellular calcium resulting in depolarisation of membrane and propagation of signal to the central nervous system. Metamorphosis on the other hand is suggested to be controlled through the phosphatidylinositol pathway involving protein kinase C (PKC) that regulates methyl farnesoate levels. Therefore, cAMP is suggested to act proximal to the receptors while phosphatidylinositol pathway is distant (Clare and Matsumura 2000).

We here report enrichment in the genes related to the "ion channel activity" GO category in the cyprid stages, among which are ionotropic receptors and transient receptor potential channels that have previously been identified in the cyprids antennules and suggested to be candidates for sensing settlement cues (Abramova et al. 2019). Both types of receptors act as ion channels and upon binding a ligand directly transduce the chemical signal into a change in flux across the membrane, acting independently of G-protein signaling. Signals from the TRP channels can also be transduced through phosphatidylinositol pathway, which has been shown to be involved in the olfactory receptor cell chemosensory transduction (Ache and Young 2005) in other crustaceans. A recent paper by Kotsiri et al., 2018 suggested that mechanosensory TRP channels mediate surface exploration in *B. amphitrite* cyprids. We observed enrichment for the genes in the "phosphatidylinositol-mediated signaling" GO category in the cyprid stages suggesting that this pathway could be also involved in the transduction of signals from settlement cues during exploratory stage of settlement. Settlement and metamorphosis in red abalone *Haliotis rufescens* are regulated by two convergent chemosensory pathways, morphogenetic through cAMP and regulatory through phosphatidylinositol pathway (Morse 1993). It has been shown that activation of the regulatory pathway increase post-receptor responsiveness to the settlement cue, suggesting interaction of different signaling pathways activated by different settlement cues (Morse 1993). Despite that our data on gene expression changes provide rather indirect evidence of similar mechanisms involved in barnacle settlement, they do support the idea that combined input from various types of receptors acting through different signalling pathways could provide a basis for fine-tuned recognition of the suitable settlement spot.

Working models for regulatory mechanisms during barnacle attachment

Signals from settlement cues in barnacles are suggested to be decoded at the site of methyl farnesoate synthesis (Kotsiri et al. 2018). Synthesised and secreted methyl farnesoate is transported to the site of ecdysteroids production, where it suppresses the surge of ecdysteroids and thereby initiates molting and metamorphosis. Hormonal regulation of molting in arthropods is initiated by a sharp increase in ecdysone titre that triggers a cascade of tissue-specific gene expression involving a hierarchy of ecdysone-responsive genes, i.e. the ecdysone cascade (Hyde et al. 2019a). In comparison to insects, the ecdysone cascade in crustaceans, and barnacles in particular, is rather poorly characterised. However, the majority of gene homologs involved in biosynthetic and signalling pathways of ecdysteroids and sesquiterpenoids can be identified in non-insect arthropods implying that these two hormonal systems were present in the last common ancestor of arthropods (Qu et al. 2015). Indeed, several of the key hormones, such as 20-hydroxyecdysone (20E) and methyl farnesoate (MF), were identified in barnacles (Yamamoto et al. 1997a, Yamamoto et al. 1997b). We here show that several of the genes encoding the involved enzymes in the production and handling of these hormones were strongly regulated during the initial stages of settling. This provides additional support that these genes are involved in attachment and early development programmes also in barnacles. Furthermore, it has also been suggested that the process of surface exploration and metamorphosis in barnacles have certain resemblance to wandering and metamorphosis in insects, with cyprids in this analogy corresponding to the pupal stage (Kotsiri et al. 2018). In the current study, we used knowledge from insects and other crustaceans as a framework and identified several key genes of the ecdysone cascade differentially regulated during *B. improvisus* settlement process.

Molting in arthropods is initiated by a sharp increase in ecdysone that bind to the heterodimer complex of the ecdysone receptor and the retinoid-X-receptor (EcR/RXR). The activated EcR/RXR complex then directly activates transcription of the early ecdysone-response genes, e.g. Broad-Complex (*BR-C*). In turn, this gene triggers tissue-specific expression of the late ecdysone-response genes coordinating molting and metamorphosis (Bayer et al. 1996). We identified and characterised several genes involved in the ecdysone cascade to be regulated during settling, including EcR, RXR, BR-C*,* FTZ-F1, HR3 and HR38 belonging to the nuclear receptors superfamily. Both Bimp_EcR and Bimp_RXR were found to be highly similar to the known sequences of EcR and RXR and appear to be a components of the ecdysone receptor complex in *B. improvisus*. It should be mentioned though, that it is characteristic for both receptors to have multiple variants originating by alternative splicing and differentially expressed in a tissue-specific manner, as has been shown in several crustaceans (Asazuma et al. 2007, Techa and Chung 2013). Therefore, further studies are required to describe the variants of the EcR and RXR and investigate their differences in ligand binding and heteromerization.

The broad complex is a key-factor in development in arthropods. We identified a BR-C-like gene that was strongly down-regulated in the attached stage in *B. improvisus.* The barnacle BR-C protein possesses high similarity to the known BR-C sequences in other organisms, and contains both

characteristic domains, i.e. the BTB/POZ and carboxyl-terminal zinc finger (ZnF) C2H2 domains. *BR-C* is differentially regulated during metamorphosis in insects and and high expression of this protein correlates with the time of pupal commitment, a state of cellular differentiation when tissues become responsive to the hormones triggering metamorphosis (Bayer et al. 1996, Zhou et al. 1998). Little is known about *BR-C* in crustaceans, however, the presence of a *BR-C* gene(s) was earlier confirmed in the shrimp *Penaeus monodon* where *BR-C* is suggested to play an important role in reproductive tissue development (Buaklin et al. 2013, Jiang et al. 2015). Importantly, *BR-C* is also responsible for initiation of expression of the glue genes in the crawling *Drosophila* larva, which it uses to attach to the surface before onset of pupation (Kaieda et al. 2017). Interestingly, the majority of cement proteins in the current experiment started to be expressed after the attachment to the surface (Figure S7). The adhesion preceding major morphological rearrangements and similarity in the hormonal regulation in both crustaceans and insects bring the possibility that cyprid cement genes might be also activated through a similar cascade involving *BRC*, however, it has to be further elucidated experimentally.

Bimp_EcR was highly expressed during all stages except in the juvenile, while the enhanced expression peak of RXR was observed in the attached stage (Figure 3). BRC was highly expressed in free-swimming and close-search stages and down-regulated after the attachment. We found several early ecdysone-responsive genes to be strongly and specifically (transiently) up-regulated in the attached stage, including the nuclear *HR3* and the *HR38*. In addition, the nuclear hormone receptor *Ftz-f1* was up-regulated (Figure 3), which has been shown to be necessary for the expression of late ecdysone-responsive genes in *Drosophila* (Yamada et al. 2000, Sultan et al. 2014). Importantly, we observed 164 genes that displayed a transient expression profile with a drastic increased expression in attached stage and then roughly cyprid-level of expression again in the juvenile stage (clusters V; Figure 1B); Thus, these genes might reflect a large-scale transcriptomic shift in *B. improvisus* in response to the ecdysone signal. It has to be emphasised that a time window from the moment of cyprid adhesion to the completion of metamorphosis takes approximately 24 hours and includes at least four morphologically different stages as described in Maruzzo et al. (2012). Therefore, our sampling in the attached stage represents only a snapshot of the events some hours after the attachment and most probably is an underestimation of the number of genes involved in the whole metamorphic transition. A more frequent, time-resolved sampling is required to track the whole repertoire of transient expression profiles of the transcriptional changes during the attachment and metamorphosis. Despite this limitation, the analysis of genes upregulated in the attached stage allows to get an overview of the early events during the morphogenetic transition. GO enrichment analysis of upregulated genes in the early attached stage showed enrichment of GO terms related to cuticle development and chitin metabolic process as well as antioxidant activity, peroxidase activity and reactive oxygen species metabolic process. A recent study showed that in the adult barnacle, molting and adhesion processes are tightly connected. In particular, the authors reported that the adhesive curing occurs under action of the same molecules and enzymes that are present in the molting fluid, including reactive oxygen species, oxidases and peroxidases (So et al. 2017). As an interesting and

related example during cyprid settling, we observed a roughly 1,000-fold upregulation of chorion peroxidase, 6-fold for dual oxidase and 56-fold for peroxidase, in the attached stage (Table 2). As was discussed in So et al. (2017), chorion peroxidases have been shown to be involved in the cross-linking of adhesives from caddisfly larva and egg shell hardening in *Aedes aegypti* (Li et al. 1996, Wang et al. 2015a) thus indicating that they could play a similar role in barnacles. Our data on extreme up-regulation of chorion peroxidase in the attached stage suggest that this enzyme take part in the adhesion process also in *B. improvisus*.

We also found several key enzymes involved in hormonal synthesis pathways. Juvenile hormone acid methyltransferase JHAMT, the main enzyme involved in the MF biosynthesis, was highly expressed in the free-swimming and close-search stages in *B. improvisus* and decrease after attachment (Figure 3). Furthermore, another gene specifically upregulated after the attachment was annotated as hemolymph juvenile hormone binding protein (JHBP) that function as a methyl farnesoate binding protein (Qu et al. 2015). These binding proteins are involved in transportation and modulation of the methyl farnesoate titre in hemolymph of the crab *Cancer magister* and the shrimp *Neocarida denticulata* (Tamone et al. 1997, Sin et al. 2015). All together, this could suggest that synthesis of methyl farnesoate and its binding proteins required for its transport to the target ecdysteroid-secreting tissues are temporarily separated, which ensures that methyl farnesoate exerts its regulatory function after the attachment. Interestingly, proteins from the cytochrome CP450 protein family, are one of the important genes in the ecdysone synthesis pathway, is upregulated specifically after the attachment.

Highly distinct settlement programmes on different types of surfaces

Previous settlement of *B. improvisus* cyprids on surface with different wettabilities revealed that settlement inversely correlates with increased wettability, and highly wettable surfaces totally inhibit cyprid settlement (Dahlström et al. 2004). It was suggested that wettability might affect settlement either through cyprid chemical recognition of the molecular groups on the surface or physical forces, such as electrostatic repulsion, can impede cyprid contact with the surface (Dahlström et al. 2004). We observed a large number of differences in the gene expression on the two surfaces suggesting that surface wettability acts as a settlement cue through sensory recognition and initiation of regulatory events that leads to changes in the physiology of *B. improvisus*, rather than purely due to electrophysical repulsion forces. This is further supported by the fact that over 30% of the annotated genes more highly expressed on the hydrophobic surface compared to on the hydrophilic surface in the close-search stage are related to signaling transduction, e.g. G-protein coupled receptors, MFS-type transporters and phosphatidylinositol kinases. In addition, the regulatory programmes initiated in the attached stage are very distinct on the two surfaces, which further support that the cyprids' physiology is highly influenced by the wettability of the surface. It should be kept in mind that the analysed cyprids on the hydrophilic surface have entered the attached stage (the same type of

morphological change appear on both surfaces), indicating that attachment and initial metamorphoses can proceed via different routes. Interestingly, later during the development when metamorphosis are partly completed and the organisms have entered the juvenile stage, the differences in gene expression on the two surfaces are less pronounced indicating that later developmental programmes proceed via similar paths.

Studies in gastropod *Haliotis asinina* settlement suggesting that settlement cues differ in terms of their inducing effectiveness, with a more effective cue triggering direct and quick pathways leading to more rapid metamorphosis than for a less effective cue (Williams and Degnan 2009). In the case of barnacle cyprid larvae, experimental evidence revealed that cyprids settle faster on hydrophobic than hydrophilic surface, suggesting that the former one is a more effective cue (Dahlström et al. 2004). Our results showed five-times more genes upregulated on the hydrophobic surface than on hydrophilic in the close-search stage, and 10-times more genes in the attached stage, suggesting that exploration of the "favourable" surface appears to induce a stronger and/or earlier transcriptomic response. The hierarchical importance of the settlement cues in barnacles is not yet clear. However, if pheromones analogous to morphogenetic cue as GABA-mimetic peptides in red abalone (Morse and Morse 1984), the surface chemistry can play a role of a regulatory cue. Therefore, detecting a hydrophobic favorable surface might act by increasing the responsiveness of larvae to other morphogenetic cues, e.g. pheromones, and increase the propensity of larvae to settle. This is supported by the stronger transcriptomic response both in number of genes and their responses on the favourable surface. This includes roughly three-fold higher expression of the key gene in the ecdysone cascade on the hydrophobic surface i.e. *BR-C* gene, as well as genes involved in major developmental pathways, including Hedgehog transcription factors and components in EGF signalling.

Conclusions

In summary, we provide the first description of the ecdysone cascade genes and their expression during settlement of *B. improvisus*. As a follow up, more time-resolved studies have to be performed in order to elucidate the exact expression patterns and roles of these genes in the barnacle molting and metamorphosis. Furthermore, we provide an overview of the global transcriptional changes during settlement, including free-swimming and close search stages. Identification of different types of receptors and signalling pathways suggests that barnacles could also have distinct morphogenetic and regulatory pathways, as was previously described for abalone. Gene expression on surfaces with different wettabilities further provide support for the hypothesis of chemical recognition of the surface chemistry by the larvae. Furthermore, the gene expression response on the hydrophobic surface suggests that exploration of the "favourable" surface might induce stronger and/or earlier transcriptomic response.

Declarations

Ethics approval

There are no specific guidelines and permits required for the use of barnacles in scientific experiments in Sweden.

Competing interests

There are no competing interests.

Funding

Funding for this project was provided to AB by the EU commission, SEAFRONT— Synergistic Fouling Control Technologies; #614034.

Acknowledgements

The authors would like to acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure, NGI, and Uppmax for providing assistance in massive parallel sequencing and computational infrastructure, respectively. The authors would also like to thank Tomas Larsson for the help with bioinformatics analysis, Martin Ogemark for maintaining the *B. improvisus* culture and providing cyprid larvae, Mattias Berglin for technical assistance with wettability measurements, and Per Jonsson for valuable comments on the manuscript.

List of abbreviations

20E	20-hydroxyecdysone
AT	Attached stage
BP	Biological process
BRC	Broad complex
BTB-ZF	Broad-complex, tramtrack and bric-à-brac - zinc finger
cAMP	Cyclic AMP
CDD	Conserved Domain Database
CS	Close-search stage
DBD	DNA binding domain
DEG	Differentially expressed gene
EcR	Ecdysone receptor
EGF	Epidermal growth factor
FC	Fold change
FDR	False discovery rate
FPKM	Fragments per Kilobase of transcript per Million mapped reads
FS	Free-swimming stage
FTZ-F1	Fushi tarazu transcription factor 1
GABA	Gamma-Aminobutyric Acid
GO	Gene ontology
GPCR	G-protein coupled receptor
HB	Hydrophobic surface
HL	Hydrophilic surface
HR3	Hormone receptor 3
HR38	Hormone receptor 38
IR	Ionotropic receptor
JHAMT	Juvenile hormove acid methyltransferase
JHBP	Juvenile hormone binding protein
JUV	Juvenile stage
LBD	Ligand binding domain
MF	Methyl farnesoate
MF	Molecular function
MFS	Major facilitator superfamily transporter
NCBI	National Center for Biotechnology Information
ORF	Open reading frame
PKC	Protein kinase C
RXR	Retinoid X receptor
TMM	Trimmed Mean of M values
TRP	Transient receptor potential channel

References

Abramova, A., M. A. Rosenblad, A. Blomberg and T. A. Larsson (2019). "Sensory receptor repertoire in cyprid antennules of the barnacle Balanus improvisus." PloS one **14**(5): e0216294.

Ache, B. W. and J. M. Young (2005). "Olfaction: diverse species, conserved principles." Neuron **48**(3): 417-430.

Aldred, N., A. Alsaab and A. S. Clare (2018). "Quantitative analysis of the complete larval settlement process confirms Crisp's model of surface selectivity by barnacles." Proc. R. Soc. B **285**(1872): 20171957.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data, Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

Asazuma, H., S. Nagata, M. Kono and H. Nagasawa (2007). "Molecular cloning and expression analysis of ecdysone receptor and retinoid X receptor from the kuruma prawn, Marsupenaeus japonicus." Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology **148**(2): 139-150.

Bayer, C. A., B. Holley and J. W. Fristrom (1996). "A Switch inBroad-ComplexZinc-Finger Isoform Expression Is Regulated Posttranscriptionally during the Metamorphosis ofDrosophilaImaginal Discs." Developmental biology **177**(1): 1-14.

Berntsson, K. M. and P. R. Jonsson (2003). "Temporal and spatial patterns in recruitment and succession of a temperate marine fouling assemblage: a comparison of static panels and boat hulls during the boating season." Biofueling **19**(3): 187-195.

Berntsson, K. M., P. R. Jonsson, M. Lejhall and P. Gatenholm (2000). "Analysis of behavioural rejection of micro-textured surfaces and implications for recruitment by the barnacle Balanus improvisus." Journal of Experimental Marine Biology and Ecology **251**(1): 59-83.

Bielecki, J., B. K. K. Chan, J. T. Hoeg and A. Sari (2009). "Antennular sensory organs in cyprids of balanomorphan cirripedes: standardizing terminology using Megabalanus rosa." Biofouling **25**(3): 203-214.

Bray, N. L., H. Pimentel, P. Melsted and L. Pachter (2016). "Near-optimal probabilistic RNA-seq quantification." Nature Biotechnology **34**(5): 525-527.

Buaklin, A., K. Sittikankaew, B. Khamnamtong, P. Menasveta and S. Klinbunga (2013). "Characterization and expression analysis of the Broad-complex (Br-c) gene of the giant tiger shrimp Penaeus monodon." Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology **164**(4): 280-289.

Chen, Z.-F., K. Matsumura, H. Wang, S. M. Arellano, X. Yan, I. Alam, J. A. Archer, V. B. Bajic and P.-Y. Qian (2011). "Toward an understanding of the molecular mechanisms of barnacle larval settlement: a comparative transcriptomic approach." PLoS One **6**(7): e22913.

Clare, A. (1996). "Signal transduction in barnacle settlement: calcium re-visited." Biofouling **10**(1-3): 141-159.

Clare, A., R. Thomas and D. Rittschof (1995). "Evidence for the involvement of cyclic AMP in the pheromonal modulation of barnacle settlement." Journal of Experimental Biology **198**(3): 655-664.

Clare, A. S. and K. Matsumura (2000). "Nature and perception of barnacle settlement pheromones." Biofouling **15**(1-3): 57-71.

Crisp, D. (1974). "Factors influencing the settlement of marine invertebrate larvae." Chemoreception in marine organisms: 177-265.

Dahlström, M., H. Jonsson, P. R. Jonsson and H. Elwing (2004). "Surface wettability as a determinant in the settlement of the barnacle Balanus improvisus (Darwin)." Journal of Experimental Marine Biology and Ecology **305**(2): 223-232.

Dahlström, M., L. G. Mårtensson, P. R. Jonsson, T. Arnebrant and H. Elwing (2000). "Surface active adrenoceptor compounds prevent the settlement of cyprid larvae of Balanus improvisus." Biofouling **16**(2-4): 191-203.

Davidson, I., C. Scianni, C. Hewitt, R. Everett, E. Holm, M. Tamburri and G. Ruiz (2016). "Mini-review: Assessing the drivers of ship biofouling management–aligning industry and biosecurity goals." Biofouling **32**(4): 411-428.

Dreanno, C., R. R. Kirby and A. S. Clare (2006a). "Smelly feet are not always a bad thing: the relationship between cyprid footprint protein and the barnacle settlement pheromone." Biology letters **2**(3): 423-425.

Dreanno, C., K. Matsumura, N. Dohmae, K. Takio, H. Hirota, R. R. Kirby and A. S. Clare (2006b). "An alpha2-macroglobulin-like protein is the cue to gregarious settlement of the barnacle Balanus amphitrite." Proc Natl Acad Sci U S A **103**(39): 14396-14401.

Furman, E. R. and A. B. Yule (1990). "Self-fertilisation in Balanus improvisus Darwin." Journal of Experimental Marine Biology and Ecology **144**(2-3): 235-239.

Haas, B. J., A. Papanicolaou, M. Yassour, M. Grabherr, P. D. Blood, J. Bowden, M. B. Couger, D. Eccles, B. Li, M. Lieber, M. D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C. N. Dewey, R. Henschel, R. D. Leduc, N. Friedman and A. Regev (2013). "De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis." Nature Protocols **8**(8): 1494-1512.

Hao, W., Y. Zhang and Y.-s. Xu (2013). "Identification of a juvenile hormone esterase binding protein gene and its developmental and hormone regulation in the silkworm, Bombyx mori." Journal of insect physiology **59**(9): 906-912.

Hyde, C. J., A. Elizur and T. Ventura (2019a). "The crustacean ecdysone cassette: a gatekeeper for molt and metamorphosis." The Journal of steroid biochemistry and molecular biology **185**: 172-183.

Hyde, C. J., Q. P. Fitzgibbon, A. Elizur, G. G. Smith and T. Ventura (2019b). "Transcriptional profiling of spiny lobster metamorphosis reveals three new additions to the nuclear receptor superfamily." BMC genomics **20**(1): 531.

Jaszczak, J. S., J. B. Wolpe, R. Bhandari, R. G. Jaszczak and A. Halme (2016). "Growth coordination during Drosophila melanogaster imaginal disc regeneration is mediated by signaling through the relaxin receptor Lgr3 in the prothoracic gland." Genetics **204**(2): 703-709.

Jiang, S., Y. Zhang, S. Sun, Y. Gong, Y. Xiong, H. Qiao, W. Zhang, S. Jin and H. Fu (2015). "Molecular cloning, characterization, and expression analysis of a Broad-Complex homolog during development in the oriental river prawn Macrobrachium nipponense." Genetics and Molecular Research **14**(2): 5141-5152.

Jindra, M., S. R. Palli and L. M. Riddiford (2013). "The juvenile hormone signaling pathway in insect development." Annual review of entomology **58**: 181-204.

Jonsson, P. R., A.-L. Wrange, U. Lind, A. Abramova, M. Ogemark and A. Blomberg (2018). "The barnacle Balanus improvisus as a marine model-culturing and gene expression." JoVE (Journal of Visualized Experiments)(138): e57825.

Judge, M. L. and S. F. Craig (1997). "Positive flow dependence in the initial colonization of a fouling community: results from in situ water current manipulations." Journal of Experimental Marine Biology and Ecology **210**(2): 209-222.

Kaieda, Y., R. Masuda, R. Nishida, M. Shimell, M. B. O'Connor and H. Ono (2017). "Glue protein production can be triggered by steroid hormone signaling independent of the developmental program in Drosophila melanogaster." Developmental biology **430**(1): 166-176.

Kotsiri, M., M. Protopapa, G. M. Roumelioti, A. Economou-Amilli, E. Efthimiadou and S. Dedos (2018). "Probing the settlement signals of Amphibalanus amphitrite (vol 34, pg 492, 2018)." Biofouling **34**(5): X-X.

Lagersson, N. C. and J. T. Hoeg (2002). "Settlement behavior and antennulary biomechanics in cypris larvae of Balanus amphitrite (Crustacea : Thecostraca : Cirripedia)." Marine Biology **141**(3): 513-526.

Lang, W. H., R. B. Forward Jr and D. C. Miller (1979). "Behavioral responses of Balanus improvisus nauplii to light intensity and spectrum." The Biological Bulletin **157**(1): 166-181.

Law, C. W., M. Alhamdoosh, S. Su, X. Dong, L. Tian, G. K. Smyth and M. E. Ritchie (2016). "RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR." F1000Research **5**.

Li, J., B. A. Hodgeman and B. M. Christensen (1996). "Involvement of peroxidase in chorion hardening in Aedes aegypti." Insect biochemistry and molecular biology **26**(3): 309-317.

Li, W. and A. Godzik (2006). "Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences." Bioinformatics **22**(13): 1658-1659.

Liu, Y., J. Zhou and K. P. White (2013). "RNA-seq differential expression studies: more sequence or more replication?" Bioinformatics **30**(3): 301-304.

Maruzzo, D., N. Aldred, A. S. Clare and J. T. Høeg (2012). "Metamorphosis in the cirripede crustacean Balanus amphitrite." PloS one **7**(5): e37408.

Maruzzo, D., S. Conlan, N. Aldred, A. S. Clare and J. T. Hoeg (2011). "Video observation of surface exploration in cyprids of Balanus amphitrite: the movements of antennular sensory setae." Biofouling **27**(2): 225-239.

Matsumura, K., M. Nagano and N. Fusetani (1998). "Purification of a larval settlement-inducing protein complex (SIPC) of the barnacle, Balanus amphitrite." Journal of Experimental Zoology **281**(1): 12-20.

Matsumura, K. and P.-Y. Qian (2014). "Larval vision contributes to gregarious settlement in barnacles: adult red fluorescence as a possible visual signal." Journal of Experimental Biology **217**(5): 743-750.

Michels, J., E. Appel and S. N. Gorb (2016). "Functional diversity of resilin in Arthropoda." Beilstein journal of nanotechnology **7**(1): 1241-1259.

Morse, A. N. and D. E. Morse (1984). GABA-mimetic molecules from Porphyra (Rhodophyta) induce metamorphosis of Haliotis (Gastropoda) larvae. Eleventh International Seaweed Symposium, Springer.

Morse, D. E. (1993). "Signalling in planktonic larvae." Nature **363**(6428): 406.

Pérez-Losada, M., M. Harp, J. T. Høeg, Y. Achituv, D. Jones, H. Watanabe and K. A. Crandall (2008). "The tempo and mode of barnacle evolution." Molecular phylogenetics and evolution **46**(1): 328-346.

Qu, Z., N. J. Kenny, H. M. Lam, T. F. Chan, K. H. Chu, W. G. Bendena, S. S. Tobe and J. H. L. Hui (2015). "How did arthropod sesquiterpenoids and ecdysteroids arise? Comparison of hormonal pathway genes in noninsect arthropod genomes." Genome biology and evolution **7**(7): 1951-1959.

Ritchie, M. E., B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi and G. K. Smyth (2015). "limma powers differential expression analyses for RNA-sequencing and microarray studies." Nucleic acids research **43**(7): e47-e47.

Robinson, M. D., D. J. McCarthy and G. K. Smyth (2010). "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data." Bioinformatics **26**(1): 139-140.

Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva and E. M. Zdobnov (2015). "BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs." Bioinformatics **31**(19): 3210-3212.

Sin, Y. W., N. J. Kenny, Z. Qu, K. W. Chan, K. W. Chan, S. P. Cheong, R. W. Leung, T. F. Chan, W. G. Bendena and K. H. Chu (2015). "Identification of putative ecdysteroid and juvenile hormone pathway genes in the shrimp Neocaridina denticulata." General and comparative endocrinology **214**: 167-176.

Smith, P. A., A. S. Clare, H. H. Rees, M. C. Prescott, G. Wainwright and M. C. Thorndyke (2000). "Identification of methyl farnesoate in the cypris larva of the barnacle, Balanus amphitrite, and its role as a juvenile hormone." Insect biochemistry and molecular biology **30**(8-9): 885-890.

Smith-Unna, R., C. Boursnell, R. Patro, J. M. Hibberd and S. Kelly (2016). "TransRate: reference-free quality assessment of de novo transcriptome assemblies." Genome Research **26**(8): 1134-1144.

So, C. R., J. M. Scancella, K. P. Fears, T. Essock-Burns, S. E. Haynes, D. H. Leary, Z. Diana, C. Y. Wang, S. North, C. S. Oh, Z. Wang, B. Orihuela, D. Rittschof, C. M. Spillmann and K. J. Wahl (2017). "Oxidase Activity of the Barnacle Adhesive Interface Involves Peroxide-Dependent Catechol Oxidase and Lysyl Oxidase Enzymes." Acs Applied Materials & Interfaces **9**(13): 11493-11505.

Sultan, A. R. S., Y. Oish and H. Ueda (2014). "Function of the nuclear receptor FTZ-F1 during the pupal stage in D rosophila melanogaster." Development, growth & differentiation **56**(3): 245-253.

Tamone, S. L., G. D. Prestwich and E. S. Chang (1997). "Identification and characterization of methyl farnesoate binding proteins from the crab, Cancer magister." General and comparative endocrinology **105**(2): 168-175.

Techa, S. and J. S. Chung (2013). "Ecdysone and retinoid-X receptors of the blue crab, Callinectes sapidus: cloning and their expression patterns in eyestalks and Y-organs during the molt cycle." Gene **527**(1): 139-153.

Tran, C. and M. G. Hadfield (2012). "Are G-protein-coupled receptors involved in mediating larval settlement and metamorphosis of coral planulae?" The Biological Bulletin **222**(2): 128-136.

Turner, A. J., R. E. Isaac and D. Coates (2001). "The neprilysin (NEP) family of zinc metalloendopeptidases: genomics and function." Bioessays **23**(3): 261-269.

Wang, C.-S., H. Pan, G. M. Weerasekare and R. J. Stewart (2015a). "Peroxidase-catalysed interfacial adhesion of aquatic caddisworm silk." Journal of The Royal Society Interface **12**(112): 20150710.

Wang, Z., D. H. Leary, J. Liu, R. E. Settlage, K. P. Fears, S. H. North, A. Mostaghim, T. Essock-Burns, S. E. Haynes and K. J. Wahl (2015b). "Molt-dependent transcriptomic analysis of cement proteins in the barnacle Amphibalanus amphitrite." BMC genomics **16**(1): 859.

Williams, E. A. and S. M. Degnan (2009). "Carry-over effect of larval settlement cue on postlarval gene expression in the marine gastropod Haliotis asinina." Molecular ecology **18**(21): 4434-4449.

Yamada, M.-a., T. Murata, S. Hirose, G. Lavorgna, E. Suzuki and H. Ueda (2000). "Temporally restricted expression of transcription factor betaFTZ-F1: significance for embryogenesis, molting and metamorphosis in Drosophila melanogaster." Development **127**(23): 5083-5092.

Yamamoto, H., S. Kawaii, E. Yoshimura, A. Tachibana and N. Fusetani (1997a). "20-Hydroxyecdysone regulates larval metamorphosis of the barnacle, Balanus amphitrite." Zoological science **14**(6): 887-893.

Yamamoto, H., T. Okino, E. Yoshimura, A. Tachibana, K. Shimizu and N. Fusetani (1997b). "Methyl farnesoate induces larval metamorphosis of the barnacle, Balanus amphitrite via protein kinase C activation." Journal of Experimental Zoology **278**(6): 349-355.

Yamamoto, H., A. Tachibana, K. Matsumura and N. Fusetani (1995). "Protein kinase C (PKC) signal transduction system involved in larval metamorphosis of the barnacle, Balanus amphitrite." Zoological science **12**(4): 391-397.

Yan, X.-C., Z.-F. Chen, J. Sun, K. Matsumura, R. S. Wu and P.-Y. Qian (2012). "Transcriptomic analysis of neuropeptides and peptide hormones in the barnacle Balanus amphitrite: evidence of roles in larval settlement." PLoS One **7**(10): e46513.

Young, M. D., M. J. Wakefield, G. K. Smyth and A. Oshlack (2012). "goseq: Gene Ontology testing for RNA-seq datasets." R Bioconductor.

Zhou, B., K. Hiruma, T. Shinoda and L. M. Riddiford (1998). "Juvenile hormone prevents ecdysteroid-induced expression of broad complex RNAs in the epidermis of the tobacco hornworm, Manduca sexta." Developmental biology **203**(2): 233-244.

Supplementary

Table S1 BUSCO assessment of the de novo transcriptome assembly

BUSCO assessment	
Gene number	72,407
Complete BUSCOs	742
Complete and single-copied BUSCOs	724
Complete and duplicated BUSCOs	18
Fragmented BUSCOs	102
Missing BUSCOs	222
Total BUSCO groups searched	1,066

Table S2 A list of annotated genes that were found to be upregulated in the attached stage and were back to roughly close-search expression level in the juvenile stage (transient expression profile cluster V; Figure 1B and 3).

Gene	**Annotation**
c151995_g1	Chitin binding Peritrophin-A domain
c155284_g1	POL4_DROME_Retrovirus-related Pol polyprotein from transposon 412
c166066_g1	FTZF1_Nuclear hormone receptor FTZ-F1
c177339_g1	FTZF1_Nuclear hormone receptor FTZ-F1
c177564_g1	Interferon-regulatory factor 3, transcription factor
c179093_g1	HEXC_BOMMO_Chitooligosaccharidolytic beta-N-acetylglucosaminidase
c183223_g1	Chitin binding Peritrophin-A
c185366_g1	structural constituent of cuticle
c186711_g1	LACH_DROME_Lachesin
c186829_g1	HR3_DROME_ nuclear hormone receptor HR3
c189272_g1	CHIA_HUMAN_Acidic mammalian chitinase
c189435_g1	VKT1_ANTEL_KappaPI-actitoxin-Ael3a, potassium channel inhibitor
c190234_g1	PTHD3_MOUSE_Patched domain-containing protein 3
c191630_g1	UN93L_DROME_UNC93-like protein
c192145_g1	CU30_BOMMO_Larval cuticle protein LCP-30
c192488_g1	SRR_BOVIN_Serine racemase
c192575_g2	SC6A5_HUMAN_Sodium- and chloride-dependent glycine transporter 2
c193994_g1	DOXA1_HUMAN_Dual oxidase maturation factor 1
c195616_g1	C6A13_DROME_Probable cytochrome P450 6a13

c196099_g1	peptidase inhibitor activity
c196406_g1	structural constituent of cuticle
c196646_g1	DLL1_HUMAN_Delta-like protein 1
c197902_g1	HR38_BOMMO_ nuclear hormone receptor HR38
c198231_g1	SCRB2_RAT_Lysosome membrane protein 2
c200060_g1	B3GT5_PANPA_Beta-1,3-galactosyltransferase 5
c200149_g1	TCB1_CAEBR_Transposable element Tcb1 transposase
c200445_g1	B3A2_HORSE_Anion exchange protein 2
c200451_g1	VA3_SOLIN_Venom allergen 3
c200595_g1	UAP1_MOUSE_UDP-N-acetylhexosamine pyrophosphorylase
c201326_g1	SC6A1_RAT_Sodium- and chloride-dependent GABA transporter 1
c202323_g1	CP4C1_BLADI_Cytochrome P450 4C1
c202457_g1	ATS16_MOUSE_A disintegrin and metalloproteinase with thrombospondin motifs 16
c202694_g1	PGBM_HUMAN_Basement membrane-specific heparan sulfate proteoglycan core protein
c203035_g1	PTHD3_MOUSE_Patched domain-containing protein 3
c203044_g1	PERC_DROME_Chorion peroxidase
c203177_g1	SKEL1_DROME_Protein Skeletor, isoforms B/C
c203228_g1	CHST3_MOUSE_Carbohydrate sulfotransferase 3
c203985_g1	PCE_TACTR_Proclotting enzyme
c204253_g1	VPS62_YEAST^_Vacuolar protein sorting-associated protein 62
c205059_g1	CUT1_CAEEL_Cuticlin-1
c205923_g1	CHS6_USTMA_Chitin synthase 6
c207089_g1	5NTD_LUTLO_Protein 5NUC
c207412_g1	ALS_MOUSE_Insulin-like growth factor-binding protein complex acid labile subunit
c207645_g1	PXDN_MOUSE_Peroxidasin homolog

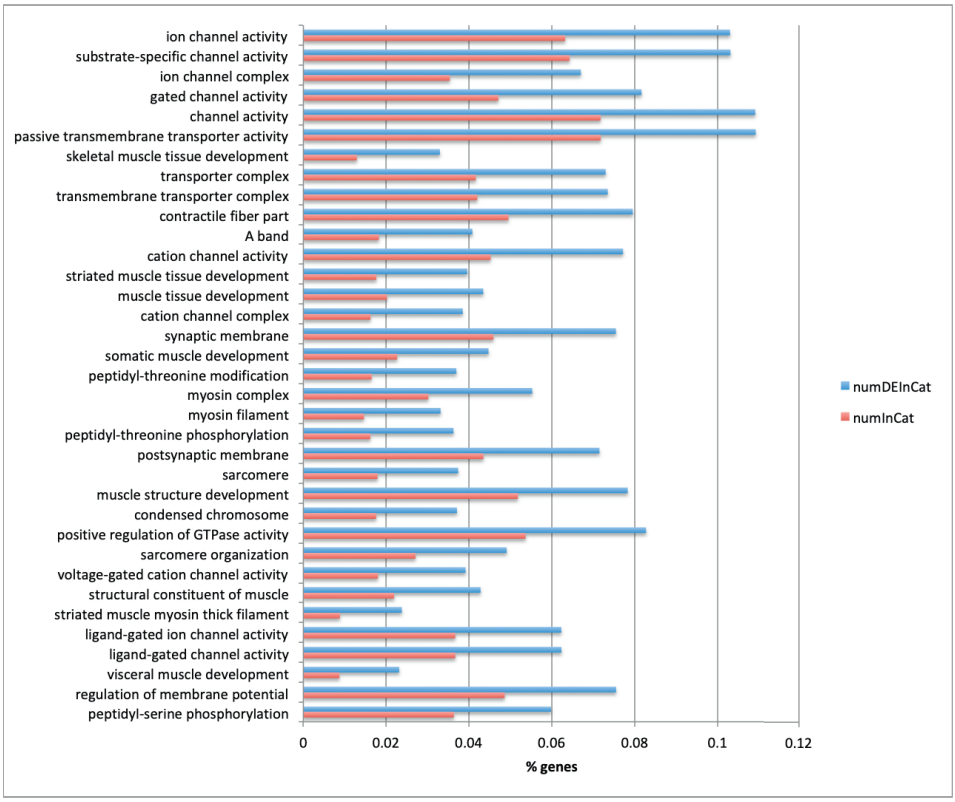


Figure S3 GO enrichment for the genes more highly expressed in the cyprids compared to in juveniles (enrichment factor 1.5; sorted by FDR). numInCat number of genes in the category, numDEInCat number of differentially expressed genes in the category.

Table S4 Top 25 genes downregulated in the attached stage compared to close-search (FDR 0.05, FC 2), sorted by Fold change.

Gene	Annotation	Description	Fold Change	FDR
c181960_g1	CP2DB_MOUSE	Cytochrome P450 2D11	0,0	0,005
c192872_g1	MLR_BOMMO	Myosin regulatory light chain 2	0,1	0,021
c185714_g2	RTXE_DROME	Probable RNA-directed DNA polymerase from transposon X-element	0,1	0,013
c199219_g1	WFDC5_AOTNA	WAP four-disulfide core domain protein 5	0,2	0,010
c205720_g2	CHIT_MANSE	Endochitinase	0,2	0,008
c204433_g1	CHIT3_DROME	Probable chitinase 10	0,2	0,003
c206419_g1	SV2A_RAT	Synaptic vesicle glycoprotein 2A	0,2	0,010
c202678_g1	PAN1_YARLI	Actin cytoskeleton-regulatory complex protein PAN1	0,2	0,004
c200100_g1	SCP1_ASTLP	Sarcoplasmic calcium-binding protein 1	0,2	0,006
c192047_g1	FTN_DROME	Flightin	0,2	0,010
c207488_g1	FOLC_BOVIN	Folylpolyglutamate synthase, mitochondrial	0,2	0,017
c206724_g1	TLCD2_DANRE	TLC domain-containing protein 2	0,2	0,029
c188359_g1	ZAN_MOUSE	Zonadhesin	0,2	0,031
c206928_g1	MCTP2_HUMAN	Multiple C2 and transmembrane domain-containing protein 2	0,2	0,031
c203145_g1	EIGER_DROME	Protein eiger	0,2	0,030
c195413_g1	SUIS_RABIT	Sucrase-isomaltase, intestinal	0,2	0,038
c205749_g1	BRC1_DROME	Broad-complex core protein isoforms 1/2/3/4/5	0,2	0,003
c179402_g1	ACE_CHICK	Angiotensin-converting enzyme	0,3	0,030
c198230_g1	ARRH_HELVI	Arrestin homolog	0,3	0,049
c203235_g1	ACE_DROME	Angiotensin-converting enzyme	0,3	0,003
c204212_g1	SL3_SISCA	Snaclec 3	0,3	0,033
c186227_g1	HDC_DROME	Headcase protein	0,3	0,031
c191897_g1	NEP_MOUSE	Neprilysin	0,3	0,047
c201191_g2	ELF1_DROME	Protein grainyhead	0,3	0,031

Table S5 Top 25 genes A) upregulated and B) downregulated in the juvenile stage compared to attached (FDR 0.05, FC 2), sorted by fold change.

A)

Gene	Annotation	Description	Fold Change	FDR
c205732_g1	EST2C_MOUSE	Acylcarnitine hydrolase	2091,3	9,22E-06
c206063_g1	PERO_DROME	Peroxidase	1679,3	9,86E-06
c189049_g1	WBP2L_HUMAN	Postacrosomal sheath WW domain-binding protein	856,7	9,62E-05
c195373_g1	TENX_HUMAN	Tenascin-X	488,3	3,74E-05
c172684_g5	CATL_SARPE	Cathepsin L	430,7	1,99E-05
c172650_g1	VA5_PACCH	Venom allergen 5	420,6	1,35E-05
c195666_g1	YR877_MIMIV	Putative lipocalin R877	384,8	2,21E-05
c188548_g1	AQP9_HUMAN	Aquaporin-9	369,0	3,26E-05
c196582_g1	FCER2_MOUSE	Low affinity immunoglobulin epsilon Fc receptor	322,5	0,00013136

c149338_g1	CALM_STRPU	Calmodulin	303,5	9,99E-05
c181977_g2	CYSP2_HOMAM	Digestive cysteine proteinase 2	281,3	9,34E-05
c176722_g3	CYSP2_HOMAM	Digestive cysteine proteinase 2	245,6	3,95E-05
c194806_g1	LYS_CRAGI	Lysozyme	238,3	8,30E-05
c193694_g1	RTBS_DROME	Probable RNA-directed DNA polymerase from transposon BS	204,1	0,000837854
c190873_g1	MET24_HUMAN	Methyltransferase-like protein 24	191,7	4,13E-05
c200079_g2	LYAM2_CANLF	E-selectin	190,5	8,59E-05
c198705_g1	LY75_HUMAN	Lymphocyte antigen 75	174,8	9,69E-05
c167311_g1	CPG2_CAEEL	Chondroitin proteoglycan-2	173,5	2,44E-05
c164567_g1	EST5_RAT	Liver carboxylesterase B-1	159,0	0,000173098
c164550_g1	ORYA_ORYSJ	Oryzain alpha chain	157,1	0,00066491
c197734_g1	DNS2A_PIG	Deoxyribonuclease-2-alpha	152,0	0,000101709
c130548_g2	LIPR3_HUMAN	Pancreatic lipase-related protein 3	151,6	8,96E-05
c174701_g1	HEXB_FELCA	Beta-hexosaminidase subunit beta	149,7	0,000902095
c200763_g1	SPB8_BOVIN	Serpin B8	146,3	0,000209888

B)

Gene	Annotation	Description	Fold Change	FDR
c207412_g1	ALS_MOUSE	Insulin-like growth factor-binding protein complex acid labile subunit	0,002	1,52E-05
c188359_g1	ZAN_MOUSE	Zonadhesin	0,003	0,000162482
c205749_g1	BRC1_DROME	Broad-complex core protein isoforms 1/2/3/4/5	0,003	6,64E-06
c202678_g1	PAN1_YARLI	Actin cytoskeleton-regulatory complex protein PAN1	0,004	9,22E-06
c201125_g2	TLL2_MOUSE	Tolloid-like protein 2	0,005	3,84E-05
c196874_g1	STUB_DROME	Serine proteinase stubble	0,005	2,52E-05
c198699_g1	TRIPC_BOVIN	E3 ubiquitin-protein ligase TRIP12	0,005	7,67E-05
c200334_g1	C209B_MOUSE	CD209 antigen-like protein B	0,005	1,46E-05
c202138_g1	PRSS8_HUMAN	Prostasin	0,005	1,69E-05
c171720_g3	NAS34_CAEEL	Zinc metalloproteinase nas-34	0,005	5,07E-05
c198699_g2	TRIPC_HUMAN	E3 ubiquitin-protein ligase TRIP12	0,006	0,000173098
c180673_g1	KCP_HALAI	BPTI/Kunitz domain-containing protein	0,006	0,000121548
c180764_g1	SRR_DICDI	Probable serine racemase	0,007	0,000122974
c199162_g2	TRIPC_XENTR	E3 ubiquitin-protein ligase TRIP12	0,007	4,13E-05
c171720_g1	NAS34_CAEEL	Zinc metalloproteinase nas-34	0,007	0,000284438
c204621_g1	POL3_DROME	Retrovirus-related Pol polyprotein from transposon 17.6	0,008	0,000397152
c206987_g1	ATS16_MOUSE	A disintegrin and metalloproteinase with thrombospondin motifs 16	0,008	6,40E-06
c207735_g2	POL3_DROME	Retrovirus-related Pol polyprotein from transposon 17.6	0,008	3,00E-05
c170345_g1	RREB1_MOUSE	Ras-responsive element-binding protein 1	0,009	0,000170621
c202338_g1	TRIPC_BOVIN	E3 ubiquitin-protein ligase TRIP12	0,009	0,00048209
c195607_g1	JHBP_BOMMO	Juvenile hormone-binding protein	0,010	0,001093588
c157458_g1	COLL7_MIMIV	Collagen-like protein 7	0,011	0,000322421

c188622_g1	C15C1_BOMMO	Farnesoate epoxidase	0,012	7,71E-05
c187948_g1	CHDH_RAT	Choline dehydrogenase, mitochondrial	0,012	0,000196211

Figure S6 Surface wettability change over the course of experiment.

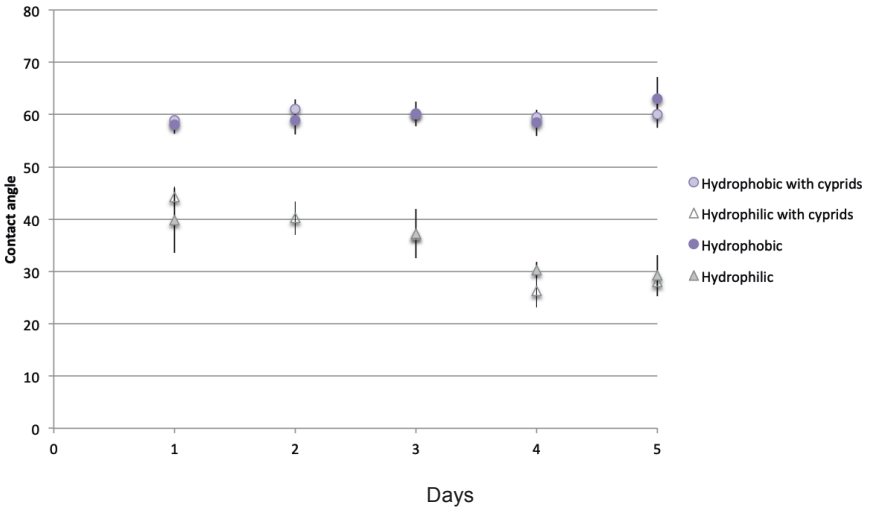


Figure S7 Cement protein expression during settlement of *B. improvisus*

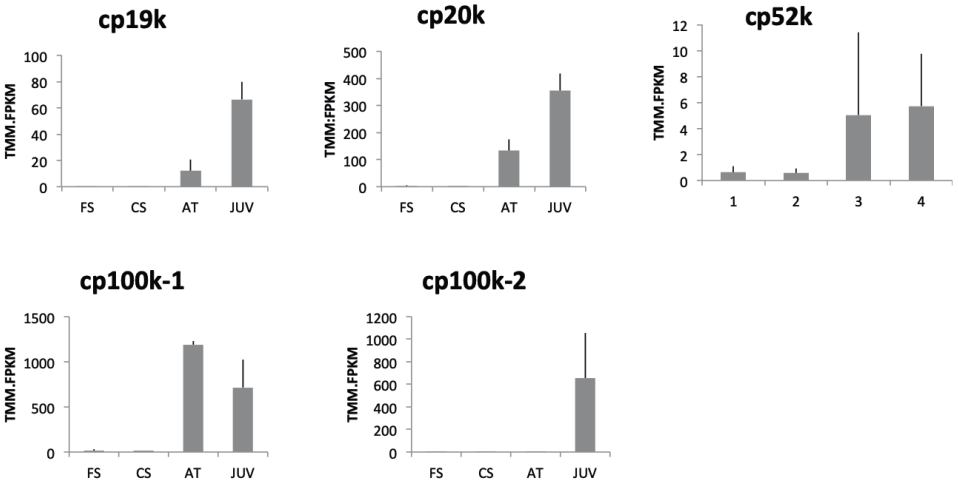


Table S8 Top 20 genes highly regulated between the two surfaces (HLvsHB) in the free-swimming stage (FDR 0.05, FC 2), sorted by fold change.

Gene	Annotation	Description	Fold Change	FDR
c161100_g1	ASCC3_CHICK	Activating signal cointegrator 1 complex subunit 3	8,1	0,001
c166497_g1	NPL21_CAEEL	Neprilysin-21	5,5	0,043
c166497_g2	NPL11_CAEEL	Neprilysin-11	5,3	0,005
c184343_g2	RTXE_DROME	Probable RNA-directed DNA polymerase from transposon X-element	5,1	0,047
c179086_g1	GCYB2_HUMAN	Guanylate cyclase soluble subunit beta-2	4,8	0,002
c173142_g1	CFA70_MACFA	Cilia- and flagella-associated protein 70	4,7	0,031
c195422_g1	RTJK_DROME	RNA-directed DNA polymerase from mobile element jockey	4,7	0,012
c193352_g1	JHBP_MANSE	Juvenile hormone-binding protein	4,6	0,021
c199508_g3	SDK_DROME	Protein sidekick	4,5	0,029
c163387_g1	RHDF1_DROME	Inactive rhomboid protein 1	-1,2	0,014
c172459_g1	SDK_DROME	Protein sidekick	-0,2	0,040
c189150_g1	FAS_CHICK	Fatty acid synthase	-0,1	0,031
c202730_g1	LONF3_MOUSE	LON peptidase N-terminal domain and RING finger protein 3	0,2	0,000
c194288_g1	UBR4_HUMAN	E3 ubiquitin-protein ligase UBR4	0,2	0,009
c180460_g3	MAGI1_HUMAN	Membrane-associated guanylate kinase, WW and PDZ domain-containing protein 1	0,4	0,018
c207089_g1	5NTD_LUTLO	Protein 5NUC	0,4	0,002
c188533_g1	UBR4_HUMAN	E3 ubiquitin-protein ligase UBR4	0,5	0,015
c183880_g1	BRC1_DROME	Broad-complex core protein isoforms 1/2/3/4/5	0,6	0,042
c187845_g1	POL4_DROME	Retrovirus-related Pol polyprotein from transposon 412	0,6	0,029
c205661_g1	S18B1_HUMAN	MFS-type transporter SLC18B1	0,4	0,000

Table S9 Top 20 genes highly regulated between the two surfaces in the close-search stage (FDR 0.05, FC 2), sorted by fold change (HL/HB).

Gene	Annotation	Description	Fold Change	FDR
c175675_g1	HAKAI_XENLA	E3 ubiquitin-protein ligase Hakai	16,9	0,029
c175186_g2	CU20_TENMO	Adult-specific cuticular protein ACP-20	10,0	0,009
c171712_g1	NPHN_RAT	Nephrin	9,4	0,044
c169266_g1	AKTIP_XENTR	AKT-interacting protein	9,0	0,010
c177579_g1	EMRE_AEDAE	Essential MCU regulator, mitochondrial	7,2	0,026
c173656_g1	RNH_SYNY3	Ribonuclease HI	4,0	0,048
c203958_g1	STUB_DROME	Serine proteinase stubble	3,7	0,011
c178050_g1	CAF17_MOUSE	Putative transferase CAF17 homolog, mitochondrial	3,0	0,034
c180792_g1	POMP_PONAB	Proteasome maturation protein	2,7	0,009
c184410_g1	DFP3_ANTMY	Putative defense protein 3	2,7	0,003
c105892_g1	RTBS_DROME	Probable RNA-directed DNA polymerase from transposon BS	0,1	0,011
c204383_g1	CLH2_ARATH	Chlorophyllase-2, chloroplastic	0,1	0,038
c164522_g1	KRBA2_HUMAN	KRAB-A domain-containing protein 2	0,1	0,040

c177790_g1	RTP2_TRYBG	Retrotransposable element SLACS 132 kDa protein	0,1	0,006
c166497_g2	NPL11_CAEEL	Neprilysin-11	0,1	0,014
c176433_g2	RTJK_DROME	RNA-directed DNA polymerase from mobile element jockey	0,1	0,006
c181575_g1	RTBS_DROME	Probable RNA-directed DNA polymerase from transposon BS	0,1	0,011
c183758_g2	RTJK_DROFU	RNA-directed DNA polymerase from mobile element jockey	0,2	0,010
c171810_g1	POL4_DROME	Retrovirus-related Pol polyprotein from transposon 412	0,2	0,044
c200803_g1	RTXE_DROME	Probable RNA-directed DNA polymerase from transposon X-element	0,2	0,000

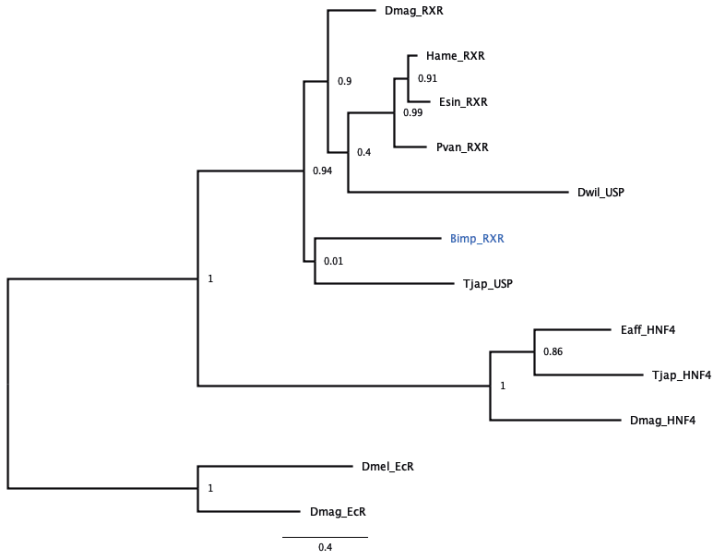
Table S10 Top 20 genes highly regulated between the two surfaces in the attached stage (FDR 0.05, FC 2), sorted by fold change.

Gene	Annotation	Description	Fold Change	FDR
c196728_g2	AQP_ANOGA	Aquaporin AQPAn.G	18,197	0,013374647
c112394_g1	CUPA4_CANPG	Cuticle protein AM1239	12,345	0,000385529
c198231_g1	SCRB2_RAT	Lysosome membrane protein 2	4,537	0,009244213
c200445_g1	B3A2_HORSE	Anion exchange protein 2	3,596	0,004861874
c182604_g1	BP10_PARLI	Blastula protease 10	0,021	0,006572355
c171806_g2	TRET1_APILI	Facilitated trehalose transporter Tret1	0,024	0,034332445
c204157_g1	PERO_DROME	Peroxidase	0,035	6,38E-05
c195460_g1	RTJK_DROME	RNA-directed DNA polymerase from mobile element jockey	0,038	0,006572355
c167443_g8	TRDMT_HUMAN	tRNA (cytosine(38)-C(5))-methyltransferase	0,048	2,46E-12
c198987_g1	OVCH1_HUMAN	Ovochymase-1	0,048	0,009339422
c155440_g1	HARB1_BOVIN	Putative nuclease HARBI1	0,052	0,018125603
c198874_g1	SPRR3_MOUSE	Small proline-rich protein 3	0,058	1,78E-08
c201374_g1	CEL2B_HUMAN	Chymotrypsin-like elastase family member 2B	0,066	0,009244213
c165745_g1	CYSP3_HOMAM	Digestive cysteine proteinase 3	0,068	0,000402109
c172827_g2	R1A_CVHN1	Replicase polyprotein 1a	0,074	0,005990281
c184945_g1	CPMD8_HUMAN	C3 and PZP-like alpha-2-macroglobulin domain-containing protein 8	0,075	0,00028871
c180303_g1	WBP2L_HUMAN	Postacrosomal sheath WW domain-binding protein	0,082	0,013294748
c187090_g1	CAH6_HUMAN	Carbonic anhydrase 6	0,087	0,00748944
c178623_g3	RTBS_DROME	Probable RNA-directed DNA polymerase from transposon BS	0,088	0,019478422
c204463_g1	CAH2_TRIHK	Carbonic anhydrase 2	0,102	0,000680052

Table S11 Top 20 genes highly regulated between the two surfaces in juvenile stage (FDR 0.05, FC 2), sorted by fold change

Gene	Annotation	Description	logFC	FDR
c199835_g1	JHAMT_DROME	Juvenile hormone acid O-methyltransferase	36,858	0,01874073
c139649_g1	EF1A1_RAT	Elongation factor 1-alpha 1	23,013	0,012065358
c157376_g1	SLPI_HUMAN	Antileukoproteinase	18,516	0,016347976
c188798_g1	TEN3_HUMAN	Teneurin-3	13,318	0,000211069
c194902_g1	CUA2B_TENMO	Larval cuticle protein A2B	12,138	0,022193484
c187327_g1	EDEM3_MOUSE	ER degradation-enhancing alpha-mannosidase-like protein 3	9,984	0,002880871
c164086_g2	SPAN_STRPU	Protein SpAN	6,821	0,010461339
c188766_g2	PXDN_CAEBR	Peroxidasin homolog	6,518	0,040356992
c185996_g1	RTJK_DROME	RNA-directed DNA polymerase from mobile element jockey	3,558	0,004997322
c207089_g1	5NTD_LUTLO	Protein 5NUC	3,428	0,001040558
c195790_g2	FBXW5_MOUSE	F-box/WD repeat-containing protein 5	0,011	0,022291378
c37620_g1	SCP20_ARATH	Serine carboxypeptidase-like 20	0,011	0,033646004
c214424_g1	RNOY_CRAGI	Ribonuclease Oy	0,011	0,040356992
c192494_g1	RTJK_DROME	RNA-directed DNA polymerase from mobile element jockey	0,011	0,029857798
c163661_g1	TBA_TETTH	Tubulin alpha chain	0,011	2,83E-08
c177468_g1	TAR1_YEAST	Protein TAR1	0,012	1,45E-05
c165997_g3	ART2_YEAST	Putative uncharacterized protein ART2	0,014	0,000639712
c142314_g2	EF1A_TETPY	Elongation factor 1-alpha	0,015	0,000233787
c210479_g1	PAPA4_CARPA	Papaya proteinase 4	0,018	0,002580607
c211321_g1	CALM_TETPY	Calmodulin	0,023	0,016347976

A)



B)

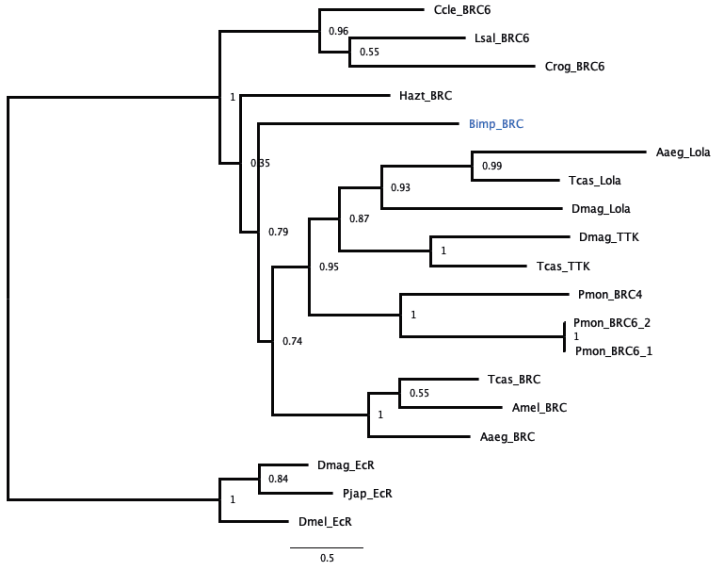


Figure S12 Maximum-likelihood phylogeny of *B. improvisus* candidate A) RXR and B) BRC. Dmel *D. melanogaster*, Dmag *D. magna*, Tjap *Tigriopus japonicus*, Pjap *Penaeus japonicus,* Esin *Eriocheir sinensis*, Mnip *Macrobrachium nipponense*, Ggal *Gallus gallus*, Omyk *Oncorhynchus mykiss.* The numbers indicate bootstrap values for each branch.

V

A pilot study for the sequencing of the barnacle *Balanus improvisus* genome reveals extreme genetic diversity

Magnus Alm Rosenblad[1], Anna Abramova[1], Ulrika Lind[1], Pall Olason[2], Stefania Giacomello[3], Björn Nystedt[2], Anders Blomberg[1*]

[1] Dept of Chemistry and Molecular Biology, University of Gothenburg.
[2] Dept of Cell and Molecular Biology, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Uppsala University, Husargatan 3, SE-752 37 Uppsala, Sweden
[3] Dept of Biochemistry and Biophysics, National Bioinformatics Infrastructure Sweden, Science for Life Laboratory, Stockholm University, Box 1031, SE-17121 Solna, Sweden.
* corresponding author: magnus.almrosenblad@cmb.gu.se

Abstract

Objective Barnacles are key marine crustaceans in several biotopes. Barnacle biofouling on man-made constructions has large economical consequences and is estimated at a global 150 billion US dollars annually in added fuel and maintenance costs. Despite ecological and economical importance, there is a surprising void of basic genomic knowledge on barnacles, and a barnacle reference genome is lacking. We here set out to characterize the genome of the bay-barnacle *Balanus improvisus* (= *Amphibalanus improvisus*) based on short-read whole genome sequencing and assembly, experimental genome size estimations, individual gene sequencing and transcriptomics.
Results We show that *B. improvisus* has a haploid genome size of ~ 740 million base-pairs, which is slightly larger than the total size of the genome assembly (~ 600Mb). The assembled genome is highly fragmented with an N50 of 2.2kb, which probably is a result of the extremely high nucleotide diversity (average ≈ 5% in coding regions), and a high repeat content (at least 40%). We argue that the genomic features described here can help explain previous contradictory and inconclusive results from genetic studies of gene families of interest for biofouling. A well-resolved reference genome is urgently needed to properly explore and understand proteins of interest in barnacle biology in general and for developing better antifouling strategies.

Introduction

Barnacles are one of the main biofouling organisms on man-made underwater constructions and ship hulls causing considerable ecological and economical impacts. Furthermore, they demonstrate one of the strongest underwater adhesions and are therefore of great interest in marine biotechnology. The bay-barnacle *Balanus improvisus* is a common barnacle species in temperate waters that is used as a model organism for studies of biofouling mechanisms in marine and brackish conditions. Substantial amount of knowledge has been accumulated over the years of research regarding settlement biology of this species, however, there is an apparent lack of genomic data. Currently, there is no reference genome available for any of the barnacle species and even the genome size is unclear.

While transcriptome sequencing is now a standard and cheap method for acquiring large-scale datasets to be used in functional studies and for the identification of interesting genes, it is not suited to gain a complete view of the proteome as the completeness is dependent on life-stage, conditions, and tissues as well as the expression-level of genes (lowly expressed genes are usually missed). Traditional transcriptome sequencing with short reads is also unsatisfying in that it depends on an assembly, at which assembly errors can occur, resulting in artificial chimeric sequences representing mixed alleles, paralogs and/or pseudogenes, as well as an overestimation of the total number of transcripts. Such errors in turn risk leading to severely limited or erroneous interpretations of functional and evolutionary studies.

A well-resolved genome assembly would be invaluable to help address many of these problems, and provide a fundamental framework in order to correctly interpret functional, evolutionary and ecological studies of this important biofouling species. As a first step, we here provide the most genomic characterization of the *B. improvisus* genome to date, providing important insights into the extreme genetic diversity of this species and setting the stage for a successful creation of a high-quality barnacle reference genome.

Materials and Methods

We here provide an overview of Material and Methods, for more detailed descriptions, see Supplementary information.

Sample preparation, sequencing and genome assembly

High quality genomic DNA was prepared from one adult individual of *B. improvisus* from Tjärnö, west-coast of Sweden, using a protocol described in Lind et al., 2010. The sample was sent to the national sequencing facility SciLifeLab (Stockholm) for paired-end Illumina HiSeq2000 sequencing (2x100 bp read length) of 150 bp and 300 bp fragment libraries producing a total of 157 Mbp. The reads were filtered, trimmed and assembled as described in Lind et al., 2013. Using different kmer sizes (k=25-57) produced similar results (Table 1).

Genome characterisation

Using sequencing reads and the assembly contigs we estimated the repeat content, genome size and nucleotide diversity, the latter two using k-mer analysis. Gene structure predictions were based on transcripts matching conserved genes from Drosophila and the assembly contigs, producing distributions of exon and intron sizes. Genetic diversity in the protein-coding regions (CDS) for both synonymous and non-synonymous mutations were calculated based on alignments of allelic contigs for 70 CEGMA genes.

Genome size experimental verification

Adult barnacle cirri and cyprid larvae were used as material for flow cytometric measurements of nuclear DNA content. Nuclear preparation was made according to Hare et al., 2011. Dissected cirri and larvae (~30 individuals) were gently homogenized in Galbraith buffer with a pestle and filtered through a filter. Chicken erythrocytes were used as a reference (C-value of 1.14 Gbp) as earlier described (Dhillon et al., 1977). Freshly prepared nuclei suspensions from barnacles were mixed with aliquots of chicken nuclei, and co-stained. The mean fluorescence was quantified using a BD FACSAria II flow cytometer (>615nm). Data were analyzed with BD FACSDiva software and DNA values for *B. improvisus* were calculated by comparison to the chicken genome.

Results

Genome sequencing and draft assembly

Adult *B. improvisus* individuals have a maximum wet weight of about 12 mg (Panova et al., 2016), limiting the amount of total DNA per individual to ~ 10-15µg. This is too low an amount to produce additional long-insert sequencing libraries (without PCR amplification) that are recommended for short-read assembly strategies. In addition, the DNA amount and quality from each individual are highly variable making DNA extraction for library preparation tedious and results quite unpredictable (Panova et a., 2016). The genomic DNA was sequenced using short-read technology (Illumina, 2 x 100bp) to a total of 157 Gbp. De novo genome assembly resulted in a remarkably fragmented assembly, with a total assembly size of ~600 Mbp and an N50 of 2.2 kbp. Only 22 Mbp (4%) of the assembly were represented in contigs longer than 10 kbp (Table 1). This is an unusually poor result given the amount of input data, indicating an underlying genome architecture that poses a considerable assembly challenge.

Genome size

Conflicting estimates of genome sizes of 740 Mbp or 1400 Mbp have been reported for *Balanus amphitrite*, with two other Balanidae species reported to have similar large sizes of 1260 Mbp (*Semibalanus cariosus*) and 1400 Mbp (*B.eburneus)* (Bachmann and Rheinsmith, 1973; Rheinsmith, 1974). Given the fragmented assembly result, the total assembly size cannot reliably predict the genome size for *B. improvisus*. We therefore performed an experimental genome size estimation of *B. improvisus* based on nuclei preparation, DNA staining and FACS, resulting in an estimated haploid genome size of 738 Mbp (+- 9 Mbp) (Figure 1A, Table S1). This is supporting the previous lower-end genome size estimate of *B.amphitrite*. We also investigated the kmer frequency spectra of the unassembled whole genome sequencing reads, using the total amount of kmers and the average expected kmer coverage to estimate the genome size, which resulted in diploid (2n) sizes of 1.37-1.46 Gbp. Although this analysis is expected to provide a more crude estimate than the experimental estimate, this method also rejected the larger of the previously reported genome sizes, and we thus conclude the *B. improvisus* haploid genome to be ~740 Mbp.

Table 1. Genome assembly statistics for *B. improvisus*

	Unscaffolded	Paired-end scaffolding	Transcriptome scaffolding
Number	587,357	545,974	530,331
Total size*	612 Mbp	612 Mbp	612 Mbp
Size in contigs/scaffolds*			
> 1 Kbp	449 Mbp	465 Mbp	468 Mbp
> 5 Kbp	119 Mbp	166 Mbp	192 Mbp
> 10 Kbp	22 Mbp	42 Mbp	74 Mbp
Max contig/scaffold size*	39 Kbp	40 Kbp	82 Kbp
Fraction of CEGMA core genes			
Complete (>70% aligned)	48%	52%	61%
Partial (>30% aligned)	84%	85%	92%
Contig N50, bp			2.2kb

* Excluding gaps

N50 length of the shortest size-ordered contig required to represent 50% of the estimated genome size.

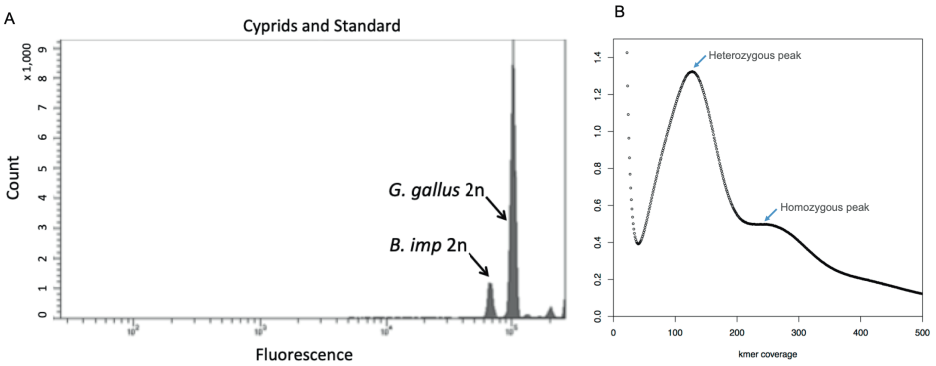


Figure 1 A) Flow cytometry of propidium iodide-stained nuclei from cyprids and cirri of *B. improvisus* and the chicken blood standard *G. gallus* (C-value of 1,14 Gbp) nuclei. Fluorescence peaks of cyprid 2n nuclei and reference *G. gallus* 2n nuclei. B) The distribution of 15-kmer frequencies of *B. improvisus* genome reads vs coverage, showing a pronounced heterozygous peak and a much smaller homozygous peak at double the coverage (256).

Repeat content

Since repeats can be expected to be collapsed or disintegrated in the draft genome assembly, it is nontrivial to estimate the true repeat content of the genome. We started by naively using RepeatMasker to align repeat libraries from other species to the draft genome assembly, resulting in only 1% of the assembly being masked. We then went on to construct a species-specific repeat library for *B. improvisus* based on non-stringent assembly of shallow coverage of sequencing reads (Supplementary Information M&M), resulting in a much larger proportion (17%) of the draft genome assembly being masked. The large difference in masking capability between using previously known repeats from other species compared to species-specific repeats, indicates that the *B. improvisus* repeats are novel or very diverged compared to previously known repeats. Since many repeats can be expected to be heavily underrepresented in the draft genome assembly, we finally used the species-specific repeat library to screen the raw genome sequencing reads, resulting in an estimated 40% repeat content. Since the detection methods can be expected to be underestimating rather than overestimating the repeat content, we conclude that at least 40% of the *B. improvisus* genome is represented by high-copy repeat families. Due to the fragmented nature of the assembly and the lack of similarity to known repeats, it was not meaningful possible to classify or quantify the abundance of different types of repeat families at this point.

Genetic diversity

The kmer frequency spectra from the raw sequence reads showed a very pronounced bimodal shape at relatively short kmers (k=15), a pattern typical for species with high heterozygosity (Figure 1B). In line with this, we also noted that most of the assembled contigs in the draft genome assembly appeared to represent a single allele rather than a merged representation of the two alleles. To further interrogate this, we first performed PCR and cloning of both alleles from an earlier reported single-copy gene, the octopamine receptor OctA (Lind et al., 2010), in four *B. improvisus* adult individuals (Table S2). The nucleotide diversity in this 1473 bp single-exon CDS was 2.6-3.9%. Adding data from an additional 27 cloned cDNA from OctA alleles from adults and cyprid larvae (representing several hundreds of individuals from one single population) estimated the OctA pairwise nucleotide diversity in *B. improvisus* for the CDS to be 3.6% (Table S2). Extending the analysis to the 5' UTR region

demonstrated an even more extreme sequence variation, with additional inserts and deletions (Figure 2). To further investigate the general level of nucleotide diversity in coding regions, we searched the genome assembly for conserved core (putative single-copy) genes where the two alleles had been split up in separate contigs in the assembly process. For a set of 137 conserved protein-coding genes as predicted from CEGMA, we identified putative split alleles for about 70 genes (51%) with a 5.3% median (range 2 - 10%) nucleotide diversity (Figure 3). Even for the 16 proteins in this set with no amino acid mismatch between the alleles, the median nucleotide diversity was still 4.0%, indicating that the nucleotide diversity is high also for genes which are under strong purifying selection against non-synonymous substitution. These results suggest that nucleotide diversity in coding regions in *B. improvisus* is among the highest reported for any species to date (Leffler et al., 2012).

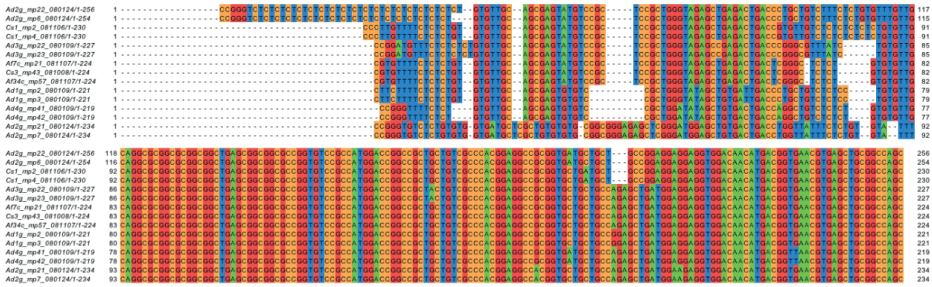


Figure 2 Example of 5' variation in OctA Sanger clones showing 7 different alleles (5' UTR + partial CDS). The top two and bottom two are from the same adult (Ad2). Ad = adult, Af= adult from field experiment, C=cyprid. Numbers refer to the separate individuals. g = genomic, c = cDNA, mpX_xxxxxx = the clone id.

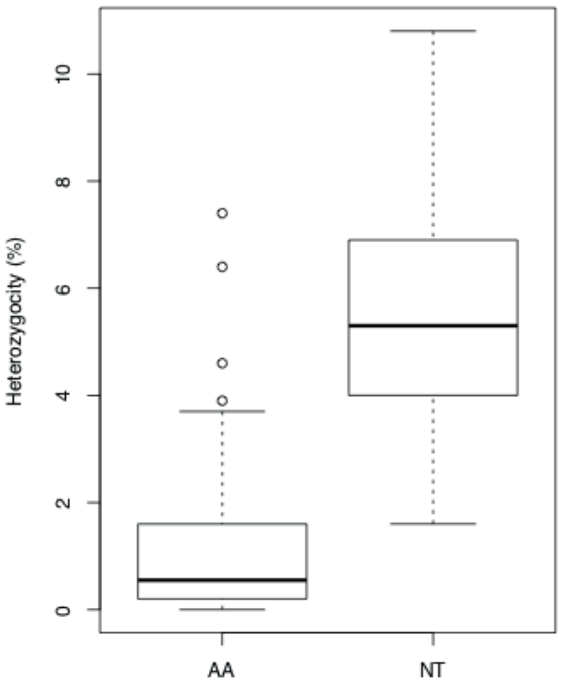


Figure 3 Nucleotide diversity in coding regions as determined from pairwise alignments of putatively split alleles in the *B. improvisus* draft assembly, based on a set of 70 predicted core conserved proteins from the CEGMA analysis. AA: amino acids, NT: nucleotides.

Gene structure and completeness

To investigate the gene structure of *B. improvisus*, we constructed a non-redundant reference set of 4,705 putative full-length genes from the *B. improvisus* transcriptome assembly (from Illumina-based deep-sequencing of RNA from one single adult individual), with clear protein homology to the *D. melanogaster* proteome (Supplementary Information). The reference genes were aligned to the draft genome assembly using Scipio, allowing transcripts to span across DNA contigs. Despite the very fragmented and incomplete genome assembly, the coding gene content of this non-redundant set of reference genes appeared to be reasonably well represented, with >85% of the reference genes being covered over at least 70% of their length. The *B. improvisus* genes typically display short exons (average 202 bp) (Figure S1). For introns, we noted that large introns of several thousand base-pairs were not uncommon on large contigs. The estimated intron size (average 706 bp) is likely heavily underestimated since many genes contain intronic assembly-gaps of unknown lengths, which cannot be properly accounted for. The longest intron found in a single contig was 13 kbp (NT-C2 domain protein).

Discussion/Conclusion

We here provide an initial genomic characterization of the bay-barnacle *B. improvisus*, including the genome size, repeat content, heterozygosity, nucleotide diversity, and gene structure. The most striking feature of the *B. improvisus* genome is the extremely high genetic diversity, with about 5% variation in coding regions and much bigger difference in the non-coding parts. This is among the highest diversity described for any species so far, but in line with very high levels of genetic diversity observed also in other marine species (Leffler et al., 2012). Because coding exons only represent a few percent of animal genomes it is highly likely that the intronic and intergenic regions display a much higher variation, which could make distinguishing between allelic and paralogous regions very hard and thus result in problems when phasing a genome assembly.

Although draft genome assemblies as the one presented here can be useful in order to investigate specific genes and gene families (Lind et al., 2010, Lind et al., 2013, Lind et al., 2016, Abramova et al., 2019), a reference barnacle genome is urgently needed to advance science in this important group of organisms. Based on the here reported extremely high nucleotide diversity, high heterozygosity, and in combination with a high repeat content, we advice to opt for high-coverage long-read sequencing when planning a barnacle genome project; this could then aim for allelic resolution and a complete diploid genome assembly (i.e. including both alleles). Although low-cost alternatives like 10X genomics could lead to an improved assembly compared to the draft presented here, it is unlikely to suffice for such a complex genome, as is also indicated by our in-house early-trial 10X genomics assemblies (data not shown). With the emergence of diploid genome assemblies, we foresee new challenges and possibilities in population genomics analyses, including proper read mapping and accurate allele frequency estimations from population re-sequencing data. While diploid genome assemblies more properly reflects the genome than traditional haploid genome assemblies, new or adapted bioinformatics tools to process re-sequencing data will be needed to take advantage of this improvement. We believe that the here reported extremely high genetic diversity in *B. improvisus* and the connected challenges in producing a good reference genome will be applicable to a large variety of marine species with large, outbred populations and corresponding high genetic diversity.

Acknowledgments

This work was supported by the Centre for Marine Evolutionary Biology at the University of Gothenburg (http://www.cemeb.science.gu.se/). MAR was partially funded by the National Bioinformatics Infrastructure Sweden (NBIS). PO, SG, BN were financially supported by the Knut and Alice Wallenberg Foundation as part of the National Bioinformatics Infrastructure Sweden at SciLifeLab. We thank Tomas Larsson and Mats Töpel for assisting in genome assembly and analysis, Martin Ogemark for providing and maintaining barnacle larvae culture, John Patrick Alao for assisting with flow cytometry measurements.

Supplementary Information

Sample preparation

Adult *B. improvisus* individuals have a maximum wet weight of about 12 mg (Panova et al., 2016), limiting the amount of total DNA per individual to ~ 10-15µg was too low an amount to produce additional long-insert sequencing libraries (without PCR amplification) that were recommended for short-read assembly strategies. In addition, the DNA amount and quality from each individual are highly variable making DNA extraction for library preparation tedious and results quite unpredictable (Panova et al., 2016).

Genome size from kmer histograms

Calculations were done by using jellyfish (v 1.1.11 and 2.1.4) to create kmer histograms and the formula: Genome size = n * (L-k+1)/C, where n= nb of reads , L= avg read length, k= kmer size and C= kmer coverage peak. Depending on kmer size the results were 1.37-1.46 Gbp which was deemed to be the diploid (2n) size after identifying a small homozygous peak at twice the coverage (256) seen when lowering the kmer to 15 (Figure 1B).

Nucleotide diversity of coding genes

To investigate the general level of heterozygosity in coding regions, we search the genome assembly for conserved core (putative single-copy) genes where the two alleles had been split up in separate contigs in the assembly process. First, we used CEGMA v 2.5 to align a set of 458 conserved protein-coding genes to the assembly, resulting in 336 predicted partial or complete proteins. Short proteins were removed, resulting in a set of 137 putative *B. improvisus* proteins of at least 300 amino acids in length. Each protein was aligned back to the assembly using Scipio v 1.4.1, with a relatively stringent Scipio score cutoff (*--min_score 0.50*) and with hits restricted to single contigs (*--single_target_hits*). In line with allele splitting being frequent in the assembly, many of the aligned proteins (70 proteins, 51%) displayed exactly two high-scoring alignments in the assembly, putatively representing the two alleles. For all such cases, we aligned the two putative alleles to each other, and calculated the fraction of mismatches for both the amino acid sequence and the nucleotide sequence, giving a median heterozygosity of 0.5% for amino acids and 5.3% for nucleotides (Figure 3). All alignments were manually inspected to rule out alignment artifacts. While it can be anticipated that split alleles will be more likely found for genes with high heterozygosity, we could still conclude that almost half (45%) of all the 137 investigated genes showed a nucleotide diversity of at least 3%, suggesting that very high heterozygosity is a common feature in *B. improvisus* also in coding regions. For 16 proteins, no amino acid mismatch was found between the alleles, but the median nucleotide diversity was still 4.0% for this set, indicating that the nucleotide heterozygosity is high also for genes which are under strong purifying selection against non-synonymous substitution. While assembly errors can not be completely ruled out, there is no reason to expect high error rates in the assembly given the sequencing technology and the assembly read depth, and it is difficult to reconcile high error rates with the observed patterns of purifying selection. It is of course also possible that a few extreme outliers represent divergent gene copies (or even pseudogenes) rather than alleles, but we consider these cases to be rare.

Genome size experimental verification

Adult barnacle tissues (cirri) and cyprid larvae were used as material for flow cytometric measurements of nuclear DNA content. All samples were frozen at -80°C prior to nuclei preparation and staining. Nuclear sample preparation protocols followed guidelines outlined by Hare et al., 2011. In short, the frozen samples were thawed on ice and immediately after dissection each tissue and larvae (~30 individuals) were collected separately into microtubes. To obtain nuclei suspensions samples were gently homogenized in 1ml of Galbraith buffer (Hare et al., 2011) with a pestle (15 strokes). Each nuclear suspension was prepared by filtering through 20µm nylon net filter (Millipore, Ireland) to remove cellular debris.

Chicken erythrocytes were used as a genome size standard (C-value of 1.14 Gbp) as earlier described (Dhillon et al., 1977). Fresh chicken blood was mixed with ACD blood anticoagulant and stored at -20°C. Cell suspension was prepared by adding 15µl of chicken blood to 1ml of Galbraith buffer, followed by homogenization and filtering procedure as described above for the barnacle samples. Aliquots of chicken nuclei were added to freshly prepared nuclei suspensions from barnacle tissues and to larvae and subsequently stained; propidium iodide (50µg/ml) and RNase (20µg/ml) were added and samples were co-stained in the dark at 4°C for up to 22hs. The mean fluorescence of nuclei in each sample was quantified in at least three replicates using a BD FACSAria II flow cytometer. The cytometer was activated for the red fluorescence detection (>615nm). At least 10,000 nuclei were measured per sample. Data were analyzed with BD FACSDiva software and DNA values were calculated by comparison to the chicken genome.

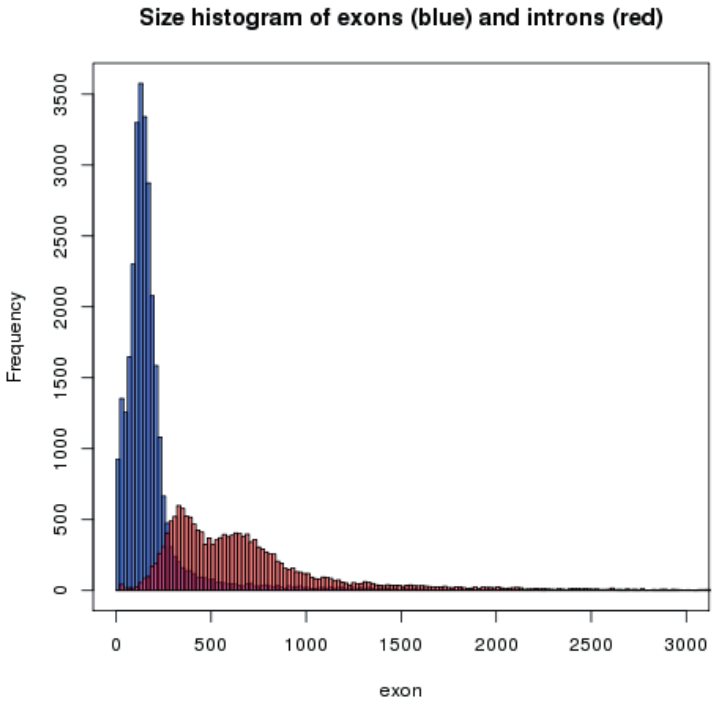


Figure S1 Exon and intron lengths distribution in *B. improvisus* genome.

Table S1 Nuclear DNA content (haploid) of *B. improvisus* obtained by propidium iodide staining and flow cytometry analysis. Chicken *Gallus gallus* red blood cells (diploid C-value 1.14Gbp) were used as a reference to calculate barnacle haploid C-value.

Sample	Tissue	C-value*	Average C-value
Cyprid larvae	30 individuals	747.0 Mbp	738 ± 9.2 Mbp
Adult	Cirri	730.5 Mbp	

*Means for each tissue type generated from three technical replicates for each nuclei preparation (CV technical replicates ≈ 2%).

Table S2 Genetic variance calculated based on sequences of both alleles from a single-copy gene, the octopamine receptor OctA (Lind et al., 2010), in four *B. improvisus* adult individuals. The nucleotide diversity in this 1473 bp single-exon CDS was 2.6-3.9%. Adding data from an additional 27 cloned cDNA from OctA alleles from adults and cyprid larvae (representing several hundreds of individuals from one single population) estimated the OctA pairwise nucleotide diversity in *B. improvisus* for the CDS to be 3.6%.

1. alfa-octopamine receptor (OctR0)

Samples Individual/populations	Genetic variance (%) CDS
Adult 1	3.5
Adult 2	3.9
Adult 3	3.7
Adult 4	2.6
Population adult/cyprid genomic DNA (n=16)	3.7
Population adult/cyprid cDNA (n=27)	3.6

2. Most highly expressed genes (RNA-seq)

≈ 5% genetic variance

References

Bachmann, K., & Rheinsmith, E. L. (1973). Nuclear DNA amounts in pacific Crustacea. *Chromosoma*, *43*(3), 225-236.

Leffler, E. M., Bullaughey, K., Matute, D. R., Meyer, W. K., Segurel, L., Venkat, A., ... & Przeworski, M. (2012). Revisiting an old riddle: what determines genetic diversity levels within species?. *PLoS biology*, *10*(9), e1001388.

Lind, U., Rosenblad, M. A., Wrange, A. L., Sundell, K. S., Jonsson, P. R., André, C., ... & Blomberg, A. (2013). Molecular characterization of the α-subunit of Na+/K+ ATPase from the euryhaline barnacle Balanus improvisus reveals multiple genes and differential expression of alternative splice variants. *PloS one*, *8*(10), e77069.

Panova, M., Aronsson, H., Cameron, R. A., Dahl, P., Godhe, A., Lind, U., ... & Blomberg, A. (2016). DNA extraction protocols for whole-genome sequencing in marine organisms. In *Marine genomics* (pp. 13-44). Humana Press, New York, NY.

Rheinsmith, E. L., Hinegardner, R., & Bachmann, K. (1974). Nuclear DNA amounts in Crustacea. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, *48*(3), 343-348.